

## Research Article

# Robust Multivariate Control Charts to Detect Small Shifts in Mean

**Habshah Midi<sup>1,2</sup> and Ashkan Shabbak<sup>2</sup>**

<sup>1</sup> Faculty of Science, University Putra Malaysia, 43400 Serdang, Selangor, Malaysia

<sup>2</sup> Laboratory of Applied and Computational Statistics, Institute for Mathematical Research, University Putra Malaysia, 43400 Serdang, Selangor, Malaysia

Correspondence should be addressed to Ashkan Shabbak, ashkan@inspem.upm.edu.my

Received 14 February 2011; Accepted 16 March 2011

Academic Editor: Alexander P. Seyranian

Copyright © 2011 H. Midi and A. Shabbak. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The classical multivariate CUSUM and EWMA charts are commonly used to detect small shifts in the mean vectors. It is now evident that those charts are easily affected by outliers which may be due to small or moderate changes in the mean vector. In this paper, we propose a robust multivariate CUSUM and Robust multivariate EWMA charts to remedy the problem of small changed in scatter outliers. Both the empirical and simulation results indicate that the proposed robust multivariate CUSUM and EWMA charts offer substantial improvement over other multivariate CUSUM and EWMA charts. This article also discussed the robustness of the proposed charts, when there is a small or moderate sustained shift in the data set.

## 1. Introduction

The overall quality of a process is often measured by the joint level of several correlated characteristics.

Although, in practice, separate univariate control chart for each characteristic is often used to detect changes in the inherent variability of a process, it can be very misleading.

One of the major drawback of using separate univariate control charts is that the overall probability of false alarm (or overall type I error) becomes large, particularly when the number of the quality characteristics increases. Moreover, using separate univariate control charts for each characteristic would not have detected out of control conditions, because it ignores the correlation between the quality characteristics. As an alternative, multivariate control charts are recommended to simultaneously monitor more than one quality characteristics [1, 2]. Generally, two phases, namely Phase I and Phase II, are considered in any multivariate statistical process control application.

Phase I of monitoring scheme of a process consists of collecting a sufficient number of data to ascertain whether or not this Historical Data Set (HDS) indicates a stable

(or in-control) process whereas, in Phase II, future observations are monitored based on the control limits calculated from Phase I to determine if the process continues to be stable or not. If the process in phase I is out of control, the deviation sources must be detected and after removing out of control points, the control charts should be re-established. In other words, phase I is to identify multivariate outliers, recognize the root cause, and finally remedy the sources. Consequently, the estimated control limits of phase II would become more accurate [1, 3].

In statistical quality control, a process changes into an out-control situation when multiple outliers exist in the data set. An outlier is defined as an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism [4]. Multiple outliers can appear either as random distributed observations within the HDS set (scatter outliers) or as sequential observations in a certain period of the HDS (sustained shift).

Therefore, it is imperative to identify outliers in phase I that may otherwise lead to model misspecification and incorrect results in phase II. phase I and phase II of monitoring scheme of a process are usually called retrospective and prospective analysis, respectively (see [1]).

The Hotelling's  $T^2$  statistic was first introduced by Hotelling [5]. It is analogous to the Shewhart control chart in the univariate cases and widely used in the multivariate control charts for both retrospective and prospective stages of the monitoring schemes [6–10].

However, the Hotelling's  $T^2$  multivariate types of the Shewhart control chart are insensitive to small and moderate shifts in the mean vector due to the fact that these charts only use the information of current sample. To rectify this problem, the MCUSUM or MEWMA charts were introduced [11, 12]. These charts use the information from both the current and past observations. Since the MCUSUM and the MEWMA charts are more sensitive to the small and moderate shifts in the mean vector, they may be effective alternative to  $T^2$  charts when detection of small or moderate shift in a process are the main concerned [1, 2].

Typically, many researchers have suggested applying the MCUSUM and MEWMA in the prospective stage of the monitoring scheme [1]. However, due to the fact that in many real cases, the presence of observation with a small (moderate) changes or special cause in the HDS can create problem for our estimation in phase II, it would be desirable if those methods can also be used in the retrospective stage of the process monitoring scheme.

Sullivan and Woodall [13] suggested to employ the MCUSUM and MEWMA charts in phase I. Ryan [2] mentioned that it is reasonable if the MCUSUM and MEWMA procedures are used in the retrospective analysis of the monitoring scheme. Koning and Does [14] also employed a CUSUM chart in phase I and stated that it can be extendable to multivariate cases. Bersimis et al. [15] also pointed out to the application of the MCUSUM and MEWMA charts in phase I by giving an overview of the multivariate statistical quality control methods [2, 13–15].

Although the use of the MCUSUM and MEWMA charts can be helpful in phase I for small or moderate shifts, little work has been focused in this area. Hence, in this article, the ideas of Sullivan and Woodall [13] and other researchers already mentioned have motivated us to investigate the performance of MCUSUM and MEWMA charts in the presence of scatter outliers. Since the MCUSUM and MEWMA charts are based on classical location and dispersion estimates, they are easily affected by multiple outliers. To rectify this problem, we propose to incorporate robust location and scale estimates of the minimum volume ellipsoid (MVE) or the minimum covariance determinant (MCD) [16] in the formulation of the robust MCUSUM and MEWMA charts.

The remaining parts of this paper are organized in the following structures. In Section 2, some backgrounds of the Hotelling's  $T^2$ , MCUSUM and MEWMA charts, will be reviewed. A brief introduction of the location and dispersion robust estimators, which are based on the MVE and the MCD, is presented in Section 3. The proposed robust MCUSUM and MEWMA control charts are introduced in Section 4. In Section 5, Monte Carlo simulation study is carried out in order to evaluate the performance of the proposed robust MCUSUM and MEWMA charts in both outlier situations, namely, the scatter outliers and the sustained shift. Numerical examples are provided in Section 6. Finally, discussions and conclusions are presented in Section 7.

## 2. Background of Multivariate Control Charts

### 2.1. Hotelling's $T^2$ Charts

The Hotelling's  $T^2$  is a very common approach in multivariate control charts which is based on mahalanobis distance. Consider a HDS in phase I of monitoring scheme, which consists of  $m$  independent time-ordered observation vectors of dimension  $p$ , where  $p$  is the number of quality characteristics that are being measured ( $p < m$ ). The mahalanobis distance is defined as

$$MD_i^2 = (X_i - T(X))'C(X)^{-1}(X_i - T(X)), \quad (2.1)$$

where  $X_i$  is a random vector containing elements for the  $i$ th time period  $T(X)$  and  $C(X)$  are the location and scale estimates, respectively. If  $X_i$  follows the  $p$  variate normal distribution with  $\mu$  as a population mean vector and  $\Sigma$  as a population  $p \times p$  covariance matrix, then  $MD_i^2$  has chi-square distribution with  $k$  degrees of freedom [17]. The general form of  $T^2$  statistic are obtained by replacing  $T(X)$  and  $C(X)$  in (2.1) with classical sample mean vector ( $\bar{X}$ ) and the classical sample covariance matrix ( $S$ ). The classical Hotelling's  $T^2$  is then defined as

$$T_i^2 = (X_i - \bar{X})'S^{-1}(X_i - \bar{X}), \quad (2.2)$$

where

$$\bar{X} = \frac{1}{m} \sum_{i=1}^m X_i, \quad S = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X})(X_i - \bar{X})'. \quad (2.3)$$

As already been mentioned, the  $T^2$  chart is used both for retrospective and prospective analysis [1]. In phase I, the parameters are estimated retrospectively based on the current HDS, so the vector  $X_i$  is not independent of the estimators  $\bar{X}$  and  $S$ . In this situation, according to Mason and Young [7] and Tracy et al. [8], the statistical distribution of  $T^2$  is given as

$$T^2 \sim \left[ \frac{(m-1)^2}{m} \right] B\left( \frac{p}{2}, \frac{m-p-1}{2} \right), \quad (2.4)$$

where  $B(p/2, (m-p-1)/2)$  represents a beta distribution with parameters  $p/2$  and  $(m-p-1)/2$ . The upper control limit (UCL) of the  $T^2$  in (2.4) is  $((m-1)^2/m)B(\alpha, p/2, (m-p-1)/2)$ ,

where  $\alpha$ , is the  $\alpha$ th upper quantile of the beta distribution, which is the probability of false alarm for each point plotted on the control chart. The lower control limit LCL is often set to zero [1, 7].

