QUALITY AND RELIABILITY ENGINEERING INTERNATIONAL
Qual. Reliab. Engng. Int. 2002; **18**: 429–441 (DOI: 10.1002/qre.487)

A MAGNITUDE-ROBUST CONTROL CHART FOR MONITORING AND ESTIMATING STEP CHANGES FOR NORMAL PROCESS MEANS

JOSEPH J. PIGNATIELLO, JR* AND JAMES R. SIMPSON
Department of Industrial Engineering, Florida State University, Florida A&M University, Tallahassee, FL 32310-6046, USA

SUMMARY

Statistical process control charts are intended to assist operators of a usually stable system in monitoring whether a change has occurred in the process. When a change does occur, the control chart should detect it quickly. If the operator can also be provided information that aids in the search for the special cause, then critical off-line time can be saved. We investigate a process-monitoring tool that not only provides speedy detection regardless of the magnitude of the process shift, but also supplies useful change point statistics. A likelihood ratio approach can be used to develop a control chart for permanent step change shifts of a normal process mean. The average run length performance for this chart is compared to that of several cumulative sum (CUSUM) charts. Our performance comparisons show that this chart performs better than any one CUSUM chart over an entire range of potential shift magnitudes. The likelihood ratio approach also provides point and interval estimates for the time and magnitude of the process shift. These crucial change-point diagnostics can greatly enhance special cause investigation. Copyright © 2002 John Wiley & Sons, Ltd.

KEY WORDS: statistical process control; change-point estimation; change-point detection; likelihood ratio; hypothesis testing; confidence region estimation; Shewhart $\overline{X}$ control chart; CUSUM control chart; average run length; quality control; time of process change; process improvement; process monitoring; special cause identification

1. INTRODUCTION

Statistical process control (SPC) charts are used to monitor for process changes by distinguishing the special causes of variation from the common causes of variation. Data from a process are collected, often in subgroups, and a control chart statistic is compared to one or more control limits. As long as the control chart statistic is within its limits, the control chart suggests that only the common causes of variation are present. When the control chart statistic exceeds a limit, the control chart signals that there may be one or more special causes present.

An important aspect of how well a control chart performs is how quickly it responds or reacts to changes in a process. The sooner a process change is detected by a control chart, the sooner the process engineers can initiate their search for the special cause. Once the special cause has been identified the appropriate action can then be taken to rectify or improve the process.

In this paper, we investigate a control chart that quickly detects changes in the mean of a normal process for any magnitude. This magnitude-robust chart also provides both point and interval estimates of the time when the process change first manifested itself and of the magnitude of that change. These estimates can be valuable diagnostic tools to help process engineers in identifying the special cause responsible for the process change. We show that the magnitude-robust control chart can be derived from a likelihood ratio test for a step change in a normal process mean. It turns out that the cumulative sum (CUSUM) chart is a special case of the magnitude-robust control chart. We compare average run length performances and show that the magnitude-robust chart has a better overall average run length performance than that of the CUSUM.

*Correspondence to: Joseph J. Pignatiello, Department of Industrial Engineering, Florida State University, Florida A&M University, Tallahassee, FL 32310-6046, USA. Email: pigna@eng.fsu.edu

Published online 13 August 2002

Received 8 January 2001
Revised 29 January 2002

2. COMMON CONTROL CHARTS FOR MONITORING NORMAL PROCESS MEANS, PROCESS BEHAVIOR MODELS AND ASSOCIATED HYPOTHESIS TESTS

We will assume that the process to be monitored can be modeled well with a normal distribution with mean μ and standard deviation σ. We will assume that subgroups consist of n independent observations, allowing for the $n = 1$ case of individual observations. When the process is in-control we will assume that $\mu = \mu_0$ and that the standard deviation is equal to its in-control value of $\sigma = \sigma_0$.

The Shewhart $\overline{X}$ chart is the most commonly used control chart for monitoring the mean of a normal process. The $\overline{X}$ chart operates by plotting $\overline{X}_T$, the average for subgroup T, against the control limits $LCL = \mu_0 - 3\sigma_{\overline{X}}$ and $UCL = \mu_0 + 3\sigma_{\overline{X}}$, where $\sigma_{\overline{X}} = \sigma_0/\sqrt{n}$. If $\overline{X}_T < LCL$ or $\overline{X}_T > UCL$ then the chart signals that the process mean has changed.

Shewhart did not initially develop the $\overline{X}$ chart from a hypothesis testing perspective. However, for all intents and purposes, with each subgroup the Shewhart $\overline{X}$ chart does conduct a test of the null hypothesis $H_0 : \mu = \mu_0$ versus the two-sided composite alternative $H_a : \mu \neq \mu_0$. With the standard control limits as given above, the probability of type I error is 0.0027.

The average run length (ARL) has been traditionally used to assess how quickly, on average, a control chart reacts to process changes. The Shewhart $\overline{X}$ chart's ARL performance for detecting a step change in the mean from μ_0 to μ is easily obtained from $\text{ARL}(\mu) = 1/(1-\beta(\mu))$, where $\beta(\mu) = \Phi(3-\delta) - \Phi(-3-\delta)$, $\delta = (\mu - \mu_0)/\sigma_{\overline{X}}$ is the standardized magnitude of the step change in the mean and Φ is the standard normal cumulative distribution function.

Although the Shewhart $\overline{X}$ chart's ARL is reasonably low for large changes in the mean (i.e. $|\delta| \geq 3$), the chart is fairly slow to react to small-to-moderately-sized changes (i.e. $|\delta| \leq 2$). To improve ARL performance for the $\overline{X}$ chart, supplemental Western Electric runs rules (WERR) are often applied. Champ and Woodall [1] investigated the ARL performance of the $\overline{X}$ chart when the supplemental WERR were applied. They found that the supplemental WERR significantly improve the ARL performance for the Shewhart $\overline{X}$ chart, but that it is still not as good as that of the CUSUM control chart.

Developed by Page [2,3], the CUSUM chart was derived from a sequential probability ratio test (SPRT) of the null hypothesis $H_0 : \mu = \mu_0$ versus the simple alternative $H_a : \mu = \mu_a$, where $\mu_a = \mu_0 + \delta_a\sigma_{\overline{X}}$ and δ_a is the specific standardized magnitude of change that one wishes to detect. The SPRT (see [4]) operates by comparing the sequential probability ratio

$$SPR(T) = \left[\prod_{t=1}^{T} \frac{1}{\sqrt{2\pi}\sigma_{\overline{X}}} \exp\left(\frac{-1}{2\sigma_{\overline{X}}^2}(\overline{x}_t - \mu_a)^2\right)\right] \times \left[\prod_{t=1}^{T} \frac{1}{\sqrt{2\pi}\sigma_{\overline{X}}} \exp\left(\frac{-1}{2\sigma_{\overline{X}}^2}(\overline{x}_t - \mu_0)^2\right)\right]^{-1} \quad (1)$$

to an appropriate constant A as each new subgroup T is formed. If $SPR(T) > A$, then the test concludes in favor of H_a after observing subgroup T.

