An exponentially weighted moving average scheme with variable sampling intervals for monitoring linear profiles

Zhonghua Li, Zhaojun Wang*

Department of Statistics, School of Mathematical Sciences, Nankai University, Tianjin 300071, PR China

Keywords: Linear profile Control chart design Statistical process control Reliability engineering

Abstract

This paper proposes an exponentially weighted moving average scheme with variable sampling intervals for monitoring linear profiles. A computer program in Fortran is available to assist in the design of the control chart and the algorithm of the Fortran program is also given. Some useful guidelines are also provided to aid users in choosing parameters for a particular application. Simulation results on the detection performance of the proposed control chart, compared with some other competing methods show that it provides quite robust and satisfactory performance in various cases, including intercept shifts, slope shifts and standard deviation shifts. A real data example from an optical imaging system is employed to illustrate the implementation and the use of the proposed control chart.

1. Introduction

Statistical process control (SPC) has been widely used to monitor various industrial processes in reliability engineering. Most of research on SPC focused on the charting techniques and it was assumed that the quality of a process or product can be adequately represented by the univariate distribution of a quality characteristic or by the multivariate distribution of a vector consisting of several quality characteristics. However, in many practical situations, the quality of a process or product is better characterized by a relationship between a response variable and one or more explanatory variables. So there has been recent interest in monitoring process or product characterized by a simple linear profile. For Phase I control charts, Mahmoud and Woodall (2004) propose a method based on using indicator variables in a multiple regression model and Mahmoud, Parker, Woodall, and Hawkins (2007) propose a change point approach based on the segmented regression technique for testing the constancy of the regression parameters. For Phase II control charts, Kang and Albin (2000) propose two monitoring approaches: one is multivariate and the other one is a combination of exponentially weighted moving average (EWMA) and R chart. Based on the transformed model, Kim, Mahmoud, and Woodall (2003) recommend use of three univariate EWMA control charts (EWMA3) for detecting shifts in the intercept, slope and the standard deviation simultaneously. Simulation results of Kim et al. (2003) and Gupta et al. (2006) show that EWMA3 has better performance in terms of average run length (ARL), which is defined as the average number of samples before the chart signals an out-of-control condition.Recently, Zhang, Li, and Wang (2009) propose a control chart based on EWMA and Likelihood Ratio test (ELR) for monitoring linear profiles. Besides simple linear profile, Zou, Tsung, and Wang (2007) propose a Multivariate EWMA (MEWMA) scheme when the quality of a process can be characterized by a general linear profile. Woodall, Spitzner, Montgomery, and Gupta (2004) discuss some of the general issues about the problems of linear profiles and encourage research in profile monitoring. In a recent review paper, Woodall (2007) shows that the profile monitoring framework includes applications such as lumber manufacturing, monitoring of shapes, organic pigments applied to cotton surfaces, shelf-life of a food product, public health surveillance, etc.

Extensive research in recent years has developed Variable Sample Rate (VSR) control charts that vary the sampling rate as a function of current and prior sample results. The advantage of using a VSR chart instead of a Fixed Sampling Rate (FSR) chart is that a VSR chart provides much faster detection of small and moderate process changes, for a given in-control ARL and a given in-control average sampling rate. There are several approaches that can be used to vary the sampling rate. One approach is a Variable Sampling Intervals (VSI) chart that varies the sampling intervals as a function of the sample results from the process. Another approach to varying the sampling rate is a Variable Sample Size (VSS) chart that varies the sample sizes as a function of the sample results from the process. The VSI and VSS features can be combined to give a Variable Sample Sizes and Sampling Intervals (VSSI) control chart.

* This manuscript was processed by Area Editor E. A. Elsayed.
* Corresponding author. Tel.: +86 22 23498233; fax: +86 22 23506423.
E-mail address: ejywang@nankai.edu.cn (Zh. Wang).
that allows the sample sizes and sampling intervals to vary. There have been lots of research on conventional control chart using VSR features in the literature, for the $X$ control chart, see Prabhhu, Montgomery, and Runger (1998), Costa (1998), Chen and Chiu (2005) and Celano et al. (2006); for the cumulative sum (CUSUM) control chart, see Reynolds, Amin, and Arnold (1990); Zhang and Wu (2007) and Wu et al. (2007); for the EWMA control chart, see Reynolds (1996) and Reynolds et al. (2001); for the cumulative count of conforming chart, see Liu, Xie, Goh, Liu, and Yang (2006) and for the Hotelling’s $T^2$ control chart, see Aparisi and Haro (2001).

For the $X$ control chart, Chen and Liao (2004) study multi-criteria design and show that the design parameters of sample sizes and sampling intervals need to be adjusted. Shamsuzzaman, Wu, and Elias (2009) propose an algorithm for deploying manpower to an $X$ & $S$ control chart which minimizes the expected total cost.

Montgomery (2007) shows that one important area of SPC research continues to be the use of control charts with variable sample sizes and/or sampling intervals. Although there have been lots of research on conventional control charts using VSR features in the literature, there is little work on linear profile monitoring schemes, except the VSI MEWMA of Zou et al. (2007) and the VSI ELR of Zhang et al. (2009), against which we compare our proposed chart. As Woodall (2007) points out, profile monitoring is very useful in an increasing number of practical applications, the objective of this paper is to perform a detailed investigation of an EWMA scheme with variable sampling intervals (VSI EWMA) for monitoring linear profiles in which it is desirable to determine the sampling interval for the next sample before sampling is started for this sample. As Zou et al. (2007) point out, the design of a combination of VSI EWMA is not at all trivial, because three warning limits and three control limits must be chosen simultaneously so that all charts have the same individual in-control average sampling rate and average false alarm rate. We overcome this difficulty by transforming the three EWMA statistics to one omnibus statistic.

