Journal of Quality Technology
A Quarterly Journal of Methods, Applications and Related Topics

ISSN: 0022-4065 (Print) 2575-6230 (Online) Journal homepage: https://www.tandfonline.com/loi/ujqt20

Self-Starting Multivariate Control Charts for Location and Scale

Edgard M. Maboudou-Tchao & Douglas M. Hawkins

Self-Starting Multivariate Control Charts for Location and Scale

EDGARD M. MABOUDOU-TCHAO

University of Central Florida, Orlando, Florida 32816

DOUGLAS M. HAWKINS

University of Minnesota, Minneapolis, Minnesota 55455

Multivariate control charts are advisable when monitoring several correlated characteristics. The multivariate exponentially weighted moving average (MEWMA) is ideal for monitoring the mean vector, and the multivariate exponentially weighted moving covariance matrix (MEWMC) detects changes in the covariance matrix. Both charts were established under the assumption that the parameters are known a priori. This is seldom the case, and Phase I data sets are commonly used to estimate the chart's in-control parameter values. Plugging in parameter estimates, however, fundamentally changes the run-length distribution from those assumed in the known-parameter theory and diminishes chart performance, even for large calibration samples. Self-starting methods, which correctly studentize the incoming stream of process readings, provide exact control right from start up. We extend the existing multivariate self-starting methodology to a combination chart for both the mean vector and the covariance matrix. This approach is shown to have good performance.

Key Words: Average Run Length (ARL); Cholesky Decomposition; Multistandardization; Recursive Residual; Regression Adjustment.

Introduction and Existing Work

PROCESS parameters are unknown in most applications and are estimated for use in control charts. This practice is detrimental to chart performance (Jones et al. (2001)) because obtaining good estimates in the multivariate setting is not an easy task, especially in higher dimensions. Another possible solution to the problem is a "self-starting" methodology, which allows us to monitor the process from start up. With that methodology, process monitoring may start without requiring large preliminary Phase I samples. In self-starting charts, successive process readings are used for updating the parameter estimates and simultaneously checking for process stability. The self-starting methodology operates by transforming the incoming stream of process readings into a stream of mutually independent identically distributed data of which the in-control distribution, including its parameter values, is known. In the univariate case, Hawkins (1987) discussed a self-starting cumulative sum (CUSUM) and Quesenberry (1991, 1995, 1997) discussed Shewhart equivalents. Schaffer (1998), Quesenberry (1997), Sullivan and Jones (2002) have proposed a variety of self-starting multivariate charts for monitoring the mean vector. Hawkins and Maboudou-Tchao (2007) proposed a self-starting multivariate exponentially weighted moving average (SSMEWMA) chart for control of the mean. Capizzi and Masarotto (2010) presented a self-starting cumulative score (CUSCORE) control chart for monitoring the mean vector. In addition, some authors apply the self-starting methodology for monitoring both location and scale. Zamba and Hawkins (2009) presented a multivariate change point model for changes in the mean vector and covariance matrix. Zou and Tsung (2010) discussed self-starting features for monitoring the location and scale of a univariate process. In this article, we de-

Dr. Maboudou-Tchao is an Assistant Professor in the Department of Statistics. He is a member of ASQ. His email address is emaboudo@mail.ucf.edu.

Dr. Hawkins is a Professor in the School of Statistics. He is an ASQ Fellow. His email address is dhawkins@umn.edu.

velop a parallel multivariate self-starting methodology for the covariance matrix, which will be used in conjunction with the SSMEWMA to monitor both the mean vector and covariance matrix when the process parameters are unknown.

Monitoring for a Shift in the Mean Vector

Methods for monitoring conventional control charts may be characterized by whether they use individual rational subgroups ("Shewhart-type" charts) or are accumulative techniques that combine information from successive observations. Shewhart-type charts are effective for detecting large shifts, and accumulative methods work better for smaller but persistent shifts. The best known work in multivariate control charting is that of Hotelling (1947), which is a direct multivariate extension of the univariate Shewhart $\overline{X}$ chart.

Many people are attracted to the Shewhart-type chart because of its simplicity. However, it has serious limitations. The chart has no memory—it uses only the current sample and fails to accumulate evidence of change. This means that the chart does not perform well against sustained small changes in the mean vector.

The ability of the T^2-chart to detect smaller shifts is increased by applying it to the mean vectors of rational subgroups of readings. As in the univariate parallels of I and $\overline{X}$ charting, this leads to a trade-off of the additional sensitivity against the increased cost of the additional readings or the time delay in holding individual readings over until the rational subgroup is complete.

Following the accepted wisdom that accumulative methods work better with individual observations than with rational subgroups, our method uses individual process vectors. Let $\mathbf{x}_t \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $t = 1, 2, 3, \ldots$, be the successive process reading vectors. One multivariate control chart that is well suited to detecting persistent small changes is the multivariate exponentially weighted moving averages (MEWMA), due to Lowry et al. (1992), which uses the recursion

$$\mathbf{z}_t = \lambda \mathbf{x}_t + (1 - \lambda)\mathbf{z}_{t-1} \tag{1}$$

with $0 < \lambda \leq 1$ and $\mathbf{z}_0 = \boldsymbol{\mu}$. The quantity plotted on the control chart is

$$M_t = (\mathbf{z}_t - \boldsymbol{\mu})' \boldsymbol{\Sigma}_{\mathbf{z}_t}^{-1} (\mathbf{z}_t - \boldsymbol{\mu}), \tag{2}$$

where the covariance matrix $\boldsymbol{\Sigma}_{\mathbf{z}_t} = [\lambda/(2-\lambda)][1 - (1-\lambda)^{2t}]\boldsymbol{\Sigma}$.

The MEWMA chart gives an out-of-control signal if M_t exceeds some positive constant h, chosen to achieve a specified in-control average run length (IC ARL)

Monitoring for a Shift in the Covariance Matrix

There are two current standard Shewhart-type methodologies for monitoring the covariance matrix $\boldsymbol{\Sigma}$. One, due to Montgomery (2001), uses the generalized variance; the other, due to Alt (1984), uses the generalized likelihood ratio. As usual for Shewhart charts, these are effective for large transient changes in the covariance matrix but are less suitable for smaller persistent shifts.

Recent proposals by Huwang et al. (2006) and Hawkins and Maboudou-Tchao (2008) defined an analogous chart, the multivariate exponentially weighted moving covariance matrix (MEWMC). In the MEWMC, the process readings $\mathbf{x}_t$ are first multistandardized to $\mathbf{u}_t = \boldsymbol{\Sigma}^{-1/2}(\mathbf{x}_t - \boldsymbol{\mu})$.

While in control, the $\mathbf{u}_t$ are N($\mathbf{0}$,$\mathbf{I}_p$). The standardized MEWMC is defined by the recursion

$$\mathbf{S}_t = (1 - \lambda)\mathbf{S}_{t-1} + \lambda \mathbf{u}_t \mathbf{u}_t', \tag{3}$$

where $\mathbf{S}_0 = \mathbf{I}_p$ and λ is the smoothing constant.

Define

$$C_t = \mathrm{tr}(\mathbf{S}_t) - \log|\mathbf{S}_t| - p,$$

where tr is the trace operator and $|\cdot|$ represents the determinant. The MEWMC control chart then consists of plotting C_t against t and signaling a loss of control if $C_t > h$, where h is chosen to achieve a specified in-control ARL.

