Integration of Sequential Process Adjustment and Process Monitoring Techniques

Rong Pan

Enrique del Castillo

Department of Industrial and Manufacturing Engineering

The Pennsylvania State University

University Park, PA 16802

July, 2001

Abstract

When a manufacturing process is subject to random shocks, detecting the changes in the process and adjusting an out-of-target process are two essential functions of process quality control. Traditional SPC techniques emphasize process change detection, but do not provide an explicit process adjustment method. This paper discusses a general sequential adjustment procedure based on Stochastic Approximation techniques and combines it with several commonly used control charts. The performance of these methods depends on the sensitivity of the control chart to detect shifts in the process mean, on the accuracy of the initial estimate of shift size, and on the number of sequential adjustments that are made. It is shown that sequential adjustments are superior to single adjustment strategies for almost all types of process shifts and magnitudes considered. A combined CUSUM chart used in conjunction with our sequential adjustment approach can improve the average squared deviations, the performance index considered herein, more than any other combined scheme unless the shift size is very large. The proposed integrated approach is compared to always applying a standard integral (EWMA) controller with no monitoring component. The number of adjustments in the proposed approach is justified by comparing the cost and the benefit of the adjustment. We show that this strategy – combining control charts and sequential adjustments – is recommended for monitoring and adjusting a process when random shocks occur infrequently in time.

1 Introduction

In traditional statistical process control (SPC) it is frequently assumed that an initially in-control process is subject to random shocks, which may shift the process mean to an off-target value. Different types of control charts are then employed to detect such shifts in mean, since the time of the shift is not predictable. However, SPC techniques do not provide an explicit process adjustment method. Process adjustment is usually regarded a function pertaining to engineering process control (EPC) or automatic process control (APC), areas that traditionally have belonged to process engineers rather than to quality engineers. The lack of adjustments that exists in the SPC applications may cause a large quality off-target cost – a problem of particular concern in a short-run manufacturing process. Therefore, it is important to explore some on-line adjustment methods that are able to keep the process quality characteristic on target with relatively little effort. Integrating EPC and SPC techniques for process quality control has been discussed by several authors in the recent literature, e.g., Box and Kramer [3], MacGregor [18], Montgomery *et al.* [19], Tsung *et al.* [29], Tucker *et al.* [30].

This paper focuses on studying methods for sequential process adjustment based on Stochastic Approximation (SA) techniques for the purpose of quality control. Corresponding to the conditions prevalent in short-run manufacturing systems, small-sample properties of different sequential adjustment techniques will be investigated. We assume a univariate process that consists of a measurable quality characteristic y and a single controllable factor x. The process mean is defined as the expectation of y. The process is initially in a stationary and uncorrelated state, but random shocks could shift the process mean off-target. A control chart is in use to monitor this process and the chart is applied to individual samples because it is assumed the production lot size is small. Whenever the control chart signals an "out-of-control" alarm, we suspect that a mean shift has occurred and proceed with sequentially adjusting the process.

Let μ_t be the process mean at sample or part t and let $\{\varepsilon_t\}$ be a sequence of iid random errors that models both process inherent variation and measurement error. Then the process model is given by the following difference equation:

$$y_t = x_{t-1} + \mu_t + \varepsilon_t \tag{1}$$

where $\varepsilon_t \sim (0, \sigma^2)$. In the simulation presented later, normality of the errors is assumed.

Without loss of generality, the target of y_t is assumed to be zero, thus y can be understood as a deviation from target. The process starts in the in-control state which is assumed to be such that the mean of the process equals the target, i.e., $\mu_1 = 0$ and

$$\mu_t = \mu_{t-1} + \delta(t), \quad \text{for } t = 2, 3, ... \tag{2}$$

with

$$\delta(t) = \begin{cases} 0 & \text{if } t < t_0, \\ \delta \sim N(\mu_s, \sigma_s^2) & \text{if } t \geq t_0, \quad \text{where } t_0 = \text{shift time.} \end{cases}$$

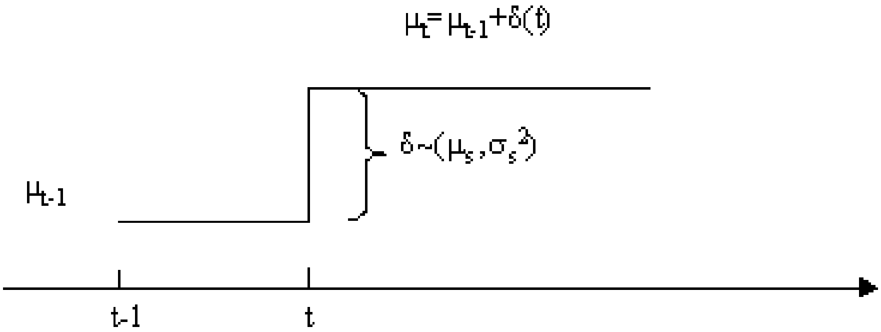

Figure 1: Step-type disturbance on the process mean

Here, t_0, δ, μ_s and σ_s^2 are assumed all unknown, see Figure 1. In this simple model one can see that the random shift in the process mean can be compensated by varying the controllable factor x after the shift is detected. In recent work (see, e.g, Chen and Elsayed [6], Crowder and Eshleman [7], Yashchin [34]), there is considerable emphasis on estimating a time-varying process mean instead of adjusting for such variability. Because the true process mean is not observable directly, adjustments based only on one estimate are almost always biased. This paper discusses a general sequential adjustment procedure and combines it with several commonly used control charts. In order to simplify the set-up of the control chart, the process variance is assumed known in advance. It will be shown that this strategy – combining control charts and sequential adjustments – is good at monitoring and adjusting a process under infrequent random shocks and it also simplifies estimating the current process mean.

In the following sections, some commonly used control charts will be reviewed and their properties on estimating shift sizes will be discussed. A sequential method for process adjustment is then proposed. The performance of various combinations of this adjustment method and different control charts is then evaluated.

2 Control charts and shift size estimates

Shewhart control charts with $\pm 3\sigma$ control limits are the most common type of process monitoring scheme in industry. In this type of chart, sub-group (sample) means or individual observations of the quality characteristic are plotted and any point that is out of the control limits indicates strong evidence that a change in the process mean has occurred. But it is well known that the Shewhart chart is insensitive to small or moderate mean shifts (Montgomery [20]). In order to detect small shifts more quickly, CUSUM (cumulative sum) and EWMA (exponentially weighted moving average) charts are recommended instead.

In particular, a CUSUM chart can be shown to be the generalized likelihood ratio test for the hypothesis $H_0: \ \mu = 0$ versus $H_1: \ \mu = \mu_0$ where μ_0 is a predetermined out of control mean (Lorden [16]). The test statistics for two-sided tabular CUSUM chart are

$$c_t^+ = max\{0, y_t - K + c_{t-1}^+\}$$

and

$$c_t^- = max\{0, -y_t - K + c_{t-1}^-\} \tag{3}$$

where $K = \frac{s}{2}\sigma$ and s is the shift size that one wishes to detect (Woodall and Adams [33]). The control limit of the CUSUM statistics is defined as $H = h\sigma$, where h is another design parameter. Whenever c^+ or c^- exceeds H, an out-of-control alarm is signaled.

