# A Power Analysis of Variable Deletion Within the MEWMA Control Chart Statistic

Jay R. Schaffer & Shawn VandenHul

University of Northern Colorado
McKee 530
Greeley, CO 80639

jay.schaffer@unco.edu
gathen9@hotmail.com

ABSTRACT

One of the drawbacks to multivariate charting schemes is their inability to identify which variable was the source of the signal. The multivariate exponentially weighted moving average (MEWMA) developed by Lowry, et al (1992) is an example of a multivariate charting scheme whose monitoring statistic is unable to determine which variable caused the signal. The purpose of this paper is to examine the use of variable deletion within the MEWMA control chart statistic. By deleting variables systematically, the re-calculated MEWMA statistic is reduced. Critical values are developed to determine if the reduction is statistically significant. A power analysis is performed to determine the likelihood of a correct identification. In addition, a two-variable deletion scheme is also studied.

## INTRODUCTION

With today's use of 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 the variables being monitored. Alt

(1984), Jackson (1985), Wierda (1994), Lowry and Montgomery (1995), and Mason et al. (1997) discuss much of the literature on this topic.

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 better than Hotelling's (1947) $\chi^2$ chart when detecting a shift in the mean vector of a multivariate normal distribution.

Woodall and Montgomery (1999) discussed 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 (see Jackson (1985)) using principle component analysis. Mason et al. (1995) proposed to decompose Hotelling's $T^2$ statistic by removing individual variables from its calculation. Woodall and Montgomery (1999) discussed that more work is needed on graphical methods for data visualization when interpreting signals from multivariate control charts.

The purpose of this paper is to present a graphical approach to identify the source of a signal from the MEWMA control chart. This paper will examine 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. This paper deletes variables in a multivariate process instead of deleting individual observations in a data set. In addition, this paper will estimate the probability of correctly identifying the source using various simulations.

## MEWMA CHART

Suppose we have 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 $r^{th}$ observation, that is,

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

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 paper examines removing 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 source of the signal.

ONE-VARIABLE DELETION

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

$\mathbf{x}_i' = (x_{i1}, x_{i2}, \ldots, x_{i,j-1}, x_{i,j+1}, \ldots, x_{ip}, x_{ij}) = (\mathbf{x}_{i(j)}', x_{ij})$, where $\mathbf{x}_{i(j)}'$ is a (p-1)x1 vector excluding the j^th^ variable. In addition, let $\mathbf{S}_{(j)}$ be the (p-1)x(p-1) principal sub-matrix of $\mathbf{S}$ 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 $\mathbf{z}_{0(j)} = \mathbf{0}$ and

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

and

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

The calculation of $T_{i(j)}^2$ is continued 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 the user identifying the cause of the signal. The smallest reduced

MEWMA statistic may indicate which variable contributed to the signal. For example, if the 1st variable shifts, the reduced MEWMA statistics may resemble Figure 1.

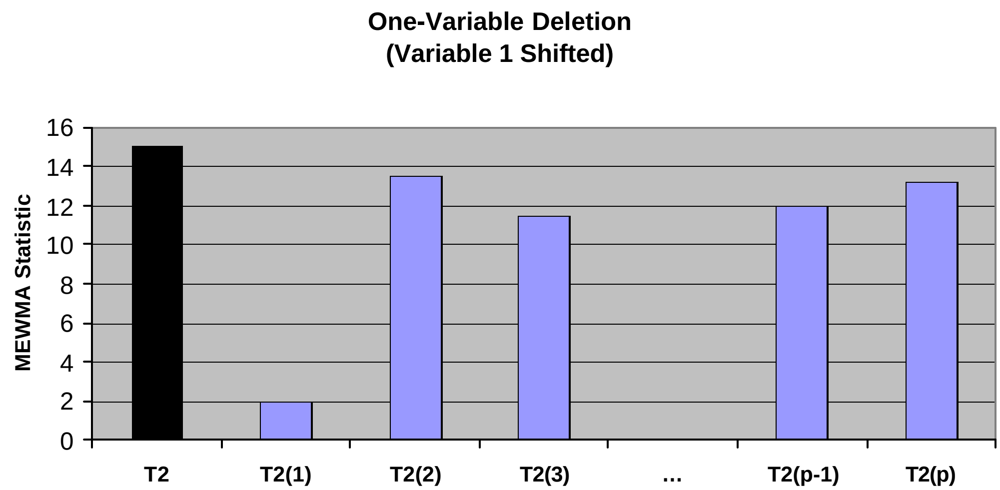


Figure 1. A General Representation of the Reduced MEWMA Statistics if Variable 1 Shifted.

A similar analysis is required if more than two variables change. For example, the reduced MEWMA statistics may resemble Figure 2, if the 1st and the 2nd variables shift or may resemble Figure 3, if the 1st, 2nd, and 3rd variables shift.

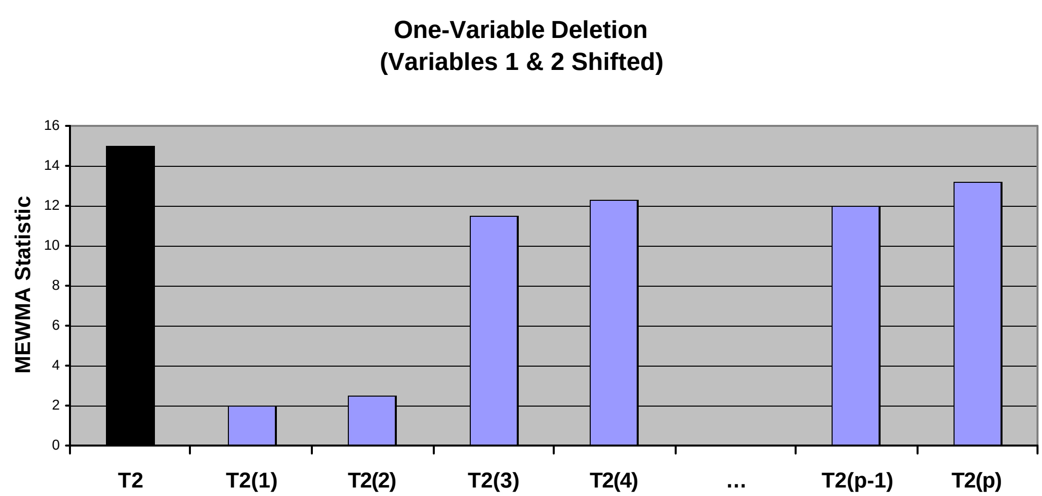


Figure 2. A General Representation of the Reduced MEWMA Statistics if Variables 1 and 2 Shifted.

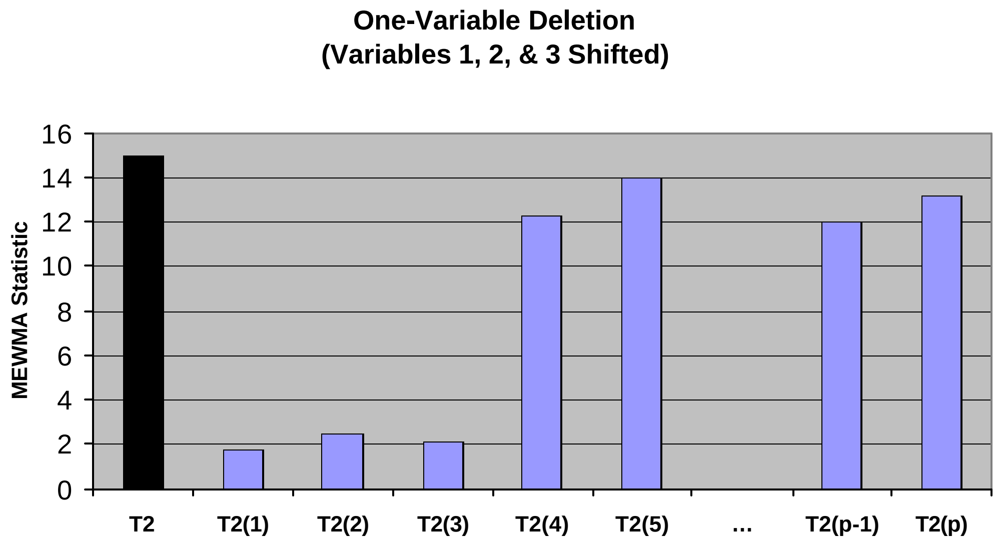


Figure 3. A General Representation of the Reduced MEWMA Statistics if Variables 1, 2, and 3 Shifted.