For $\delta_a > 0$, it can be shown that the SPRT leads to the following CUSUM procedure for detecting an increase in the mean. For subgroup T, compute $Z_T = (\overline{X}_T - \mu_0)/\sigma_{\overline{X}}$ and then compare

$$S_T^+ = \max\{0, Z_T - k^+ + S_{T-1}^+\} \quad (2)$$

to h^+. If $S_T^+ > h^+$, then signal that the process mean has increased. The reference value k^+ is often taken to be 0.5 [5]. The choice of k^+ is related to the specific simple alternative hypothesis H_a through δ_a. In particular, the SPRT leads to choosing $k^+ = \delta_a/2$. Thus, the common choice of $k^+ = 0.5$ (see [5] or [6]) comes from the specific simple alternative $H_a : \mu = \mu_a$, where $\mu_a = \mu_0 + 1\sigma_{\overline{X}}$.

It also follows that for $\delta_a < 0$, the CUSUM for detecting decreases in the mean is

$$S_T^- = \max\{0, -Z_T - k^- + S_{T-1}^-\} \quad (3)$$

and a signal is generated when $S_T^- > h^-$. Usually, $h^+ = h^- = h$ and $k^+ = k^- = k$. To detect changes in either direction, two one-sided CUSUM charts are monitored simultaneously.

Stoumbos and Reynolds [7] proposed a control chart scheme that involves the SPRT. They suggested using the SPRT to sequentially determine the required sample size at fixed sampling intervals.

Note that the SPRT only considers two specific states. That is, the SPRT considers the process to be either in-control with $\mu = \mu_0$ or else the process is out-of-control with $\mu = \mu_a$. The SPRT does not *explicitly* account for a possible change from one state to another. That is, the SPRT's underlying process behavior model assumes that the process has been and still is in the state of control (H_0) or else the process has been in the pre-specified out-of-control state (H_a) since the process monitoring began. Also, this process behavior model does not explicitly consider any out-of-control value for the process mean other than the prespecified μ_a.

An alternative to the SPRT is to consider a change-point model and a likelihood ratio test. The change-point model assumes that the process is in-control for a period of time before shifting to an out-of-control state. Hinkley [8,9] considered some asymptotic results related to this general approach. Pollak and Siegmund [10] proposed a likelihood-based control chart for situations involving an unknown initial mean. Sullivan and Woodall [11] use the likelihood ratio test to determine whether preliminary observations were collected from an in-control or out-of-control process. A generalized CUSUM was proposed by Lorden [12] and generalized likelihood ratio control charts were proposed by Basseville and Nikiforov [13]. Lai [14] presented a summary of these methods as well as a class of sequential detection rules. The strategy of using the likelihood ratio for detecting a change in a process is also mentioned in Crowder *et al.* [15]. The purpose of this paper is to evaluate the likelihood ratio strategy for detecting and characterizing step changes in the mean.

3. NORMAL PROCESS LOCATION STEP CHANGE MODEL

Consider a permanent step change model for the behavior of a normal process mean. The model assumes that the process is initially in-control with independent observations coming from a normal distribution with a known mean μ_0 and a known standard deviation σ_0. After an unknown point in time $\tau \geq 0$ (known as the process change point), the process location abruptly changes from μ_0 to $\mu_a = \mu_0 + \delta\sigma_{\overline{x}}$, where δ is the unknown magnitude of the change. The model also assumes that once this step change in the process location occurs, the process remains at the new level μ_a until the special cause has been identified and removed. It is in this sense that we say that the location step change is permanent.

The location step change model can be parameterized as follows. During the formation of subgroups $t = 1, 2, \ldots, \tau$, the process mean μ_t is equal to its known in-control value of μ_0. For subgroups $t = \tau+1, \tau+2, \ldots$, the process mean $\mu_t = \mu_a$ where μ_a is an unknown value of the mean when the process is out-of-control. The unknown change point τ is the last subgroup formed from the in-control process. The special case $\tau = 0$ indicates that the process was out-of-control when the first subgroup was formed. Letting T refer to the current or most recent subgroup, then $\tau \geq T$ indicates that a subgroup from the changed process has not yet been formed.

Traditionally, the step change model has been implicitly assumed when evaluating and comparing the ARL performances of various control charts. In some analyses [16,17], the change point is assumed to be $\tau = 0$. In such cases, the change point model reduces to the process behavior model that is associated with the SPRT. The resulting ARL refers to the performance of a control chart which is applied to a process that is already out-of-control. This ARL performance is referred to as the control chart's initial or zero-state ARL performance [18]. For an arbitrarily large $\tau > 0$, the ARL performance of a control chart is called its asymptotic or steady-state ARL performance [18].

4. THE LIKELIHOOD RATIO TEST AND CONTROL CHART FOR A NORMAL PROCESS LOCATION STEP CHANGE MODEL

After observing T subgroups, the null hypothesis of interest is that the process has been and still is in-control. Thus, $H_0 : \mu_t = \mu_0$ for $1 \leq t \leq T$, where μ_0 is the known value for the mean when the process is in control. The alternative hypothesis is that the process was initially in-control, but following some change point τ, the process mean changed from μ_0 to μ_a where μ_a is unknown. Thus, the alternative hypothesis can be stated as $H_a : \mu_t = \mu_0$ for $0 \leq t \leq \tau$ and $\mu_t = \mu_a$ for $\tau + 1 \leq t \leq T$ where both τ and μ_a are unknown.

Given the T subgroup averages, the likelihoods for H_0 and H_a and their ratio can be considered for this hypothesis testing situation. Implemented as a control chart, this likelihood ratio test is conducted as each new subgroup T is formed until a signal is issued. The likelihood under the null hypothesis $H_0 : \mu_t = \mu_0$ for $t = 1, 2, \ldots, T$ is

$$L_0(\mathbf{x}) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi}\sigma_{\overline{x}}} \exp\left(-\frac{1}{2}\left(\frac{\overline{x}_t - \mu_0}{\sigma_{\overline{x}}}\right)^2\right) \quad (4)$$

Assuming a location change point at τ, the likelihood under the alternative hypothesis is

$$L_a(\tau, \mu_a|\mathbf{x}) = \prod_{t=1}^{\tau} \frac{1}{\sqrt{2\pi}\sigma_{\overline{x}}} \exp\left(-\frac{1}{2}\left(\frac{\overline{x}_t - \mu_0}{\sigma_{\overline{x}}}\right)^2\right) \times \prod_{t=\tau+1}^{T} \frac{1}{\sqrt{2\pi}\sigma_{\overline{x}}} \times \exp\left(-\frac{1}{2}\left(\frac{\overline{x}_t - \mu_a}{\sigma_{\overline{x}}}\right)^2\right) \quad (5)$$

The natural logarithm of the ratio of L_a to L_0 can be expressed as

$$R(\tau, \mu_a|\mathbf{x}) = \log_e \frac{L_a(\tau, \mu_a|\mathbf{x})}{L_0(\mathbf{x})} = \frac{1}{2\sigma_{\overline{x}}^2}\left(\sum_{t=\tau+1}^{T} (\overline{x}_t - \mu_0)^2 - \sum_{t=\tau+1}^{T} (\overline{x}_t - \mu_a)^2 \right) \quad (6)$$

Sufficiently large values of this ratio favor the alternative hypothesis that a step change in the mean occurred.