EWMA charts use the EWMA smoothing method to predict the process mean. This utilizes not only the current measurement but the discounted historical data as well. The EWMA statistic is defined as

$$z_t = \lambda y_t + (1 - \lambda) z_{t-1}, \quad 0 < \lambda < 1, \tag{4}$$

where λ is the smoothing parameter or weight. The EWMA chart control limits are $\pm L$ $\sigma\sqrt{\frac{\lambda}{(2-\lambda)}[1-(1-\lambda)^{2t}]}$.

A control chart can be used not only to detect the time of the mean shift, but also to estimate the magnitude of the shift. For instance, the EWMA statistic is widely used as the estimate of the current process mean when an EWMA chart detects a shift. In addition, the following equation is used for the CUSUM estimate of the mean:

$$\hat{\mu} = \begin{cases} K + \frac{c_t^+}{N^+} & \text{if } c_t^+ > H \\ -K - \frac{c_t^-}{N^-} & \text{if } c_t^- > H \end{cases} \tag{5}$$

where N^+ and N^- are the number of periods in which a run of non zero values of c^+ or c^-, respectively, were observed (Montgomery [20]). Shift detection and shift size estimation are valuable for process adjustment purposes. If the shift size is precisely known, it is obvious that by letting $x_{t+1} = -\mu_t$ the process will be reset back to its target in view of equation (1). Nevertheless, due to the process disturbances $\{\varepsilon_t\}$, the process mean is not directly observable.

Several process adjustment methods for quality improvement purposes have been proposed in the literature. Taguchi [27] emphasized the importance of adjustments and recommended adjusting by the opposite deviation ($-y_t$) whenever y_t exceeds the control limit of a Shewhart chart. This means that the process mean at the time of the out-of-control is estimated by the last observed data point. This estimate always gives a large shift size, thus is significantly biased when the actual shift size is small. Adams and Woodall [1] also showed that the optimal control parameters and loss functions given by Taguchi are severely misleading in many situations. An alternative feedback adjustment method recommended by Luceño [17] is to use an EWMA statistic of the past data collected from the process. It has been shown that if the disturbance is an integrated moving average (IMA(1,1)) process with parameter θ, the EWMA statistic is optimal in the sense of mean square error when its smoothing parameter λ is equal to $(1-\theta)$. Kelton *et al.* [14] suggested that continuously observing (without adjustment) several data after receiving the first "out-of-control" alarm will benefit the process mean estimation. For instance, they suggest that the average of 10 deviations y_t after an alarm occurs is a good estimate of a shift size of 1.5σ. Delaying the mean estimation was also recommended by Ruhhal *et al.* [22], although they dealt with a generalized IMA process model. Evidently, this method is only acceptable for a manufacturing process with high process capabilities and long production runs. For a short-run process, this approach may produce a high scrap rate.

Wiklund [31, 32] proposed a maximum likelihood estimation (MLE) method based on a truncated normal probability density function. His argument relies on the fact that the estimation of the process mean is made on the condition that one finds a point exceeding the control limit of Shewhart chart. He also discusses other estimation methods based on using CUSUM and EWMA control charts. Table 1 provides results of the estimated shift size by different methods computed from partial results in Wiklund's simulation study, where the standard errors of the estimates appear in parenthesis. In Wiklund's study, he assumes $\mu_s > 0$ but $\sigma_s = 0$. One can see that Taguchi's method is very misleading on small to moderate shifts, that the EWMA is not a sensitive estimator of the shift sizes, and that the MLE and CUSUM perform comparatively better, but they are still inefficient when the shift size becomes large.

μ_s	Taguchi's method	Wiklund's MLE method	CUSUM ($h = 5$, $k = 0.5$)	EWMA ($\lambda = 0.25$, $L = 3$)
0	0 (3.30)	0 (1.38)	0 (2.39)	0 (1.22)
0.5σ	3.1 (1.28)	1.1 (1.11)	1.0 (0.67)	1.24 (0.14)
1σ	3.3 (0.71)	1.3 (1.10)	1.3 (0.54)	1.27 (0.14)
1.5σ	3.4 (0.54)	1.5 (1.22)	1.6 (0.55)	1.31 (0.14)
2σ	3.5 (0.50)	1.8 (1.33)	1.9 (0.67)	1.36 (0.17)
3σ	3.8 (0.60)	2.5 (1.50)	2.6 (0.77)	1.44 (0.26)
4σ	4.3 (0.78)	3.5 (1.60)	3.2 (0.82)	1.55 (0.32)

Table 1: Shift size estimates (and their standard errors) obtained using different methods

More accurate estimation methods of the process mean has appeared in recent research. Chen and Elsayed [6] provide a Bayesian estimation method for detecting the shift size and estimating the time of the shift (t_0). Crowder and Eshleman [7] applied a Bayesian method to the short-run process control. They assumed that the process mean is subject to small frequent changes that result in serial autocorrelation, so the hypothesis test of whether a major shift in the mean occurred is not relevant. Yashchin [34] proposed an adaptive EWMA estimator of the process mean, and he showed that his estimator is good at detecting abrupt mean changes. However, this method require extensive computational effort for estimating the process mean at each step and is unsuitable for on-line process control.

3 Sequential adjustments

Since Taguchi's single observation estimation method is inaccurate in most of the cases, conducting only one adjustment is insufficient to bring the shifted quality characteristic back to the target. A better adjustment scheme can be derived from Stochastic Approximation techniques. The basic idea is that it is better to adjust the off-target process sequentially and estimate the current process mean simultaneously over several time periods.

Suppose a shift occurs at time t_0, and the input variable x_t for $t > t_0$ is varied according to the following equation:

$$x_t = x_{t-1} - a_t y_t \tag{6}$$

where $x_{t_0} = 0$ and $\{a_t\}$ is a series such that $\sum_{t=t_0}^{\infty} a_t = \infty$ and $\sum_{t=t_0}^{\infty} a_t^2 < \infty$. Then, x_t will converge in mean square to the value x^* such that $E[y_t|x_t = x^*] = 0$ (Robbins and Monro [21]). For a process model as simple as equation (1), $a_t = 1/(t - t_0)$ provides the fastest convergence rate (Ruppert [23]). In reality, the time t_0 is unobservable, so it is replaced by t', the time when the shift is detected by the control chart. The setting x_t can be viewed as the negative estimate of the process mean at time t; therefore, $-x_{t'}$ is the first estimate of the shift size and it is recursively updated using equation (6). The actual shift size will be eventually compensated for if adjustments of any size are allowed. However, this is not realistic for the purpose of controlling a short-run manufacturing process. Given the resolution of a machine and the smallest magnitude of the feasible adjustment, one can assign the number of sequential adjustments in advance.