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

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

such that:

$$\boldsymbol{\mu}_{\mathbf{o}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{\mu}_{\mathbf{1}} = \begin{bmatrix} \sqrt{2} \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{S} = \begin{bmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{bmatrix}$$

The reader will note a shift of $d = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \mathbf{S}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)} = 3$ occurred on the $16^{th}$ sample. Table 1 displays a data simulation of the above conditions along with the corresponding MEWMA statistics, $T_i^2$. Using $l = 0.10$, and h = 10.97 (in-control ARL = 200), the MEWMA chart signaled on the $21^{st}$ 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.

| 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 1: Simulated Process With Corresponding MEWMA Statistics, $T_i^2$.

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 will be:

$$\mathbf{S}_{(1)} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}.$$

Using Equations (5)-(7), the reduced MEWMA statistic, $T_{i(1)}^2$, is calculated for i = 1, 2, …, 21 and displayed in Table 2. The reader will note on the 21$^{st}$ sample, $T_{21(1)}^2$ = 0.9358 represents the reduced MEWMA statistic with the contribution of the first variable removed.

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

Table 2: Reduced MEWMA Statistics, $T^2_{i\,(1)}$.

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

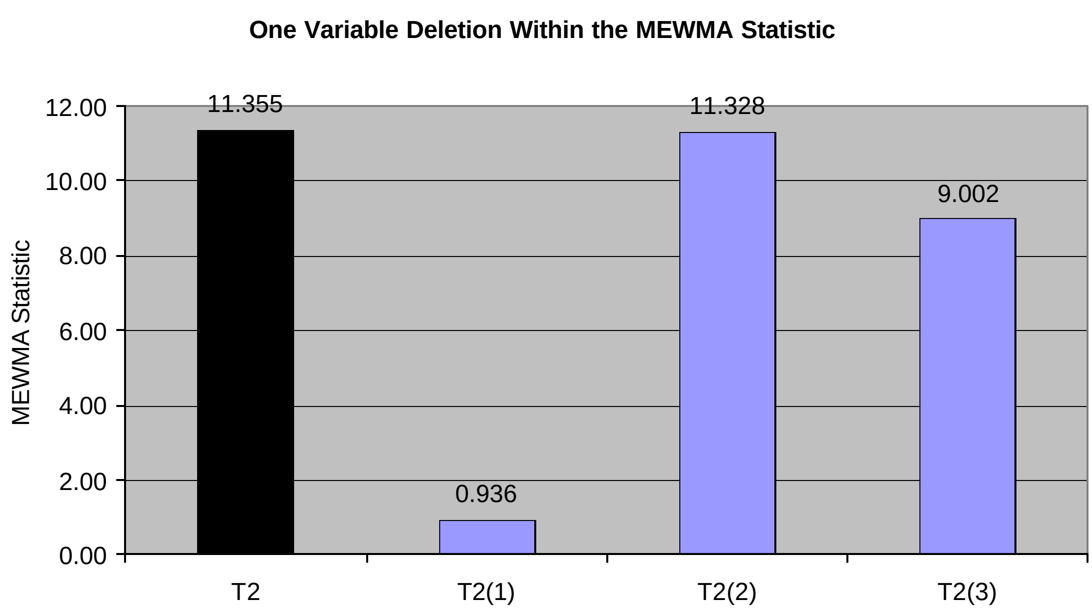


Figure 4: MEWMA $T^2$ Statistic and the Three Reduced MEWMA Statistics.

TWO-VARIABLE DELETION

Assume on the s$^{th}$ sample, the MEWMA chart signaled a change ($T_s^2 > h$). The p variables are now removed, two at a time, from the calculation of $T_s^2$. Assume the j$^{th}$ and k$^{th}$ variables are to be removed. Now let

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

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

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

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

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

and

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

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

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 aid the user 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 1$^{st}$ 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.

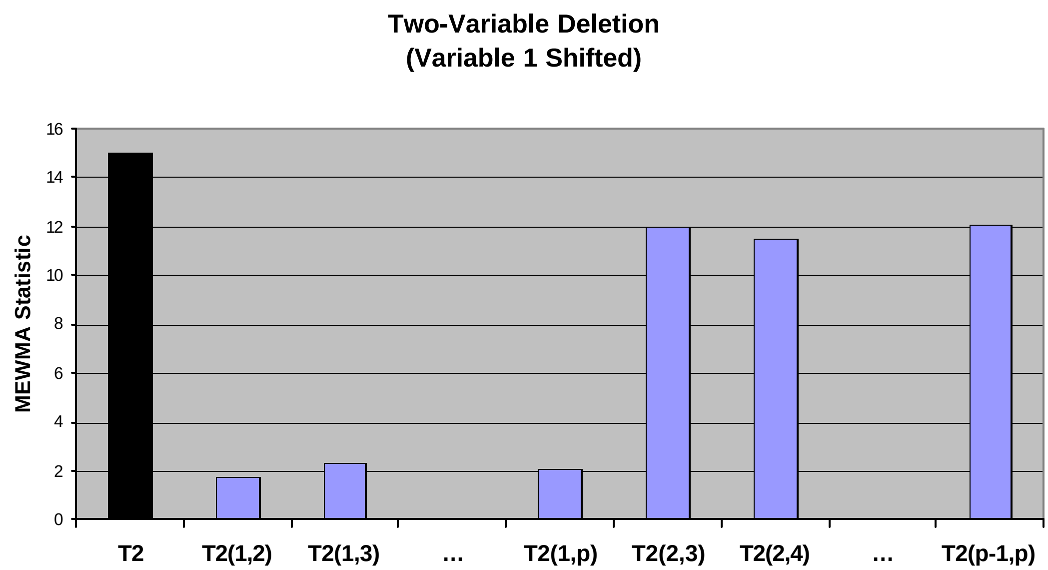


Figure 5. A General Representation of the Reduced MEWMA Statistics if Variable 1 Shifted.

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 1$^{st}$ and the 2$^{nd}$ variables shift. A similar analysis is required if three variables have shifted. The reduced MEWMA statistics may resemble Figure 7, if the 1$^{st}$, 2$^{nd}$, and 3$^{rd}$ variables shift.

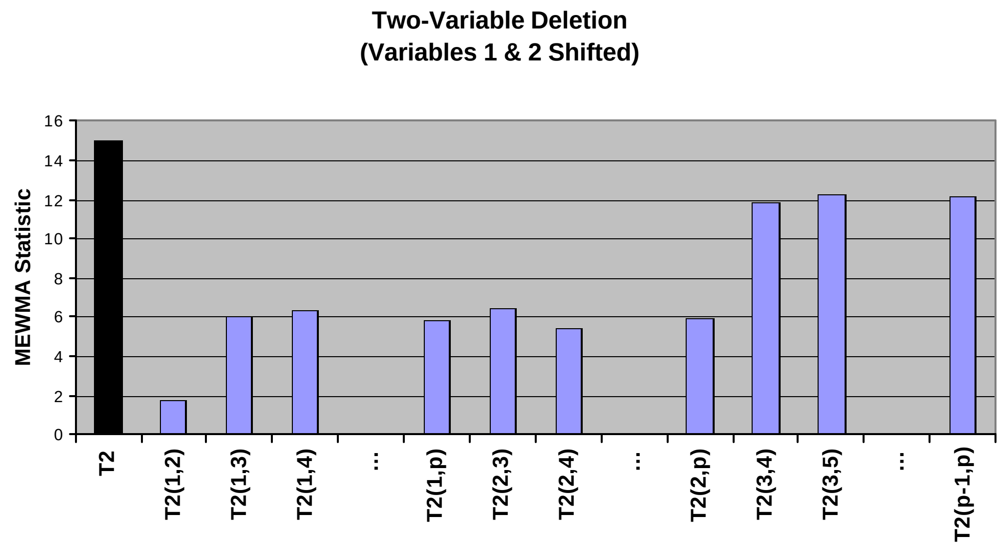


Figure 6. A General Representation of the Reduced MEWMA Statistics if Variables 1 and 2 Shifted.

