



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

**Nonlinear
Analysis**

Real World Applications

Nonlinear Analysis: Real World Applications 6 (2005) 817–844

www.elsevier.com/locate/na

Economic statistical design of adaptive control schemes for monitoring the mean and variance: An application to analyzers

Zachary G. Stoumbos^{a,*},¹, Marion R. Reynolds Jr.^b

^a*Department of Management Science and Information Systems, Rutgers Business School, Rutgers University, 94 Rockefeller Road, Piscataway, NJ 08854-8054, USA*

^b*Departments of Statistics and Forestry, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0439, USA*

Received 17 May 2000; accepted 4 May 2005

Abstract

Several individuals control chart schemes are contrasted for the problem of monitoring the mean and variance of a normal process variable, with special consideration given to monitoring process analyzers, such as electrochemical devices, chromatographs, potentiometers, refractometers, and spectrometers. The combination of the exponentially weighted moving average (EWMA) chart and the Shewhart X chart that uses a variable sampling interval (VSI) policy is shown to be very effective for this problem. We develop a comprehensive economic model for the design of control schemes based on this chart combination. The economic model expresses the long-run cost per time unit of operating the combined VSI EWMA and VSI X chart scheme as a function of its design parameters, the parameters that describe the behavior of the process, and the cost parameters associated with the operation of the scheme. This economic model can be used to quantify the cost reduction that can be achieved by using the combined VSI scheme instead of traditional control schemes that use fixed sampling rates. We show that the reduction in cost as well as gains in performance are substantial.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Analyzers; Calibration; Economic design; EWMA control charts; Loss function; Markov chains; Shewhart control charts; Standards; Statistical process control; Stochastic optimization; Variable sampling interval control charts

* Corresponding author. Tel.: +1 732 445 6849; fax: +1 732 445 6329.

E-mail address: stoumbos@andromeda.rutgers.edu (Z.G. Stoumbos).

¹ Zachary Stoumbos' work was supported in part by 2002 and 2003 Rutgers Business School Research Grants.

1. Introduction

Statistical process control (SPC) refers to a collection of statistical methods used extensively to monitor and improve the quality and productivity of industrial processes and service operations. SPC primarily involves the implementation of control charts, which are graphical devices widely used to monitor manufacturing processes to quickly detect any change in a process that may affect the quality of the output. The most commonly used types of control charts are the Shewhart charts, proposed by Shewhart [46], the cumulative sum (CUSUM) charts, initially investigated by Page [28], and the exponentially weighted moving average (EWMA) charts, originating in the work of Roberts [40]. Lai [15] and Stoumbos et al. [54] give reviews of developments in the theory and application of control charting methods.

The traditional practice when applying a control chart is to use a fixed sampling rate (FSR) and take samples of fixed size at fixed sampling intervals. In recent years, however, it has been shown that the statistical performance of control charts can be greatly improved by varying the sampling intervals or sample sizes as a function of the data taken from the process. The basic idea is that sampling should be more intense whenever there is an indication of a problem with the process and less intense when there is no such indication. These adaptive or variable sampling rate (VSR) control charts are more efficient than traditional FSR charts in that they provide faster detection of process changes for a given average sampling rate. The increased efficiency of VSR charts can also be used to reduce the sampling effort necessary to provide a required detection capability. For a survey of work on adaptive control charts, see Tagaras [56].

Process analyzers, such as electrochemical devices, gas and liquid chromatographs, potentiometers, refractometers, and spectrometers, constitute a vital component of the electronics, chemical, petrochemical, polymer, and semiconductor industries. Monitoring the performance of analyzers so to ensure that they are calibrated, validated, and operating as expected is critical in meeting international standardization requirements, such as those set forth by the American Society for Testing and Materials (ASTM), the American National Standards Institute (ANSI), and the International Organization for Standardization (ISO). For additional information on process analyzers and standards, see [5,6,18,27].

The objective of this paper is to develop a very effective, adaptive control chart scheme for monitoring the mean μ and variance σ^2 of a normal random variable, with special consideration given to the problem of monitoring the performance of process analyzers. The paper is divided into two major parts. In the first part, several individual FSR and VSR control charts are reviewed and contrasted for the problem of monitoring μ and σ^2 . The combination of the EWMA chart and the Shewhart X chart that uses a variable sampling interval (VSI) policy is shown to be very effective for this problem from the perspective of statistical performance.

In the second part, a comprehensive economic model is developed for the design of combined VSI EWMA and VSI X control chart schemes. The economic model expresses the long-run cost per time unit of operating the combined scheme as a function of its design parameters, the parameters that describe the behavior of the process, and the cost parameters associated with the operation of the scheme. This economic model can be used to quantify the cost reduction that can be achieved by using the combined VSI scheme, instead of

traditional control schemes that use fixed sampling rates. We show that the reduction in cost as well as gains in statistical performance are substantial and illustrate these facts with numerical results and a detailed example.

Although there is extensive literature on the economic design of traditional FSR control charts (see [12,14,23], and numerous references therein), very little work has been done on the economic modeling and design of adaptive control charts (see [7,29,30]). This is especially true for the more complex and generally more realistic problem of simultaneously monitoring μ and σ^2 that we investigate here, with special consideration given to monitoring the performance of process analyzers.

2. Development of control schemes and statistical performance evaluation

2.1. The relationships between different types of adaptive control schemes

The output of a process analyzer is usually represented by a normal random variable X . The mean μ (level of bias) and variance σ^2 (level of precision) of X are commonly used to determine the performance of an analyzer. This performance can be measured during several different types of evaluation, such as calibration, validation, and comparison with laboratory results. The way in which an observation is taken depends on the type of evaluation, but for the common types of evaluation a single observation is typically taken at each sampling point. In practice, the minimum time before another analyzer observation can be taken is determined by scheduling considerations and is usually on the order of one day. The process observations taken at each sampling time will be assumed here to be independent. The control scheme often used to monitor the mean μ of an analyzer is a combination of a Shewhart chart of the individual observations (X chart) used together with an overlay of an EWMA chart for μ . A Shewhart moving range chart (MR chart) is often used to monitor the variance σ^2 of an analyzer.

In general, the VSR idea can be applied to monitor the mean μ and variance σ^2 of a process using several different adaptive approaches. One approach is to use a VSI chart, which varies the time interval between samples as a function of the process data. A second approach is to use a variable sample size (VSS) chart, which varies the size of the samples taken. A third approach is to apply a sequential probability ratio test (SPRT) at specified sampling times. These adaptive approaches can also be used simultaneously in various combinations, depending on the practical sampling limitations associated with the application under consideration (see [56] and references therein).

A VSS control chart would generally not be appropriate for monitoring analyzer performance. This type of adaptive control chart assumes that samples of size $n > 1$ can be taken at some of the sampling points. With advance notice, it might be possible to obtain more than one observation at a sampling point, but it is not clear that this could generally be justified. Thus, if it is assumed that a single observation will be obtained at each sampling point, then a VSS chart could not be applied to monitor analyzer performance.

Ideally, for an SPRT or generalized SPRT (GSPRT) control chart to be at its best, a sample of more than one observation should also be taken at each sampling point (see [34,49–51,53]). If a relatively long time period is required to obtain each observation or

sample, then the performance advantage in applying an SPRT or GSPRT chart may be reduced in that the SPRT-based chart would behave and perform similarly to a VSI control chart (see [49,53]).

With a single observation at each sampling time, a VSI control scheme appears to be the best way to gain the advantages of the VSR feature. After each observation is obtained, a decision is made about the time duration until the next observation. If the current observation is close to the target, then the decision will be to wait a relatively long time until the next observation. On the other hand, if the current observation is not close to the target (but still within the control limits), then the decision will be to take the next observation as soon as is feasible. The integration of the adaptive, VSI feature into the combination of the Shewhart X chart and the EWMA chart will be discussed after the issues of optimality and of monitoring the process variance have been considered.

2.2. *The relationship of optimality results to the process monitoring problem*

There are a number of optimality results associated with statistical tests and process monitoring. A question of interest is how these optimality results relate to the process-monitoring problem being considered here. For the problem of hypothesis testing, the optimal test of a simple null hypothesis versus a simple alternative hypothesis is the SPRT (see [47]). This means that among all tests that have a given probability of a type I error and a given probability of a type II error, the SPRT will, on average, require the fewest number of observations to reach a decision.

The application of a CUSUM control chart is equivalent to applying a sequence of SPRTs. Each of these SPRTs in the CUSUM chart tests the null hypothesis that the process parameter being monitored is at its in-control or target value against the alternative hypothesis that the process parameter has shifted to a specified out-of-control value. When an SPRT accepts the null hypothesis that the process parameter is on target, then the CUSUM chart starts another SPRT at the next sampling point. However, if an SPRT rejects the hypothesis that the parameter is on target, then this is taken as a signal generated by the CUSUM chart. The FSR CUSUM chart is the optimal FSR control chart in the sense that it provides the fastest detection of a specified process shift among all control charts that have the same long-run average false-alarm rate (or simply, false-alarm rate) when the process is on target.

Most optimality results for adaptive control charts are concerned with the optimal choice of sampling intervals in VSI charts. It has been shown that an optimal choice of sampling intervals in a VSI control chart is to use only two possible sampling intervals (see [31,32,48]). This choice is best in the sense that it minimizes the average time required to detect a specified process shift, for a given false-alarm rate and a given in-control average sampling rate.

As developed by Stoumbos and Reynolds [50], the SPRT chart applies an SPRT to the individual observations taken at each sampling point, employing a fixed sampling interval (FSI) between the sampling points. The fact that an SPRT is the optimal test implies that the SPRT chart is the optimal adaptive FSI control chart. The SPRT chart is optimal in the sense that it minimizes the average time required to detect a given process shift among all control charts that have the same false-alarm rate and the same in-control average sample size at each sampling point. However, as discussed in the previous section, the SPRT chart

does not generally apply to the analyzer-monitoring problem, because a single observation is usually taken from an analyzer at each sampling point.

For the case of a single observation taken at each sampling point, it appears that an optimal adaptive control chart would be a VSI CUSUM chart or a GSPRT chart. However, the VSI EWMA chart of the process observations will be considered in this paper. It has been shown that an EWMA chart has approximately the same ability to detect process shifts as a CUSUM chart (see, for example, [37,43]). Thus, it appears that the VSI EWMA chart will have statistical properties that are very close to the best attainable properties. In the context of FSI charts, Lucas and Saccucci [20] and Stoumbos et al. [55] reported that the properties of the EWMA and CUSUM charts were close enough that a choice between them could be based on other considerations, such as ease of interpretation. The EWMA control chart is naturally two-sided and perceived to be easier to interpret by many practitioners.