## 2.2. Multivariate Cumulative Sum Charts

The Cumulative Sum Chart (CUSUM) was first introduced by Page in 1954 as a powerful method in the monitoring small or moderate shifts in the mean process. The CUSUM charts directly combine all information from a dataset in the sequence of sample values by computing the cumulative sums of deviations of a sample values from a target value, that is, the CUSUM charts are based on using a sequential probability ratio tests. An excellent overview of CUSUM methods are given by Hawkins and Olwell [18] and Montgomery [1].

An inclusive review of the generalization approaches of the univariate CUSUM charts to multivariate CUSUM can be found in Bersimis et al. [15].

Another generalization approaches was proposed by Pignatiello and Runger in 1990. Let  $X_i$  be a random vector containing elements for the  $i$ th observation which derives from the  $p$  variate normal distribution with an in-control population mean vector  $\mu_0$  and an in-control common  $p \times p$  covariance matrix,  $\Sigma_0$ . Pignatiello and Runger [19] defined  $C_i$  as

$$C_i = \max \left\{ \left[ D_i' \sum_0^{-1} D_i \right]^{-1/2} - kn_i, 0 \right\} \quad i = 1, 2, \dots, m, \quad (2.5)$$

where  $C_0 = 0$  and  $k$  is the reference value based on the out-control value of  $\mu$  ( $\mu_1$ ) which is also generally used in the univariate CUSUM schemes. A good choice of  $k$  is often chosen about halfway between  $\mu_0$  and  $\mu_1$ . Additionally,  $D_i$  is obtained by

$$D_i = \sum_{j=i-n_i+1}^i (X_j - \mu_0). \quad (2.6)$$

And  $n_i$  is the number of subgroups (if any). It is formally defined as

$$n_i = \begin{cases} n_{i-1} + 1 & \text{if } C_i > o, \\ 1 & \text{otherwise.} \end{cases} \quad (2.7)$$

In this paper, we only focus on the individual observations, that is, ( $n_i = 1$ ).

Although most approaches of the MCUSUM procedures are designed for phase II, the approach given in (2.5) can be viewed as a logical basis of phase I. This is due to the fact that it is based on the multivariate distance of related observation from in-control mean vector, somehow similar to  $T^2$  charts. Sullivan and Woodall [13] suggested setting  $k$  equals to zero for phase I when employing the CUSUM in (2.5) [13, 19].

### 2.3. Multivariate Exponentially Weighted Moving Average

The Exponentially Weighted Moving Average (EWMA) is another alternative to the Shewhart chart when small or moderate shifts occur in the process mean [1]. This chart is due to Roberts [12], who introduced a geometric moving average chart.

It is worth mentioning that the CUSUM and EWMA charts differ from each other in how we use the obtained information from observations in the HDS, that is, EWMA statistic assigns less and less weight to the past observations than the current observation, whereas all observations in the CUSUM procedure are weighted equally. Lowry et al. [20] introduced a Multivariate EWMA as a natural direct extension of the univariate case [20].

Let  $X_i$  be a random vector from the  $p$  variate normal distribution with known in-control  $\mu_0$  as a population mean vector and  $\Sigma_0$  as a known in-control common  $p \times p$  covariance matrix. Lowry et al. [20] defined MEWMA denoted as  $Z_i$

$$Z_i = RX_i + (I - R)z_{i-1} = \sum_{j=1}^i R(I - R)^{i-j} X_j \quad Z_0 = 0_p = (0, \dots, 0)', \quad i = 1, 2, \dots, m, \quad (2.8)$$

where  $R$  is a  $p \times p$  diagonal matrix with diagonal elements  $r_k$ ,  $0 \leq r_k \leq 1$ ,  $k = 1, 2, \dots, p$  and  $I$  is the identity matrix. Often all  $r_k$  set to be equal if there is no any reason to weight past observation differently. The MWEMA chart signals if  $Z_i' \Omega_i^{-1} Z_i$  exceeds a related UCL, Where  $\Omega_i$  is the covariance matrix of  $Z_i$  and defined as

$$\Omega_i = \frac{(1 - (1 - r)^{2i})r}{(2 - r)\Sigma_0}, \quad \forall r_k = r, \quad k = 1, \dots, p. \quad (2.9)$$

When subgroup size is large enough, the MEWMA control chart can be calculated as

$$E_i = Z_i' \left( \frac{r}{2 - r} \Sigma_0 \right)^{-1} Z_i. \quad (2.10)$$

Choosing the values of  $r$  determines the sensitivity of the charts to different magnitudes of shifts. A smaller weight makes the MEWMA chart more sensitive to small shifts and larger  $r$  leads the chart to be more sensitive for larger shifts. If  $r = 1$ , then the MEWMA is equivalent to  $T^2$  control charts.

Montgomery [1] suggested choosing  $r$  from the range  $0.05 \leq r \leq 0.25$ . In practice, a popular choice are  $r = 0.05$ ,  $r = 0.10$ , and  $r = 0.20$ . As a good compromise, Croux et al. [21] and Gelper et al. [22] recommended  $r = 0.20$ .

Similar to MCUSUM in Section 2.2, the MEWMA chart defined in (2.10) can be used in phase I with some modifications. Sullivan and Woodall [13] suggested choosing  $r = 0.05$  because  $r$  smaller value is more sensitive to a small shift. They also recommended centering the observation by modifying (2.8) to

$$Z_i = R(X_i - \mu_0) + (I - R)z_{i-1}, \quad Z_0 = 0_p. \quad (2.11)$$

In phase I, the in-control mean vector and covariance matrix are not known and should be estimated. Sullivan and Woodall [13] recommended to apply the MWEMA chart for both

reverse time order and regular time order of data to achieve equal detection probability for the same magnitudes of shift which start from the same number of observations from either end of the HDS.

### 3. High Breakdown Estimators based on the MVE and MCD

It is now evident that the classical sample mean or classical sample covariance (like  $T^2$ ) are easily affected by abnormal observations [9, 23]. One simple way to overcome this problem in univariate cases is to use the sample median and MAD (median absolute deviation) instead of the sample mean and variance, respectively [24].

Although this strategy cannot be applied to multivariate data, the idea can be used in this case, by replacing the location and the scale estimators with measures of central tendency and dispersion which are more resistant to atypical data and make the multivariate approaches more robust. Minimum volume ellipsoid (MVE) and minimum covariance determinant (MCD) are frequently employed in the detection of multivariate outliers in regression contexts due to their high breakdown point and affine equivariant properties [16, 23, 25–28]. These two high breakdown estimators were initiated first by Rousseeuw [16]. The finite sample breakdown point of an estimator shows that how many outlier, need to ruin the estimation of the estimator completely.