Sequential adjustments applied on machining were first proposed by Grubbs [12] and recently discussed by Trietsch [28]. They dealt with a machine setup error problem as follows: due to a bad setup operation, the output quality characteristic of the machine has an offset value d (unknown) from its target. Suppose the error can be adjusted directly, then the solution which minimizes the variance of the estimate of the offset $\hat{d}$ is obtained by equation (6) with $a_t = 1/t$, which constitutes a harmonic sequence. Clearly, this solution is exactly the same as the sequential adjustment solution of a mean shift except that the shift time is precisely known ($t_0 = 0$). The relation between Grubbs' rule and stochastic approximation was recognized recently by del Castillo and Pan [11]. Their results also show that the harmonic rule performs better for a wide range of shift sizes compared to other rules derived from Kalman filter theory, but it may be inferior to an integral control rule (i.e., an EWMA controller) if the shift size is small.

The EWMA controller is a very popular adjustment device used in the semiconductor industry, but one difficulty in applying this type of controller is selecting a proper control parameter, especially when the manufacturing process is subject to random shocks. To deal with this, Guo *et al.* [13] proposed a dynamic EWMA control scheme, in which the EWMA control parameter is switched to a sequence similar to the harmonic rule when the shock is detected. Guo's method will be discussed in more detail in the next section.

In this paper, the performance of an adjustment scheme is evaluated by the scaled Average Integrated Squared Deviation (AISD) of the process output, which is defined as

$$\frac{\text{AISD}(n)}{\sigma^2} = \frac{1}{n\sigma^2}\sum_{t=1}^{n} y_t^2. \tag{7}$$

For a process having m sequential adjustments after the shift is detected, with adjustments following equation (6) with $a_t = 1/t$, the expectation of this index after the shift detection of the shift equals

$$E\left[\frac{\text{AISD}(m)}{\sigma^2}\right] = \frac{\mu_s^2}{m\sigma^2} + \frac{1}{m}\left[\sum_{t=2}^{m}\frac{t}{t-1} + 1\right]. \tag{8}$$

To derive expression (8) from the process equation (1), assume that a shift occurs at time t_0, i.e., $\mu_t = \delta$ for $t \geq t_0$, $\delta \sim N(\mu_s, \sigma_s^2)$. Let $K_t = \frac{1}{t-t_0}$, so the sequential adjustment scheme is of the form:

$$x_t = x_{t-1} - K_t y_t,$$

with $x_{t_0} = 0$ and the adjustments start at time $t_0 + 1$. After some algebraic manipulations, we can get:

$$y_{t_0} = \delta + \varepsilon_{t_0}$$

and for $t > t_0$

$$y_t = -\frac{1}{t - t_0} \sum_{i=t_0}^{t-1} \varepsilon_i + \varepsilon_t.$$

Therefore,

$$E[y_t] = \begin{cases} \mu_s & \text{for } t = t_0, \\ 0 & \text{for } t > t_0 \end{cases}$$

and

$$Var(y_t) = \begin{cases} \sigma^2 + \sigma_s^2 & \text{for } t = t_0, \\ \frac{t-t_0+1}{t-t_0}\sigma^2 & \text{for } t > t_0 \ . \end{cases}$$

Without loss of generality, we let t_0 be 1. If the magnitude of the shift size is assumed to be a constant, i.e., $\sigma_s = 0$, by substituting the previous expressions into the expectation of (7), equation (8) is obtained.

The performance of the different sequential adjustment rules depends on the number of adjustments and on the precision and accuracy of the initial estimate of the shift size. Tables 2 and 3 give the expected scaled AISD of a shifted process without any adjustment and with several adjustments, respectively, if only one shift occurred. As it can be seen, when the shift size is smaller than or equal to 0.5σ and only a few (less than 10) adjustments are allowed, there is no adjustment scheme which can reduce the AISD of the process.

μ_s	0	0.5σ	1σ	1.5σ	2σ	3σ	4σ
E[AISD(m)]/σ^2	1	1.25	2	3.25	5	10	17

Table 2: E[AISD] of a shifted process without adjustments

E[AISD(m)]/σ^2	$\mu_s = 0$	1σ	2σ	3σ
No. of adj. m=5	1.42	1.62	2.22	3.22
10	1.28	1.38	1.68	2.18
20	1.18	1.23	1.38	1.63

Table 3: E[AISD] of a shifted process with sequential adjustments

4 Integration of EPC and SPC

The proposed integrated process monitoring and adjustment scheme consists of three steps: monitor the process using a control chart, estimate the shift size when a shift in the process mean is detected, and finally apply the sequential adjustment procedure to bring the process mean back to target. To compare the performance of various combinations of control charts and adjustment methods, we first simulate a manufacturing process (1) for a total of 50 observations, and monitor and adjust it using one of the six methods listed on Table 4.

Method	Shift detection	Shift size estimation	Adjustment
1	Shewhart chart for individuals (3σ limits)	Last observation (Taguchi's method)	one adjustment after an out-of-control alarm
2	Shewhart chart for individuals (3σ limits)	Maximum Likelihood Estimate (Wiklund's method)	one adjustment according to the MLE value
3	CUSUM chart for individuals (k=0.5 h=5)	CUSUM estimate (equation (5))	one adjustment according to the CUSUM estimate
4	Shewhart chart for individuals (3σ)	last observation (Taguchi's method)	5 sequential adjustments following (6) with $a_t = 1/(t - t')$
5	Shewhart chart for individuals (3σ)	MLE (Wiklund's method)	5 sequential adjustments following (6) with $a_t = 1/(t - t')$
6	CUSUM chart for individuals (k=0.5 h=5)	CUSUM estimate (equation (5))	5 sequential adjustments following (6) with $a_t = 1/(t - t')$

Table 4: Six methods of integrating control charts and sequential adjustments

We assume that a shift in the mean occurs after the fifth run and adjustments are conducted immediately after the shift is detected. The mean value of 10,000 simulation results are illustrated in Figure 2. The y axis in the figure represents the percentage improvement in the AISD of using some adjustment method compared to the AISD without adjustment, i.e., $\frac{\text{AISD}_{no\ adjust} - \text{AISD}_{method\ i}}{\text{AISD}_{no\ adjust}} \times 100$, so this is a "larger the better" value. This value is plotted with respect to the actual shift size which was varied from 0 to 4σ. Here the shift sizes are constant (i.e., $\sigma_s = 0$). One can see that the sequential adjustment methods (4 to 6) are superior to the one-step adjustment methods (1 to 3) for almost all shift sizes. More specifically, using a CUSUM chart and sequential adjustments (Method 6) has significant advantage over other methods when the shift size is small or moderate, and using a Shewhart chart and sequential adjustments (Method 4) is better for large shifts. Moreover, one-step adjustment methods, especially the Taguchi's method, may dramatically deteriorate a process when the shift size is small. No method can improve AISD when the shift size is very small, but comparatively Method 6 is still better than others.