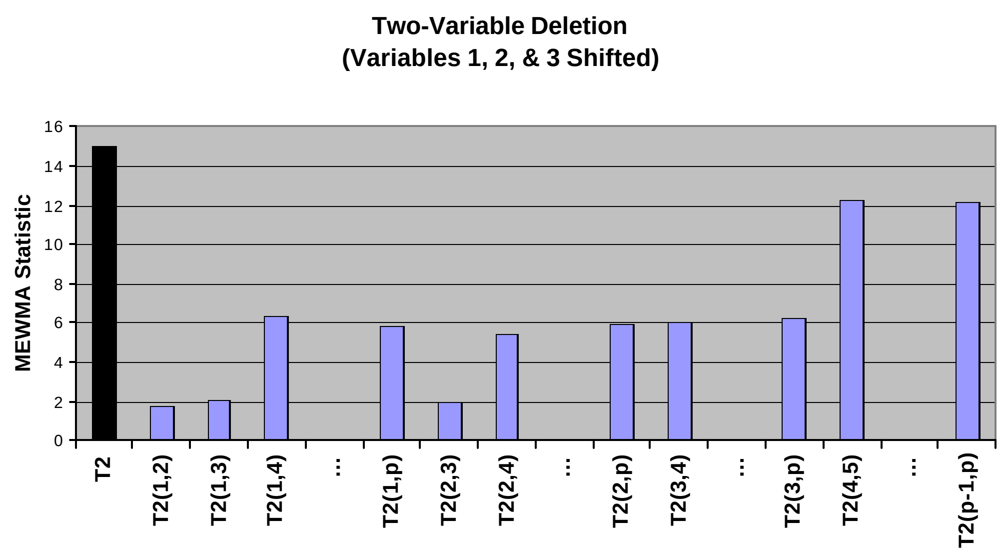


Figure 7. A General Representation of the Reduced MEWMA Statistics if Variables 1, 2, and 3 Shifted.

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

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

such that:

$$\boldsymbol{\mu}_\mathbf{o} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{\mu}_\mathbf{1} = \begin{bmatrix} \sqrt{3/2} \\ \sqrt{3/2} \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{S} = \begin{bmatrix} 1 & 0.5 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 & 0.5 \\ 0.5 & 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1 \end{bmatrix}.$$

The reader will note a shift of $d = \sqrt{(\boldsymbol{\mu}_\mathbf{1} - \boldsymbol{\mu}_\mathbf{0})' \mathbf{S}^{-1} (\boldsymbol{\mu}_\mathbf{1} - \boldsymbol{\mu}_\mathbf{0})}$ = 3 occurred on the 16th sample. Table 3 displays a data simulation of the above conditions along with the corresponding MEWMA $T_i^2$ statistics. Using $l = 0.10$ and h = 12.93 (in-control ARL = 200), the MEWMA chart signaled on the 20th observation such that $T_{20}^2$ = 13.793.

| 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 |

Table 3. Simulated Process With Corresponding MEWMA Statistics, $T_i^2$.

Using Equations (9)-(11), the reduced MEWMA statistics are $T^2_{20(1,2)}$ = 0.296, $T^2_{20\,(1,3)}$ = 4.771, $T^2_{20\,(1,4)}$ = 5.213, $T^2_{20\,(2,3)}$ = 10.481, $T^2_{20\,(2,4)}$ = 11.246, and $T^2_{20\,(3,4)}$ = 9.674. Figure 8 displays the reduced MEWMA statistics. The reader will note $T^2_{20\,(1,2)}$ = 0.296 indicates variables 1 and 2 likely contributed to the signal.

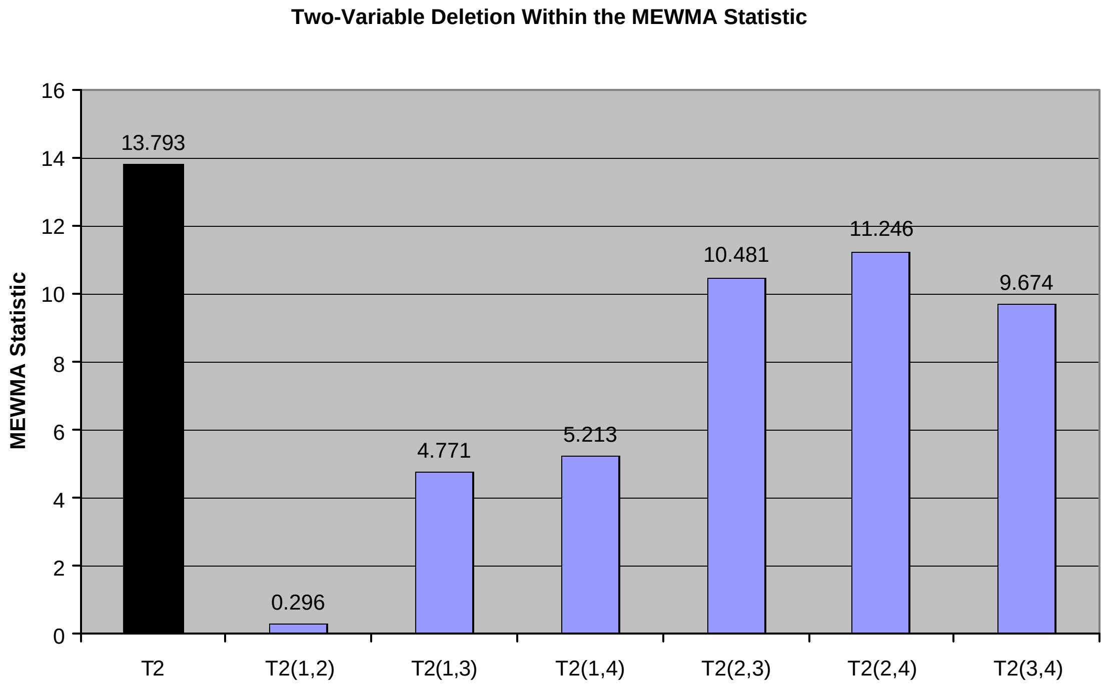


Figure 8. MEWMA $T^2$ Statistic and the Reduced MEWMA Statistics.

CRITICAL VALUES

Along with the variable deletion technique discussed above, an operator will need to be able to determine if the reduced MEWMA statistics are significantly small. Simulations were conducted to identify a suitable critical value (CV). 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,

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

where

$$\mathbf{S} = \begin{bmatrix} 1 & r & r & \cdots & r \\ r & 1 & r & \cdots & r \\ r & r & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 & r \\ r & r & \cdots & r & 1 \end{bmatrix}_{pxp} .$$

Eighty conditions were examined using p = 3, 4, 5, and 10; ρ = 0.5 and 0.8; five different $\boldsymbol{\mu}_1$ such that δ = 1; and five different $\boldsymbol{\mu}_1$ such that δ = 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 when ρ = 0.5 such that δ = 1 and δ = 3 respectively. When p = 10, approximate decimal values were used in place of cumbersome exact fractions. Tables for ρ = 0.8 have been excluded, but are constructed in a similar manner as Tables 4 and 5.

| 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 | $\mu_1 = \begin{bmatrix} \sqrt{0.55000} \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} \sqrt{0.30556} \\ \sqrt{0.30556} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} 2\sqrt{0.11957} \\ \sqrt{0.11957} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} \sqrt{0.22917} \\ \sqrt{0.22917} \\ \sqrt{0.22917} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} 3\sqrt{0.04661} \\ 2\sqrt{0.04661} \\ \sqrt{0.04661} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ |

Table 4. Twenty Conditions Examined When $\delta = 1$.