The basic strategy for obtaining the MVE estimator is to search ellipsoid from all ellipsoids which covers at least  $h$  points of dataset, where  $h$  can be taken equals to  $[m/2] + 1$ . Subsets of size  $h$  are called halfsets. After finding that ellipsoid, the usual sample mean and sample covariance of the corresponding ellipsoid are computed and declared as the MVE estimators. Rousseeuw and Leroy [23] proved that the finite sample breakdown point of these estimators is 0.5.

Due to the difficulties of computing the MVE, there are some proposed algorithms to calculate their values in the literatures. The interested reader can refer to Jensen et al. [29] to see the complete reference list of these algorithms. One of these algorithms is a subsampling method which is based on taking a fixed number of subsets randomly from the data set each of size  $p + 1$ . These randomly chosen subsets, determine the shape of an ellipsoid. Finally, the size of this ellipsoid inflates by multiplying it by a constant until covers  $h$  subsets of data set [23, 27]. In this study, the subsampling algorithm is utilized because of its wide availability in statistical software packages such as SAS, S-Plus, and R. Another popular approach for finding robust estimates for mean vector and the covariance matrix is the MCD. The main idea is the same as MVE, but, here, MCD is obtained by finding the half set that gives the minimum value of the determinant of the covariance matrix. Its breakdown point is similar to the MVE. Analogous to the MVE, due to the difficulties in the calculations of MCD, there are some proposed algorithms to calculate its values [27, 30, 31]. Rousseeuw and Van Driessen [31] proposed a Fast-MCD algorithm which is easily available in the software packages and is used in this article. It is worth mentioning here that the reweighted MVE and MCD were employed in this study to increase the efficiency of both the MVE and MCD estimators [23, 32, 33].

### 4. Robust Multivariate Control Charts

As already been mentioned, the Hotelling's  $T^2$  statistic based on classical sample mean and sample covariance matrix is badly affected by outliers [7, 9, 34, 35].

Sullivan and Woodall [9] recommended replacing  $S$  in (2.2) with a covariance matrix which is estimated by using the vector difference between successive observations of the HDS. This covariance matrix is defined as

$$S_{sd} = \frac{1}{2(m-1)} \sum_{i=1}^{m-1} V_i V_i', \quad i = 1, 2, \dots, m, \quad (4.1)$$

where  $V_i = X_{i+1} - X_i$  is the difference between successive observations. Henceforth, the  $T^2$  statistic which is calculated based on (4.1) is referred, as  $T_{sd}^2$ . Another approach to make  $T^2$  chart more robust is to use the MVE or the MCD estimators instead of location and dispersion estimators of the Hotelling's  $T^2$  statistic which is applied first by Vargas in 2003. The application of robust control charts for individual observations based on the MVE and MCD have also been discussed extensively by Jensen et al. [29]. According to them, after obtaining the robust multivariate location and scale estimates given by either the MVE or the MCD, the robust Hotelling  $T^2$  statistics are defined as follows:

$$\begin{aligned} T_{MVE,i}^2 &= \left( X_i - \bar{X}_{iMVE} \right)' S_{MVE}^{-1} \left( X_i - \bar{X}_{iMVE} \right), \\ T_{MCD,i}^2 &= \left( X_i - \bar{X}_{iMCD} \right)' S_{MCD}^{-1} \left( X_i - \bar{X}_{iMCD} \right), \end{aligned} \quad (4.2)$$

where  $i = 1, 2, \dots, m$ . It is important to note that there is no guarantee that the assumptions of the mathematical distribution of (2.2) preserved when replacing the location and scale estimators with robust versions. Although Wilimas et al. [35] were successful in the determination of the mathematical distribution of the Hotelling's  $T^2$  based on the  $S_d$ , still there is no exact known mathematical distribution for the Hotelling's  $T^2$  based on the MVE or the MCD. Hence, to overcome this problem, empirical distribution of the robust  $T^2$  statistic is used for calculating the empirical control limits. Hereinafter, the robust version of Hotelling's  $T^2$  statistic based on the MVE or the MCD is indicated by  $RT_i^2$ . For applying the MCUSUM and MEWMA charts in the retrospective analysis, due to the fact that usually in phase I of the monitoring scheme, the  $\mu_0$  and  $\Sigma_0$  are unknown, the classical sample mean vector and covariance matrix are utilized to estimate them [13]. Since the MCUSUM and MEWMA procedures in (2.5), (2.6), (2.10), and (2.11) utilize the classical covariance matrix, those procedures may suffer from outliers. Hence, we propose using the robust covariance matrix from the MVE or the MCD estimators to formulate the robust version of the MCUSUM chart for individual observations, which are defined as follows:

$$\begin{aligned} C_{MVE,i}^2 &= D'_{MVE,i} S_{MVE}^{-1} D_{MVE,i}, \quad \text{where } D_{MVE,i} = \sum_{j=1}^i \left( X_j - \bar{X}_{MVE} \right), \\ C_{MCD,i}^2 &= D'_{MCD,i} S_{MCD}^{-1} D_{MCD,i}, \quad \text{where } D_{MCD,i} = \sum_{j=1}^i \left( X_j - \bar{X}_{MCD} \right), \end{aligned} \quad (4.3)$$



for  $i = 1, 2, \dots, m$  and  $C_0 = 0$ . Likewise, the proposed robust version of the MEWMA chart for individual observations are defined as

$$E_{MVE,i}^2 = Z'_{MVE,i} \left( \frac{r}{2-r} S_{MVE}^{-1} \right)^{-1} Z_{MVE,i}, \quad \text{where } Z_{MVE,i} = R(X_i - \bar{X}_{MVE}) + (I - R)Z_{MVE,i-1},$$

$$E_{MCD,i}^2 = Z'_{MCD,i} \left( \frac{r}{2-r} S_{MCD}^{-1} \right)^{-1} Z_{MCD,i}, \quad \text{where } Z_{MCD,i} = R(X_i - \bar{X}_{MCD}) + (I - R)Z_{MCD,i-1},$$
(4.4)

for  $i = 1, 2, \dots, m$  and  $Z_0 = 0$ .

In the next section, we investigate the performance of our proposed robust MCUSUM and MEWMA charts for small and moderate shifts in the mean vector in the presence of both scatter and sustained shift outliers.

## 5. Simulation Study

In this section, a Monte Carlo simulation study is carried out to assess the performance of the control charts already discussed. The simulation is designed based on three subgroups each of size  $m = 30, 50,$  and  $100$  with the number of characteristics  $p = 2, 3, 5,$  and  $10$ . Let us assume that the in-control process is a  $p$ -variate Normal distribution with the mean vector  $\mu_0$  and the variance-covariance matrix  $\Sigma_0$ . The out-control process is a  $p$ -variate Normal distribution with the same variance-covariance matrix but with shifted mean vector as  $\mu_1$ . The amount of the shift is measured by a scalar that is defined as

$$(\mu_1 - \mu_0)^T \sum_0^{-1} (\mu_1 - \mu_0). \quad (5.1)$$

This measure is called noncentrality parameter and, hereafter, is referred as ncp. It is clear from (5.1), that the severity of the shift only depends on the values of  $\mu_1$ . Additionally, it can be assumed that  $\mu_0$  is a zero vector and  $\Sigma$  is a  $p \times p$  identity matrix.

