

The Performance of the MEWMA Control Chart When Parameters Are Estimated

Mahmoud A Mahmoud, Petros Maravelakis

▶ To cite this version:

Mahmoud A Mahmoud, Petros Maravelakis. The Performance of the MEWMA Control Chart When Parameters Are Estimated. Communications in Statistics - Simulation and Computation, 2010, 39 (09), pp.1803-1817. 10.1080/03610918.2010.518269. hal-00637026

HAL Id: hal-00637026

https://hal.science/hal-00637026

Submitted on 29 Oct 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



The Performance of the MEWMA Control Chart When Parameters Are Estimated

Journal:	Communications in Statistics - Simulation and Computation
Manuscript ID:	LSSP-2010-0138.R1
Manuscript Type:	Original Paper
Date Submitted by the Author:	06-Aug-2010
Complete List of Authors:	Mahmoud, Mahmoud; Faculty of Economics and Political Science, Statistics Maravelakis, Petros; University of the Aegean
Keywords:	EWMA, Estimation Effect, Average Run Length, Statistical Process Control
Abstract:	In the present article, we study the effect of estimating the vector of means and variance-covariance matrix in the performance of the MEWMA control chart. We show through simulation that the performance of the MEWMA control chart is affected when the parameters are estimated compared to the known parameters case. The control limit that achieves the desired in control ARL when estimating the parameters is computed in several cases and formulas used to find approximately its values are given. Finally, an optimal design strategy for the MEWMA chart with estimated parameters is presented.



URL: http://mc.manuscriptcentral.com/lssp E-mail: comstat@univmail.cis.mcmaster.ca

The Performance of the MEWMA Control Chart When Parameters Are Estimated

Mahmoud¹, M. A. and Maravelakis², P.E.

¹ Cairo University,

Faculty of Economics and Political Science, Department of Statistics,

Cairo, Egypt

² University of the Aegean,

Department of Statistics and Actuarial-Financial Mathematics

Karlovasi, 83200, Samos

Abstract: In the present article, we study the effect of estimating the vector of means and variance-covariance matrix in the performance of the multivariate exponentially weighted moving average (MEWMA) control chart. We show through simulation that the performance of the MEWMA control chart is affected when the parameters are estimated compared to the known parameters case. We show also that larger number of Phase I samples are required to achieve the expected statistical performance of the MEWMA chart when smaller smoothing constants are used in designing it. Under some sampling scenarios, more than 2500 samples are needed to estimate the unknown parameters to satisfy the intended statistical performance. The control limit that achieves the desired in control ARL when estimating the parameters is computed in several cases and formulas used to find approximately its values are given. Finally, an optimal design strategy for the MEWMA chart with estimated parameters is presented.

Keywords: Average Run Length, EWMA, Estimation Effect, Multivariate Control charts, Statistical Process Control.

1. Introduction

Control charts are a valuable tool for detecting an out of control process. During the last decades there is an increasing interest for multivariate control chart techniques due to the fact that many processes involve several quality characteristics (variables) that are correlated. One of the most known control charts that belongs to this category is the multivariate exponentially weighted moving average (MEWMA) chart. The MEWMA control chart was introduced by Lowry et al. (1992) and it is used to detect fast small to moderate shifts in a process. Prabhu and Runger (1997) showed how to design this chart.

Generally speaking, the monitoring of a multivariate process includes two phases: Phase I and Phase II. In Phase I, one checks the stability of the process and estimates the unknown parameters. However, the main purpose in Phase II is to detect out-of-control conditions as quickly as possible. The performance of a Phase I control chart is measured in terms of the probability of chart signals. On the other hand, the performance of a Phase II control chart is usually measured in terms of the parameters of the run length, where the run length is defined as the number of samples taken until the chart produces a signal. In order to use the MEWMA control chart during a Phase II application, the existing procedures for designing it are based on the assumption of known process mean vector μ and variance covariance matrix Σ . In practice, however, μ and Σ are usually unknown and we have to estimate these parameters using a historical data set (Phase I). Using parameters estimates with design procedures intended for known parameters can result in significantly deteriorated chart performance since these estimators add extra variability in the process.

Several authors have dealt with this problem for different univariate control charts, see e.g., Chen (1997) and Jones et al. (2001, 2004). Quesenberry (1993) studied the effect of parameter estimation on the performance of the univariate Shewhart chart. This researcher suggested that at least 100 Phase I subgroups with sample size of 5 observations is needed to achieve expected statistical performance of the Shewhart chart. Jones et al. (2001) showed that a number of subgroups more than 100 is needed to achieve expected statistical performance of the univariate EWMA

chart. The problem in the multivariate case has been studied only in the case of the T^2 chart (Nedumaran and Pignatiello (1999) and Champ et al. (2005)). None of the previous studies investigated the effect of parameter estimation on the performance of the MEWMA chart. This has been pointed by Jensen et al (2006) along with the conjecture that no attention has been given to the multivariate EWMA chart a fact that motivated this work. Our main focus in this paper is to investigate the effect of estimating μ and Σ from an in-control Phase I data set on the performance of the MEWMA control chart used to monitor the process mean vector μ .

In Section 2, we present the MEWMA chart and its main properties. The effect of estimating the parameters on the performance of the proposed scheme is given in Section 3. In Section 4, we present the control limits that give the desired in-control average run length (ARL) when m samples of size n are used for estimating the parameters of the p-variate MEWMA chart. Section 5 presents the way to design the MEWMA control chart. Finally some comments and concluding remarks are given in Section 6.