The Unknown Parameter Case

The control charts described above assume that the in-control parameters are known, but in practice, this is hardly ever the case. Most practitioners use the "plug-in" method, where a large Phase I data set is gathered while the process is believed to be in control. The parameters are then estimated from this Phase I data set, and the estimates substituted for the true parameter values.

There is a considerable body of evidence that this method leads to run-length behavior considerably different from that given by known-parameter theory. The run-length distribution has heavier tails, leading to increased average run length both in- and out-of-control. In the multivariate setting of the T^2 chart, for example, this deleterious effect has been explored by Champ et al. (2005).

A different approach is based on the idea of using the regular process readings for both dynamic parameter estimation and maintaining the control of the process. This is the "self-starting" paradigm, in which each successive observation is used to update the mean and standard deviation of all observations accumulated to date. These evolving estimates are then used to studentize the next process reading for use in a statistical process control (SPC) chart. Hawkins (1987) developed a self-starting CUSUM for univariate normal data and Quesenberry (1991, 1995, 1997) discussed Shewhart equivalents. Quesenberry (1997), Schaffer (1998), and Sullivan and Jones (2002) proposed self-starting multivariate charts. A recent multivariate self-starting approach was proposed by Hawkins and Maboudou-Tchao (2007). This "vector accumulation" in the framework of self-starting multivariate charting introduced a transformation of p-dimensional unknown-parameter vectors into known-parameter vectors of the same dimensionality.

These existing self-starting multivariate charts allow one to monitor the mean vector but not the covariance matrix. In this paper, we propose a new method that will monitor changes in both the mean vector and the covariance matrix using a self-starting formulation.

The Multivariate Self-Starting Approach

From the different self-starting formulations, we focus on the proposal of Hawkins and Maboudou-Tchao (2007). Assume the process reading vectors $\mathbf{x}_t \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with unknown μ and Σ. The methodology has two stages. The first stage transforms the stream $\mathbf{x}_t$ into a sequence of mutually independent p-vectors $\mathbf{r}_t$ with zero mean vector and a covariance matrix that is diagonal, but with unknown variances. The second stage uses the double probability integral transform to transform the vectors $\mathbf{r}_t$ into a sequence of mutually independent $p-$vectors $\mathbf{u}_t$ that follow a multivariate standard normal $N(\mathbf{0}, \mathbf{I}_p)$ distribution, which can then be used in known-parameter control charts. Changes in the $\mathbf{x}_t$ process lead to changes in the $\mathbf{u}_t$ process, so diagnosing the latter points to features of the former.

The definitions of $\mathbf{r}_t$ and $\mathbf{u}_t$ are given in Hawkins and Maboudou-Tchao (2007) and are repeated in the appendix.

Suffice it to note that $\mathbf{r}_t$ are the vectors of recursive residuals of the regression of each variable in $\mathbf{x}_t$ on the variables with lower subscripts and the $\mathbf{u}_t$ are double probability integral transforms of $\mathbf{r}_t$, studentized by the preceding $\mathbf{r}_j$.

In this procedure, $p+1$ initial process reading vectors are required to set up the first nondegenerate estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, and so these initial readings are "lost". The first nontrivial $\mathbf{u}_t$ vector occurs for $t = p + 2$, and the tth data vector gives rise to the $t - p - 1$th nontrivial $\mathbf{u}_t$.

As the in-control period increases, the running mean vector and covariance matrix used in the multistandardization converge to the true mean vector and covariance matrix, so that asymptotically the $\mathbf{u}_t$ approach the known-parameter multistandardized vectors $\boldsymbol{\Sigma}^{-1/2}(\mathbf{x}_t - \boldsymbol{\mu})$.

To implement the charting, starting at $t = p + 2$, define

$$\mathbf{z}_t = (1 - \lambda)\mathbf{z}_{t-1} + \lambda \mathbf{u}_t \tag{5}$$

and

$$\mathbf{S}_t = (1 - \lambda)\mathbf{S}_{t-1} + \lambda \mathbf{u}_t \mathbf{u}_t', \tag{6}$$

where $0 < \lambda < 1$ is the smoothing constant and with the initialization $\mathbf{z}_{p+1} = \mathbf{0}$; $\mathbf{S}_{p+1} = \mathbf{I}_p$. Then compute

$$\begin{aligned} M_t &= \mathbf{z}_t' \boldsymbol{\Sigma}_{\mathbf{z}_t}^{-1} \mathbf{z}_t \\ &= \frac{2 - \lambda}{\lambda[1 - (1 - \lambda)^{2(t-p-1)}]} \|\mathbf{z}_t\|^2 \end{aligned} \tag{7}$$

and

$$C_t = \mathrm{tr}(\mathbf{S}_t) - \log|\mathbf{S}_t| - p, \tag{8}$$

where tr is the trace operator and $|\mathbf{S}_t|$ is the determinant of $\mathbf{S}_t$.

Plot M_t and C_t on two separate charts and signal an out-of-control if $M_t > h_1$ or $C_t > h_2$, where the control limits h_1, h_2 are chosen to achieve a target in-control ARL.

Values of h_1 can be obtained from the tables given in Hawkins and Maboudou-Tchao (2007) and values of h_2 can be obtained from the tables given in Hawkins and Maboudou-Tchao (2008). More extensive tables of both quantities are on the web site www.stat.umn.edu/hawkins.

We call this combination procedure a self-starting multivariate exponentially weighted moving average and moving covariance matrix (SSMEWMAC) chart. The chart starts with the $p+2$nd process reading and continues indefinitely. At each stage while in control, the chart quantity $\mathbf{u}_t$ is exactly $N(\mathbf{0}, \mathbf{I}_p)$, and so M_t

and C_t behave exactly as in the known-parameter setting used in the literature.

Performance Study

The performance of a chart is judged by its out-of-control run lengths. Several factors affect the performance of this combination chart. Some are the same as those affecting the known-parameter charts: the dimension, the in-control ARL, the size of the shift in the mean vector and that of the shift in the $p(p+1)/2$ elements of the covariance matrix, and so need little further exploration.

In addition though, another factor is specific to the unknown-parameter self-starting setting, and this is the length of time the process runs in control before a shift occurs. This time will be referred as the "learning time". As is intuitively clear, a longer initial in-control period will lead to better estimates of the in-control parameters and hence better performance of the self-starter, and, conversely, chart performance will be modest with very short learning times.

A simulation study to explore the impact of some of these factors was conducted. Previous work on the known-parameter setting reported in Hawkins and Maboudou-Tchao (2008) explored the effect of the dimension, the size of the mean shift, and the size of changes in the covariance matrix, and our new simulations explored a more modest suite of these values, along with the new element of the learning time. The number of Monte Carlo simulations for each scenario was 20,000.

To explore the impact of the dimension, we ran the analysis for a relatively small dimension, $p = 5$; a medium dimension, $p = 10$; and a larger dimension, $p = 20$.

We used $\lambda = 0.1$ and, by setting the in-control ARL of the mean and the covariance matrix portions to 500, set the in-control ARL of the combined scheme to 250. From tables on the the Web site, this is given by the settings $h_1 = 17.186$, $h_2 = 1.754$ for $p = 5$, $h_1 = 25.831$, $h_2 = 4.65$ for $p = 10$ and $h_1 = 40.806$, and $h_2 = 15.028$ for $p = 20$.

After an initial in-control period of length τ, the mean vector and/or covariance matrix was shifted, and the ARL to signal monitored. Any simulation run in which there was a signal before time τ was omitted from the calculations. The simulations used a selection of τ values ranging from 10 to 2000.

