Journal of Modern Applied Statistical Methods

Volume 11 | Issue 2 Article 3

11-1-2012

A Graphical Examination of Variable Deletion within the MEWMA Statistic

Jay R. Schaffer
University of Northern Colorado

Shawn VandenHul
University of Northern Colorado

Recommended Citation

Schaffer, Jay R. and VandenHul, Shawn (2012) "A Graphical Examination of Variable Deletion within the MEWMA Statistic," *Journal of Modern Applied Statistical Methods*: Vol. 11 : Iss. 2 , Article 3.
DOI: 10.22237/jmasm/1351742520
Available at: http://digitalcommons.wayne.edu/jmasm/vol11/iss2/3

Journal of Modern Applied Statistical Methods
November 2012, Vol. 11, No. 2, 303-318

1538 – 9472/12/$95.00

Regular Articles

A Graphical Examination of Variable Deletion within the MEWMA Statistic

Jay R. Schaffer Shawn VandenHul
University of Northern Colorado
Greeley, CO

A general procedure for identifying the variable(s) that contribute(s) to the signal of the multivariate extension of the exponentially weighted moving average (MEWMA) chart is presented. The procedure systematically removes one or two variables from the MEWMA statistic calculations. Percentages are calculated for correctly identifying various shifts.

Key words: Multivariate quality control, MEWMA, variable deletion.

Introduction

With modern computers, it is common to monitor several correlated quality characteristics simultaneously. Various types of multivariate control charts have been proposed to take advantage of the relationships among variables being monitored (Alt, 1984; Jackson, 1985; Wierda, 1994; Lowry and Montgomery, 1995; Mason, et al., 1997). Lowry, et al. (1992) proposed a multivariate extension of the exponentially weighted moving average (MEWMA) control chart. They demonstrated that the average run length (ARL) performance of the MEWMA is similar to that of the multivariate cumulative sum (MCUSUM) control charts discussed by Crosier (1988) and Pignatiello and Runger (1990) and is better than Hotelling's (1947) χ^2 chart for detecting a shift in the mean vector of a multivariate normal distribution.

Woodall and Montgomery (1999) showed that, once an out-of-control signal is given by a multivariate chart, it may be difficult to identify the variable (or variables) that contributed to the signal. Jackson (1980, 1991) proposed examining the Hotelling's T^2 statistic (Jackson, 1985) using principle component analysis (PCA). Mason, et al. (1995) suggested decomposing Hotelling's T^2 statistic by removing individual variables from its calculation. Woodall and Montgomery (1999) noted that additional work is needed on graphical methods for data visualization when interpreting signals from multivariate control charts.

This article presents a graphical approach to identify the source of a signal from the MEWMA control chart and examines the effects of systematically deleting a variable, or pairs of variables, from the calculations of the MEWMA statistic. The methodology is similar to examining the PRESS residuals (Allen, 1971) or DFBETAS (Belsley, et al., 1980) in regression analysis. Methodology used herein deletes variables in a multivariate process as opposed to deleting individual observations in a data set; in addition, the probability of correctly identifying the source using various simulations is estimated.

Jay Schaffer is an Associate Professor of Statistics and is currently the chair of Applied Statistics and Research Methods. Email him at: jay.schaffer@unco.edu. Shawn VandenHul is a graduate of the Ph.D. program in Applied Statistics and Research Methods.

MEWMA Chart

Assume a sequence of independent observations from a p-variate normal distribution whose mean vector shifts from $\boldsymbol{\mu}_0$ to $\boldsymbol{\mu}_1$ on the rth observation, that is,

$$\mathbf{x}_i \sim N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}), \quad i = 1,2,\ldots,r-1$$
$$\sim N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}), \quad i = r, r+1, r+2\ldots \tag{1}$$

Lowry, et al. (1992) defined vectors of exponentially weighted moving averages,

$$\mathbf{z}_i = \lambda \mathbf{x}_i + (1-\lambda)\mathbf{z}_{i-1} \tag{2}$$

i = 1, 2, …, where $\mathbf{z}_0 = \mathbf{0}$ and $0 < \lambda \leq 1$. The MEWMA chart would give an out-of-control signal if

$$T_i^2 = \mathbf{z}_i{}'\boldsymbol{\Sigma}_{\mathbf{z}_i}^{-1}\mathbf{z}_i > h \tag{3}$$

where h > 0 is chosen to achieve a specified in-control ARL and

$$\boldsymbol{\Sigma}_{\mathbf{z}_i} = \frac{\lambda}{2-\lambda}\left[1-(1-\lambda)^{2i}\right]\boldsymbol{\Sigma}. \tag{4}$$

The one-variable deletion within the MEWMA statistic removes variables from the T_i^2 statistic when a signal is detected. This study examines the removal of one variable and two variables at a time: one-variable deletion removes one variable at a time and recalculates the current T_i^2 statistic excluding the removed variable, two-variable deletion removes pairs of variables and recalculates the current T_i^2 statistic excluding the removed pair. Given either method, a small, reduced T_i^2 statistic would indicate a possible signal source.

One-Variable Deletion

Assume on the s^{th} sample, the MEWMA chart signaled a change ($T_s^2 > h$). The p variables are removed, one at a time, from the calculation of T_s^2. Assume the j^{th} variable is removed such that

$$\mathbf{x}_i{}' = \left(x_{i1}, x_{i2}, \ldots, x_{i,j-1}, x_{i,j+1}, \ldots, x_{ip}, x_{ij}\right)$$
$$= \left(\mathbf{x}_{i(j)}{}', x_{ij}\right),$$

where $\mathbf{x}_{i(j)}{}'$ is a (p-1)x1 vector excluding the j^{th} variable. In addition, let $\boldsymbol{\Sigma}_{(j)}$ be the (p-1)x(p-1) principal sub-matrix of $\boldsymbol{\Sigma}$ excluding the j^{th} variable. With the j^{th} variable removed, the MEWMA equations become

$$\mathbf{z}_{i(j)} = \lambda \mathbf{x}_{i(j)} + (1-\lambda)\mathbf{z}_{i-1,(j)} \tag{5}$$

i = 1, 2, …, s where $\mathbf{z}_{0(j)} = \mathbf{0}$,

$$T_{i(j)}^2 = \mathbf{z}_{i(j)}{}'\boldsymbol{\Sigma}_{\mathbf{z}_{i(j)}}^{-1}\mathbf{z}_{i(j)} \tag{6}$$

and

$$\boldsymbol{\Sigma}_{\mathbf{z}_{i(j)}} = \frac{\lambda}{2-\lambda}\left[1-(1-\lambda)^{2i}\right]\boldsymbol{\Sigma}_{(j)}. \tag{7}$$

The calculation of $T_{i(j)}^2$ continues until the s^{th} sample.

A graphical comparison of the set of reduced MEWMA statistics $\{T_{s(1)}^2, T_{s(2)}^2, \ldots, T_{s(p)}^2\}$ to T_s^2 should aid in identifying the cause of the signal. The smallest reduced MEWMA statistic may indicate which variable contributed to the signal. For example, if the 1^{st} variable shifts, the reduced MEWMA statistics may resemble Figure 1. A similar analysis is required if more than two variables change. For example, the reduced MEWMA statistics may resemble Figure 2, if the 1^{st} and the 2^{nd} variables shift or may resemble Figure 3, if the 1^{st}, 2^{nd} and 3^{rd} variables shift.

