The Parameter Selection and Average Run Length Computation for EWMA Control Charts

Sheng Shu Cheng[1], Fong-Jung Yu[2], Shih-Ting Yang[3] and Jiang-Liang Hou[4]
[1]Department of Business Administration, Yu-Da University of Science and Technology, Miaoli, Taiwan
[2]Department of Industrial Engineering and Management, Da-Yeh University, Changhua, Taiwan
[3]Department of Information Management, Nanhua University, Chiayi, Taiwan
[4]Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu, Taiwan

Keywords: Statistical Process Control, Exponentially Weighted Moving Average, Smoothing Parameter Selection, Determination of the Control Limits.

Abstract: In the Statistical Process Control (SPC) field, an Exponentially Weighted Moving Average for Stationary processes (EWMAST) chart with proper control limits has been proposed to monitor the process mean of a stationary autocorrelated process. There are two issues of note when using the EWMAST charts. These are the smoothing parameter selections for the process mean shifts, and the determination of the control limits to meet the required average run length (ARL). In this paper, a guideline for selecting the smoothing parameter is discussed. These results can be used to select the optimal smoothing parameter in the EWMAST chart. Also, a numerical procedure using an integration approach is presented for the ARL computation with the specified control limits. The proposed approach is easy to implement and provides a good approximation to the average run length of EWMAST charts.

1 INTRODUCTION

The majority of traditional control charts are based on an assumption that processed data is statistically independent; however, this assumption does not hold in certain production environments. It is well known that using traditional charts for monitoring autocorrelated processes usually results in an unnecessarily high occurrence of false positives. A common approach to handling autocorrelated data is to apply traditional control charts on the stream of residuals after the process data have been fitted to a time-series model. Several residual control charts have been proposed in recent years. Such as: Alwan and Roberts (1998) proposed a special cause chart (SCC) which uses a time-series model to obtain and monitor the residuals. Montgomery and Mastrangelo (1991) provided a procedure of plotting one-step-ahead EWMA prediction errors on a control chart (M-M chart). English *et al* (1991) and Wincek (1990) suggested Kalman filtering to obtain the residuals. The Proportional-Integral- Derivative (PID) chart by Jiang *et al* (2002) and the Dynamic T^2 chart by Tsung and Apley (2002) have been proposed for monitoring of processes with a feedback controller.

There are some problems in the above control charts. Zhang (1998) proposed an EWMAST control chart to monitor the original autocorrelated data. The EWMAST chart is very similar to the traditional EWMA chart, except that it is designed to be applied to the monitoring of stationary, autocorrelated data. Both the EWMA and EWMAST charts are used for charting the same statistic, but the EWMAST chart is used in conjunction with modified control limits to account for the additional variation within an autocorrelated process. The major advantages in using an EWMAST chart are: there is no requirement for an operator to use time-series techniques; and this method has a comparable ability to the residual-based charts for detecting large mean shifts. Specifically, when we set the smoothing parameter, λ =1, when using an EWMAST chart, then it is equivalent to the traditional Shewhart chart. Therefore, the EWMAST chart can be more flexible than the Shewhart chart as its smoothing parameter can be set according to the magnitude of the mean shift in the stationary autocorrelated process. According to the simulation results in Zhang (1998), the EWMAST chart is more sensitive than residual-

Cheng S., Yu F., Yang S. and Hou J..
The Parameter Selection and Average Run Length Computation for EWMA Control Charts.
DOI: 10.5220/0005146902940299
In *Proceedings of the International Conference on Neural Computation Theory and Applications* (NCTA-2014), pages 294-299
ISBN: 978-989-758-054-3

based charts for positive autocorrelated data. For negative autocorrelated data, the performance of the EWMAST chart is still superior, but is not as significant as the positively correlated cases.

Regarding the selection of an optimal smoothing parameter for the EWMAST chart, Zhang's (2000) suggestion was to set the smoothing parameter λ equal 0.2 for most applications, and provided the ARLs for many autocorrelated processes. In general, the simulation technique is rather costly, especially in the case of on-target analysis. Therefore, it is not a suitable approach when investigating the ARL performance in practical applications.

In order to set the parameters of EWMAST charts more easily, a computing algorithm is presented and tabulated parameters, which yield the shortest out-of-control ARL for EWMAST charts, are provided in this article.