Performance Under a Mean Shift

If the process shifts from $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to $N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$, the chart behavior is determined by the noncentrality $\delta^2 = (\boldsymbol{\mu}_1 - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu})$ and so the performance following any step change in the mean can be modeled by changing a single component of the mean vector as follows: Set the in-control mean vector to $\mathbf{0}$ and the covariance matrix to the identity matrix. Thereafter, by adding a shift δ to the first component of each vector later than time τ, we can explore all possible step changes in the mean vector.

Figure 1 shows the resulting ARL for the choice $\delta = 1$ and $\delta = 0.5$. In order to show the main features, a logarithmic scale is used on both axes.

As the learning time increases, the SSMEWMAC converges to the known-parameter setting, and so the asymptote of the ARL curve is the ARL of the known-parameter MEWMAC. This chart's out-of-control ARL for this choice of δ would be 13.28 for $p = 5$, 16.68 for $p = 10$, and 21.36 for $p = 20$. These asymptotic values are also shown as horizontal lines in Figure 1. The figure shows that, with a short initial learning period, the chart takes a long time to detect the shift, but improves substantially with increasing τ.

The plots show that the SSMEWMAC effectively reaches its asymptote in some 200 (for $\delta = 1$) to 1000 (for $\delta = 0.5$) readings, doing so faster in smaller than in larger dimension.

Exploring the impact for δ further, we ran more simulations in which we fixed the learning period τ at 50, 200, and 500. Again, we set the in-control mean vector to $\mathbf{0}$ and the covariance matrix to the identity matrix and added a shift δ to the first component of each vector after time τ, varying δ from 0 to 3 in steps of 0.5. Dimensions 5, 10, and 20 were used. The results are shown in Figure 2.

These plots show that the SSMEWMAC reacts quickly to large shifts, even if τ is modest. There is substantial separation between the curves for $\tau = 50$ and those for the higher values, but little separation between $\tau = 200$ and 500. This is true in all three dimensions.

Performance Under a Covariance Matrix Shift

A full study of the chart performance is infeasible, as it would involve arbitrary changes in the $p(p+1)/2$ elements of the covariance matrix. Hence, we sketch just a scenario that changes the covariance matrix

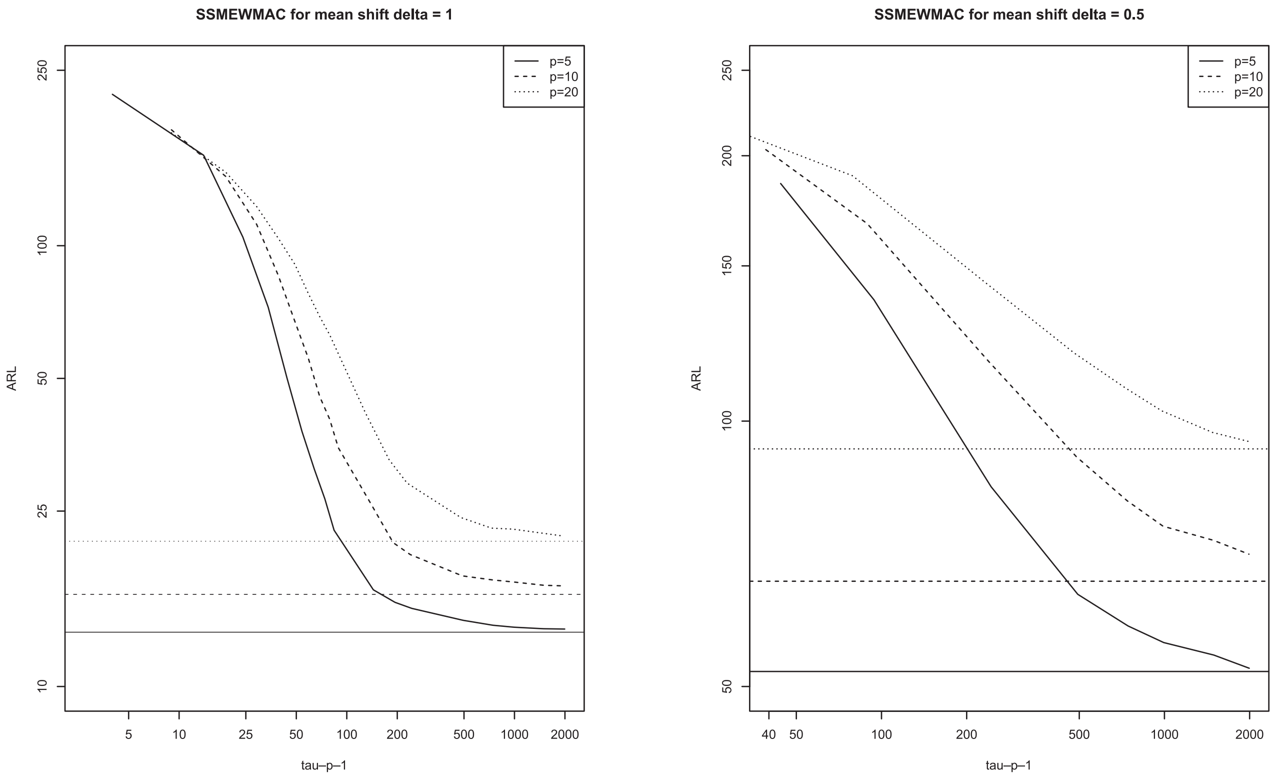

FIGURE 1. SSMEWMAC for Change in Mean. Left: $\delta = 1$, right: $\delta = 0.5$. The horizontal lines represent OOC ARL of the known parameter MEWMAC.

from the identity matrix to a matrix with σ^2 in the (1, 1) position while the other elements remain unchanged, and that leaves the mean vector at $\mathbf{0}$. We look separately at a doubling and at a halving in σ^2. The ARLs for various learning periods τ are shown in Figure 3.

Starting with the variance increase, in the known parameter case, the out-of-control ARL is 41.70 for $p = 5$, 68.55 for $p = 10$, and 110.8 for $p = 20$, giving the asymptotes to which the three graphs will tend as τ increases. A glance at Figure 3 shows that coming close to the known-parameter ARLs requires several thousand readings, an order of magnitude longer learning period than suffices for the mean vector. This is not just a commentary on the performance of the self-starter, but may be a salutary warning on the traditional practice of Phase I studies to estimate the covariance matrix. As the required Phase I sample size for the traditional approach would be at least as large as that needed for the self-starting method to approach its asymptote, this means that the traditional approach requires surprisingly large samples for the effect of the parameter estimation to be negligible.

Turning to variance decreases, the known-parameter out-of-control ARL is 164.6 for $p = 5$, 206.2 for $p = 10$, and 230.2 for $p = 20$. The SSMEWMAC achieves known parameter behavior for a learning time close to 3000. Two striking features of this graph are the large value of the asymptote and the flat initial portion. This implies that, not only is a variance decrease much harder to detect than an equal-sized increase, but a very long initial learning period is required to get even modest shift detectability.

Next, to explore the effect of the size of the variance shift, we fixed τ and varied σ from 0.1 to 10, covering both increases and decreases in variance. The results are shown in Figure 4, which uses a log scale for σ.