The rest of this paper is organized as follows. In the next section, some competitive schemes are briefly introduced. Our proposed VSI EWMA and its designing strategies are presented in Section 3. The numerical comparisons with the EWMA of Kim et al. (2003), the VSI MEWMA of Zou et al. (2007) and the VSI ELR of Zhang et al. (2009) are carried out in Section 4. A real data example from an optical imaging system is used to illustrate the VSI EWMA in Section 5. Some computation aspects are presented in Section 6. Several remarks conclude this paper in Section 7.

2. EWMA control charts for monitoring linear profiles

In this section, the EWMA of Kim et al. (2003), the MEWMA of Zou et al. (2007) and the ELR of Zhang et al. (2009) are briefly introduced.

2.1. The EWMA$ _3$ chart

Assume that the $j$th random sample collected over time is $(x_i, y_{ij})$, $i = 1, 2, \ldots, n_j$. When process is in control, the relationship between the response variable and the explanatory variables is assumed to be

$$y_{ij} = \beta_0 + \beta_1 x_i + \varepsilon_{ij}, i = 1, 2, \ldots, n_j,$$

where $\varepsilon_{ij}$ is an independently identically distributed (i.i.d.) as standard normal random variable. We assume that $n_j$ are all equal to $n$. This is usually the case in many practical applications. When the parameters $A_0$, $A_1$ and $\sigma^2$ are unknown, they can be estimated from Phase 1 sample data. Suppose that there are, in total, $m (m \geq 1)$ in-control samples of size $n_i$, $(x_i, y_{ij}), i = 1, 2, \ldots, n_j; j = 1, 2, \ldots, m$. The most often used unbiased estimators of $A_0$, $A_1$, and $\sigma^2$ are the average of the $m$ least square estimators, $\hat{A}_0$, $\hat{A}_1$, and $\hat{\sigma}^2$, which are given by

$$\hat{A}_0 = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} y_{ij}}{m \sum_{i=1}^m n_i},$$

$$\hat{A}_1 = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{A}_0) x_i}{m \sum_{i=1}^m n_i},$$

$$\hat{\sigma}^2 = \frac{1}{m \sum_{i=1}^m n_i - 2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{A}_0 - \hat{A}_1 x_i)^2,$$

where

$$\bar{y}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} y_{ij}, \quad \bar{x}_i = \frac{1}{n_j} \sum_{i=1}^{n_j} x_i, \quad s_{X_1} = \left( \frac{1}{n_j} \sum_{i=1}^{n_j} (x_i - \bar{x}_i)^2 \right)^{1/2},$$

$$S_{xy} = \frac{1}{n_j} \sum_{i=1}^{n_j} (x_i - \bar{x}_i) y_{ij}.$$

After the estimations are determined, the parameters are assumed to be known and the Phase II monitoring could be started. In the following of this paper, we focus on Phase II monitoring linear profiles.

Kim et al. (2003) code the explanatory values and obtain the following alternative form of the relationship model:

$$y_{ij} = \beta_0 + \beta_1 x_i + \varepsilon_{ij}, \quad i = 1, 2, \ldots, n,$$

where

$$\beta_0 = A_0 + A_1 \bar{x}, \quad \beta_1 = A_1, \quad x_i = x_i - \bar{x}.$$

For the $j$th sample, the least square estimators of $\beta_0$, $\beta_1$ and $\sigma^2$ are

$$\hat{\beta}_0 = \bar{y}_j - \hat{A}_1 \bar{x},$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n_j} \left( y_{ij} - \hat{A}_0 - \hat{A}_1 x_i \right)}{\sum_{i=1}^{n_j} \left( x_i - \bar{x}_i \right)^2},$$

$$\hat{\sigma}^2 = \frac{1}{n_j - 2} \sum_{i=1}^{n_j} \left( y_{ij} - \hat{\beta}_0 - \hat{\beta}_1 x_i \right)^2.$$

Note that these three estimators are independent. Kim et al. (2003) propose using three EWMA charts (EWMA$_0$, EWMA$_1$, EWMA$_2$) to monitor the changes of $Y$-intercept ($\beta_0$), the slope ($\beta_1$) and the standard deviation ($\sigma$), respectively. They are

$$E_0(j) \triangleq EWMA_0(j) = \theta \hat{\beta}_0 + (1 - \theta) EWMA_0(j - 1),$$

$$E_1(j) \triangleq EWMA_1(j) = \theta \hat{\beta}_1 + (1 - \theta) EWMA_1(j - 1),$$

$$E_2(j) \triangleq EWMA_2(j) = max(0, \ln(MSE_j)) + (1 - \theta) \text{EWMA}_2(j - 1), \ln(\sigma^2),$$

where $0 < \theta \leq 1$ is a smoothing parameter, $E_0(0) = B_0$, $E_1(0) = B_1$ and $E_2(0) = \ln(\sigma^2)$. The smoothing parameter $\theta$ is set at 0.2, which is a typical choice in the literature (Kang & Albin, 2000; Kim et al., 2003; Zou et al., 2007; Zhang et al., 2009). The three EWMA charts are used jointly.

