Key Words: Variable Sampling Rate, Economic Design, Markov Chain

INTRODUCTION

In continuous production processes control charts are widely used in monitoring changes in process parameters. The traditional approach to sampling for a control chart is to take the fixed sampling rate (FSR) in which we take a fixed sample size (FSS) with a fixed sampling interval (FSI) between samples. In recent years there have been investigations of control charts in which the sampling rate is varied during the operations of the chart as a function of the data from the process. Control charts to vary the sampling interval are called variable sampling interval (VSI) charts, in which a short sampling interval is used when there is an indication of a problem and a long sampling interval is used when there is no indication of a problem.

Another way to vary the sampling rate as a function of the process data is to vary the sample size. Control charts to vary the sample size are called variable sample size (VSS) charts, in which a large sample size is used when there is an indication of a problem and a small sample size is used when there is no indication of a problem.

Variable sampling rates (VSR) charts allow both the sample size and the sampling interval to vary depending on the previous value of the control statistic. The idea of the VSR chart is to combine the VSI and VSS features. VSR $\bar{X}$ charts were considered by Prabhu, Montgomery and Runger (1994, 1997), Costa (1999), and Park and Reynolds (1999). VSR CUSUM charts were considered Rendtel (1990), and Arnold and Reynolds (1994, 2001). Reynolds and Arnold (2001) considered VSR EWMA charts.

In the traditional FSR control charts the performance of the chart is measured by the average run length (ARL) for comparing the efficiencies of the control schemes. Since the sample size and the lengths of the sampling intervals are varying during the process in the VSR charts, it is necessary to develop alternative measures for the performance of the chart. An example of the measures is the average time to signal (ATS) which represents the expected length of time from a specified time point to the time that the chart signals. Arnold and Reynolds (2001), and Reynolds and Arnold (2001) considered a statistical design of VSR CUSUM charts and VSR EWMA charts, respectively. They developed a method to select design parameters by fixing the in-control ATS and the average in-control sampling rate. Prabhu, Montgomery and Runger (1997), and Park and Reynolds (1999) considered an economic design of VSR $\bar{X}$ charts.

The objective of this paper is to develop the VSR EWMA chart and investigate the effectiveness of VSR EWMA charts relative to FSR EWMA charts in the context of an economic model. We use the same economic model as the one developed by Park and Reynolds (1999). Also design parameters of the economic model, which can achieve the optimal economic performance, are selected for some given mean shifts. The effectiveness of VSR EWMA charts is also compared to VSR $\bar{X}$ charts in the economic model.

DESCRIPTION OF THE VSR EWMA CHART

Consider the problem of monitoring a process. Let $X$ be the measurable quality characteristic of interest. Assume that $\{X_{t1}, X_{t2}, \cdots, X_{tn}\}$ is a random sample at time $t$ obtained from a normal distribution with mean, $\mu_t$, and standard deviation, $\sigma$, and the objective is to detect shifts in $\mu_t$ from a target value $\mu_0$. Suppose the random samples of variable size are taken at intervals of variable length during the operation of this process.

Let $N_t$ be the sample size used at the $t$-th sampling time and $H_t$ be the sampling interval used between sampling times $t-1$ and $t$. Let $Z_t = \sqrt{N_t} (\bar{X}_t - \mu_0)/\sigma$ be the standardized sample mean for sample $t$. Then the standardized VSR EWMA
The sampling interval $h$ can be chosen to be between a minimum feasible sampling interval $h_{\text{min}}$ and a maximum feasible sampling interval $h_{\text{max}}$. Then for $t \geq 2$, the sampling interval $H_t$ can be represented as

$$H_t = \left\{ \begin{array}{ll} h_1 & \text{if } |E_{t-1}| < c t \\ h_2 & \text{if } c t \leq |E_{t-1}| < c \end{array} \right.,$$

where $h_{\text{min}} \leq h_2 \leq h_1 \leq h_{\text{max}}$, and $c t$ denotes the threshold limit to switch between the two sampling intervals.