The SSMEWMAC reacts quickly to a shift (either increase or decrease) when τ is large except in a neighborhood of $\sigma = 1$. Around $\sigma = 1$, we observe that the OOC ARL exceeds the IC ARL for small variance decreases, meaning that the SS-MEWMAC chart is slightly ARL-biased. This situation was discussed by Pignatiello et al. (1995) and is not surprising, as Hawkins and Maboudou-Tchao

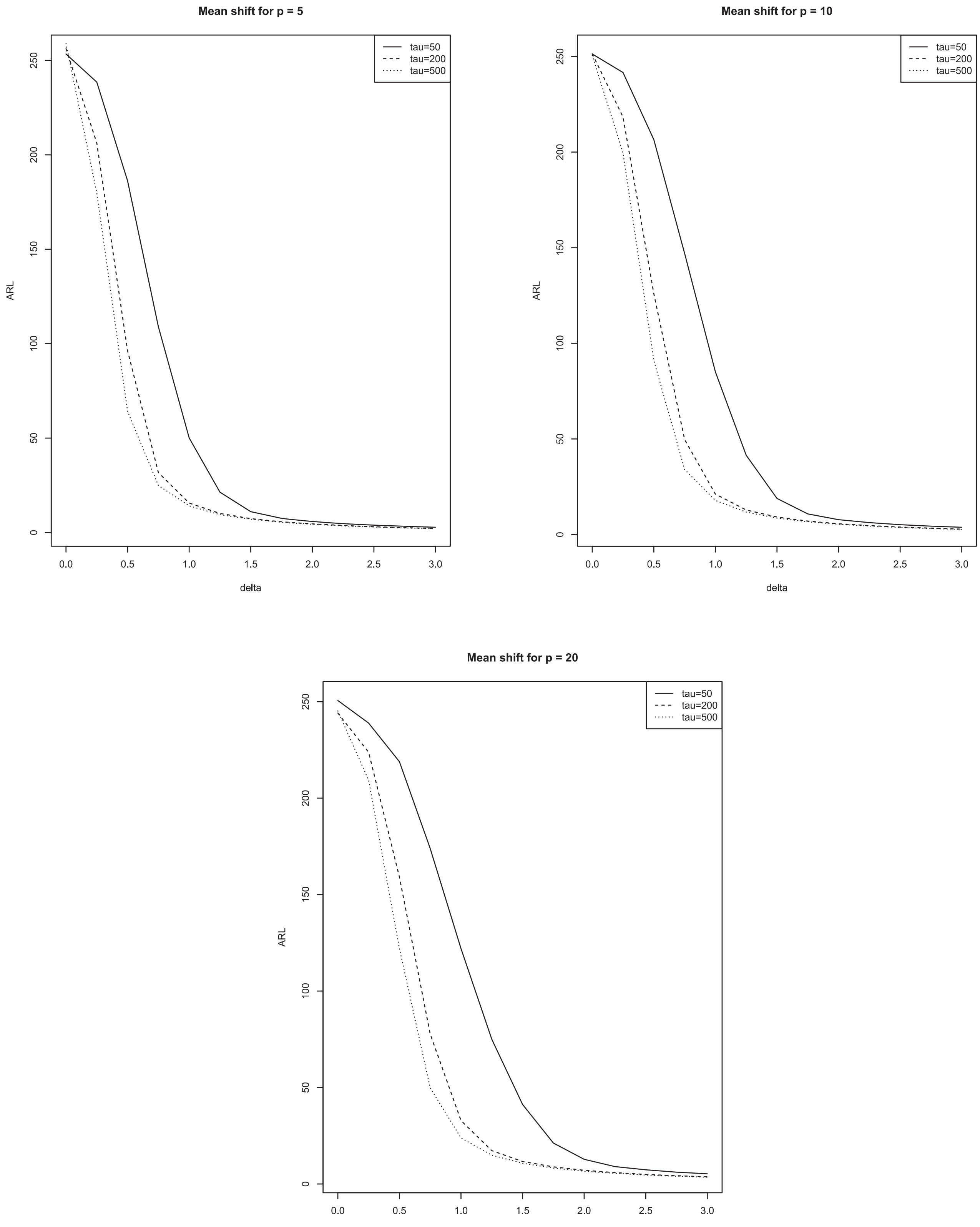

FIGURE 2. SSMEWMAC for Change in Mean. Top left: $p = 5$, top right: $p = 10$, bottom: $p = 20$.

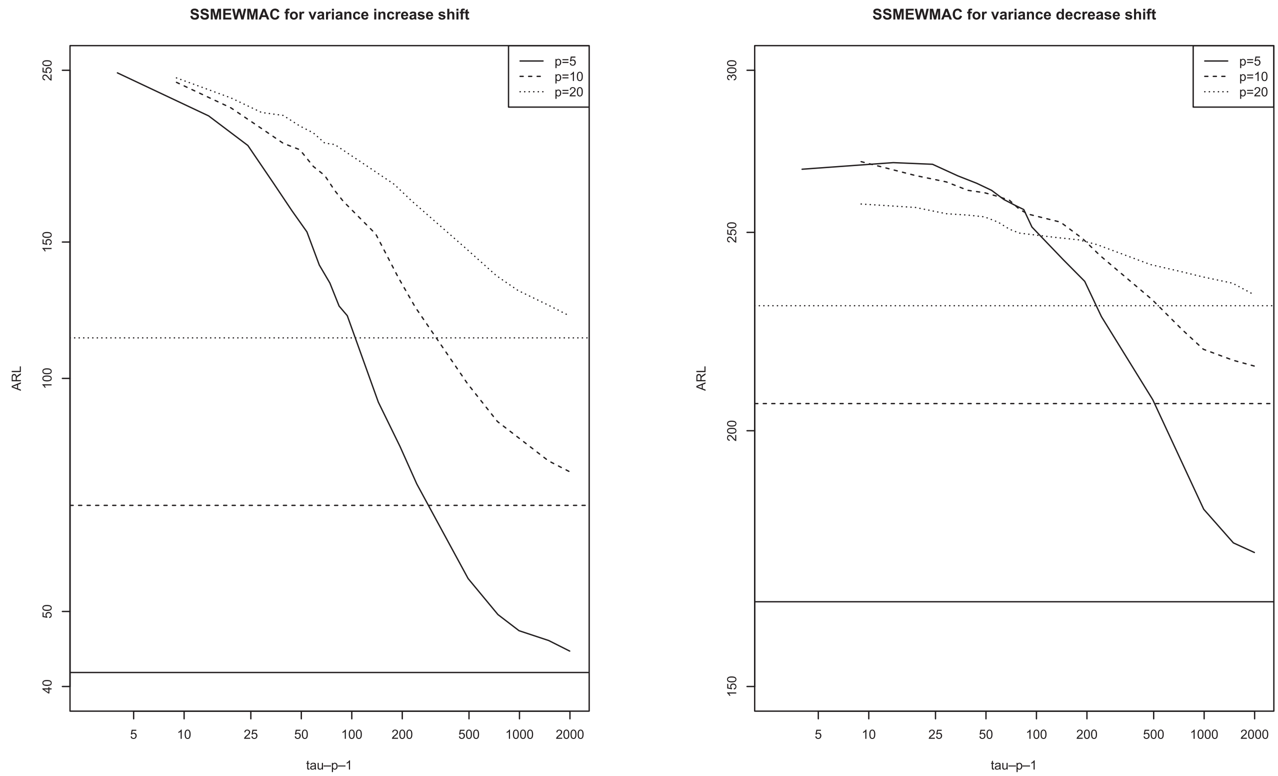

FIGURE 3. SSMEWMAC for Change in Variance. Left: shift increase, right: shift decrease. The horizontal lines represent OOC ARL of the known parameter MEWMAC.

(2008) pointed out that the underlying MEWMC is ARL-biased. The effect of the ARL bias is reduced in higher dimension.