The EWMA$_0$ gives a signal as soon as one or more of the following three conditions hold.

$$|E_0(j) - \hat{\sigma}| > L_0 \sigma \left( \frac{\theta}{(2 - \theta)m} \right) \frac{\sqrt{\theta}}{1 - \theta},$$

$$|E_1(j) - \hat{\sigma}| > L_0 \sigma \left( \frac{\theta}{(2 - \theta)m} \right),$$

$$|E_2(j) - \hat{\sigma}^2| > L_2 \sqrt{\frac{\theta}{2 - \theta}} \text{Var}_x[\text{ln}(\text{MSE}_j)],$$

where $L_0$, $L_1$, and $L_2$ are chosen to give a specified in-control ARL and $\text{Var}_x[\text{ln}(\text{MSE}_j)] \approx \frac{2}{n_j - 2} + \frac{2}{(n_j - 2)^2} + \frac{4}{3(n_j - 2)^3} - \frac{16}{15(n_j - 2)^5}$. As Liu et al. (2006) indicate, in practical applications, people are more concerned at the process deterioration rather than improvement, so we focus on the statistic $E_0$, designed to detect an increase in variance.
2.2. The MEWMA chart

Assume that the jth random sample is \((X_j, Y_j)\), where \(Y_j\) is \(n_j\)-variate vector and \(X_j\) is a \(n_j \times p(n_j > p)\) matrix. When the process is in-control, the underlying model is

\[
Y_j = X_j\beta + \epsilon_j,
\]

where \(\beta = (\beta_1, \beta_2, \ldots, \beta_p)\) is the \(p\)-dimensional coefficient vector and the \(\epsilon_j\) are i.i.d. as an \(n_j\)-variate multivariate normal random vector with mean 0 and \(\Sigma_1\) covariance matrix. The \(n_j\)s are assumed to be equal and \(X_j\) is assumed to be fixed for different \(j\), denoted as \(n\) and \(X\), respectively. Define

\[
Z_j(\hat{\beta}) = (\hat{\beta} - \beta)/\sigma, \quad Z_j(\sigma) = \Phi^{-1}\left\{F\left((n-p)/\sigma^2; n-p\right)\right\},
\]

where \(\hat{\beta} = (XX')^{-1}XY\), \(\sigma_j^2 = (Y_j - X\hat{\beta})'(Y_j - X\hat{\beta})/\hat{\beta}\), \(\Phi^{-1}\) is the inverse of the standard normal cumulative distribution function, and \(F(\cdot; v)\) is the chi-squared distribution function with \(v\) degrees of freedom.

Zou et al. (2007) defines the MEWMA charting statistic as

\[
W_j = \theta Z_j + (1 - \theta) W_{j-1}, \quad j = 1, 2, \ldots,
\]

where \(Z_j = (Z_j(\hat{\beta}), Z_j(\sigma'))\). The MEWMA signals when

\[
U_j = W_j \Sigma^{-1} W_j > L \frac{\theta}{1 - \theta},
\]

where \(L > 0\) is chosen to achieve a specified in-control ARL and \(\Sigma = \left(XX'^{-1} 0 \right)\) is the in-control covariance matrix of \(Z_j\).

2.3. The ELR chart

Zhang et al. (2009) obtain the generalized likelihood ratio statistic for sample \(j\) as follows

\[
LR_j = C_j - n \log \sigma_j^2 - n,
\]

where

\[
C_j = \sum_{i=1}^{n} (y_{ij} - B_0 - B_1 x_i)^2, \quad \sigma_j^2 = \frac{1}{n} \sum_{i=1}^{n} (y_{ij} - B_1 x_i - B_0)^2.
\]

Subsequently, four EWMA statistics are introduced by

\[
EI_j = \theta b_{0j} + (1 - \theta) EI_{j-1}, \quad ES_j = \theta b_{1j} + (1 - \theta) ES_{j-1}, \quad EE_j = \theta \Sigma_{ij} + (1 - \theta) EE_{j-1}, \quad EC_j = \theta c_{ij} + (1 - \theta) EC_{j-1},
\]

where \(b_{0j} = B_0, \quad b_{1j} = B_1, \quad \Sigma_{ij} = \frac{1}{n} \sum_{i=1}^{n} (y_{ij} - ES_j x_i - EE_j) \Sigma_{ij}^{-1} \)

and \(ES_0 = 0\), \(EE_0 = 0\), \(EC_0 = 0\), \(S_j = \frac{1}{n} \sum_{i=1}^{n} (y_{ij} - ES_j x_i - EE_j) \Sigma_{ij}^{-1} \)

Finally, they substitute \(EC_j\) and \(EE_j\) for \(C_j\) and \(\sigma_j^2\) in \(LR_j\) and obtain the charting statistics

\[
ELR_j = LR_j - n \log EE_j - n.
\]

If \(ELR_j > L\), an alarm is triggered.

3. The VSI EWMAs control chart

The VSI EWMAs is the EWMAs control chart, defined in Eqs. (2) and (3), using a longer sampling interval as long as the sample point is close to the target so that there is no indication of process changes. However, if the sample point is far from the target, but still within the action limits, a shorter sampling interval is used. If a sample point falls in the action region, then the process is considered to be out-of-control.

Assume that in the conventional linear profile monitoring, the fixed sampling interval is \(d_0\). In general, for our VSI EWMAs, the sampling interval function \(d(\cdot)\) can be of any form, but previous research on VSI control charts has shown that it is sufficient to use only two possible values for the sampling intervals to achieve good statistical properties in VSI control charts, see, for example, Costa (1998), Wu, Zhang, and Wang (2007) and Reynolds et al. (2001).

Let \(d_1\) and \(d_2\) represent these two possible sampling intervals, where \(0 < d_1 < d_2\). Then the sampling interval function \(d(\cdot)\) can be defined by partitioning \(C\) of the continuation or in-control region, into two regions, say warning region \(R_w\) and central region \(R_c\), such that

\[
d(\cdot) = \begin{cases} d_1, & \text{if the monitor statistic falls into } R_w, \\ d_2, & \text{if the monitor statistic falls into } R_c. \end{cases}
\]

It is advisable to start the control with the shorter sampling interval, \(d_1\), so the first sample is taken quickly after the process is started in case of start-up problems.

Unlike VSI control charts for \(X\) control charts, CUSUM control charts, or EWMAs control charts, it is quite difficult to specify the regions \(R_w\) and \(R_c\) for the VSI EWMAs. The reason is that the monitoring statistics \(E(j), E_j(j)\) and \(E(j)\) do not have an explicit in-control distribution, although \(b_0\) and \(B_1\) are known to be normally distributed with means \(B_0\) and \(B_1\) and variances \(\sigma^2/n\) and \(\sigma^2/S_n\), respectively.