Consider a modified example from Lowry, et al. (1992). Assume

$$\mathbf{x}_i \sim N_3(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}), \quad i = 1,2,\ldots,15$$
$$\sim N_3(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}), \quad i = 16,17,18,\ldots \tag{8}$$

such that:

$$\boldsymbol{\mu}_\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

Figure 1: A General Representation of the Reduced MEWMA Statistics if Variable 1 Shifted

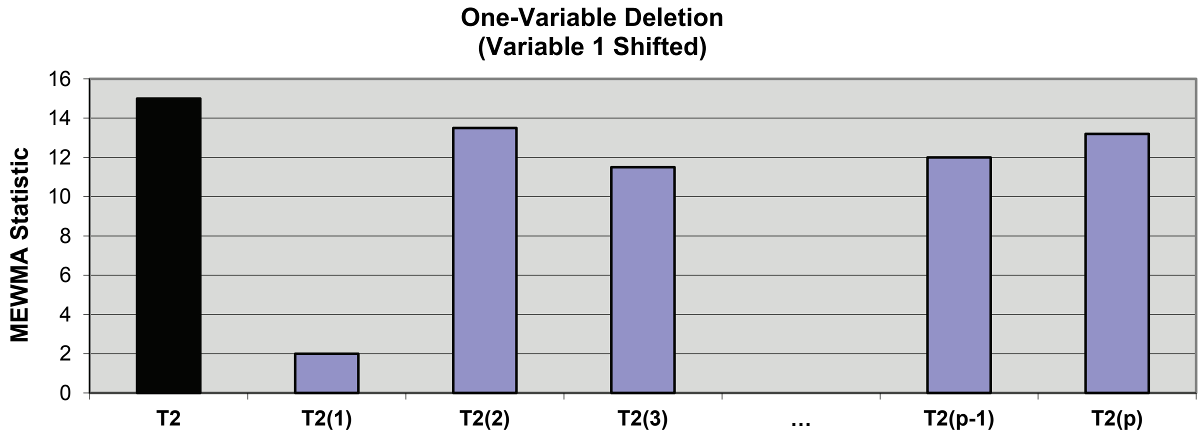


Figure 2: A General Representation of the Reduced MEWMA Statistics if Variables 1 and 2 Shifted

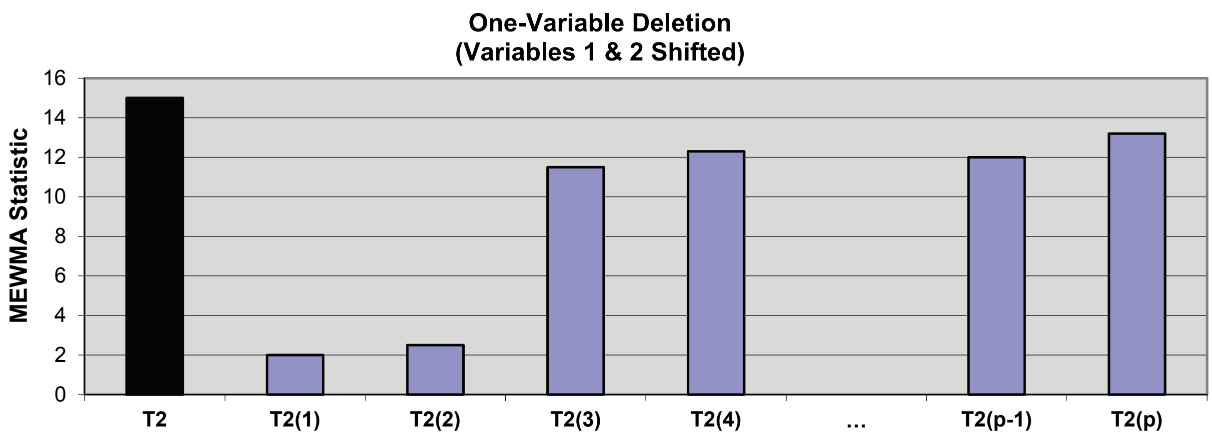
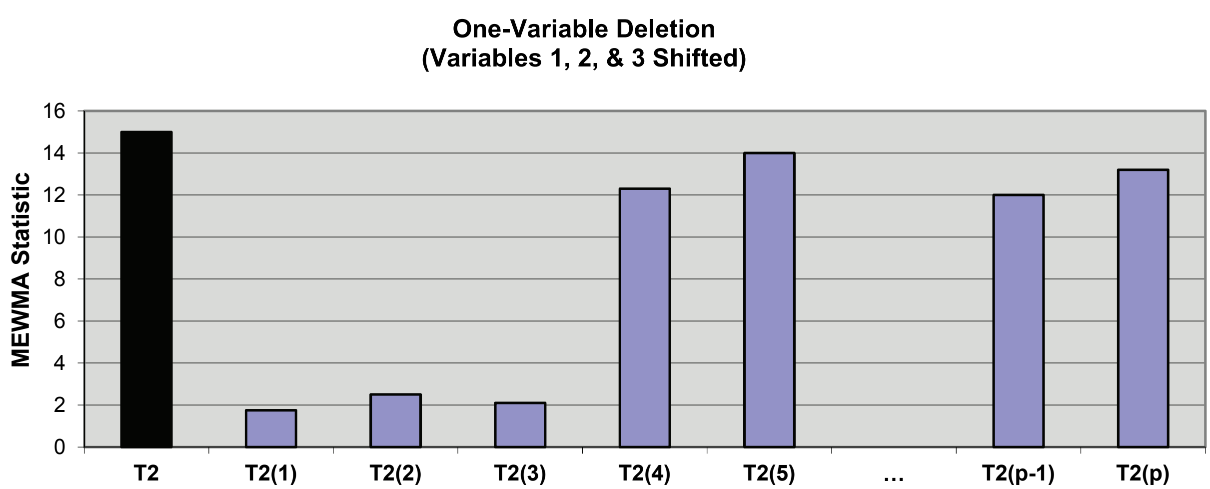


Figure 3: A General Representation of the Reduced MEWMA Statistics if Variables 1, 2 and 3 Shifted



$$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{2} \\ 0 \\ 0 \end{bmatrix},$$

and

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{bmatrix}.$$

Note that a shift of

$$\delta = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)} = 3$$

occurred on the 16th sample. Table 1 displays a data simulation of these conditions along with the corresponding MEWMA statistics, T_i^2. Using $\lambda = 0.10$, and h = 10.97 (in-control ARL = 200), the MEWMA chart signaled on the 21st observation such that T_{21}^2 = 11.3551. However, it is not apparent which variable changed through an examination of the data or the MEWMA chart.

Using the data from Table 1, the first variable is removed from the calculation of the MEWMA statistic. Variables 2 and 3 are used to recalculate a reduced MEWMA statistic, $T_{i(1)}^2$. The reduced covariance matrix is then:

$$\boldsymbol{\Sigma}_{(1)} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}.$$

The reduced MEWMA statistic, $T_{i(1)}^2$, is calculated for i = 1, 2, …, 21 and displayed in Table 2. Note that, on the 21st sample, $T_{21(1)}^2$ = 0.9358 represents the reduced MEWMA statistic with the contribution of the first variable removed.

Repeating the one-variable deletion procedure for the remaining two variables, the reduced MEWMA statistic excluding variable 2 is $T_{21(2)}^2 = 11.3282$ and the reduced MEWMA statistic excluding variable 3 is $T_{21(3)}^2 = 9.0015$. Comparing the three reduced MEWMA statistics to the MEWMA statistic $T_{21}^2 = 11.3551$, it is likely variable 1 contributed to the signal. Figure 4 displays the MEWMA statistic along with the three reduced MEWMA statistics.