2. The MEWMA control chart

Let X_i i=1,2,... be $(p \times 1)$ random vectors taken at regular time intervals and assumed to follow a p-variate normal distribution $N_p(\mu_0, \Sigma_0)$, where μ_0 is the known p-dimensional vector of means and Σ_0 is the known variance-covariance matrix. Each vector records the values of p correlated quality characteristics that we want to monitor simultaneously and we further assume that these X_i 's, i=1, 2,... are independent. The random vector X_i can be either the sample mean vector or individual observation vector at time i. The main concern is to detect shifts in the process mean vector μ from the target vector μ_0 . We assume without loss of generality that $\mu_0 = 0$. The multivariate EWMA control chart statistic given by Lowry et al. (1992) is

$$\mathbf{Z}_{i} = \mathbf{R}\mathbf{X}_{i} + (\mathbf{I} - \mathbf{R})\mathbf{Z}_{i-1}, i = 1, 2, 3,...$$
 (1)

where \mathbf{R} =diag($r_1, r_2, ..., r_p$), $0 < r_j \le 1$, j= 1, 2,..., p, and \mathbf{I} is the identity matrix. Usually the observations are weighted using the same value r, that is $r_1 = r_2 = ... = r_p = r$. The initial value \mathbf{Z}_0 is usually set equal to the in-control mean vector of the process (i.e., $\mathbf{Z}_0 = \boldsymbol{\mu}_0$). A straightforward note is that if \mathbf{R} = \mathbf{I} then the MEWMA control chart is equivalent to the T^2 Chart.

The MEWMA chart gives an out-of-control signal if $\mathbf{Z}_i' \boldsymbol{\Sigma}_{\mathbf{Z}_i} \mathbf{Z}_i > h$ where $\boldsymbol{\Sigma}_{\mathbf{Z}_i}$ is the variance-covariance matrix of \mathbf{Z}_i and h is the control limit used to achieve the specified in-control ARL. The variance-covariance matrix of \mathbf{Z}_i when $r_1 = r_2 = ... = r_p = r$ is equal to

$$\Sigma_{\mathbf{Z}_i} = \frac{r[1 - (1 - r)^{2i}]}{(2 - r)} \Sigma$$

The MEWMA chart statistic is usually constructed in terms of the asymptotic covariance matrix

$$\Sigma_{\mathbf{z}_i} = \frac{r}{(2-r)} \Sigma. \tag{2}$$

When μ_0 and Σ_0 are unknown, then *m* in-control Phase I samples, each of size *n* are used to estimate them. The in-control process mean vector μ_0 is estimated by

$$\overline{\overline{\mathbf{X}}} = \frac{\sum_{j=1}^{m} \overline{\mathbf{X}}_{j}}{m},\tag{3}$$

where $\overline{\mathbf{X}}_j$ is the j^{th} sample mean vector. Also, the unknown variance-covariance matrix Σ is estimated by

$$\overline{\mathbf{S}} = \frac{\sum_{j=1}^{m} \mathbf{S}_{j}}{m},\tag{4}$$

where S_i is the within sample variance-covariance matrix.

For simplicity, we consider only charts with equal smoothing parameters and chart statistic calculated based on the asymptotic covariance matrix in Eq. (2). Lowry

et al. (1992) showed that the MEWMA chart with equal smoothing parameters is directionally invariant. Champ et al. (2005) showed that the in-control run length distribution of the directionally invariant T^2 chart statistic does not depend on the unknown process parameters. The in-control run length distribution of the T^2 chart statistic depends only on the number of variables p, the number of Phase I groups m, and the sample size within each group n. The out-of-control run length distribution depends on these constants as well as the statistical distance between the in-control and out-of-control mean vectors. This distance is usually referred to as the non-centrality parameter. Thus, the performance analysis of the T^2 chart can be given without knowledge of the in-control values of the process parameters or their estimates.

Lowry et al. (1992) showed that the ARL performance of the MEWMA control chart with equal smoothing parameters (i.e., $r_1 = r_2 = ... = r_p = r$) depends only on the non-centrality parameter. This means that the MEWMA chart has the property of directional invariance. The use of unequal weights cancels out this property since in this case the ARL depends on the direction of the shift. It can be shown that the run length distribution of the directionally invariant MEWMA control chart with estimated parameters depends only on the values of p, m, n, and the smoothing parameter r. The proof is given in Champ and Jones-Farmer (2007).

Rigdon (1995a,b) gave an integral and a double integral equation for the calculation of in-control and out-of-control ARLs of the MEWMA chart. Runger and Prabhu (1996) used a Markov chain approximation to determine the run length performance of the MEWMA chart. Evaluating the run length performance of the MEWMA chart using the double integral equations or Markov chain approximation when the parameters are estimated from m in-control Phase I samples is intractable. Thus, the results we present in this paper are rather based on extensive Monte Carlo simulation.

Prabhu and Runger (1997) provided recommendations for the selection of parameters for a MEWMA chart. To detect quickly small sustained shifts in the process mean vector, it is usually recommended that one design the MEWMA chart

using small values of the smoothing parameter r; see Lowry et al. (1992). To achieve robustness to violations of the multivariate normality assumption, Stoumbos and Sullivan (2002) recommended the use of smoothing parameters in the range from 0.02–0.05 for the MEWMA chart. However, as mentioned in Mahmoud and Zahran (2010), a very large sudden shift in the process mean vector may not be detected immediately if small values of r are used. When using a small value of r, the current observation receives a small weight r, and an observation far from the target value may not result in an immediate out-of-control signal. Moreover, the MEWMA charts with small values of r can build up an arbitrary large amount of inertia; see for example Woodall and Mahmoud (2005) and Mahmoud and Zahran (2010). In this paper, we investigate the performance of the MEWMA chart with some different recommended values of the smoothing parameters r=0.05, 0.10, and 0.20 when the incontrol parameters are estimated.

3. Effect of estimated parameters on the performance of the MEWMA chart

When estimates are used instead of unknown parameters, the sampling distribution of the MEWMA chart statistics should account for the variability in the estimators. If the researcher does not take this variability into account, the in-control and out-of-control performance of the control charts can be strongly affected. The direct result of not taking this variability into account is the significant increase of the number of false alarms, especially when the sample size of the phase I data set is small. As the sample size of the phase I data set increases, the estimation error decreases. In this section we investigate using simulation the performance of the MEWMA chart with estimated parameters.