To overcome these problems, we divide the three monitor statistics \(E_j(j), E_j(j)\) and \(E(j)\) by the three corresponding control limits. That is, what we plot are

\[
\begin{align*}
SE_j(j) &\equiv \frac{E_j(j) - B_0}{L_2 \sigma_2^{1/2}}, \\
SE_j(j) &\equiv \frac{E_j(j) - B_1}{L_2 \sigma_2^{1/2}}, \\
SE_j &\equiv \frac{E_j(j)}{L_2 \sigma_2^{1/2} \text{Var}(\ln(MSE_j))}.
\end{align*}
\]

In this condition, the VSI EWMAs control chart signals as soon as \(A(j) > 1\), where

\[
A(j) = \max(SE_j(j), SE_j(j), SE_j(j)).
\]

Then our VSI EWMAs can use warning limit \(\omega\) and control limit \(h = 1\) to divide the chart into the central region \(R_c = (0, \omega)\), warning region \(R_w = (\omega, 1)\) and action region \(R_a = [1, +\infty)\).

To facilitate the derivation of \(\omega\), define \(p_0\) as the conditional probability of a sample point \(A(j)\) falling in the central region given that this point does not fall in the action region, i.e.,

\[
p_0 = P(A(j) < 1 \mid A(j) < 1).
\]

A large value of \(p_0\) indicates that the value \(\omega\) is close to 1, and a large number of samples is taken using the longer sampling interval \(d_2\). Due to the intricacy of the distribution of \(A(j)\), we can only find the warning limit \(\omega\) corresponding to different \(p_0\) through Monte Carlo simulation. As Kim et al. (2003), it is assumed in this paper that the underlying in-control linear profile model is

\[
y_{ij} = 1 + 2x_i + \epsilon_j, \quad \text{where the } \epsilon_j's \text{ i.i.d. as standard normal random variables. The } \chi_i \text{ values are } -3(2)3 \text{ with } \chi = 0. \text{ As expected, } \omega \text{ is related to the sample size } n, \text{ the explanatory variables } x_1, \ldots, x_n \text{ and the control limits } L_1, L_2, L_3. \text{ Given these, a Fortran program to find corresponding } \omega \text{ is available from the authors upon request.}
\]

The corresponding algorithm is deferred in Section 6 to avoid practitioners.

Although it is needed to change the original three monitoring statistics \(E_j, E_j, E_j\) into one statistic \(A(j)\), it does not hamper the diagnostic ability of the VSI EWMAs at all. Once \(A(j)\) gives an out-of-control signal for some \(j\), all needs to be done is check which one of the three monitoring statistics \(E_j, E_j, E_j\) is the maximum. If this maximum is \(E_j, E_j, E_j\), there is significant cause to generate a process change in the intercept, the slope or the standard
deviation. We do not need separate parametric tests as in Zou et al. (2007), which is appealing for the practitioners because they may be reluctant to take extra effort to diagnose which parameter has changed.

3.1. The sensitivity analysis of VSI EWMA$_3$ chart

Traditionally, the ARL has been generally employed as a performance indicator to evaluate the effectiveness of various control schemes, provided that the sampling interval remains constant. However, when the sampling interval is variable, the time to signal is not a constant multiple of the ARL, and thus ARL is not appropriate for evaluating the effectiveness of VSI control charts. The widely used performance indicators for control charts with VSI are the average time to signal (ATS), which is defined as the expected value of time from the start of the process to the time when the charts indicate an out-of-control signal, and the adjusted average time to signal (AATS), which is defined as the expected value of time from the occurrence of an assignable cause to the time when the charts indicate an out-of-control signal. The AATS is also called the steady-state ATS (SSATS).

When the process is in-control, the ATS may be used to develop the measures of the false alarm rate for a chart. A chart with a larger in-control ATS indicates a lower false alarm rate than other charts. When the process is out-of-control, the AATS may be used to measure the performance of a chart. A chart with a smaller out-of-control AATS indicates a better detection ability of process shifts than other charts. To make the linear profile monitoring schemes with and without VSI comparable, the same in-control average sample rate is used, i.e.,

$$1 - p_0|d_1 + p_0d_2 = d_0.$$

The performance of the VSI EWMA$_3$ is related to determination of the following parameters: the warning limit $w$, the sampling interval $d_1$ and $d_2$. Of course, the parameters $L_1$, $L_2$ and $L_3$ involved in computing the statistics $A(j)$ should be specified first to give a specified in-control ATS. They can be determined by simulation according to Kim et al. (2003).

In this paper, it is assumed that $d_0 = 1$ without loss of generality. Otherwise, the results can be obtained multiplied by $d_0$. Kim et al. (2003) show that $L_1 = 3.0156$, $L_2 = 3.0109$ and $L_3 = 1.3723$ give an overall in-control ARL of roughly 200. Thus, the ARL is 200 under the assumption that $d_0 = 1$.

To facilitate the determination of $w$, the conditional probability $p_0$, which can be considered as the proportion of samples taken in the sampling interval $d_2$ when the process is in-control, needs to be specified. Fig. 1a–c provide the ATS and AATS for several VSI linear profile monitoring schemes with various $p_0$ when the intercept $B_0$, the slope $B_1$ and the standard deviation $\sigma$ changes to $B_0 + \delta \sigma$, $B_1 + \delta \sigma$ and $\gamma \sigma$, respectively. These are on a log scale for a clearer comparison. In Fig. 1, $d_1 = 0.1$ and $d_2$ computed from Eq. (7) are used. Fig. 1 shows that the VSI EWMA$_3$ with smaller $p_0$ value has a smaller out-of-control AATS when the in-control ATS is approximately 200, which results in a quicker detection of process shifts. It seems that $p_0$ should be as small as possible from statistical point of view. However, too small a $p_0$ gives too large a $d_2$, which implies that once $A(j)$ falls in $R_3$, too long a sampling interval will be used. This is not realistic in practice. For EWMA and $\bar{X}$ control chart, Reynolds (1996) and Lin and Chou (2005a) suggest that the $p_0$ value should be in the vicinity of 0.8. Therefore, we suggest to use a $p_0$ value in the vicinity of 0.8 for the VSI EWMA$_3$.