| p | 1 variable shift | 2 variable shift (equal) | 2 variable shift (unequal) | 3 variable shift (equal) | 3 variable shift (unequal) |
|---|---|---|---|---|---|
| 3 | $\mu_1 = \begin{bmatrix} \sqrt{2} \\ 0 \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} \sqrt{3/2} \\ \sqrt{3/2} \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} 2\sqrt{6/11} \\ \sqrt{6/11} \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \\ \sqrt{2} \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} 3\sqrt{3/10} \\ 2\sqrt{3/10} \\ \sqrt{3/10} \end{bmatrix}$ |
| 4 | $\mu_1 = \begin{bmatrix} \sqrt{15/8} \\ 0 \\ 0 \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} \sqrt{5/4} \\ \sqrt{5/4} \\ 0 \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} 2\sqrt{15/32} \\ \sqrt{15/32} \\ 0 \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} \sqrt{5/4} \\ \sqrt{5/4} \\ \sqrt{5/4} \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} 3\sqrt{15/68} \\ 2\sqrt{15/68} \\ \sqrt{15/68} \\ 0 \end{bmatrix}$ |
| 5 | $\mu_1 = \begin{bmatrix} \sqrt{9/5} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} \sqrt{9/8} \\ \sqrt{9/8} \\ 0 \\ 0 \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} 2\sqrt{9/17} \\ \sqrt{9/17} \\ 0 \\ 0 \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} 3\sqrt{1/2} \\ 2\sqrt{1/2} \\ \sqrt{1/2} \\ 0 \\ 0 \end{bmatrix}$ |
| 10 | $\mu_1 = \begin{bmatrix} \sqrt{1.65000} \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} \sqrt{0.91667} \\ \sqrt{0.91667} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} 2\sqrt{0.35870} \\ \sqrt{0.35870} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} \sqrt{0.68750} \\ \sqrt{0.68750} \\ \sqrt{0.68750} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ | $\mu_1 = \begin{bmatrix} 3\sqrt{0.13983} \\ 2\sqrt{0.13983} \\ \sqrt{0.13983} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ |

Table 5. Twenty Conditions Examined When $\delta = 3$.

CRITICAL VALUES: ONE VARIABLE DELETION

Ten thousand simulations were conducted, for each of the previously discussed 80 conditions, to approximate the distributions of the reduced MEWMA statistics. Consider a process containing p = 5 variables, in which variable 1 has shifted and the MEMWA chart has signaled. Figure 9 displays the results of 10,000 reduced MEWMA statistics when $\rho = 0.5$, and $\delta = 3$. Using the distributions of $T^2_{(2)}$, $T^2_{(3)}$, … $T^2_{(5)}$, which will be identical, a CV can be obtained by finding the 5th percentile of the distribution. In this particular case, the 5th percentile = CV = 9.6933.

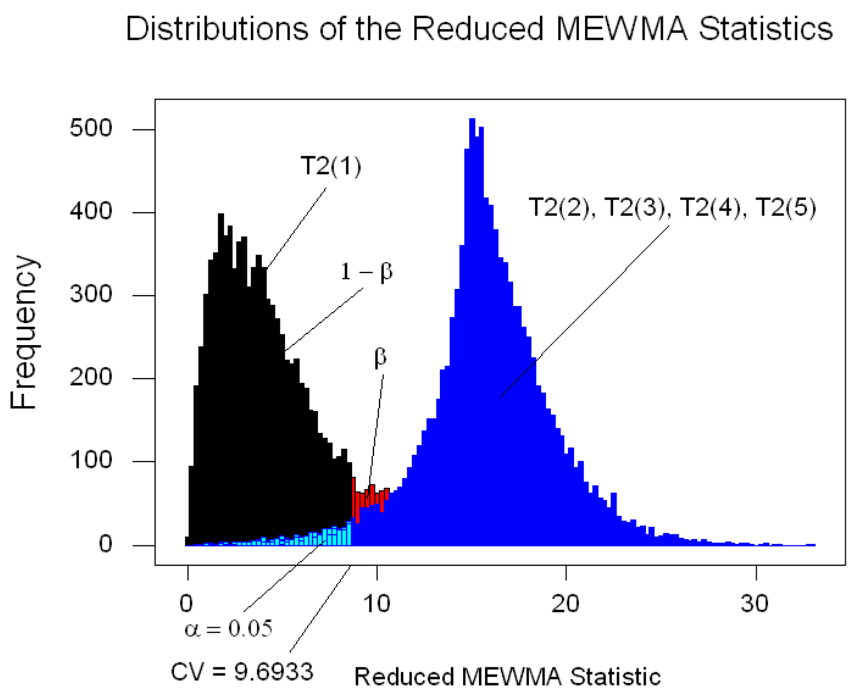


Figure 9. Distributions of the Reduced MEWMA Statistics.

Unfortunately, the CV changes for each of the eighty conditions. For example, when p = 5, $\rho = 0.5$, and $\delta = 3$, the CV = {9.6933, 8.6806, 8.7078, 7.1151, 8.3502} when one variable shifts, two variables shift equally, two variables shift unequally, three variables shift equally, and three variables shift unequally, respectively.

In an effort to obtain a single CV that an operator may employ in practice, the 20 CV's for $p = 5$ were averaged. The average CV would take into account the different correlations within the process ($\rho = 0.5$ and 0.8), size of shift ($\delta = 1$ and 3), and 5 types of shift. In addition, the average CV would be approximately at the 5 percent level of significance, $\alpha$.

When $p = 5$, the average CV = 8.3561. In a similar manner, the average CV's were calculated when $p = 3$, 4, and 10. The average CV's = {2.7435, 5.0600, 18.9512} respectively. Reviewing the previously discussed example in Figure 4, the operator can conclude that it is likely that variable 1 caused the MEWMA chart to signal since the reduced MEWMA statistic $T^2_{(1)} = 0.936 < 2.7435 = CV$.

CRITICAL VALUES: TWO VARIABLE DELETION

Ten thousand simulations were conducted, for each of the previously discussed 80 conditions, to approximate the distributions of the reduced MEWMA statistics when two variables are deleted. Consider a process containing $p = 10$ variables such that $\rho = 0.5$, and $\delta = 3$, and variable 1 has shifted causing the MEMWA chart to signal. Using the simulated distributions of $T^2_{(2,3)}$, $T^2_{(2,4)}$, … $T^2_{(9,10)}$, which will be identical, a CV can be obtained by finding the $5^{th}$ percentile of the distribution. In this particular case, the $5^{th}$ percentile = CV = 16.9041.

As before, the CV changes for each of the eighty conditions. For example, when $p = 10$, $\rho = 0.5$, and $\delta = 3$, the CV = {16.9041, 16.4368, 16.6303, 15.5254, 15.8896} when one variable shifts, two variables shift equally, two variables shift unequally, three variables shift equally, and three variables shift unequally, respectively.

In an effort to obtain a single CV that an operator may employ in practice, the 20 CV's for $p = 10$ were averaged. The average CV would take into account the different correlations within the process ($\rho = 0.5$ and $0.8$), size of shift ($\delta = 1$ and $3$), and 5 types of shift. In addition, the average CV would be approximately at the 5 percent level of significance, $\alpha$.

When $p = 10$, the average CV =15.8410. In a similar manner, the average CV's were calculated when $p = 3, 4,$ and $5$. The average CV's = {0.8805, 2.6408, 4.7210} respectively.

## ONE-VARIABLE DELETION ANALYSIS

Traditionally, in hypothesis testing, power is defined as the probability of rejecting $H_0$ when $H_0$ is false. Power can be calculated by comparing the distribution of the test statistic when $H_0$ is false to the CV created through an examination of the distribution of the test statistic when $H_0$ is true. In a similar manner, define $\alpha = P(T^2_{(i)} < CV$ when variable i has not shifted) and $\beta = P(T^2_{(i)} > CV$ when variable i has shifted). Thus, power could be calculated by $1 - \beta = P(T^2_{(i)} < CV$ when variable i has shifted). Figure 9 displays a graphical representation of how power is computed when $p = 5$ and the first variable has shifted.

A power analysis was conducted to examine the effectiveness of the one-variable deletion technique. The previously discussed simulated distributions and average CV's were used to calculate power. Figure 10 displays the power for variables 1 and 2 (i.e. the probability that $T^2_{(i)} <$ average CV = 2.7435 when variable i has shifted) in a 3 variable process. In addition, the levels of significance, $\alpha$, using the average CV, are also displayed for each condition. The reader will note that simulations were not conducted when all three variables shift due to the fact that no "null" distribution would be available to determine CV's. From Figure 10, the reader can see that if $\delta = 3$, $\rho = 0.5$, and variable one shifts, then $P(T^2_{(1)} < \text{average CV}) = 0.6524$ with $\alpha = 0.0157$.

