College of Business Administration
University of Rhode Island

William A. Orme

WORKING PAPER SERIES
encouraging creative research

Sensitivity ratios as a measure of the effects of the mean shift and dispersion shift for Multivariate EWMA monitoring

Xia Pan, Jeffrey Jarrett

2008/2009 No. 16

This working paper series is intended to facilitate discussion and encourage the exchange of ideas. Inclusion here does not preclude publication elsewhere. It is the original work of the author(s) and subject to copyright regulations.

Office of the Dean
College of Business Administration
Ballentine Hall
7 Lippitt Road
Kingston, RI 02881
401-874-2337
www.cba.uri.edu
Sensitivity ratios as a measure of the effects of the mean shift and dispersion shift for Multivariate EWMA monitoring

Xia Pan, Ph.D.
Lingnan College, Sun Yat-sen University
Guangzhou, China
pan1xia2@gmail.com

Jeffrey Jarrett
College of Business Administration
University of Rhode Island
Kingston, RI 02881
jejarrett@mail.uri.edu*

* Corresponding author
ABSTRACT
Previously, studies of the multivariate exponentially weighted moving average (EWMA) control processes centered on methods for constructing quality control charts for average run length (ARL). Simulation and Markov Chain analysis produced methods for constructing such quality control techniques for serially correlated processes. In this paper, we focus on aspects of the distribution of the chart statistic. Based on the distribution of the chart statistics for in-control and out-of-control situations, we propose to use sensitivity ratios to measure the effects of shifts in both mean and dispersion. Using sensitivity measure, we investigate the importance of serial correlation in quality monitoring and its impact on the sensitivity of performance. This allows for adjustments in the optimal exponential weight factor as it relates to serial correlation.

Keywords: Multivariate EWMA, Quality, Sensitivity Analysis.
Introduction


Shewhart quality control charts for multivariate processes employ the Hotelling $T^2$ statistic for a current sample. Since, univariate Shewhart control charts are not sensitive to small and moderate shifts in process parameters, one often employs other methods. Exponentially weighted moving average (EWMA) charts are more sensitive (Roberts, 1959; Crowder, 1989; Lucas and Suscatici, 1990). Hence, LWCR extended the univariate EWMA control chart to the multivariate case by simulation. They considered that the multivariate EWMA chart has greater sensitivity to shifts in the mean than more traditional Hotelling $T^2$ control methods.

LWCR proposed to build the EWMA quantity $\tilde{z}_i = r\tilde{x}_i + (1 - r)\tilde{z}_{i-1}$ for each of the $p$ variables, then form the quadratic Hotelling $T^2_i = \tilde{z}_i'\Sigma^{-1}_{\tilde{z}}\tilde{z}_i$ where the covariance used in the $T^2$ is the covariance of the EWMA vector. We will call their multivariate EWMA scheme as EWMA-M, in accordance with the order of these two steps where the EWMA is built before the Hotelling $T^2$. (An alternative multivariate EWMA scheme is M-
EWMA Pan (2005), which builds the Hotelling $T^2$ of the variables before the formation of the EWMA of the $T^2$'s. Lui (1996) presented an improvement for EWMA-M. Runger and Prahu (1996) used Markov chain analysis to calculate the ARL for EWMA-M and Prahu and Runger (1997) discussed the design of the same scheme. However, all these studies assumed the processes to be serially independent.

Others chose to study the usefulness of multivariate EWMA methods as well. Stoumbus and Sullivan investigated the effects of non-normality on the performance of the multivariate EWMA control chart, and its special case, the Hotelling’s Chi-Squared control chart when applied to individual observations. The purpose in this case was to monitor the mean vector of a multivariate process variable. Khoo studied the sensitivity of multivariate EWMA control charts under other circumstances. In addition, Lee and Khoo (2006) explored a method for optimally designing multivariate EWMA charts based on the measures of average run length (ARL) and median run length (MRL).

Another approach to the multivariate EWMA charts, M-EWMA, was previously proposed by Pan (2005). The M-EWMA scheme builds the Hotelling $T^2$ of the variables before the formation of the EWMA of the $T^2$'s. Specifically, for $p$-dimensional multivariate normal processes $\bar{x}_i \sim N(0, \Sigma)$ at the $i^{th}$ observation, the Hotelling $T_i^2 = \bar{x}_i' \Sigma^{-1} \bar{x}_i$ is built at first. Then, the EWMA of the $T_i^2$'s, denoted as $Q_i$, is $Q_i = rT_i^2 + (1-r)Q_{i-1}$, according to the order of construction steps is the statistic of M-EWMA chart. Pan (2005) used integral equation method to compute the ARL's of M-EWMA charts for in-control and out-of-control situations without the presence of serial correlation. Both EWMA-M and M-EWMA are multivariate EWMA schemes.
The above schemes have a common problem, that is, they cannot be directly employed when the processes are serially correlated. An indirect way to apply the multivariate EWMA schemes for serially correlated processes is to adopt Alwan and Roberts’ (1988) approach. They suggest estimating the residuals, i.e., one-step-ahead forecasting errors, of the autocorrelated process. In turn, they apply traditional control charts for the residuals. Extending this approach to multivariate cases, one can apply the above EWMA-M or M-EWMA scheme to the residuals of the serially correlated multivariate processes, until the processes are modeled properly and the initial number of observations is sufficiently large and the residuals are asymptotically independent over time. At this point, we determine the sensitivity of these approaches to changes in process parameters in the presence of serial correlation. Since the process parameters are usually unknown, the appropriate estimation and use of the covariance matrix is vital for correct execution of multivariate EWMA. This may occur if the direct sample variance is a biased estimate of the population variance for a serially correlated process.