Two-Variable Deletion

Assume on the sth sample, the MEWMA chart signaled a change (T_s^2 > h). The p variables are removed, two at a time, from the calculation of T_s^2. Assume the jth and kth variables are to be removed. Now let

$$\mathbf{x}_i' = \begin{pmatrix} x_{i1}, x_{i2}, \ldots, x_{i,j-1}, x_{i,j+1}, \ldots, \\ x_{i,k-1}, x_{i,k+1}, \ldots x_{ip}, x_{ij}, x_{ik} \end{pmatrix}, \\ = \left(\mathbf{x}_{i(j,k)}', x_{ij}, x_{ik}\right)$$

where $\mathbf{x}_{i(j,k)}'$ is a (p-2)x1 vector excluding the jth and kth variables. In addition, let $\boldsymbol{\Sigma}_{(j,k)}$ be the (p-2)x(p-2) principal sub-matrix of $\boldsymbol{\Sigma}$ excluding the jth and kth variables. With the jth and kth variables removed, the MEWMA equations become

$$\mathbf{z}_{i(j,k)} = \lambda \mathbf{x}_{i(j,k)} + (1-\lambda)\mathbf{z}_{i-1,(j,k)} \tag{9}$$

i = 1, 2, …, s where $\mathbf{z}_{0(j,k)} = \mathbf{0}$,

$$T_{i(j,k)}^2 = \mathbf{z}_{i(j,k)}' \boldsymbol{\Sigma}_{\mathbf{z}_{i(j,k)}}^{-1} \mathbf{z}_{i(j,k)} \tag{10}$$

and

$$\boldsymbol{\Sigma}_{\mathbf{z}_{i(j,k)}} = \frac{\lambda}{2-\lambda}\left[1-(1-\lambda)^{2i}\right]\boldsymbol{\Sigma}_{(j,k)}. \tag{11}$$

The calculation of $T_{i(j,k)}^2$ is continued until the sth sample.

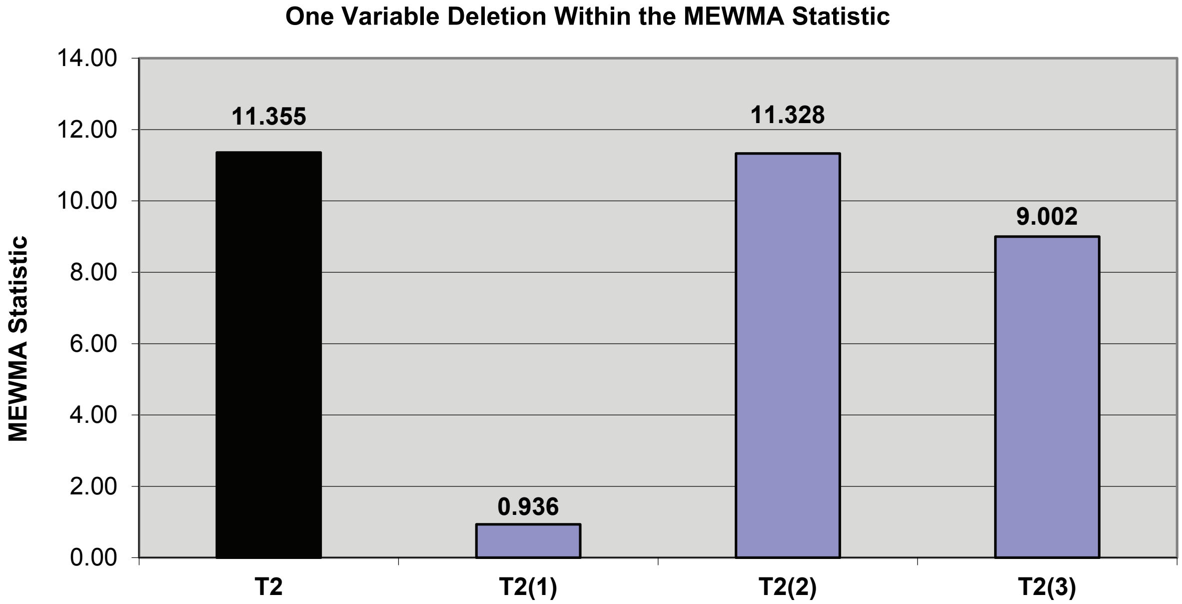
Table 1: Simulated Process with Corresponding MEWMA Statistics, T_i^2

i	x_1	x_2	x_3	T_i^2
1	0.1307	0.5629	-0.7255	0.7203
2	1.5662	-0.3972	0.5767	2.3382
3	0.5733	1.4400	1.4343	1.9161
4	-0.0342	-0.0966	0.8100	1.6907
5	0.2922	0.0853	-0.3257	1.2270
6	-0.2988	-0.7700	0.3948	1.1400
7	0.1389	0.4851	0.1806	0.8966
8	-0.0184	-0.5328	0.4871	1.3231
9	0.6751	-0.3919	-1.4367	1.1445
10	-2.5591	-1.4792	-2.3697	1.0214
11	-1.8930	0.4438	-0.9319	2.2448
12	-0.4950	0.4710	-0.0471	2.6071
13	-1.1572	0.8478	-0.5695	5.2338
14	0.2098	-0.8472	0.1777	2.7816
15	0.0101	0.1780	0.9616	2.0170
16	1.1233	-0.6925	-1.2685	1.1097
17	0.8364	-1.5027	-0.1821	1.6985
18	0.6587	1.0085	0.5520	0.8009
19	2.3631	2.1432	0.9458	2.0496
20	2.4894	0.2182	-0.2358	6.7361
21	2.3260	0.7702	0.5218	11.3551

Table 2: Reduced MEWMA Statistics, $T_{i(1)}^2$

i	1	2	3	4	5	6	7
$\mathbf{T_{i\,(1)}^2}$	1.6690	0.0192	1.1522	1.4192	0.7580	0.9119	0.7640
i	**8**	**9**	**10**	**11**	**12**	**13**	**14**
$\mathbf{T_{i\,(1)}^2}$	1.2226	0.0973	0.9894	1.5824	1.4426	2.5004	1.2420
i	**15**	**16**	**17**	**18**	**19**	**20**	**21**
$\mathbf{T_{i\,(1)}^2}$	0.2939	1.0651	1.3763	0.4595	0.5040	0.6895	0.9358

Figure 4: MEWMA T^2 Statistic and the Three Reduced MEWMA Statistics



A graphical comparison of the set of reduced MEWMA statistics $\{T_{s(1,2)}^2, T_{s(1,3)}^2, \ldots, T_{s(1,p)}^2, T_{s(2,3)}^2, T_{s(2,4)}^2, \ldots, T_{s(p-1,p)}^2\}$ to T_s^2 should aids in identifying the cause of the signal. The smallest group of reduced MEWMA statistics may indicate which variable contributed to the signal. For example, if the 1st variable shifts, the reduced MEWMA statistics may resemble Figure 5, such that the group of reduced MEWMA statistics associated with the first variable is uniformly smaller than the others.

A more detailed analysis is required if two variables have shifted. The smallest reduced MEWMA statistic may indicate which pair of variables changed. In addition, any reduced MEWMA statistic associated with one of the pair of variables that shifted may be slightly larger, yet smaller than any other reduced MEWMA statistic not associated with the pair that changed. The reduced MEWMA statistics may resemble Figure 6, if the 1st and the 2nd variables shift. A similar analysis is required if three variables have shifted. The reduced MEWMA statistics may resemble Figure 7, if the 1st, 2nd and 3rd variables shift.