Since τ and μ_a are unknown, the test is conducted by determining the value of $R(\tau, \mu_a|\mathbf{x})$ maximized over all possible values of μ_a and τ given the observed subgroup averages, $\mathbf{x}' = [\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_T]$. This maximum value is denoted as R_T.

For any given value of τ, it is easy to show that the value of μ_a which maximizes $R(\tau, \mu_a|\mathbf{x})$ is

$$\widehat{\mu}_a(\tau) = \overline{\overline{X}}_{T,\tau} = \frac{1}{T-\tau} \sum_{i=\tau+1}^{T} \overline{X}_i \quad (7)$$

the average of the $T - \tau$ most recent subgroup averages. Substituting $\widehat{\mu}_a(\tau)$ for μ_a in $R(\tau, \mu_a|\mathbf{x})$ in (6) and reducing,

$$R(\tau, \widehat{\mu}_a(\tau)|\mathbf{x}) = \frac{T-\tau}{2\sigma_{\overline{x}}^2}(\overline{\overline{X}}_{T,\tau} - \mu_0)^2 \quad (8)$$

can be expressed as a function of τ. Thus, for each new subgroup T, the maximum of the log-likelihood ratio is

$$R_T = R(\widehat{\tau}, \widehat{\mu}_a(\widehat{\tau})|\mathbf{x}) = \max_{0 \le \tau < T} R(\tau, \widehat{\mu}_a(\tau)|\mathbf{x}) \quad (9)$$

where $\widehat{\tau}$ is the value of τ which maximizes $R(\tau, \widehat{\mu}_a(\tau)|\mathbf{x})$ in (8). If R_T is sufficiently large, say $R_T > B$ where B is an appropriate constant, then the test concludes in favor of H_a.

Consider the implementation of this hypothesis test as a control chart. Once each new subgroup T is formed, the statistic R_T would be computed and plotted against an upper control limit of B. If $R_T > B$ then this control chart would signal at subgroup T that there is evidence of a step change in the mean. Furthermore, when this control chart signals,

$$\widehat{\tau} = \arg \max_{0 \le t < T} \frac{T-t}{2\sigma_{\overline{x}}^2}(\overline{\overline{X}}_{T,t} - \mu_0)^2 \quad (10)$$

and

$$\widehat{\mu}_a(\widehat{\tau}) = \overline{\overline{X}}_{T,\widehat{\tau}} = \frac{1}{T-\widehat{\tau}} \sum_{i=\widehat{\tau}+1}^{T} \overline{X}_i \quad (11)$$

provide (maximum-likelihood) point estimates of the change point τ and the new process mean μ_a, respectively.

It should be noted that the standard one-sided CUSUM charts given in Equations (2) and (3) are special cases of the magnitude-robust chart when $\delta = \delta_a$ is a known fixed quantity. If a specific $\delta_a > 0$ is given, then it can be shown that maximizing the log-likelihood ratio in (6) results in the use of S_T^+ in Equation (2) with $k = \delta_a/2$. Thus, the standard CUSUM with $k = 0.50$ results from fixing $\delta_a = 1$. Similarly, S_T^- in Equation (3) follows when the specified $\delta_a < 0$ and the reference value is taken to be $k = -\delta_a/2$. The chart which we are considering uses the same likelihood approach, but with an unknown μ_a (or, equivalently, δ_a). So, an advantage of this control chart is that prior knowledge or specification of δ_a is not required to fine tune the chart.

Another distinct advantage of using a control chart derived from the likelihood ratio approach is that it provides valuable diagnostic tools which can help process engineers focus their search for the special cause. Along with the signal of a process change, the chart being investigated here additionally provides process engineers with point and interval estimates of both the magnitude of the change and when that change first manifested itself in the data.

Samuel *et al.* [19] considered some of the point estimation properties of $\widehat{\tau}$. Pignatiello and Samuel [20] showed that the point estimate $\widehat{\tau}$ is much less biased than the CUSUM chart's built-in estimate of the change point. They also showed that a confidence set for τ (based on [21]) which provides at least 90% coverage for $\delta \ge 1$ can be constructed as

$$\mathbb{C} = \{t \mid R(t, \widehat{\mu}_a(t)) > R_T - 2.97\} \quad (12)$$

where T represents the subgroup when the control chart signaled and R_T is the maximum value of the log-likelihood function from (9). Process engineers could consider any $t \in \mathbb{C}$ as a possible last subgroup from the in-control process. Finally, a 90% confidence interval on μ_a can be given as $\widehat{\mu}_a(\widehat{\tau}) \pm 1.645\sigma_{\overline{x}}/\sqrt{T-\widehat{\tau}}$.

5. AVERAGE RUN LENGTH COMPARISON

Understanding that additional benefits of this method are estimates of the time and magnitude of the process change, we now focus attention on the chart's detection performance relative to some common alternatives. Specifically, we compare the ARL performance of the likelihood-ratio-based chart to that of three CUSUM charts, which differ in the reference values k and corresponding decision intervals h. Each CUSUM chart is designed to detect different-sized shifts in the mean. Although the $k = 0.50$ CUSUM is predominately used in practice, it is also insightful to compare the ARL performances of the $k = 0.25$ and $k = 1.00$ CUSUM charts. Later, we also consider the combined Shewhart–CUSUM chart [22] with Shewhart limits placed at $\pm 3.5\sigma_{\overline{X}}$.

The performance comparison assumes that a step change of magnitude δ occurs following subgroup $\tau \geq 0$. Although several methods including integral equations [16,17] and Markov chains [23] exist for computing ARLs for CUSUM charts, these methods cannot be extended for the likelihood-ratio-based chart. Thus, Monte Carlo simulation is used to estimate ARLs for all of the control charts. To partially verify this approach, simulations of the CUSUM charts were run for $\tau = 0$ and essentially the same results given by the integral equations approach were obtained. These simulation results also agreed with published steady-state ARL results for $\tau > 0$.

