

Stochastics and Statistics

Joint economic design of EWMA control charts for mean and variance

Doğan A. Serel^{a,*}, Herbert Moskowitz^b

^a Faculty of Business Administration, Bilkent University, 06800 Bilkent, Ankara, Turkey

^b Krannert Graduate School of Management, Purdue University, West Lafayette, IN 47907-2056, USA

Received 17 January 2006; accepted 30 September 2006

Available online 18 December 2006

Abstract

Control charts with exponentially weighted moving average (EWMA) statistics (mean and variance) are used to jointly monitor the mean and variance of a process. An EWMA cost minimization model is presented to design the joint control scheme based on pure economic or both economic and statistical performance criteria. The pure economic model is extended to the economic-statistical design by adding constraints associated with in-control and out-of-control average run lengths. The quality related production costs are calculated using Taguchi's quadratic loss function. The optimal values of smoothing constants, sampling interval, sample size, and control chart limits are determined by using a numerical search method. The average run length of the control scheme is computed by using the Markov chain approach. Computational study indicates that optimal sample sizes decrease as the magnitudes of shifts in mean and/or variance increase, and higher values of quality loss coefficient lead to shorter sampling intervals. The sensitivity analysis results regarding the effects of various inputs on the chart parameters provide useful guidelines for designing an EWMA-based process control scheme when there exists an assignable cause generating concurrent changes in process mean and variance.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Quality control; Average run length; Economic design; EWMA control chart; Markov chain

1. Introduction

Control charts are used for monitoring the level of variation in a production process over time with the objective of reacting quickly to harmful deviations from the normal operating conditions (as well

as indicating the effects of process changes and improvements).¹ To implement control charts in practice, control limits, sample size, and sampling frequency must be specified. Once these design parameters are determined, the processes can be controlled against assignable (special) causes leading to undesirable process output. The sample

* Corresponding author. Tel.: +90 312 290 2415; fax: +90 312 266 4958.

E-mail address: serel@bilkent.edu.tr (D.A. Serel).

¹ Control charts are also used to monitor the impact of process improvements. Its design parameters can then be readjusted to reflect their effects, and then the process remonitored for out-of-control conditions associated with special causes.

statistics computed based on random samples taken from the process are compared against control limits, and a decision is made regarding whether the process is currently in-control or out-of-control. False alarms occur when an in-control process is erroneously classified as out-of-control (type 1 error) and an assignable cause is searched. There also exists the risk of concluding that the process is in-control based on the sample test statistic although the process is actually out-of-control (type 2 error).

Various types of quality control charts have been proposed to monitor the process mean and/or variability. Although the Shewhart \bar{X} -bar chart is the most common chart applied in practice to monitor the process mean, in the last 20 years exponentially weighted moving average (EWMA) charts have been offered as a suitable alternative to the \bar{X} -bar chart for detecting small to moderate shifts in the process mean. Similarly, various alternative control charts exist for monitoring process variation (Acosta-Mejia et al., 1999). Generally control charts for mean and dispersion are used simultaneously since, due to a special cause, one of these two parameters may deviate from its in-control value while the other parameter remains unchanged. Sometimes a single assignable cause may result in changes in both process mean and variation. For example, in integrated circuit manufacturing, the solder paste is printed onto the printed circuit board (PCB) before the mounting of circuit components. The thickness of the solder paste influences the solderability of circuit components on the board. When the process goes out of control, the thickness of the paste is off-target and at the same time the process variance is large since the solder paste thickness is not uniformly distributed over the board (Gan et al., 2004). Thus, process mean and variance are simultaneously affected by the same special cause in this manufacturing setting.

The statistical performance of EWMA based control charts for monitoring the process mean and variation jointly have been studied in several papers (e.g. Morais and Pacheco, 2000; Reynolds and Stoumbos, 2001; Knoth and Schmid, 2002; Reynolds and Stoumbos, 2004). The *statistical design* of control charts takes into account the in-control and out-of-control average run lengths resulting from the sample size and the control limits chosen by the user. Average run length (ARL) is a measure of the expected number of consecutive samples taken until the sample statistic falls outside

the control limits. Since ARL is a function of the prevailing process mean and standard deviation, its value depends on whether the process is in-control or out-of-control. When multiple charts are used jointly for monitoring the process, the investigation for an assignable cause is initiated when at least one of the charts triggers an out-of-control signal. Hence, not the ARLs of the individual charts but the joint ARL of the overall control scheme is the relevant performance measure when multiple charts are used simultaneously. Since the statistical design of control charts does not explicitly take into account the dependence of the sampling, inspection, and defective product costs on the chart parameters selected, some researchers have suggested the formulation of economic models as an alternative method to a purely statistical approach for designing the control charts.

The *economic design* approach to control charts advocates the determination of the control chart design parameters based on a cost-minimization model that takes into account all costs affected by the choice of these parameters (Lorenzen and Vance, 1986). The economic design of control charts for monitoring the process mean has been investigated extensively in the literature (see, e.g. Montgomery, 1980; Ho and Case, 1994a). In the literature on control charts employing an EWMA type statistic, several authors have explored the economic design of EWMA control charts to monitor the process mean (Ho and Case, 1994b; Montgomery et al., 1995). Park et al. (2004) extended the traditional economic design of an EWMA chart to the case where the sampling interval and sample size may vary depending on the current chart statistic. Tolley and English (2001) studied the economic design of a control scheme combining both EWMA and \bar{X} -bar charts.

Another research stream in the literature has considered the joint economic design of mean and variation control charts. Although the joint economic design of \bar{X} -bar and R (range) or \bar{X} -bar and S (standard deviation) control charts has been studied by several researchers (e.g., Saniga, 1979; Rahim, 1989; Rahim and Costa, 2000; McWilliams et al., 2001), joint economic design of EWMA based control charts for monitoring both process mean and dispersion does not appear to have been previously investigated. We propose such a model that can be used to design a joint EWMA-based mean (μ) and variance (σ^2) control scheme. Our model is built upon the general cost function of Lorenzen

and Vance (1986) which is applicable to different types of control charts. In addition to pure economic design of control charts, we also consider *statistically constrained economic designs*. Since the in-control ARL of economically designed charts often can be significantly less than what users prefer, some researchers have proposed to reformulate the pure economic model by incorporating additional constraints on the statistical performance of the model (e.g., McWilliams et al., 2001). In our statistically constrained economic design, the solution is required to satisfy two constraints specified by the user: (a) the in-control ARL should be greater than a lower bound, and (b) the out-of-control ARL should be less than an upper bound.