Although no optimality results are known for the choice of the sample size, we consider only two possible sample sizes, $n_1$ and $n_2$ for simplicity. It is assumed that there is a minimum feasible sample size $n_{\text{min}}$ and a maximum feasible sample size $n_{\text{max}}$. The minimum sample size can usually be set as $n_{\text{min}} = 1$, but if a variance estimate is required at each sample then the minimum sample size can be set as $n_{\text{min}} = 2$. Then for $t \geq 2$, the sample size $N_t$ can be represented as

$$N_t = \left\{ \begin{array}{ll} n_1 & \text{if } |E_{t-1}| < c S \\ n_2 & \text{if } c S \leq |E_{t-1}| < c \end{array} \right.,$$

where $n_{\text{min}} \leq n_1 \leq n_2 \leq n_{\text{max}}$, and $c S$ denotes the threshold limit to switch between the two sample sizes.

The sampling interval $H_t$ and the sample size $N_t$ can be represented together as

$$(H_t, N_t) = \left\{ \begin{array}{ll} (h_1, n_1) & \text{if } 0 \leq |E_{t-1}| < \min\{c t, c S\} \\ (h_*, n_*) & \text{if } \min\{c t, c S\} \leq |E_{t-1}| \leq \max\{c t, c S\} \\ (h_2, n_2) & \text{if } \max\{c t, c S\} \leq |E_{t-1}| < c \end{array} \right.,$$

where $h_1 = h_{\text{min}}$ and $h_2 = h_{\text{max}}$, and $c S$ is assumed that there is a minimum feasible sample size $n_{\text{min}}$ and a maximum feasible sample size $n_{\text{max}}$. The minimum sample size can usually be set as $n_{\text{min}} = 1$, but if a variance estimate is required at each sample then the minimum sample size can be set as $n_{\text{min}} = 2$. Then for $t \geq 2$, the sample size $N_t$ can be represented as

$$N_t = \left\{ \begin{array}{ll} n_1 & \text{if } |E_{t-1}| < c S \\ n_2 & \text{if } c S \leq |E_{t-1}| < c \end{array} \right.,$$

where $n_{\text{min}} \leq n_1 \leq n_2 \leq n_{\text{max}}$, and $c S$ denotes the threshold limit to switch between the two sample sizes.

The sampling interval $H_t$ and the sample size $N_t$ can be represented together as

$$(H_t, N_t) = \left\{ \begin{array}{ll} (h_1, n_1) & \text{if } 0 \leq |E_{t-1}| < \min\{c t, c S\} \\ (h_*, n_*) & \text{if } \min\{c t, c S\} \leq |E_{t-1}| \leq \max\{c t, c S\} \\ (h_2, n_2) & \text{if } \max\{c t, c S\} \leq |E_{t-1}| < c \end{array} \right.,$$

where $h_1 = h_{\text{min}}$ and $h_2 = h_{\text{max}}$, and $c S$ is the threshold limit to switch between the two sampling intervals.

In the economic design of a control scheme, the effectiveness of a control procedure can be measured by the long-run average cost per unit time, which considers various costs incurred in a process cycle (see, e.g., Lorenzen and Vance (1986)). In evaluating the long-run average cost, a cycle of the process is first defined and the expected cost per cycle is divided by the expected cycle length.

Suppose the process starts out with the process mean $\mu_1$ at the target value $\mu_0$ and $\mu$ remains at this target until a special cause occurs and produces a shift to some value $\mu_1(\neq \mu_0)$. We assume that the time, say, $T_0$, until a special cause occurs follows an exponential distribution with mean $1/\lambda$. When the special cause occurs and shifts the mean to $\mu_1$, it is assumed that $\mu$ remains at $\mu_1$ until the control chart signals and the special cause is found and removed. Let $T_1$ be the length of time that $\mu$ remains at the shifted value. A cycle of process is defined here as the time from the start of the process to the true out-of-control signal. Then the length of a cycle, say, $T$, can be expressed as $T_0 + T_1$.