2 DESCRIPTION OF THE EWMAST CHART

To illustrate the autocorrelated process, we consider the important case of monitoring the process mean μ_0 of a stationary AR(1) process and defined as:

$$X_t - \mu_0 = \phi(X_{t-1} - \mu_0) + e_t, \; t = 1,2,... \quad (1)$$

where $\{X_t\}$ represents the process output, $|\phi| < 1$ is a constant representing the stationary process, and $\{e_t\}$ is a normally distributed white noise with finite variance, σ_e^2. The EWMAST chart is constructed by charting the EWMA control statistic under the stationary process. The chart statistic Z_t is defined as:

$$Z_t = (1-\lambda)Z_{t-1} + \lambda X_t, \quad t = 1,2,..., \quad (2)$$

where $Z_0 = \mu_0$ is the on-target process mean, and λ is a smoothing parameter within the range $(0 < \lambda \le 1)$. Assume the process $\{X_t\}$ with stationary variance σ_X^2 undergoes a single assignable cause that shifts the process mean to $\mu_t = \mu_0 + \delta\sigma_X$ at time *t*. To monitor $\{\mu_t\}$, a plot of Z_t is made by selecting suitable values of the smoothing parameter λ and the width of the control limits $\mu_0 \pm L\sigma_{Z_t}$, where $L > 0$. The parameter *L* serves as a width adjustment for control limits to meet the required in-control ARL. Zhang (1998) indicates that the control statistic Z_t has variance:

$$\sigma_{z_t}^2 = \frac{\lambda}{2-\lambda}\sigma_x^2 \left\{ \begin{array}{l} 1-(1-\lambda)^{2t} + \\ 2\sum_{k=1}^{t-1} \rho(k)(1-\lambda)^k [1-(1-\lambda)^{2(t-k)}] \end{array} \right\} \quad (3)$$

where $\rho(k)$ is the autocorrelation function of $\{X_t\}$ at time lag k, and $\rho(k)$ can simplify to ϕ^k for the AR(1) processes. The limiting value of the variance of Z_t in equation (3) as t increases to infinity is given in (4):

$$\sigma_z^2 = \frac{\lambda}{2-\lambda}\sigma_x^2 \left\{ 1 + 2\sum_{k=1}^{\infty} \rho(k)(1-\lambda)^k \right\} \quad (4)$$

Hence, the control limits are taken to be:

$$\begin{cases} LCL = \mu_0 - L\sigma_z \\ UCL = \mu_0 + L\sigma_z. \end{cases} \quad (5)$$

3 DESIGN OF THE EWMAST CHART

Constructing an EWMAST chart requires the specification of λ and the constant *L*. In general, the choice of parameters (λ , *L*) is based on a compromise for certain statistic or economic constraints. It is well known that small values of λ are better for detecting small shifts in the mean and large values of λ are better for detecting large shifts. Although the suggestion for designing the EWMAST smoothing parameter ($\lambda = 0.2$) by Zhang (1998) is given in practical terms, there remain multiple issues that have not been clearly specified. The $\lambda = 0.2$ can only be viewed as a heuristic suggestion when the magnitude of the shift is unknown. Another concern is that Zhang's suggestion does not consider the in-control ARL.

To select optimal parameters λ and L for an EWMAST chart, an extensive simulation, with 10,000 runs per parameter setting, was implemented. Thus, the standard deviation of the ARL estimation error is less than 1% of the actual ARL. For each simulation run, the $\{e_t\}$ are generated as an independent sequence of random numbers, based upon the standard normal distribution utilized in the IMSL® software (1989). The value of $\{X_t\}$ is generated from equation (1). The charting statistic is calculated via equation (2), with Z_0 initialized at 0 and terminated when Z_t exceeds the control limits.

Table 1: Optimal EWMAST control schemes.