The rest of the paper is organized as follows. After describing the EWMA control charts for mean and variance in Section 2, we present the cost model used in our research in Section 3. Subsequently, the main aspects of the numerical search method used in finding the optimal solution to the cost model are discussed in Section 4. In Section 5, some numerical examples of EWMA control charts that are designed based on either an economic or economic-statistical criterion are presented. Finally, we offer a summary of the results and concluding remarks in Section 6.

2. EWMA control charts

We assume that the observations for the process variable X are independent and normally distributed. When the process is in control, the mean and variance of X is μ_0 and σ_0^2 , respectively. At any sampling instant t , the sample mean and variance are computed from

$$\bar{X}_t = \sum_{i=1}^n X_{it}/n \quad \text{and}, \quad (1)$$

$$S_t^2 = \sum_{i=1}^n (X_{it} - \bar{X}_t)^2 / (n - 1), \quad (2)$$

where \bar{X}_t and S_t^2 are the sample mean and variance at time t , and n is the fixed sample size, $n \geq 2$. Using \bar{X}_t and S_t^2 , the chart statistics are calculated as

$$Z_t = \lambda_m \bar{X}_t + (1 - \lambda_m) Z_{t-1}, \quad (3)$$

$$Y_t = \max\{\ln(\sigma_0^2), \lambda_v \ln(S_t^2) + (1 - \lambda_v) Y_{t-1}\}, \quad (4)$$

where λ_m and λ_v are the smoothing constants associated with the EWMA chart for mean (EWMA-m) and variance (EWMA-v), respectively, $0 < \lambda_m, \lambda_v \leq$

1, $Z_0 = \mu_0$, $Y_0 = \ln(\sigma_0^2)$. The statistic Z_t is used in the EWMA-m chart, and Y_t is associated with the EWMA-v chart.

When EWMA schemes are used for process monitoring, not only the current observations of X but also the observations from previous samples are taken into account. In the computation of the test statistic, more recent samples are given a larger weight than the ones taken earlier. The user can increase the weight given to the last sample by increasing the value of the smoothing constant.

Lucas and Saccucci (1990) describe the properties of the EWMA-m chart in detail. We use the lower and upper control limits (LCL_m and UCL_m) computed based on the asymptotic in-control standard deviation of the EWMA chart statistic Z such that

$$LCL_m = \mu_0 - L_m \sigma_z, \quad (5)$$

$$UCL_m = \mu_0 + L_m \sigma_z, \quad (6)$$

where $\sigma_z = \sigma_0(\lambda_m / (2 - \lambda_m)n)^{0.5}$, and L_m is the control limit parameter. Thus, whenever Z_t is outside the interval (LCL_m, UCL_m) , the process is considered to be out of control and a search for assignable cause is conducted. Due to the natural variation of the process, out-of-control signals may also occur while the process is in control. However, when there is a shift in process mean and/or variance, the chart will generate an out-of-control signal much more quickly.

To monitor the process dispersion, a number of authors have previously studied the control charts based on EWMA of $\ln S^2$ (Crowder and Hamilton, 1992; Gan, 1995; Acosta-Mejia et al., 1999). The particular dispersion chart we adopt in this study is referred to as EWMA-v which has the lower and upper control limits as follows:

$$LCL_v = \ln(\sigma_0^2), \quad (7)$$

$$UCL_v = \ln(\sigma_0^2) + L_v \sigma_y, \quad (8)$$

where $\sigma_y^2 = \lambda_v \psi'[(n - 1)/2] / (2 - \lambda_v)$, $\psi'(\cdot)$ is the trigamma function, and L_v is the control limit parameter. The trigamma function is the second logarithmic derivative of the gamma function $\Gamma(\cdot)$, i.e. $\psi'(u) = d^2 \ln[\Gamma(u)] / du^2$, $u > 0$. We consider an upper one-sided EWMA-v chart which is well-suited for detecting the increases in the process standard deviation. An increase in the process standard deviation would either reflect an undesirable special cause or the impact of an undesirable

process change (either purposeful or unpurposeful); a decrease in process variation would indicate the effect of a process improvement. EWMA-v chart yields an out-of-control signal when Y_t exceeds UCL_v . The statistical properties of the combined EWMA-m/EWMA-v control scheme have been explored in [Morais and Pacheco \(2000\)](#).

To illustrate the smoothing effect of exponential weighting, we plot the values of \bar{X}_t and Z_t for 40 samples generated via simulation in [Fig. 1](#). The observations constituting the first 20 samples are generated from the in-control process distribution with $\mu_0 = 0, \sigma_0 = 1$. The samples 21 through 40 are based on the out-of-control process distribution with $\mu_1 = 0.5, \sigma_1 = 1.5$; we also set $n = 4, \lambda_m = 0.2$ in this simulation experiment. The values of $\ln(S_t^2)$ and Y_t (with $\lambda_v = 0.2$) obtained from the same simulation run are displayed in [Fig. 2](#). The lower and upper control limits shown in these charts are based on $L_m = L_v = 2.5$. [Figs. 1 and 2](#) show that the time series of exponentially weighted sample statistics (Z_t and Y_t) exhibit less variability than the original series (\bar{X}_t and $\ln(S_t^2)$) from which they are derived.

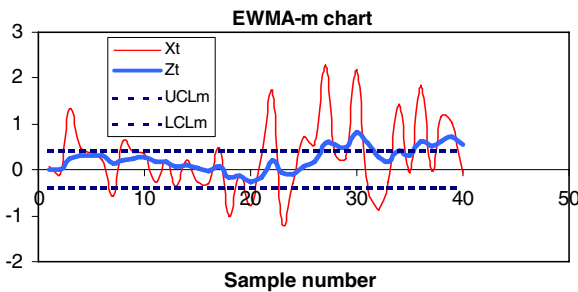


Fig. 1. EWMA control chart for mean – Plot of \bar{X}_t and test statistic Z_t when step changes in both mean and variance occur at sample 21.