In the next section, simulation modeling of the step change is described. Some issues related to the handling of false alarms when $\tau > 0$ and making fair ARL comparisons is then discussed. Finally, ARL results are presented and compared with those of some other control charts.

5.1. *Simulation modeling of a step change*

Monte Carlo simulation was used to estimate the ARL performances of the various control charts. The simulation study was conducted as follows. Observations were generated from an in-control normal distribution for subgroups $t = 1, 2, \ldots, \tau$. Starting with subgroup $\tau + 1$, observations were generated from a normal distribution with mean $\mu_a = \mu_0 + \delta\sigma_{\overline{X}}$. Without loss of generality, $\mu_0 = 0$ and $\sigma_{\overline{X}} = 1$. That is, the simulations focussed on the $Z_t = (\overline{X}_t - \mu_0)/\sigma_{\overline{X}}$ which are $N(0, 1)$ when the process is in-control and $N(\delta, 1)$ after the change point. Step changes of standardized magnitudes $\delta = 0.25, 0.50, \ldots, 5.00$ were examined. Observations on Z_t were collected until the control chart issued a signal at subgroup T. The length for that run was recorded as $T - \tau$. This procedure was then repeated for a total of N independently-seeded runs for each value of δ. A total of $N = 100\,000$ independently-seeded runs were used for each estimated ARL.

5.2. *False alarms*

The simulation modeling of false alarms needs to be carefully addressed. When $\tau > 0$ and a control chart issues a signal at subgroup T where $T \leq \tau$, then the signal is a false alarm since the signal was given before the simulated process change could occur. When a false alarm was encountered in a simulation run, it was treated in the same way that a false alarm would be treated on an actual process. Namely, if one determines that a signal is indeed a false alarm, then one is affirming that the process is currently in-control and could restart their monitoring of the process. Thus, when a false alarm was encountered at subgroup T, the control chart was restarted at subgroup $T+1$ while not altering the scheduled change point.

For example, if the change point was $\tau = 25$ and a false alarm was issued at subgroup 10, then the appropriate statistics would be zeroed out and that simulation run would continue as if subgroup 11 was the *first* one from an in-control process. The process change would still occur at $\tau = 25$. Thus, if there were no other false alarms on this particular run, there would have been 15 subgroups observed from the restarted in-control process when the first subgroup from the changed process (i.e. subgroup 26) is observed. With this approach, the number of subgroups since the chart was started or restarted that have been observed from the in-control process at the time of the process change will not necessarily be fixed at τ, but will instead be random and less than or equal to τ.

5.3. *ARL calibration of control charts*

Each control chart was calibrated so that when the process was in-control, the ARL was equal to about 168. This specific ARL value was chosen since the standard $k = 0.50$ CUSUM with a commonly-recommended decision interval of $h = 4$ has an in-control ARL of 167.7 according to Vance [16] and Gan [17]. For the other two CUSUM charts, decision intervals of $h = 6.53$ ($k = 0.25$) and $h = 2.129$ ($k = 1.00$) produced estimated in-control ARLs of 167.9 and 168.4, respectively. Simulation revealed that $B = 4.87$ resulted in an estimated in-control ARL of 167.6 for the likelihood-ratio-based chart. This estimate was based on $N = 100\,000$ independently-seeded runs, yielding an approximate 95% confidence interval of [166.6, 168.7].

5.4. *Initial ARL performance comparisons*

The first scenario considers control chart ARL performance for processes that are out-of-control when the charts are first applied, i.e. when $\tau = 0$.

Table 1. ARL comparison, $\tau = 0$. Each estimated ARL is based on $N = 100\,000$ independently-seeded runs. Rounded standard errors greater than or equal to 0.01 are shown in parentheses

		CUSUM		
δ	Magnitude-robust	$k = 0.25$	$k = 0.50$	$k = 1.00$
0.25	68.51 (0.18)	58.48 (0.15)	74.26 (0.22)	103.98 (0.32)
0.50	26.57 (0.06)	22.94 (0.04)	26.72 (0.07)	44.24 (0.13)
0.75	14.20 (0.03)	13.44 (0.02)	13.29 (0.03)	20.28 (0.06)
1.00	8.92 (0.02)	9.42 (0.01)	8.42 (0.01)	10.83 (0.03)
1.25	6.24 (0.01)	7.26 (0.01)	6.06 (0.01)	6.74 (0.02)
1.50	4.69 (0.01)	5.94 (0.01)	4.75 (0.01)	4.70 (0.01)
1.75	3.68 (0.01)	5.02	3.91	3.56 (0.01)
2.00	3.01 (0.01)	4.37	3.35	2.87
2.25	2.52	3.88	2.93	2.41
2.50	2.16	3.50	2.62	2.08
2.75	1.89	3.19	2.39	1.83
3.00	1.68	2.94	2.20	1.65
3.25	1.52	2.73	2.04	1.50
3.50	1.38	2.54	1.92	1.38
3.75	1.28	2.38	1.81	1.27
4.00	1.19	2.25	1.71	1.20
4.25	1.13	2.15	1.61	1.13
4.50	1.08	2.07	1.50	1.09
4.75	1.05	2.02	1.40	1.05
5.00	1.03	1.98	1.31	1.03
	$B = 4.87$	$h = 6.53$	$h = 4.00$	$h = 2.129$

Table 1 shows the corresponding estimated ARL values and their associated standard errors.

The results indicate that the CUSUM charts perform well when detecting the shifts which are close to the ones for which they were specifically designed. For example, the $k = 0.25$ CUSUM has the lowest ARL for changes of magnitude $\delta \leq 0.50$. For changes of magnitude of $0.75 \leq \delta \leq 1.25$, the standard $k = 0.50$ CUSUM provides the lowest ARL. For changes of magnitude of $\delta \geq 1.50$, the $k = 1.00$ CUSUM provides the lowest ARL. Although each CUSUM chart performs the best in the δ region close to $\delta = 2k$, no single CUSUM control chart is uniformly best.

It can be seen that although the $k = 0.25$ CUSUM performs well for small changes, it does not perform as well as any of the other charts at detecting large changes. Conversely, the $k = 1.00$ CUSUM which performs quite well at detecting the large changes does not perform particularly well for small changes. The $k = 0.50$ CUSUM can be considered somewhat of a compromise in that it does not perform badly for small or large changes, and it does expectedly well for changes close to $\delta = 1$ in magnitude.