We will in the next section, consider variation in the chart statistic determined by EWMA-M methods for the absence of serial correlation. The sensitivity measure will then aid the choice of the optimal EWMA weighting factor and we will better understand the effects of serial correlation in the EWMA-M structure. Later, we examine the situation for the presence of serial correlation.

**EWMA-M for Serially Independent Processes**

For a $p$-dimensional multivariate i.i.d. processes $\bar{x}_i = (x_1, x_2, \cdots, x_p)'$ at time point $i$, constructing the EWMA quantities based on the previous observations, we have

$$\bar{z}_i = r\bar{x}_i + (1-r)\bar{z}_{i-1}, \quad i=1, 2 \ldots$$
Without losing generality the mean vector of $\bar{x}_i$ is set at zero. We then construct the quadratic Hotelling $T^2$ of $\bar{z}_i$ as the chart statistic:

$$T_i^2 = \bar{z}_i \Sigma^{-1} \bar{z}_i$$  \quad (2)$$

where the covariance used in the $T^2$ is the covariance of $\bar{z}_i$, and $r$ was chosen as a scalar weighting parameter. It seems obvious that under the assumption of i.i.d. normal distribution $\bar{x}_i \sim N(\bar{0}, \Sigma)$ with known parameters, the EWMA-M statistic $T_i^2 = \bar{z}_i \Sigma^{-1} \bar{z}_i$ follows chi-square distribution with order of $p$, because each $\bar{z}_i$ is still normal. However, omitted here is that initial value of the EWMA, $\bar{z}_0$. As $\bar{z}_0$ has to be set as a certain value (say, 0), this may make the $T_i^2$ a little different from the exact $\chi^2(p)$.

To investigate the details of the distribution, we denote $\bar{x} = [\bar{x}_i, \bar{x}_{i-1}, \ldots, \bar{x}_1]'$ be an ip×1 column vector, $\bar{r}$ be a 1×i row vector $\bar{r} = [r, r(1-r), \ldots, r(1-r)^{i-1}]$, and $I_p$ are a $p \times p$ unity matrix. For independent observations, we have $\bar{x} \sim N(0,V)$ with $V = (I_i \otimes \Sigma)$. Setting the initial EWMA vector $\bar{z}_0 = \bar{0}$, we have,

$$\bar{z}_i = (\bar{r} \otimes I_p) \bar{x}$$  \quad (3)$$

And $\Sigma_{z_i} = E(\bar{z}_i \bar{z}_i') = (\bar{r} \otimes I_p) E(\bar{x} \bar{x}') (\bar{r} \otimes I_p)'$  

$$= (\bar{r} \otimes I_p)(I_i \otimes \Sigma)(\bar{r} \otimes I_p)'$$  

$$= (\bar{r} \bar{r}') \otimes \Sigma = \frac{r}{2-r} (1-(1-r)^{2i}) \Sigma$$  \quad (4)$$

where $E(\bar{x} \bar{x}') = (I_i \otimes \Sigma)$ because of the independence for observations at different time points ($I_i$ is a $i \times i$ unity matrix), and that $(\bar{r} \bar{r}')$ is scalar. The result in (4) is the same as given in LWCR. Letting $R_r$ be a normalized i×i matrix that $R_r = (\bar{r} \bar{r}) (\bar{r} \bar{r})^{-1}$, the EWMA-M chart statistic is
\[ T_i^2 = \hat{z}_i^{'} \Sigma_{z_i}^{-1} \hat{z}_i = \hat{x}^{'} (\tilde{r} \otimes I_p) \Sigma_{z_i}^{-1} (\tilde{r} \otimes I_p) \tilde{x} \]
\[ = \hat{x}^{'} (\tilde{r} \otimes I_p) ((\tilde{r} \tilde{r}^{'} )^{-1} \otimes \Sigma^{-1}) (\tilde{r} \otimes I_p) \tilde{x} \]
\[ = \hat{x}^{'} (R_r \otimes \Sigma^{-1}) \tilde{x}. \quad (5) \]

We derive the distribution of (5) by applying the theorems on quadratic form discussed in Box (1954) and others (see Appendix 1). According to Appendix 1, the EWMA-M in (5) is approximately distributed as \( T_i^2 \sim \chi^2 (\rho) \) which was expected. Thus, the distribution of (5) is irrelevant to the exponential weighting factor \( r \). This result is consistent with the reported simulation results in LWCR. Therefore, the in-control ARL's and the control limits in their table 1 are the same for varying values of \( r \).