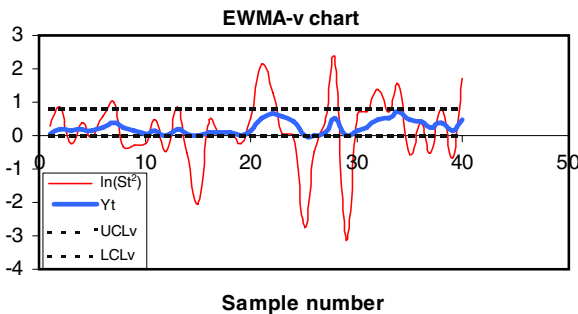


Fig. 2. EWMA control chart for variance – Plot of $\ln(S_t^2)$ and test statistic Y_t when step changes in both mean and variance occur at sample 21.

3. Cost model

3.1. Total cost components

We denote the time between two consecutive samples (sampling interval) by h . It is assumed that the in-control time for the process is distributed exponentially with mean $1/\theta$. The assumption of exponential distribution for failure time is not as restrictive as it seems since previous research has found that results in this kind of economic design models are relatively insensitive to the type of probability distribution used for failure time ([McWilliams, 1989](#)). We allow the possibility that both the process mean and variance may change when an assignable cause occurs. When the process is out of control, the mean of X becomes $\mu_1 = \mu_0 + \delta\sigma_0$ and the standard deviation of X shifts to σ_1 . Using the [Lorenzen and Vance \(1986\)](#) framework, the expected cost per unit time (hour), C , associated with the control scheme consisting of EWMA-m and EWMA-v charts is (cf. Eq. (10) in Lorenzen and Vance):

$$C = \{C_0/\theta + C_1(-\tau + nE + h(ARL_1) + \gamma_1 T_1 + \gamma_2 T_2) + sF/ARL_0 + W\} / \{1/\theta + (1 - \gamma_1)sT_0/ARL_0 - \tau + nE + h(ARL_1) + T_1 + T_2\} + \{[(a + bn)/h][1/\theta - \tau + nE + h(ARL_1) + \gamma_1 T_1 + \gamma_2 T_2]\} / \{1/\theta + (1 - \gamma_1)sT_0/ARL_0 - \tau + nE + h(ARL_1) + T_1 + T_2\}. \tag{9}$$

The parameters in (9) are defined as follows.

- C_0 cost per hour due to nonconformities produced while the process is in control
- C_1 cost per hour due to nonconformities produced while the process is out of control
- τ expected time between the occurrence of the assignable cause and the time of the last sample taken before the assignable cause = $[1 - (1 + \theta h)\exp(-\theta h)] / [\theta(1 - \exp(-\theta h))]$
- E time to sample and chart one item
- ARL_0 average run length while in control
- ARL_1 average run length while out of control
- T_1 expected time to discover the assignable cause
- T_2 expected time to repair the process
- γ_1 = 1 if production continues during searches, 0 if production ceases during searches
- γ_2 = 1 if production continues during repair, 0 if production ceases during repair
- s expected number of samples taken while in control = $\exp(-\theta h) / [1 - \exp(-\theta h)]$

F	cost per false alarm
W	cost to locate and repair the assignable cause
T_0	expected search time when the signal is a false alarm
a	fixed cost per sample
b	cost per unit sampled

The traditional approach to the economic design of control charts involves calculation of the expected cost per hour by dividing the expected cost per cycle by the expected cycle length. Each cycle is made up of two parts: (a) an in-control interval, and (b) an out-of control interval following that. The cost function in (9) is derived by dividing the sum of costs incurred during the in-control and out-of-control segments by the expected cycle length. The expected lengths of the in-control and out-of-control intervals, $E(I_{in})$ and $E(I_{out})$, are

$$E(I_{in}) = 1/\theta + (1 - \gamma_1)sT_0/ARL_0, \tag{10}$$

$$E(I_{out}) = -\tau + nE + h(ARL_1) + T_1 + T_2. \tag{11}$$

The in-control interval in (10) is composed of the expected time until failure and the expected time spent for investigating false alarms. The expected length of the in-control interval depends on whether the production continues during searches or not. The right hand side of (10) follows from the memoryless property of the exponential distribution. The out-of-control interval in (11) includes the time from the occurrence of the assignable cause to the next sampling instant ($h - \tau$), the time until an out-of-control signal ($h(ARL_1 - 1)$), the time to collect and chart a sample (nE), the time to discover the assignable cause (T_1), and the time to repair the process (T_2). After an assignable cause is found and corrective action is taken, the process mean and variance are restored to their in-control values μ_0 and σ_0^2 , and the cycle restarts.

The decision variables are $n, h, L_m, L_v, \lambda_m,$ and λ_v . As defined previously, average run length is the expected number of samples taken before an out of control signal is observed. To minimize false alarms and react swiftly to out-of-control conditions, large values for ARL_0 and small values for ARL_1 are desirable. ARL_0 and ARL_1 depend on all decision variables except h . Since the total cost function is not analytically very tractable, numerical optimization methods must be used to determine the optimal solution.

3.2. Quadratic loss function

Some researchers have suggested the computation of production costs based on a quadratic loss function for economically designing the X-bar control charts (Moskowitz et al., 1994; Elsayed and Chen, 1994; Linderman et al., 2005). In accordance with Taguchi’s quality loss concept, in this approach the quality costs increase quadratically as the deviation of the quality characteristic from its target value increases. Let T be the target value for the quality characteristic monitored, and K be the Taguchi loss coefficient. Although it is desirable that the ideal value (design specification) of the quality characteristic T is equal to the in-control mean μ_0 , under some real-world operating conditions, it may be difficult or costly to adjust the process mean to its target value. If μ_0 is different from T , a fixed bias impacts all manufactured items. Denoting the probability density function of the quality characteristic X by $f(x)$, the expected quality cost per unit of product when the process is in control, J_0 , is

$$\begin{aligned} J_0 &= \int_{-\infty}^{\infty} K(x - T)^2 f(x) dx \\ &= \int_{-\infty}^{\infty} K(x - \mu_0 + \mu_0 - T)^2 f(x) dx \\ &= \int_{-\infty}^{\infty} K(x - \mu_0)^2 f(x) dx + \int_{-\infty}^{\infty} K(\mu_0 - T)^2 f(x) dx \\ &= K[\sigma_0^2 + (\mu_0 - T)^2]. \end{aligned} \tag{12}$$