As mentioned in Section 2, when the mean vector and the covariance matrix are unknown they are usually estimated using an in-control reference sample consisting of m subgroups of size n from an in-control $N_p(\mu_0, \Sigma_0)$ where $m \ge 1$ and m(n-1) > p. Then the estimators used are $\hat{\mu}_0 = \overline{\overline{X}}$ and $\hat{\Sigma}_0 = \overline{S}$, where $\overline{\overline{X}}$ and \overline{S} are as defined in Eqs (3) and (4), respectively. Anderson (1984) showed that if a random sample $X_1, X_2, ..., X_n$ is taken from $N_p(\mu_0, \Sigma_0)$ at sampling stage j, then the sample

mean vector $\overline{\mathbf{X}}_j \sim N_p(\boldsymbol{\mu}_0, \frac{\boldsymbol{\Sigma}_0}{n})$ and the matrix $(n-1)\mathbf{S}_j \sim Wishart_p(\boldsymbol{\Sigma}_0, n-1)$, where $Wishart_p(\boldsymbol{\Sigma}_0, n-1)$ is the p-variate Wishart distribution with parameters $\boldsymbol{\Sigma}_0$ and n-1 degrees of freedom. Accordingly, the in-control estimator of the process mean vector $\overline{\overline{\mathbf{X}}} \sim N_p(\boldsymbol{\mu}_0, \frac{\boldsymbol{\Sigma}_0}{nm})$ and the matrix $[m(n-1)\overline{\mathbf{S}}] \sim Wishart_p(\boldsymbol{\Sigma}_0, m(n-1))$.

As mentioned in Section 2, the in-control run length distribution of the MEWMA chart statistic depends only on p, m, n and r. Since the ARL performance of the MEWMA chart depends on the in-control parameters or their estimates only through the non-centrality parameter, we assume without loss of generality that $\mu_0 = 0$ and $\Sigma_0 = \mathbf{I}$. In our simulation studies we considered five different values for p (p=2, 3, 4, 5, or 6), nine values for p (p=3, 40, 50, 70, 100, 150, 200, 300, or 500), four values for p (p=3, 5, 10, or 15), and three values for p (p=3, 5, 10, or 15), and three values for p (p=3). The values used for p, p, and p were chosen in order to represent small to large values of these parameters. The three selected values for the smoothing parameter p are the most recommended values in the literature.

In our simulation study, the procedure used to investigate the effect of p, m, n and r on the in-control ARL performance of the MEWMA chart can be described as follows:

- 1- Using 50,000 simulation runs, we first estimated the upper control limit *h* needed to produce an in-control ARL of 200 when different values of the smoothing parameter *r* is used in designing the MEWMA chart statistic with known parameters. The performance and design of a multivariate control charts based on an in-control ARL of 200 were studied by several researchers; see for example Lowry et al. (1992), Champ et al. (2005), and Mahmoud and Zahran (2010). Table 1 gives the values of *h* according to different values of *p* and *r*. For example, when *r*=0.05 the upper control limit that gives the desired in-control ARL is *h*=7.36 in case of *p*=2, *h*=9.41 in case of *p*=3 and *h*=12.97 in case of *p*=5.
- 2- Then, we generated a random vector $\overline{\overline{\mathbf{X}}}$ from $N_p(\boldsymbol{\mu}_0, \frac{\boldsymbol{\Sigma}_0}{nm})$ and matrix $\overline{\mathbf{S}}$ from $Wishart_p(\boldsymbol{\Sigma}_0, m(n-1))$. The random vector $\overline{\overline{\mathbf{X}}}$ and the matrix $\overline{\mathbf{S}}$ were

- generated to represent the random information given in the parameter estimates stage (Phase I).
- 3- Then we generated a random vector \mathbf{X}^{i} to represent the new Phase II process information observed at the i^{th} sampling stage.
- 4- The MEWMA statistic defined in Eq. (1) was then calculated based on the estimated parameters $\overline{\overline{X}}$ and \overline{S} and compared to the upper control limit h calculated in Step 1.
- 5- Steps 3-4 are repeated until a signal is given. When the signal is given then we record the run length.
- 6- Steps 2-5 are repeated 50,000 times and the ARL is then calculated. The procedure IML in the statistical package SAS was used to generate the random vectors and matrices and to estimate the ARL values.

(Insert Table 1 about here)

Tables 2-4 give the in-control ARL values of the MEWMA chart for different values of m, n, p, and r. As shown in these tables, the in-control performance of the MEWMA control chart with estimated parameters is strongly affected if the researcher uses the upper control limit h designed based on known parameters. For example, as shown in Table 2, when the smoothing parameter r=0.05 and h=7.36 are used in designing the MEWMA chart for monitoring bivariate quality characteristics (p=2) and the parameters are estimated using m=30 Phase I samples with n=3observations, the in-control ARL produced by the chart is only 85.82. The in-control ARL of this chart when the parameters are known is 200. This example shows that the in-control run length distribution is seriously affected by ignoring the variability added by estimating the unknown parameters. As expected, as m and/or n increases, the number of false alarms produced by the MEWMA chart decreases, as shown in Tables 2-4. For instance, in the previous example if m increases from 30 to 500, the ARL increases from 85.82 to 177.15. On the contrary, the number of false alarms produced by the MEWMA chart increases as p increases. Compared to the previous example, when m=30, n=3, r=0.05, and p=6, the in-control ARL produced by the MEWMA chart is reduced to 41.49.

(Insert Tables 2-4 about here)

As for the effect of the smoothing parameter r used in designing the MEWMA chart, the results in Tables 2-4 show that as r increases the number of false alarms produced by the MEWMA chart decreases. For example, if m=30, n=3, p=2 and r increases from 0.05 to 0.10, the ARL increases from 85.82 to 94.68. If the same m, n, and p are used and r increases from 0.10 to 0.20, the in-control ARL increases from 94.68 to 111.18. Notice that the T^2 chart is a special case of the MEWMA chart when r=1. Our simulation study shows that if r increases from 0.20 to 1, the in-control ARL increases from 111.18 to 164.02.