Due to the unknown distribution of the proposed robust statistics when replacements of estimators were made, the approximated empirical control limits are obtained by simulation. The control limits are exact when applied to a single point in phase I, despite this phase is a retrospective analysis of all observations. In this situation,  $\alpha$  is the probability of false alarm for each point plotted on the control chart. Therefore, if all of the statistics were distributed independently, the probability of false alarm would be  $1 - (1 - \alpha)^m$ , where  $m$  denotes the number of subgroups. In this case, the probability of false alarm is called the overall false alarm.

Practically, it is preferred to determine the control limits by simulation to give a specified overall false alarm even though the distributions of statistics are known [2, 13, 35]. Hereafter, we refer to  $\alpha$  as the overall false alarm. Since, the lower control limits for all control charts which are discussed earlier are often set to zero [1, 7, 13, 29], the simulated empirical upper control limits (UCLs) are obtained by generating 5000 in-control datasets for each combination of  $m$  and  $p$ .

Due to the affine equivariant property of all preceding statistics, these limits are applicable to any values of  $\mu$  and  $\Sigma$ . In this article, the overall false alarm is given by  $\alpha = 0.05$ .



**Table 1:** The empirical simulated UCLs for different types of Hotelling’s  $T^2$ .

$p$	$m$	$T^2$ charts			
		$T^2$	$T^2_{MVE}$	$T^2_{MCD}$	$T^2_{sd}$
2	30	10.561	24.351	58.812	12.284
	50	12.302	21.042	34.615	13.443
	100	14.124	20.478	24.441	14.712
3	30	12.202	29.028	108.530	15.038
	50	13.976	25.145	50.430	15.984
	100	16.503	23.011	30.371	17.115
5	30	14.950	49.371	319.625	19.930
	50	17.581	33.792	111.725	20.496
	100	20.054	28.448	41.639	21.798
10	30	20.033	80.693	1555.711	36.008
	50	24.250	63.427	393.625	31.856
	100	27.964	44.287	84.655	31.902

**Table 2:** The empirical simulated UCLs for different types of MCUSUM and MEWMA.

$p$	$m$	MCUSUM charts				MEWMA charts			
		$C^2$	$C^2_{sd}$	$C^2_{MVE}$	$C^2_{MCD}$	$E^2$	$E^2_{sd}$	$E^2_{MVE}$	$E^2_{MCD}$
2	30	60.71	82.80	172.93	977.19	3.70	5.17	9.28	39.39
	50	110.08	132.04	264.99	867.38	5.57	6.85	10.52	24.55
	100	225.50	249.03	427.95	872.99	8.52	9.66	12.57	17.77
3	30	72.11	109.30	200.86	1829.22	4.43	6.81	10.82	73.41
	50	127.32	165.45	252.84	1370.61	6.43	8.45	11.14	38.35
	100	278.79	317.81	468.20	1244.82	9.85	11.63	13.55	22.44
5	30	93.39	174.05	337.15	4930.11	5.63	10.61	17.59	194.71
	50	169.22	244.99	309.70	3290.83	8.27	12.35	13.67	88.67
	100	360.53	426.13	525.88	1972.71	12.40	15.24	16.24	32.64
10	30	134.60	417.99	290.08	15236.53	7.99	24.69	17.60	673.30
	50	250.92	473.44	558.16	10122.93	11.71	22.92	24.20	272.41
	100	541.31	739.27	747.96	4812.36	17.79	24.95	22.88	69.34

The empirical simulated UCLs were determined by calculating all the related statistics for each observation in the generated datasets and record the maximum value for the corresponding dataset. In order to retain  $\alpha = 0.05$ , the 95th percentile of the recorded maximum values is declared as the simulated UCLs.

The obtained simulated UCLs for different types of  $T^2$ , MCUSUM and MEWMA charts, are exhibited in Tables 1 and 2 for different sizes of  $m$  and  $p$ . Following the idea of Sullivan and Woodall [13], the value of  $r$  is chosen to be 0.05.

The estimated value, of  $C^2$ ,  $E^2$ , and  $T^2$  in all tables refer to the MCUSUM, MEWMA and Hotelling’s  $T^2$  charts, respectively which are computed by using the classical covariance matrix. Likewise, the  $C^2_{sd}$ ,  $E^2_{sd}$ , and  $T^2_{sd}$  denote the MCUSUM, MEWMA, and Hotelling’s  $T^2$  charts respectively which are based on the successive covariance matrix, respectively.

**Table 3:** The probability of signals for scatter outliers, where  $p = 2$  and  $m = 30$ .

$\epsilon$	ncp	MCUSUM charts				MEWMA charts			$T^2$ charts		
		$C^2$	$C_{sd}^2$	$C_{MVE}^2$	$C_{MCD}^2$	$E^2$	$E_{sd}^2$	$E_{MVE}^2$	$E_{MCD}^2$	$T_{MVE}^2$	$T_{MVE}^2$
5%	1	0.05	0.04	0.05	0.04	0.06	0.04	0.05	0.04	0.05	0.05
	3	0.03	0.04	0.05	0.05	0.04	0.04	0.05	0.06	0.06	0.07
	5	0.04	0.05	0.09	0.05	0.05	0.05	0.1	0.06	0.10	0.08
	10	0.04	0.04	0.12	0.07	0.04	0.04	0.14	0.08	0.25	0.17
10%	1	0.07	0.06	0.06	0.04	0.06	0.06	0.06	0.04	0.05	0.05
	3	0.06	0.06	0.08	0.06	0.06	0.05	0.08	0.07	0.08	0.08
	5	0.05	0.04	0.10	0.08	0.05	0.04	0.09	0.07	0.09	0.09
	10	0.04	0.05	0.23	0.12	0.05	0.05	0.23	0.12	0.24	0.17
15%	1	0.05	0.06	0.06	0.05	0.06	0.06	0.06	0.05	0.05	0.06
	3	0.05	0.05	0.09	0.08	0.06	0.06	0.10	0.10	0.06	0.06
	5	0.06	0.05	0.13	0.11	0.06	0.05	0.13	0.11	0.07	0.08
	10	0.05	0.04	0.31	0.20	0.04	0.04	0.29	0.22	0.17	0.14
20%	1	0.05	0.06	0.05	0.04	0.05	0.05	0.05	0.03	0.05	0.04
	3	0.05	0.04	0.10	0.10	0.06	0.05	0.10	0.10	0.05	0.06
	5	0.05	0.06	0.16	0.16	0.06	0.06	0.15	0.15	0.06	0.06
	10	0.04	0.05	0.38	0.31	0.04	0.05	0.37	0.32	0.16	0.14

Traditionally, the performance of control charts is evaluated by using the Average Run Length (ARL) index. However, in the retrospective analysis, the ARL is not meaningful because a signal is generated for all observations in the HDS. Instead, for comparing control charts in phase I, the signal probability is used to achieve a specified overall false alarm. For estimating the probability of signals in the existence of scatter outliers, four percentage of outliers, denoted as  $\epsilon = 5\%, 10\%, 15\%$ , and  $20\%$ , are generated for each dataset. Then a contaminated dataset of size  $m$  in the  $p$  dimensions is generated for different values of ncp. Each generated dataset is compared with the corresponding empirical UCL in Tables 1 or 2.