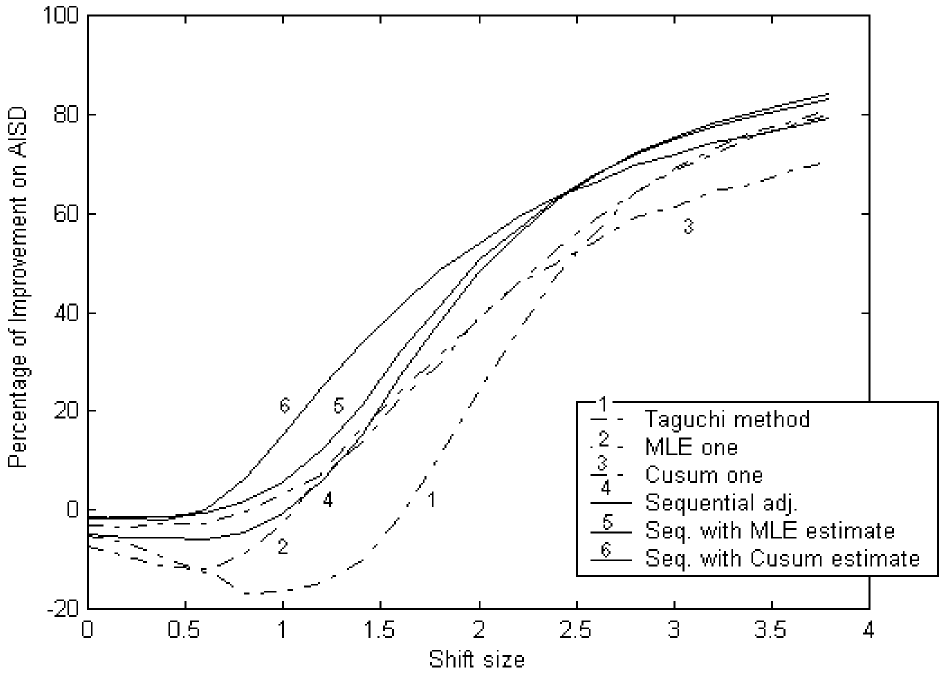

Figure 2: Performance of six integrated methods of control charts and adjustments (the process mean was shifted after the 5th observation)

To study a general shifting process, the mean shift in the following simulation is changed to a stochastic process in which shifts occur randomly in time according to a geometric distribution. Specifically, the occurrence of a shift at each run is a Bernoulli trial with probability $p = 0.05$ and the shift size is normally distributed as $N(\mu_s, \sigma_s^2) = N(\mu_s, 1)$. Besides the previous six methods, an integral control scheme (i.e., an EWMA controller) was studied for comparison purposes. The convergence of EWMA schemes with a small control parameter for adjusting a step type disturbance has been shown by Sachs *et al.* [24]. The EWMA control scheme takes the same form as equation (6) except that the sequence $\{a_t\}$ is a constant λ; here, we set this control parameter at 0.2. There is no process monitoring needed for the integral control scheme because the controller is always in action. The simulations were repeated 10,000 times.

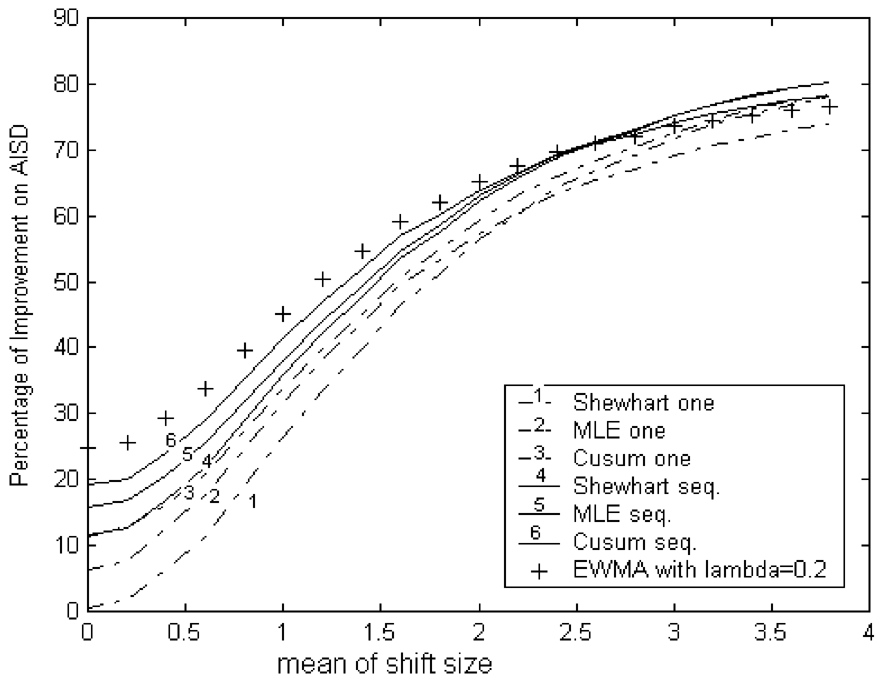

Figure 3: Performance of EPC and SPC integration for a more general shift model (the shift occurs with probability $p = 0.05$ at each observation)

For this general shift model, it can be seen from Figure 3 that sequential adjustment methods still out-perform any one-step adjustment method. Evidently, the EWMA controller performs better than any other sequential method when the shift size mean is small, which explains the popularity of EWMA controllers. However, one main advantage of the proposed SPC/EPC integrated methods is that they detect process changes using common SPC charts whereas the EWMA controller alone does not have this SPC function, in other words, there is no possibility for process improvement through correction of assignable causes if only an

EWMA controller is utilized. Process improvement through human intervention is facilitated by having a monitoring (SPC) mechanism that triggers the adjustment procedure and keeps a time-based record of alarms useful for process diagnostics.

The step-type random shift model considered in this paper is similar to Barnard's model (Barnard, [2]). When the shift occurs more and more frequently, it approaches an IMA (integrated moving average) disturbance. As it is well-known, the EWMA controller is the minimum variance controller for a responsive process with IMA disturbance (Box and Luceño, [5]). This explains why the EWMA chontroller works better when shifts occur more frequently (larger p) than less frequently (smaller p). However, since this scheme conducts adjustments at every process run, it cannot be suitable for a process where adjustment costs must be considered. The economic consideration of adjustments will be discussed in the next section.

Another drawback of the EWMA controller is that one has to decide what value of the control parameter λ to use. It is recommended that this parameter should be small in order to maintain the stability of the process (Sachs *et al.* [24]), but small parameter values may not be optimal from an AISD point of view, especially when the mean shift size is large. Moreover, the high performance of the EWMA scheme comes from the frequent random shifts modeled in the previous simulation study (an average of 2.5 shifts per 50 runs). If the chance of shifts decreases, the inflation of variance which is caused by adjusting an on-target process will deteriorate the effectiveness of this scheme. The inflation in variance for discrete integral (EWMA) controllers has been studied by Box and Luceño [5] and del Castillo [10] who provided asymptotic results. The small-sample properties of the variance provided by EWMA and harmonic adjusting rules are given by del Castillo and Pan [11].

In Figure 4, the probability of random shifts p was decreased to 0.01 and the same simulation as in Figure 3 was conducted. Under these conditions, the EWMA method cannot compete well with the sequential adjustment methods combined with CUSUM or Shewhart chart monitoring. More simulation results for different probabilities of shifts p are listed in Table 5. It is found that the EWMA adjustment method is better for small shifts and Method 4 is better for large shifts when p is large; as p gets smaller ($p < 0.02$), i.e., the process is subject to infrequent random shocks, Method 6 gets harder to beat. Therefore, the proposed SPC/EPC integrated methods work better when p is small, which is relevant in the microelectronics industry where process upsets occur very rarely.