In order to reduce the variability in the sampling distribution of the estimates so that the ARL performance of a chart with estimated parameters is not significantly affected, the number of Phase I samples m should increase to an appropriate level. For example, Champ et al. (2005) recommended a number of Phase I samples m>900/n when designing the 2-variate T^2 control chart with estimated parameters. Several researchers gave the appropriate number of samples required to achieve a desired incontrol ARL when a control chart with estimated parameters is used when the sample size n=5; see for example Quesenberry (1993) and Jones (2002). Our simulation results show that the sufficient number of Phase I samples depends on the smoothing parameter r. These results show that larger number of Phase I samples are required when smaller smoothing constants are used. For example, while at least 2000 samples of size 5 is needed to achieve the expected statistical performance of the 2-variate MEWMA chart with estimated parameters when r=0.05, only 1400 and 1000 samples are sufficient to achieve it when r=0.10 and 0.20, respectively. As expected, monitoring a mean vector with a dimension more than 2 using the MEWMA chart with estimated parameters requires larger number of Phase I samples than those listed for p=2 to achieve the desired ARL performance. For example, when p=6 and r=0.20, more than 2500 samples of size 5 are needed so that the in-control ARL produced by the MEWMA chart is 200.

4. Corrected Limits for the MEWMA chart with estimated parameters

In some practical applications the data are plentiful and hence waiting until 1000 or more Phase I samples to accumulate is feasible. However, in most real-life

applications such large number of Phase I samples is usually not available. As shown in Section 3, if the practitioner uses a control chart with parameters estimated based on small number of samples, the statistical performance of the chart can be very poor. On the other hand, waiting a long time to obtain the necessary data may result in the presence of undetected parameter shifts during this time. Therefore, several researchers have developed design procedures for some quality control charts that do not require the assumption of known parameters; see Quesenberry (1993), Jones (2002), and Champ et al. (2005). The main idea of these procedures was to use wider control limits for the chart in order to reflect the variability of the parameter estimates used in constructing the chart statistic. Champ et al. (2005) referred to these limits as the "corrected limits". In this section we develop using simulation the corrected limits of the MEWMA chart that give exactly an in-control ARL of 200 when m Phase I samples of size n are used in estimating the unknown parameters.

Tables 5-7 give the simulated corrected control limits for the p-variate MEWMA chart with parameters estimated from m samples of size n that give incontrol ARL of 200 when r=0.05, 0.10, and 0.20, respectively. These limits were obtained by the simulation algorithm (steps 2-6) described in Section 3. An additional binary search similar to that described in Champ et al. (2005) was used to obtain these limits. The main idea of the binary search was to adjust the limits until the desired ARL value was obtained.

(Insert Tables 5-7 about here)

The results in Tables 5-7 show that the corrected limits depend on the size of m, n, p, and r. Smaller values m and/or n require wider corrected limits to achieve the desired in-control ARL. In practice, if m and/or n are different from the values given in Tables 5-7, the practitioner may use interpolation to determine an appropriate corrected control limit. Alternatively, we give estimated corrected limits for the p-variate MEWMA charts in Tables 8-10 when r=0.05, 0.10, and 0.20, respectively. The linear equation models given in these tables are least squares estimates of the corrected limits using the logarithm of m (with base=10) as the independent variable in simple linear regression models. For example, when n=5, p=2, and r=0.05, Table 8 gives the estimate of the corrected limits as a linear function of the logarithm of m by:

$$h = 13.055 - 2.118 \times \log_{10} m$$
, $30 \le m \le 500$. (5)

When n=5, p=2, and r=0.05, if m=210 samples are available, the estimated corrected limits calculated from Equation (5) is 8.14. Using 50,000 simulation runs, this estimated corrected limit produces an in-control ARL of 208. Each of the least squares estimates of the corrected limits listed in Tables 8-10 produced a coefficient of determination R^2 of at least 90%.

(Insert Tables 8-10 about here)

5. Optimal design of the MEWMA chart with estimated parameters

Although the corrected control limits given in Section 4 result in a chart with the desired in-control ARL, a subsequent increase in the out-of-control ARL of the MEWMA chart with estimated parameters over the chart with known parameters is expected. The optimal statistical design of an MEWMA chart requires not only determining the control limit h that satisfies the desired in-control ARL, but also the smoothing parameter r that minimizes the out-of-control ARL for a shift of a specified size measured in terms of the non-centrality parameter. As mentioned before, the out-of-control run length performance of the MEWMA chart does not depend on the parameters or their estimates. This performance rather depends on the values of m and n through the value of the corrected control limit, the smoothing parameter r, the number of monitored variables p, and the size of the shift measured in terms of the non-centrality parameter λ .

In this section, we compute the value of the smoothing parameter r and the corresponding corrected limit h that optimize the out-of-control performance of the p-variate MEWMA chart for different values of m of size n=5. The MEWMA chart is usually recommended for detecting optimally small sustained shifts. In our study, two sizes of the process shift were considered; a standardized shift in the non-centrality parameter from λ =0 to $\lambda = \delta / \sqrt{n}$ was considered, where δ =0.5 or 1. Three values were considered for p (p=2, 3, or 4) and 4 values for m (m=30, 50, 100, or 200). The procedure used to find the optimal design of the MEWMA chart with estimated parameters can be described as follows:

- 1. For given values of *m* and *p*, all the combinations of *r* and *h* that produce an in-control ARL of 200 were obtained based on Steps 2-6 in the procedure described in Section 3.
- 2. For each combination of r and h, a standardized shift in the non-centrality parameter of size δ/\sqrt{n} is introduced and the out-of-control ARL is then estimated. This shift was made by changing the process mean vector from $\mu = 0$ to $\mu = \mu_1$ such that $\sqrt{\mu_1' \Sigma^{-1} \mu_1} = \delta/\sqrt{n}$. The optimal values of r and h are those that minimize the out-of-control ARL for the specified shift.

Tables 11 and 12 give the optimal combination of r and h for different values of m and p when $\delta = 0.5$ and 1, respectively. The second column in both tables gives the optimal values of r in detecting shifts of various sizes for an in-control ARL of 200. The following columns give the values of the upper control limit h that achieve the desired in-control ARL of 200 when m samples of size n=5 are used in estimating the unknown parameters and the out-of-control ARL corresponding to the specified shift. The last two columns in these tables give the upper control limits and the out-of-control ARLs when the MEWMA statistic is calculated based on known parameters.