Consider a modified example from Lowry, et al. (1992). Assume

$$\begin{aligned} \mathbf{x}_i &\sim N_4(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}), \quad i = 1, 2, \ldots, 15 \\ &\sim N_4(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}), \quad i = 16, 17, 18, \ldots \end{aligned} \tag{12}$$

such that:

$$\boldsymbol{\mu}_\mathbf{o} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\boldsymbol{\mu}_\mathbf{1} = \begin{bmatrix} \sqrt{3/2} \\ \sqrt{3/2} \\ 0 \\ 0 \end{bmatrix},$$

and

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.5 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 & 0.5 \\ 0.5 & 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1 \end{bmatrix}.$$

Figure 5: A General Representation of the Reduced MEWMA Statistics if Variable 1 Shifted

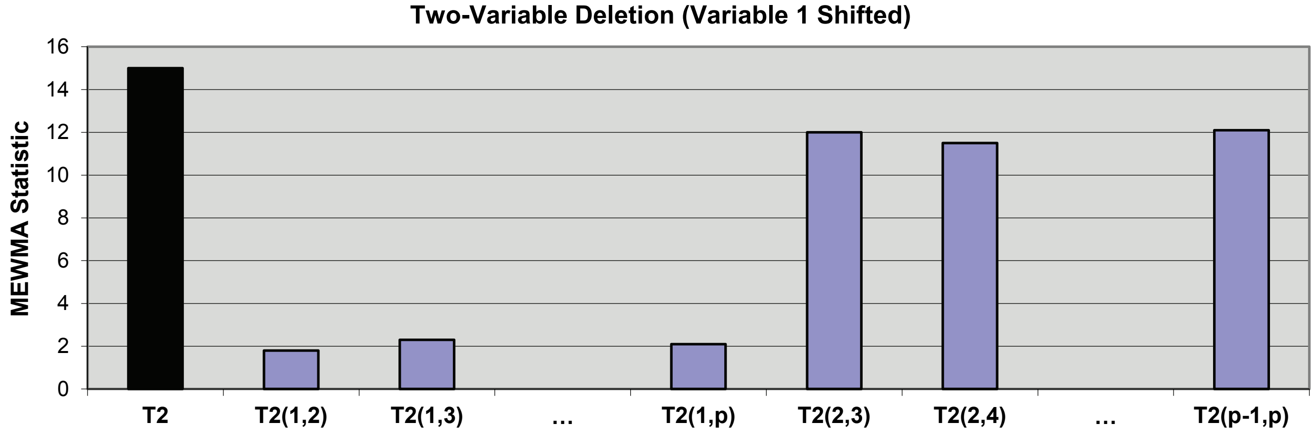


Figure 6: A General Representation of the Reduced MEWMA Statistics if Variables 1 and 2 Shifted

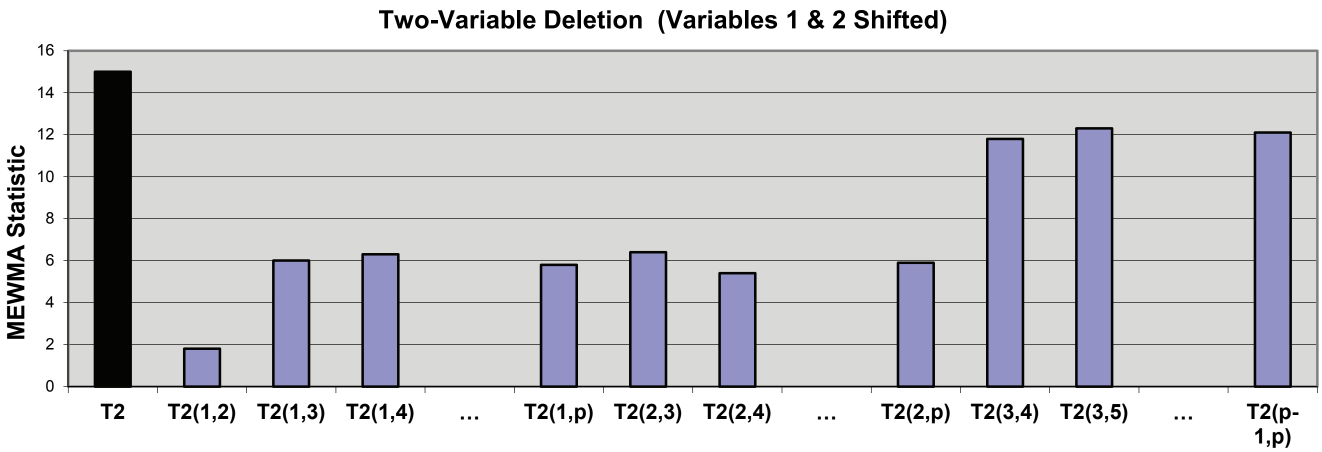


Figure 7: A General Representation of the Reduced MEWMA Statistics if Variables 1, 2 and 3 Shifted

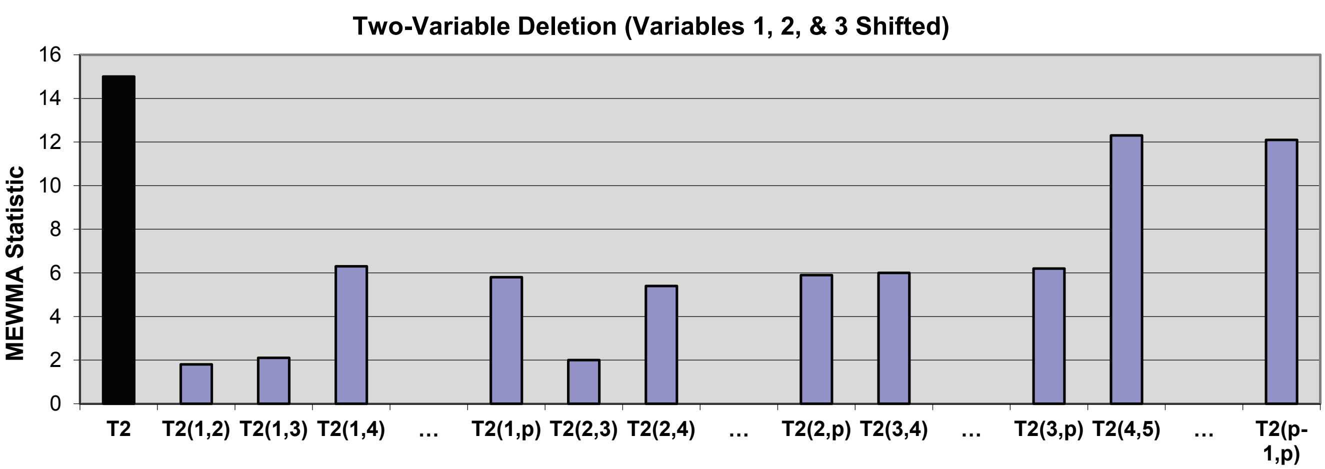


Table 3: Simulated Process with Corresponding MEWMA Statistics, T_i^2

i	X_1	X_2	X_3	X_4	T_i^2
1	0.502	0.130	0.150	0.086	0.296
2	-0.862	-0.877	-0.515	0.025	0.531
3	0.630	1.914	1.396	2.179	2.605
4	0.448	-0.422	-0.036	0.692	2.816
5	-0.995	-0.605	-0.772	-1.700	0.402
6	-0.090	1.305	-1.037	-0.812	1.540
7	-0.951	-1.808	-0.142	0.197	0.772
8	-0.549	-0.136	-0.350	0.671	1.857
9	0.068	-0.312	-2.316	0.680	5.749
10	2.132	0.072	-1.062	-1.362	6.477
11	-0.738	0.141	0.030	1.026	5.355
12	1.293	-1.380	-0.687	0.953	9.455
13	-0.249	-0.954	-1.079	-0.001	11.282
14	0.733	-1.432	0.480	-0.406	10.842
15	0.704	-0.170	-0.120	0.159	10.970
16	2.036	2.011	1.985	1.179	7.910
17	0.950	3.354	1.741	1.824	5.742
18	2.044	0.054	0.281	-0.287	8.560
19	0.858	0.593	0.151	-0.932	9.069
20	1.231	2.576	-0.213	-0.428	13.793