When the VSI EWMA$_3$ is used, the sampling intervals may vary. The detection ability depends on the sampling intervals $d_1$ and $d_2$.

![Fig. 1. Log(AATS) for $p_0=0.5(0.1)0.9$, $d_1 = 0.1$.](image-url)

Fig. 2a–c provide the ATS and AATS (on a log scale) for several VSI EWMA$_3$ with various $d_1$ and $d_2$ computed from Eq. (7) when the intercept $B_0$, the slope $B_1$ and the standard deviation $\sigma$ changes to $B_0 + \delta \sigma$, $B_1 + \delta \sigma$ and $\gamma \sigma$, respectively. In Fig. 2, $p_0 = 0.8$ is used based on the discussions above. Fig. 2 shows that the VSI EWMA$_3$ with smaller $d_1$ value has a smaller out-of-control AATS. It seems that $d_1$ should be as small as possible. However, $d_1$ depends on the shortest time required to sample each item in practice. Thus, $d_1$ may be equal to the shortest time to sample each item.
3.2. The design procedure of VSI EWMA\textsubscript{3} chart

For designing the VSI EWMA\textsubscript{3} chart, we should find a combination of parameters \((p_0, \omega, d_1, d_2, l_0, l_3, l_4)\) that achieves the specified in-control ATS and the chart signals quickly when the process undergoes a shift. From the performance of the VSI EWMA\textsubscript{3} chart shown in Figs. 1 and 2, the following design procedure is recommended.

1. Set the in-control ATS according to the administrative consideration and determine the values of \(l_0, l_3\) and \(l_4\), which can be obtained through simulation. They are selected under the criterion that the statistics \(E_x, E_z\) and \(E_E\) are equally important, i.e., the control charts with \(E_x, E_z\) and \(E_E\) give nearly the same in-control ATS individually. See more details in Kim et al. (2003). If a manager, however, does not think the statistics \(E_x, E_z\) and \(E_E\) are equally important, we can control individual in-control average false alarm rate by adjusting \(L_0, L_3\) and \(L_4\).

2. Choose the conditional probability \(p_0\). It may be in the vicinity of 0.8 if there is no special requirement.

3. Choose the warning limit \(\omega\) by a computer program, which guidelines are detailed in Section 6. A Fortran program is also available from the authors upon request.

4. Determine the shorter sampling interval \(d_1\). It may be the shortest time to sample each item. In the literature, \(d_1 = 0.01\) and 0.1 for X control chart in Prabhuh (1994), \(d_1 = 0.1\) and 0.5 for CUSUM charts in Reynolds et al. (1990), \(d_1 = 0.1\) and 0.25 for EWMA control charts in Reynolds et al. (2001), \(d_1 = 0.1\) and 0.2 for Hotelling’s \(T^2\) control chart in Aparisi and Haro (2001) and \(d_1 = 0.1, 0.25\) and 0.5 for MEWMA control chart in Zou et al. (2007). So \(d_1 = 0.1\) is a typical choice if it is feasible to sample again after the current sample is obtained.

5. \(d_2\) is computed from Eq. (7).

4. Comparisons

In this section, a comparative study is conducted by Monte Carlo simulation to evaluate the performance of the VSI EWMA\textsubscript{3}. The simulations are sufficiently long, 10,000 replications, such that the standard errors of the estimates are less than 2%, enabling us to draw reasonable conclusions. The in-control ATS of each chart is set to be equal, and so is the in-control average false alarm rate, such that the comparisons can be conducted under the same criteria.

4.1. Comparison with EWMA\textsubscript{3}

To compare with the result of Kim et al. (2003), note that the standard deviation \(\sigma = 1\) and the in-control ATS is approximately 200 under the assumption that \(d_0 = 1\). From the design strategies in the previous section, \(p_0 = 0.8\) and \(d_1 = 0.1, 0.25\) and 0.5 are employed for the VSI EWMA\textsubscript{3}. After \(p_0\) and \(d_1\) are determined, \(d_2\) is computed from Eq. (7). From our computer program, the warning limit \(\omega\) corresponding to \(p_0 = 0.8\) is 0.56.

TheRequested mismatch is fixed here. The AATS comparisons under intercept shifts from \(B_0\) to \(B_0 + \lambda \sigma\), under slope shifts from \(B_1\) to \(B_1 + \delta \sigma\), under combinations of intercept and slope shifts, and under deviation shifts from \(\sigma\) to \(\gamma \sigma\) are shown in Tables 1 and 2, respectively. These shift patterns cover a wide range of shifts in practice. In these tables, the rows labeled “EWMA\textsubscript{3}” are the AATS of the traditional linear profile monitoring method of Kim et al. (2003) and the rows labeled “VSI EWMA\textsubscript{3}” are the AATS of our proposed VSI EWMA\textsubscript{3}.

From Tables 1 and 2, we have the following observations:

1. It is clear that adding the VSI feature to the linear profile monitoring method of Kim et al. (2003) substantially improves the efficiency of the chart. The out-of-control AATS can have a 57–98%, 54–98% and 24–82% off for intercept shifts, slope shifts and standard deviation shifts, respectively. For the combinations of intercept and slope shifts, the out-of-control AATS can have a 46–98% off.
Table 1  
AATS comparisons with EWMA/ under intercept and slope shifts.