2.3. Monitoring the process variance

In SPC, the traditional approach to monitoring the process variance is based on the ranges of the samples taken from the process. When only one observation is taken at each sampling point, it is not possible to use ranges computed within the samples, so the moving range (*MR*) of two successive individual observations is frequently used. In particular, the Shewhart *MR* chart is based on plotting the control statistic

$$R_k = |X_k - X_{k-1}|, \quad k = 2, 3, \dots,$$

where X_k is the observation taken at sampling point k , $k = 1, 2, \dots$. A signal is generated that the standard deviation σ has increased if R_k exceeds the upper control limit

$$h_R \sigma_0,$$

where σ_0 is the in-control value of σ . In most applications, the chart parameter h_R is determined to yield specified statistical properties when the process is in control. A lower control limit can be introduced if it is desirable to detect decreases in σ (see [35,36]).

For the problem of monitoring the process mean μ and variance σ^2 , a number of researchers: Nelson [26], Roes et al. [41], Rigdon et al. [39], Albin et al. [2], Amin and Ethridge [3], Stoumbos and Reynolds [52], and Reynolds and Stoumbos [35–37] have argued that there is essentially no advantage to using the Shewhart *MR* chart with the Shewhart *X* chart. The *MR* chart will be considered in this section because it is the traditional control chart for monitoring σ^2 and is currently in widespread use. It will be argued below that the *X* chart is better than the *MR* chart for detecting changes in σ^2 .

The Shewhart *X* chart is based on plotting X_k versus k , $k = 1, 2, \dots$, and a signal is given at sampling point k if X_k falls outside of control limits constructed at

$$\mu_0 \pm h_X \sigma_0,$$

where μ_0 denotes the target value for μ . In practical applications, the chart parameter h_X is usually taken to be equal to 3, to give the standard “three-sigma” control limits.

From the definition of its control limits, it follows that the Shewhart *X* chart signals at sampling point k if $|X_k - \mu_0| > h_X \sigma_0$. That is, the *X* chart is equivalent to a control chart that

signals at point k if $(X_k - \mu_0)^2 > h_X^2 \sigma_0^2$. The statistic $(X_k - \mu_0)^2$ has a nice interpretation as the squared deviation of the observation from the target. For the problem of monitoring σ , this statistic can also be justified from the statistical theory of hypothesis testing. When a single observation is available from a normal distribution with specified mean μ_0 , the uniformly most powerful test for the variance is based on the statistic $(X_k - \mu_0)^2$ (see [17]). Thus, it follows that the best individual Shewhart control chart for detecting a shift in the process variance is the control chart based on $(X_k - \mu_0)^2$, and this chart is equivalent to the X chart. That is, the X chart is the best Shewhart chart for detecting a shift in σ , even though the X chart is usually regarded as a control chart for detecting shifts in μ . The superiority of the X chart over the MR chart will be illustrated with some numerical results given below.

The performance of a control chart is traditionally evaluated using the average run length (ARL). The ARL is the expected number of samples taken until the chart generates a signal. When the process is in control, the ARL should be large so that the rate of false alarms is low. When the process changes to an out-of-control state, the ARL should be small so that this out-of-control state is quickly detected. For example, when the process is in-control, a Shewhart X chart with three-sigma limits has an ARL of 370.4. This means that if the process remains in control, the false-alarm rate will be one false alarm in every 370.4 samples, which in the case of the X chart would be individual observations. This can be expressed as a false-alarm rate of $1/370.4 = 0.0027$ false alarms per sample. On the other hand, if the process mean μ shifts from μ_0 to $\mu_0 + 2\sigma_0$, then the ARL of the X chart is 6.3. This means that it will take on average 6.3 samples to detect this shift in μ . If the process standard deviation σ increases from the in-control value of σ_0 to $2\sigma_0$, then the ARL of the X chart is 7.5. Thus, on average, it will take 7.5 samples to detect a 100% increase in the process standard deviation.

When attempting to choose one of several different control chart schemes to use in a particular application, it is useful to determine which scheme will be fastest at detecting process changes that are of interest. In this case, the control chart schemes can be compared by adjusting their individual control limits so they all have the same in-control ARL. Then, the ARLs can be compared for various out-of-control situations to see which chart or chart combination will be faster at detecting these out-of-control situations. The in-control ARLs of the control schemes should be matched to be the same to ensure that the schemes have the same false-alarm rates, so that a fair comparison of their out-of-control ARL performance can be made.

In order to compare the performance of the X chart to that of the MR chart for detecting shifts in σ , the control limits of the MR chart were adjusted to give the same in-control ARL of 370.4 as the X chart. Columns 2 and 3 of Table 1 give ARL values for the Shewhart X and MR charts for a wide range of values of σ/σ_0 . The ratio σ/σ_0 expresses the units-free size of a shift in σ , with $\sigma/\sigma_0 = 1$ corresponding to the in-control case $\sigma = \sigma_0$. These ARL values were computed as described in [35,52]. The results in Table 1 show that the X chart will detect shifts in σ faster than the MR chart. Thus, when the X chart is being used there is no good reason to add the MR chart for purposes of detecting increases in σ .

When the X chart is applied together with an EWMA chart, the EWMA chart will be sensitive to shifts in μ and the X chart will be sensitive to shifts in σ . The X chart will also be sensitive to large shifts in μ . Thus, the question arises of how to determine whether a signal by the X chart is due to a shift in μ , a shift in σ , or possibly a shift in both μ and σ .

Table 1
 ARL values with chart parameter values for Shewhart and EWMA charts for monitoring σ

Size of shift σ/σ_0	Shewhart		EWMA-SDT	
	X chart	MR chart	$\lambda = 0.1$	$\lambda = 0.2$
1.00	370.4	370.4	370.4	370.4
1.50	22.0	25.2	14.4	16.0
2.00	7.5	9.3	6.0	6.1
2.50	4.3	5.6	3.9	3.9
3.00	3.2	4.1	3.0	2.9
5.00	1.8	2.3	1.8	1.8
8.00	1.4	1.6	1.4	1.4
h_X	3.000	—	—	—
h_R	—	4.215	—	—
h_S	—	—	3.432	4.112

In the X chart, the pattern of points produced by a shift in μ would tend to be different from the pattern produced by a shift in σ . In particular, a shift in μ will result in observations that tend to fall on one side of μ_0 , while an increase in σ will result in observations that tend to fall on both sides of μ_0 , but at an increased distance from μ_0 . Thus, the pattern of points on the X chart can be used as a graphical diagnostic aid. For extensive discussions on such diagnostic issues, including discussions based on probabilistic arguments, see [35,37].

The X chart is a Shewhart chart, and thus it will not be as sensitive to small shifts in μ or σ as a CUSUM or EWMA control chart. If the X chart is used with an EWMA chart of the observations, then this combination will be effective for detecting small shifts in μ . However, if it is important to detect quickly small shifts in σ , then a CUSUM or EWMA statistic could be considered (see, for example, [37,38]). One such statistic for σ is the EWMA of the squared deviations from target (EWMA-SDT) (see, for example, [9,22,45]). In particular, the upper one-sided control statistic of the EWMA-SDT chart for detecting increases in σ can be written as

$$S_k = (1 - \lambda) \max\{S_{k-1}, \sigma_0^2\} + \lambda(X_k - \mu_0)^2, \quad k = 1, 2, \dots,$$

where λ is a smoothing parameter satisfying $0 < \lambda \leq 1$ and the starting value is usually $S_0 = \sigma_0^2$. A signal is given at sample k if S_k exceeds an upper control limit set at $\sigma_0^2 + h_S \sigma_0^2 \sqrt{2\lambda/(2 - \lambda)}$, where h_S is the chart parameter that determines the distance of the control limit from σ_0^2 .

Some ARL values were computed for two matched EWMA-SDT charts with $\lambda = 0.1$ and 0.2, respectively, and are given in columns 4 and 5 of Table 1. These ARL values were computed as described in [35,52]. Compared to the X chart, the EWMA-SDT chart is faster at detecting increases in σ , especially for small shifts. Thus, if the X chart is used with an EWMA chart, as is often the case in monitoring analyzer performance, the question arises of whether also adding the EWMA-SDT chart would be worthwhile. When another chart

is added to a monitoring scheme, there will always be an increase in the false-alarm rate. If such an increase can be tolerated in exchange for faster detection of small shifts in σ , then in many cases one could “tighten” the control limits of the X chart by a small amount to improve its ability to detect shifts in σ . Thus, it appears that the addition of a specific third chart for detecting small shifts in σ , such as the EWMA-SDT chart, may not be necessary in many practical situations. For extensive investigations of the properties of EWMA-STD charts in the context of the monitoring μ and/or σ , see [35–37,52].

2.4. Combining the X chart with the EWMA chart

The control chart scheme that we will consider here for monitoring analyzer performance is the combination of the X chart and EWMA chart of the observations implemented with the VSI feature. The VSI feature can be used with one of the charts or with both charts. The use of the VSI feature with the EWMA chart will be investigated first. Then, the use of the VSI feature with both charts will be considered. Using the VSI feature with the X chart alone does not appear to be the best use of this feature, so this possibility will not be investigated.

Applying the X chart and EWMA chart together means that after an observation is taken, this individual observation is plotted on the X chart and the weighted average of the observation and past observations is plotted on the EWMA chart. If desired, one of these control charts can be superimposed on the other so that the information in the two charts can be displayed concurrently. The control statistic plotted on the EWMA chart is

$$Y_k = (1 - \lambda)Y_{k-1} + \lambda X_k, \quad k = 1, 2, \dots,$$

where λ is a smoothing parameter with $0 < \lambda \leq 1$ and the starting value is usually taken to be $Y_0 = \mu_0$. The EWMA chart has control limits at $\mu_0 \pm h_E \sigma_Y$, where $\sigma_Y = \sigma_0 \sqrt{\lambda / (2 - \lambda)}$ is the asymptotic in-control standard deviation of Y_k , and h_E is the chart parameter that determines the distance of the control limits from μ_0 . A number of papers in the last 20 years have considered the design and implementation of EWMA control charts (see, for example, [8,11,20,25,55,57]). When the X chart and the EWMA chart are used together, a signal is given if X_k falls outside of $\mu_0 \pm h_X \sigma_0$, or if Y_k falls outside of $\mu_0 \pm h_E \sigma_Y$.

2.5. Applying the VSI feature to the combination of the EWMA and X charts

Suppose that the EWMA chart and X chart are used together and consider first the situation in which the VSI feature is applied only to the EWMA chart. After observation k is taken, the time until the next observation is taken depends on the position of the point Y_k plotted on the EWMA chart. As discussed above, it is recommended that two possible sampling intervals, one short and the other long, be used. Let d_1 represent the short sampling interval and d_2 the long sampling interval, where $0 < d_1 \leq d_2$. For example, if the minimum time required to obtain another observation on an analyzer is one day, then d_1 could be one day and d_2 might be seven days.