(Insert Tables 11-12 about here)

Our simulation results show that the optimal value of r does not depend on the number m of Phase I samples or on whether the MEWMA chart statistic is calculated based on known or estimated parameters. Also, this optimal value does not vary much as p changes. The optimal values of r given in this paper are very close to the optimal values given in Table 2 in Lowry et al. (1992, p. 49).

The results in Tables 11-12 show that the out-of-control ARL performance improves significantly as the number of samples m increases. For example, when p=2, the out-of-control ARL for detecting a shift of size δ =0.5 decreases from 46.41 when m=30 to 29.72 when m=200. When known parameters are used in constructing the 2-variate MEWMA chart statistic, the out-of-control ARL for detecting this size of shift is 26.50, as shown in Table 11.

5. Conclusions

In this paper we have studied the effect of estimating the parameters on the MEWMA control chart. Specifically, we have shown through simulation that the performance of the MEWMA chart is seriously affected if the vector of means and the covariance matrix are estimated based on small number of Phase I samples. The value of the control limit is computed for several combinations of the number of samples and sample size so that the in-control performance of the MEWMA chart in the case of known and estimated parameters is equivalent. Finally, an optimal design of the MEWMA chart is presented providing to the practitioners the suitable smoothing parameter for fast detection of an out of control situation for different shift sizes.

References

Anderson, T.W. (1984). *An Introduction to Multivariate Statistical Analysis* (2nd edition), New York: Wiley.

Champ, C.W. and Jones-Farmer, L.A. (2007). Properties of Multivariate Control Chart with Estimated Parameters, Sequential Analysis, 26: 153-169.

- Champ, C.W., Jones-Farmer, L.A. and Rigdon, S.E. (2005). Properties of the T^2 Control Chart when Parameters are Estimated, *Technometrics*, 47: 437-445.
- Chen, G. (1997). The mean and standard deviation of the run length distribution of \overline{X} charts when control limits are estimated, *Statistica Sinica*, 7: 789-798.
- Jensen, W.A., Jones-Farmer, L.A., Champ, C.W. and Woodall, W.H. (2006). Effects of parameter estimation on control chart properties: A literature review, *Journal of Quality Technology*, 38: 349-364.
- Jones, L.A. (2002). The statistical design of the EWMA control charts with estimated parameters, *Journal of Quality Technology*, 34: 277-288.
- Jones, L.A., Champ, C.W. and Rigdon, S.E. (2001). The performance of exponentially weighted moving average charts with estimated parameters, *Technometrics*, 43: 156-167.
- Jones, L.A., Champ, C.W. and Rigdon, S.E. (2004). The run length distribution of the CUSUM with estimated parameters, *Technometrics*, 43: 156-167.

- Lowry, C.A., Woodall, W.H., Champ, C.W. and Rigdon, S.E. (1992). A multivariate exponentially weighted moving average control chart, *Technometrics*, 34: 46-53.
- Mahmoud, M.A., and Zahran, A. (2010). A multivariate adaptive exponentially weighted moving average control chart, *Communication in Statistics: Theory and Methods*, 39: 606-625.
- Nedumaran, G. and Pignatiello, J. (1999). On constructing T^2 control charts for online process monitoring, *IIE Transactions*, 31: 529-536.
- Prabhu, S.S. and Runger, G.C. (1997). Designing a multivariate EWMA control chart, *Journal of Quality Technology*, 29: 8-15.
- Quesenberry, C.P. (1993). The effect of sample size on estimated limits for \overline{X} and X control charts, *Journal of Quality Technology*, 25: 237-247.
- Rigdon, S.E. (1995a). A double-integral equation for the average run length of a MEWMA control chart, *Statistics and Probability Letters*, 24: 365-373.
- Rigdon, S.E. (1995b). An integral equation for the in-control average length of a MEWMA Control Chart, *Journal of Statistical Computation and Simulation*, 52: 351-365.
- Runger, G.C. and Prabhu, S.S. (1996). A Markov chain model for the multivariate EWMA control chart, *Journal of the American Statistical Association*, 91: 1701-1706.
- Stoumbos, Z.G. and Sullivan, J.H. (2002). Robustness to non-normality of the multivariate EWMA control chart, *Journal of Quality Technology*, 34: 260-276.
- Woodall, W. H. and Mahmoud, M. A. (2005). The inertial properties of quality control charts, *Technometrics*, 47: 425-436.

Table 1: The upper control limit value that produce an in-control ARL of 200 when known parameters are used in constructing the MEWMA chart statistic, given for different values of p and r.

p	r						
P	0.05	0.10	0.20				
2	7.36	8.67	9.67				
3	9.41	10.77	11.87				
4	11.21	12.72	13.86				
5	12.97	14.52	15.75				
6	14.59	16.29	17.48				

Table 2: The ARL values when m Phase I samples with p variables, each of size n are used to estimate the unknown parameters and the UCL used is the one that gives incontrol ARL of 200 when the parameters are known. The smoothing parameter used in calculating the MEWMA chart statistic is r=0.05.

n	n					m				
p	11	30	40	50	70	100	150	200	300	500
	3	85.82	97.16	105.79	116.00	132.83	146.09	154.80	164.75	177.15
2	5	87.05	97.31	106.19	116.12	133.46	148.18	155.68	166.05	177.52
	10	89.38	99.00	107.04	119.90	134.34	148.79	156.16	168.32	179.61
	15	89.40	99.33	107.99	122.56	135.93	149.46	156.51	168.72	180.42
	3	65.82	76.80	86.22	101.01	115.54	133.40	146.50	155.21	174.51
3	5	69.21	81.05	91.00	104.37	120.61	134.99	147.27	160.15	174.97
	10	71.58	81.86	91.33	104.73	121.31	136.83	147.40	162.83	175.29
	15	71.93	82.75	91.51	105.07	121.45	137.26	147.49	162.90	175.32
	3	55.11	65.69	74.75	89.62	104.48	120.48	135.41	149.26	164.57
4	5	58.78	69.17	78.43	92.66	108.16	124.97	137.90	151.95	166.56
	10	60.23	71.39	79.62	95.91	110.04	125.70	138.80	152.58	167.75
	15	61.08	71.62	80.33	96.87	110.22	125.99	138.88	152.67	167.63
	3	46.80	56.12	65.24	79.51	96.15	115.84	127.03	145.00	164.05
5	5	50.79	60.12	69.05	84.82	100.49	119.05	131.22	148.97	164.35
	10	52.79	63.23	72.10	86.25	102.94	121.01	134.31	151.11	164.76
	15	54.07	64.64	73.89	87.44	103.78	122.04	134.20	151.94	164.82
	3	41.49	49.97	58.45	71.52	87.10	109.71	120.25	139.96	157.06
6	5	45.05	55.06	62.75	78.01	94.82	113.90	125.96	143.64	161.12
	10	47.93	57.47	65.97	81.07	95.74	115.46	126.11	143.91	161.23
	15	49.83	58.54	67.85	81.42	96.39	115.86	127.71	144.00	161.38