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$B_0 + \alpha \delta$</th>
<th>$B_0 \times \alpha^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>EWMA $\delta$</td>
<td>197.8</td>
</tr>
<tr>
<td></td>
<td>VSI EWMA $\delta$</td>
<td>197.5</td>
</tr>
<tr>
<td>0.05</td>
<td>EWMA $\delta$</td>
<td>176.5</td>
</tr>
<tr>
<td></td>
<td>VSI EWMA $\delta$</td>
<td>173.8</td>
</tr>
<tr>
<td>0.10</td>
<td>EWMA $\delta$</td>
<td>132.1</td>
</tr>
<tr>
<td></td>
<td>VSI EWMA $\delta$</td>
<td>126.0</td>
</tr>
</tbody>
</table>

(2) The time reduction becomes larger as the shifts get larger. This implies that adding VSI feature to the linear profile monitoring method has more benefit for larger process shifts. This is different from other control charts, such as $\chi^2$, CUSUM and EWMA control charts, where adding VSI feature is more effective for small to moderate process shifts. The reason may be that when a process has a larger process shift, the shorter sampling interval $d_1$ is used most of the time.

Table 2  
AATS comparisons with EWMA/ under deviation shifts from $\sigma$ to $\gamma \sigma$.

<table>
<thead>
<tr>
<th>$\gamma\sigma$</th>
<th>$B_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>33.9</td>
</tr>
<tr>
<td>1.4</td>
<td>27.4</td>
</tr>
<tr>
<td>1.6</td>
<td>28.6</td>
</tr>
<tr>
<td>1.8</td>
<td>30.2</td>
</tr>
</tbody>
</table>

(3) With some exceptions, such as $\lambda = 0$, $\delta = 0.025$ and 0.05, control charts with smaller $d_1$ have smaller AATS, which is consistent with previous research on VSI control charts.

4.2. Comparison with VSI MEWMA and VSI ELR

To compare with the results of Zou et al. (2007) and Zhang et al. (2009), the underlying model is the same with the model used by Kim et al. (2003). The AATS comparisons under intercept shifts
from \( B_0 \) to \( B_0 + h \sigma \) and under deviation shifts from \( \sigma \) to \( \sigma \gamma \) are shown in Tables 3 and 4, respectively, with the lines labeled “VSI MEWMA” being the AATS of the VSI MEWMA of Zou et al. (2007) and the lines labeled “VSI ELR” being the AATS of the VSI ELR of Zhang et al. (2009). Note that the smoothing parameter \( \theta = 0.2 \) in Zou et al. (2007) and Zhang et al. (2009). We use the same smoothing parameter to make fair comparisons, although we can use different \( \theta \) for different shifts to get optimal results.

From Tables 3 and 4, we have the following observations.

1. For intercept shifts, for \( \lambda > 0.1 \), the VSI EWMA3 slightly outperforms both the VSI MEWMA and the VSI ELR. In particular, with respect to the VSI MEWMA, the advantage of the VSI EWMA3 is more evident for a smaller value of \( d_1 \). We will not report our result under shifts in slope because it is, however, similar to the case under shifts in intercept based on our simulation results.
2. For deviation shifts, our proposed VSI EWMA3 is designed for detecting up-sided shifts. When the shift size is larger than 1, our method has uniformly better performance than VSI MEWMA with only one exception when \( \gamma = 3.0 \), \( d_1 = 0.5 \). However, for \( \gamma > 1.1 \), the VSI EWMA3 is never better than the VSI ELR throughout the three values of \( d_1 \).

Although our proposed VSI EWMA3 chart does not globally perform better than the VSI MEWMA and VSI ELR, it has its own advantages. First, the identification of the out-of-control profile parameters of VSI MEWMA is based on other three parametric tests, which makes the application so complex that practitioners may be reluctant to take extra effort to diagnose which parameter has changed after they get an out-of-control signal. While our VSI EWMA3 inherits the advantage of EWMA3 that the three EWMA statistics are independent, so it does not need extra diagnostic aids. We can have some idea of the cause of a shift by seeing which one of the three monitoring statistics is the maximum. Second, there are four EWMA statistics involved in VSI ELR and three in our VSI EWMA3, which makes the computation of our VSI EWMA3 a little easier. Third, VSI MEWMA and VSI ELR cannot control individual in-control average false alarm rate as our VSI EWMA3 because their control charts employ one control limit \( L \) while our VSI EWMA3 can use \( L_1, L_2 \) and \( L_3 \) to control average false alarm rate for intercept, slope and standard deviation, respectively, if a manager does not think the statistics \( E_1 \), \( E_2 \) and \( E_3 \) are equally important. We think these features of our VSI EWMA3 could be appealing for managers and practitioners.

### 5. A real data example

In this section, the VSI EWMA3 is illustrated by a real data example.


gupta, montgomery, and woodall (2006) use to compare the performance of two methods. the data set consists of line widths of photo masks reference standards on 10 units (40 measurements) used for monitoring linear calibration profiles of an optical imaging system. the line widths are used to estimate the parameters of the linear calibration profile, \( y_i = 0.2817 + 0.9767x_i \) with a residual standard deviation of 0.06826 \( \mu \). the data set is presented in Table 7 of Gupta et al. (2006).

In order to be consistent with Eq. (1), the original data has been standardized by

\[
y_0 = \frac{y_i - 0.2817}{0.06826} = \left( \frac{0.9767}{0.06826} \right) x_i + c_y
\]

where \( c_y \sim N(0, 1) \). The parameters \( p_0 = 0.8 \) and \( d_1 = 0.1 \) are used for the VSI EWMA3. The results are listed in Table 5. The columns labeled “E1,” “E2,” “E3,” “E4,” “SEE,” “SEE,” “SEE,” and “A” are defined in Eqs. (2), (5) and (6), respectively. The last column is the sampling interval employed. Note that there is an out-of-control signal for the 4th sample because \( A(4) = 1.178 > 1 \) (labeled by *** for significance to indicate when the chart gives an out-of-control signal). This is consistent with the result of Gupta et al. (2006). The time used to detect this shift is 0.100 + 1.225 \times 3 = 3.775 for the VSI EWMA3.