For quick detection of step changes regardless of the magnitude, the magnitude-robust chart should be considered, since it has a strong ARL performance for changes of *all* magnitudes. It outperforms the $k = 0.50$ CUSUM for changes of magnitude $\delta = 0.25$, essentially performs the same for changes of magnitude $0.50 \leq \delta \leq 1.25$ and performs better for changes of magnitude $\delta \geq 1.50$. Although it does not have the best ARL performance at any specific value

Table 2. ARL comparison, $\tau = 50$. Each estimated ARL is based on $N = 100\,000$ independently-seed runs. Rounded standard errors greater than or equal to 0.01 are shown in parentheses

		CUSUM		
δ	Magnitude-robust	$k = 0.25$	$k = 0.50$	$k = 1.00$
0.25	65.41 (0.17)	53.85 (0.15)	72.75 (0.22)	103.46 (0.33)
0.50	24.73 (0.06)	20.25 (0.04)	25.28 (0.07)	43.74 (0.13)
0.75	13.17 (0.03)	11.64 (0.02)	12.41 (0.03)	19.85 (0.06)
1.00	8.28 (0.02)	8.10 (0.01)	7.72 (0.01)	10.60 (0.03)
1.25	5.85 (0.01)	6.21 (0.01)	5.55 (0.01)	6.52 (0.02)
1.50	4.42 (0.01)	5.07 (0.01)	4.33 (0.01)	4.55 (0.01)
1.75	3.49 (0.01)	4.30 (0.01)	3.57	3.44 (0.01)
2.00	2.87 (0.01)	3.74	3.05	2.78
2.25	2.42	3.32	2.68	2.33
2.50	2.09	3.00	2.40	2.01
2.75	1.83	2.75	2.18	1.78
3.00	1.63	2.53	2.01	1.60
3.25	1.48	2.36	1.87	1.46
3.50	1.36	2.20	1.76	1.34
3.75	1.26	2.08	1.65	1.25
4.00	1.18	1.97	1.56	1.18
4.25	1.12	1.89	1.47	1.12
4.50	1.08	1.82	1.38	1.07
4.75	1.05	1.76	1.29	1.05
5.00	1.03	1.71	1.22	1.03
	$B = 4.87$	$h = 6.53$	$h = 4.00$	$h = 2.129$

of δ, there is no range of δ where it does not perform well and it performs nearly the best for all values of δ. In this sense, the magnitude-robust control chart is robust to uncertainty in the magnitude of the change in the mean.

5.5. *Steady-state ARL performance comparisons*

The second ARL performance study considers control charts which are applied on processes that are initially in-control, but experience a step change of magnitude δ following the formation of subgroup $\tau > 0$. The results shown in Table 2 are for a change point of $\tau = 50$. From simulation experiments not reported on here, the ARL performances of the control charts for $\tau = 20$ were approximately the same as for larger values of τ such as $\tau = 100$ and 200. Thus, the results reported here are indicative of a wide range of values of the change point.

Table 2 shows results similar to the $\tau = 0$ case. The $k = 0.25$ CUSUM has the lowest ARLs for changes of magnitude $\delta \leq 0.75$. The standard $k = 0.50$ CUSUM has the lowest ARLs for magnitudes of change in the $1.00 \leq \delta \leq 1.50$ range. The $k = 1.00$ CUSUM has the lowest ARLs when the magnitude of change is $\delta \geq 1.75$. Again, there is no single best control chart for all values of δ. The control chart that has the best ARL performance depends upon the magnitude of the change, δ.

Again, the $k = 0.50$ CUSUM does reasonably well for both small and larger changes. However, the magnitude-robust chart has a better ARL performance than the standard $k = 0.50$ CUSUM for changes of magnitude $\delta \leq 0.50$ as well as for changes of

Table 3. ARL comparison with combined Shewhart–CUSUM scheme and FIR CUSUM, $\tau = 0$

δ	Magnitude-robust	Combined Shewhart-CUSUM	FIR CUSUM
0.00	167.6	159.2	278.3
0.25	68.5	72.3	60.7
0.50	26.6	26.3	17.8
0.75	14.2	13.2	9.4
1.00	8.9	8.3	6.4
1.50	4.7	4.7	4.0
2.00	3.0	3.2	2.9
2.50	2.2	2.4	2.2
3.00	1.7	1.9	1.8
4.00	1.2	1.3	1.2
5.00	1.0	1.1	1.0
	$B = 4.87$	$h = 4.00$	$h = 8.00$
		$k = 0.5$	$k = 0.25$
		$z = 3.5$	$z = 3.5$, $S_0 = 4$

```
MACRO

LIKELIHD X RT
# Develops the control chart statistic RT(T) for the
magnitude-robust chart

MCOLUMN X MU R TEMP RT
MCONSTANT I N T MUNOT

LET N = COUNT(X)
LET MUNOT = 0

DO T = 1: N
    DO I = 1:T
      COPY X TEMP;
        USE I:T
      LET MU(I)=MEAN(TEMP)
      LET R(I)=(T-(I-1))/2*(MU(I)-MUNOT)**2
  ENDDO
    LET RT(T) = MAX(R)
ENDDO

ENDMACRO
```

Figure 1. MINITAB macro for computing the magnitude-robust control chart statistic as well as the change point estimaters

magnitude $\delta \geq 1.75$. Over the range of magnitudes $0.75 \leq \delta \leq 1.50$, the difference in the ARLs of the magnitude-robust chart and that of the standard $k = 0.50$ CUSUM is not more than 0.76. Although the magnitude-robust control chart does not have the best ARL performance for any specific value of δ, it nearly has the best ARL performance for all values of δ. Thus, unless the magnitude of the change is known *a priori*, we conclude that the magnitude-robust chart has a better ARL performance over the entire range of δ values.

5.6. *Comparison with a combined Shewhart–CUSUM scheme*

Another alternative procedure that one might consider is the combined Shewhart–CUSUM scheme (see [22]), which augments a standard CUSUM chart with a Shewhart $\overline{X}$ chart whose limits are placed at $\pm 3.5\sigma_{\overline{X}}$. The ARLs of the magnitude-robust procedure and the combined Shewhart–CUSUM scheme are given in Table 3. The magnitude-robust chart and the combined Shewhart–CUSUM scheme have similar