		δ =0.5					δ =1.5		
ϕ	ARL_0	λ	L	ARL_{min}	ϕ	ARL_0	λ	L	ARL_{min}
0.25	100	0.05	1.815	23.36	0.25	100	0.3	2.388	5.28
	370	0.05	2.432	37.99		370	0.2	2.800	7.21
	500	0.05	2.555	41.90		500	0.2	2.903	7.68
	1000	0.05	2.828	52.42		1000	0.2	3.133	8.83
0.50	100	0.05	1.730	32.40	0.50	100	0.2	2.185	7.42
	370	0.05	2.356	58.17		370	0.2	2.714	10.93
	500	0.05	2.483	65.72		500	0.1	2.670	11.73
	1000	0.05	2.759	87.24		1000	0.1	2.925	13.59
0.75	100	0.05	1.585	49.05	0.75	100	0.2	2.013	11.92
	370	0.05	2.230	104.07		370	0.1	2.410	19.14
	500	0.05	2.358	121.86		500	0.1	2.532	21.01
	1000	0.05	2.641	177.68		1000	0.05	2.641	25.44
0.90	100	0.05	1.372	68.04	0.90	100	1.0	2.185	17.24
	370	0.05	2.041	176.47		370	0.05	2.041	35.88
	500	0.05	2.175	217.04		500	0.05	2.175	40.07
	1000	0.05	2.470	349.61		1000	0.05	2.470	51.46
		δ =2.5					δ =3.0		
ϕ	ARL_0	λ	L	ARL_{min}	ϕ	ARL_0	λ	L	ARL_{min}
0.25	100	0.8	2.552	2.25	0.25	100	1.00	2.570	1.60
	370	0.5	2.954	3.04		370	0.70	2.976	2.12
	500	0.5	3.041	3.22		500	0.70	3.071	2.25
	1000	0.4	3.230	3.63		1000	0.70	3.284	2.59
0.50	100	1.0	2.538	2.67	0.50	100	1.00	2.538	1.72
	370	0.5	2.887	4.13		370	1.00	2.980	2.52
	500	0.5	2.985	4.45		500	1.00	3.071	2.76
	1000	0.4	3.173	5.18		1000	0.80	3.259	3.41
0.75	100	1.0	2.430	3.30	0.75	100	1.00	2.430	1.95
	370	1.0	2.903	5.97		370	1.00	2.903	3.19
	500	1.0	3.000	6.79		500	1.00	3.000	3.59
	1000	0.3	3.008	8.92		1000	1.00	3.218	4.68
0.90	100	1.0	2.185	4.02	0.90	100	1.00	2.185	2.16
	370	1.0	2.711	8.54		370	1.00	2.711	4.18
	500	1.0	2.821	10.03		500	1.00	2.821	4.89
	1000	1.0	3.064	14.46		1000	1.00	3.064	6.88

The same program was used for an out-of-control ARL with a mean shift $\delta\sigma_X$ added to $\{X_t\}$ at time t =1. The design of an EWMAST chart consists of the selection of charting parameters (λ, L) that satisfy certain ϕ , δ , and in-control ARL. For most practical purposes, Table I show the particular values of λ and L on the EWMAST chart for certain ϕ (ϕ = 0.25,0.5,0.75 and 0.9) to provide the required in-control ARLs (ARL_0 =100, 370, 500 and 1000).

Table I shows that, as we expected, when ϕ =0.25 and ARL_0 =100 the optimal λ value is 0.3 for detecting a 1.5- σ_X shift; and λ =0.8 is optimal for detecting a 2.5-σ_X shift, but in the case of ϕ=0.25 and ARL_0=500 the optimal λ value is 0.2 for detecting a 1.5-σ_X shift and λ=0.5 for a 2.5-σ_X shift. This phenomenon is the same as which occurs in EWMA charts that have been studied by Lucas and Saccucci (1990). Another test, with specified ARL_0 =370 and a three-σ_X shift determined that the optimal value of λ is 0.7 for ϕ =0.25, but the optimal value of λ is 1.0 for ϕ=0.75. Therefore, the optimal λ of an EWMAST chart may be affected by the different values of ϕ, the magnitude of the mean shift ($\delta\sigma_X$) and the in-control average run length (ARL_0).

The optimal λ values should be dependent on the magnitude of the mean shift, the autoregressive parameter, and the in-control average run length. This is helpful for the operator to set an appropriate value of λ when an EWMAST chart is chosen to monitor a stationary process. The following steps are recommended as guidelines for designing an optimal EWMAST chart.

Step 1. First, specify the desired in-control ARL_0, the autocorrelation coefficient ϕ, and decide upon the smallest process mean shift, in terms of $\delta\sigma_X$, that must be rapidly detected..

Step 2. Next, select the optimal parameters (λ, L) from Table I.

Step 3. Finally, evaluate the entire ARL performance for this EWMAST chart to determine whether the chart provides suitable protection against other shifts.

4 THE NUMERICAL PROCEDURE FOR FINDING THE ARL OF THE EWMAST CHART

In this section, a numerical procedure is presented for the investigation of the ARL of EWMAST charts. Knowledge of the run length is important and permits us to illustrate the performance of a chart in terms of average run length. In general, since the run length of a chart is a nonnegative, random, integer variable, chart performance can be evaluated by averaging:

$$\mathrm{ARL} = E(\mathrm{RL}) = \sum_{k=0}^{\infty} k \times p(\mathrm{RL} = k) \tag{6}$$