Gupta et al. (2006) replace the EWMA charts of EWMA3 by \( X \) charts to monitor the intercept and slope and by an \( S^2 \) chart to monitor the error variance, which we call EWMA3 – Shewhart. Although all the EWMA3 – Shewhart of Gupta et al. (2006), the ELR of Zhang et al. (2009) and the VSI EWMA3 give an out-of-con-

**Table 3**

<table>
<thead>
<tr>
<th>( \lambda )</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.8</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
<th>3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( d_1 = 0.5 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VSI MEWMA</td>
<td>124.4</td>
<td>51.9</td>
<td>23.3</td>
<td>12.6</td>
<td>8.0</td>
<td>5.7</td>
<td>3.6</td>
<td>2.6</td>
<td>1.6</td>
<td>1.1</td>
<td>0.7</td>
</tr>
<tr>
<td>VSI ELR</td>
<td>122.8</td>
<td>52.1</td>
<td>23.9</td>
<td>13.1</td>
<td>8.4</td>
<td>6.0</td>
<td>3.8</td>
<td>2.8</td>
<td>1.6</td>
<td>1.0</td>
<td>0.5</td>
</tr>
<tr>
<td>VSI EWMA3</td>
<td>124.3</td>
<td>50.8</td>
<td>22.8</td>
<td>12.4</td>
<td>7.9</td>
<td>5.8</td>
<td>3.8</td>
<td>2.9</td>
<td>1.9</td>
<td>1.5</td>
<td>1.1</td>
</tr>
<tr>
<td>( d_1 = 0.25 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VSI MEWMA</td>
<td>122.2</td>
<td>48.3</td>
<td>20.4</td>
<td>10.6</td>
<td>6.6</td>
<td>4.8</td>
<td>3.1</td>
<td>2.3</td>
<td>1.4</td>
<td>1.0</td>
<td>0.8</td>
</tr>
<tr>
<td>VSI ELR</td>
<td>117.1</td>
<td>47.6</td>
<td>20.7</td>
<td>10.9</td>
<td>6.7</td>
<td>4.8</td>
<td>3.0</td>
<td>2.2</td>
<td>1.4</td>
<td>0.9</td>
<td>0.5</td>
</tr>
<tr>
<td>VSI EWMA3</td>
<td>121.5</td>
<td>47.5</td>
<td>19.1</td>
<td>10.0</td>
<td>6.1</td>
<td>4.5</td>
<td>2.9</td>
<td>2.2</td>
<td>1.4</td>
<td>1.1</td>
<td>0.8</td>
</tr>
<tr>
<td>( d_1 = 0.1 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VSI MEWMA</td>
<td>120.0</td>
<td>45.2</td>
<td>18.1</td>
<td>9.2</td>
<td>5.8</td>
<td>4.2</td>
<td>2.8</td>
<td>2.1</td>
<td>1.4</td>
<td>1.1</td>
<td>0.9</td>
</tr>
<tr>
<td>VSI ELR</td>
<td>112.2</td>
<td>44.2</td>
<td>18.0</td>
<td>9.2</td>
<td>5.7</td>
<td>4.2</td>
<td>2.8</td>
<td>2.2</td>
<td>1.4</td>
<td>0.9</td>
<td>0.5</td>
</tr>
<tr>
<td>VSI EWMA3</td>
<td>119.1</td>
<td>42.1</td>
<td>16.4</td>
<td>7.7</td>
<td>4.8</td>
<td>3.4</td>
<td>2.2</td>
<td>1.7</td>
<td>1.0</td>
<td>0.8</td>
<td>0.6</td>
</tr>
</tbody>
</table>

**Table 4**

<table>
<thead>
<tr>
<th>( \gamma )</th>
<th>1.1</th>
<th>1.2</th>
<th>1.4</th>
<th>1.8</th>
<th>2.2</th>
<th>2.6</th>
<th>3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( d_1 = 0.5 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VSI MEWMA</td>
<td>68.9</td>
<td>27.4</td>
<td>14.1</td>
<td>8.8</td>
<td>3.2</td>
<td>1.9</td>
<td>1.1</td>
</tr>
<tr>
<td>VSI ELR</td>
<td>68.3</td>
<td>24.1</td>
<td>7.1</td>
<td>2.6</td>
<td>1.6</td>
<td>1.1</td>
<td>0.9</td>
</tr>
<tr>
<td>VSI EWMA3</td>
<td>66.5</td>
<td>27.0</td>
<td>9.5</td>
<td>3.5</td>
<td>2.3</td>
<td>1.8</td>
<td>1.5</td>
</tr>
<tr>
<td>( d_1 = 0.25 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VSI MEWMA</td>
<td>66.3</td>
<td>25.2</td>
<td>12.5</td>
<td>7.7</td>
<td>2.7</td>
<td>1.7</td>
<td>1.1</td>
</tr>
<tr>
<td>VSI ELR</td>
<td>65.7</td>
<td>22.1</td>
<td>6.1</td>
<td>2.2</td>
<td>1.4</td>
<td>1.1</td>
<td>0.9</td>
</tr>
<tr>
<td>VSI EWMA3</td>
<td>61.7</td>
<td>25.0</td>
<td>8.0</td>
<td>2.7</td>
<td>1.6</td>
<td>1.3</td>
<td>1.1</td>
</tr>
<tr>
<td>( d_1 = 0.1 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VSI MEWMA</td>
<td>63.9</td>
<td>23.4</td>
<td>11.4</td>
<td>6.9</td>
<td>2.6</td>
<td>1.7</td>
<td>1.2</td>
</tr>
<tr>
<td>VSI ELR</td>
<td>63.6</td>
<td>20.7</td>
<td>5.6</td>
<td>2.2</td>
<td>1.4</td>
<td>1.1</td>
<td>0.9</td>
</tr>
<tr>
<td>VSI EWMA3</td>
<td>57.5</td>
<td>21.8</td>
<td>6.1</td>
<td>1.9</td>
<td>1.2</td>
<td>0.9</td>
<td>0.8</td>
</tr>
</tbody>
</table>
control signal for the 4th sample, there are some differences in identifying causes for the shift. For the VSI EWMA, we can easily find this shift is mainly caused by the shift of the slope because $SE_E(4) = 1.178$ is the maximum. Note that the ERL chart of Zhang et al. (2009) concludes that this shift signal is caused by a deviation shift. Gupta et al. (2006), on the other hand, find that there are problems for both slope and deviation. But they also notice that the deviation values for the 5th and 6th sample are below the lower control limit, which implies the signal for deviation is not so strong. So the conclusion made by our VSI EWMA that the shift is mainly caused by slope seems reasonable. Considering this sample data set is small, further research may be needed to identify the causes of the shift for this real data application.