The decision rule for the VSI feature is based on the distance of Y_k from the target μ_0 . The basic idea is that if Y_k is close to μ_0 , then the next observation should be taken in d_2 time units, but if Y_k is not close to μ_0 (but still within the control limits), then the next

observation should be taken in d_1 time units. To make this sampling decision rule precise, two additional limits are set at $\mu_0 \pm g_E \sigma_Y$, where $g_E < h_E$. Then, the decision rule is to use the long sampling interval d_2 if Y_k falls in $\mu_0 \pm g_E \sigma_Y$, and use the short sampling interval d_1 if Y_k falls between $\mu_0 - h_E \sigma_Y$ and $\mu_0 - g_E \sigma_Y$ or between $\mu_0 + g_E \sigma_Y$ and $\mu_0 + h_E \sigma_Y$.

The sampling rule described above bases the decision about the next sampling interval on the value of the EWMA statistic Y_k , so that the VSI feature is used only in the EWMA chart. However, as discussed previously, it seems reasonable that the VSI feature could also be used in the X chart. This means that the decision about the next sampling interval would also depend on the value of the statistic X_k plotted on the X chart. When it is desirable to also add the VSI feature to the X chart, a reasonable sampling rule must be specified for combining the VSI X chart and the VSI EWMA chart into a single control scheme.

When the VSI feature is used with the X chart, two additional limits are constructed at $\mu_0 \pm g_X \sigma_0$, where $g_X < h_X$. Then, the decision rule is to use the long sampling interval d_2 if X_k falls in $\mu_0 \pm g_X \sigma_0$, and use the short sampling interval d_1 if X_k falls between $\mu_0 - h_X \sigma_0$ and $\mu_0 - g_X \sigma_0$ or between $\mu_0 + g_X \sigma_0$ and $\mu_0 + h_X \sigma_0$.

When the VSI EWMA chart is combined with the VSI X chart, then after each observation, both charts will specify a sampling interval to use next. A reasonable decision rule to use in this case is to use the long interval d_2 if both charts specify d_2 , and use the short sampling interval d_1 if either one (or both) of the charts specifies d_1 . This means that d_1 is used if either Y_k or X_k is not reasonably close to the target μ_0 . For given values of g_E and g_X , this decision rule results in d_1 being used more often than it would be if the VSI feature was used with only one of the two charts. Thus, when using the VSI feature with both charts it may be necessary to increase g_E and g_X somewhat to avoid using d_1 too often when the process is in control.

2.6. Evaluating the statistical properties of combined VSI EWMA and VSI X charts

The ability of a control chart to detect special causes is typically evaluated by how fast the chart detects the special causes after they occur. When a control chart uses a fixed-length sampling interval, the expected time until the chart signals is the product of the sampling interval and the ARL. However, when a VSI chart is being used, the expected time for the chart to signal cannot be determined using just the ARL. Thus, for VSI charts, a separate, direct measure of signal delay must be used. Let the average time to signal (ATS) be the expected length of time from the start of process monitoring until a signal is generated.

The ATS provides a measure of the time required to detect a parameter change when the change is present at the start of process monitoring. When the process is out of control, there are two problems with using the ATS as a measure of signal-delay performance. One problem is that the change may occur at a random time in the future, after the process has been running for some time. In this case, the control statistic Y_k of the EWMA chart at the time of the change will usually not be at its starting value Y_0 . A second problem is that a change that takes place in the future may occur somewhere within a sampling interval between successive observations.

One approach to a more realistic computation of the ATS is based on the assumption that a control statistic has reached its steady-state or stationary distribution (conditional on no false alarms) by the time of the process change. The ATS computed under this assumption and

from the random point in time that the change occurs is called the steady-state ATS (SSATS). The SSATS allows for the possibility that the process change can occur between successive sampling times. As in previous work (see, for example, [32,33,48,49]), we assume that when a change occurs within a particular sampling interval, the time of the change is uniformly distributed over the interval.

For an individual Shewhart X chart, the ATS and SSATS can simply be expressed in terms of probabilities involving the normal distribution. For more complex charts such as the EWMA chart, numerical methods based on modeling the control statistic as a Markov process, such as the integral equation method or the Markov chain method, can usually be used. The combined scheme based on using the EWMA chart together with the X chart can also be modeled as a Markov process. When applicable, the integral equation method offers higher accuracy for the same computational effort and is easy to apply to the VSI EWMA chart alone. However, when the EWMA chart is combined with the X chart, the transition density of the resulting Markov process is not continuous, and this causes problems with the application of the integral equation method. Thus, the best method for evaluating the ATS and SSATS for the schemes of interest here is the Markov chain method.

The Markov chain method for finding the properties of the combined VSI EWMA and VSI X charts is based on discretizing the possible values for the EWMA statistic Y_k . Each discrete value of Y_k corresponds to a state of the Markov chain. The accuracy of the Markov chain method depends on the number, say r , of states used. In general, the larger the value of r the higher the accuracy, but large values of r require more computational effort. For detailed discussions of the ATS and SSATS and their evaluation, see [32,33,48,49].

2.7. The accuracy of the Markov chain method for evaluating statistical properties

In applying the Markov chain method for evaluating the statistical properties of control charts, the number of states r in the Markov chain that are necessary in order to achieve an acceptable level of accuracy must be determined. Thus, we will next investigate the effect of the choice of r .

As an example of the effect of r , consider a combination of the VSI EWMA and FSI X charts $h_X = 3.0$ in the X chart. This choice of three-sigma limits will give an in-control ARL of 370.4 when the X chart is used alone. For the VSI EWMA chart, consider the parameters $h_E = 2.7015$, $g_E = 1.5704$, $\lambda = 0.1$, $d_1 = 0.1$, and $d_2 = 1.1$. The parameters $h_E = 2.7015$ and $\lambda = 0.1$ will give an in-control ARL of 370.4 when the EWMA chart is used alone. The choice of $g_E = 1.5704$ will insure that when the VSI EWMA chart is used alone and the process is in control, the short sampling interval d_1 will be used only 10% of the time, and the average sampling interval will be 1.0 time unit. That is, the in-control ATS of the VSI EWMA chart used alone will be 370.4 time units. These properties of the VSI EWMA chart were evaluated using the highly accurate integral equation method. Note that, in general, a time unit could be any appropriate length of time, such as an hour, a workshift, a day, or even a week.

The properties of the above control scheme that uses the VSI EWMA chart in combination with the FSI X chart cannot be determined from the properties of the individual charts used alone. For this control scheme and eight choices of r , Table 2 gives in-control ARL, in-control ATS, and out-of-control SSATS values for $\delta = 0.5$, where $\delta = |(\mu - \mu_0)/\sigma_0|$ denotes

Table 2

The effect of the number of states r on the computation of the ARL, ATS, and SSATS of the combined VSI EWMA and FSI X charts

r	ARL when $\delta = 0$	ATS when $\delta = 0$	SSATS when $\delta = 0.5$
11	176.3680	173.0352	19.4633
21	191.9770	191.1533	19.4663
31	195.2933	195.1078	19.4685
41	196.5570	196.6170	19.4716
61	197.5251	197.7791	19.4765
101	198.0571	198.3948	19.4789
201	198.2670	198.6368	19.4793
399	198.3233	198.6984	19.4795

the size of a standardized shift in μ . Note that $\delta = 0$ corresponds to the in-control case $\mu = \mu_0$.

In Table 2, the ARL, ATS, and SSATS values stabilize as r becomes larger. In general, the Markov chain method tends to be least accurate when λ is very small and g_E is very large (see [32]). In many practical applications, the value of λ would not be chosen much below the value 0.1 used here, and the value of g_E would not be above the value used here. Thus, for other choices of λ and g_E , the accuracy of the Markov chain method should not be worse than for the case considered here. In fact, for many applications, a value of r between 31 and 61 would give sufficient accuracy for constructing an FSI or VSI control scheme based on the EWMA and X charts.

In Table 2, the ARL, ATS, and SSATS values increase toward respective asymptotic values, as r increases. This suggests using the calculated values to fit curves that can be used to estimate the asymptotic values. We show below that the curves can be fitted using relatively small values of r , so that it is possible to get very accurate estimates of the ARL, ATS, and SSATS without using large values of r . To explain this method in more detail, let $ARL(r)$ be the ARL calculated for a specific value of r (the same idea applies to the ATS and SSATS). Then, calculate $ARL(r)$ for several values of r and use regression to fit the model

$$ARL(r) = \beta_0 + \beta_1 \frac{1}{r^2} + \beta_2 \frac{1}{r^4}, \tag{1}$$

where β_0 , β_1 , and β_2 are the regression coefficients. This model will generally fit well, so that letting $r \rightarrow \infty$, we obtain $ARL(\infty) = \beta_0$. That is, the β_0 from the regression model is an estimate of the ARL for an infinite number of states. The effectiveness of the regression model in (1) is illustrated next for the evaluation of the ARL, ATS, and SSATS of the combined VSI EWMA and FSI X Charts.

Consider the results given in Table 3, where the last row of this table gives the values for $r = 399$ from Table 2. The results in Table 3 show that very accurate ARL, ATS, and SSATS values can be obtained using a relatively small number of states and the regression model in Eq. (1). The relatively small disadvantage to this approach is that the quantities of interest must be calculated for several values of r , which slightly increases the complexity

Table 3

The effect of using the regression model in Eq. (1) to compute the ARL, ATS, and SSATS of the combined VSI EWMA and FSI X charts

Values of r	ARL when $\delta = 0$	ATS when $\delta = 0$	SSATS when $\delta = 0.5$
15, 19, 23, 27	198.5485	197.6949	19.4247
17, 25, 33, 41	198.2138	198.5408	19.4671
23, 29, 35, 41	198.2217	198.1796	19.4458
25, 29, 33, 37, 41	198.4807	198.1638	19.4443
11, 21, 31, 41, 51	198.2268	198.6233	19.4748
21, 31, 41, 51, 61	198.3238	198.7349	19.4799
61, 81, 101, 121, 141	198.3293	198.6730	19.4753
399	198.3233	198.6984	19.4791

of the operation.