Let ρ be the ratio of the out-of-control standard deviation to in-control standard deviation, i.e. $\rho = \sigma_1/\sigma_0$. When the process is out of control, the expected quality cost per unit, J_1 , is (cf. Elsayed and Chen, 1994)

$$\begin{aligned} J_1 &= \int_{-\infty}^{\infty} K(x - \mu_1 + \mu_1 - T)^2 f(x) dx \\ &= \int_{-\infty}^{\infty} K(x - \mu_1)^2 f(x) dx + \int_{-\infty}^{\infty} K(\mu_0 + \delta\sigma_0 - T)^2 f(x) dx \\ &= K[\sigma_1^2 + (\mu_0 + \delta\sigma_0 - T)^2] \\ &= K[\rho^2\sigma_0^2 + (\mu_0 - T)^2 + \delta^2\sigma_0^2 - 2\delta\sigma_0(\mu_0 - T)]. \end{aligned} \tag{13}$$

Taguchi loss coefficient K can be estimated by using

$$K = A/(x - T)^2, \tag{14}$$

where A is the cost to rework or scrap one unit of product when the value of quality characteristic is x , and the deviation from target $|x - T|$ is acceptable (Elsayed and Chen, 1994). Let p be the

production rate (units produced per hour). To incorporate the production costs based on quadratic loss into the economic design, we substitute $C_0 = J_0p$, and $C_1 = J_1p$ into (9). After this step we minimize the total hourly cost C to find the best values for the chart parameters. Note that the mean and variance of the quality characteristic directly influence the total cost in the model with quadratic loss.

4. Computational optimization procedure

We use the Markov chain method to compute the average run length of the combined control scheme when the process is in control and out of control. This approach is commonly used in calculating the ARL of control charts (e.g., Saccucci and Lucas, 1990), and is described in more detail in Appendix A. In order to find the optimal solution, we first fix the sample size n and optimize (9) with respect to the other five decision variables. Then we repeat the previous step for all values of n between two and twenty, which is a reasonable upper limit for the sample size from a practical perspective. Finally the best n value and corresponding values of the remaining variables are identified by inspecting the minimum total costs associated with each n value. Thus, the sample size is actually constrained to be less than 20 in our economic design. A similar approach is used in Torng et al. (1995).

We use the Nelder–Mead downhill simplex method in order to find the most economic design for a specific sample size and a given set of input parameters. This widely known numerical method, starting with a set of points defining an initial simplex, searches for an optimal solution to a multivariable minimization problem by using only the function evaluations, and does not require the computation of derivatives (Press et al., 1992). If the number of variables is m , the simplex consists of $m + 1$ vertices. At each iteration, the vertex with the highest cost is removed, and replaced by a new vertex. The search terminates when either the improvement in the objective function value in the last iteration becomes less than a prespecified small threshold value, or the limit on the number of iterations is reached. Based on our numerical experiment, the limit on the number of iterations in the search procedure is set as 300. Although the Nelder–Mead method does not guarantee convergence to the global optimal solution, as discussed in Section 5, in our experiment we have observed that

the cost values resulting from implementations with different initial points are close to each other.

The statistically constrained economic design is found by employing essentially the same algorithm with a slight modification such that the search is forced back into the feasible region when the statistical constraints are violated by the current set of design parameters. This is accomplished by imposing a large penalty cost to the objective function when any ARL constraint is not satisfied. In the statistically constrained optimization case, we minimize (9) subject to two additional constraints: $ARL_0 \geq LB$, and $ARL_1 \leq UB$, where LB and UB are the desired lower and upper bounds on the ARL of the scheme. Note that when statistical constraints exist, the numerical search should be started with a feasible set of vertices.

We also define the search space by placing the following upper limits on the variables: $h \leq 20$, $\lambda_m, \lambda_v \leq 0.99$, $L_m, L_v \leq 4$. From a practical viewpoint, we also set the minimum allowable values for λ_m and λ_v to 0.05.

5. Numerical examples

In the following examples we assume production continues during the search for an assignable cause, but it ceases during repair, i.e. $\gamma_1 = 1, \gamma_2 = 0$. We use the following values of parameters: $\theta \in \{0.01, 0.05\}$, $F = 500$, $W = 250$, $a = 5$, $b = 1$, $E = 0.5$, $T_0 = 0$, $T_1 = 20$, $T_2 = 0$, $p = 200$ per hour, $\sigma_0^2 = 1$, and $\mu_0 = T = 0$. Let ρ be the ratio of the out-of-control standard deviation to in-control standard deviation, i.e. $\rho = \sigma_1/\sigma_0$. The optimal values of design parameters for given different shift values δ and ρ are listed in Table 1 for the joint EWMA scheme with $K = 0.1$. Note that in the economic design model using quadratic loss, J_1 increases with δ and ρ , as defined in (13), implying that cost of a defective unit is higher with greater shifts in mean and/or variance.

Table 2 shows the best statistically constrained designs for the EWMA scheme when $LB = 250$ and $UB = 20$. Economic-statistical designs for $LB = 100$ and $UB = 10$ are displayed in Table 3. Comparison of Table 1 with Tables 2 and 3 indicates that for small shifts in mean and/or variance (that is, $\delta, \rho \leq 1.5$) economic-statistical designs have shorter sampling intervals (h) compared to the economic designs. It is also observed that economic-statistical designs are more costly than pure economic designs, and the cost difference between

Table 1
Optimal economic designs ($K = 0.1$)