During an in-control period there are costs due to sampling and false alarms, and during an out-of-control period there are costs due to sampling and operating with $\mu$ off target. The model contains cost parameters which allow the specification of the costs associated with sampling, false alarms, and operating off target. In calculating the expected cost per cycle we define the following cost parameters.
$a$: the fixed cost for taking a sample

$b$: the additional cost for each individual observation in the sample

$C_F$: the cost of a false alarm

$C_T$: the cost per unit time due to operating off target

Let $S$ and $O$ be the number of samples, and the total number of individual observations, respectively, taken during a cycle. And let $T_0$ be the number of false alarms occurred during an in-control period, $T_0$. Then the expected cost per cycle is calculated as

$$L = \frac{a \cdot E(S) + b \cdot E(O) + C_F \cdot E(F_0) + C_T \cdot E(T_1)}{E(T)}$$

(3)

The long-run cost per unit time, say, $L$, can be expressed as the ratio of the expected cost per cycle to the expected cycle length, that is,

$$L = \frac{a \cdot E(S) + b \cdot E(O) + C_F \cdot E(F_0) + C_T \cdot E(T_1)}{E(T)}$$

(3)

To evaluate the long-run cost per unit time, we should obtain the expectations $E(S)$, $E(O)$, $E(T)$, and $E(F_0)$. Note that $E(T_1) = E(T) - 1/\lambda$. The required expectations are evaluated by using the Markov chain approach. The continuation region, $(-c, c)$, of the standardized VSR EWMA chart statistic in equation (1) is divided into several subregions and each subregion is represented by a transient state while the signal region is represented by an absorbing state. Details about the Markov chain approach and derivations of the required expectations are omitted in this paper.

The economic model in equation (3) can be refined by considering additional factors, such as positive times associated with sampling, searching for special causes, and repair. But it was known that these additional factors should not have a large effect on the final results from the model unless the values chosen for the corresponding parameters are large (see, e.g., Park and Reynolds (1999)).

In designing an economic VSR EWMA chart, the objective is to find the set of eight chart parameters \(\{c, c_S, c_f, n_1, n_2, h_1, h_2, r\}\) which minimize the long-run cost per unit time in equation (3) with constraints $n_{min} \leq n_1 \leq n_2 \leq n_{max}$, $h_{min} \leq h_2 \leq h_1 \leq h_{max}$ for given process and cost parameters. This nonlinear optimization problem with constraints was solved by using the generalized reduced gradient (GRG) procedure with finite difference approximations to partial derivatives (see, e.g., Lasdon et al. (1978)).

It is assumed that the sample sizes must be between $n_{min} = 1$ and $n_{max} = 50$, the sampling intervals must be between $h_{min} = 0.1$ and $h_{max} = 10$, and an hour is used as the unit of time. Although the sample sizes $n_1$ and $n_2$ of the VSR case are always integer values, they are first treated as continuous variables in finding optimal values for the eight chart design parameters. In order to complete the optimization procedure four cases are considered where $n_1$ and $n_2$ are set to be the larger and smaller nearest integers to the previously obtained sample size values and the values previously obtained for the other six design parameters are used as the starting point to re-optimize with the new integer values for the sample sizes. Optimal parameters of the FSR case are obtained similarly by restricting $n_1 = n_2$ and $h_1 = h_2$. For a given set of input parameters, the optimization problem for the FSR case is solved first. The following values are used as the starting point.

$$c_f = 3.0, \ n_f = 10.0/\delta, \ h_f = 0.5, \ r_f = 0.5,$$

where $c_f, n_f, h_f$, and $r_f$ denote the control limit, the sample size, the sampling interval, and the weight of the EWMA chart for the FSR case, respectively, and $\delta$ denotes the amount of mean shift in units of the process standard deviation. After the optimal design parameters for the FSR case are obtained, the optimization problem for the VSR case is solved by using the following values as the starting point.