The ARL, ATS, and SSATS values in Tables 2 and 3 were presented for purposes of accuracy assessment. Their interpretation in terms of process monitoring is as follows. When the process is in control, a false alarm will occur on average about every 198 samples. This corresponds to an average of about 198 time units between false alarms (recall that the chart parameters were chosen to give an average sampling interval of 1.0 time unit when the process is in control). If either the VSI EWMA chart or the FSI X chart had been used alone, the false-alarm rate for either one of these charts would have been one false alarm in 370.4 time units. This means that combining the two charts into a single control scheme will result in an increase in the expected rate of false alarms. The SSATS of the scheme is the expected detection time for a process shift. In particular, when the process mean shifts by one half of the process standard deviation ($\delta = 0.5$), it will take on average about 19.5 time units to detect this shift.

2.8. Comparisons of combined FSI and VSI EWMA and X charts

In trying to decide which process-monitoring scheme to recommend for practical applications, several questions arise. One question concerns the benefit that can be gained by using a VSI scheme instead of an FSI scheme. Another question is whether it is better to use the VSI feature with only the EWMA chart or with both the EWMA chart and the X chart. A third question concerns the choice of chart parameters, such as the λ , d_1 , d_2 , h_E , g_E , h_X , and g_X . In this section, we will investigate these questions, allowing the VSI feature to be used in both the EWMA chart and the X chart.

Table 4 gives in-control ATS values and out-of-control SSATS values for ten FSI and VSI control schemes that are combinations of an EWMA chart and an X chart. The out-of-control SSATS values are given for six values of $\delta = |(\mu - \mu_0)/\sigma_0|$ (corresponding to shifts in μ with $\sigma = \sigma_0$) and for six values of σ/σ_0 (corresponding to shifts in σ with $\mu = \mu_0$). The chart parameter values h_E , g_E , h_X , and g_X are given in the last four rows of Table 4. The regression method for the model in Eq. (1) and $r = 21, 31, 41, 51,$ and 61 Markov chain

Table 4

In-control ATS and out-of-control SSATS values with chart parameter values for FSI and VSI combinations of EWMA and X charts

		FSI EWMA FSI X chart	VSI EWMA VSI X chart	VSI EWMA VSI X chart	VSI EWMA FSI X chart	VSI EWMA VSI X chart	FSI EWMA FSI X chart	VSI EWMA FSI X chart	VSI EWMA VSI X chart	VSI EWMA FSI X chart	VSI EWMA VSI X Chart
Size of shift		$\lambda = 0.1$	$\lambda = 0.1$	$\lambda = 0.1$	$\lambda = 0.1$	$\lambda = 0.1$	$\lambda = 0.2$	$\lambda = 0.2$	$\lambda = 0.2$	$\lambda = 0.2$	$\lambda = 0.2$
δ	σ/σ_0	—	$d_1 = 0.1$ $d_2 = 1.1$	$d_1 = 0.1$ $d_2 = 1.1$	$d_1 = 0.1$ $d_2 = 1.6$	$d_1 = 0.1$ $d_2 = 1.6$	—	$d_1 = 0.1$ $d_2 = 1.1$	$d_1 = 0.1$ $d_2 = 1.1$	$d_1 = 0.1$ $d_2 = 1.6$	$d_1 = 0.1$ $d_2 = 1.6$
		$d = 1.0$	$d = 1.0$	—	$d = 1.0$	—	$d = 1.0$	$d = 1.0$	—	$d = 1.0$	—
0.00	1.00	198.32	198.73	198.74	199.78	199.78	203.88	204.13	204.13	204.48	204.49
0.50	1.00	25.35	19.48	21.53	15.38	17.07	32.32	26.26	28.24	21.21	23.22
1.00	1.00	8.52	5.75	6.21	4.66	4.78	8.70	5.64	6.20	4.31	4.65
1.50	1.00	4.67	3.21	3.11	2.72	2.45	4.31	2.73	2.77	2.23	2.19
2.00	1.00	2.94	2.14	1.82	1.90	1.52	2.65	1.77	1.61	1.52	1.39
3.00	1.00	1.28	1.11	0.80	1.15	0.90	1.21	0.97	0.77	0.96	0.88
5.00	1.00	0.52	0.56	0.55	0.77	0.77	0.52	0.56	0.55	0.75	0.77
0.00	1.50	17.09	16.43	15.16	16.19	13.66	16.39	15.35	14.42	14.64	12.77
0.00	2.00	6.24	5.97	5.20	6.08	4.75	6.02	5.55	4.97	5.38	4.45
0.00	2.50	3.59	3.47	2.94	3.65	2.81	3.49	3.23	2.84	3.22	2.66
0.00	3.00	2.53	2.47	2.08	2.68	2.07	2.48	2.31	2.02	2.37	1.99
0.00	5.00	1.30	1.31	1.12	1.53	1.27	1.29	1.25	1.11	1.38	1.24
0.00	8.00	0.90	0.94	0.83	1.16	1.02	0.90	0.90	0.83	1.07	1.01
h_E		2.7015	2.7015	2.7015	2.7015	2.7015	2.8593	2.8593	2.8593	2.8593	2.8593
g_E		—	1.5704	1.8713	0.8113	1.1532	—	1.6097	1.8715	0.8276	1.1251
h_X		3.0000	3.0000	3.0000	3.0000	3.0000	3.0000	3.0000	3.0000	3.0000	3.0000
g_X		—	—	1.8713	—	1.1532	—	—	1.8715	—	1.1251

states was used to compute the ATS and SSATS values in order to obtain highly accurate results.

The control limits of the ten control schemes were determined so that all EWMA charts and all X charts have in-control ARLs of 370.4 when considered as individual charts. This means that for all X charts, $h_X = 3.0$, but the value of h_E for each EWMA chart depends on the choice of the smoothing parameter λ . The in-control ARLs of the ten combined schemes are close to 200 (see Tables 2 and 3). The short interval d_1 for the VSI charts was chosen to be 0.1 time units, and the long interval d_2 was chosen to be either 1.1 or 1.6 time units. For each individual VSI EWMA chart, the value of g_E was determined to attain an in-control average sampling interval of 1.0 time unit. Similarly, for each individual VSI X chart, the value of g_X was determined to attain an in-control average sampling interval of 1.0 time unit. This results in an in-control average sampling interval of approximately 1.0 time unit when the EWMA and X charts are used in combination. The parameter λ in the EWMA charts was chosen to be either 0.1 or 0.2. The fixed-length sampling interval d for all FSI control charts was taken to be 1.0 time unit.

The first combined scheme in Table 4 (column 3) is an FSI control scheme with $\lambda = 0.1$ in the EWMA chart. Results for this FSI scheme were included for purposes of comparison. Column 4 contains results for a combination scheme with the VSI feature in the EWMA chart but without the VSI feature in the X chart. This scheme has $\lambda = 0.1$ in the EWMA chart, $d_1 = 0.1$, and $d_2 = 1.1$. Column 5 contains results for a scheme with the same λ , d_1 , and d_2 , as in column 4, but with the VSI feature in both of the charts. For the scheme in column 5, the value of g_X was taken to be equal to g_E and both parameters were increased, so that the in-control ATS is the same as that of the corresponding control scheme in column 4 that does not use the VSI feature in the X chart.

Columns 6 and 7 of Table 4 give SSATS values for the same control schemes as columns 4 and 5 except that $d_2 = 1.6$. The pattern of schemes in columns 3–7 is repeated in columns 8–12, except that $\lambda = 0.2$ in the EWMA charts. The in-control ATS values for the 10 combination schemes in Table 4 are not precisely equal but are close enough that comparisons between the schemes can be made.

Comparing the control schemes in Table 4 with the VSI feature to the corresponding schemes without the VSI feature shows that the VSI feature provides a substantial reduction in the SSATS for small and moderate shifts in μ . For detecting increases in σ , the VSI schemes without the VSI feature in the X chart provide only a modest decrease in SSATS for small and moderate shifts in σ compared to the corresponding FSI schemes. However, if the X chart is used with the VSI feature, then the ability to detect increases in σ is improved considerably. The reason for this is that the X chart is the primary chart for detecting an increase in σ . If the X chart does not use the VSI feature, then the VSI feature will not help much in detecting increases in σ . Using the VSI feature in the X chart also affects the ability to detect shifts in μ ; the time required to detect small shifts is increased slightly, while the time required to detect moderately large shifts is decreased slightly. This occurs because the addition of the VSI feature to the X chart requires that g_E be increased in order to maintain the same in-control average sampling interval. The EWMA chart is the primary control chart for detecting small increases in μ , and this increase in g_E reduces the effectiveness of the EWMA chart.

The conclusions from Table 4 about the VSI feature are that adding the VSI feature to

the EWMA chart enhances the ability to detect shifts in μ . The benefits of also adding the VSI feature to the X chart are improved ability to detect increases in σ and improved ability to detect moderately large shifts in μ . The disadvantages are reduced ability to detect small shifts in μ and somewhat increased complexity of the monitoring scheme. In practical applications, the decision about adding the VSI feature to the X chart will depend on an evaluation of the particular tradeoffs involved.

The results in Table 4 can be used to assess the effects of the choice of λ and d_2 . As expected, using $\lambda = 0.1$ is better for detecting small shifts in μ , while $\lambda = 0.2$ is better for detecting large shifts in μ . However, $\lambda = 0.2$ is slightly better than $\lambda = 0.1$ for detecting increases in σ . This result for σ occurs because the EWMA chart will be most responsive to increases in σ when λ is large, and the increased variability in the observations is not averaged out as much in the EWMA control statistic. The EWMA chart is not the primary chart for detecting increases in σ , so the choice of λ should not be dictated by concerns for monitoring σ .

The results in Table 4 show that for detecting shifts in μ , using $d_2 = 1.6$ is better than using $d_2 = 1.1$, except for very large shifts. The same is true for detecting increases in σ for the charts that use the VSI feature in the X chart. The average sampling interval was taken to be 1.0 when in control, so $d_2 = 1.6$ corresponds to lower values of g_E and g_X than does $d_2 = 1.1$. Lower values for g_E and g_X mean that it will be easier to use the short sampling interval when there is a small shift in μ or σ . The drawback to low values of g_E and g_X is that d_1 will be used more frequently when the process is in control. In some practical situations, this may be undesirable from a psychological point of view, if d_2 is regarded to be the “normal” sampling interval, and d_1 is regarded to be the sampling interval that signifies a potential problem with the process.

3. An economic model for the design of combined VSI EWMA and VSI X control schemes

3.1. Background information

Control charts are designed to detect changes in the system being monitored. It is desirable that a control chart detects changes quickly so that the system does not operate in an undesirable state for a long period of time. It is also desirable that a control chart does not produce a large number of false alarms, because false alarms lead to increased cost and loss of confidence in the control scheme. In addition, it is desirable that the sampling cost associated with operating a control chart be kept to a reasonable level. Designing a control chart scheme for use in a particular application requires finding an acceptable balance between the conflicting requirements of fast detection of system changes, a low false-alarm rate, and a reasonably low sampling rate.