Examples

We use two real data sets from the paper of Porzio and Ragozini (2003). The data derive from a car production line. Bodies of vehicles are monitored through optical electronic devices taking real-time measurements on 68 key points on the surfaces. The data consist of 1000 cars, so yielding 1000 cases with 68 variables $(X_1, X_2, \ldots, X_{68})$. In light of the substantial missing information, we will illustrate using two small subsets of the dimensions.

Example 1

The first subset is Porzio and Ragozini's variables $(X_1, X_2, X_3, X_4, X_8)$. We set the smoothing constant λ to 0.1 and each in-control ARL to 500. As we have five variables, this corresponds (using the Web tables) to $h_1 = 17.186$, $h_2 = 1.754$. Calculating the SSMEWMAC for the first 55 process readings gives the results shown in Figure 5. The SSMEWMA part of the chart goes outside its control limit at the 53rd reading, while the SSMEWMC chart remained inside its control limits during the whole sequence.

Diagnosing the shift consists of estimating when it occurred and of what the change consisted. Unlike the CUSUM, an EWMA does not give a direct estimate of the time of the shift, but one can be found visually by thinking about the EWMA's expected value. While the process is in control, the EWMA has a constant expectation, but if a step change occurs, it starts to rise, ultimately stabilizing at a constant out-of-control level. Thus, the visual diagnosis consists of judging where the EWMA changed from a broadly horizontal meander to a path moving toward some new asymptote, the time of this change then defining the change point.

Applying this reasoning to Figure 5, the SSMEWMA appears to be drifting near horizontally until about reading 45, after which it shoots up. This suggests that something changed about the 46th observation, and so we proceed on the basis that the first 45 observations are the in-control portion and the subsequent readings the out-of-control portion.

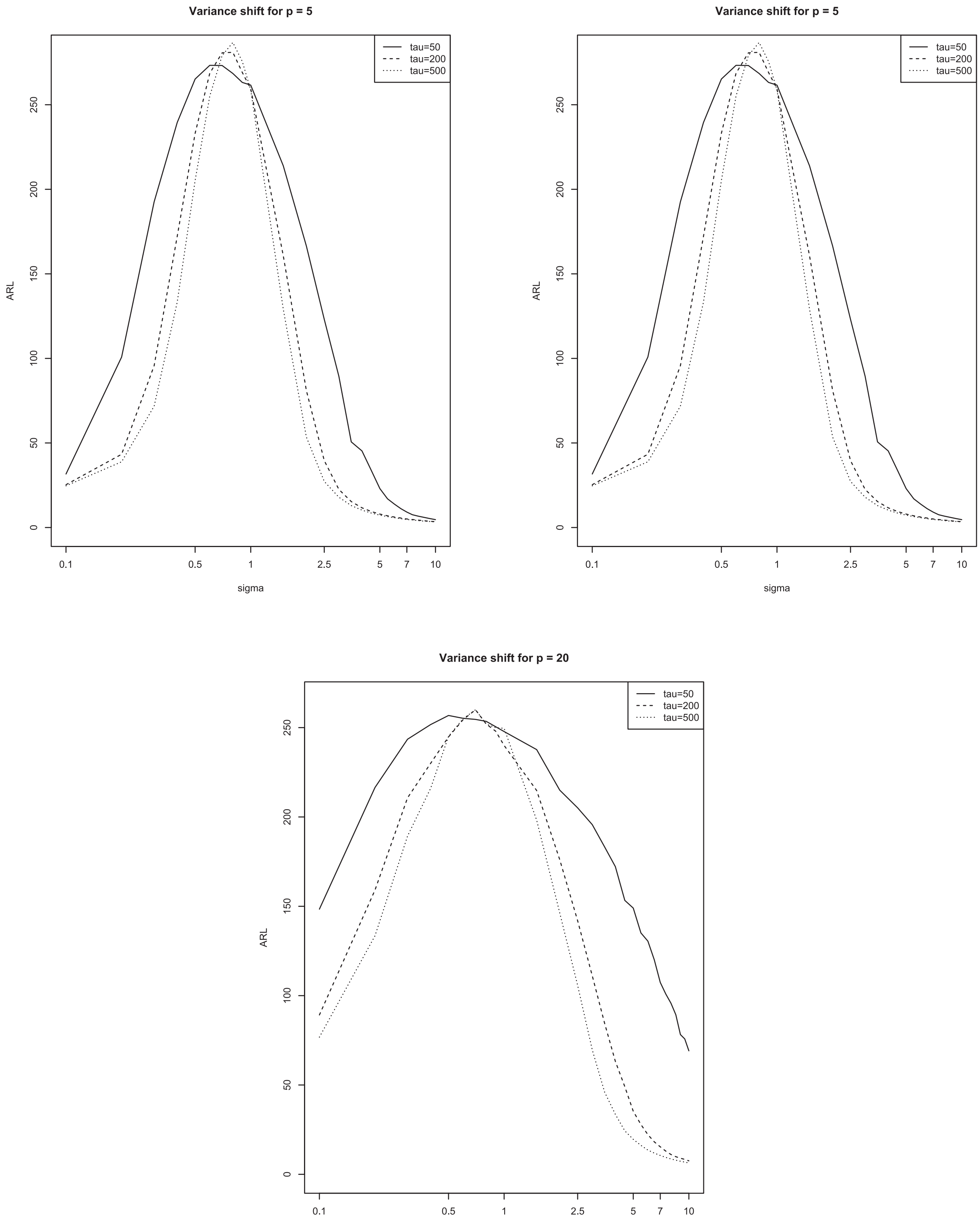

FIGURE 4. SSMEWMAC for Change in Variance. Top left: $p = 5$, top right: $p = 10$, bottom: $p = 20$.

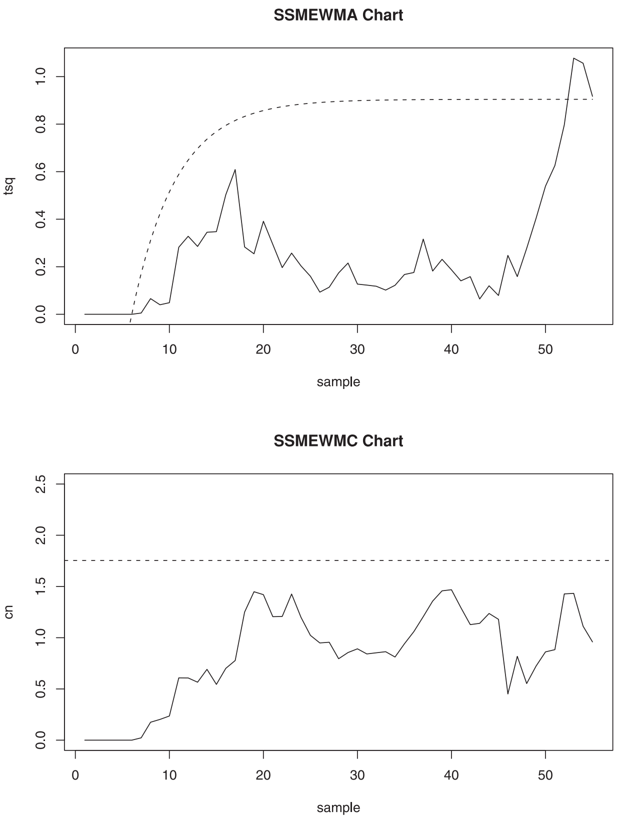

FIGURE 5. SSMEWMAC Plots.