θ	δ	ρ	C	h	n	λ_m	λ_v	L_m	L_v	
0.01	0.5	1	24.51	20.00	7	0.29	0.11	2.45	2.67	
		1.5	32.26	10.65	11	0.80	0.77	2.80	1.83	
		2	39.10	6.14	6	0.99	0.86	3.13	1.69	
	1	1	28.54	15.63	10	0.83	0.15	2.53	3.12	
		1.5	34.98	8.10	7	0.76	0.99	2.67	1.88	
		2	41.92	5.19	5	0.81	0.84	3.09	1.69	
	1.5	1	33.64	9.40	6	0.88	0.05	2.73	2.80	
		1.5	39.43	5.53	5	0.85	0.81	2.75	2.04	
		2	46.45	4.14	4	0.84	0.86	3.02	1.63	
	2	1	40.21	5.27	4	0.85	0.11	2.96	2.27	
		1.5	45.70	4.41	4	0.82	0.41	3.00	3.98	
		2	52.68	3.43	3	0.83	0.89	2.94	1.53	
	0.05	0.5	1	25.36	20.00	2	0.68	0.11	3.90	1.38
			1.5	45.60	15.57	9	0.85	0.98	2.28	1.37
			2	65.77	4.87	5	0.94	0.74	2.90	1.52
		1	1	38.38	19.98	8	0.73	0.09	2.20	3.88
			1.5	54.33	7.64	6	0.80	0.94	2.34	1.65
			2	73.92	3.89	4	0.83	0.92	2.74	1.44
1.5		1	53.15	7.41	5	0.82	0.86	2.44	3.19	
		1.5	67.90	4.26	4	0.87	0.93	2.56	1.80	
		2	87.20	3.03	3	0.93	0.88	2.77	1.41	
2		1	72.10	4.51	3	0.86	0.96	2.65	2.24	
		1.5	86.35	3.37	3	0.78	0.66	2.63	1.78	
		2	105.52	2.63	3	0.77	0.99	2.86	1.45	

Table 2
Optimal statistically constrained economic designs ($ARL_0 \geq 250$, $ARL_1 \leq 20$, $K = 0.1$)

θ	δ	ρ	C	h	n	λ_m	λ_v	L_m	L_v	
0.01	0.5	1	24.89	12.20	6	0.35	0.35	3.25	2.15	
		1.5	32.63	6.80	9	0.63	0.59	3.35	2.10	
		2	39.47	6.49	8	0.50	0.72	3.35	2.10	
	1	1	29.12	5.45	4	0.34	0.20	2.90	2.00	
		1.5	35.37	4.20	5	0.50	0.20	3.00	1.90	
		2	42.46	5.45	7	0.49	0.68	3.38	2.14	
	1.5	1	34.10	4.45	4	0.50	0.15	3.00	1.90	
		1.5	39.63	3.95	4	0.60	0.15	3.00	1.90	
		2	46.94	3.20	4	0.55	0.35	3.00	1.90	
	2	1	40.46	3.95	3	0.62	0.20	2.90	2.00	
		1.5	45.75	3.70	3	0.67	0.20	2.90	2.00	
		2	52.98	3.20	3	0.76	0.20	2.90	2.00	
	0.05	0.5	1	26.84	12.20	6	0.35	0.35	3.25	2.15
			1.5	46.28	8.20	9	0.61	0.56	3.35	2.10
			2	66.45	3.70	4	0.55	0.69	3.40	1.70
		1	1	38.88	11.45	5	0.39	0.35	3.25	2.15
			1.5	54.91	3.70	5	0.63	0.55	2.90	2.30
			2	74.50	3.70	4	0.55	0.75	3.40	1.70
1.5		1	53.66	4.36	3	0.55	0.55	2.90	2.11	
		1.5	68.15	3.70	5	0.75	0.55	2.90	2.30	
		2	87.55	2.70	4	0.85	0.85	3.00	1.90	
2		1	72.43	3.70	3	0.55	0.65	2.90	2.60	
		1.5	86.51	2.70	3	0.85	0.85	3.00	1.90	
		2	105.83	2.51	3	0.85	0.85	3.00	1.90	

Table 3
Optimal statistically constrained economic designs ($ARL_0 \geq 100$, $ARL_1 \leq 10$, $K = 0.1$)

θ	δ	ρ	C	h	n	λ_m	λ_v	L_m	L_v	
0.01	0.5	1	24.59	18.20	7	0.35	0.35	2.75	2.15	
		1.5	32.41	9.00	11	0.72	0.64	2.85	2.10	
		2	39.41	6.60	8	0.63	0.75	3.01	2.10	
	1	1	28.68	13.00	10	0.65	0.50	2.85	2.31	
		1.5	35.02	6.60	7	0.70	0.75	2.85	2.10	
		2	42.28	6.60	7	0.69	0.75	2.87	2.10	
	1.5	1	33.66	8.20	6	0.78	0.50	2.85	2.44	
		1.5	39.49	6.20	5	0.78	0.52	2.75	2.15	
		2	46.89	3.45	4	0.65	0.39	2.92	2.10	
	2	1	40.22	6.20	4	0.89	0.35	2.84	2.15	
		1.5	45.75	3.70	3	0.67	0.20	2.90	2.00	
		2	52.97	3.20	3	0.72	0.35	2.92	2.10	
	0.05	0.5	1	26.18	19.40	5	0.50	0.50	2.85	2.10
			1.5	46.04	9.80	10	0.72	0.61	2.85	2.10
			2	65.95	3.92	5	0.75	0.66	2.90	1.70
1		1	38.58	13.70	7	0.57	0.35	2.75	2.15	
		1.5	54.56	6.20	7	0.69	0.61	2.75	2.15	
		2	74.06	3.83	5	0.70	0.75	2.90	1.70	
1.5		1	53.21	6.95	5	0.80	0.35	2.75	2.38	
		1.5	68.00	3.76	4	0.72	0.75	2.90	1.70	
		2	87.35	3.70	4	0.75	0.75	2.90	1.70	
2		1	72.21	3.76	3	0.83	0.55	2.90	2.00	
		1.5	86.49	3.70	3	0.80	0.70	2.90	1.70	
		2	105.77	2.33	3	0.70	0.55	2.90	1.70	

the economic and economic-statistical design decreases as the magnitudes of shifts increase. The similar patterns have also been observed previously in the economic design of EWMA charts for mean (Montgomery et al., 1995).