In addition, LWCR noted that the EWMA weighting parameter is a diagonal matrix with different elements, although they only reported the simulation results for scalar \( r \). If the weights are different for different variables, the EWMA weighting parameter is a matrix

\[
W = \begin{pmatrix}
  r_1 \\
  r_2 \\
  \vdots \\
  r_p
\end{pmatrix}
\quad (6)
\]

We define a new \( p \times ip \) matrix, \( R = [W, W(I_p - W), \cdots, W(I_p - W)^{i-1}] \), to replace the previous Kronecker product \( (\tilde{r} \otimes I_p) \) in (3), so that (3) becomes \( \tilde{z}_i = R \tilde{x} \). Then, from the diagonal property of matrix \( RR' \), it is not difficult to see that

\[
\Sigma_{z_i} = R(I \otimes \Sigma)R' = (RR')\Sigma 
\quad (7)
\]

\[
T_i^2 = \hat{z}_i^{'} \Sigma_{z_i}^{-1} \hat{z}_i = \hat{x}^{'} R' \Sigma^{-1} (RR')^{-1} R \tilde{x} 
\quad (8)
\]
If the $r_1, r_2, \ldots, r_p$ in $W$ is the same $r$, (7) and (8) reduce to (4) and (5) respectively.

Appendix 1 indicates (8) follows $\chi^2(p)$, a chi-square determined by the number of dimensions of the system, and irrelevant to the EWMA weighting parameters for the variables. Since the scalar weight is just a special case of the matrix weight, (3) is a special case of (7).

We note that the result $T_i^2 \sim \chi^2(p)$ is based on the approximation that it has the same first two moments as (A-1). The exact in-control distribution of the chart statistic, (A-1), is determined by the eigenvalues of $U$. From the structure of $U$ in (A-3), it is conceivable that the eigenvalues of $U$ depend on $r$ (or $W$) and $i$. (Later we verify this by computer aided computation.) Therefore, the weighting parameters and the time length from the initial point influence the higher moments of the chart statistic and make it differ from $\chi^2(p)$. MacGregor and Harris (1993), when studying exponentially weighted moving square (EWMS), pointed out that the approximation is better for large $i$, i.e., the greater the distance from the initial time point the better.

**Shift Effects and the Measure of Sensitivity - Mean Shift**

If a mean shift $\ddot{\delta}_0$ occurs at $i_0 = i - k$, $0 \leq k \leq i$, then we write $\bar{x} = \bar{u} + \ddot{\delta}$, where $\bar{u}$ is denoted as the in-control processes and $\ddot{\delta}$ is a $i \times 1$ column vector that has $\delta_0$ in the first $kp$ elements and zeros in the last $i_0p$ elements. Taking the scalar $r$ case as example (the case of (6) can be discussed in the same way), the statistic of (5) becomes

$$T_i^2 = \bar{u}'(R_r \otimes \Sigma^{-1})\bar{u} + 2\tilde{\delta}'(R_r \otimes \Sigma^{-1})\bar{u} + \tilde{\delta}'(R_r \otimes \Sigma^{-1})\tilde{\dot{\delta}}.$$  \hspace{1cm} (9)
This is a noncentral chi-square distribution. The first item in (9) is the in-control chi-square (note the mean vector of in-control processes $\bar{u}_i$ is assumed to be zeros without losing generality). The second item is a normal distribution with mean of zero and variance of $4\delta^2 (R_r \otimes \Sigma^{-1})\bar{\delta}$. The second and the third items are all determined by the value of the third item which is a constant scalar. Defining the third item as $\Delta$,

$$
\Delta = \bar{\delta}^t (R_r \otimes \Sigma^{-1}) \bar{\delta} = \bar{\delta}_0^t (R_{rk} \otimes \Sigma^{-1}) \bar{\delta}_0
$$

(10)

where $R_{rk}$, the upper-left $k \times k$ partition matrix is $R_r$, partitioned according to shift occurrence time $i_0$. Obviously, (10) is relevant to $r$ through $R_{rk}$. Therefore, the choice of the value of the exponential weight $r$ in EWMA-M does matter for the out-of-control distributions of the chart statistic.

By reexamining (10), it is actually $\delta_0^2$ times the sum of all the elements of $R_{rk} \otimes \Sigma^{-1}$. Hence, it is not difficult to show that (10) is also $\bar{\delta}_0^t \Sigma^{-1} \bar{\delta}_0'$ times the sum of all the elements of $R_{rk}$. Denoting both $\Delta_0 = \bar{\delta}_0^t \Sigma^{-1} \bar{\delta}_0'$ and the sum of all the elements of $R_{rk}$ as $S_{rk}$, we have $\Delta = S_{rk} \Delta_0$. Note $\Delta_0^{1/2} = (\bar{\delta}_0^t \Sigma^{-1} \bar{\delta}_0')^{1/2}$ is only the noncentrality parameter in LWCR. $\Delta_0$ (or $\Delta_0^{1/2}$) measures the size of the shift in the process itself instead of incorporating the control chart, while the measure $\Delta$ is chart-specific. It can be shown that the $S_{rk}$ is

$$
S_{rk} = \frac{(2 - r)(1 - (1 - r)^k)^2}{r(1 - (1 - r)^{2k})}
$$

(11)

Since the mean of (5) is $\sqrt{p}$ and the mean of (9) can be viewed as $\Delta + \sqrt{p}$, the difference of them, $\Delta$, can be viewed as the measure for the difference in distribution.
Comparing this difference with the in-control dispersion measured by the in-control standard deviation, $\sqrt{2p}$, we can design the sensitivity ratio

$$\eta_i = \frac{\Delta}{\sqrt{2p}} = \frac{S_{rk}\Delta_0}{\sqrt{2p}}$$

(12)

to measure the performance of the control chart for detecting mean shift. As $S_{rk}$ is the only factor depending on $r$, we choose $r$ for the optimal sensitivity through $S_{rk}$. Figure 1 shows how the $S_{rk}$ changes with $r$ for different occurrence time of the shift. For $k=1$, indicating a recent shift, exponential weighting will lower the sensitivity. Later, when $k=2$, weighting at $r=0.5$ gives the best sensitivity. For $k$ is 3, $r$ should be chosen at about 0.3. For $k$ is around 5, $r$ should be 0.2. For $k$ is about 10, $r$ should be 0.1. For $k$ is as large as 30, $r$ should be as small as 0.4. In general, if the soon detection after the shift is wanted (say less than 5 time intervals), larger $r$ is needed. This is consistent with the result based on ARL criteria in LCWR and Prabhu and Runger (1997).