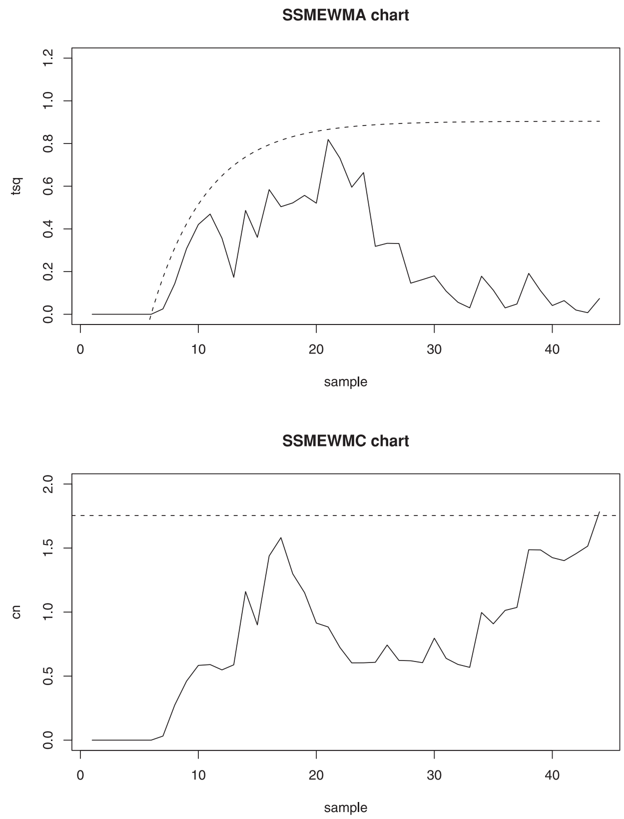

FIGURE 6. SSMEWMAC Plots.

Looking more closely at the issue of diagnosis, matters are complicated by the fact that a signal in the SSMEWMA component of the scheme does not necessarily mean that the mean vector has shifted. The signal can be triggered by several factors.

- A change in the mean vector,
- A change in the relationship between the variables,
- A change in the variability.

Expanding on the latter two possibilities, in the known-parameter setting, the location and the scale components of the combined chart do not give "clean" unambiguous signals. A change in the covariance matrix of the data impacts the run-length behavior of the location chart, so a signal from the SSMEWMA could be an indirect indicator of an increase in the variability of the data. On the other side, as the SSMEWMC is implicitly centered at the in-control mean vector, a shift in the mean vector accelerates signaling in the SSMEWMC. This means that an SSMEWMA signal could be caused by a variance increase and an SSMEWMC signal could be caused by a shift in mean. Thus careful interpretation is necessary after a signal.

Several methods to perform diagnostics following a shift are available—see, for example, Mason et al. (2002) or Hawkins and Maboudou-Tchao (2008).

Recall the meaning of the five U variables:

- U_1 is the variable X_1 transformed to an exact in-control N(0, 1) distribution.
- U_2 is the residual of the regression of X_2 on X_1, similarly transformed to an exact N(0, 1) in-control distribution.
- U_3, U_4, and U_5 are residuals of the regression of X_3, X_4, and X_5 on the X's of lower subscript, similarly transformed to N(0, 1).
- All five U variables are uncorrelated.

Thus, changes in the mean vector of $\mathbf{x}$ will tend to be reflected as the mean vector of $\mathbf{u}$ becoming nonzero; changes in the regression coefficients of the X_j on their predecessors will tend to be be reflected in nonzero covariances between the U_j and changes

TABLE 1. Diagnosis Results for the OOC Sample Using Hawkins Regression Adjusted Follow Up

	Dependent variables				
Constant	U_1	U_2	U_3	U_4	U_5
Coeff	−0.278	−1.396	−0.012	1.104	−0.569
s.e.	0.613	0.286	0.519	1.008	1.457
t	−0.454	−4.887	−0.022	1.095	−0.391
P-value	0.663	0.00275**	0.9831	0.335	0.722

in the regression residual variance of the X_j will tend to be reflected in the variances of the U_j departing from the in-control value 1.

While this problem is not really a 'cascade' problem, we think this ordering of the variables is not unreasonable, and we will apply the general diagnostic approach proposed by Fatti and Hawkins (1986), whose SPC application was outlined in Hawkins and Maboudou-Tchao (2008).

The Porzio and Ragozini data set showed no hint of a shift on the covariance matrix of this sequence of readings; the SSMEWMC remains well inside its control limits. Our interpretation will therefore be confined to a study of the "intercept" component of the $\mathbf{u}$ vectors, which diagnoses shifts in the mean vector. Table 1 shows the estimate and standard error of these intercepts, along with t-values and nominal P-values, for the out-of-control readings 46 through 55.

This table shows just one striking feature—the highly significant value of the U_2 intercept in the out-of-control period. This suggests that the shift that occurred was the rather straightforward one of a change in the mean of X_2 after observation 45.

Summary statistics of the in- and out-of-control sequences support this interpretation:

$$\bar{\mathbf{x}}_{\text{IC}} = (0.068, 0.202, 0.647, 0.389, -0.345)$$

and

$$\bar{\mathbf{x}}_{\text{OOC}} = (0.044, 0.068, 0.752, 0.394, -0.405).$$

The pooled standard deviations of the $\mathbf{x}$ are $(0.085, 0.126, 0.192, 0.179, 0.170)$, showing that the mean of X_2 dropped by 1.5 standard deviations after reading 45, while the remaining measurements stayed close to their earlier means.

Example 2

Next, we use a different subset of variables (X_9, X_{10}, X_{12}, X_{13}, X_{16}) from the same data, which we will denote $(X_6, X_7, X_8, X_9, X_{10})$ for simplicity.

The charts for this data set are shown in Figure 6; the chart parameters are the same as those used in Example 1.

Unlike the first example, the SSMEWMA remained inside its control limit for the entire sequence, but the SSMEWMC signaled an out-of-control at reading 44. This gives an indication that the covariance matrix changed at some point during the sequence. Visually, the C-trace started a steep rise around observation 32, so we tentatively divide the series into the in-control sequence 1–32 and the out-of-control 33–44.

The diagnosis of the presumed covariance matrix shift consists of fitting the regressions of each U_j on $U_1, \ldots, U_{j-1}$ in the out-of-control data. Each of these regressions should be null, with no nonzero true coefficients, and with residual variance equal to 1. These targets correspond to testable hypotheses giving rise to nominal P-values. These P-values should be viewed with some caution because we look at the follow-up diagnostics only after receiving a signal that the collection of null hypotheses is implausible. Still, clearly large and clearly small P-values provide helpful guidance on the source of the signal.

Table 2 summarizes the results of this testing. Each column shows the fitted coefficients of the regression of that U_j on its predecessors. The last row of the table gives the results of testing whether the residual variance equals 1. Along with each fitted quantity is the nominal P-value testing whether the quantity is what it should be under the in-control situation.