$$c_I = (0.4) \ c_f, \ c_S = (0.6) \ c_f, \ c = (1.1) \ c_f,$$

$$h_1 = h_f, \ h_2 = h_{min}, \ n_1 = (0.5) \ n_f, \ n_2 = (1.5) \ n_f,$$

$$r = \begin{cases} 0.3 & \text{if } \delta \leq 1.5, \\ 1.0 & \text{if } \delta > 1.5 \end{cases}$$

The optimal design parameters for the VSR case are easily obtained by use of the GRG procedure in most cases. But occasionally the optimal design parameters are not obtained correctly, which can be verified when the minimum expected cost per hour of the VSR case is larger than the FSR case or the VSR case with a restriction of $c_S = c_I$. In such cases a new set of the starting point, which is set by changing the previously obtained design parameters, is used to find the correct optimal design parameters.

**COMPARISON OF OPTIMAL VSR AND FSR EWMA CHARTS**

In this section the optimal FSR and VSR designs are obtained for some sets of values of process and cost parameters which might be typical cases of applications. The objective of examining these designs is to compare the designs and costs of the FSR and VSR charts and to gain insight into the nature of the optimal design for the VSR EWMA chart.
We select process and cost parameters as follows:

\[ \delta \in \{0.5, 2.0\}, \ C_F = 200, \ C_T = 2 \ C_F, \]
\[ \lambda \in \{0.01, 0.001\}, \ a \in \{0, 1\}, \ b \in \{0.1, 1\}. \]

The optimal chart parameters and the properties of the FSR and VSR charts for the set of parameters are obtained in Table 1 to 2. The two tables are set up in the same way for different combinations of \( \delta \) and \( C_F \). Similar tables are also obtained for cases when \( \delta = 1.0, 3.0, \) and \( C_F = 50 \), but these tables are not shown here. The entries in the first row are for the FSR EWMA chart, the entries in the second row are for the VSR EWMA chart and the entries in the third row are for the VSR EWMA chart with the restriction that \( c_S = c_f \). The first three columns in the table give the values of \( \lambda, a, \) and \( b \), respectively, and the optimal design parameters of the charts are given in the following columns. The column labeled \( \text{Obs./hr.} \) gives \( E(O_0)/E(T_0) \), that is, the long-run average number of observations taken per hour during in-control periods of the process. This column allows for a comparison of the average sampling rate of the charts when the process is in control. The column labeled \( \text{F.A./hr.} \) gives \( E(F_0)/E(T_0) \) times 1000, that is, 1000 times the long-run average number of false alarms per hour during in-control periods. This column allows for a comparison of the false alarm rates of the charts. The column labeled \( E(T_1) \) gives the expected detection time for a given shift, that is, the expected time from the process shift until the signal.

The last column labeled \( L(PR) \) denotes the minimum expected cost per hour for each chart with, in parentheses, the percent reduction on cost from using the optimal VSR EWMA chart instead of the optimal FSR EWMA chart. The percent reduction on cost is calculated as

\[ PR = \frac{L_{FE} - L_{VE}}{L_{FE}} \times 100, \]

where \( L_{FE} \) and \( L_{VE} \) denote the minimum expected cost per hour of the FSR and VSR EWMA charts, respectively.

For the parameter combinations considered in Table 1 to 2, the percent reduction on cost ranges from a modest 5.6% to a substantial 36.3%. Restricting the VSR EWMA chart to have \( c_S = c_f \) does not result in large changes in the other design parameters and gives almost as large as or at most a modestly larger than the VSR chart without restriction. It appears that restricting \( c_S = c_f \) may be used for applications in which administrative convenience is more important than a modest reduction in the long-run average cost. Discussions of the numerical results in this section are for the VSR EWMA chart without restriction \( c_S = c_f \).