There are three basic approaches that have been used in designing control charts. The first approach is based on heuristics that have been developed from past experience in particular application areas. The second approach uses statistical properties of a control chart, such as the in-control ATS (false-alarm rate) and SSATS, and designs the chart to give reasonable values for these properties. This approach was used throughout Section 2.

The third approach uses statistical properties of a control chart but also explicitly models the costs associated with false alarms, sampling, and failing to detect system changes. This third, economic design approach has a control chart constructed to minimize the long-run average cost associated with the control chart's operation.

An economic model for a control chart involves three different types of parameters. The first type is system parameters describing system behavior, such as the expected time until a system change will occur. The second type is cost parameters, such as the cost of a false alarm. The third type is control chart parameters, such as the values of sampling intervals and the control limits. For a given set of system parameters and a given set of cost parameters, the economic model can be used to determine the optimal set of control chart parameters.

In some cases, the control chart design that is optimal according to an economic model may not be so convenient to use. For example, a sampling interval of 1.873 time units may not be as convenient to use as a sampling interval of 2.0 time units. In other cases, there may be subjective factors associated with the monitoring operation that are not captured in the economic model. In these cases, the system engineers may want to evaluate a set of designs that have been identified as being acceptable. The economic model can be used to assess the costs associated with each design in the set, so that the tradeoffs between costs and subjective factors may be evaluated.

Another approach to accounting for subjective factors for which it is difficult to assign explicit costs is to place constraints on the control chart design in the optimization process. For example, it may be difficult to assess the impact of the loss of confidence due to a control chart that produces too many false alarms. In this case, a constraint could be placed on the false-alarm rate of the chart. Then, the optimization procedure would find the minimal-cost design that satisfies this constraint (see, for example, [24,44]).

3.2. *The general structure of the economic model*

The economic structure of the model that will be used here is analogous to that of some models used in previous economic design studies of control charts (see, for example, [10,12,19,23,24,44]). However, the model must be nontrivially extended to account for the adaptive, VSI feature and the two control charts used in combination to monitor the process mean and variance. Here, we will develop an economic model for the combined VSIEWMA and VSI X charts that were investigated above from the perspective of statistical performance.

The objective of the current study is to develop efficient monitoring procedures for both the process mean μ and the process standard deviation σ . To avoid having to write both μ and σ when referring to the process parameters being monitored, we define the vector θ to be $\theta = (\mu, \sigma)$, and let $\theta_0 = (\mu_0, \sigma_0)$ represent its target value. Suppose that the process starts out with θ at the target value θ_0 and θ remains at this target until a special cause occurs and produces a shift to some other value of θ . The time, say T_0 , until a special cause occurs is a random variable with some specified distribution. The size of the shift in θ that is produced by the special cause is also a random variable with some specified distribution. When the special cause occurs and produces a shift in θ , it is assumed that θ remains at the shifted value until the control chart scheme signals and the special cause is found and removed. Let T_1 be the length of time that θ remains at the shifted value. The distribution of T_1 will depend on how fast the control chart scheme is able to detect shifts in θ produced by the

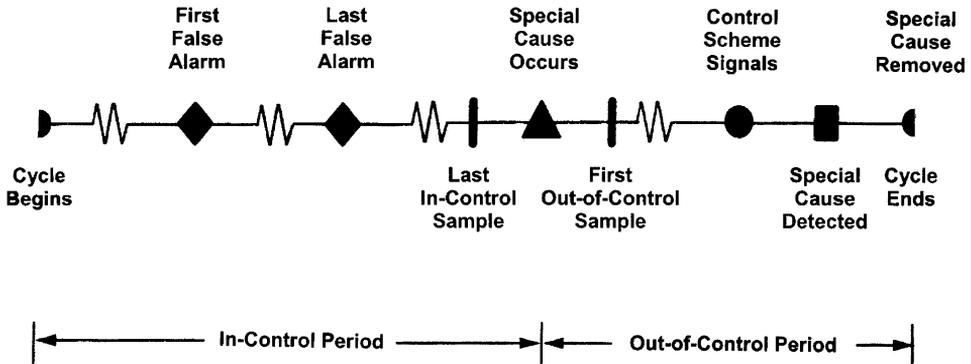


Fig. 1. A diagram of an in-control period followed by an out-of-control period in a process cycle.

special cause, as well as the time required to remove the special cause once it is detected. The time required to remove the special cause may include repair and/or adjustment time. Once the special cause is detected and removed and θ is returned to θ_0 , the process continues at θ_0 until another special cause occurs. The time until this special cause occurs is assumed to have the same distribution as T_0 . The size of the shift produced by this special cause has the same distribution as the previous shift and persists until it is detected and θ is returned to the target. This sequence of alternating in-control periods and out-of-control periods is assumed to continue as the process operates over time (see Fig. 1). Apart from the control chart scheme, the behavior of the process is characterized by the distribution of T_0 , the length of an in-control period, and by the distribution of the shift in θ that can be produced by special causes.

In economic design models, the operation of the control chart scheme is usually viewed as a series of cycles, where a cycle consists of an in-control period followed by an out-of-control period (see Fig. 1). During an in-control period, there are costs due to sampling and false alarms. During an out-of-control period, there are costs due to sampling and operating with θ out of control. The model contains cost parameters that allow the specification of the costs associated with sampling, false alarms, and operating out of control. The long-run average cost per unit time can be obtained as the expected cost per cycle divided by the expected length of a cycle. This is justified by renewal-reward process theory (see, for example, [42]). In particular, this long-run cost per unit time, say L , can be expressed as the ratio

$$L = \frac{L_S + L_F + L_O}{E(T_0) + E(T_1)},$$

where L_S is the expected cost of sampling in a cycle, L_F the expected cost due to false alarms during the in-control period in a cycle, and L_O the expected cost due to producing off-target during the out-of-control period in a cycle.

The expression for L contains five components that are functions of the design parameters of the control chart scheme, the parameters associated with the behavior of the process, and the cost parameters. Once explicit expressions for the components of L are obtained, a numerical procedure can be used to find the values of the control chart design parameters that will minimize L for given values of the process parameters and cost parameters. These values for the chart design parameters are then used to construct a control chart scheme that is optimal in an economic sense for the given values of the process and cost parameters.

The above expression for L is a general expression and additional assumptions about the structure of the model need to be made in order to develop explicit expressions for the five components of L . The model for the behavior of the process will be considered first.

3.3. The process model

It will be assumed that the time until the occurrence of a special cause has an exponential distribution with parameter ρ . This means that the expected time until the occurrence of a special cause is $1/\rho$. Thus, the expected length of the in-control period is

$$E(T_0) = 1/\rho.$$

Suppose that when a special cause occurs it can produce a shift to any one of η possible values of θ represented by $\theta_1, \theta_2, \dots, \theta_\eta$. Each of these possible shifted values of θ could involve a shift in μ , a shift in σ , or a shift in both μ and σ at the same time. When a shift in θ occurs as a result of a special cause, let p_i be the probability that it is a shift to θ_i , for $i = 1, 2, \dots, \eta$. That is, when a special cause occurs, the shift will be to θ_i with probability p_i .

When a special cause occurs, it will eventually be detected by the control chart scheme. After the special cause is detected by the control scheme, assume that R time units are required for repair and adjustment of the process before θ is returned to the target θ_0 .

3.4. The cost parameters

The costs associated with sampling in a cycle will depend on the number of observations taken during the in-control period and during the out-of-control period. Let O_0 be the number of observations taken during the in-control period, and let O_{1i} be the number of observations taken during the out-of-control period when the shift in θ during this out-of-control period is to θ_i . A Markov chain method for determining $E(O_0)$ and $E(O_{1i})$ will be given below in Section 3.5. If

$$c_1 = \text{cost of taking one observation,}$$

then the expected cost of sampling in a cycle will be

$$L_S = c_1 \left[E(O_0) + \sum_{i=1}^{\eta} p_i E(O_{1i}) \right].$$

If F_0 denotes the number of false alarms that occur during an in-control period and each false alarm produces a cost of c_2 , then the expected cost of false alarms in a cycle is

$$L_F = c_2 E(F_0).$$

A Markov chain method for finding $E(F_0)$ will be given in the following section.

The cost associated with operating with θ out of control due to a special cause will depend on the size and type of the shift in θ produced by the special cause. Let c_{3i} be the cost per unit time due to operating out of control because of a shift to θ_i . When there is a shift to θ_i , let T_{1i} be the time from the shift to the signal by the control chart scheme. A Markov chain method for evaluating $E(T_{1i})$ will be given in the next section. The expected cost during an out-of-control period is then

$$L_O = \sum_{i=1}^{\eta} c_{3i} p_i [E(T_{1i}) + R].$$

The expected value $E(T_{1i})$ is the average time required for the control chart scheme to detect a special cause that produces a shift to θ_i . Therefore, it follows that the expected length of an out-of-control period is

$$E(T_1) = \sum_{i=1}^{\eta} p_i E(T_{1i}) + R.$$

Thus, the expected length of a cycle is

$$E(T_0) + E(T_1) = \frac{1}{\rho} + \sum_{i=1}^{\eta} p_i E(T_{1i}) + R.$$

3.5. A Markov chain model for the terms of the economic model

In this section, we develop an exact Markov chain model for the terms needed in the economic model for the combination of the VSI X and VSI EWMA charts. To simplify the notation, the control statistics X_k and Y_k for these two charts will be expressed as the vector $\mathbf{W}_k = (X_k, Y_k)$. The control chart scheme using the two-dimensional control statistic \mathbf{W}_k can be modeled as a Markov chain. The Markov chain model for \mathbf{W}_k is developed by partitioning the regions within the control limits into t subregions. The control statistic is then replaced by a discretized version that can assume only one representative value in each subregion. Let s_i be the representative value for region i , for $i = 1, 2, \dots, t$. Then, state i of the Markov chain corresponds to the discretized version of \mathbf{W}_k that is equal to s_i . In what follows, the notation will not make a distinction between \mathbf{W}_k and its discretized version. An additional state of the Markov chain, state $t + 1$, is needed to correspond to \mathbf{W}_k in the signal region outside of the control limits. Let s_{t+1} be the representative value for this region outside the control limits.