6. Computational aspects

The Fortran programs used in previous sections are available from the authors upon request. A user can also code his own computer program by the guidelines below for finding $\omega$:

1. Input parameters to be specified by the practitioners, including $x_1, x_2, \ldots, x_n, p_0, d_1, L_1, L_2, L_3$.
2. Set the maximum of $\omega$ to $\omega_{\text{max}} = 1$ and the minimum of $\omega$ to $\omega_{\text{min}} = 0$ and set $\omega = (\omega_{\text{max}} + \omega_{\text{min}})/2$.
3. Set $\text{ATS} = 0$. Generate standard normal random variables to compute $y_j$ and obtain $A(j)$. If $A(j) > \omega$, update ATS to $\text{ATS} = \text{ATS} + d_1$ and update ATS to $\text{ATS} = \text{ATS} + d_2$ otherwise.
4. Repeat step (3) 8 times which is the Monte Carlo sample size and obtain the average of the ATS. If ATS is in-control ARL, update $\omega_{\text{max}}$ to $\omega$; otherwise, update $\omega_{\text{min}}$ to $\omega$.
5. The algorithm is not stopped until the absolute difference of ATS and in-control ARL is less than a prefixed value $e$.

Based on the algorithm above, all the results are implemented in Fortran 95 with IMSL package. Routine “rnnor” is used to generate standard normal random variables. The computation time is highly correlated to the total Monte Carlo sample size $B$ and the prefixed value $e$. Hence larger value of $B$ and smaller values of $e$ will lead the result with higher accuracy, but more time consuming, and vice versa. To balance the time and accuracy, we recommend that $B = 5000$ and $e = 1$. The execution time is <10 s on a Pentium 4 with CPU processor 3.00 GHz.

7. Conclusion and extension

In this paper, we propose an EWMA scheme with variable sampling intervals for monitoring linear profiles. The proposed chart has the following positive features: (1) it can be designed and constructed with just two additional parameters, the conditional probability $p_0$ and the shorter sampling interval $d_1$; (2) it is quite robust and sensitive to various types of shifts, including intercept shifts, slope shifts, standard deviation shifts and combinations of intercept and slope shifts; (3) it does not need extra diagnostic aids and (4) it can control the individual in-control average false alarm rate.

The properties of VSI EWMA$_3$ have been in this paper under the assumption that the observations from the process are normally distributed. For some processes, this assumption may not be realistic. The VSI feature could of course be used with non-normal observations, but in this case, the properties and design strategies would need to be developed under a model which allows for non-normality. Lin and Chou (2005b) study the design of VSS and VSI $\chi$ charts under non-normality based on Burr distribution. This warrants further research.

Acknowledgement

The authors are grateful to the editor and the anonymous referees for their valuable comments that helped clarify the main points and greatly improve this paper. This paper is supported by Natural Sciences Foundation of China (10771107).

References


Table 5

<table>
<thead>
<tr>
<th>No.</th>
<th>$E_1$</th>
<th>$E_2$</th>
<th>$E_3$</th>
<th>$SE_E(1)$</th>
<th>$SE_E(2)$</th>
<th>$SE_E(3)$</th>
<th>$A$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.510</td>
<td>0.979</td>
<td>0.123</td>
<td>0.398</td>
<td>0.164</td>
<td>0.130</td>
<td>0.398</td>
<td>0.100</td>
</tr>
<tr>
<td>2</td>
<td>4.502</td>
<td>0.977</td>
<td>0.079</td>
<td>0.194</td>
<td>0.004</td>
<td>0.084</td>
<td>0.194</td>
<td>1.225</td>
</tr>
<tr>
<td>3</td>
<td>4.504</td>
<td>0.978</td>
<td>0.000</td>
<td>0.234</td>
<td>0.101</td>
<td>0.000</td>
<td>0.234</td>
<td>1.225</td>
</tr>
<tr>
<td>4</td>
<td>4.524</td>
<td>0.990</td>
<td>0.453</td>
<td>0.736</td>
<td>0.117</td>
<td>0.575</td>
<td>1.178</td>
<td>1.225</td>
</tr>
<tr>
<td>5</td>
<td>4.522</td>
<td>0.991</td>
<td>0.238</td>
<td>0.684</td>
<td>1.232</td>
<td>0.252</td>
<td>1.232</td>
<td>0.100</td>
</tr>
<tr>
<td>6</td>
<td>4.522</td>
<td>0.989</td>
<td>0.000</td>
<td>0.692</td>
<td>1.088</td>
<td>0.000</td>
<td>1.088</td>
<td>0.100</td>
</tr>
</tbody>
</table>