



Journal of Applied Sciences

ISSN 1812-5654

science
alert

ANSI*net*
an open access publisher
<http://ansinet.com>

The Performance of Robust Multivariate Statistical Control Charts based on Different Cutoff-points with Sustained Shift in Mean

¹Ashkan Shabbak, ^{1,2}Habshah Midi and ²Mohd Nooh Hassan

¹Laboratory of Applied and Computational Statistics, Institute for Mathematical Research,
University Putra Malaysia, 43400 Serdang, Selangor, Malaysia

²Department of Mathematics, Faculty of Science,
University of Putra Malaysia, 43400 Serdang, Selangor, Malaysia

Abstract: A separate univariate control chart for each characteristic is often used to detect changes in the inherent variability of a process because of ease of computation. Nevertheless, the shortcoming of using separate individual charts would not have detected out of control condition when the characteristics of interests are correlated. The problems get more complicated in the existence of sustained shift in the process mean. In this study, we proposed a robust multivariate control chart which is less sensitive to the sustained shift in mean process. However, the theoretical cutoff-points of the proposed charts are intractable. As an alternative, we proposed two different procedures of empirical cutoff-points. The performance of these two cutoff-points in detecting a step shift in the mean vector is investigated extensively by real example and monte carlo simulation.

Key words: T^2 statistic, Median Absolute Deviation (MAD), Minimum Covariance Determinant (MCD), Minimum Volume Ellipsoid (MVE), probability of signals, mahalanobis distance

INTRODUCTION

The quality of products (whether in industry or service sectors) is the first priority for producers. Essentially, the quality of goods and services can be improved by collecting data that enable one to evaluate the capability and stability of a process.

Shewart (1931), introduced control charts, which play an important role in statistical quality control to improve the quality of product and services. Thereafter, the application of control charts was popularized worldwide by Deming (2000) and Shewart (1931)

The purpose of any control chart is to identify occurrences of special causes of variation that come from outside of the usual process. In real situation, more than one important quality characteristics of a process or products and services are considered. The overall quality of a process is often measured by the joint level of several correlated characteristics. In this situation, the usage of multivariate control chart is more suitable to simultaneously monitor more than one quality characteristic. In any multivariate statistical process control application, generally two phases are considered (Alt, 1985; Montgomery, 2006).

The phase I of the monitoring scheme consists of collecting a sufficient number of data to ascertain whether or not historical data indicate 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 a stable process or not. The phase I analysis is sometime called retrospective analysis. 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 reconstructed. The aim of phase I is to identify multivariate outliers and recognize the cause of that for remedying sources so that phase II estimated control limits become more accurate (Alt and Smith, 1988; Tracy *et al.*, 1992). That is effective control limits based on the phase I scheme are very crucial to spot any instability in a future production process. Subsequently our analysis should be free from any misleading influence which is caused by atypical observations that are called outliers. In statistical quality control concepts, 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 (Hawkins, 1980).

The detection of outliers in phase I is very crucial because it is responsible for causing model misspecification and incorrect results. Outlier can be detected by using univariate or multivariate methods. The identification of outliers in multivariate cases is more difficult than in the univariate case. For instance, the simple graphical methods that can be used to detect outliers in a single dimension are often not available in higher dimensions.

The Hotelling's T^2 statistic which was first introduced by Hotelling (1931) is the most popular statistic used in multivariate control chart. This T^2 statistic is equivalent to the squared mahalanobis distance and its computation is based on the classical sample mean vector and classical sample variance-covariance matrix (Hotelling, 1931).

It is now evident that this statistic, which is based on the classical estimators, is easily affected by outliers (Rousseeuw and Leroy, 2003; Sullivan and Woodall, 1996).

When, there is more than one outlier or the existence of multiple outliers in the data set, the detection situation becomes more difficult due to masking and swamping (Rousseeuw and van Zomeren, 1990). Masking occurs when we fail to detect the outliers (false negative) while swamping occurs when observations are incorrectly declared as outliers.

The modern strategy for dealing with masking in the univariate case is to substitute the sample mean and variance with median and MAD (median absolute deviation), respectively (Abu-Shawiesh *et al.*, 2009; Wilcox and Keselman, 2003).

Nonetheless, this strategy cannot be applied to multivariate data. To rectify this problem, a popular strategy is to make multivariate approaches more robust by replacing the location and the scale estimators with measures of central tendency and dispersion that are resistant to outliers, such as the Minimum Volume Ellipsoid (MVE) and the Minimum Covariance Determinant (MCD) which are proposed by Rousseeuw and Leroy (2003). It is worth mentioning here that the sampling distribution of the modified Hotelling's T^2 is intractable. In this situation, we proposed to use the empirical cutoff-points instead of using the Chi-Squares cutoff-points (Hadi, 1992; Jensen *et al.*, 2007; Vargas and Lagos, 2007).