Let \mathbf{W}_0 denote the starting value for the control statistic, which may, for example, be used to start the EWMA before the first observation is sampled. Suppose that the control chart scheme is normally started with $\mathbf{W}_0 = s_{i_0}$, so that i_0 is a potential starting state for the

Markov chain. The first sampling interval for the VSI scheme is of a specified size, say d_0 . For example, in some applications, it would be reasonable to take d_0 to be a short sampling interval, to guard against problems that might be present at startup. The sampling interval b_{i_0} , which is specified by the value $\mathbf{W}_0 = s_{i_0}$, may not be equal to d_0 . In this case, the state i_0 will not serve as the starting state of the Markov chain and another state must be designated or defined for the starting state. This issue will be further discussed later in this section, after some more structure is developed for the Markov chain model.

For $i, j = 1, 2, \dots, t + 1$, let q_{ij} be the transition probability from state i to state j . Assume that after a signal by the control chart scheme, the control statistic is immediately restarted at s_{i_0} . This means that the transition probabilities from state $t + 1$ are the same as from state i_0 ; i.e., $q_{t+1,j} = q_{i_0,j}$, for $j = 1, 2, \dots, t + 1$. Also, let b_i be the sampling interval to be used next, when the control statistic is currently in state i .

If \mathbf{Q} is used to represent the transition probability matrix for the first t states, then many properties of the control chart can be computed in terms of \mathbf{Q} . Let $\mathbf{M} = (\mathbf{I} - \mathbf{Q})^{-1}$, and let m_{ij} be the element (i, j) of \mathbf{M} . If the Markov chain is started in some state i , then the ARL for the control chart is $\sum_{j=1}^t m_{ij}$ and the ATS is $\sum_{j=1}^t m_{ij} b_j$.

To find several of the expectations that are needed in the economic model, it is necessary to know the state of the Markov chain at the sample immediately before the shift in θ . The determination of the probability that the Markov chain is in a particular state immediately before the shift can be done by developing an expanded Markov chain that jointly models the state of the control statistic and the state of θ in the intervals between samples. Let V_k be defined as 1 if $\theta = \theta_0$ both at sample k and in the interval after sample k , as 2 if $\theta = \theta_0$ at sample k but shifts from θ_0 in the interval after sample k , and 3 if $\theta \neq \theta_0$ both at sample k and in the interval after sample k . The expanded Markov chain will be used to model (\mathbf{W}_k, V_k) , as described below.

For $i = 1, 2, \dots, t + 1$, state i in the Markov chain corresponds to $\mathbf{W}_k = s_i$ and $V_k = 1$. When $\mathbf{W}_k = s_i$, the next sampling interval will be b_i and the probability of no shift in θ in this interval is $e^{-\rho b_i}$. For $i = t + 2, t + 3, \dots, 2t + 2$, state i in the Markov chain corresponds to $\mathbf{W}_k = s_i$ and $V_k = 2$. When $\mathbf{W}_k = s_i$, the probability of a shift in the next interval is $1 - e^{-\rho b_i}$. If q_{ij}^* is the transition probability from state i to state j of the Markov chain, then for $i, j = 1, 2, \dots, t + 1$,

$$q_{ij}^* = q_{ij} e^{-\rho b_j},$$

and

$$q_{i,t+1+j}^* = q_{ij} (1 - e^{-\rho b_j}).$$

Also, $q_{ij}^* = 0$, for $i = t + 2, t + 3, \dots, 2t + 2$ and $j = 1, 2, \dots, 2t + 2$, because it is not possible to return to the states corresponding to $\theta = \theta_0$ without a signal by the control scheme. Note that there are additional states in the Markov chain corresponding to $V_k = 3$, but for the purposes here, it is not necessary to explicitly define these states.

If the Markov chain starts in one of the first $t + 1$ states and remains in these states for some period of time, then during this time the value of θ is θ_0 . When the Markov chain moves to one of the states, $t + 2, t + 3, \dots, 2t + 2$, then the shift in θ will occur in the next interval, so the Markov chain will spend one time period in these states. It follows that one of these states will determine the value of \mathbf{W}_k immediately before the shift in θ .

If W_k falls outside of the control limits when $\theta = \theta_0$, then this is a false alarm and the Markov chain is in state $t + 1$. If after a false alarm the chart is restarted at s_{i_0} and d_0 is used as the next sampling interval, then state $t + 1$ corresponds to s_{i_0} and d_0 . Thus, it follows that state $t + 1$ might serve as a starting state for the Markov chain. If there is no shift in θ before the first sample is taken ($V_0 = 1$), then the starting state for the Markov chain can be taken to be $t + 1$. However, if there is a shift in θ before the first sample ($V_0 = 2$), then the starting state for the Markov chain can be taken to be $2t + 2$. Therefore, the Markov chain will start in state $t + 1$ with probability $e^{-\rho d_0}$ and in state $2t + 2$ with probability $1 - e^{-\rho d_0}$.

Let Q^* be the transition probability matrix for the first $t + 1$ states of the Markov chain. Let $M^* = (I - Q^*)^{-1}$ and let m_{ij}^* be element (i, j) of M^* . Let $p_{t+1,t+1+j}$ be the probability that the Markov chain is in state $t + 1 + j$ at the sample immediately before the shift in θ , given that the starting state is $t + 1$. Then, it follows that

$$p_{t+1,t+1+j} = \sum_{i=1}^{t+1} m_{i+1,i}^* q_{i,t+1+j}^* = \sum_{i=1}^{t+1} m_{t+1,i}^* q_{ij} (1 - e^{-\rho b_j}),$$

since $m_{t+1,i}^*$ is the expected number of times in state j , and $q_{i,t+1+j}^*$ is the probability of a transition from state i to state $t + 1 + j$.

To find $E(O_0)$, the expected number of observations in a given in-control period, note that if the starting state is $2t + 2$, then there are no observations before the shift. But if the starting state is $t + 1$, then the expected number of observations is $\sum_{j=1}^{t+1} m_{t+1,j}^*$. Thus, it follows that

$$E(O_0) = e^{-\rho d_0} \sum_{j=1}^{t+1} m_{t+1,j}^*.$$

To find the expected number of false alarms during an in-control period, note that if the starting state is $2t + 2$, then there are no observations before the shift, and thus there can be no false alarms. If the starting state is $t + 1$, then the initial visit to state $t + 1$ should not be counted as a false alarm, but there will be a false alarm every time that the Markov chain returns to state $t + 1$. In addition, there will be a false alarm if the Markov chain goes to state $2t + 2$ after starting in state $t + 1$. Thus, it follows that

$$E(F_0) = e^{-\rho d_0} (m_{t+1,t+1}^* - 1 + p_{t+1,2t+2}).$$

The expected time out of control depends on the starting state for the in-control period and the shift in θ that occurs. For $j = 1, 2, \dots, t + 1$, let $ATS(j, l)$ be the ATS when the control chart starts with $W_0 = s_j$ and with $\theta = \theta_l$. The computation of the expected time out of control depends on $ATS(j, l)$ and on $p_{t+1,t+1+j}$, the probability that $W_k = s_j$ immediately before the shift. If it is assumed that an observation is taken instantaneously, then it follows that the shift must occur in the interval between two observations. Given that the shift occurs in an interval of length d , Duncan [10] has shown that the expected distance into the interval is

$$\tau(d) = \frac{1}{\rho} - \frac{d}{e^{d\rho} - 1}.$$

For $j = 1, 2, \dots, t + 1$, let τ_j be defined by $\tau_j = \tau(b_j)$, so that τ_j is the expected distance from the shift to the previous observation, when the control statistic is in state $t + 1 + j$ at this previous observation. If the shift occurs before the first observation and the Markov chain starts in state $2t + 2$, then this is equivalent to starting the control chart with $\mathbf{W}_0 = s_{t+1}$. In this case, the expected time from the shift to the signal is $\text{ATS}(t + 1, l) - \tau_{t+1}$. If the shift does not occur before the first observation, and the Markov chain starts in state $t + 1$, then the expected time from the shift to the signal is $\sum_{j=1}^{t+1} p_{t+1,t+1+j} [\text{ATS}(j, l) - \tau_j]$. Thus, it follows that the expected time required to detect a shift to θ_l is

$$E(T_{1l}) = (1 - e^{-\rho d_0})[\text{ATS}(t + 1, l) - \tau_{t+1}] + e^{-\rho b_0} \sum_{j=1}^{t+1} p_{t+1,t+1+j} [\text{ATS}(j, l) - \tau_j].$$

The expected number of observations taken during an out-of-control period can be obtained using the same argument as in the derivation of $E(T_{1l})$. In particular, if $\text{ARL}(j, l)$ is the ARL when the control chart starts with $\mathbf{W}_0 = s_j$ and with $\theta = \theta_l$, then the expected number of observations required to detect a shift to θ_l is

$$E(O_{1i}) = (1 - e^{-\rho d_0})\text{ARL}(t + 1, l) + e^{-\rho b_0} \sum_{j=1}^{t+1} p_{t+1,t+1+j} \text{ARL}(j, l).$$

3.6. Optimization of the economic model

As discussed in Section 3.2, in the economic modeling approach, a control chart scheme is designed to minimize the long-run cost per unit time L associated with the operation of the scheme. For a given set of system parameters and a given set of cost parameters, the economic model can be used to determine the optimal vector of control chart parameters $(h_E, g_E, h_X, g_X, \lambda, d_2)$. That is, the economic-model approach involves a multivariate nonlinear optimization problem, which has the form

$$\text{minimize } L(h_E, g_E, h_X, g_X, \lambda, d_2).$$

The short sampling interval d_1 is not involved in the above minimization, because of the established fact that the shortest feasible interval d_1 gives the best performance (see [31,32,48]). Thus, d_1 will be taken to be the shortest feasible sampling interval as determined by administrative considerations and process constraints. For example, in many practical applications of analyzer calibration, d_1 would be on the order of one workshift or one day.

Multivariate search methods for optimization problems are classified into two general categories: Gradient methods and derivative-free methods (see [4,21]). Gradient methods demand function and partial derivative evaluations, while derivative-free methods only employ function evaluations. Gradient methods would be expected to be more efficient in the majority of cases, due to the additional information utilized. If analytical partial derivatives can be derived, however, the question arises of whether a search technique should be used at all. If derivative approximations based on finite difference equations are used, the efficiency of gradient methods should be similar to that of the derivative-free

methods. Gradient methods incorporating derivative approximations usually present some numerical problems in the vicinity of the optimum, where the approximations become very small. Since no closed-form expressions exist for the cost function L or for the statistical properties of the combined VSI EWMA and VSI X charts, a gradient-based minimization procedure must involve derivative approximations with the associated drawbacks.