**Shift Effects and the Measure of Sensitivity - Dispersion Shift**

A shift in the dispersion parameter is a change in the process covariance matrix. Yeh, Huwang, and Wu (2004) studied a likelihood-ratio-based EWMA approach for multivariate variability. We now consider the EWMA-M chart by building the Hotelling $T^2$ on the in-control covariance matrix. For out-of-control processes, the theoretical distribution which is the basis for constructing the chart is not clear. Moreover, if the mean of $\bar{x}_i$ is still zero, the theorems on quadratic form (see Appendix1) are directly applicable. Suppose a shift from $\Sigma$ to $\Sigma_1$ occurs at $t_0 = i - k$, $(0 \leq k \leq i)$. Since the covariance matrix is positively definite, there exists a matrix $C$ so that $\bar{x}_i = C\bar{u}_i$ and
\[ \Sigma_1 = C \Sigma C' \] after the time point \( i_0 \). The chart statistic is still expressed in the form of (4) and (5) or (7) and (8) except \( \bar{x} \) is actually shifted. Since the covariance of \( \bar{x} \) is

\[ E(\bar{x}\bar{x}') = \begin{pmatrix} I_k \otimes \Sigma_1 & \Sigma \otimes I_k \otimes I_k \\ I_k \otimes \Sigma & I_k \otimes \Sigma \otimes I_k \end{pmatrix} . \] (13)

Referring to Appendix 1, the matrix \( U \) is (for common weight \( r \), i.e. (4) and (5))

\[ U = E(\bar{x}\bar{x}') (R_r \otimes \Sigma^{-1}) = \begin{pmatrix} R_{rk} \otimes \Sigma \otimes \Sigma^{-1} & R_{rk} \otimes I_p \\ R_{rk} \otimes I_p & R_{rk} \otimes I_p \end{pmatrix} \] (14)

where \( R_{rk}, R_{ri}, R_{rkj}, \) and \( R_{rkj} \) are partition matrices of \( R_r \) which is partitioned according to shift occurrence time \( i_0 \). The eigenvalues of \( U \) depends on \( R_{rk}, R_{ri}, \) and \( \Sigma_1^{-1}\Sigma \), and the eigenvalues of \( UU \) depends on all the four partitioned matrices of \( R_r \).

We can easily verify through simulation that the eigenvalues of the partitioned \( R_r \) does differ for varying values of \( r \). Therefore, the distribution of the chart statistic is still chi-square but with parameter shift that depend on the exponential weight \( r \) (i.e., \( g \) and \( v \) will be dependent on \( r \) instead of being 1 and \( p \)). The departure from the in-control statistic’s distribution determines how quickly on average the control chart detects the shift.

To design a sensitivity measure for the detection of the out-of-control status due to dispersion change, we suggest the use of the ratio between the variance for the out-of-control and the variance for the in-control chart statistics. Since the in-control chart statistic is \( \chi^2(p) \), its variance is \( 2p \). The variance of the out-of-control chart statistic is \( 2\text{trace}(UU) \), where \( U \) is given in (14). We have the sensitivity

\[ \eta_2 = \frac{\text{trace}(UU)}{p} \] (15)
Alternatively, we use the ratio of the standard deviations, which is the square root of (15), to measure sensitivity.

To show the effect of $r$ on sensitivity $\eta_2$, we consider some particular cases that the 3-dimensional processes’ covariance matrix shifts from $\Sigma$ to $\Sigma_1$ at time point $i_0 = i - k$, ($0 \leq k \leq i$). One of the cases is the dispersion shift is in such a way that

$$
\Sigma_1\Sigma^{-1} = \begin{bmatrix}
1 & 0.5 & -1.5 \\
0.5 & 2 & 0.75 \\
-1.5 & 0.75 & 3
\end{bmatrix}
$$

Substituting (16) in (14), we calculate $\eta_2$ for different $r$ and $k$. The result, shown in Figure 2, indicates that larger $r$ yields higher sensitivity $\eta_2$. We also examined several cases of $\Sigma_1\Sigma^{-1}$. As long as the variance for each process increases (i.e. the diagonal elements of $\Sigma_1\Sigma^{-1}$ are great than one), larger $r$ is preferred. This indicates that the current sample information is more important. If at least one of the process variances decreases rather than increases, smaller $r$ is preferred (the decrease in $\eta_2$ is monotone when $r$ increases). In general, if not-very-small $k$ is allowed ($k>10$), $r$ values greater than 0.2 do not make much difference in $\eta_2$. In other words, larger $r$ over 0.2 only gives better $\eta_2$ for $k<10$, i.e. when the shift just happened not long ago.