The relatively stable values of decision variables across different shift sizes in Tables 2 and 3 can be explained by two factors. First, the feasible region in the statistically constrained economic design is smaller than that in the pure economic design. Second, the values of decision variables at the start of the search influence the resulting values at the termination of the algorithm. Although this implies that the results found may not be globally optimal, in our experiment we have observed only minor differences in the optimal cost values obtained using different seeds of variables. We also remark that, we have also identified the optimal solution by using exhaustive grid search approach in our set of numerical examples for economic and economic-statistical design. We have found that the results obtained via Nelder–Mead method and those obtained via grid search are very close in all cases considered. Since we have also observed similar patterns in other test problems (not reported here), it can be stated that using an alternative search

method is not likely to yield significantly different results. We also note that, since it requires a relatively high computation time, exhaustive grid search is not a practical solution approach in the current problem.

Regarding the decision variables, especially for large shifts, sample size and sampling interval have been found to be relatively more robust than other variables to changes in initial values. One of the reasons behind the observed sensitivity of control limits and smoothing constants to starting values may be the additional flexibility provided by using two charts rather than a single chart. The change in one variable, say L_m , is compensated by a change in another variable, say L_v , and hence, different combinations of variables essentially lead to the same impact on the total cost. If only a single chart was used, due to a smaller set of decision variables, the number of alternative combinations of variables resulting in approximately same value of the total cost would probably be less, and therefore, the search would be more likely to converge to the same values of decision variables at termination regardless of the initial values.

We investigate the impact of the Taguchi loss coefficient K on the economic design parameters

Table 4
Optimal economic designs ($K = 0.4$)

θ	δ	ρ	C	h	n	λ_m	λ_v	L_m	L_v		
0.01	0.5	1	90.42	18.81	18	0.60	0.05	2.06	3.78		
		1.5	114.79	3.10	6	0.71	0.56	2.93	1.68		
		2	140.22	2.19	4	0.99	0.72	3.04	1.56		
		1	1	103.40	6.22	8	0.73	0.24	2.56	3.86	
			1.5	125.04	2.68	5	0.71	0.78	2.70	1.82	
			2	150.89	1.64	3	0.69	0.79	3.09	1.41	
	1.5	1	122.05	3.10	4	0.73	0.15	2.75	2.49		
		1.5	142.33	1.74	3	0.65	0.99	2.78	1.66		
		2	168.43	1.51	3	0.84	0.82	3.13	1.43		
	2	1	147.02	2.34	3	0.84	0.29	2.81	2.38		
		1.5	166.70	1.41	2	0.72	0.37	2.96	1.58		
		2	192.53	1.30	2	0.77	0.83	3.09	1.12		
	0.05	0.5	1	99.13	20.00	7	0.52	0.81	1.54	3.35	
			1.5	160.94	3.53	6	0.92	0.89	2.49	1.36	
			2	235.15	1.54	3	0.96	0.87	2.81	1.19	
			1	1	135.41	5.09	6	0.79	0.92	2.19	2.88
				1.5	192.32	2.13	4	0.73	0.99	2.41	1.57
				2	266.46	1.34	3	0.72	0.73	2.71	1.32
1.5		1	189.06	2.93	4	0.78	0.15	2.57	2.21		
		1.5	244.02	1.56	3	0.67	0.05	2.57	1.52		
		2	318.05	1.15	2	0.52	0.76	2.73	0.94		
2		1	261.90	1.70	2	0.79	0.89	2.50	2.22		
		1.5	315.75	1.27	2	0.61	0.56	2.57	1.44		
		2	389.92	0.95	2	0.61	0.42	2.75	1.14		

Table 5
Optimal economic designs ($K = 0.7$)

θ	δ	ρ	C	h	n	λ_m	λ_v	L_m	L_v		
0.01	0.5	1	155.39	9.79	12	0.47	0.32	2.10	2.71		
		1.5	195.06	1.90	5	0.65	0.62	2.83	1.62		
		2	239.03	1.42	3	0.99	0.89	2.88	1.33		
		1	1	176.90	3.83	6	0.62	0.05	2.52	2.61	
			1.5	212.94	1.57	4	0.63	0.68	2.81	1.67	
			2	257.38	1.27	3	0.89	0.80	2.91	1.38	
	1.5	1	208.89	2.42	4	0.78	0.05	2.71	3.02		
		1.5	243.05	1.46	3	0.75	0.57	2.79	1.82		
		2	287.57	1.01	2	0.86	0.82	2.90	1.10		
	2	1	252.27	1.98	3	0.84	0.16	2.75	2.29		
		1.5	285.22	1.04	2	0.67	0.55	2.96	2.13		
		2	329.97	1.03	2	0.92	0.72	2.87	1.08		
	0.05	0.5	1	169.96	19.74	12	0.60	0.05	1.36	2.34	
			1.5	273.02	2.24	5	0.80	0.80	2.39	1.42	
			2	401.12	0.96	2	0.91	0.75	2.66	0.94	
			1	1	230.26	3.20	5	0.55	0.55	2.30	3.10
				1.5	327.20	1.16	3	0.49	0.54	2.61	1.42
				2	455.16	0.82	2	0.95	0.64	2.75	0.93
1.5		1	322.85	2.11	4	0.78	0.05	2.53	2.15		
		1.5	416.80	0.88	2	0.58	0.75	2.56	1.85		
		2	544.71	0.89	2	0.74	0.71	2.82	0.97		
2		1	449.21	1.03	2	0.81	0.55	2.51	2.57		
		1.5	542.09	0.95	2	0.81	0.81	2.52	2.01		
		2	670.77	0.81	2	0.66	0.70	2.58	1.25		

by using $K = 0.4$, and $K = 0.7$ in Tables 4 and 5, respectively. Based on the results displayed in Tables 1, 4, and 5, the optimal values of the sampling interval (h) exhibit a decreasing trend as K increases. Thus, more frequent sampling is required when the cost incurred due to defective products is high.