Major differences found in comparing the optimal VSR EWMA chart with the optimal FSR EWMA chart are that the control limit \( c \) of the VSR chart is considerably higher than the control limit \( c_f \) of the FSR chart, and that the long sampling interval \( h_I \) and the weight \( r \) of the VSR chart are considerably smaller than the fixed sampling interval \( h_f \) and the weight \( r_f \) of the FSR chart. We could not find any clear trend of the sample size of the VSR chart when compared to the FSR chart. The VSR chart has a much smaller false alarm rate (F.A./hr.) and modestly smaller value of \( E(T_1) \) than the FSR chart, while the long-run average number of observations (Obs./hr.) is more or less the same. It seems that the much smaller false alarm rate of the VSR EWMA chart is caused from the larger control limit and the smaller weight of the VSR chart compared to the FSR case. Comparing with the FSR EWMA chart, the efficiency of the VSR EWMA chart is mainly achieved by a smaller number of false alarm rate, a smaller detection time, and depending more on the past data.

The weight of the optimal FSR EWMA chart is larger than 0.9 for many cases. This fact implies that using the EWMA chart of the FSR scheme in an economic model is not as much effective as in a statistical model. In general, the weight of the VSR EWMA chart is less than that of the FSR EWMA chart and this means that EWMA scheme can be more efficiently adopted to the VSR scheme than the FSR scheme.

**THE OPTIMAL VSR EWMA CHART DESIGN**

When a VSR EWMA chart is being designed for a particular application, it is better to use information from the FSR EWMA chart if available. In case where there is no FSR EWMA chart in current use, it is useful to know how the VSR chart is designed in relation to the various process and cost parameters.

The process parameter \( \lambda \) determines how often a special cause occurs. A small value of \( \lambda \) indicates a less often occurrence of a special cause in the process. The main effect of changing \( \lambda \) is on the sampling interval as might be expected. Increasing \( \lambda \) decreases the long sampling interval \( h_I \), but the optimal value for the short sampling interval is the minimum possible sampling interval. The threshold limit for sampling interval \( c_f \) is always smaller than the threshold limit for sample size \( c_S \). This means that the short sampling interval is more often used than the large sample size is.
The process parameter $\delta$ determines the likelihood of the location shift with larger values of $\delta$ corresponding to large shifts being more likely to occur. Increasing $\delta$ tends to decrease $h_1$ in order to detect large shifts quickly and also tends to decrease the two sample sizes $n_1$ and $n_2$ due to easiness of detecting large shifts. Increasing $\delta$ also increases the limits $c_1$ and $c$. The effects of increasing $\delta$ are to decrease the false alarm rate and to make the chart less likely switch from a low sampling intensity to a high sampling intensity.

When the sampling cost for observation is large ($b = 1$) the two sample sizes are equal or at least closer to each other than when the sampling cost for observation is small ($b = 0.1$). This fact indicates that the VSS scheme is not very useful when the sampling cost for observation is large. The general effect of increasing sampling costs $a$ and/or $b$ is to increase the long sampling interval $h_1$ and to decrease the control limit $c$. Increasing the fixed cost of a sample $a$ tends to increase the two sample sizes, and increasing the additional cost for each individual observation $b$ tends to decrease the two sample sizes. The joint effect of increasing the false alarm cost $C_F$ and $C_T$ is to increase the average sampling rate, that is, decreases $h_1$ and increases the two sample sizes, and to increase the control limit $c$.

The facts described for the VSR EWMA charts are mostly consistent with the conclusions obtained in previous studies of the VSR $X$ chart by Park and Reynolds (1999).

**CONCLUSIONS**

An economic design model has been developed by Park and Reynolds (1999) for evaluating the long-run cost per hour associated with the operation of the VSR and FSR $X$ charts. This economic design model has been applied here again for evaluating the long-run cost per hour associated with the operation of the VSR and FSR EWMA charts. The optimal parameters of the VSR and FSR EWMA charts can be obtained in this economic model for given values of the process and cost parameters.

For most of the parameter combinations considered here, it is shown that the control limit of the VSR chart is considerably higher, and the weight of the VSR chart is considerably smaller than that of the FSR chart. This results in a much smaller false alarm rate of the VSR chart. The percent reduction on cost presented show that applying VSR scheme in place of FSR scheme to the EWMA chart design can result in moderate to substantial cost savings. Using the same threshold limit for both the sample size and the sampling interval does not lead to a large increase in cost, which suggests to implement the VSR EWMA chart with just one threshold limit in most applications.