**EWMA-M and Serially Correlated Processes**

When the above EWMA-M scheme is applied to serially correlated processes, we note the following. First we focus on estimation, since in practice the process parameters are usually unknown. Second, we more completely comprehend the impact of serial correlation in the processes on the control charts.
For estimation, we note that one cannot simply obtain an estimate of a covariance matrix even if is estimated correctly, and, in turn construct the EWMA-M statistic and set control limits according to ARL.

The estimation issue is related to difficulties concerning the distribution. From the previous discussion, the EWMA-M statistic $T_i^2 = \bar{z}_i'\Sigma_z^{-1}\bar{z}_i$ is $\chi^2(p)$ when the $p$-dimensional processes are serially correlated as long as $\Sigma_z$ is known. This can be also seen from $\bar{z}_i = (\bar{\rho} \otimes I_p)\bar{x}$, $\Sigma_z = (\bar{\rho} \otimes I_p)E(\bar{\bar{x}}')(\bar{\rho}' \otimes I_p)$ and

$$T_i^2 = \bar{z}_i'\Sigma_z^{-1}\bar{z}_i$$

$$= \bar{x}'(\bar{\rho}' \otimes I_p)(\bar{\rho} \otimes I_p)E(\bar{\bar{x}}')(\bar{\rho} \otimes I_p)^{-1}(\bar{\rho}' \otimes I_p)\bar{x} \sim \chi^2(p) \quad (17)$$

In the derivation, the serial correlation in $\bar{x} = [\bar{x}_i', \bar{x}_{i-1}', \ldots, \bar{x}_1']'$ only affects $E(\bar{\bar{x}}')$, this is finally deleted when calculating $U$ for quadratic form. $\Sigma_z$ is not directly known and it is obtained from the process parameter,

$$E(\bar{\bar{x}}') = \Gamma_i = \begin{bmatrix} \Gamma(0) & \Gamma(1)' & \cdots & \Gamma(i-1)' \\
\Gamma(1) & \Gamma(0) & \cdots & \Gamma(i-2)' \\
\vdots & \vdots & \ddots & \vdots \\
\Gamma(i-1) & \Gamma(i-2) & \cdots & \Gamma(0) \end{bmatrix}. \quad (18)$$

where the $\Gamma(l)$s are the autocovariance matrices of the processes for lag $l$. When serial correlation exists, (4) or (7) are not available anymore, because the $\Gamma(l)$s for $l \neq 0$ are not zero.
When the process parameters are unknown, the estimation can be through either of the following two ways in the initial phase of constructing control charts. First, $\Gamma(l)$ is estimated through to form $\Gamma_i$. Second, construct $\bar{z}_i = r\bar{x}_i + (1-r)\bar{z}_{i-1}$ from the initial phase of construction of $\bar{x}_i$ data from observation 1 to observation $i$, and calculate the sample variance-covariance based on the values of $\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_i$ to estimate $\Sigma_{\bar{z}}$.

A second approach called the “Z approach” appears promising $\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_i$ are not random, since limitation that the initial value is certain (for example, $\bar{z}_0 = 0$). Alternatively, the autocovariance approach is promising when the process model is moving-average. Specifically, if the underlying process is a vector moving average of order $q$ (denoted as VMA(q)), the autocovariance matrices become zero after $q$ lags. That is, $\Gamma(l) = 0$ for $l > q$, and, in turn,

$$
\Gamma_i = \begin{pmatrix}
\Gamma(0) & \Gamma(1)' & \cdots & \Gamma(q)' & \cdots & 0 \\
\Gamma(1) & \Gamma(0) & \Gamma(1)' & \ddots & \vdots \\
\vdots & \Gamma(1) & \ddots & \ddots & \ddots \\
\Gamma(q) & \ddots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & \Gamma(q) & \cdots & \Gamma(1) & \Gamma(0)
\end{pmatrix}
$$

(19)

If the order $q$ is small, we need only a few estimates, $\hat{\Gamma}(0), \hat{\Gamma}(1), \ldots, \hat{\Gamma}(q)$. For example, VMA(1), we have

$$
\Sigma_{\bar{z}_i} = (\bar{r} \otimes I_n) \Gamma_i (\bar{r} \otimes I_n)
$$

$$
= \frac{r}{2-r} (1-(1-r)^{2i}) \Gamma(0) + \frac{r(1-r)}{2-r} (1-(1-r)^{2i-2}) (\Gamma(1) + \Gamma(1)')
$$

(20)

We need only $\hat{\Gamma}(0)$ and $\hat{\Gamma}(1)$ to calculate from the initial phase of quality control chart construction.
The autocovariance approach is then suitable for VMA processes with known coefficient matrices $\Theta_k$ and error term covariance matrix $\Omega$, and $\Gamma(l)$ is calculated by

$$\Gamma(l) = \sum_{h=0}^{q-l} \Theta_h \Omega \Theta_{h+l}'.$$

For processes with autoregressive terms, $\Gamma(l)$ is non zero even when the lag $l$ is large, (i.e., $\Gamma(l) = \sum_{j=1}^{p} \Gamma(l-j) \Phi_j'$, see Reinsel, 1993). Hence the autocovariance approach is not suitable. (Note, Appendix 2 gives $\Sigma_{z_i}$ for VMA(q) of higher orders of q.)