According to the results in Tables 1–5, in general, as the sizes of the shifts increase, both the sampling interval (h) and sample size (n) appear to decrease. The inverse relationship between the optimal sample size and the magnitude of the shift has also been observed in the economic design of X-bar charts (Montgomery, 1980). The results given in Park et al. (2004) for EWMA mean charts also reveal a similar pattern. The reduction in sample size can be explained by the fact that large shifts in mean and/or variance can be detected more easily than small shifts, implying a reduced need for a large sample size. On the other hand, higher sensitivity of the charts to large shifts does not lead to a lower sampling frequency; hence, with regard to the sampling interval, the effect of cost of defectives appears to outweigh the effect of the improved statistical power of the control scheme. The decrease in sampling interval reduces the average time for receiving an out-of-control signal when a shift occurs. Hence, a shorter sampling interval helps to constrain the quality related costs which increase with shift sizes. We also remark that, in previous research optimal sampling interval is observed to decrease with shift size for an EWMA mean chart (Park et al., 2004), which is in consistency with our results for the EWMA-m/EWMA-v scheme.

Comparing results for $\theta = 0.01$ versus $\theta = 0.05$ in Tables 1–5, we observe that as the failure arrival rate (θ) increases, the total cost per hour (C) increases. This behavior is due to the fact that as θ decreases the fraction of time that the process will be in state of control increases, which results in a lower cost per hour. According to results, it also appears that in general an increase in θ causes the sampling interval (h) to decrease. Again, this impact is similar to earlier research findings regarding X-bar chart.

6. Conclusion

We have studied the joint economic design of EWMA control charts for monitoring the mean and variance of a process. The average run length of the joint control scheme is computed by using

the Markov chain method. The quadratic loss function approach has been used to represent the cost of producing defective products in the objective function of the model. In this approach, costs due to off-target performance are computed based on Taguchi's loss function which quadratically penalizes the deviations from the target of the quality characteristic. In our numerical examples we have observed that, in general, both the optimal sample size and sampling interval decrease as the size of shifts in mean and/or variance increases. Similarly, increases in the Taguchi quality loss coefficient K result in shorter time intervals between samples. We have also explored statistically constrained economic designs which are sometimes preferred over pure economic designs by the users who desire the control scheme to achieve certain statistical performance targets. It is found that including constraints on average run length of the scheme leads to a decrease in the optimal sampling interval when shifts in mean and/or variance are small.

In the literature, the joint economic design of control charts for process mean and dispersion has been investigated for Shewhart control charts employing a test statistic based on the current sample. In addition to the X-bar/Range and X-bar/Standard deviation schemes analyzed in previous research and the EWMA-m/EWMA-v scheme investigated in this paper, future research may consider joint economic design of other commonly used control charts for mean and dispersion such as cumulative sum (CUSUM) charts, as well as their multi attributed variants.

Acknowledgement

The authors thank the referees for their helpful comments.

Appendix A. Computing the ARL using the Markov chain method

In the Markov chain method we divide the interval between the UCL and LCL in each chart into k equally spaced subintervals (k must be an odd integer). For the EWMA-m chart, the subintervals are $R_1 = [u_0, u_1]$, $R_2 = [u_1, u_2]$, \dots , $R_i = [u_{i-1}, u_i]$, \dots , $R_k = [u_{k-1}, u_k]$ where $u_i = LCL_m + i\Delta u$ and $\Delta u = (UCL_m - LCL_m)/k$. The subintervals correspond to the transitional states in the Markov chain and the transition probabilities $p_{i,j}$ are found by setting

the EWMA statistic Z_t to the midpoint of the sub-interval R_i when $u_{i-1} < Z_t \leq u_i$. Hence,

$$p_{i,j} = P(Z_t \in R_j \mid Z_{t-1} \in R_i) \\ = P(u_{j-1} < Z_t \leq u_j \mid Z_{t-1} = (u_{i-1} + u_i)/2).$$

Note that $u_{(k+1)/2} = \mu_0$. Let $\Phi(\cdot)$ denote the cumulative distribution function (cdf) for the standard normal probability distribution. The transition probabilities can be computed iteratively by using

$$p_{i,j} = f_{i,j} - f_{i,j-1}, \quad i, j = 1, \dots, k,$$

where

$$f_{i,j} = \Phi\{[2L_m(j - (1 - \lambda_m)(i - 0.5) - 0.5\lambda_mk)]/(\rho k[\lambda_m(2 - \lambda_m)]^{0.5}) - \delta n^{0.5}/\rho\}, \\ i = 1, \dots, k, \quad j = 0, \dots, k.$$

Recall that $\rho = \sigma_1/\sigma_0$; for the in-control case, $\rho = 1$ (Morais and Pacheco, 2000). The transient states in the Markov chain are the in-control states and the EWMA statistic Z_t moves to the absorbing state if Z_t falls outside the control limits. The run length distribution of the EWMA-m chart can be found by using the initial probability vector and transition probability matrix. The initial probability vector contains the probabilities of Z starting in each state of the Markov chain. In this paper we use the zero state ARL, i.e. the starting state for the EWMA statistic is the in-control mean with probability one (Lucas and Saccucci, 1990). Note that we express the shift $\delta = (\mu_1 - \mu_0)/\sigma_0$ differently from Morais and Pacheco (2000) who define the shift in terms of the standard deviation of the sample mean.

The run length distribution of the EWMA-v chart can be determined following a similar procedure. Let $\Delta v = (UCL_v - LCL_v)/k$, and $\chi^2_{n-1}(\cdot)$ be the cdf for the chi-squared probability distribution with $n - 1$ degrees of freedom. The transition probabilities $q_{i,j}$ in the Markov chain, associated with the EWMA-v chart, can be computed recursively from

$$q_{i,j} = h_{i,j} - h_{i,j-1}, \quad i, j = 1, \dots, k,$$

where

$$h_{i,j} = \chi^2_{n-1}\{(n - 1) \exp[(j - 1) - (1 - \lambda_v)(i - 1.5)]\Delta v/\lambda_v/\rho^2\}, \\ i = 2, \dots, k, \quad j = 1, \dots, k, \\ h_{1,j} = \chi^2_{n-1}\{(n - 1) \exp[(j - 1)\Delta v/\lambda_v/\rho^2]\}, \\ j = 1, \dots, k$$

and $h_{i,0} = 0, i = 1, \dots, k$.