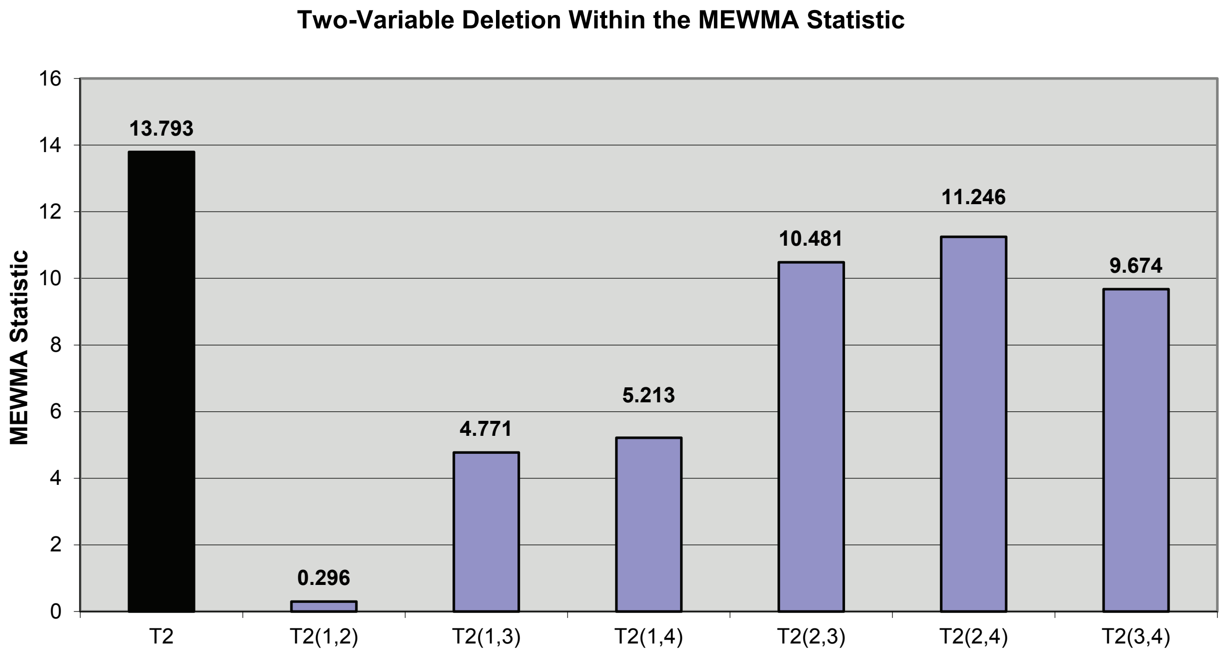
Note that a shift of

$$\delta = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)} = 3$$

occurred on the 16th sample. Table 3 displays a data simulation of these conditions along with the corresponding MEWMA T_i^2 statistics. Using $\lambda = 0.10$ and h = 12.93 (in-control ARL = 200), the MEWMA chart signaled on the 20th observation such that $T_{20}^2 = 13.793$.

Using equations (9)-(11), the reduced MEWMA statistics are $T_{20(1,2)}^2$ = 0.296, $T_{20(1,3)}^2$ = 4.771, $T_{20(1,4)}^2$ = 5.213, $T_{20(2,3)}^2$ = 10.481, $T_{20(2,4)}^2$ = 11.246, and $T_{20(3,4)}^2$ = 9.674. Figure 8 displays the reduced MEWMA statistics. Note that $T_{20(1,2)}^2$ = 0.296 indicates variables 1 and 2 likely contributed to the signal.

Figure 8: MEWMA T^2 Statistic and the Reduced MEWMA Statistics



Methodology

Simulations were conducted to estimate the probability of correctly identifying the source of the MEWMA chart's signal. Consider a sequence of independent observations from a p-variate normal distribution whose mean vector shifts from $\boldsymbol{\mu}_0 = \mathbf{0}$ to $\boldsymbol{\mu}_1$ on the 16th observation, that is,

$$\begin{aligned} \mathbf{x}_i &\sim N_p(\mathbf{0}, \boldsymbol{\Sigma}), \quad i = 1, 2, \ldots, 15 \\ &\sim N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}), \quad i = 16, 17, 18 \ldots \end{aligned} \tag{13}$$

where

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.5 & 0.5 & \ldots & 0.5 \\ 0.5 & 1 & 0.5 & \ldots & 0.5 \\ 0.5 & 0.5 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 & 0.5 \\ 0.5 & 0.5 & \ldots & 0.5 & 1 \end{bmatrix}_{pxp} .$$

Forty conditions were examined using p = 3, 4, 5 and 10; five different $\boldsymbol{\mu}_1$ such that $\delta = 1$; and five different $\boldsymbol{\mu}_1$ such that $\delta = 3$. The vectors $\boldsymbol{\mu}_1$ are constructed such that (1) one variable shifts, (2) two variables shift equally, (3) two variables shift unequally, (4) three variables shift equally or (5) three variables shift unequally. Tables 4 and 5 display the conditions examined such that $\delta = 1$ and $\delta = 3$ respectively. When p = 10, approximate decimal values were used in place of exact fractions.

Results

One-Variable Deletion Analysis

If one variable shifts, a one-variable deletion is considered to be a success if the smallest reduced statistic correctly identified the variable that changed. If two variables shift, the one-variable deletion is considered to be a success if the two smallest reduced statistics correctly identify the two variables that changed. If three variables shift, the one-variable deletion is considered to be a success if the three smallest reduced statistics correctly identify the three

Table 4: Twenty Conditions Examined when δ = 1

p	1 Variable Shift	2 Variable Shift (Equal)	2 Variable Shift (Unequal)	3 Variable Shift (Equal)	3 Variable Shift (Unequal)
3	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{2/3} \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{1/2} \\ \sqrt{1/2} \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 2\sqrt{2/11} \\ \sqrt{2/11} \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{2/3} \\ \sqrt{2/3} \\ \sqrt{2/3} \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 3\sqrt{1/10} \\ 2\sqrt{1/10} \\ \sqrt{1/10} \end{bmatrix}$
4	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{5/8} \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{5/12} \\ \sqrt{5/12} \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 2\sqrt{5/32} \\ \sqrt{5/32} \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{5/12} \\ \sqrt{5/12} \\ \sqrt{5/12} \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 3\sqrt{5/68} \\ 2\sqrt{5/68} \\ \sqrt{5/68} \\ 0 \end{bmatrix}$
5	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{3/5} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{3/8} \\ \sqrt{3/8} \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 2\sqrt{3/17} \\ \sqrt{3/17} \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{1/3} \\ \sqrt{1/3} \\ \sqrt{1/3} \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 3\sqrt{1/6} \\ 2\sqrt{1/6} \\ \sqrt{1/6} \\ 0 \\ 0 \end{bmatrix}$
10	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{0.55000} \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{0.30556} \\ \sqrt{0.30556} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 2\sqrt{0.11957} \\ \sqrt{0.11957} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{0.22917} \\ \sqrt{0.22917} \\ \sqrt{0.22917} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 3\sqrt{0.04661} \\ 2\sqrt{0.04661} \\ \sqrt{0.04661} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