In the iid case, a traditional Shewhart chart can be easily calculated. The $p(\mathrm{RL} = k)$ are evaluated directly and summed to obtain the ARL. However, in charting procedures that use recursive charting statistics, such as EWMA charts and CUSUM charts, equation (6) is not easily evaluated. In this situation, equation (6) can be rewritten as:

$$\mathrm{ARL} = \sum_{k=0}^{\infty} p(\mathrm{RL} > k) = \sum_{k=0}^{\infty} P_k \,. \tag{7}$$

Let the random variable N be the time of the first passage of the process $\{Z_t\}$ exceeding the control limits given by equation (5). The probability of an in-control process at time k, given the initial condition Z_0 ($LCL \le Z_0 \le UCL$), is written as:

$$P_k = p(\mathrm{RL} > k) = P_{k-1} \times r_k , \tag{8}$$

where $P_0 = 1.0$, and r_k is the shrink ratio of P_k relative to P_{k-1}. By using equation (8) to establish a recursive formulation for the probability of a ruined problem, the average run lengths can be found by directly summing the P_k terms. These computations demonstrate the behavior of run length distributions over various AR(1) parameters.

To investigate the behavior of P_k and r_k, our analysis shows that there is a linear relationship between P_k and r_k with increasing k.

The following steps constitute the ARL computation method for an EWMAST chart applied to the AR(1) process.

Step 1. Given λ and L values, and letting k be the time index, set k =1 as the beginning of the process;

Step 2. Calculate the $\sigma^2_{z_k}$ using equation (3) for each k, if $k \ge 2$, and also calculate $\mathrm{Cov}(Z_i, Z_j)$ for $1 \le i, j \le k$ using the equation (A.3) in Zhang's (1998) studies ;

Step 3. Use Alan's (1998) algorithm to find the probability of $p(\mathrm{RL} > k)$;

Step 4. Find the shrink ratio of $r_k = \dfrac{p(\mathrm{RL} > k)}{p(\mathrm{RL} > k-1)}$;

Step 5. Collect $r_1, r_2, \ldots, r_k$, if $\{r_k\}$ is converging to a constant p or $p(\mathrm{RL} > k) \le 10^{-5}$; then, set $k = N$ and go to Step 6; otherwise, set $k = k+1$ and go to Step 2;

Step 6. Compute

$$\mathrm{ARL} \approx \sum_{i=0}^{N-1} P(\mathrm{RL} > i) + \left(\frac{P_N}{1-p}\right).$$

Table 2: ARLs for the EWMAST chart applied to an AR(1) process with $\varphi > 0$.

(ϕ, *L*)	shift	(1)[a]	(2)[b]	*r.d.* =[(1)-(2)]/(2)*100%
ϕ=0.25, *L*=2.80	0.0	373.41	379.49	-1.60%
	0.5	56.31	56.57	-0.46%
	1.0	14.72	14.74	-0.14%
	2.0	4.676	4.68	-0.09%
	3.0	2.797	2.73	2.45%
ϕ=0.50, *L*=2.72	0.0	378.15	372.43	1.54%
	0.5	88.97	86.28	3.12%
	1.0	24.29	23.22	4.61%
	2.0	6.580	6.33	3.95%
	3.0	3.571	3.50	2.03%
ϕ=0.75, *L*=2.59	0.0	390.41	378.90	3.04%
	0.5	149.71	146.07	2.49%
	1.0	47.40	45.45	4.29%
	2.0	11.279	10.80	4.44%
	3.0	5.129	5.03	1.97%
ϕ=0.90, *L*=2.43	0.0	423.87	377.46	12.30%
	0.5	228.02	201.73	13.03%
	1.0	89.14	82.62	7.89%
	2.0	20.791	19.26	7.95%
	3.0	7.615	7.09	7.40%

[a] Computational results.
[b] Zhang's (2000) simulation results.

Table 3: ARLs for the EWMAST chart applied to an AR(1) process with $\phi < 0$.

(ϕ, *L*)	shift	(1)[a]	(2)[b]	*r.d.* =[(1)-(2)]/(2)*100%
ϕ=-0.25, *L*=2.92	0.0	384.96	374.52	2.79%
	0.5	23.31	22.35	4.30%
	1.0	6.892	6.72	2.56%
	2.0	2.862	2.85	0.42%
	3.0	1.935	1.94	-0.26%
ϕ=-0.50, *L*=2.94	0.0	366.15	374.90	-2.33%
	0.5	14.056	13.97	0.62%
	1.0	4.840	4.64	4.31%
	2.0	2.239	2.27	-1.37%
	3.0	1.590	1.62	-1.85%
ϕ=-0.75, *L*=2.94	0.0	347.80	375.88	-7.47%
	0.5	8.010	8.25	-2.91%
	1.0	3.319	3.38	-1.80%
	2.0	1.704	1.78	-4.27%
	3.0	1.291	1.24	4.11%
ϕ=-0.90, *L*=2.90[c]	0.0	376.05	377.17	-0.30%
	0.5	5.336	5.37	-0.63%
	1.0	2.480	2.53	-1.98%
	2.0	1.480	1.45	2.07%
	3.0	1.147	1.07	7.20%

[a] Computational results.
[b] Zhang's (2000) simulation results.
[c] Result from Zhang's (2000) use of *L*=2.65.