A general approach to the minimization of L could involve using numerical derivative approximations in a procedure based on the generalized reduced gradient (GRG) method, which was first developed by [1] (see also [16]). However, because of the complexity of the GRG procedure, the implementation of a GRG-based algorithm for the nonlinear minimization of L must involve an interface with a commercial software package. For practical applications where this complication is not desirable and in view of the discussion given in the previous paragraph, the simple but effective derivative-free optimization procedure of Hooke and Jeeves [13] can be used for the design of the combined VSI EWMA and VSI X control chart scheme. The optimization procedure of Hooke and Jeeves [13] was also used by Montgomery et al. [24] in the design of an approximate economic model for an individual FSR EWMA chart for monitoring μ alone.

In some applications, the control scheme design that is optimal according to the unconstrained economic model may be inconvenient to use and/or subjective factors associated with the monitoring operation may exist that are not captured in the economic model. In order to address such potential complications, Saniga [44] considered an economic statistical model for the design of FSR Shewhart \bar{X} and R charts by introducing constraints on certain statistical properties, including the false-alarm rate. Such constraints can be introduced in the above optimization procedure for L by adding a large penalty cost to the objective function L if a certain constraint is violated, thus forcing the search back into the feasible region (see [4,21]).

In the following section, we will illustrate with an analyzer application example the use of unconstrained and constrained economic modeling for the design of a combined VSI EWMA and VSI X control chart scheme. This illustration quantifies the substantial cost reduction and gains in performance that can be achieved by using the combined VSI scheme instead of its FSR counterpart.

3.7. An illustrative example

Consider an industrial process where an electrochemical device is to be calibrated on a regular basis in order to ensure that it meets certain regulatory standards. Suppose that the device is currently being monitored using a combination of an FSI EWMA chart and an FSI X chart, with a fixed-length sampling interval of $d = 1.0$ h. In the past, there have been occasional changes in the device's mean readings and/or increases in the standard deviation of the readings. The primary objective of applying the control chart scheme is to quickly detect changes in the mean μ and/or increases in the standard deviation σ of the device's readings, so that the problem can be eliminated before an excessive amount of time goes by and a large quantity of nonconforming product has been produced.

The process engineers are interested in improving the performance of the current control scheme and plan to implement the VSI feature in both the EWMA chart and the X chart. The equipment used for testing is in close proximity to the electrochemical device, so that

the short sampling interval was chosen to be $d_1 = 0.1$ h. After some careful analysis, the engineers identified the following cost and process parameters:

Cost of taking an observation, $c_1 = \$50$.

Cost of a false alarm, $c_2 = \$2000$.

Cost of a shift in μ of size $\delta = 0.5$ (with no shift in σ), $c_{3,1} = \$100/\text{h}$.

Cost of a shift in μ of size $\delta = 2.0$ (with no shift in σ), $c_{3,2} = \$500/\text{h}$.

Cost of a shift in σ of size $\sigma/\sigma_0 = 1.5$ (with no shift in μ), $c_{3,3} = \$200/\text{h}$.

Cost of a shift in μ of size $\delta = 2.0$ concurrent with a shift in σ of size $\sigma/\sigma_0 = 2.0$, $c_{3,4} = \$500/\text{h}$.

Expected time until the occurrence of a special cause, $E(T_0) = 1/\rho = 100.0$ h.

Search, repair, and recalibration time, $R = 1.0$ h.

Probability of a shift in μ of size $\delta = 0.5$ (with no shift in σ), $p_1 = 0.5$.

Probability of a shift in μ of size $\delta = 2.0$ (with no shift in σ), $p_2 = 0.2$.

Probability of a shift in σ of size $\sigma/\sigma_0 = 1.5$ (with no shift in μ), $p_3 = 0.2$.

Probability of a shift in μ of size $\delta = 2.0$ concurrent with a shift in σ of size $\sigma/\sigma_0 = 2.0$, $p_4 = 0.1$.

Applying the GRG minimization algorithm, the optimal economic design for the combined VSI EWMA and VSI X chart scheme was determined to use $\lambda = 0.09$, $d_1 = 0.10$ h, $d_2 = 4.06$ h, $h_E = 2.19$, $g_E = 0.98$, $h_X = 2.61$, and $g_X = 1.17$. This VSI scheme has an in-control ATS of 156.2 h and a long-run cost of $L = \$54.70/\text{h}$. In comparison, the optimal economic design for the combined FSI EWMA and FSI X chart scheme was determined to use $\lambda = 0.15$, the fixed-length sampling interval of $d = 6.00$ h, $h_E = 1.38$, and $h_X = 1.88$. This VSI scheme has an in-control ATS of only 70.7 h and a long-run cost of $L = \$63.24/\text{h}$. In Table 5, the columns labeled 1 and 2 give in-control ATS and out-of control SSATS values for these two optimal FSI and VSI chart combinations. In particular, column 1 gives values for the FSI combination and column 2 for the VSI combination. The optimal VSI chart combination provides much better SSATS performance than the optimal FSI chart combination, for all four combinations of shifts in μ and/or σ , at a much lower false-alarm rate and long-run cost.

After closely considering the optimal FSI and VSI schemes in columns 1 and 2 of Table 5, the process engineers decided that sampling intervals much longer than 2 h would be too risky, and an in-control ATS less than the traditional 370.4 h may cause an excessive number of disruptive false alarms. They also decided that the SSATS should not exceed 40 h for shifts in μ of size $\delta = 0.5$, or larger. Thus, to determine economic statistical designs for the combined FSI and VSI EWMA and X chart schemes that account for these concerns, the engineers introduced the following three constraints:

(C1) $\text{ATS} \geq 370.4$ h, for $\delta = 0.0$ and $\sigma = \sigma_0$.

(C2) $\text{SSATS} \leq 40.0$ h, for $\delta = 0.5$ and $\sigma = \sigma_0$.

(C3) For the FSI scheme, $d \leq 2.0$ hours; for the VSI scheme, $d_2 \leq 2.0$ hours.

Columns 3 and 4 of Table 5 give in-control ATS and out-of control SSATS values for two FSI and VSI combinations of the EWMA and X Charts that have been optimized subject to constraints (C1)–(C3). In particular, column 3 gives values for the FSI combination,

Table 5

In-control ATS and out-of-control SSATS values with chart parameter values and long-run costs for optimal FSI and VSI combinations of EWMA and X charts

			Optimal economic design		Optimal economic statistical design			
Constraint on ATS, for $\delta = 0.00$ and $\sigma = \sigma_0$			None	None	ATS ≥ 370.4	ATS ≥ 370.4	ATS ≥ 370.4	
Constraint on SSATS, for $\delta = 0.50$ and $\sigma = \sigma_0$			None	None	SSATS ≤ 40.0	SSATS ≤ 40.0	SSATS ≤ 20.0	
Constraint on sampling interval			None	None	$d \leq 2.0$	$d_2 \leq 2.0$	$d_2 \leq 2.0$	
Optimal EWMA parameter λ			$\lambda = 0.15$	$\lambda = 0.09$	$\lambda = 0.13$	$\lambda = 0.09$	$\lambda = 0.08$	
Optimal sampling interval(s)			$d = 6.00$	$d_1 = 0.10$ $d_2 = 4.06$	$d = 1.35$	$d_1 = 0.10$ $d_2 = 2.00$	$d_1 = 0.10$ $d_2 = 1.90$	
Minimal long-run cost per hour			\$63.24	\$54.70	\$76.73	\$62.33	\$71.10	
Size of shift	Probability of shift	Cost of shift per hour	FSI EWMA FSI X chart Column 1	VSI EWMA VSI X chart Column 2	FSI EWMA FSI X chart Column 3	VSI EWMA VSI X chart Column 4	VSI EWMA VSI X chart Column 5	
δ	σ/σ_0							
0.00	1.00	—	70.7	156.2	371.3	370.8	383.0	
0.50	1.00	0.5	\$100	35.2	29.2	39.0	27.7	18.9
2.00	1.00	0.2	\$500	6.9	3.4	4.1	2.4	1.9
0.00	1.50	0.2	\$200	22.1	20.4	28.2	23.7	20.2
2.00	2.00	0.1	\$500	19.2	15.0	17.8	14.3	11.2
Optimal chart parameters			h_E	1.38	2.19	2.85	2.93	2.91
			g_E	—	0.98	—	0.93	0.99
			h_X	1.88	2.61	3.08	3.19	3.19
			g_X	—	1.17	—	0.95	1.02

and column 4 gives values for the VSI combination. Once again, as compared to the FSI chart combination in column 3, the VSI combination in column 4 gives much better SSATS performance for all four combinations of shifts, at a much lower false-alarm rate and long-run cost. In particular, the long-run cost for the FSI scheme in column 3 is \$76.73/h, as compared to the long-run cost for the VSI scheme in column 4, which is only \$62.33/h. This cost is even lower than the long-run cost for the unconstrained, optimal FSI scheme in column 1, which is \$63.24/h. Since the optimal long sampling interval for the VSI scheme in column 4 is $d_2 = 2.0$ h, which is convenient from an administrative standpoint, the process engineers decided to implement this scheme for improved performance in monitoring the electrochemical device as well as a reduction in long-run cost.

It is interesting to note that the SSATS performance of the VSI scheme in column 4 can be substantially improved at a long-run cost still lower than that of the FSI scheme in column 3, but higher than the cost of the VSI scheme in column 4 chosen for implementation by the engineers. For example, the VSI chart combination in column 5 is an optimal, economic statistical design satisfying constraints (C1) and (C3), though subject to the much more restrictive constraint, $SSATS \leq 20.0$ h, for $\delta = 0.5$ and $\sigma = \sigma_0$. The long-run cost of the VSI

scheme in column 5 is \$71.10/h, still less than the cost of \$76.73/h of the FSI scheme in column 3, but higher than that of the VSI scheme in column 4. The VSI scheme in column 5 uniformly outperforms all other chart combinations in columns 1–4 at a lower false-alarm rate. For convenience, the long sampling interval of $d_2 = 1.9$ h for the VSI scheme in column 5 could be rounded up to $d_2 = 2.0$ h with a small increase in long-run cost of about 3%, which is still less than the long-run cost of the FSI scheme in column 3.

For purposes of exposition, the time and monetary units used in this paper were taken to be the hour and the dollar. However, the numerical results presented here extend to any other time or monetary units. In general, the full use of the economic design approach for practical applications will require using a computer program to find the optimal design. However, the relative performance and cost conclusions reached here for the combined FSI and VSI EWMA and X charts generally extend to many other cost and process parameter values, and in some cases, the numerical results presented in this paper may be sufficient for finding a design that is close to optimal.

4. Conclusions

In this paper, various individuals control chart schemes were contrasted for the problem of monitoring the mean μ and variance σ^2 of a normal process variable, with special consideration given to the problem of monitoring the performance of process analyzers, such as electrochemical devices, gas and liquid chromatographs, potentiometers, refractometers, and spectrometers. The combination of the VSI EWMA chart and the VSI Shewhart X chart was shown to be a very effective control scheme for this problem, from the perspective of statistical performance. From a practical perspective, this control scheme also has the convenient feature that the two charts can be shown concurrently, on the same plot, while displaying the original process data.