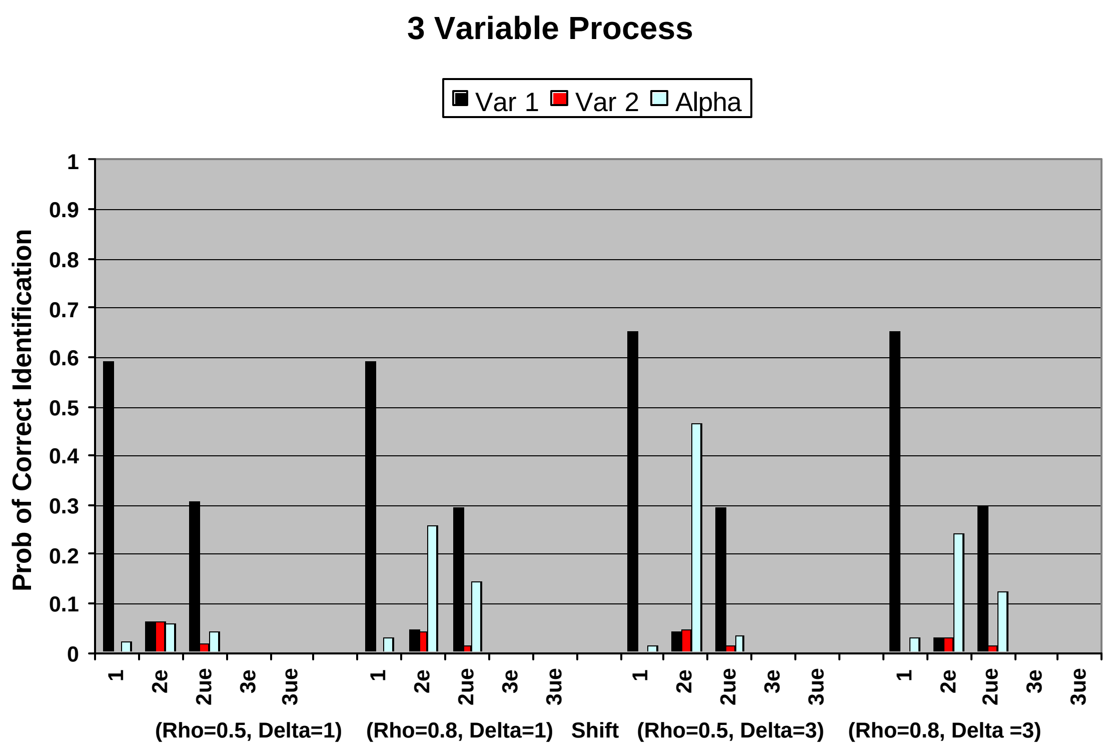


Figure 10. Power and $\alpha$ Levels for a 3 Variable Process.

Figure 11 displays the power for variables 1,2, and 3 (i.e. the probability that $T^2_{(i)} < \text{average CV} = 5.0600$ when variable i has shifted) in a 4 variable process. In addition, the levels of significance, $\alpha$, using the average CV, are also displayed for each condition. From Figure 11, the reader can see that if $\delta = 3$, $\rho = 0.5$, and variable one shifts, then $P(T^2_{(1)} < \text{average CV}) = 0.7269$ with $\alpha = 0.0158$.

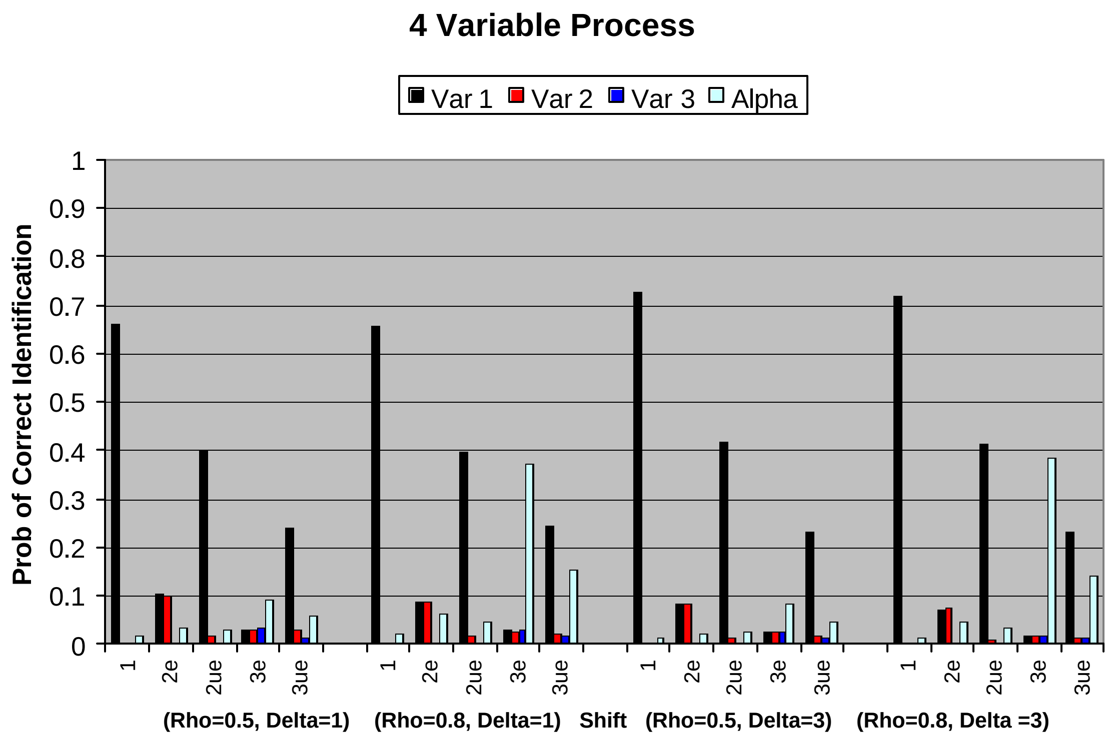


Figure 11. Power and α Levels for a 4 Variable Process.

Figure 12 displays the power for variables 1,2, and 3 (i.e. the probability that $T^2_{(i)}$ < average CV = 8.3561 when variable i has shifted) in a 5 variable process. In addition, the levels of significance, α, using the average CV, are also displayed for each condition. From Figure 12, the reader can see that if $\delta = 3$, $\rho = 0.5$, and variable one shifts, then $P(T^2_{(1)} < \text{average CV}) = 0.8459$ with $\alpha = 0.0200$.