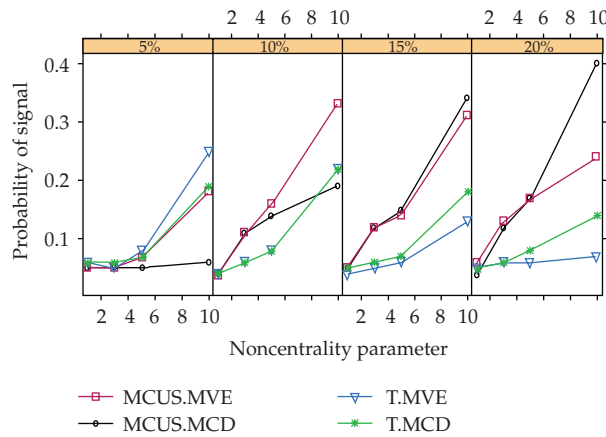
This process is repeated for 5000 times, and eventually the proportion of datasets that have at least one point greater than the UCLs is reported as the probability of signal for each respective chart. Due to space constraint, only the results for  $m = 30$ , and  $m = 100$  and  $p = 2$  are exhibited in Tables 3 and 4. Other results are consistent.

As shown in Tables 3 and 4, both proposed Robust MCUSUM and MEWMA charts outperform the  $RT_i^2$  based on the MVE or the MCD estimators except for  $\epsilon = 5\%$ . At 5% outliers, the  $RT_i^2$  have higher probability of signals than MCUSUM and MEWMA charts for all ncp. However, as the value of  $\epsilon$  increases, the probability of detecting shift for the Robust MCUSUM and MEWMA increases. It is interesting to note that by increasing the number of subgroup size both MCUSUM and MEWMA charts based on MCD is slightly better than that when based on MVE. The values for  $T_{sd}^2$  are even worse than those for  $RT_i^2$  based on the MVE or the MCD in all cases (the results are not shown here).

The results for other levels of  $m$  and  $p$  are consistent but are not reported here due to space limitations. However, when the dimension increases, the subgroup size should also be increased to keep the consistency due to the fact that as  $p$  rises for a fixed value of  $m$ , the breakdown point of the MVE and MCD decreases. For example, Figures 1 and 2 illustrate the probability of signals in the case of scatter outliers for MCUSUM and MEWMA procedures,

**Table 4:** The probability of signals for scatter outliers, where  $p = 2$  and  $m = 100$ .

$\epsilon$	ncp	MCUSUM charts				MEWMA charts				$T^2$ charts	
		$C^2$	$C^2_{sd}$	$C^2_{MVE}$	$C^2_{MCD}$	$E^2$	$E^2_{sd}$	$E^2_{MVE}$	$E^2_{MCD}$	$T^2_{MVE}$	$T^2_{MVE}$
5%	1	0.06	0.06	0.06	0.06	0.05	0.05	0.05	0.06	0.06	0.06
	3	0.06	0.05	0.07	0.06	0.04	0.03	0.06	0.06	0.10	0.10
	5	0.06	0.06	0.11	0.08	0.04	0.04	0.08	0.07	0.18	0.17
	10	0.07	0.06	0.31	0.17	0.05	0.04	0.24	0.18	0.51	0.46
10%	1	0.04	0.05	0.05	0.06	0.04	0.04	0.05	0.05	0.06	0.06
	3	0.05	0.05	0.15	0.16	0.04	0.04	0.14	0.14	0.08	0.08
	5	0.06	0.05	0.24	0.21	0.05	0.05	0.22	0.21	0.15	0.17
	10	0.05	0.06	0.64	0.51	0.06	0.05	0.60	0.50	0.54	0.54
15%	1	0.04	0.05	0.04	0.06	0.05	0.05	0.05	0.06	0.05	0.05
	3	0.05	0.05	0.17	0.16	0.04	0.05	0.16	0.17	0.07	0.07
	5	0.06	0.06	0.27	0.27	0.05	0.05	0.25	0.28	0.09	0.10
	10	0.05	0.05	0.68	0.70	0.05	0.05	0.66	0.68	0.39	0.43
20%	1	0.05	0.04	0.05	0.07	0.04	0.04	0.04	0.06	0.06	0.06
	3	0.06	0.06	0.17	0.19	0.04	0.04	0.17	0.18	0.06	0.06
	5	0.05	0.05	0.31	0.33	0.04	0.04	0.29	0.33	0.09	0.10
	10	0.06	0.05	0.69	0.76	0.04	0.04	0.66	0.74	0.30	0.34

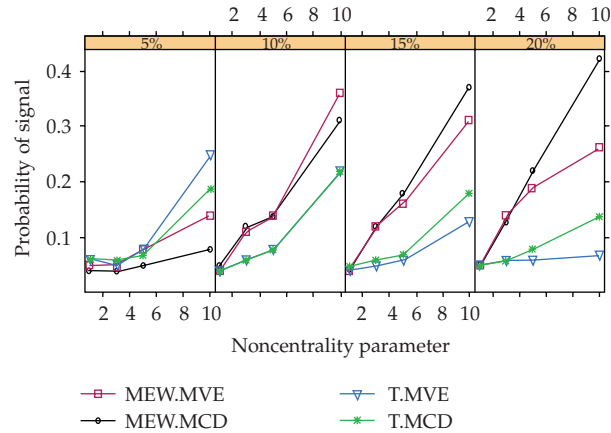


**Figure 1:** MCUSUM chart for different percentage of scatter outliers, where  $p = 5$  and  $m = 100$ .

for  $p = 5$  and  $m = 100$ . It is apparent from these figures that although  $m$  is large, for fixed values of ncp and  $\epsilon$ , the performance of the respective control charts for  $p = 2$  is better than when  $p = 5$ .

The situation in which the mean vector after  $k$  observations changes from in-control value to out-control value is called sustained shifts and is modeled as

$$\mu_i = \begin{cases} \mu_0, & \text{for } i = 1, 2, \dots, k, \\ \mu_1, & \text{for } i = 1, 2, \dots, m, \end{cases} \quad (5.2)$$



**Figure 2:** MEWMA chart for different percentage of scatter outliers, where  $p = 5$  and  $m = 100$ .

where  $\mu_0$  and  $\mu_1$  are in-control and out-control mean vector, respectively. In this paper, to assess the performance of our proposed control charts in comparison with other existing control charts, two situations are considered for the sustained shift case: first, when a small or moderate change occurs midway in the HDS and, secondly, when a small or moderate change happens near the end or at the beginning of data. It is important to note that for phase I of monitoring scheme of a process, the probability of detection shift, which appears either at the beginning or at the end of the HDS, should be the same due to the fact that the  $\mu_0$  is unknown in this phase. That is, in both cases some observations may come from one statistical distribution and some other observations may come from another distribution [13].

For the shift in the middle of the HDS, the required datasets are obtained by generating 50% of  $m$  observations from in-control process and the other 50%, from the shifted observations for different values of the ncp. The shifted generated observations are compared with the corresponding empirical UCL from Tables 1 or 2 and the signal is recorded for the first shifted observation that exceeds the respective UCL. This process is repeated 5000 times for each combination of  $m$  and  $p$ , and eventually the proportion of signals is declared as the signal probability.

For the case of existing shift at the end or at the beginning of a data, for simplicity, we assume shift occurs at the end of the HDS. Then, the data sets are simulated consequently in the form of generating 80% of  $m$  observations from the in-control process and 20% from the shifted process. Subsequently, the signal probability is calculated for each control chart. Due to space limitation, only the results of  $m = 30, 50$  and  $100$  for  $p = 2$  are presented in Tables 5, 6, and 7.

When the MEWMA control chart is applied, the time order of observations in the HDS should be taken into consideration. This is because of the weight effect of the previous observations. This effect causes two sustained shift situations with the similar magnitude in which one located near the first observation in the HDS and the other is at the same distance from the last observation, which have different probability of signal. Nonetheless, it is desirable to have the same probability of signal for the same shift magnitude, in which the number of observations should be the same, starting from either end of the data.