TABLE 2. Diagnosis Results for the OOC Sample Using Hawkins Regression Adjusted Follow Up

		Dependent variables				
	Constant	U_6	U_7	U_8	U_9	U_{10}
Constant	Coeff	0.103	0.185	−0.134	0.086	0.105
	s.e.	0.207	0.215	0.241	0.562	0.181
	t	0.498	0.796	−0.555	0.154	0.580
	P-value	0.628	0.445	0.592	0.881	0.579
U_6	Coeff	—	−0.67	−0.521	1.812	0.239
	s.e.	—	0.335	0.399	0.997	0.381
	t	—	−2.002	−1.305	1.818	0.630
	P-value	—	0.0732	0.224	0.107	0.549
U_7	Coeff	—	—	−0.134	0.836	0.497
	s.e.	—	—	0.318	1.134	0.378
	t	—	—	−3.572	0.737	1.319
	P-value	—	—	0.006**	0.482	0.229
U_8	Coeff	—	—	—	0.166	0.103
	s.e.	—	—	—	0.764	0.246
	t	—	—	—	0.217	0.419
	P-value	—	—	—	0.833	0.687
U_9	Coeff	—	—	—	—	−0.294
	s.e.	—	—	—	—	0.114
	t	—	—	—	—	−2.590
	P-value	—	—	—	—	0.036*
Variances	σ^2	0.513	0.632	0.641	3.367	0.348
	df	11	10	9	8	7
	χ^2 stat	5.646	6.325	5.768	26.934	2.434
	P-value	0.208	0.425	0.474	0.0014**	0.136

The regressions include an intercept, which, in the absence of a signal of a mean shift from the MEWMA component, is logically redundant, and indeed, the first row of Table 2 does not suggest any differences. Turning to the remainder of Table 2—the slopes—the significant coefficient of U_8 on U_7 ($P = 0.006$) suggests a change in the relationship between X_8 and X_7, and the marginally significant ($P = 0.036$) coefficient of U_{10} on U_9 hints at a change in the relationship between X_{10} and X_9.

Much more striking is the test of the residual variance of U_9. In the post-32 segment, this variance is more than three times higher than it is up to reading 32, and this ratio is highly significant ($P = 0.0014$). Either the marginal variance of X_9 has increased substantially or the relationship of X_9 with its predecessors has weakened.

The step-down diagnostics used here do not indicate which of these two is the root cause. A quick check is provided in Table 3, which shows the estimates of the coefficients (with their standard errors) of the regression of X_8 on X_6 and X_7 as well as the residual variances before and after the shift was detected. The coefficient of X_7 after the shift differs from the preshift value by some four standard errors. This means that there is a change in the relationship between X_8 and X_7. There is, however, no significant difference in the residual variances before and after

TABLE 3. Regression of X_8 on X_6 and X_7

	Before shift	After shift
$\hat{\beta}_6$	−0.16 (0.63)	−1.89 (1.34)
$\hat{\beta}_7$	0.28 (0.46)	−2.52 (0.78)*
Residual MS	0.17	0.11

the shift. This agrees with the conclusion obtained based on the $\mathbf{u}$'s.

Next, Table 4 displays the estimates of the coefficients (with their standard errors) of the regression of X_9 on X_6, X_7, and X_8 as well as the residual variances before and after the shift was detected. There is no significant change for the slope portion of the regression, but the residual variance of X_9 increases after the shift. So the shift observed at observation 32 is apparently due to the substantial increase in the marginal variance of X_9.

Diagnosis Following a Shift

The two examples presented above exemplify diagnosis when one of the chart signals without a signal on the other chart.

More generally, when one or both charts signal, the first step is a diagnosis of when the process appears to have shifted. A visual approach to this task is to decide when the charts went from a broadly horizontal drift to where one or both of them appeared to show an upward trend. This defines the putative changepoint.

Let $\boldsymbol{\mu}_L$ and $\boldsymbol{\Sigma}_L$ denote the parameters of the left or preshift sequence. Similarly, $\boldsymbol{\mu}_R$ and $\boldsymbol{\Sigma}_R$ denote the parameters of the right or postshift sequence. The diagnosis of the shift can be thought of as testing the hypotheses

1. H_0: $\boldsymbol{\mu}_L = \boldsymbol{\mu}_R$ vs H_1: $\boldsymbol{\mu}_L \neq \boldsymbol{\mu}_R$.
2. H_0: $\boldsymbol{\Sigma}_L = \boldsymbol{\Sigma}_R$ vs H_1: $\boldsymbol{\Sigma}_L \neq \boldsymbol{\Sigma}_R$.

The hypothesis of equal covariance matrices can be tested by the Bartlett likelihood ratio test, and to the extent that the covariance matrices are comparable, the hypothesis of equal mean vectors can be tested by a two-sample Hotelling T^2 test.

As illustrated in the two examples, the mean vector and covariance matrix of the postshift $\mathbf{u}$ can be helpful in shedding light on exactly what it was that changed in the mean vector or the covariance matrix, identifying shifts in the "covariate adjusted" mean, for example, or pointing to shifts in the regression relationships among the variables.

TABLE 4. Regression of X_9 on X_6, X_7 and X_8

	Before shift	After shift
$\hat{\beta}_6$	0.81 (0.48)	5.41 (2.55)
$\hat{\beta}_7$	0.72 (0.35)	2.23 (1.98)
$\hat{\beta}_8$	−0.27 (0.14)	−0.14 (0.57)
Residual MS	0.10	0.31
Total variance	0.12	0.41

Cautions and Pitfalls with Self-Starting Charts

Self-starting charts have a number of features that require attention. One is the reaction to a signal. When the self-starter signals, it is essential to respond to the signal by diagnosing what changed and when the change occurred and to quarantine the data gathered during the altered regime, keeping them out of the updating process. Failure to do so will result in contaminated estimates that distort the running means and inflate the variances.

Another feature requiring attention is the reaction toward very early shifts. If the process runs at its in-control level for just a handful of readings and then shifts or if the process starts in the wrong location so that even the first observations are already out-of-control, then one never gets adequate information on the in-control parameters and the chart may fail to indicate that there are any problems. It is a prudent precaution against an early shift, once one is past the very early phase of data gathering, to "run the chart backwards"—that is, to start with the most recent process reading and successively add the earlier readings, going back to the start of the sequence.

This attempted solution does not always work. In particular, it will fail if the process starts in the wrong location and stays there. However, this same criticism can be made, even more strongly, for the traditional approach, which requires that a large Phase I sample (rather than the self-starter's more modest learning requirement) be gathered while the process is in control.

A perhaps more serious blind spot is that the self-starting approach is unable to detect slow drifts for which cumulative effect on the running means and, more important, variances may mask their presence.

Conclusion

Traditional control charts are designed under the assumption that the in-control process parameters are known exactly. Powerful charts that accumulate information across successive samples, as do MEWMA or MEWMC charts, are inherently unable to distinguish between random errors in the estimates of the parameters and genuine shifts. This is ameliorated, but not solved, by collecting larger Phase I data sets, which can be daunting. The unknown-parameter self-starting formulation removes the need for a large Phase I exercise. This is of particular benefit for short-run and start-up processes, as it allows for control schemes that monitor the process mean vector and covariance matrix as soon as enough observations have been accumulated to get an initial estimate of the mean vector and covariance matrix and have a single remaining degree of freedom for error.

The proposal includes a control chart for the mean vector and a separate one for the covariance matrix. There is some cross chatter between the two component charts: a shift in the covariance matrix affects the run behavior of the mean chart, and a shift in the mean vector shortens the runs of the covariance matrix chart. Thus, follow-up diagnostics are needed to identify the shift. Using an example, we illustrate the step-down approach that is particularly relevant to cascade processes, but that may be applied in any setting.