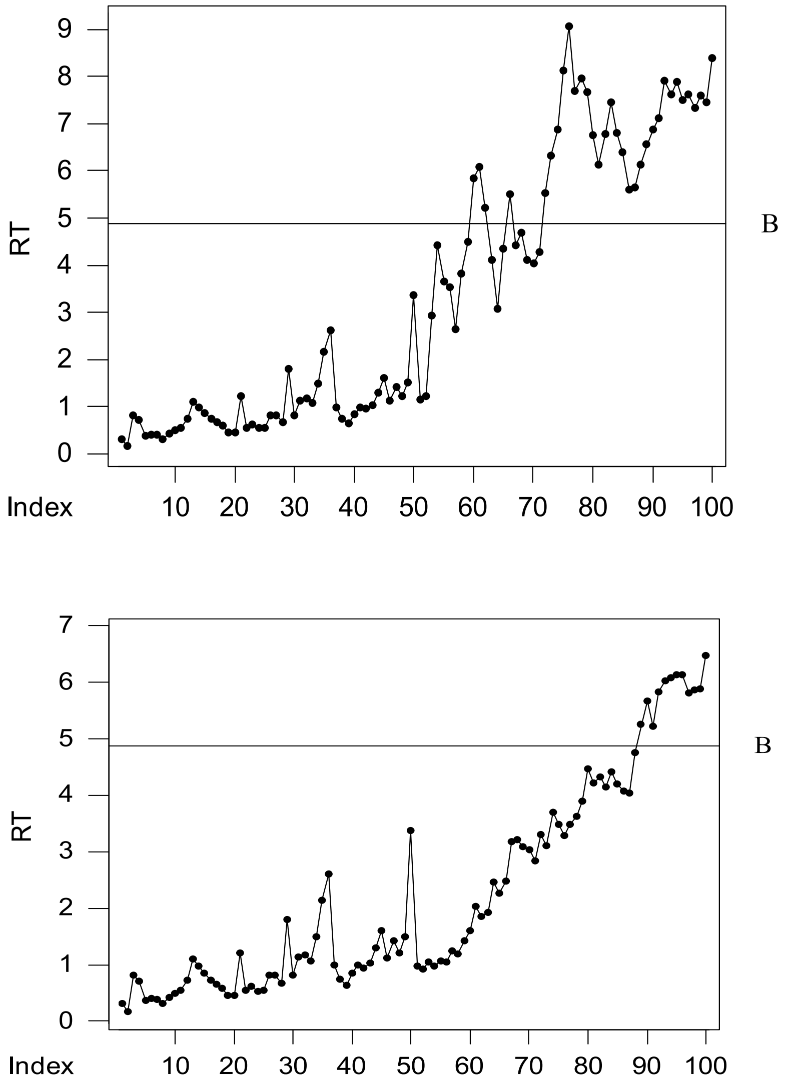


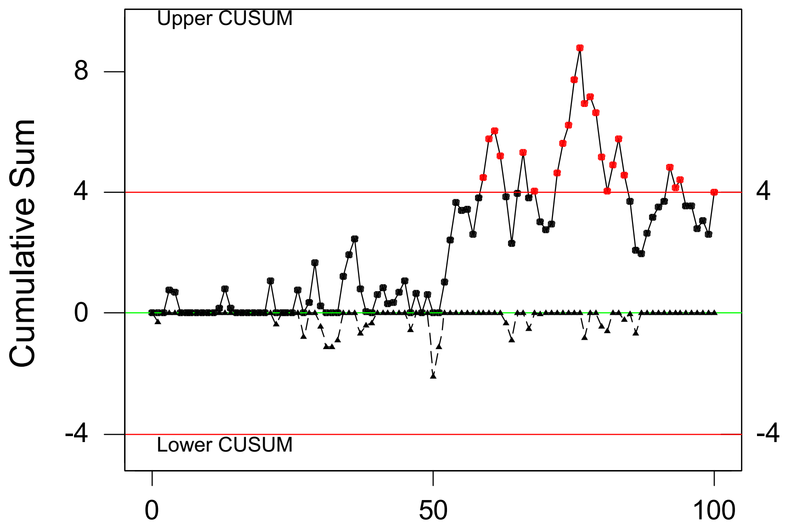
Figure 2. Magnitude-robust control charts for data experiencing a $0.50\sigma_{\overline{X}}$ shift in the mean at $t = 51$. Realization 1 signals at $t = 60$, while realization 2 signals at $t = 89$

ARL performances while that of the FIR CUSUM is superior.

Certainly, the improved statistical performance of a control chart is desirable. However, when selecting a control chart to implement on a process, there are other characteristics of the control chart to consider. For example, the combined Shewhart–CUSUM scheme requires simultaneous monitoring of three control chart statistics and the FIR adds yet another stipulation. Also, neither of these combined schemes readily provide the user with any change-point diagnostics. On the other hand, the magnitude-robust chart requires that only a single statistic be monitored against an upper control limit and readily provides information on the time and magnitude of the change point. In the next section, we use an example to illustrate the design and use of the magnitude-robust chart.

6. IMPLEMENTATION ISSUES

The magnitude-robust chart can be built in a similar manner to most standard control charts for variables. The R_T statistic, which is the maximum of all $R(\tau, \widehat{\mu}_a(\tau)|\mathbf{x})$ values over $0 \leq \tau < T$, can be plotted on a chart against an upper control limit of B. Determining R_T requires T calculations of R (Equation (8)), which increases the amount of computation over most standard control charts. However, a simple program can be written using a



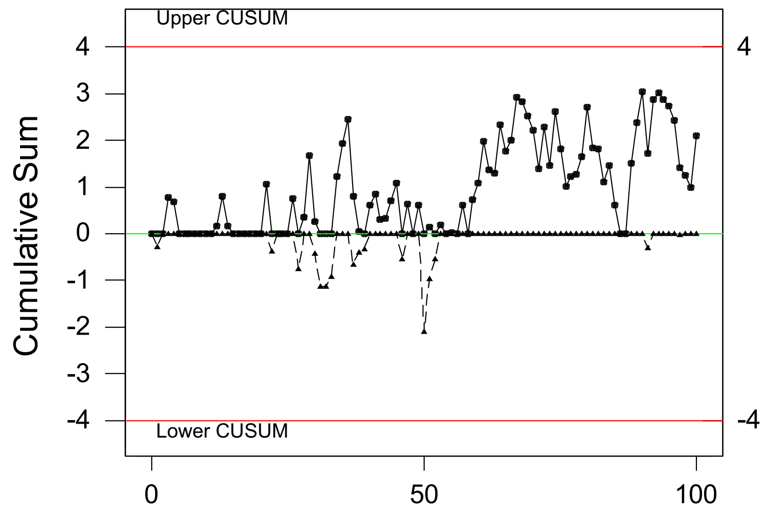


Figure 3. CUSUM control charts with $h = 4$, $k = 0.5$ for the same two realizations of data experiencing a $\delta = 0.50$ in the mean at $t = 51$. Realization 1 signals at $t = 58$, while realization 2 does not signal

common programming language or even a spreadsheet to compute R for each subgroup. Computing R_T for relatively large values of T hardly takes any time even on a modest personal computer.

Values of R_T that exceed B are worthy of special cause investigation. When $R_T > B$, estimates for the time and magnitude of the change (both point estimates and confidence sets) can be provided to aid the search for the special cause. Since point estimates of the time and magnitude of the change are arguments of the R_T statistic, they are available immediately. Another couple of lines of code (or cells in a spreadsheet) can be added to calculate the confidence set for the time of the change and the confidence interval for the magnitude of the change. An example Minitab macro (see Figure 1) demonstrates the relative ease in coding the chart and the change point diagnostics.