Table 5: Twenty Conditions Examined when δ = 3

p	1 Variable Shift	2 Variable Shift (Equal)	2 Variable Shift (Unequal)	3 Variable Shift (Equal)	3 Variable Shift (Unequal)
3	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{2} \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{3/2} \\ \sqrt{3/2} \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 2\sqrt{6/11} \\ \sqrt{6/11} \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \\ \sqrt{2} \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 3\sqrt{3/10} \\ 2\sqrt{3/10} \\ \sqrt{3/10} \end{bmatrix}$
4	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{15/8} \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{5/4} \\ \sqrt{5/4} \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 2\sqrt{15/32} \\ \sqrt{15/32} \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{5/4} \\ \sqrt{5/4} \\ \sqrt{5/4} \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 3\sqrt{15/68} \\ 2\sqrt{15/68} \\ \sqrt{15/68} \\ 0 \end{bmatrix}$
5	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{9/5} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{9/8} \\ \sqrt{9/8} \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 2\sqrt{9/17} \\ \sqrt{9/17} \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 3\sqrt{1/2} \\ 2\sqrt{1/2} \\ \sqrt{1/2} \\ 0 \\ 0 \end{bmatrix}$
10	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{1.65000} \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{0.91667} \\ \sqrt{0.91667} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 2\sqrt{0.35870} \\ \sqrt{0.35870} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} \sqrt{0.68750} \\ \sqrt{0.68750} \\ \sqrt{0.68750} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$	$\boldsymbol{\mu}_1 = \begin{bmatrix} 3\sqrt{0.13983} \\ 2\sqrt{0.13983} \\ \sqrt{0.13983} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

variables that changed. These definitions are similar to the examples shown in Figures 1-3.

MEWMA simulations were conducted using $\delta = 1$ and a one variable shift such that 10,000 out of control signals were obtained. Using $p = 3$, $h = 10.97$, and $\lambda = 0.10$, it was found that the smallest reduced MEWMA statistic correctly identified the variable that changed in 85.24% of the simulations. Using $p = 4$, $h = 12.93$, $\lambda = 0.10$, it was found that the smallest reduced MEWMA statistic correctly identified the variable that changed in 84.73% of the simulations. In addition, using $p = 5$, $h = 14.74$, and $\lambda = 0.10$, showed that the smallest reduced MEWMA statistic correctly identified the variable that changed in 83.84% of the simulations. Using $p = 10$, $h = 22.91$, and $\lambda = 0.10$, it was found that the smallest reduced MEWMA statistic correctly identified the variable that changed in 81.85% of the simulations. Simulations using $\delta = 3$ and a one variable shift produced better results such that when p = 3, 4, 5 and 10, the smallest reduced MEWMA statistic successfully identified the variable that changed in 88.53%, 87.58%, 87.18% and 86.43% of the simulations respectively.

The success rate of one-variable deletion correctly identifying the source of the signal declines as the number of variables shifting increases. For example, using $p = 3$, $h = 10.97$, $\lambda = 0.10$, $\delta = 1$ and an equal-sized two variable shift, the two smallest reduced MEWMA statistics correctly identify the two variables that shifted in 26.89% of the 10,000 simulations. The success rate rapidly declines when three variables shift and the three smallest reduced MEWMA statistics are used to identify the variables that changed. Figures 9 and 10 display the success rates of one-variable deletion when $\delta = 1$ and $\delta = 3$ respectively.

Two-Variable Deletion Analysis

If one variable shifts, the two-variable deletion is considered to be a success if the $(p-1)$ smallest reduced MEWMA statistics correctly identify the variable that changed. If two variables shift, the two-variable deletion is considered to be a success if the smallest reduced MEWMA statistics correctly identify the two variables that changed. If three variables shift, the two-variable deletion is considered to be a success if the three smallest reduced MEWMA statistics correctly identify the three variables that changed. These definitions are similar to the examples shown in Figures 5-7.

MEWMA simulations were conducted using $\delta = 1$ and a one variable shift such that 10,000 out of control signals were obtained. Using $p = 3$, $h = 10.97$, and $\lambda = 0.10$, it was found that the two smallest reduced MEWMA statistics correctly identified the variable that changed in 87.05% of the simulations. Using $p = 4$, $h = 12.93$, $\lambda = 0.10$, it was found that the three smallest reduced MEWMA statistics correctly identified the variable that changed in 77.58% of the simulations. In addition, using $p = 5$, $h = 14.74$, and $\lambda = 0.10$, showed that the four smallest reduced MEWMA statistics correctly identified the variable that changed in 76.33% of the simulations. Using $p = 10$, $h = 22.91$, and $\lambda = 0.10$, it was found that the nine smallest reduced MEWMA statistics correctly identified the variable that changed in 74.28% of the simulations. Simulations using $\delta = 3$ and a one variable shift produced similar or better results such that when p = 3, 4, 5 and 10, the $(p-1)$ smallest reduced MEWMA statistics successfully identified the variable that changed in 85.61%, 81.77%, 84.61% and 80.38% of the simulations respectively.

The success rate of two-variable deletion correctly identifying the source of the signal decreases when two variables shift. However, the decrease is not as pronounced as the one-variable deletion. For example, using $p = 3$, $h = 10.97$, $\lambda = 0.10$, $\delta = 1$ and an equal-sized two variable shift, the smallest reduced MEWMA statistic correctly identifies the two variables that shifted in 77.02% of the 10,000 simulations. Figures 11 and 12 display the success rates of two variable deletion when $\delta = 1$ and $\delta = 3$ respectively. In addition, there tends to be a slight decrease in the success rate when comparing an unequal shift to an equal shift. The success rate rapidly declines when three variables shift and the three smallest reduced MEWMA statistics are used to identify the variables that changed.