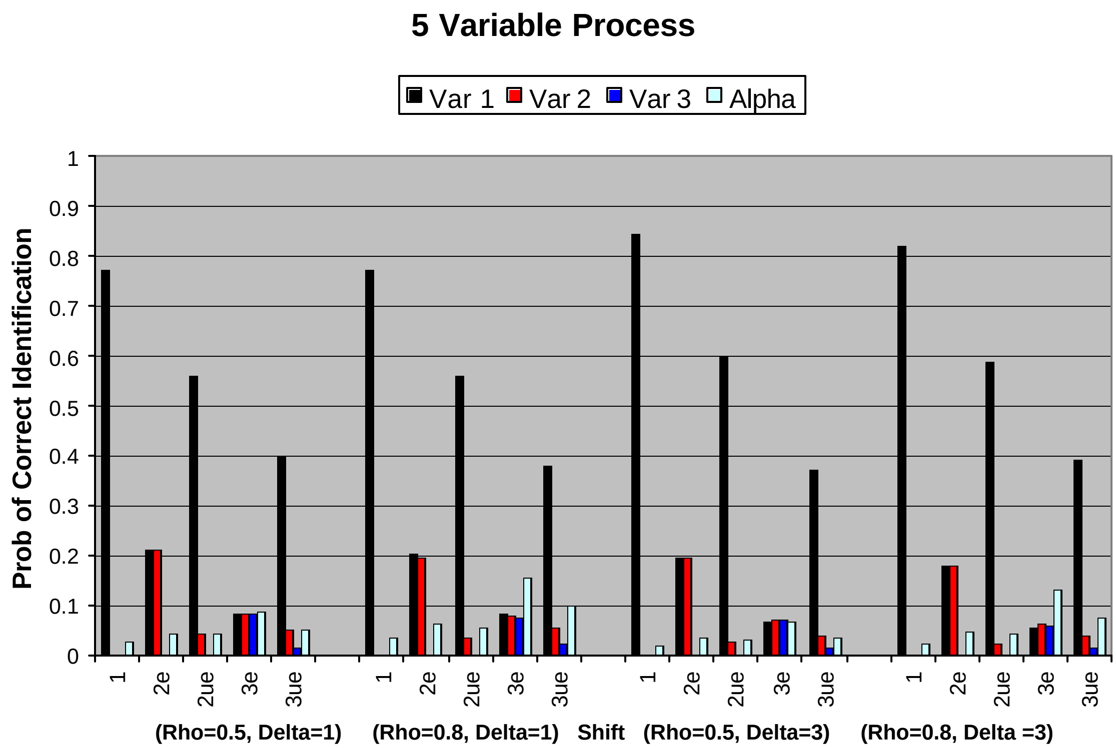


Figure 12. Power and α Levels for a 5 Variable Process.

Figure 13 displays the power for variables 1,2, and 3 (i.e. the probability that $T^2_{(i)}$ < average CV = 18.9512 when variable i has shifted) in a 10 variable process. In addition, the levels of significance, α, using the average CV, are also displayed for each condition. From Figure 13, the reader can see that if $\delta = 3$, $\rho = 0.5$, and variable one shifts, then $P(T^2_{(1)} <$ average CV$) = 0.8795$ with $\alpha = 0.0338$. The reader will also note that if $\delta = 3$, $\rho = 0.5$, and variable one and two shift equally, then $P(T^2_{(1)} <$ average CV$) = 0.4643$ and $P(T^2_{(2)} <$ average CV$) = 0.5028$ with $\alpha = 0.0431$. In addition, if $\delta = 3$, $\rho = 0.5$, and variable one and two shift unequally, then $P(T^2_{(1)} <$ average CV$) = 0.8049$ and $P(T^2_{(2)} <$ average CV$) = 0.1262$ with $\alpha = 0.0378$.

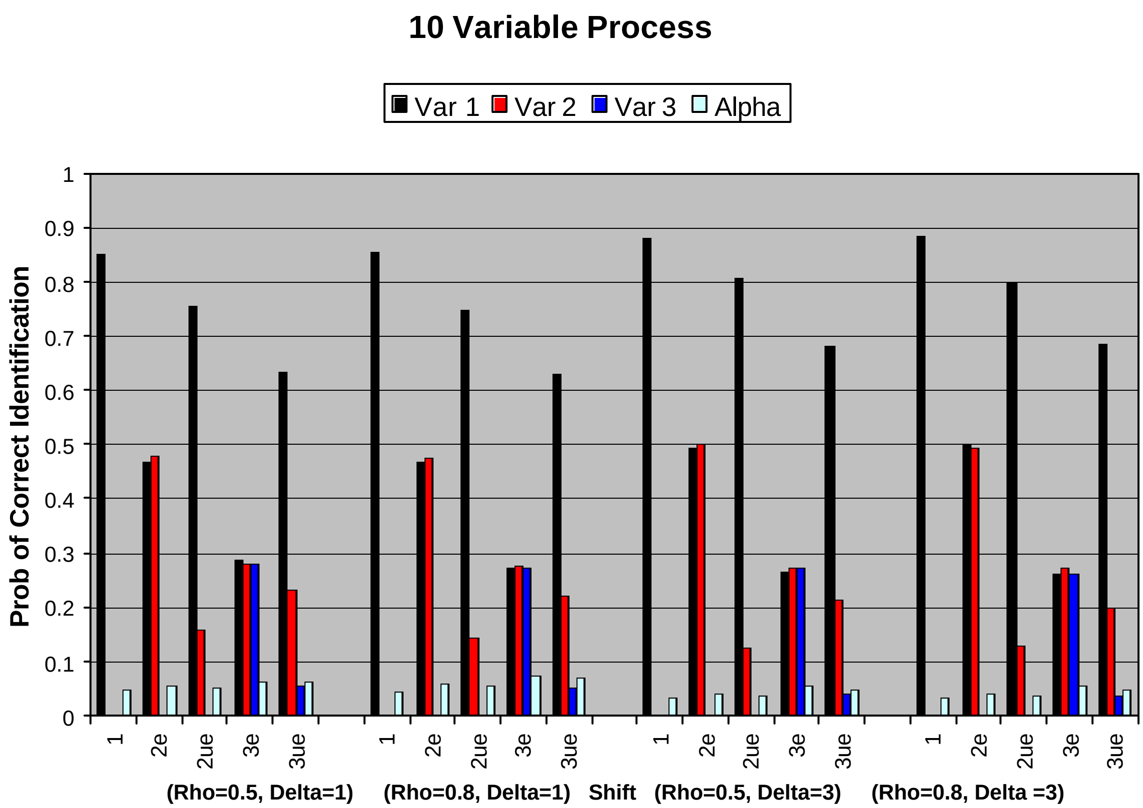


Figure 13. Power and $\alpha$ Levels for a 10 Variable Process.

TWO-VARIABLE DELETION ANALYSIS

Power can be calculated by comparing the distribution of the test statistic when $H_0$ is false to the CV created through an examination of the distribution of the test statistic when $H_0$ is true. In a similar manner, define $\alpha = P(T^2_{(i,j)} < CV$ when variable i and j have not shifted) and $\beta = P(T^2_{(i,j)} > CV$ when variable i or j or both have shifted). Thus, power could be calculated by $1 - \beta = P(T^2_{(i,j)} < CV$ when variable i or j or both have shifted).

Given the nature of the two-variable deletion technique, the power analysis becomes more complex. Every simulated distribution of $T^2_{(i,j)}$ was examined and compared to the respective average CV for p = 3, 4, 5, and 10. The simulations involving ρ = 0.8 were excluded from the two-variable deletion power analysis since there was little difference between ρ = 0.5 and ρ = 0.8 scenarios.

Figure 14 displays the P($T^2_{(i,j)}$ < Average CV) for p = 3, 4, 5, and 10 when variable one has shifted δ = 1 and 3. For example, when p = 5 and δ = 3, P($T^2_{(1,\,j)}$ < Average CV = 4.7210) = 0.7115 for j = 2, 3, 4, 5.