The magnitude-robust chart will be illustrated using simulated data generated from known distributions. Initially the data is drawn from random normal variates from an in-control process with mean $\mu = 0.0$ and standard deviation $\sigma_{\overline{X}} = 1.0$. Starting at $t = 51$, the data is drawn from a normal process with a slightly higher mean, $\mu = 0.50$, and the same standard deviation ($\sigma_{\overline{X}} = 1.0$). As a result, the example dataset is in-control for $n = 50$ and experiences a 0.50 $\sigma_{\overline{X}}$ shift in the mean at $t = 51$. Based on an average run

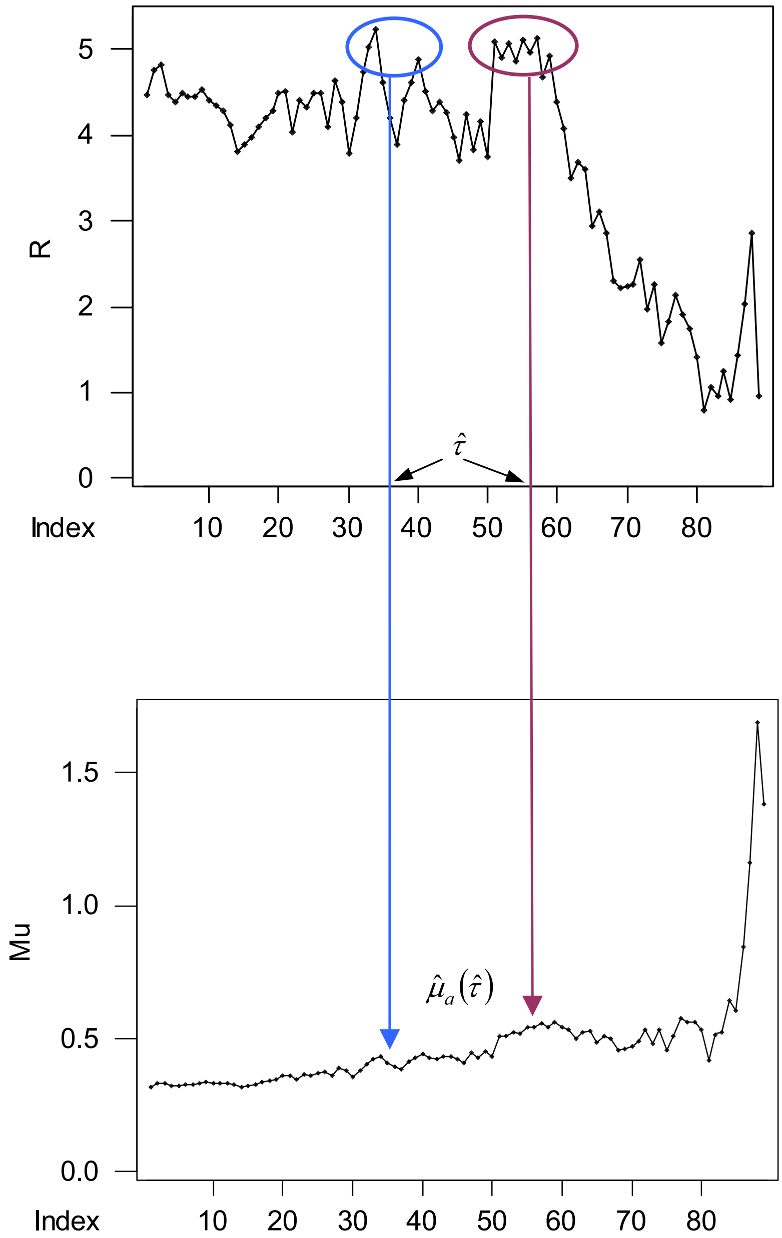


Figure 4. Change-point estimates for dataset realization 2. (a) R versus subgroup index. Large R values correspond to most likely times for the change point. The circled points show the eleven most likely times for the shift change to occur. (b) The estimated size of the shift

length for the magnitude-robust chart with $B = 4.87$ of about 25, we would expect the chart to signal out-of-control sometime around $t = 75$, depending on the actual dataset used.

Two realizations of the above-described data were generated and the resulting control chart statistic R_T is plotted against the upper control limit B (Figure 2). In each case, the control chart signals although the time of signal varies. For comparison, the same data was charted using a CUSUM with similar in-control properties. The time of the signal was similar for realization 1, but the CUSUM did not signal for any of the 100 observations generated in realization 2 (Figure 3).

Once the magnitude-robust control chart signals, diagnostics are immediately available regarding the time and magnitude of the change. Figure 4 shows change-point estimation plots from the realization 2 dataset, where the magnitude-robust chart signals at $t = 89$. The top change-point plot shows computed

Table 4. In-control ARLs for various values of B

B	ARL
4.00	78.626
4.25	97.282
4.50	123.666
4.75	152.392
4.87	167.626
5.00	187.604
5.25	232.366
5.50	292.361
5.75	357.632
6.00	457.914

values for R at each sampling point up to the signal at $t = 89$. The largest values for R indicate the most likely time ($\widehat{\tau}$) the change occurred. A confidence set of these most likely times can also be developed. The estimates $\widehat{\mu}$ are also calculated for each possible change-point estimate. The $\widehat{\mu}$ values associated with the most likely time of the change can be easily identified.

To design the magnitude-robust control chart, one must select an upper control limit B to obtain a desired in-control ARL of ARL_0. We conducted additional simulation runs (see Table 4) of the magnitude-robust control chart with $\tau = 0$, $\delta = 0$ and $N = 10\,000$. In-control ARLs were obtained for various values of B. Applying ordinary least squares to those simulation results yields the following approximately linear relationship between the natural logarithm of ARL_0 and B

$$\log_e \text{ARL} = 0.8728 + 0.8732B \tag{13}$$

Thus, to obtain a value of B for the magnitude-robust control chart for a given ARL_0, an approximate value of B would be

$$\widehat{B} = \frac{\log_e \text{ARL}_0 - 0.8728}{0.8732} \tag{14}$$

assuming that one is interpolating within the range of these data.

7. DISCUSSION

The magnitude-robust control chart offers some significant advantages over existing control charts. In addition to having better overall ARL performance, the magnitude-robust control chart also provides valuable information to process engineers concerning the time of the change and the magnitude of the change. Although the CUSUM chart also provides a change-point estimator (but not a confidence interval), Pignatiello and Samuel [23] showed that it is inferior to the MLE $\widehat{\tau}$ that the magnitude-robust control chart provides.