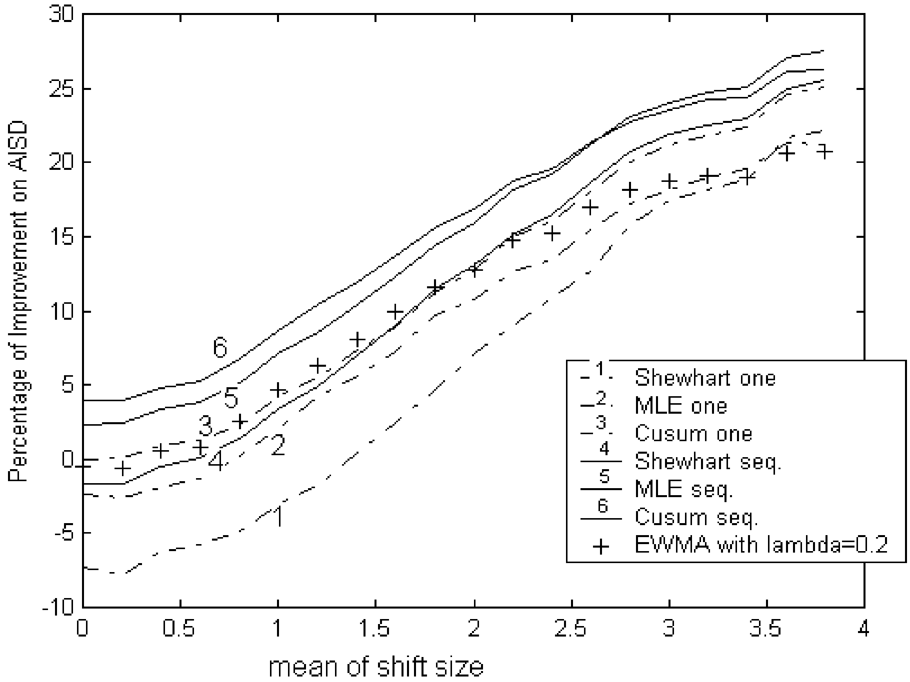

Figure 4: Performance of EPC and SPC integration for the general shift model, less frequent shifts (p=0.01).

The performance of the different SPC/EPC integration methods studied herein depends on 1) their ability to detect a shift and 2) their ability to estimate the process mean. When a process is in the in-control state, an out-of-control alarm signaled by the control chart is called a false alarm. Adjustments triggered by false alarms will inflate the variance of the in-control process, although the inflation will converge to zero when sequential adjustments are used. On the other hand, if the control chart cannot signal an alarm quickly after a real shift has occurred, it will also impede a quick recovery through adjustment.

Since the detection properties of a CUSUM chart can be tuned by modifying its design

% improvement on AISD		Mean of shift size: 0	1σ	2σ	3σ	4σ
p=0.05	Method 4	11.20	36.05	62.27	**74.93**	**81.04**
		(0.30)	(0.38)	(0.36)	(0.31)	(0.29)
	Method 6	18.89	41.50	64.07	73.90	78.73
		(0.28)	(0.35)	(0.33)	(0.30)	(0.29)
	EWMA controller	**24.91**	43.11	60.47	67.76	70.71
	($\lambda = 0.1$)	(0.27)	(0.32)	(0.30)	(0.28)	(0.28)
	EWMA controller	24.51	**45.32**	65.26	73.31	76.68
	($\lambda = 0.2$)	(0.30)	(0.36)	(0.33)	(0.31)	(0.30)
	EWMA controller	21.16	44.02	**65.59**	74.21	78.38
	($\lambda = 0.3$)	(0.33)	(0.39)	(0.36)	(0.33)	(0.32)
p=0.035	Method 4	6.65	24.31	47.76	**62.41**	**68.85**
		(0.26)	(0.37)	(0.40)	(0.39)	(0.38)
	Method 6	13.80	30.35	50.56	61.91	66.68
		(0.25)	(0.33)	(0.36)	(0.36)	(0.36)
	EWMA controller	**18.31**	32.18	48.68	56.01	59.58
	($\lambda = 0.1$)	(0.25)	(0.32)	(0.34)	(0.34)	(0.34)
	EWMA controller	16.82	**32.81**	51.21	61.09	64.35
	($\lambda = 0.2$)	(0.29)	(0.36)	(0.39)	(0.38)	(0.39)
	EWMA controller	13.13	30.33	**51.76**	60.82	65.48
	($\lambda = 0.3$)	(0.32)	(0.40)	(0.41)	(0.41)	(0.42)
p=0.02	Method 4	1.48	11.85	28.86	41.60	**48.34**
		(0.24)	(0.32)	(0.39)	(0.43)	(0.45)
	Method 6	8.07	17.52	**32.53**	**41.94**	47.20
		(0.21)	(0.29)	(0.36)	(0.39)	(0.41)
	EWMA controller	**10.37**	**18.86**	30.68	38.05	41.06
	($\lambda = 0.1$)	(0.22)	(0.29)	(0.35)	(0.38)	(0.39)
	EWMA controller	7.35	17.09	31.40	39.49	43.57
	($\lambda = 0.2$)	(0.26)	(0.33)	(0.40)	(0.43)	(0.45)
	EWMA controller	2.16	13.03	28.90	38.19	42.28
	($\lambda = 0.3$)	(0.28)	(0.37)	(0.44)	(0.47)	(0.49)
p=0.005	Method 4	-3.36	-1.02	3.64	9.02	12.57
		(0.18)	(0.21)	(0.28)	(0.34)	(0.37)
	Method 6	**1.32**	**3.60**	**7.88**	**11.77**	**14.37**
		(0.12)	(0.16)	(0.23)	(0.28)	(0.32)
	EWMA controller	-0.36	1.55	5.53	7.72	9.95
	($\lambda = 0.1$)	(0.13)	(0.17)	(0.24)	(0.27)	(0.30)
	EWMA controller	-5.55	-2.91	1.42	4.89	7.19
	($\lambda = 0.2$)	(0.16)	(0.21)	(0.27)	(0.32)	(0.35)
	EWMA controller	-11.25	-8.47	-2.94	0.42	3.15
	($\lambda = 0.3$)	(0.18)	(0.23)	(0.31)	(0.36)	(0.39)

Table 5: Performance of SPC/EPC integrated adjustment schemes and EWMA scheme when varying the probability of a shift. The numbers are the mean values and standard errors (in parenthesis) of the percentage improvement on AISD (compared to the process without adjustment) computed from 10,000 simulations. Bold numbers are largest improvement for each p and mean shift size combination.