Furthermore, a comprehensive economic model was developed for the design of control schemes based on the combination of the VSI EWMA and VSI X charts. The economic model relates the long-run cost per time unit of operating this chart combination as a function of the charts' parameters, the system behavior, and various cost factors. The economic model can be used to quantify the reduction in cost that can be achieved by using the VSI control scheme, instead of traditional control schemes that use fixed sampling rates. The numerical results in this paper demonstrated that the cost reduction as well as the gains in statistical performance are substantial. In fact, in many practical applications, the high efficiency of the combined VSI EWMA and VSI X charts could be used to reduce the sampling effort and cost necessary to insure a required detection capability.

Finally, it should be stressed that the numerical results and conclusions of this paper are not only relevant to the problem of monitoring the performance of process analyzers. The results and conclusions generally apply to any SPC application concerned with monitoring μ and σ^2 using individual observations.

References

- [1] J. Abadie, J. Carpentier, Generalization of the Wolfe reduced gradient method to the case of nonlinear constraints, in: R. Fletcher (Ed.), *Optimization*, Academic Press, New York, NY, 1969, pp. 34–37.

- [2] S.L. Albin, L. Kang, G. Shea, An X and EWMA chart for individual observations, *J. Quality Technol.* 29 (1997) 41–48.
- [3] R.W. Amin, R.A. Ethridge, A note on individual and moving range control charts, *J. Quality Technol.* 30 (1998) 70–74.
- [4] M.S. Bazaraa, H.D. Sherali, C.M. Shetty, *Nonlinear Programming*, second ed., Wiley, New York, NY, 1993.
- [5] W. Boyes, *Instrumentation Reference Book*, third ed., Elsevier Science, Burlington, MA, 2003.
- [6] K.J. Clevert, *Process Analyzer Technology*, Wiley, New York, NY, 1986.
- [7] A.F.B. Costa, M.A. Rahim, Economic design of \bar{X} charts with variable parameters: a Markov chain approach, *J. Appl. Stat.* 28 (2001) 875–885.
- [8] S.V. Crowder, A simple method for studying run-length distributions of exponentially weighted moving average charts, *Technometrics* 29 (1987) 401–407.
- [9] R. Domangue, S.C. Patch, Some omnibus exponentially weighted moving average statistical process monitoring schemes, *Technometrics* 33 (1991) 299–313.
- [10] A.J. Duncan, The economic design of \bar{X} charts used to maintain current control of a process, *J. Am. Stat. Assoc.* 51 (1956) 228–242.
- [11] F.F. Gan, Joint monitoring of process mean and variance using exponentially weighted moving average control charts, *Technometrics* 37 (1995) 446–453.
- [12] C. Ho, K.E. Case, Economic design of control charts: a literature review for 1981–1991, *J. Quality Technol.* 26 (1994) 39–53.
- [13] R. Hooke, T.A. Jeeves, Direct search solution of numerical and statistical problems, *J. Assoc. Comput. Mach.* 8 (1961) 212–229.
- [14] J.B. Keats, E. Del Castillo, E. von Collani, E.M. Saniga, Economic modeling for statistical process control, *J. Quality Technol.* 29 (1997) 144–147.
- [15] T.L. Lai, Sequential changepoint detection in quality control and dynamical systems (with discussion), *J. Roy. Stat. Soc. Ser. B* 57 (1995) 613–658.
- [16] L.S. Lasdon, A.D. Waren, A. Jain, M. Ranter, Design and testing of generalized reduced gradient code for nonlinear programming, *ACM Trans. Math. Software* 4 (1978) 34–50.
- [17] E.L. Lehmann, *Testing Statistical Hypotheses*, second ed., Springer, New York, NY, 1985.
- [18] B.G. Liptak, *Analytical Instrumentation*, CRC Press, Boca Raton, FL, 1994.
- [19] T.J. Lorenzen, L.C. Vance, The economic design of control charts: a unified approach, *Technometrics* 28 (1986) 3–10.
- [20] J.M. Lucas, M.S. Saccucci, Exponentially weighted moving average control schemes: properties and enhancements, *Technometrics* 32 (1990) 1–12.
- [21] D.G. Luenberger, *Linear and Nonlinear Programming*, second ed., Addison-Wesley, Reading, MA, 1989.
- [22] J.F. MacGregor, T.J. Harris, The exponentially weighted moving variance, *J. Quality Technol.* 25 (1993) 106–118.
- [23] D.C. Montgomery, The economic design of control charts: a review and literature survey, *J. Quality Technol.* 12 (1980) 75–87.
- [24] D.C. Montgomery, J.C. Torng, J.K. Cochran, F.P. Lawrence, Statistically constrained economic design of EWMA control charts, *J. Quality Technol.* 27 (1995) 250–256.
- [25] M.C. Morais, A. Pacheco, On the performance of combined EWMA schemes for μ and σ : a Markovian approach, *Commun. Stat. Part B—Simulation Comput.* 29 (2000) 153–174.
- [26] L.S. Nelson, Control charts for individual measurements, *J. Quality Technol.* 14 (1982) 172–173.
- [27] G.D. Nichols, *On-Line Process Analyzers*, Wiley, New York, NY, 1988.
- [28] E.S. Page, Continuous inspection schemes, *Biometrika* 41 (1954) 100–114.
- [29] C. Park, M.R. Reynolds Jr., Economic design of a variable sampling rate \bar{X} chart, *J. Quality Technol.* 31 (1999) 427–443.
- [30] S.S. Prabhu, D.C. Montgomery, G.C. Runger, Economic–statistical design of an adaptive \bar{X} chart, *Int. J. Prod. Econom.* 49 (1997) 1–15.
- [31] M.R. Reynolds Jr., Optimal variable sampling interval control charts, *Sequential Anal.* 8 (1989) 361–379.
- [32] M.R. Reynolds Jr., Evaluating properties of variable sampling interval control charts, *Sequential Anal.* 14 (1995) 59–97.
- [33] M.R. Reynolds Jr., Variable sampling interval control charts with sampling at fixed times, *IIE Trans.* 28 (1996) 497–510.

[34] M.R. Reynolds Jr., Z.G. Stoumbos, The SPRT chart for monitoring a proportion, *IIE Trans. Quality Rel. Eng.* 30 (1998) 545–561.

[35] M.R. Reynolds Jr., Z.G. Stoumbos, Monitoring the process mean and variance using individual observations and variable sampling intervals, *J. Quality Technol.* 33 (2001) 181–205.

[36] M.R. Reynolds Jr., Z.G. Stoumbos, Individuals control schemes for monitoring the mean and variance of processes subject to drifts, *Stochast. Anal. Appl.* 19 (2001) 863–892.

[37] M.R. Reynolds Jr., Z.G. Stoumbos, Control charts and the efficient allocation of sampling resources, *Technometrics* 46 (2004) 200–214.

[38] M.R. Reynolds Jr., Z.G. Stoumbos, Should observations be grouped for effective process monitoring?, *J. Quality Technol.* 36 (2004) 343–366.

[39] S.E. Rigdon, E.N. Cruthis, C.W. Champ, Design strategies for individuals and moving range control charts, *J. Quality Technol.* 26 (1994) 274–287.

[40] S.W. Roberts, Control charts based on geometric moving averages, *Technometrics* 1 (1959) 239–250.

[41] K.C.B. Roes, R.J.M.M. Does, Y. Schurink, Shewhart-type control charts for individual observations, *J. Quality Technol.* 25 (1993) 188–198.

[42] S.M. Ross, *Applied Probability Models with Optimization Applications*, Holden-Day, San Francisco, CA, 1970.

[43] M.S. Saccucci, R.W. Amin, J.M. Lucas, Exponentially weighted moving control schemes with variable sampling intervals, *Commun. Stat.—Simul. Comput.* 21 (1992) 627–657.

[44] E.M. Saniga, Economic statistical control charts design with an application to \bar{X} and R charts, *Technometrics* 31 (1989) 313–321.

[45] S.E. Shamma, R.W. Amin, An EWMA quality control procedure for jointly monitoring the mean and variance, *Int. J. Quality Rel. Manage.* 10 (1993) 58–67.

[46] W.A. Shewhart, *Economic Control of Quality of Manufactured Product*, Van Nostrand, New York, NY, 1931.

[47] D. Siegmund, *Sequential Analysis*, Springer, New York, NY, 1985.

[48] Z.G. Stoumbos, J. Mittenthal, G.C. Runger, Steady-state-optimal adaptive control charts based on variable sampling intervals, *Stochast. Anal. Appl.* 19 (2001) 1025–1057.

[49] Z.G. Stoumbos, M.R. Reynolds Jr., Control charts applying a general sequential test at each sampling point, *Sequential Anal.* 15 (1996) 159–183.

[50] Z.G. Stoumbos, M.R. Reynolds Jr., Control charts applying a sequential test at fixed sampling intervals, *J. Quality Technol.* 29 (1997) 21–40.

[51] Z.G. Stoumbos, M.R. Reynolds Jr., Corrected diffusion theory approximations in evaluating properties of SPRT charts for monitoring a process mean, *Nonlinear Anal.—Theory, Methods Appl.* 30 (1997) 3987–3996.

[52] Z.G. Stoumbos, M.R. Reynolds Jr., Robustness to non-normality and autocorrelation of individuals control charts, *J. Stat. Comput. Simul.* 66 (2000) 145–187.

[53] Z.G. Stoumbos, M.R. Reynolds Jr., The SPRT control chart for the process mean with samples starting at fixed times, *Nonlinear Anal.—Real World Appl.* 2 (2001) 1–34.

[54] Z.G. Stoumbos, M.R. Reynolds Jr., T.P. Ryan, W.H. Woodall, The state of statistical process control as we proceed into the 21st century, *J. Am. Stat. Assoc.* 95 (2000) 992–998.

[55] Z.G. Stoumbos, M.R. Reynolds Jr., W.H. Woodall, Control chart schemes for monitoring the mean and variance of processes subject to sustained shifts and drifts, in: C.R. Rao, R. Khattree (Eds.), *Handbook of Statistics: Statistics in Industry*, vol. 22, Elsevier Science, Amsterdam, Netherlands, 2003, pp. 553–571.

[56] G. Tagaras, A survey of recent developments in the design of adaptive control charts, *J. Quality Technol.* 30 (1998) 212–231.

[57] E. Yashchin, Statistical control schemes: methods, applications and generalizations, *Int. Stat. Rev.* 61 (1993) 41–66.