Table 3: The ARL values when m Phase I samples with p variables, each of size n are used to estimate the unknown parameters and the UCL used is the one that gives incontrol ARL of 200 when the parameters are known. The smoothing parameter used in calculating the MEWMA chart statistic is r=0.10.

n	n					m				
p	11	30	40	50	70	100	150	200	300	500
	3	94.68	105.37	117.50	131.88	141.05	156.04	163.53	175.83	182.53
2	5	97.64	107.94	118.49	132.88	142.27	157.93	165.22	176.79	182.97
	10	99.05	108.90	119.13	133.17	143.68	158.35	166.13	176.87	183.07
	15	100.05	109.41	120.24	134.38	144.45	158.76	166.56	176.88	183.18
	3	72.58	84.98	96.48	109.84	123.76	142.82	153.68	162.20	175.07
3	5	76.10	88.75	99.88	114.06	129.55	146.33	156.01	165.74	175.58
	10	78.93	90.38	101.81	115.78	130.13	148.35	158.65	166.38	176.38
	15	80.94	92.07	103.00	116.16	130.56	148.90	158.97	166.52	177.03
	3	57.12	69.73	79.38	97.49	113.07	133.59	145.66	155.98	171.34
4	5	64.07	74.23	86.30	102.24	116.60	137.67	148.39	158.57	175.72
	10	66.42	78.72	90.55	106.12	122.32	140.77	149.54	160.42	175.75
	15	68.02	80.15	91.10	106.49	123.60	141.14	150.64	161.16	175.77
	3	47.12	59.59	73.14	85.40	104.13	121.62	134.37	151.85	168.59
5	5	55.36	67.32	77.72	93.64	110.77	127.72	140.27	155.50	170.07
	10	60.11	70.52	80.49	95.90	112.82	131.77	143.87	157.64	170.23
	15	61.92	74.68	83.05	96.44	113.34	134.96	145.84	158.41	170.86
	3	40.95	52.88	62.87	80.70	98.35	118.29	135.17	150.76	171.17
6	5	48.96	59.98	70.63	87.59	105.97	124.67	137.46	154.31	172.25
	10	53.60	66.50	76.31	92.44	108.24	129.14	142.55	156.44	172.36
	15	55.26	68.29	78.57	93.57	111.34	131.42	143.18	157.52	172.44

Table 4: The ARL values when m Phase I samples with p variables, each of size n are used to estimate the unknown parameters and the UCL used is the one that gives incontrol ARL of 200 when the parameters are known. The smoothing parameter used in calculating the MEWMA chart statistic is r=0.20.

n	n					m				
p	n n	30	40	50	70	100	150	200	300	500
	3	111.18	123.14	130.29	144.76	155.72	163.87	171.10	180.10	186.79
2	5	111.87	124.23	133.75	145.64	156.41	164.38	173.79	182.78	190.02
	10	112.31	125.58	135.51	146.16	158.13	166.77	175.06	184.26	191.10
	15	113.33	126.44	136.68	147.95	159.70	168.85	179.52	187.21	193.14
	3	83.04	98.25	107.89	122.11	140.72	155.51	162.12	172.88	183.90
3	5	88.71	103.68	113.69	129.19	144.20	157.60	165.87	174.87	183.08
	10	94.36	107.27	117.52	131.73	147.15	158.91	168.20	176.42	184.04
	15	95.85	108.64	118.30	133.46	148.49	159.18	169.88	177.89	184.11
	3	65.67	79.35	90.72	108.96	126.54	141.24	154.43	169.87	178.05
4	5	74.44	89.60	100.95	116.56	135.59	150.42	162.37	171.57	180.24
	10	82.32	97.49	106.64	122.08	139.88	153.87	163.31	173.04	182.48
	15	86.23	100.44	109.67	126.05	139.27	154.37	164.43	174.40	183.09
	3	52.73	66.84	79.18	97.48	115.81	140.31	149.07	163.73	175.88
5	5	66.54	80.35	93.03	107.26	128.87	146.09	157.31	169.66	179.82
	10	74.93	87.30	98.99	111.42	135.49	151.88	162.37	170.30	182.14
	15	77.38	90.46	103.05	115.86	137.51	154.05	163.63	171.60	182.55
	3	43.89	57.01	67.38	86.24	106.19	127.25	139.82	153.76	171.65
6	5	56.18	71.26	82.32	98.99	117.66	138.86	149.88	161.15	175.50
	10	66.01	80.47	92.07	107.81	126.42	142.00	156.00	165.97	176.61
	15	68.42	81.77	95.81	111.59	129.89	144.83	158.26	166.86	178.42

Table 5: The UCL values that produce in-control ARL of 200 when m Phase I samples with p variables, each of size n are used to estimate the unknown parameters. The smoothing parameter used in calculating the MEWMA chart statistic is r=0.05.