**The presence of serial correlation in EWMA-M chart**

To consider the choice of the weighting factor $r$ or $R$ for EWMA-M on serially correlated process, we note that the EWMA-M statistic is serially correlated by the weighting factor $r$ or $R$. Thus, running the chart on serially correlated processes is equivalent to varying the weighting factor’s value(s). For example, consider a simplified example of sensitivity for mean shift where the process is a VMA(1) with a scalar coefficient $\theta$, and the weighting factor $r$ is also a scalar. In this case, it is not difficult to see (in Appendix 3) that the original vector $\tilde{r} = [r, r(1-r), \cdots, r(1-r)^{i-1}]$ that forms $\tilde{z}_i = (\tilde{r} \otimes I_p) \tilde{x}$ and $R_r = (\tilde{r}^{'} \tilde{r})(\tilde{r}^{'} \tilde{r})^{-1}$ in (5) becomes $\theta$–related,

$$\tilde{r}(\theta) = [r, r(1-r) + r\theta, \cdots, r(1-r)^{i-1} + r\theta(1-r)^{i-2}].$$

So $R_r$ becomes $\theta$–related $R_r(\theta)$. The other steps to derive EWMA-M do not change. The in-control chart statistic is still $\chi^2(p)$. Hence, the sensitivity $\eta_1$ in (12) becomes

$$\eta_1(\theta) = \frac{S_r(\theta) \Delta_0}{\sqrt{2p}}$$

(22)
where the $S_{rk}(\theta)$ is the sum of all the elements of the $k$-order upper-left partition matrix of $R_r(\theta)$.

Figure 3 shows how the $\theta$ affects $S_{rk}(\theta)$ at different $r$ (hence, how $\theta$ affects $\eta_i$). In Figure 3, the plots drawn in dots correspond to the case of $\theta$ being zero, which can be a reference and compared with the plots at other $\theta$. We see from the figures that 1) larger positive $\theta$ yields larger sensitivity ($S_{rk}(\theta)$, actually, same as bellow), and negative $\theta$ with smaller absolute value yields lower sensitivity. This is generally true regardless of whether the chart is at starting state (small $i$) or at steady state (large $i$), or whether the shift occurred long ago (large $k$) or just occurred (small $k$). 2) A positive $\theta$ has relatively less impact on the sensitivity than a negative $\theta$. Hence, the positive $\theta$ changes the sensitivity in smaller amounts than negative $\theta$. This is especially true for the steady state. Also, 3) the impact of $\theta$ is larger for the starting state (small $i$) than for the steady state (large $i$). It is large for shifts occurred recently (small $k$) than for earlier times (large $k$). Last, 4) positive $\theta$ requires the use of larger $r$ to achieve the optimal sensitivity, while negative $\theta$ requires to use a smaller $r$ to achieve the optimal sensitivity. There is little impact when early time periods (see Figure 3(d)). Finally, the findings in 4) may allow us to adjust the optimal weighting factor for the EWMA-M chart if the information on the serial correlation in the process is available.

**Summary**

In this paper, we studied multivariate EWMA control charts constructed by the method of EWMA-M. Applying Box quadratic form, we investigated in detail the distribution of the chart statistic. Based on the distribution of the chart statistics for in-
control and out-of-control situations, we proposed sensitivity ratios as a measure of the effects of the mean shift and dispersion shift. Using this sensitivity measure, we designed the optimal exponential weighting factor, which is consistent to results reported before.

Although ARL is the usual measure for SPC chart performance, it is by no means the only criterion, and it has shortcomings. Our proposed sensitivity measure has certain advantages. It is directly derived from the distributions of the chart statistic, hence, it is not constrained to where the control limit is located. This makes the sensitivity measure have a broader range to fit varying situations.

We discussed the EWMA-M chart on processes in presence of serial correlation. Based on the in depth knowledge on the distribution of the chart statistic, we suggest a special way of constructing the variance-covariance matrix for the EWMA-M scheme on multivariate MA processes. Using the sensitivity measure, we investigated the role of serial correlation of the process in the structure of the chart statistic, and its impact on the sensitivity performance for a special process pattern (VMA (1)). This allows us to consider adjusting the optimal exponential weighting factor according to the information on serial correlation.
APPENDIX 1. The distribution of EMWA-M

To derive the distribution of EWMA-M in (3), we apply the theorems quadratic form and its approximation in Box (1954). According to Box (1954), if \( \bar{x} = (x_1, \cdots x_m)' \) is multivariate normal \( N(0, V) \) with order of \( m \), then the quantity \( Q = \bar{x}'A\bar{x} \) with rank \( i \leq m \) is distributed as

\[
Q_i = \sum_{j=1}^{i} \lambda_j x_j^2(1)
\]

(A-1)

where \( \lambda_j \)'s are the latent roots of \( U = VA \). An approximation to this distribution that gives the same first two moments is

\[
Q_i \sim g \chi^2(v),
\]

(A-2a)

where

\[
g = \sum_{j=1}^{i} \lambda_j^2 \left/ \sum_{j=1}^{i} \lambda_j \right. = \text{trace}(UU)/\text{trace}(U)
\]

and

\[
v = \left( \sum_{j=1}^{i} \lambda_j \right)^2 \left/ \sum_{j=1}^{i} \lambda_j^2 \right. = [\text{trace}(U)]^2 / \text{trace}(UU).
\]

(A-2b)

An application of this quadratic form in univariate SPC can be found in MacGregor and Harris (1993), where they examined exponentially weighted moving square (EWMS) and exponentially weighted moving variance (EWMV).