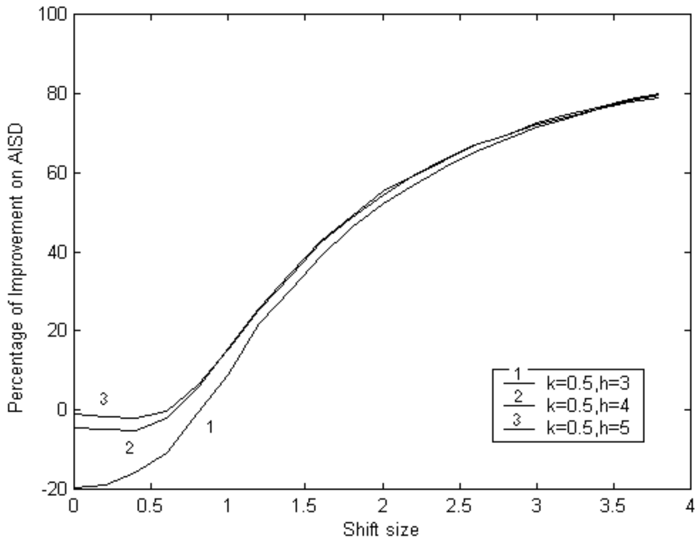

Figure 5: Performance of Method 6 with different parameters in the CUSUM chart (the process mean was shifted after the 5th observation)

parameters h and k, it is of interest to study Method 6 with different CUSUM chart parameters. In Figure 5, several different values of h were tried while fixing k at 0.5 to make the chart sensitive to small shifts. It was found that when h is small, the process will suffer from a large number of false alarms generated by the control chart; when h is large, the improvement in AISD will be limited for large shift sizes due to the lack of sensitivity that the CUSUM chart has to large shifts. A CUSUM chart with $h = 5$ seems to be the best choice since it gives fewer false alarms for a normal process and has comparatively short ARLs for large shift sizes.

In order to improve further the performance of Method 6 for large frequent shifts, we propose a hybrid monitoring scheme combined with a sequential adjustment scheme. A combined CUSUM-Shewhart chart is used, where the parameters on the CUSUM are $k = 0.5$ and $h = 5$ and the control limits on the Shewhart chart are set at $\pm 3.5\sigma$. Whenever the combined chart signals an alarm, the initial estimate of the shift size will be given by the CUSUM estimate if it is smaller than 1.5σ; otherwise, it will be the negative value of y_t (Taguchi's method). The average run lengths of this combined monitoring approach are

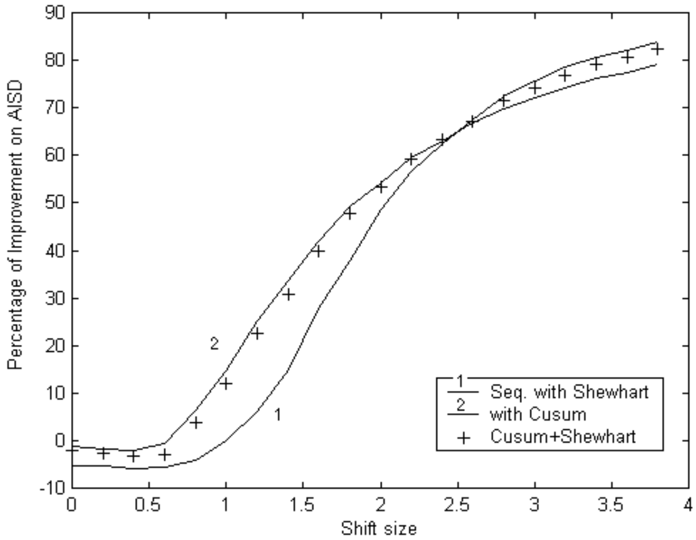

Figure 6: Performance of a hybrid monitoring and adjusting method (the process mean was shifted after the 5th observation)

ARL	Shift size				
	0	1σ	2σ	3σ	4σ
CUSUM (h=3)	59	6.36	2.56	1.59	1.15
CUSUM (h=4)	169	8.34	3.22	1.98	1.44
CUSUM (h=5)	469	10.34	3.89	2.39	1.72
CUSUM-Shewhart	391	10.20	3.77	2.10	1.34
Shewhart (3σ)	370	43.96	6.30	2.00	1.19

Table 6: ARLs of CUSUM and CUSUM-Shewhart charts. The value $k = 0.5$ was used in the CUSUM charts and $k = 0.5$, $h = 5$ and $c = 3.5$(Shewhart control limit) were used in the CUSUM-Shewhart chart. The ARLs of CUSUM charts are approximated using equations given by Siegmund [25] and the ARLs of CUSUM-Shewhart chart are from Montgomery [20].

contrasted with those of a CUSUM chart in Table 6. Comparing this new method to Methods 4 and 6 (see Figure 6), one can see that the new method makes a considerable improvement on the large shift size while sacrificing a little for small shift sizes. This trade-off cannot be avoided due to the nature of this hybrid monitoring method.

We finally point out in this section that a method for sequentially adjusting the parameter of an EWMA controller was recently proposed by Guo *et al.* [13]. They use two EWMA control charts for detecting moderate (2σ) and large (3σ) shifts. After detection, a harmonic adjustment sequence is triggered when either chart signals an alarm. In Figure 7, the two-EWMA method with the suggested chart parameters by Guo *et al.* is compared with Method 4, Method 6 and with the hybrid monitoring method proposed before by using the general shift model with the shift probability p equals 0.05. Clearly, the two-EWMA method performs worse than other methods, especially on large shift sizes. This can be explained by the insensitivity of EWMA chart on estimating a general shift size (Table 1).

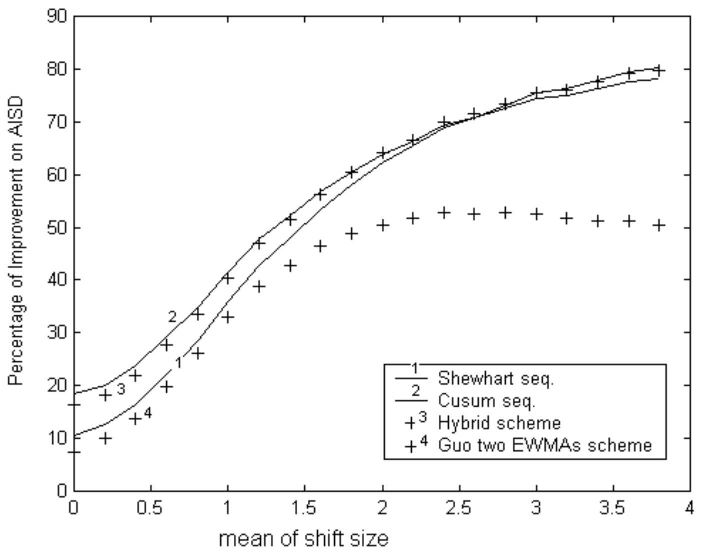

Figure 7: Comparing the two-EWMA method with other SPC/EPC integrated schemes. $\lambda_1 = 0.6$, $L_1 = 3.285$ and $\lambda_2 = 0.33$, $L_2 = 3.25$ were used for the two EWMA charts. Shifts occur with $p = 0.05$.

5 Cost justification

In quality control, the cost of adjusting a process usually can not be ignored, because the adjustment is a set of decisions and actions such as stopping the process, investigating the causes of out-of-control and resetting the process. In the simulations shown in the previous section, the number of sequential adjustments was arbitrarily selected as five. The economic consideration of the number of adjustments will be discussed in this section.