p	n					m				
P	11	30	40	50	70	100	150	200	300	500
	3	10.25	9.82	9.45	8.98	8.62	8.28	8.08	7.85	7.66
2	5	10.23	9.78	9.41	8.96	8.60	8.26	8.06	7.84	7.65
	10	10.20	9.74	9.37	8.96	8.60	8.26	8.06	7.83	7.65
	15	10.17	9.68	9.35	8.95	8.59	8.25	8.05	7.83	7.65
	3	14.17	13.25	12.70	11.95	11.30	10.75	10.45	10.13	9.86
3	5	13.95	13.10	12.54	11.80	11.22	10.68	10.41	10.12	9.85
	10	13.73	12.95	12.40	11.68	11.21	10.65	10.39	10.11	9.85
	15	13.63	12.84	12.33	11.65	11.20	10.65	10.37	10.10	9.85
	3	18.10	16.63	15.80	14.70	13.85	13.10	12.65	12.23	11.85
4	5	17.45	16.25	15.40	14.43	13.63	12.93	12.55	12.17	11.82
	10	17.00	15.97	15.26	14.35	13.57	12.90	12.53	12.15	11.80
	15	16.90	15.85	15.20	14.30	13.55	12.89	12.53	12.15	11.80
	3	21.97	20.00	18.73	17.40	16.23	15.27	14.70	14.10	13.70
5	5	20.83	19.30	18.28	17.00	15.97	15.10	14.58	14.00	13.68
	10	20.31	18.88	18.10	16.80	15.85	14.98	14.53	13.98	13.65
	15	20.22	18.70	17.93	16.77	15.80	14.95	14.50	13.98	13.65
	3	25.95	23.40	21.65	20.00	18.61	17.33	16.72	16.05	15.48
6	5	24.22	22.35	21.08	19.44	18.20	17.08	16.53	15.95	15.44
	10	23.45	21.75	20.65	19.15	18.00	16.97	16.44	15.93	15.43
	15	23.30	21.49	20.47	19.10	17.93	16.94	16.40	15.92	15.43

Table 6: The UCL values that produce in-control ARL of 200 when m Phase I samples with p variables, each of size n are used to estimate the unknown parameters. The smoothing parameter used in calculating the MEWMA chart statistic is r=0.10.

n	n					m				
p	11	30	40	50	70	100	150	200	300	500
	3	10.84	10.44	10.17	9.84	9.58	9.30	9.13	8.99	8.86
2	5	10.82	10.40	10.14	9.81	9.55	9.27	9.11	8.97	8.85
	10	10.81	10.37	10.12	9.79	9.53	9.26	9.10	8.97	8.85
	15	10.80	10.36	10.11	9.77	9.52	9.26	9.10	8.97	8.85
	3	14.47	13.72	13.25	12.70	12.18	11.81	11.53	11.29	11.10
3	5	14.12	13.45	13.08	12.54	12.09	11.76	11.47	11.27	11.09
	10	13.93	13.36	12.98	12.48	12.04	11.72	11.46	11.26	11.09
	15	13.87	13.32	12.93	12.42	12.02	11.69	11.46	11.25	11.09
	3	18.00	16.75	16.12	15.30	14.62	14.03	13.70	13.40	13.14
4	5	17.32	16.38	15.80	15.07	14.40	13.92	13.65	13.35	13.11
	10	16.88	16.12	15.56	14.92	14.33	13.86	13.63	13.32	13.10
	15	16.79	16.01	15.47	14.82	14.28	13.83	13.62	13.31	13.09
	3	21.37	19.80	18.94	17.86	16.91	16.18	15.77	15.39	15.10
5	5	20.24	19.16	18.35	17.40	16.63	15.99	15.65	15.35	15.06
	10	19.73	18.70	18.02	17.15	16.46	15.90	15.62	15.32	15.03
	15	19.58	18.55	17.85	17.08	16.40	15.87	15.60	15.30	15.00
	3	24.85	22.91	21.75	20.25	19.17	18.27	17.77	17.27	16.90
6	5	23.25	21.78	20.82	19.71	18.72	17.96	17.59	17.18	16.82
	10	22.41	21.10	20.31	19.28	18.48	17.82	17.53	17.14	16.77
	15	22.23	20.96	20.19	19.22	18.43	17.76	17.50	17.11	16.74

Table 7: The UCL values that produce in-control ARL of 200 when m Phase I samples with p variables, each of size n are used to estimate the unknown parameters. The smoothing parameter used in calculating the MEWMA chart statistic is r=0.20.

p	n		m									
		30	40	50	70	100	150	200	300	500		
2	3	11.21	10.91	10.69	10.50	10.27	10.13	10.00	9.84	9.80		
	5	11.17	10.87	10.66	10.47	10.26	10.11	9.98	9.83	9.79		
	10	11.14	10.85	10.65	10.45	10.25	10.10	9.97	9.83	9.79		
	15	11.11	10.84	10.64	10.44	10.24	10.10	9.96	9.83	9.79		
3	3	14.50	14.00	13.68	13.20	12.83	12.58	12.31	12.21	12.06		
	5	14.18	13.76	13.41	13.02	12.74	12.50	12.28	12.18	12.05		
	10	13.98	13.60	13.28	12.95	12.69	12.44	12.27	12.17	12.04		
	15	13.90	13.54	13.23	12.92	12.66	12.40	12.26	12.16	12.04		
4	3	17.78	16.93	16.33	15.70	15.23	14.78	14.59	14.34	14.16		
	5	17.08	16.39	15.94	15.42	15.02	14.67	14.48	14.28	14.12		
	10	16.65	16.10	15.70	15.25	14.87	14.60	14.42	14.24	14.11		
	15	16.53	16.00	15.63	15.18	14.80	14.55	14.38	14.22	14.10		
5	3	20.92	19.73	19.00	18.08	17.45	16.92	16.63	16.35	16.13		
	5	19.78	18.86	18.31	17.67	17.14	16.67	16.47	16.28	16.06		
	10	19.15	18.44	18.00	17.38	16.99	16.57	16.40	16.23	16.02		
	15	18.97	18.27	17.87	17.30	16.90	16.52	16.36	16.20	16.00		
6	3	24.17	22.50	21.48	20.43	19.64	18.95	18.61	18.23	17.90		
	5	22.46	21.33	20.70	19.84	19.16	18.65	18.39	18.06	17.86		
	10	21.51	20.73	20.15	19.44	18.95	18.50	18.31	17.96	17.83		
	15	21.32	20.53	20.02	19.33	18.85	18.43	18.25	17.92	17.82		