**Table 5:** The probability of signals for sustained shift, where  $p = 2$  and  $m = 30$ .

Shift	ncp	MCUSUM charts				MEWMA charts				$T^2$ charts		
		$C^2$	$C_{sd}^2$	$C_{MVE}^2$	$C_{MCD}^2$	$E^2$	$E_{sd}^2$	$E_{MVE}^2$	$E_{MCD}^2$	$T_{MVE}^2$	$T_{MVE}^2$	$T_{sd}^2$
20%	1	0.03	0.20	0.06	0.05	0.05	0.26	0.09	0.06	0.04	0.04	0.10
	2	0.08	0.40	0.09	0.07	0.13	0.47	0.14	0.11	0.06	0.06	0.15
	3	0.14	0.54	0.06	0.06	0.23	0.62	0.15	0.10	0.05	0.06	0.22
	4	0.22	0.72	0.08	0.05	0.40	0.81	0.21	0.12	0.06	0.05	0.26
	5	0.33	0.83	0.10	0.08	0.57	0.90	0.29	0.16	0.07	0.08	0.35
	10	0.81	0.99	0.22	0.16	0.96	1.00	0.59	0.42	0.14	0.13	0.67
50%	1	0.24	0.50	0.10	0.06	0.22	0.51	0.14	0.11	0.06	0.06	0.12
	2	0.60	0.83	0.11	0.07	0.54	0.81	0.18	0.12	0.04	0.04	0.16
	3	0.84	0.95	0.10	0.05	0.78	0.93	0.20	0.10	0.04	0.04	0.25
	4	0.93	0.98	0.16	0.07	0.88	0.98	0.26	0.13	0.04	0.04	0.31
	5	0.98	1.00	0.20	0.09	0.97	1.00	0.30	0.14	0.04	0.05	0.36
	10	1.00	1.00	0.38	0.09	1.00	1.00	0.50	0.16	0.03	0.04	0.57

In order to achieve this property, the MEWMA procedure is used with the respective simulated UCL for both original time order and reverse time order of the generated observations in the HDS.

It is worth mentioning that this idea is not helpful for the scatter outlier situation since the distribution for an outlier is different immediately before and in the succeeding observations. Thus, including successive observations contribute no useful information, whereas with the sustained shifts, there are several observations in a sequence from the same distribution.

The results obtained from Tables 5–7 reveal that the performance of the proposed MCUSUM and MEWMA charts, based on the MVE and MCD, are not effective in both sustained shift situations in comparison with the MCUSUM and MEWMA procedures based on the successive covariance matrix. Under such situation, it appears that the MCUSUM and MEWMA procedures, based on the classical covariance matrix, are even better than the proposed charts in detecting small and moderate shifts. It may happen due to the fact that the design of both MCUSUM and MEWMA charts are tuned to be the best for a specific sustained shift.

It can also be seen from the Tables 5–7 that for both shifts, that is, in the midway and at the last 20% observations of data, the MCUSUM and MEWMA charts based on the successive covariance matrix ( $C_{sd}^2$  and  $E_{sd}^2$ ) outperform the classical covariance matrix ( $C^2$  and  $E^2$ ).

For the midway shift, the MCUSUM charts are slightly better than the MEWMA and the MEWMA is somewhat better in the case of a shift at the last 20% observations of the data. Both the MCUSUM and MEWMA charts have less sensitivity when a shift occurs at the last 20% observations of the generated data compared with the same value of ncp at the midpoint of the data except for the maximum severity of shift (ncp = 10). However, by increasing  $m$ , their performances are fairly closed for smaller severity of shift. Although the performance of the  $T_{sd}^2$  is better than the  $T^2$  statistic based on the MVE or MCD, it is completely apparent from these tables that the performances of all  $RT_i^2$ s are significantly less efficient than the MCUSUM and MEWMA charts. Moreover, it is interesting to note that by increasing  $m$ , the performance of the  $T_{sd}^2$  is also increased. These results are consistent for other levels of  $p$  and  $m$  which are not presented here due to space limitations.

**Table 6:** The probability of signals for sustained shift, where  $p = 2$  and  $m = 50$ .

Shift	ncp	MCUSUM charts				MEWMA charts				$T^2$ charts		
		$C^2$	$C^2_{sd}$	$C^2_{MVE}$	$C^2_{MCD}$	$E^2$	$E^2_{sd}$	$E^2_{MVE}$	$E^2_{MCD}$	$T^2_{MVE}$	$T^2_{MVE}$	$T^2_{sd}$
20%	1	0.18	0.36	0.07	0.06	0.24	0.45	0.16	0.12	0.05	0.06	0.08
	2	0.43	0.67	0.09	0.07	0.55	0.77	0.32	0.18	0.06	0.04	0.16
	3	0.75	0.91	0.12	0.06	0.86	0.94	0.55	0.28	0.06	0.05	0.26
	4	0.89	0.98	0.16	0.10	0.95	0.99	0.74	0.40	0.08	0.07	0.30
	5	0.96	1.00	0.16	0.10	0.99	1.00	0.86	0.46	0.09	0.08	0.41
	10	1.00	1.00	0.42	0.38	1.00	1.00	1.00	0.86	0.20	0.22	0.79
50%	1	0.65	0.76	0.17	0.07	0.55	0.73	0.32	0.18	0.04	0.04	0.09
	2	0.96	0.98	0.46	0.11	0.91	0.96	0.60	0.28	0.04	0.04	0.18
	3	1.00	1.00	0.71	0.12	0.99	1.00	0.82	0.30	0.04	0.03	0.26
	4	1.00	1.00	0.88	0.13	1.00	1.00	0.92	0.35	0.04	0.03	0.32
	5	1.00	1.00	0.95	0.15	1.00	1.00	0.97	0.38	0.03	0.04	0.42
	10	1.00	1.00	1.00	0.20	1.00	1.00	1.00	0.48	0.03	0.05	0.67

**Table 7:** The probability of signals for sustained shift, where  $p = 2$  and  $m = 100$ .

Shift	ncp	MCUSUM charts				MEWMA charts				$T^2$ charts		
		$C^2$	$C^2_{sd}$	$C^2_{MVE}$	$C^2_{MCD}$	$E^2$	$E^2_{sd}$	$E^2_{MVE}$	$E^2_{MCD}$	$T^2_{MVE}$	$T^2_{MVE}$	$T^2_{sd}$
20%	1	0.66	0.76	0.31	0.13	0.69	0.80	0.60	0.42	0.05	0.06	0.09
	2	0.98	0.99	0.57	0.16	0.98	0.99	0.94	0.82	0.05	0.06	0.18
	3	1.00	1.00	0.74	0.20	1.00	1.00	1.00	0.95	0.05	0.06	0.28
	4	1.00	1.00	0.87	0.24	1.00	1.00	1.00	1.00	0.06	0.06	0.41
	5	1.00	1.00	0.92	0.31	1.00	1.00	1.00	1.00	0.06	0.09	0.50
	10	1.00	1.00	1.00	0.73	1.00	1.00	1.00	1.00	0.30	0.37	0.91
50%	1	0.98	0.99	0.84	0.34	0.88	0.94	0.79	0.60	0.05	0.05	0.12
	2	1.00	1.00	0.99	0.76	1.00	1.00	0.99	0.89	0.03	0.04	0.19
	3	1.00	1.00	1.00	0.94	1.00	1.00	1.00	0.98	0.04	0.04	0.30
	4	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	0.02	0.04	0.36
	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.03	0.05	0.45
	10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.03	0.05	0.81

### 6. Numerical Example

In this section, a real dataset taken from Quesenberry [36] is used to evaluate the performance of the proposed charts. The original data set consists of measurements of eleven quality characteristics of 30 products, but here we only consider the first two characteristics. The data is presented in Table 8.