When the harmonic adjustment rule is applied on a shifted process with shift size μ_s, $(x_t + \mu_s)$ is asymptotically normally distributed [23], that is

$$(x_t + \mu_s) \to^D N(0, \sigma^2/t).$$

Therefore, x_t will be likely in the interval of $\pm 3\sigma/\sqrt{t}$ around $-\mu_s$ and the effect of an adjustment will decrease rapidly when the number of adjustments grows. Figure 8 presents the results of the simulation studies where different numbers of adjustments are applied. The process was assumed to have a mean shift after the 5th run and the simulation was repeated 1,000 times. It was found that after four or five adjustments the AISDs of the process can not be further improved significantly. In Figure 9, a 3-D plot of the AISD improvement function of Method 6 (integrated CUSUM chart and sequential adjustments) is shown as a function of the mean shift size (μ_s) and the number of adjustments. The AISD improvement function is very flat on the adjustment number axis, compared to on the shift size axis, so it is not worth to do many adjustments.

The optimal number of adjustments can be obtained if the cost elements of the off-target process and cost of adjustments are known. By marginal analysis, adjustments should be conducted as long as the adjustment cost is lower than the savings obtained from decreasing the AISD by adjusting the process one more time. Suppose a process is going to be run for N observations or parts, and it will be adjusted sequentially for the first n parts. Then the expected AISD is defined as:

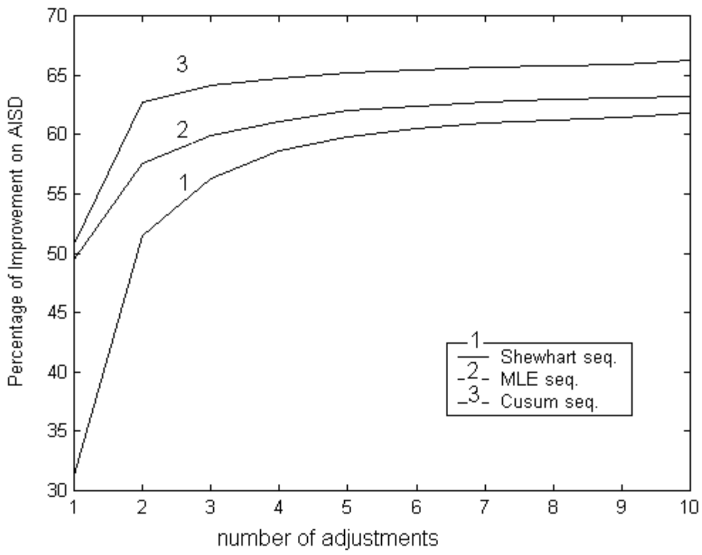

Figure 8: Improvements on AISDs when the adjustment number increases (a general shift model with p=0.05, $\mu_s = 2$ and $\sigma_s = 1$ was used.)

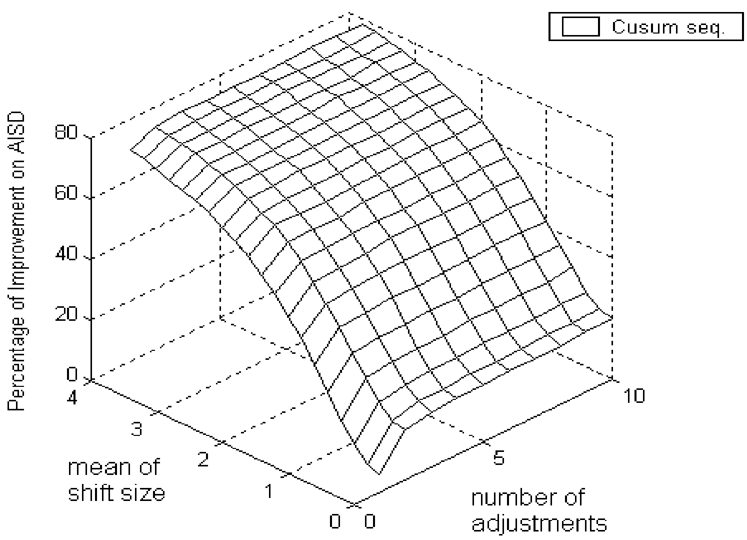

Figure 9: A 3-D view of the AISD improvements from Method 6 when both shift size and adjustment number change ($\sigma_s = 1$)

$$E[AISD(n,N)] = \frac{nE[AISD(n)] + (N-n)(Var(y_{n+1}) + E[y_{n+1}]^2)}{N} \quad (9)$$

So the adjustment is only profitable when $\Omega N\{E[AISD(n,N)] - E[AISD(n+1,N)]\} > M$, where Ω is the unit off-target cost and M is the adjustment cost. By using equations (8) and (9), we get

$$n < \frac{\sqrt{(M+\Omega\sigma^2)^2 + 4(N-1)M\Omega\sigma^2} - (M+\Omega\sigma^2)}{2M} \quad (10)$$

For example, with $N = 50$ and $\sigma = 1$, the optimal number of adjustments computed by equation (10) is given in Table 7.

M/Ω	1	2	5	10
n	6	4	2	1

Table 7: Optimal number of adjustments

6 Conclusions and future work

In this paper, several combinations of process monitoring and adjusting methods were studied. The performance of these methods depends on the sensitivity of the control chart to detect shifts in the process mean, the accuracy of the initial estimate of the shift size, and the number of sequential adjustments. Sequential adjustments are superior to single adjustment strategies for almost all types of process shifts and magnitudes considered, and a CUSUM chart used together with a simple sequential adjustment scheme can reduce the average squared deviations of a shifted process more than any other combined scheme when the shift size is not very large. We further propose a hybrid monitoring method, which, when coupled with a sequential adjustment scheme, has a more competitive performance on both small and large shift sizes.

Unlike some commonly used automatic process control methods, the integrated SPC/EPC schemes that we proposed do not require continuous adjustments on the process. Therefore, these methods are suitable for process control when the process is subject to infrequent random shocks. The number of adjustments can be justified by comparing the cost and the benefit of the adjustment. Since sequential adjustments are applied, the effect of the initial estimate of the process mean is not as critical as in the single adjustment method, so this method requires much less computation effort and is easy to be implemented on the manufacturing floor.

In this study, the process was assumed to be stable and to have no autocorrelation, and only a simple step-type change was investigated. The effectiveness of the integrated SPC/EPC schemes on autocorrelated processes and on other types of changes needs to be studied further. With respect to process monitoring, English *et al.* [8] showed that for an AR (autoregressive) time series process, the EWMA chart is preferred over a Shewhart chart in detecting mean shifts and changes in AR parameters. With respect to automatic control, Box and Luceño [5], Luceño [17] and Srivastava [26] examined the impact of control actions on IMA (integrate moving average) time series process and proposed optimal control limits if the adjustment cost is not trivial. The adjustment strategy recommended in the literature is either single adjustment based on the one-step estimate of the process change or consistent adjustment like the EWMA rule. As demonstrated in this paper, integrating control charts and sequential adjustments, which combines monitoring, estimation and adjustment, can be a better alternative for controlling a process with infrequent sudden changes, and it is worth to be explored for more general process models and types of changes.

References

[1] Adams, B.M. and Woodall, W.H. (1989). An analysis of Taguchi's on-line process control procedure under a random-walk model. *Technometrics*, **31**: 401-413.