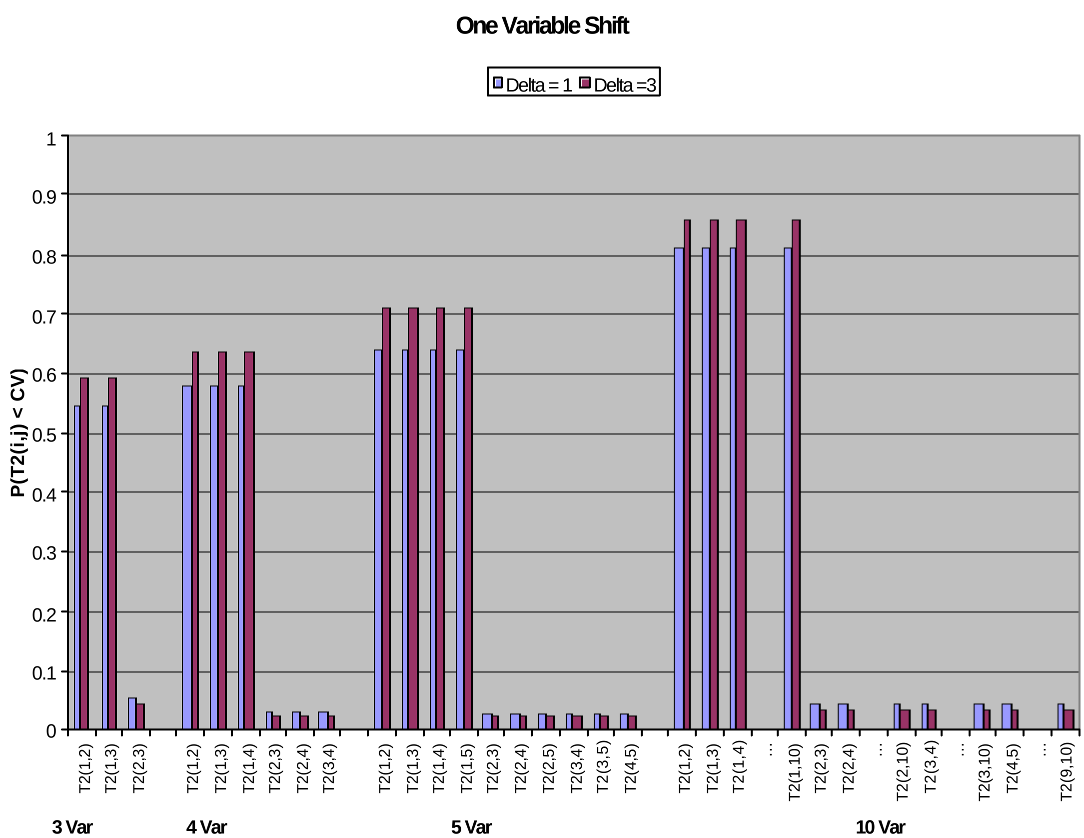


Figure 14. P($T^2_{(i,j)}$ < Average CV) for p = 3, 4, 5, and 10 When Variable One Shifts.

Figure 15 displays the P($T^2_{(i,j)}$ < Average CV) for p = 3, 4, 5, and 10 when variables one and two have shifted equally such that δ = 1 and 3. Figure 16 displays the P($T^2_{(i,j)}$ < Average CV) for p = 3, 4, 5, and 10 when variables one and two have shifted unequally such that δ = 1 and 3. Figure 17 displays the P($T^2_{(i,j)}$ < Average CV) for p = 3, 4, 5, and 10 when variables one, two, and three have shifted equally such that δ = 1 and 3. Figure 18 displays the $P(T^2_{(i,j)}$ < Average CV) for p = 3, 4, 5, and 10 when variables one, two, and three have shifted unequally such that δ = 1 and 3.

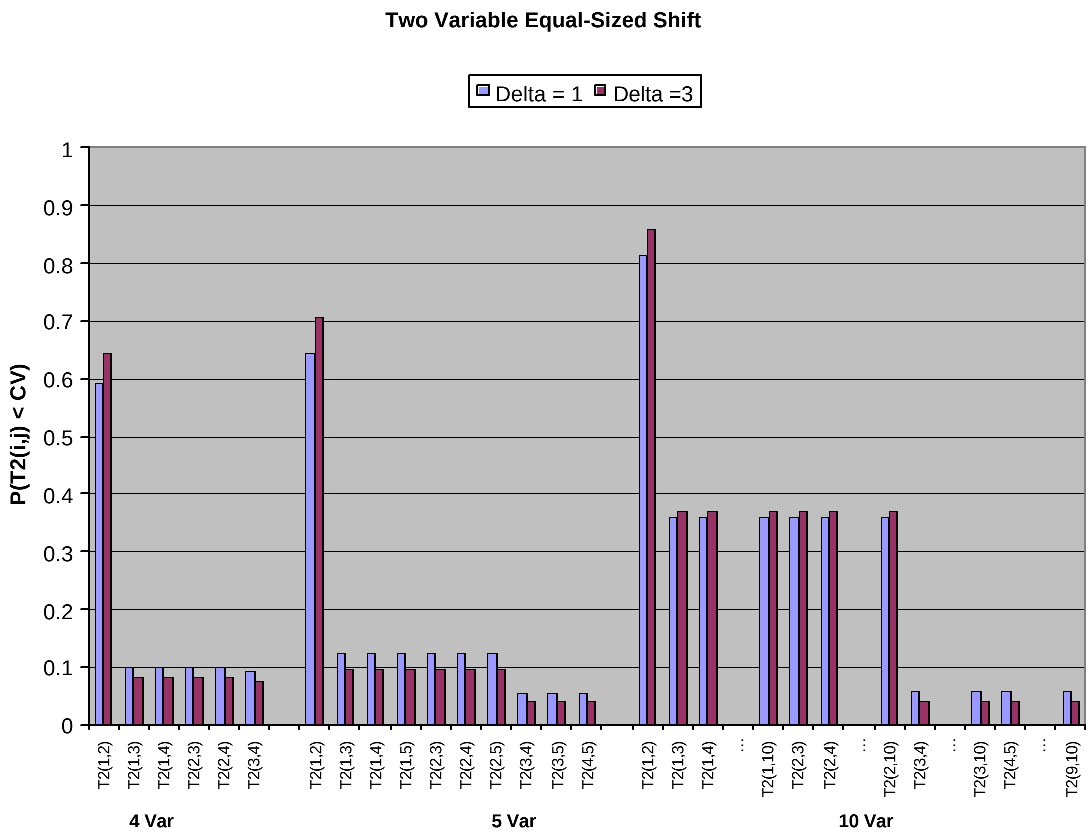


Figure 15. $P(T^2_{(i,j)}$ < Average CV) for p = 4, 5, and 10 When Variables One and Two Shift Equally.

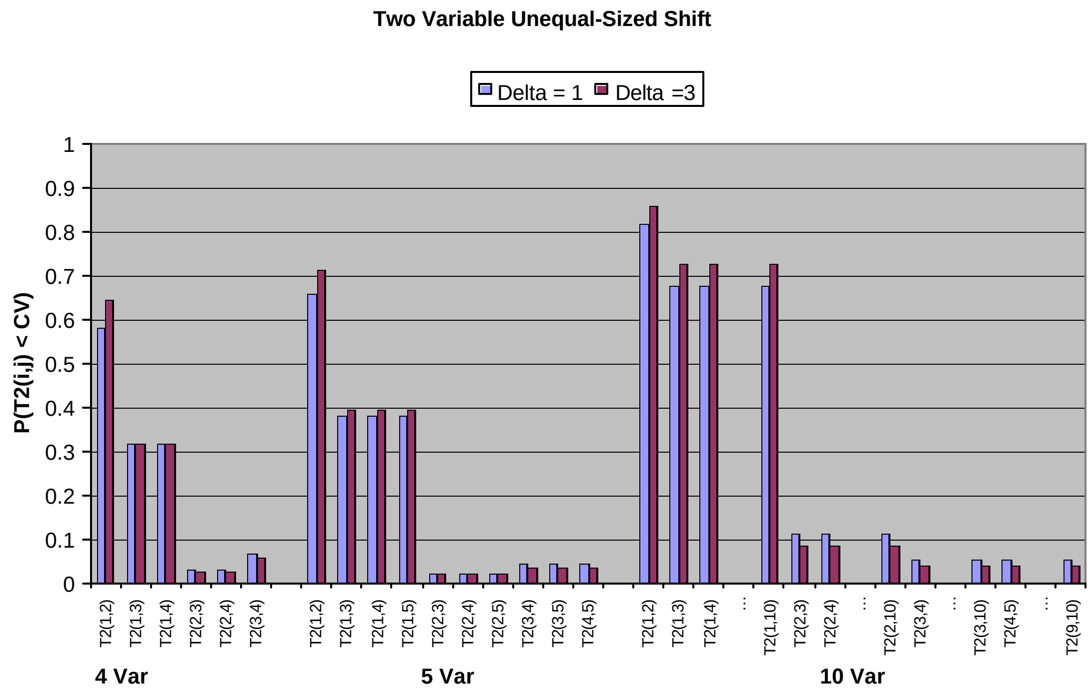


Figure 16. $P(T^2_{(i,j)}$ < Average CV) for p = 4, 5, and 10 when Variables One and Two Shift Unequally.