The use of the likelihood function also suggests a more efficient search strategy for identifying the special cause. The traditional search strategy for finding the special cause is to start with the time of the signal, T, and work backwards in time. That is, at the time of the signal, T, process engineers would examine their logbooks and records corresponding to the time frame of subgroup T. Assuming that nothing was found corresponding to subgroup T, they would then consider the time frames for subgroups $T-1$, $T-2$, etc. This procedure would continue until either: (1) the special cause was correctly identified; (2) an incorrect cause was mistakenly identified as the special cause; or (3) the search was (prematurely) terminated since nothing was found. Obviously, both of the last two outcomes result in some loss of faith in using statistical methods and could lead to abandonment of SPC altogether.

In contrast to the traditional search strategy, the likelihood function can be used to help guide the search by first examining the logbooks and records for the time frame associated with subgroup $\widehat{\tau}$. Assuming that the special cause is not identified there, the next subgroup to consider would be the one with the next-largest log-likelihood ratio. That is, the subgroups t could be searched in order $t_{[1]}, t_{[2]}, \ldots$ according to their associated log-likelihood ratios $R(t_{[1]}, \widehat{\mu}_a(t_{[1]})) \geq R(t_{[2]}, \widehat{\mu}_a(t_{[2]})) \geq \cdots$. Thus, the process engineers could focus their search on those subgroups that are most likely to be associated with the change in the process. Also, associated with each subgroup t is the estimated changed mean, $\widehat{\mu}_a(t)$.

The benefit of using the likelihood-ratio approach is that it not only produces a control chart with good ARL properties, but it also provides an estimate of the magnitude of the change and a confidence set for when the process change first manifested itself in the data. Identifying which combination of the many process variables is responsible for a change in a process allows engineers to improve quality by preventing or avoiding changes in those variables which lead to poor quality and by perpetuating those changes and optimizing those variables which can lead to better quality. Knowing when a process has changed and by how much would simplify the search for the special cause. If the time of the change could be determined, process engineers would have a smaller search window within which to look for the special cause. Consequently, the special cause can be identified more quickly and the appropriate actions needed to improve quality can be implemented sooner.

REFERENCES

1. Champ CW, Woodall WH. Exact results for Shewhart control charts with supplementary runs rules. *Technometrics* 1987; **20**:393–399.
2. Page ES. Continuous inspection schemes. *Biometrika* 1954; **41**:100–115.
3. Page ES. Cumulative sum control charts. *Technometrics* 1961; **3**:1–9.
4. Wetherill GB. *Sequential Methods in Statistics*. Chapman and Hall: London, 1975.
5. Montgomery DC. *Introduction to Statistical Quality Control* (4th edn). John Wiley: New York, 2001.
6. Ryan TP. *Statistical Methods for Quality Improvement* (2nd edn). John Wiley: New York, 1999.
7. Stoumbos ZG, Reynolds MR Jr. Control charts applying a sequential test at fixed sampling intervals. *Journal of Quality Technology* 1997; **29**:21–40.
8. Hinkley DV. Inferences about the change-point in a sequence of random variables. *Biometrika* 1970; **57**:1–17.
9. Hinkley DV. Inferences about the change-point from cumulative sum tests. *Biometrika* 1971; **58**:1–17.
10. Pollak M, Siegmund D. Sequential detection of a change in a normal mean when the initial value is unknown. *Annals of Statistics* 1991; **23**:394–416.
11. Sullivan JH, Woodall WH. A control chart for preliminary analysis of individual observations. *Journal of Quality Technology* 1996; **28**:265–278.
12. Lorden G. Procedures for reacting to a change in distribution. *Annals of Mathematical Statistics* 1971; **41**:520–527.
13. Basseville M, Nikiforov IV. *Detection of Abrupt Changes: Theory and Applications*. Prentice-Hall: Englewood Cliffs, NJ, 1993.
14. Lai TL. Sequential changepoint detection in quality control and dynamical systems. *Journal of the Royal Statistical Society, Series B* 1995; **57**:613–658.
15. Crowder SV, Hawkins DM, Reynolds MR Jr, Yashchin E. Process control and statistical inference. *Journal of Quality Technology* 1997; **29**:134–139.
16. Vance LC. Average run lengths of cumulative sum control charts for controlling normal means. *Journal of Quality Technology* 1986; **18**:189–193.
17. Gan FF. The run length distribution of a cumulative sum control chart. *Journal of Quality Technology* 1993; **25**:205–215.
18. Lucas JM, Saccucci MS. Exponentially weighted moving average control schemes: Properties and enhancements. *Technometrics* 1990; **32**:1–12.
19. Samuel TR, Pignatiello JJ Jr, Calvin JA. Identifying the time of a step change with $\overline{X}$ control charts. *Quality Engineering* 1998; **10**:521–527.
20. Pignatiello JJ Jr, Samuel TR. Estimation of the change point of a normal process mean in SPC applications. *Journal of Quality Technology* 2001; **33**:82–95.
21. Siegmund D. Boundary crossing probabilities and statistical applications. *Annals of Statistics* 1986; **14**:361–404.
22. Lucas JM. Combined Shewhart-CUSUM quality control schemes. *Journal of Quality Technology* 1982; **14**:51–59.
23. Brook D, Evans DA. An approach to the probability distribution of run length. *Biometrika* 1972; **59**:539–549.

Authors' biographies:

Joseph J. Pignatiello, Jr is an Associate Professor in the Department of Industrial Engineering at Florida State University and Florida A&M University. He earned a BS from the University of Massachusetts and an MS and PhD from The Ohio State University. His interests are in quality engineering and include statistical process control, process capability, design and analysis of experiments, robust design and engineering statistics. Dr Pignatiello won the Shewell Award from the American Society for Quality (ASQ) Chemical and Process Industries Division and the Craig Award from ASQ's Automotive Division for his research publications on process capability. He has twice won awards from the Ellis R. Ott Foundation for his publications on statistical process control and the methods of Taguchi. Dr Pignatiello also won an IIE Transactions Best Paper Award for Quality and Reliability. He serves on the editorial boards of IIE Transactions, Journal of Quality Technology and Quality Engineering.

James R. Simpson is an Associate Professor in the Department of Industrial Engineering at Florida State University and Florida A&M University. His research interests include design and analysis of experiments, simulation, response surface methods, applied optimization, regression analysis, and quality control. He has been active in industry working with various organizations including NASA, Department of Defense, Boeing, Ford, Florida Department of Transportation, Lockheed Martin, and Bristol-Myers Squibb. He is an Associate Editor for Quality Engineering. He earned a BS in Operations Research from the United States Air Force Academy, an MS in Operations Research from the Air Force Institute of Technology, and a PhD in Industrial Engineering from Arizona State University. He is a member of ASA, ASQ, and IIE and his email address is simpson@eng.fsu.edu.