[2] Barnard, G.A. (1959). Control charts and stochastic processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, **21**(2): 239-271.

[3] Box, G.E.P. and Kramer, T. (1992). Statistical process monitoring and feedback adjustment - a discussion. *Technometrics*, **34**: 251-267.

[4] Box G.E.P., Jenkins, G.M. and Reinsel, G. (1994) *Time Series Analysis, Forecasting, and Control*, 3rd edtion, Englewood Cliffs: Prentice Hall.

[5] Box, G.E.P. and Luceño, A. (1997) *Statistical Control by Monitoring and Feedback Adjustment*, John Wiley & Sons, New York, NY.

[6] Chen, A. and Elsayed, E.A. (2000) An alternative mean estimator for processes monitored by SPC charts. *Int. J. Prod. Res.*, **38** (13): 3093-3109.

[7] Crowder, S.V. and Eshleman, L. (2001) Small smaple properties of an adaptive filter appled to low volume SPC. *Journal of Quality Technology*, **33** (1): 29-45.

[8] English, J.R., Lee, S., Martin, T.W., and Tilmon, C. (2000) Detecting changes in autoregressive processes with X-bar and EWMA charts. *IIE Transactions*, **32** (12): 1103-1113.

[9] Del Castillo, E. and Hurwitz, A. (1997) Run to Run Process Control: a Review and Some Extensions. *Journal of Quality Technology*, **29** (2): 184-196.

[10] Del Castillo, E. (2001) Some properties of EWMA feedback quality adjustment schemes for drifting disturbances. *Journal of Quality Technology*, **33** (2): 153-166.

[11] Del Castillo, E. and Pan, R. (2001) An unifying view of some process adjustment methods. under review in the *Journal of Quality Technology*.

[12] Grubbs, F.E. (1954) An optimum procedure for setting machines or adjusting processes. *Industrial Quality Control*, July, reprinted in *Journal of Quality Technology*, 1983, **15** (4): 186-189.

[13] Guo, R.S., Chen, A. and Chen, J.J. (2000) An enhanced EWMA controller for processes subject to random disturbances. In *Run to run control in semiconductor manufacturing*, Moyne, J., Del Castillo, E., and Hurwitz, A., eds., CRC press.

[14] Kelton, W.D., Hancock, W.M. and Bischak, D.P. (1990) Adjustment rules based on quality control charts. *Int. J. Prod. Res.*, **28** (2): 365-400.

[15] Kramer, T.T. (1989) Process control from an economic point of view. Ph.D. Dissertation, Statistics, University of Wisconsin - Madison.

[16] Lorden, G. (1971) Procedures for reacting to a change in distribution. *Annals of Mathematical Statistics*, **41** (2): 520-527.

[17] Luceño, A. (1993) Performance of EWMA versus last observation for feedback control. *Commu. Statist. - Theory Meth.*, **22** (1): 241-255.

[18] MacGregor, J.F. (1990) A different view of the funnel experiment. *Journal of Quality Technology*, **22**: 255-259.

[19] Montgomery, D.C., Keats, J.B., Runger, G.C., Messina, W.S. (1994) Integrating statistical process control and engineering process control. *Journal of Quality Technology*, **26**: 79-87.

[20] Montgomery, D.C. (1996) *Introduction to Statistical Quality Control*, 3rd edition, New York, Wiley.

[21] Robbins, H. and Monro, S. (1951) A stochastic approximation method. *Annals of Mathmatical Statistics*, **22**: 400-407.

[22] Ruhhal, N.H., Runger, G.C. and Dumitrescu, M. (2000) Control chanrts and feedback adjustments for a jump disturbance model. *Journal of Quality Technology*, **32** (4): 379-394.

[23] Ruppert, D. (1991) Stochastic approximation. In *Handbook of Sequential Analysis*, B. K. Ghosh and P. K. Sen, Marcel Dekker, NY. 503-527.

[24] Sachs, E., Hu, A. and Ingolfsson, A. (1995) Run by Run Process Control: Combining SPC and Feedback Control. *IEEE Transactions on Semiconductor Manufacturing*, **8** (1): 26-43.

[25] Siegmund, D. (1985) *Sequential Analysis: Tests and Confidence Intervals*, Springer-Verlag, New York.

[26] Srivastava, M.S. (1999) On-line control procedures for integrated moving average process for order one. *Commu. Statist. - Theory Meth.*, **28** (8): 1857-1882.

[27] Taguchi, G. (1985) Quality engineering in Japan. *Commu. Statist. - Theory Meth.*, **14**: 2785-2801.

[28] Trietsch, D. (1998) The harmonic rule for process setup adjustment with quadratic loss. *Journal of Quality Technology*, **30** (1): 75-84.

[29] Tsung, F., Wu, H. and Nair, V. (1998) On the efficiency and robustness of discrete proportional-integral control schemes. *Technometrics*, **40** (30): 214-222.

[30] Tucker, W.T., Faltin, F.W. and Vander Weil, S.A. (1993) Algorithmic statistical process control: an elaboration. *Technometrics*, **35**(4): 363-375.

[31] Wiklund, S.J. (1992) Estimating the process mean when using control charts. *Economic Quality Control*, **7**: 105-120.

[32] Wiklund, S.J. (1993) Adjustment strategies when using shewhart charts. *Economic Quality Control*, **8**: 3-21.

[33] Woodall, W.H., and Adams, B.M. (1993) The statistical design of CUSUM charts. *Quality Engineering*, **5**(4): 559-570.

[34] Yashchin, E. (1995) Estimating the current mean of a process subject to abrupt changes. *Technometrics*, **37** (3): 311-323.

Why this paper is important

In this paper, we present a novel way of integrating process monitoring (SPC) and process adjustment (EPC) for controlling shifts in the mean value of the quality characteristic in a process.

It is known that traditional SPC techniques emphasize process change detection, but do not provide an explicit process adjustment method. The adjustment strategy recommended in the literature is either to perform a single adjustment based on the a single estimate of the process mean at an alarm time or to use continuous adjustments like in an EWMA (or integral) controller. This paper discusses a general sequential adjustment procedure based on **Stochastic Approximation** techniques and combines it with several commonly used control charts. The performance of these combinations are compared with the performance of the single adjustment method and with an EWMA controller through simulation studies. We show that sequential adjustments are superior to single adjustment strategies for almost all types of process shifts and magnitudes considered. A standard CUSUM chart used in conjunction with a sequential adjustment technique can reduce the squared deviations of a shifted process more than any other combined scheme when the shift size is not very large. Furthermore, a hybrid monitoring method coupled with sequential adjustment is proposed, which has a balanced performance on small and large shift sizes.

Unlike some commonly used automatic process control methods, the integrated SPC/EPC scheme that we proposed does not require continuous adjustments on the process, and allows for process improvement by generating SPC alarms when shifts occur. These methods are suitable when the process is subject to infrequent random shocks, as often encountered in the microelectronics industry.

The proposed SPC/EPC methods are easy to use on the manufacturing floor since 1) they are based on standard SPC charts, and 2) the adjustment rule is extremely simple, allowing even for manual implementation.