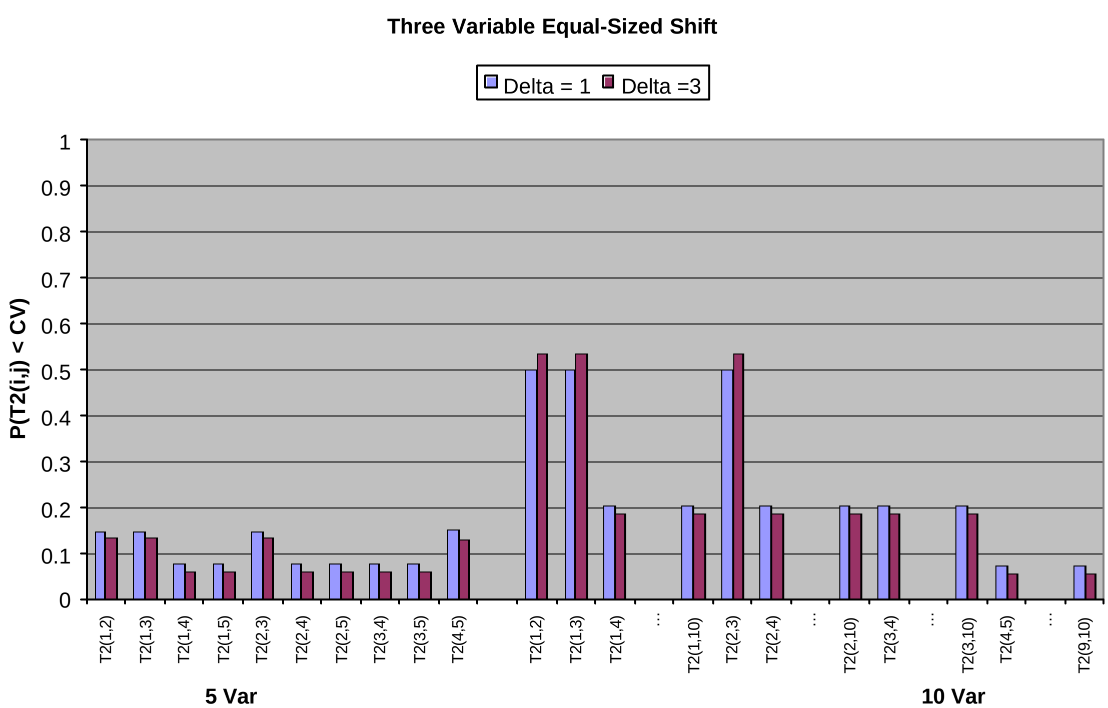


Figure 17. $P(T^2_{(i,j)}$ < Average CV) for p = 5 and 10 When Variables One, Two, and Three Shift Equally.

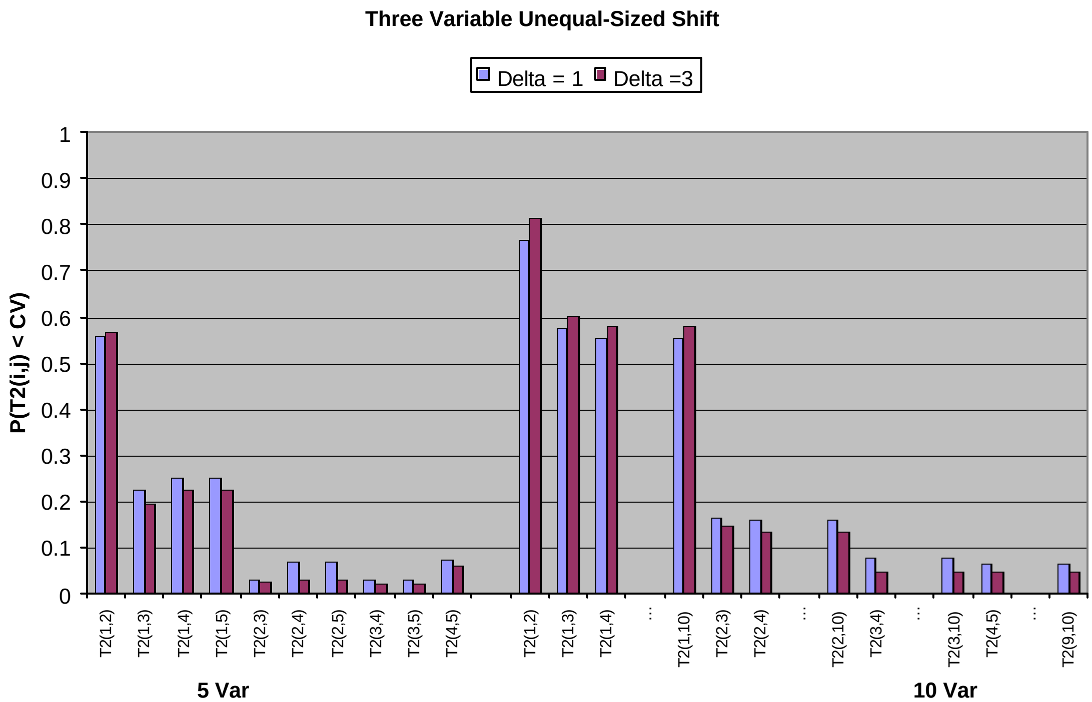


Figure 18. $P(T^2_{(i,j)} <$ Average CV) for p = 5 and 10 When Variables One, Two, and Three Shift Unequally.

## CONCLUSIONS

One-variable deletion and two-variable deletion become more powerful as the number of variables in the process increases. In addition, power increases as the size of the shift increases. Correlations between the variables have little effect on power.

One-variable deletion is able to detect a one variable shift best. The power of the technique decreases as the number of variables shifting increases. However, one-variable deletion can detect the largest "offending" variable in a multi-variable shift with reasonably high power when the number of variables in the process is large.

Two-variable deletion is able to detect a one variable shift best. However, the power of two-variable deletion is slightly lower than the power of one-variable deletion when one variable shifts. Two-variable deletion can detect two variable shifts (equal or unequal sizes) more effectively than one-variable deletion. In addition, two-variable deletion can detect the largest "offending" variable in a multi-variable shift with reasonably high power when the number of variables in the process is large.

Given the power and increasing use of modern computers in a manufacturing process, the MEWMA chart can effectively monitor a multivariate process. When the chart signals, the computer can easily create a graph of the reduced MEWMA statistics, with corresponding critical value, allowing the operator to readily identify the cause of the signal.

## REFERENCES

Allen, D. M. (1971), "Mean Square Error of Prediction as a Criterion for Selecting Variables," *Technometrics*, 13, pp. 469-475.

Alt, F. B. (1984). "Multivariate Quality Control" in *The Encyclopedia of Statistical Sciences*, eds. S. Kotz, N. L. Johnson, and C. R. Read, New York: John Wiley, pp.110-122.

Belsley, D. A., E. Kuh, and R. E. Welsch (1980), *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity,* Wiley, New York.

Crossier, R. B. (1988). "Multivariate Generalization of Cumulative Sum Quality-Control Schemes". *Techonometrics* 30, pp. 291-303.

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

Jackson, J. E. (1980). "Principal Components and Factor Analysis: Part I-Principal Components". *Journal of Quality Technology* 12, pp. 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*. John Wiley & Sons. Inc., New York, NY.

Lowry, C. A.; Woodall, W. H.; Champ, C. W.; and Rigon, S. E. (1992). "A Multivariate Exponentially Weighted Moving Average Control Chart", *Technometrics* 34, pp. 46-53.

Lowry, C. A. and Montgomery, D. C. (1995). "Review of Multivariate Control Charts". *IIE Transactions* 27, pp. 800-810.

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.

Mason, R. L.; Champ, C. W.; Tracy, N. D.; Wierda, S. J.; and Young, J. C. (1997). "Assessment of Multivariate Process Control Techniques". *Journal of Quality Technology* 29 pp. 140-143.

Pignatiello, J. J. and Runger, G. C. (1990). "Comparison of Multivariate CUSUM Charts". *Journal of Quality Control* 22, pp. 173-186.

Wierda, S. J. (1994). "Multivariate Statistical Process Control—Recent Results and Directions for Futture Research". *Statistica Neerlandica* 48, pp. 147-168

Woodall, W. H. and Montgomery, D. C. (1999). “Research Issues and Ideas in Statistical Process Control”. *Journal of Quality Technology* 31, pp.376-386.

KEYWORDS: Multivariate Quality Control, MEWMA, Variable Deletion