Figure 9: Success Rate of One-Variable Deletion when δ = 1

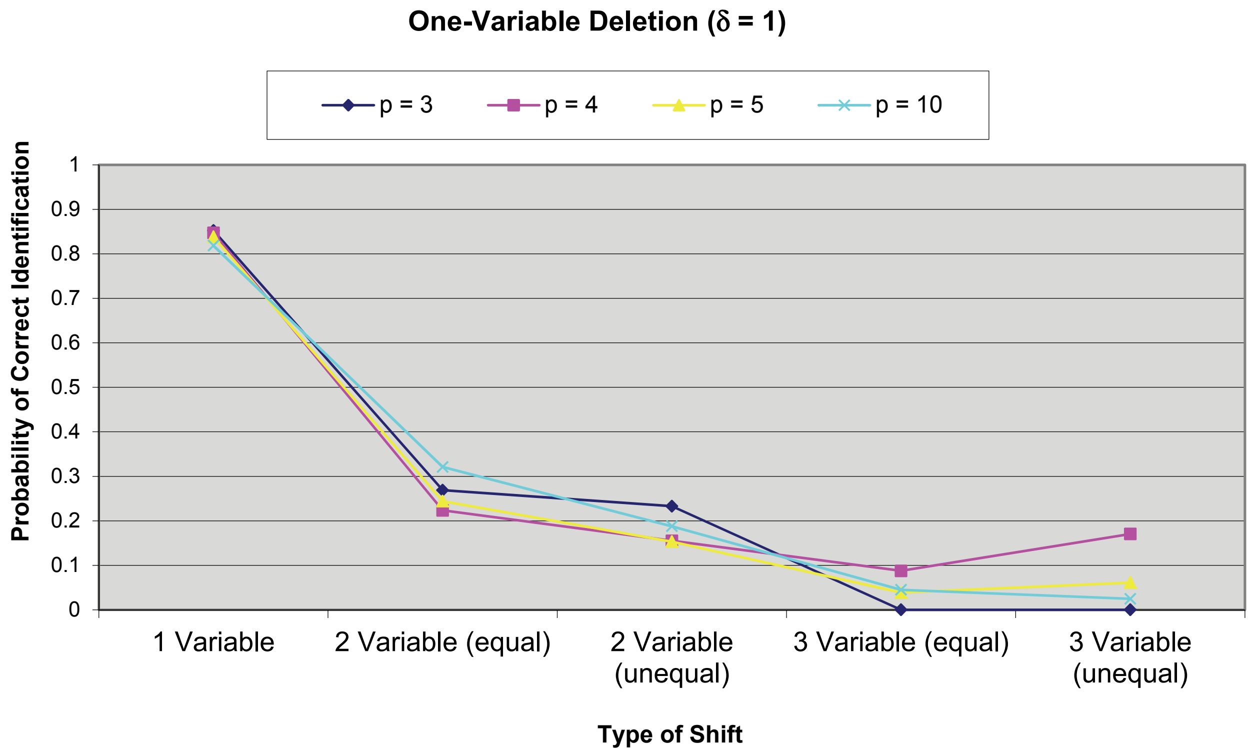


Figure 10: Success Rate of One-Variable Deletion when δ = 3

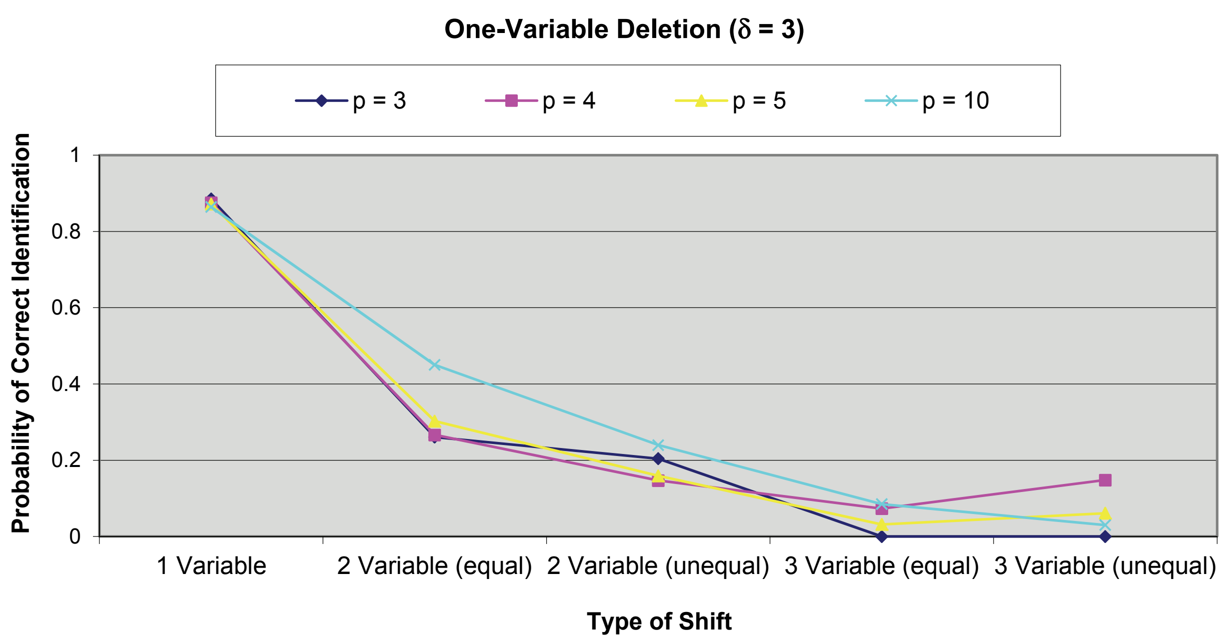


Figure 11: Success Rate of Two-Variable Deletion when δ = 1

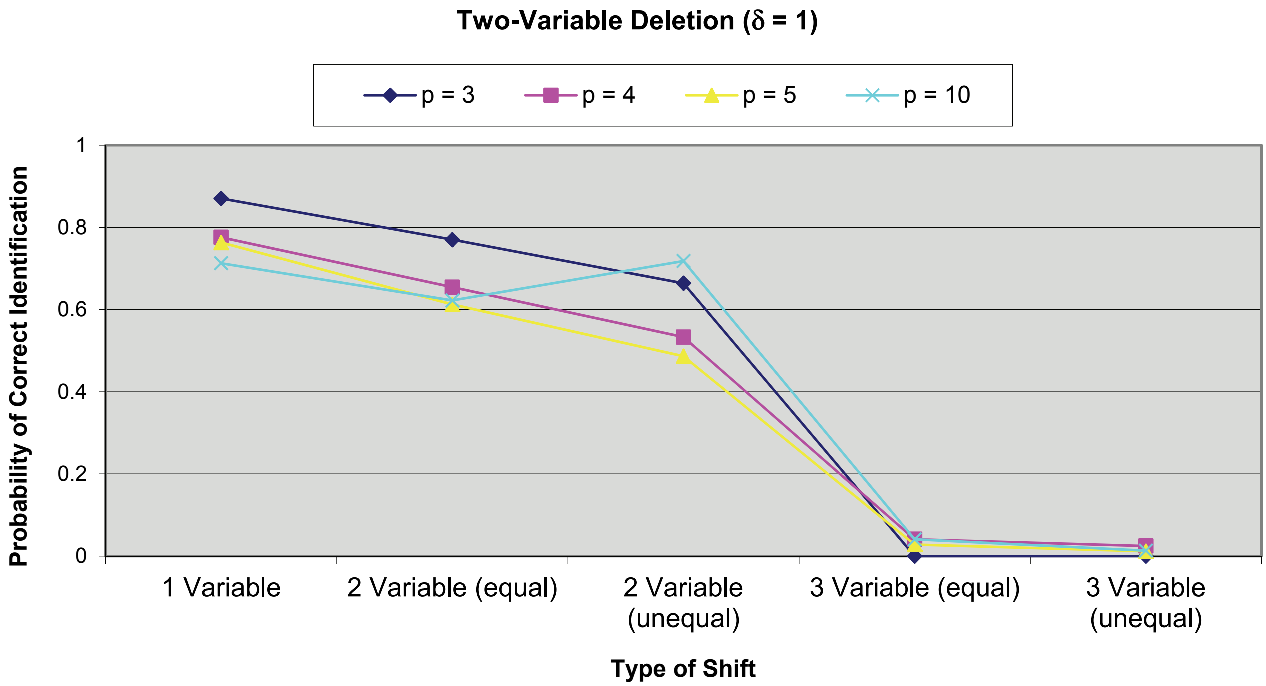
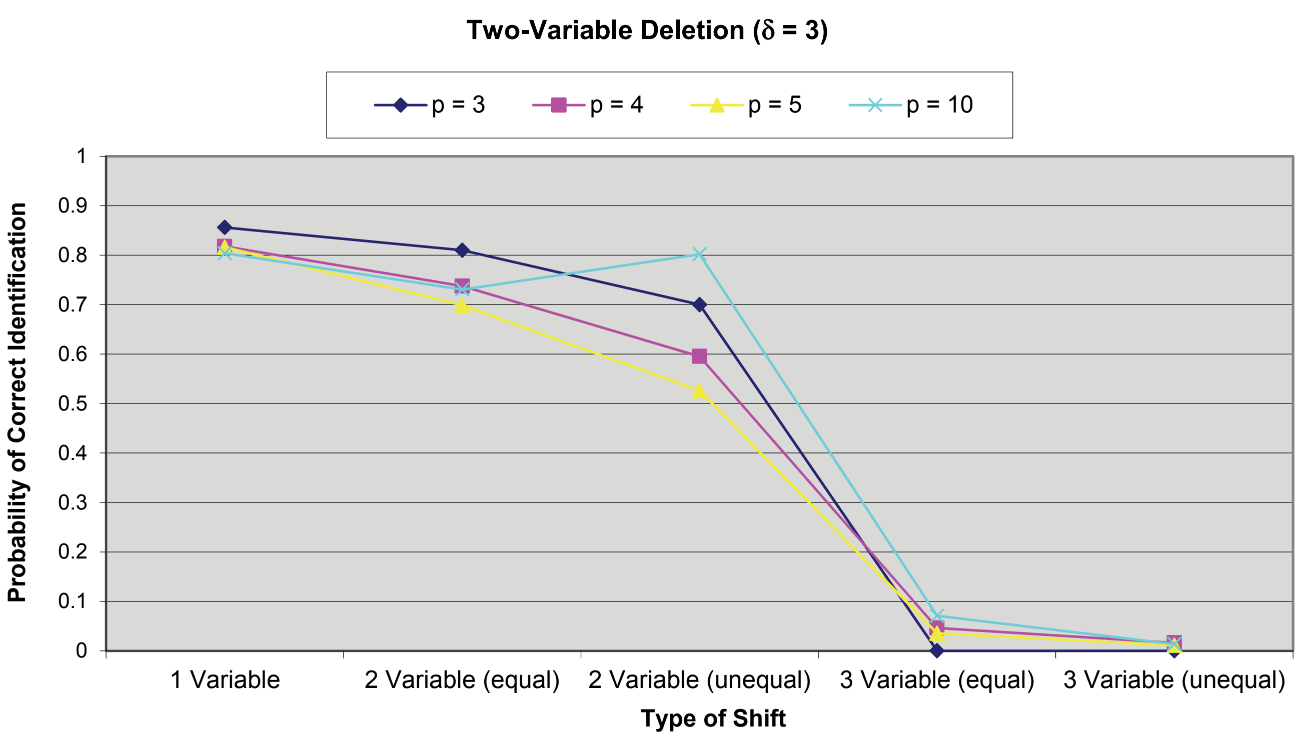