The MCUSUM, MEWMA procedures, and  $RT^2$  control charts are shown in Figures 3, 4, and 5, respectively. It is worth mentioning that from the plot of Figure 5, there is a sustained shift in the data set. By referring to the control limits of Table 1, it can be seen from Figure 3 that the MCUSUM chart based on successive covariance matrix, identified observations 21 to 26 as outliers. However, the MCUSUM chart based on the MVE or the MCD, failed to detect any outlier.

The MEWMA procedures shown in Figure 4 identify observations 22–27 as outliers, based on the respective UCLs in Table 2, for the  $E^2_{sd}, E^2_{MVE}, E^2_{MCD}$ . The results based on time

**Table 8:** The Quesenberry data set.

Product no.	$X_1$	$X_2$
1	0.567	60.558
2	0.538	56.303
3	0.53	59.524
4	0.562	61.102
5	0.483	59.834
6	0.525	60.228
7	0.556	60.756
8	0.586	59.823
9	0.547	60.153
10	0.531	60.64
11	0.581	59.785
12	0.585	59.675
13	0.54	60.489
14	0.458	61.067
15	0.554	59.788
16	0.469	58.64
17	0.471	59.574
18	0.457	59.718
19	0.565	60.901
20	0.664	60.18
21	0.6	60.493
22	0.586	58.37
23	0.567	60.216
24	0.496	60.214
25	0.485	59.5
26	0.573	60.052
27	0.52	59.501
28	0.556	58.467
29	0.539	58.666
30	0.554	60.239

reverse order of data, also exhibited on the right-hand side of Figure 4. In this case, for example, observations number 1 and 2 describe observations number 30 and 29 in the original time order of the dataset. It is clear from this figure that nothing is gained by the MVE or the MCD estimators. According to the UCLs for  $T_{sd}^2$ ,  $T_{MVE}^2$ ,  $T_{MCD}^2$  of Table 1, Figure 5 shows that only one observation (observation number 2) is detected as outlier by  $T^2$  statistics based on the MVE and the successive covariance estimators while the  $T^2$  based on the MCD estimator unable to detect any point as outlier.

## 7. Conclusion

In the retrospective analysis of a process, detecting small or moderate shifts in the mean vector is important in order to correctly estimate the parameters of a process. Since the CUSUM or the EWMA procedures are more sensitive to small and moderate changes in the dataset than  $T^2$  charts, this study is designed to investigate the application of these control



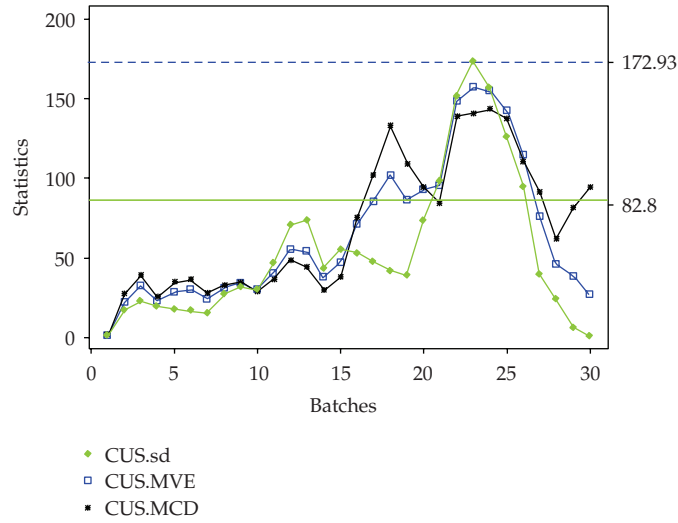


Figure 3: MCUSUM control charts.

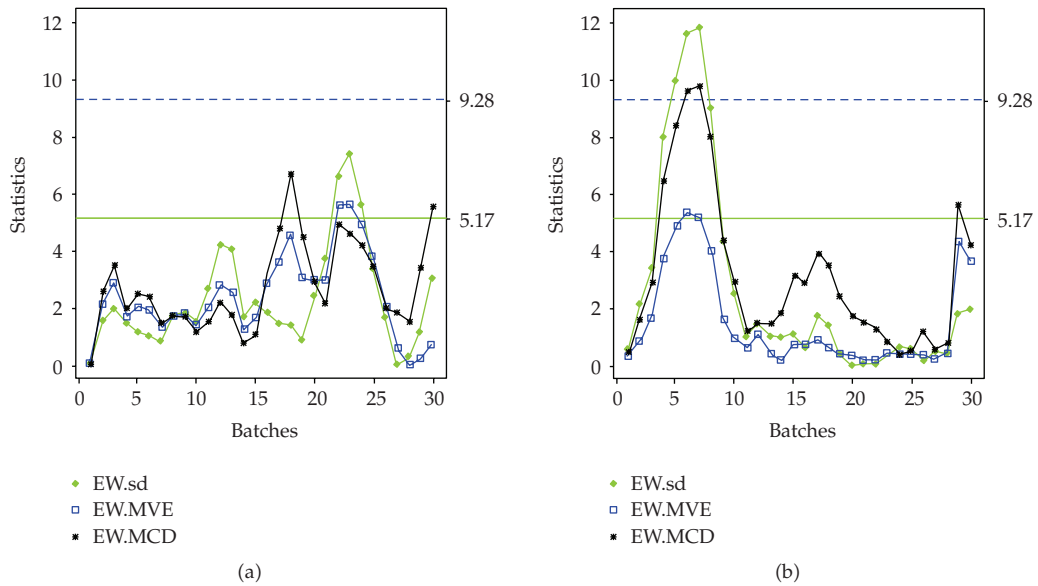


Figure 4: MWMA control charts; (b) is plotted in the original time order and (a) is plotted in the reverse time order.

procedures in multivariate data for the retrospective analysis of a process monitoring scheme, in two situations, namely, in the presence of scatter outliers and the sustained shift in the historical dataset.

We proposed robust MCUSUM and robust MEWMA procedures to remedy the problems of outliers by incorporating the MVE and the MCD in the development of such robust charts.

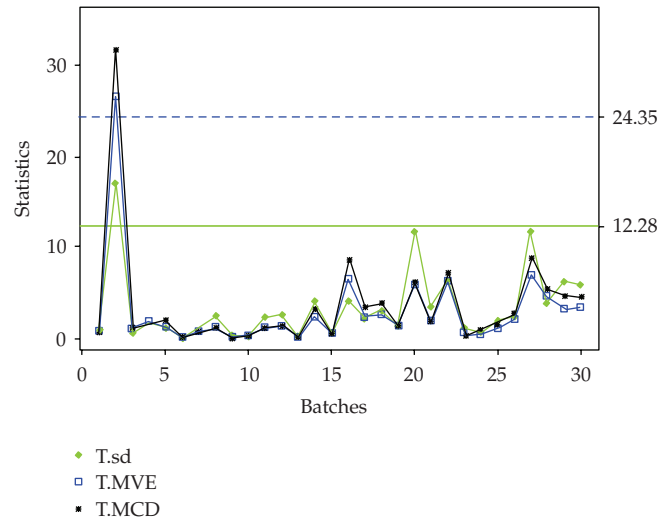


Figure 5:  $RT^2$  control charts.