Table 8: The least square estimate of the UCL as a function of the logarithm of m. The estimate produces in-control ARL close to 200 when the smoothing parameter used in calculating the p-variate MEWMA chart statistic is r=0.05.

n		ľ	ı	
p	3	5	10	15
2	13.116–2.137xlog(<i>m</i>)	13.055–2.118xlog(<i>m</i>)	12.979–2.087xlog(<i>m</i>)	12.893–2.053xlog(<i>m</i>)
3	18.709-3.492xlog(m)	18.282–3.328xlog(<i>m</i>)	17.828–3.145xlog(m)	17.608-3.056xlog(m)
4	24.423-4.974xlog(<i>m</i>)	23.241–4.519xlog(<i>m</i>)	22.551-4.237xlog(<i>m</i>)	22.312–4.138xlog(<i>m</i>)
5	$30.237 - 6.567 \times \log(m)$	28.317 - 5.806 x log(m)	27.382 - 5.436 xlog(m)	27.055–5.307xlog(<i>m</i>)
6	36.167–8.225xlog(<i>m</i>)	33.328-7.095xlog(<i>m</i>)	31.823–6.491xlog(<i>m</i>)	31.372–6.312xlog(<i>m</i>)

Table 9: The least square estimate of the UCL as a function of the logarithm of m. The estimate produces in-control ARL close to 200 when the smoothing parameter used in calculating the p-variate MEWMA chart statistic is r=0.10.

	n									
p	3	5	10	15						
2	12.956–1.612xlog(<i>m</i>)	12.908–1.600xlog(<i>m</i>)	12.851–1.579xlog(<i>m</i>)	12.819-1.567xlog(<i>m</i>)						
3	17.953–2.710xlog(<i>m</i>)	17.282–2.445xlog(m)	16.970–2.320xlog(<i>m</i>)	16.836–2.269xlog(<i>m</i>)						
4	22.795–3.835xlog(<i>m</i>)	21.653–3.385xlog(m)	20.875–3.073xlog(<i>m</i>)	20.629–2.979xlog(<i>m</i>)						
5	27.606–4.987xlog(<i>m</i>)	25.617–4.198xlog(<i>m</i>)	24.541–3.772xlog(<i>m</i>)	24.198–3.641xlog(<i>m</i>)						
6	32.715–6.308xlog(<i>m</i>)	29.755–5.141xlog(<i>m</i>)	28.057-4.469xlog(m)	27.741-4.351xlog(<i>m</i>)						

Table 10: The least square estimate of the UCL as a function of the logarithm of m. The estimate produces in-control ARL close to 200 when the smoothing parameter used in calculating the p-variate MEWMA chart statistic is r=0.20.

, n		P	ı	
p	3	5	10	15
2	12.699–1.146xlog(<i>m</i>)	12.623–1.120xlog(<i>m</i>)	12.569–1.099xlog(<i>m</i>)	12.525–1.082xlog(<i>m</i>)
3	17.093-1.995xlog(m)	16.414–1.729xlog(m)	16.020-1.573xlog(m)	15.867–1.515xlog(<i>m</i>)
4	21.395–2.866xlog(<i>m</i>)	20.043-2.352xlog(<i>m</i>)	19.250–2.042xlog(<i>m</i>)	19.021–1.959xlog(<i>m</i>)
5	25.950–3.781xlog(<i>m</i>)	23.435–2.937xlog(<i>m</i>)	22.309–2.495xlog(<i>m</i>)	21.944–2.358xlog(<i>m</i>)
6	30.046–4.842xlog(<i>m</i>)	27.040–3.661xlog(<i>m</i>)	25.314–2.979xlog(<i>m</i>)	24.908–2.829xlog(<i>m</i>)

Table 11: The optimal value of the smoothing parameter r and the corresponding UCL that give an in-control ARL of 200 when using m samples of size n=5 in estimating the parameters of the p-variate MEWMA chart used to detect a sustained shift in the non-centrality parameter from λ to $\lambda + 0.5/\sqrt{n}$.

			m										
		30		5	50		100		200		∞		
p	r	h	ARL										
2	0.06	10.37	46.41	9.55	38.05	8.86	32.31	8.37	29.72	7.72	26.50		
3	0.05	13.95	54.43	12.54	44.46	11.22	36.93	10.41	33.57	9.37	29.78		
4	0.05	17.45	59.31	15.40	48.60	13.63	41.36	12.55	36.63	11.20	32.23		

Table 12: The optimal value of the smoothing parameter r and the corresponding UCL that give an in-control ARL of 200 when using m samples of size n=5 in estimating the parameters of the p-variate MEWMA chart used to detect a sustained shift in the non-centrality parameter from λ to $\lambda + 1/\sqrt{n}$.

		m									
		30		50		100		200		∞	
p	r	h	ARL	h	ARL	h	ARL	h	ARL	h	ARL
2	0.13	10.99	12.67	10.39	11.59	9.85	10.86	9.47	10.41	9.06	9.96
3	0.13	14.22	14.77	13.24	13.12	12.35	12.28	11.86	11.76	11.23	11.10
4	0.13	17.22	16.58	15.87	14.82	14.67	13.63	14.00	12.88	13.19	11.99

Authors' response to the referees' comments for the article "The Performance of the MEWMA Control Chart When Parameters Are Estimated" by Professor Mahmoud A. Mahmoud and Petros E. Maravelakis.

Reviewer 1.

Pages 3, 4, 5. All corrections suggested by the referee have been incorporated in the revised version.

Page 7. The referee correctly pointed out that this proposition is wrong. We have deleted the proposition.

Pages 8, 9, 10, 12, 13, 14. All corrections suggested by the referee have been incorporated in the revised version.

For the referee's note in Page 9: The MEWMA chart produces an in-control ARL of around 197.2 when m=2500, p=6, r=0.2 and n=5.

Reviewer 2.

Comment 1. The correction suggested by the referee has been incorporated in the revised version.