Applying the above quadratic form into (3), we treat \( U = (I_i \otimes \Sigma)(R_r \otimes \Sigma^{-1}) = R_r \otimes I_p \), and \( UU = (R_r \otimes I_p)(R_r \otimes I_p) = R_r^2 \otimes I_p \). Since \( R_r^2 = \bar{r}'\bar{r}r'\bar{r}^{-1} = \bar{r}'(\bar{r}'\bar{r})^{-1} = R_r \), then \( UU = R_r \otimes I_p \). Since \( \text{trace}(R_r) = 1 \) and \( \text{trace}(I_p) = p \), and since the eigenvalues of a Kronecker product are the products of the eigenvalues of the respective matrices that form the Kronecker product (also, note that
the product of eigenvalues of a matrix is the trace of the matrix), we have

\[ \text{trace}(U) = \text{trace}(R_r) = p \quad \text{and} \quad \text{trace}(UU) = p. \]

From (A-2b), it is easy to see that \( g = 1 \) and \( v = p \). Therefore, the quadratic form in (3) follows \( T_i^2 \sim \chi^2(p) \).

When the exponential weighting factor is matrix given in (6), We can use the \( p \times ip \) matrix \( R = [W, W(I_p - W), \ldots, W(I_p - W)^{i-1}] \), and apply the Box quadratic form with (7) and (8). Then,

\[
U = (I \otimes \Sigma)(R’ \Sigma^{-1}(RR')^{-1} R) = R'(RR')^{-1} R \quad \text{(A-3)}
\]

\[
UU = (R'(RR')^{-1} R)(R'(RR')^{-1} R) = R'(RR')^{-1} R = U \quad \text{(A-4)}
\]

Since

\[
\text{trace}(R'(RR')^{-1} R) = \text{trace}(RR'(RR')^{-1}) = \text{trace}(I_p) = p, \quad \text{(A1-5)}
\]

we have again, \( \text{trace}(U) = \text{trace}(UU) = p \), hence \( g = 1 \) and \( v = p \). The chart statistic is then still \( T_i^2 \sim \chi^2(p) \), according to (A-2).

**APPENDIX 2. The Covariance Matrix of the EWMA Vector**

Given the form of \( R_r \) (again, with scalar \( r \)), following the algorithm rules of Kronecker matrix, it is easy to get the results of \( \Sigma_{z_i} \) results for higher VNA(q) orders. For VMA(2) processes the \( \Sigma_{z_i} \) is,

\[
\Sigma_{z_i} = \frac{r}{2 - r} (1 - (1 - r)^{2i}) \Gamma(0) + \frac{r(1 - r)}{2 - r} (1 - (1 - r)^{2i-1}) (\Gamma(1) + \Gamma(1)') + \]

\[
+ \frac{r(1 - r)^2}{2 - r} (1 - (1 - r)^{2i-2}) (\Gamma(2) + \Gamma(2)'). \quad \text{(A2-1)}
\]
and for VMA(3),

\[
\Sigma_{z_i} = \frac{r}{2-r}(1-(1-r)^2)\Gamma(0) + \frac{r(1-r)}{2-r}(1-(1-r)^2)\Gamma(1) + \Gamma(1)' + \frac{r(1-r)^2}{2-r}(1-(1-r)^2)\Gamma(2) + \Gamma(2)' + \frac{r(1-r)^3}{2-r}(1-(1-r)^3)\Gamma(3) + \Gamma(3)'.
\]  

(A2-2)

Finally for VMA(q),

\[
\Sigma_{z_i} = \frac{r}{2-r}(1-(1-r)^2)\Gamma(0) + \frac{r(1-r)}{2-r}(1-(1-r)^2)\Gamma(1) + \Gamma(1)' + \frac{r(1-r)^q}{2-r}(1-(1-r)^2)\Gamma(q) + \Gamma(q)'.
\]  

(A2-3)

When the weighting factor is a matrix \(W\), it is also not difficult to derive the relationship between \(\Sigma_{z_i}\) and

**APPENDIX 3. The role of MA(1) term for EWMA.**

As we are talking about a special example that the MA(1) coefficient is a scalar, we treat the multivariate processes just like a univariate process. For the process \(x_t = \varepsilon_t + \theta \varepsilon_{t-1}\), the EWMA is \(z_t = rx_t + (1-r)z_{t-1}\). Hence,

\[
z_t = re_t + r\theta \varepsilon_{t-1} + (1-r)z_{t-1}
\]

\[
= re_t + r\theta \varepsilon_{t-1} + (1-r)(rx_{t-1} + (1-r)z_{t-2})
\]

\[
= re_t + r\theta \varepsilon_{t-1} + (1-r)(re_{t-1} + r\theta \varepsilon_{t-2} + (1-r)z_{t-2})
\]

\[
= re_t + r\theta \varepsilon_{t-1} + r(1-r)\varepsilon_{t-1} + r(1-r)\theta \varepsilon_{t-2} + (1-r)^2 z_{t-2}
\]

\[
= re_t + (r\theta + (1-r))\varepsilon_{t-1} + r(1-r)\theta \varepsilon_{t-2} + (1-r)^2 (rx_{t-2} + (1-r)z_{t-3})
\]