Figure 12: Success Rate of Two-Variable Deletion when δ = 3



Conclusion

A general procedure for identifying the variables that contribute to the signal of the MEWMA chart was presented. One-variable deletion correctly identified a one variable shift in 82-89% of the simulations. Two-variable deletion correctly identified a one variable shift in 71-87% of the simulations, an equal-sized two variable shift in 61-81% of the simulations, and an unequal-sized two variable shift in 49-80% of the simulations.

The success rate decreases rapidly when more variables shift than are removed from the MEWMA statistic. However, examining the reduced MEWMA statistics indicated that the criteria employed herein for a successful identification may not immediately identify the variables that contributed to the signal; however, they did lead to a significantly reduced set of variables to search for the cause of the signal.

This study used only one defined covariance matrix such that the correlation between each pair of variables was 0.5. It is suspected that an increase in the success rates would be observed if the correlation between the variables is small. In several of the simulations it was noted that, when a variable would shift, it would drag other variables along with it. This in turn clouded the reduced MEWMA statistics making it more difficult to identify the variable that changed using the previously discussed definitions of a success. Further study is required using different covariance matrices.

In addition, the reported success rates assumed if q-variables shifted, then the corresponding definition of a success was used. Further study is required to examine the success rates using various definitions of a success. One suggestion might be that critical values be established to indicate to the operator that a reduced MEWMA statistic is significantly small. Additional simulations should examine the entire distribution of the reduced MEWMA statistics. Critical values could be obtained by examining the distribution of the reduced MEWMA statistics whose variables had not shifted.

Given the power of today's modern computers, variable deletion could be extended to more than two variables being removed from the calculations. Computers could provide a sequential method of analysis in which an operator examines one variable deletion results, then two-variable deletion results, three-variable deletion results, etc. Using such a method, it is anticipated that a reasonable success rate for identifying a q-variable shift using a q-variable deletion would be determined. However, this success rate would likely decrease as more variables are added to the process. Additional research is required in this area.

Using variable deletion in conjunction with the MEWMA control chart should enable a user to employ an efficient multivariate control chart with an effective post hoc analysis. In addition, it provides a helpful and easy to understand graphical solution to the problem of identifying which variable(s) contributed to the signal.

References

Allen, D. M. (1971). Mean square error of prediction as a criterion for selecting variables. *Technometrics*, *13*, 469-475.

Alt, F. B. (1984). Multivariate quality control. In *The encyclopedia of statistical sciences*, S. Kotz, N. L. Johnson and C. R. Read, Eds., 110-112. New York, NY: John Wiley.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980), *Regression diagnostics: Identifying influential data and sources of collinearity*. New York, NY: Wiley.

Crossier, R. B. (1988). Multivariate generalization of cumulative sum quality-control schemes. *Techonometrics*, *30*, 291-303.

Hotelling, H. (1947). Multivariate quality control illustrated by the air testing of sample bombsights. In *Techniques of statistical analysis*, C. Eisenhart, M. W. Hastay, and W. A. Wallis, Eds., 111-184. New York, NY: McGraw-Hill.

Jackson, J. E. (1980). Principal components and factor analysis: Part I-principal components. *Journal of Quality Technology*, *12*, 201-213.

Jackson, J. E. (1985). Multivariate quality control. *Communications in Statistics – Theory and Methods*, *14*, 2657-2688.

Jackson, J. E. (1991). *A user's guide to principal components*. New York, NY: John Wiley & Sons, Inc.

Lowry, C. A., Woodall, W. H., Champ, C. W. & Rigon, S. E. (1992). A multivariate exponentially weighted moving average control chart. *Technometrics*, *34*, 46-53.

Lowry, C. A., & Montgomery, D. C. (1995). Review of multivariate control charts. *IIE Transactions*, *27*, 800-810.

Mason, R. L., Tracy, N. D., & Young, J. C. (1995). Decomposition of T^2 for multivariate control chart interpretation. *Journal of Quality Technology*, *27*, 99-108.

Mason, R. L., Champ, C. W., Tracy, N. D., Wierda, S. J., & Young, J. C. (1997). Assessment of multivariate process control techniques. *Journal of Quality Technology*, *29*, 140-143.

Pignatiello, J. J., & Runger, G. C. (1990). Comparison of multivariate CUSUM charts. *Journal of Quality Control*, *22*, 173-186.

Wierda, S. J. (1994). Multivariate statistical process control: Recent results and directions for future research. *Statistica Neerlandica*, *48*, 147-168

Woodall, W. H., & Montgomery, D. C. (1999). Research issues and ideas in statistical process control. *Journal of Quality Technology*, *31*, 376-386.