In order to find the ARL of the combined EWMA scheme, we first determine the run length

distributions of the EWMA-m and EWMA-v charts, and then use these results to obtain the complementary cumulative distribution function of the run length of the combined scheme (Morais and Pacheco, 2000). The joint control scheme generates an out-of-control signal when one of the two charts yields an out-of-control signal. Then the probability that the run length of the joint scheme is greater than n , $P(RL_J > n)$, is

$$P(RL_J > n) = P(RL_m > n)P(RL_v > n), \quad n = 0, 1, 2, \dots,$$

where $P(RL_m > n)$ and $P(RL_v > n)$ are the probabilities that the run length of the EWMA-m and EWMA-v chart is greater than n , respectively. We can use

$$P(RL_m > n) = \mathbf{e}_i^T [\mathbf{P}]^n \mathbf{1}, \quad n = 1, 2, \dots, \tag{A1}$$

where $\mathbf{P} = [p_{i,j}]$ is the $k \times k$ matrix of transition probabilities $p_{i,j}$, \mathbf{e}_i is the i th unit vector with all elements zero except the i th element which is 1, and $\mathbf{1}$ is a column vector of ones. In (A1), $i = (k + 1)/2$. $P(RL_v > n)$ is computed similarly from

$$P(RL_v > n) = \mathbf{e}_i^T [\mathbf{Q}]^n \mathbf{1}, \quad n = 1, 2, \dots,$$

where $\mathbf{Q} = [q_{i,j}]$ is the $k \times k$ matrix of transition probabilities $q_{i,j}$, and $i = 1$ in \mathbf{e}_i . Note that

$$P(RL_m > 0) = P(RL_v > 0) = 1.$$

Finally, the ARL of the joint control scheme, ARL_J , is found from

$$ARL_J = \sum_{n=0}^{\infty} P(RL_J > n).$$

Our numerical experiments indicate that in our solution space the error associated with computing ARL_J by excluding the values of $P(RL_J > n)$ for $n > 2000$ has an insignificant impact on the optimal cost (with $k = 51$), hence, we approximate ARL_J by restricting the values of n to less than 2000.

References

Acosta-Mejia, C.A., Pignatiello, J.J., Rao, B.V., 1999. A comparison of control charting procedures for monitoring process dispersion. *IIE Transactions* 31, 569–579.
 Crowder, S.V., Hamilton, M.D., 1992. An EWMA for monitoring a process standard deviation. *Journal of Quality Technology* 24, 12–21.
 Elsayed, E.A., Chen, A., 1994. An economic design of X-bar control chart using quadratic loss function. *International Journal of Production Research* 32, 873–887.
 Gan, F.F., 1995. Joint monitoring of process mean and variance using exponentially weighted moving average control charts. *Technometrics* 37, 446–453.

- Gan, F.F., Ting, K.W., Chang, T.C., 2004. Interval charting schemes for joint monitoring of process mean and variance. *Quality and Reliability Engineering International* 20, 291–303.
- Ho, C., Case, K.E., 1994a. Economic design of control charts: A literature review for 1981–1991. *Journal of Quality Technology* 26, 39–53.
- Ho, C., Case, K.E., 1994b. The economically-based EWMA control chart. *International Journal of Production Research* 32, 2179–2186.
- Knoth, S., Schmid, W., 2002. Monitoring the mean and variance of a stationary process. *Statistica Neerlandica* 56, 77–100.
- Linderman, K., McKone-Sweet, K.E., Anderson, J.C., 2005. An integrated systems approach to process control and maintenance. *European Journal of Operational Research* 164, 324–340.
- Lorenzen, T.J., Vance, L.C., 1986. The economic design of control charts: a unified approach. *Technometrics* 28, 3–10.
- Lucas, J.M., Saccucci, M.S., 1990. Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics* 32, 1–12.
- McWilliams, T.P., 1989. Economic control chart designs and the in-control time distribution: A sensitivity study. *Journal of Quality Technology* 21, 103–110.
- McWilliams, T.P., Saniga, E.M., Davis, D.J., 2001. Economic-statistical design of X-bar and R or X-bar and S charts. *Journal of Quality Technology* 33, 234–241.
- Montgomery, D.C., 1980. The economic design of control charts: a review and literature survey. *Journal of Quality Technology* 12, 75–87.
- Montgomery, D.C., Torng, J.C.C., Cochran, J.K., Lawrence, F.P., 1995. Statistically constrained economic design of the EWMA control chart. *Journal of Quality Technology* 27, 250–256.
- Morais, M.C., Pacheco, A., 2000. On the performance of combined EWMA schemes for μ and σ : A Markovian approach. *Communications in Statistics, Part B-Simulation and Computation* 29, 153–174.
- Moskowitz, H., Plante, R., Chun, Y.H., 1994. Effect of quality loss functions on the economic design of X-bar process control charts. *European Journal of Operational Research* 72, 333–349.
- Park, C., Lee, J., Kim, Y., 2004. Economic design of a variable sampling rate EWMA chart. *IIE Transactions* 36, 387–399.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, second ed. Cambridge University Press, UK.
- Rahim, M.A., 1989. Determination of optimal design parameters of joint X-bar and R charts. *Journal of Quality Technology* 21, 65–70.
- Rahim, M.A., Costa, A.F.B., 2000. Joint economic design of X-bar and R charts under Weibull shock models. *International Journal of Production Research* 38, 2871–2889.
- Reynolds, M.R., Stoumbos, Z.G., 2001. Monitoring the process mean and variance using individual observations and variable sampling intervals. *Journal of Quality Technology* 33, 181–205.
- Reynolds, M.R., Stoumbos, Z.G., 2004. Control charts and the efficient allocation of sampling resources. *Technometrics* 46, 200–214.
- Saccucci, M.S., Lucas, J.M., 1990. Average run lengths for exponentially weighted moving average control schemes using the Markov chain approach. *Journal of Quality Technology* 22, 154–162.
- Saniga, E., 1979. Joint economic design of X-bar/R control charts with alternate process models. *AIIE Transactions* 11, 254–260.
- Tolley, G.O., English, J.R., 2001. Economic designs of constrained EWMA and combined EWMA-X-bar control schemes. *IIE Transactions* 33, 429–436.
- Torng, J.C.C., Cochran, J.K., Montgomery, D.C., Lawrence, F.P., 1995. Implementing statistically constrained economic EWMA charts. *Journal of Quality Technology* 27, 257–264.