5 COMPUTATIONAL RESULTS

Zhang (2000) estimated these ARLs at $\delta = 0$, 0.5, 1, 2 and 3.0 in units of σ_X on simulations utilizing at least 4,000 realizations from the AR(1) processes with $\phi = \pm 0.25$, ± 0.5, ± 0.75 and ± 0.9. In contrast to the proposed methodology, we do the same parameter combinations in Zhang's studies. The results are listed in Table II for $\phi > 0$ and in Table III for $\phi < 0$.

As indicated in Table II, it is clear that when the process is positively autocorrelated, the ARLs of our computational results are in agreement with those obtained by Zhang's simulation results. Let the relative difference (*r.d.*) represent the difference between Zhang's results and those obtained by the proposed method. Table II also shows that when mean shifts are small and ϕ is large, the simulation results deviate more from the computational results. This phenomenon indicates that due to the inflation of $\sigma_{z_t}^2$, the larger the ϕ, the more simulation runs are required.

As indicated in Table III, it is clear that when the process is negatively autocorrelated, the ARLs are also in agreement with Zhang's results. Table III also shows an interesting phenomenon: when ϕ becomes increasingly negative and large, the EWMAST chart becomes more sensitive. This property is completely opposite to a positive autocorrelated process. As for the *r.d.* index, we can also observe that the results of a simulation with few realizations results in an unstable estimate of ARL, especially in the case of an in-control situation with highly correlated data.

6 CONCLUSIONS

In this research, the performance of an EWMAST chart has been investigated for various parameter settings when the AR(1) process is utilized. These results demonstrate guidelines for parameter (λ, L) selection when the in-control ARL and the autogressive parameter are specified. A numerically analytical expression was also used to evaluate the ARLs of the EWMAST chart, in the important special case of an autocorrelated process. Importantly, this method enables the assessment of the run-length distribution of an EWMAST chart using underlying data from an AR(1) process. For an application of the results, the ARL algorithm can be extended to calculate run-length distribution and ARLs for other stationary-process data with determined parameters. Although these results are relatively narrow in scope when compared to the results in Lucas and Saccucci (1990), they are still helpful to the operator for setting parameters when using an EWMAST chart. As for the other requirements of in-control ARLs and the different ϕ values listed in Table I, a larger table covering a wider range of ARL values is available from the authors on request.

ACKNOWLEDGEMENTS

The authors wish to express their appreciation for the financial support by National Sciences Council of the Republic of China, Grant No NSC 102-2221-E-212-018, MOST 103-2221-E-343 -002, NSC 101-2221-E-007-045-MY3 and Da-Yeh University, Taiwan,

REFERENCES

Alwan, L.C., Roberts, H.V., 1998. Time series modeling for statistical process control. *Journal of Business and Economic statistics,* 6, 87-95.

English, J.R., Krishnamurthi, M., Sastri, T., 1991, Quality Monitoring of Continuous Flow Processes. *Computers and Industrial Engineering,* 20, 251-260.

IMSL, 1989, *STAT/LIBRARY manual.* IMSL, Inc: Houston.

Jiang, W., Wu, H., Tsung, F., Nair V.N., Tsui K.L., 2002, PID Charts for Process Monitoring. *Technometrics*, 44, 205-214.

Lucas, J.M., Saccucci, M.S., 1990, Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics*, 32, 1-12.

Montgomery, D.C, Mastrangelo, C.M., 1991, Some statistical process control methods for autocorrelated data. *Journal of Quality Technology,* 23, 179-193.

Tsung, F, Apley, D.W., 2002, The dynamic T-squared chart for monitoring feedback-controlled processes. *IIE Transactions*, 34, 1043-1054.

Wincek, M.A., 1990, Industrial strength time series: Effect of new technologies on application and practice, unpublished paper presented at the 5th annual conference on making statistics more effective in schools of business, University of Kansas, June.

Zhang, N.F., 1998, A statistical control chart for stationary process data. *Technometrics*, 40, 24-28.

Zhang, N.F., 2000, Statistical control charts for monitoring the mean of a stationary process. *Journal of Statistic Computation and Simulation*, 66, 249-258.