In this paper we use a single special cause process model. For a further study we will consider the economic model and the economic performance of VSR EWMA charts under a process model with multiple special causes.

**REFERENCES**


### TABLE 1. Optimal FSR and VSR EWMA chart designs for $\delta = 0.5$, $C_F = 200$, and $C_T = 400$

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$\alpha$</th>
<th>$c_T$</th>
<th>$c_f$</th>
<th>$c_S$</th>
<th>$c$</th>
<th>$h_f$</th>
<th>$n_f$</th>
<th>$r_f$</th>
<th>$Q_{obs\over hr}$</th>
<th>$FA_{hr}$</th>
<th>$E(T_1)$</th>
<th>$L(PR)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.1</td>
<td>0.95</td>
<td>2.04</td>
<td>3.71</td>
<td>1.46</td>
<td>3.51</td>
<td>0.75</td>
<td>1</td>
<td>0.196</td>
<td>3.80</td>
<td>50</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.78</td>
<td>1.76</td>
<td>3.08</td>
<td>1.35</td>
<td>3.41</td>
<td>1.62</td>
<td>1</td>
<td>0.909</td>
<td>3.93</td>
<td>32</td>
<td>0.88</td>
</tr>
<tr>
<td>0.001</td>
<td>0.1</td>
<td>0.95</td>
<td>2.04</td>
<td>3.71</td>
<td>1.46</td>
<td>3.51</td>
<td>0.75</td>
<td>1</td>
<td>0.196</td>
<td>3.80</td>
<td>50</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.78</td>
<td>1.76</td>
<td>3.08</td>
<td>1.35</td>
<td>3.41</td>
<td>1.62</td>
<td>1</td>
<td>0.909</td>
<td>3.93</td>
<td>32</td>
<td>0.88</td>
</tr>
</tbody>
</table>

### TABLE 2. Optimal FSR and VSR EWMA chart designs for $\delta = 2.0$, $C_F = 200$, and $C_T = 400$

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$\alpha$</th>
<th>$c_T$</th>
<th>$c_f$</th>
<th>$c_S$</th>
<th>$c$</th>
<th>$h_f$</th>
<th>$n_f$</th>
<th>$r_f$</th>
<th>$Q_{obs\over hr}$</th>
<th>$FA_{hr}$</th>
<th>$E(T_1)$</th>
<th>$L(PR)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.1</td>
<td>1.08</td>
<td>1.90</td>
<td>4.19</td>
<td>1.62</td>
<td>4.06</td>
<td>0.27</td>
<td>1</td>
<td>0.104</td>
<td>3.76</td>
<td>1.34</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1.04</td>
<td>2.00</td>
<td>3.76</td>
<td>1.36</td>
<td>3.59</td>
<td>0.68</td>
<td>1</td>
<td>0.623</td>
<td>2.92</td>
<td>0.92</td>
<td>7</td>
</tr>
<tr>
<td>0.001</td>
<td>0.1</td>
<td>1.04</td>
<td>2.00</td>
<td>3.76</td>
<td>1.36</td>
<td>3.59</td>
<td>0.68</td>
<td>1</td>
<td>0.623</td>
<td>2.92</td>
<td>0.92</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1.04</td>
<td>2.00</td>
<td>3.76</td>
<td>1.36</td>
<td>3.59</td>
<td>0.68</td>
<td>1</td>
<td>0.623</td>
<td>2.92</td>
<td>0.92</td>
<td>7</td>
</tr>
</tbody>
</table>

The tables provide the optimal FSR and VSR EWMA chart designs for different values of $\delta$, with specific parameters for $C_F = 200$ and $C_T = 400$. Each row in the tables corresponds to a different set of parameters, including $\lambda$, $\alpha$, $c_T$, $c_f$, $c_S$, $h_f$, $n_f$, $r_f$, $Q_{obs\over hr}$, $FA_{hr}$, $E(T_1)$, and $L(PR)$. The values are calculated based on the given formulas and constraints.