Self-starting methods do not solve all problems. They are inherently unable to detect situations in which a process starts up out of control or moves out of control very soon after start up. Thus, while self-starting methods do not require the traditional approach's large Phase I data gathering exercise during which the process is known to be in control, they do require some modicum of control at the start of data gathering. It is also essential to react to signals given by self-starting charts. If the mean shifts but no effort is made to diagnose and correct the shift, in due course, the estimated mean vector will adjust to the new level, implicitly turning that into the new in-control setting. We believe these are small prices to pay for the benefit of exact control over run-length behavior with minimal startup requirements,

Appendix
Definition of the Vectors $\boldsymbol{r}_t$ and $\boldsymbol{u}_t$

We now give more details of the multivariate self-starting approach of Hawkins and Maboudou-Tchao (2007). Suppose the p-component process reading vectors are $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

The ultimate objective is to transform $\mathbf{x}$ to a standard normal vector $\mathbf{z}$ using a transformation that tends to the linear transformation $\mathbf{z} = \mathbf{A}(\mathbf{x} - \boldsymbol{\mu})$, where the matrix $\mathbf{A}$ satisfies $\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}' = \mathbf{I}_p$ and $\mathbf{z} \sim N(\mathbf{0}, \mathbf{I}_p)$. Hawkins and Maboudou-Tchao (2007) suggested the choice of the triangular Cholesky inverse root of $\boldsymbol{\Sigma}$ as the matrix $\mathbf{A}$.

For $t \geq i + 1$, a predicted value $\hat{x}_{t,i}$ for the tth reading on the ith variable is obtained by repeatedly fitting the regression of x_i on its $x_1, \ldots, x_{i-1}$. The recursive residual, $r_{t,i}$, is the difference between the observed and the predicted value of the ith variable rescaled using its leverage to constant variance. After t process vectors have been observed, there will be $t - p$ $\mathbf{r}_t$ vectors (the start-up first p being not fully defined), leading to a matrix $\mathbf{R}$ of order $(t - p) \times p$.

The standardized recursive residual is defined as

$$t_{t,i} = \frac{r_{t,i}}{\sqrt{\sum_{k=i+1}^{t-1} r_{k,i}^2/(t - i - 1)}} \tag{9}$$

and follows a t-distribution with $t - i - 1$ degrees of freedom.

Next, a double probability integral transform is applied to the standardized recursive residuals to obtain independent standard normal statistics $u_{t,i}$ as

$$u_{t,i} = \Phi^{-1}[F_{t-i-1}(t_{t,i})], \tag{10}$$

where Φ^{-1} stands for the inverse normal and F_{t-i-1} the cumulative distribution function of t with $t-i-1$ degrees of freedom.

These independent standard normals $u_{t,i}$ define the matrix $\mathbf{U}$ of independent $N(\mathbf{0}, \mathbf{I}_p)$ vectors and so, after gathering t process readings, there will be $t - p - 1$ standard normal $\mathbf{u}_t$ vectors.

Acknowledgment

The authors are grateful to the Editor and referees for a number of constructive suggestions for improvement.

References

ALT, F. A. (1984). "Multivariate Quality Control". *The Encyclopedia of Statistical Sciences*, Vol. 6, N. L. Johnson, S. Kotz, and C. R. Read, eds., pp. 110–122. New York, NY: Wiley.

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

BROWN, R. L.; DURBIN, J.; and EVANS, J. M. (1975). "Techniques for Testing the Constancy of Regression Relationship Over Time (with Discussion)". *Journal of the Royal Statistical Society, Series B* 37, pp. 149–192.

CAPIZZI, G. and MASAROTTO, G. (2010). "Self-Starting CUSCORE Control Charts for Individual Multivariate Observations". *Journal of Quality Technology* 42, pp. 136–151.

CHAMBERS, J. M. (1971). "Regression Updating". *Journal of the American Statistical Association* 66, pp. 744–748.

CHAMP, C. W.; JONES-FARMER, L. A.; and RIGDON, S. E. (2005). "Properties of the T^2 Control Chart When the Parameters Are Estimated". *Technometrics* 47, pp. 437–445.

FATTI, P. L. and HAWKINS, D. M. (1986). "Variable Selection in Heteroscedastic Discriminant Analysis". *Journal of the American Statistical Association* 81, pp. 494–500.

HAWKINS, D. M. (1987). "SELF-STARTING CUSUM CHARTS FOR LOCATION AND SCALE". *The Statistician* 36, PP. 299–315.

HAWKINS, D. M. and MABOUDOU-TCHAO, E. M. (2007). "Self-Starting Multivariate Exponentially Weighted Moving Average Control Charting". *Technometrics* 49, pp. 199–209.

HAWKINS, D. M. and Maboudou-Tchao, E. M. (2008)."Multivariate Exponentially Weighted Moving Covariance Matrix". *Technometrics* 50, pp. 155–166.

HOTELLING, H. (1947). "Multivariate Quality Control, Illustrated by the Air Testing of Sample Bombsights". In *Techniques of Statistical Analysis*, C. Eisenhart, M. Hastay, and W. A. Wallis, eds., pp. 111–184. New York, NY: McGraw-Hill.

HUWANG, L.; YEH, A. B.; and WU, C. W. (2007). "Monitoring Multivariate Proess Variability for Individual Observations". *Journal of Quality Technology* 39, pp. 258–278.

JONES, L. A.; CHAMP, C. W.; and RIGDON, S. E. (2001). "The Performance of Exponentially Weighted Moving Average Charts with Estimated Parameters". *Technometrics* 43, pp. 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, pp. 46–53.

MASON, R. L.; TRACY, N. D.; and YOUNG, J. C. (1995). "Decomposition of T^2 for Multivariate Control Chart Interpretation". *Journal of Quality Technology* 27, pp. 99–108.

MONTGOMERY, D. C. (2001). *Introduction to Statistical Quality Control*, 4th edition. New York, NY: Wiley.

MONTGOMERY, D. C. and WADSWORTH, H. M. (1972). *Some Techniques for Multivariate Quality Control Applications.* Washington, DC: Transactions of the ASQC.

PIGNATIELLO, J. J.; ACOSTA-MEJIA, C. A.; and RAO, B. V. (1995). "The Performance of Control Charts for Monitoring Process Dispersion". *Proceeding of the 4th Industrial Engineering Reseacrch Conference*, pp. 320–328.

PORZIO, G. C. and RAGOZINI, G. (2003). "Visually Mining Off-Line Data for Quality Improvements". *Quality and Reliability Engineering International* 19, pp. 273–283.

QUESENBERRY, C. P. (1991). "SPC Q Charts for Start-Up Processes and Short or Long Runs". *Journal of Quality Technology* 23, pp. 213–224.

QUESENBERRY, C. P. (1993). "The Effects of Sample Size on Estimated Limits for $\overline{X}$ and X Control Charts." *Journal of Quality Technology* 25, pp. 237–247.

QUESENBERRY, C. P. (1995). "On Properties of Q Charts for Variables." *Journal of Quality Technology* 27, pp. 184–203.

QUESENBERRY, C. P. (1997). *SPC Methods for Quality Improvement.* New York, NY: Wiley.

SCHAFFER, J. R. (1998). " A Multivariate Application of the Q Chart". Presented at the 1998 Joint Statistical Meetings.

SULLIVAN, J. H. and JONES, L. A. (2002). "A Self-Starting Control Chart for Multivariate Individual Observations". *Technometrics* 44, pp. 24–33.

ZAMBA, K. D. and HAWKINS, D. M. (2009). "A Multivariate Change Point Model for Change in Mean Vector and/or Covariance Structure". *Journal of Quality Technology* 41, pp. 285–303.

ZOU, C. L. and TSUNG, F. (2010). "Likelihood Ratio-Based Distribution-Free EWMA Control Charts". *Journal of Quality Technology* 42, pp. 174–196.