(A3-1)
Therefore, setting $z_0 = 0$, we have

$$z_i = [r, r(1-r) + r\theta, \ldots, r(1-r)^{i-1} + r\theta(1-r)^{i-2}] = [\varepsilon_i, \varepsilon_{i-1}, \ldots, \varepsilon_1]'$$ (A3-2)

where $\varepsilon_i, \varepsilon_{i-1}, \ldots, \varepsilon_1$ are independent. Comparing with the case that serial correlation is absent in the process $x_i$, what is different for the structure of EWMA-M statistic is just in the vector

$$\tilde{r}(\theta) = [r, r(1-r) + r\theta, \ldots, r(1-r)^{i-1} + r\theta(1-r)^{i-2}]$$ (A3-3)

which becomes $\theta$-related (here in (A3-2) we change denotation $t$ in (A3-2) back to $i$ as we used in this paper.)
REFERENCE


Jackson, J.E. (1959) “Quality Control Methods for Several Related Variables,” Technometrics, 1, 4, 359-377


Figure 1 (a) The Sensitivity Factor $S_{rk}$ for Mean Shift, large $i$ ($i=200$)

- '•': $k=1$;  '-•': $k=2$;  '*': $k=3$;  '+': $k=5$;  '-—': $k=7$

Figure 1 (b) The Sensitivity Factor $S_{rk}$ for Mean Shift, large $i$ ($i=200$)

- '•': $k=5$;  '-•': $k=10$;  '*': $k=15$;  '+': $k=20$;  '-—': $k=30$
Figure 1 (c). The Sensitivity Factor $S_{rk}$ for Mean Shift, small $i$ ($i=10$)

- '●': $k=1$
- '-'$: $k=2$
- '★': $k=3$
- '+': $k=5$
- '--': $k=7$

Figure 2. The Sensitivity $\eta_2$ of EWMA-M Chart for Dispersion Shift, ($i=40$)

- '●': $k=5$
- '-'$: $k=10$
- '★': $k=15$
- '+': $k=20$
- '--': $k=30$
Figure 3 (a) $S_{s\theta}(\theta)$ for Sensitivity $\eta_2$ of EWMA-M Chart for Dispersion Shift, $k=2$
Staring State ($i=10$)
'-.': $\theta = -0.8$; '•': $\theta = -0.5$; '+'$: \theta = -0.2$; '●': $\theta = 0$; '---': $\theta = 0.2$; '×': $\theta = 0.5$; '---': $\theta = 0.8$;

Figure 3 (b) $S_{s\theta}(\theta)$ for Sensitivity $\eta_2$ of EWMA-M Chart for Dispersion Shift, $k=5$
Staring State ($i=10$)
'-.': $\theta = -0.8$; '•': $\theta = -0.5$; '+'$: \theta = -0.2$; '●': $\theta = 0$; '---': $\theta = 0.2$; '×': $\theta = 0.5$; '---': $\theta = 0.8$;
Figure 3 (c) $S_{rh}(\theta)$ for Sensitivity $\eta_2$ of EWMA-M Chart for Dispersion Shift, $k=10$

Steady State ($i=100$)

'|': $\theta=-0.8$; '*': $\theta=-0.5$; '+': $\theta=-0.2$; '•': $\theta=0$; '-': $\theta=0.2$; '×': $\theta=0.5$; '- -': $\theta=0.8$;

Figure 3 (d) $S_{rh}(\theta)$ for Sensitivity $\eta_2$ of EWMA-M Chart for Dispersion Shift, $k=30$

Steady State ($i=100$)

'|': $\theta=-0.8$; '*': $\theta=-0.5$; '+': $\theta=-0.2$; '•': $\theta=0$; '-': $\theta=0.2$; '×': $\theta=0.5$; '- -': $\theta=0.8$;
Founded in 1892, the University of Rhode Island is one of eight land, urban, and sea grant universities in the United States. The 1,200-acre rural campus is less than ten miles from Narragansett Bay and highlights its traditions of natural resource, marine and urban related research. There are over 14,000 undergraduate and graduate students enrolled in seven degree-granting colleges representing 48 states and the District of Columbia. More than 500 international students represent 59 different countries. Eighteen percent of the freshman class graduated in the top ten percent of their high school classes. The teaching and research faculty numbers over 600 and the University offers 101 undergraduate programs and 86 advanced degree programs. URI students have received Rhodes, Fulbright, Truman, Goldwater, and Udall scholarships. There are over 80,000 active alumnae.

The University of Rhode Island started to offer undergraduate business administration courses in 1923. In 1962, the MBA program was introduced and the PhD program began in the mid 1980s. The College of Business Administration is accredited by The AACSB International - The Association to Advance Collegiate Schools of Business in 1969. The College of Business enrolls over 1400 undergraduate students and more than 300 graduate students.

**Mission**

Our responsibility is to provide strong academic programs that instill excellence, confidence and strong leadership skills in our graduates. Our aim is to (1) promote critical and independent thinking, (2) foster personal responsibility and (3) develop students whose performance and commitment mark them as leaders contributing to the business community and society. The College will serve as a center for business scholarship, creative research and outreach activities to the citizens and institutions of the State of Rhode Island as well as the regional, national and international communities.

The creation of this working paper series has been funded by an endowment established by William A. Orme, URI College of Business Administration, Class of 1949 and former head of the General Electric Foundation. This working paper series is intended to permit faculty members to obtain feedback on research activities before the research is submitted to academic and professional journals and professional associations for presentations.

An award is presented annually for the most outstanding paper submitted.