The results of this study indicate that the robust MCUSUM and MEWMA procedures, based on the MVE or the MCD estimators, improve the detection probability of scatter outliers with a small shift in the mean vector as compared to the robust  $T^2$  charts except at 5% percentage of contaminated data. Moreover, the results indicate that the MCD-based estimators are more superior than the MVE-based estimators when the subgroup size of observation is increased.

Nonetheless, the proposed MCUSUM and MEWMA charts which are based on the MVE or the MCD are not effective in detecting sustained shift in the dataset.

When the shift occurs midway in the dataset, the performance of the MCUSUM are somewhat superior to the MEWMA. On the other hand, the MEWMA is slightly better than the MCUSUM, in the situation when the shifts happen either near the beginning or near the end of the dataset. However, by increasing the subgroup size, the performance of both charts is fairly closed.

Both the MCUSUM and MEWMA control charts are more sensitive to a small or moderate shift in the sustained shift situation, when it occurs in the middle of the dataset compared with the case that a shift occurs near the beginning or the end of data.

## References

- [1] D. C. Montgomery, *Introduction to Statistical Quality Control*, John Wiley & Sons, New York, NY, USA, 2005.
- [2] T. P. Ryan, *Statistical Methods for Quality Control*, John Wiley & Sons, New York, NY, USA, 2000.
- [3] F. B. Alt, *Multivariate Quality Control in Encyclopedia of Statistical Sciences*, John Wiley & Sons, New York, NY, USA, 1985.
- [4] D. M. Hawkins, *Identification of Outliers*, Chapman & Hall, London, UK, 1980.
- [5] H. Hotelling, "The generalization of Student's ratio," *The Annals of Mathematical Statistics*, pp. 360–378, 1931.
- [6] F. B. Alt, "17 Multivariate process control," *Handbook of Statistics*, vol. 7, pp. 333–351, 1988.
- [7] R. L. Mason and J. C. Young, *Multivariate Statistical Process Control with Industrial Applications*, vol. 9 of *ASA-SIAM Series on Statistics and Applied Probability*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa, USA, 2002.

- [8] N. D. Tracy, J. C. Young, and R. L. Mason, "Multivariate control charts for individual observations," *Journal of Quality Technology*, vol. 24, pp. 88–95, 1992.
- [9] J. H. Sullivan and W. H. Woodall, "A comparison of multivariate control charts for individual observations," *Journal of Quality Technology*, vol. 28, no. 4, pp. 398–408, 1996.
- [10] J. A. N. Vargas, "Robust estimation in multivariate control charts for individual observations," *Journal of Quality Technology*, vol. 35, no. 4, pp. 367–376, 2003.
- [11] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, pp. 100–115, 1954.
- [12] S. W. Roberts, "Control chart tests based on geometric moving averages," *Technometrics*, vol. 42, no. 1, pp. 97–101, 2000.
- [13] J. H. Sullivan and W. H. Woodall, "Adapting control charts for the preliminary analysis of multivariate observations," *Communications in Statistics. Simulation and Computation*, vol. 27, no. 4, pp. 953–979, 1998.
- [14] A. J. Koning and R. J. M. M. Does, "CUSUM charts for preliminary analysis of individual observations," *Journal of Quality Technology*, vol. 32, no. 2, pp. 122–132, 2000.
- [15] S. Bersimis, S. Psarakis, and J. Panaretos, "Multivariate statistical process control charts: an overview," *Quality and Reliability Engineering International*, vol. 23, no. 5, pp. 517–543, 2007.
- [16] P. J. Rousseeuw, "Least median of squares regression," *Journal of the American Statistical Association*, vol. 79, no. 388, pp. 871–880, 1984.
- [17] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, Prentice Hall, New Jersey, NJ, USA, 2002.
- [18] D. M. Hawkins and D. H. Olwell, *Cumulative Sum Charts and Charting for Quality Improvement*, Statistics for Engineering and Physical Science, Springer, New York, NY, USA, 1998.
- [19] J. J. Pignatiello and G. C. Runger, "Comparisons of multivariate CUSUM charts," *Journal of Quality Technology*, vol. 22, pp. 173–186, 1990.
- [20] C. A. Lowry, W. H. Woodall, C. W. Champ, and S. E. Rigdon, "A multivariate exponentially weighted moving average control chart," *Technometrics*, vol. 34, pp. 46–53, 1992.
- [21] C. Croux, S. Gelper, and K. Mahieu, "Robust exponential smoothing of multivariate time series," *Computational Statistics and Data Analysis*, vol. 54, no. 12, pp. 2999–3006, 2010.
- [22] S. Gelper, R. Fried, and C. Croux, "Robust forecasting with exponential and holt-winters smoothing," *Journal of Forecasting*, vol. 29, no. 3, pp. 285–300, 2010.
- [23] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons, New York, NY, USA, 1987.
- [24] R. R. Wilcox and H. J. Keselman, "Modern robust data analysis methods: measures of central tendency," *Psychological Methods*, vol. 8, no. 3, pp. 254–274, 2003.
- [25] A. H. M. R. Imon, "Identification of high leverage points in logistic regression," *Pakistan Journal of Statistics*, vol. 22, no. 2, pp. 147–156, 2006.
- [26] R. A. Maronna, R. D. Martin, and V. J. Yohai, *Robust Statistics*, Wiley Series in Probability and Statistics, John Wiley & Sons, Chichester, UK, 2006.
- [27] P. J. Rousseeuw and B. C. van Zomeren, "Unmasking multivariate outliers and leverage points," *Journal of the American Statistical Association*, pp. 633–663, 1990.
- [28] M. Habshah, M. R. Norazan, and A. H. M. R. Imon, "The performance of diagnostic-robust generalized potentials for the identification of multiple high leverage points in linear regression," *Journal of Applied Statistics*, vol. 36, no. 5, pp. 507–520, 2009.
- [29] W. A. Jensen, J. B. Birch, and W. H. Woodall, "High breakdown estimation methods for phase I multivariate control charts," *Quality and Reliability Engineering International*, vol. 23, pp. 615–629, 2007.
- [30] D. M. Rocke and D. L. Woodruff, "Robust estimation of multivariate location and shape," *Journal of Statistical Planning and Inference*, vol. 57, no. 2, pp. 245–255, 1997.
- [31] P. J. Rousseeuw and K. Van Driessen, "Fast algorithm for the minimum covariance determinant estimator," *Technometrics*, vol. 41, no. 3, pp. 212–223, 1999.
- [32] G. Willems, G. Pison, P. J. Rousseeuw, and S. Van Aelst, "A robust Hotelling test," *Metrika*, vol. 55, no. 1-2, pp. 125–138, 2002.
- [33] D. L. Woodruff and D. M. Rocke, "Computable robust estimation of multivariate location and shape in high dimension using compound estimators," *Journal of the American Statistical Association*, vol. 89, no. 427, pp. 888–896, 1994.

- [34] J. H. Sullivan, "Detection of multiple change points from clustering individual observations," *Journal of Quality Technology*, vol. 34, no. 4, pp. 371–383, 2002.
- [35] J. D. Williams, W. H. Woodall, J. B. Birch, and J. H. Sullivan, "Distribution of hotelling's T2 statistic based on the successive differences estimator," *Journal of Quality Technology*, vol. 38, no. 3, pp. 217–229, 2006.
- [36] C. P. Quesenberry, "The multivariate short-run snapshot Q chart," *Quality Engineering*, vol. 13, no. 4, pp. 679–683, 2001.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