In this study, we will compare the robust multivariate control charts with the classical control chart based on two different cutoff-points. We confine our study only in detecting sustained shift in the process mean. Simulation study and numerical examples are used to compare the performance of the two control charts based on two different cutoff-points.

ROBUST MULTIVARIATE CONTROL CHARTS

The Hotelling's T^2 is a very common approach in multivariate control charts which is based on mahalanobis distance. However, it is not robust. The mahalanobis distance is defined as:

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

where, X_i is a random vector, $T(X)$ and $C(X)$ are the location and scale estimates respectively and $i = 1, 2, \dots, m$. Consider that the Phase I historical dataset consists of m time-ordered vectors such that each vector is of dimension P and $n = 1$. Therefore, X_i is a vector containing elements for the i th time period. For the in-control process, it is generally assumed that each X_i comes from a p dimensional multivariate normal distribution with μ as the population mean vector that determines the location and Σ as the $p \times p$ positive definite covariance matrix that determines the dispersion ($X_i \sim MNV(\mu, \Sigma)$)

The general form of the T^2 statistic is:

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

where, \bar{X} and S are the classical sample mean vector and the classical sample variance-covariance matrix respectively which are defined as:

$$\bar{X} = \frac{1}{m} \sum_{i=1}^m X_i, \quad S = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X})(X_i - \bar{X})' \quad (3)$$

where, $i = 1, 2, \dots, m$ For sufficiently large m , the T^2 Statistics is distributed as chi-squared.

Johnson and Wichern (2002) verified that the upper control limit is $UCL = \chi_{p,\alpha}^2$ and the lower control limit is usually set to zero. Due to the non robustness of the Hotelling's T^2 statistic (Jensen *et al.* 2007; Sullivan and Woodall, 1996; Goegebeur *et al.*, 2005), the classical sample mean vector and sample-covariance matrix in Eq. 2 and 3 are replaced by the respective statistics of the MVE and the MCD. 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. Then replace the location and scale estimator of the Hotelling's T^2 with the mean and variance obtained from the MVE or MCD. Rousseeuw and Leroy (2003) proofed that the finite sample breakdown point of this estimators are approximately 0.5.

Due to the difficulties of computing the MVE, there are some proposed algorithms to calculate their values (Jensen *et al.*, 2007; Rousseeuw and Leroy, 2003).

In this study, a subsampling algorithm which is introduced by Rousseeuw and Leroy (2003) and based on taking a fixed number of subsets randomly from the data set, each of size, is utilized. These randomly chosen subsets, determine the shape of an ellipsoid. The size of this ellipsoid inflates by multiplying it by a constant until covers h subsets of data set. This algorithm is widely used because of its availability in statistical software packages such as S-Plus and R.

Another popular approach for finding robust estimates for mean vector and the variance 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 variance-covariance matrix. Its breakdown point is similar to the MVE and due to the difficulties in the calculation of the MCD, a Fast-MCD algorithm initially was proposed by Rousseeuw and van Driessen (1999).

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:

$$\begin{aligned} T_{mve,i}^2 &= (X_i - \bar{X}_{mve})' S_{mve}^{-1} (X_i - \bar{X}_{mve}) \\ T_{mcd,i}^2 &= (X_i - \bar{X}_{mcd})' S_{mcd}^{-1} (X_i - \bar{X}_{mcd}) \end{aligned} \quad (4)$$

Hereinafter, the RT_i^2 is used to indicate the robust version of Hotelling's T^2 statistic. Rousseeuw and Leroy (2003) suggested a cutoff-point for Rt_i^2 as $\chi_{p,0.5}^2$ by assuming that the dimensional variables follow a multivariate normal distribution.

However, the exact mathematical distribution of the new statistics RT_i^2 is unknown. Due to the unknown exact mathematical distribution of the RT_i^2 , the empirical cutoff- points are proposed. Another cutoff- point which are based on the median and the MAD are proposed as:

$$\text{Median}(RT_i^2) + 3\text{MAD}(RT_i^2)$$

Hadi (1992) proposed this cutoff-point for the identification of potential high leverage points in a regression model. Subsequently, Habshah *et al.* (2009) used the same cutoff-points in their Diagnostic Robust Generalized Potential method (Habshah *et al.*, 2009).

SIMULATION STUDY

Here, we report a Monte Carlo simulation study that is designed to assess the performance of the Hotelling's T^2 and the RT^2 control charts based on different

cutoff-points. We consider five sub-samples, each of size $m = 20, 30, 40, 50$ and 100 and confine the study to only 2 variables ($p = 2$). All computations and simulations in this study were carried out by using S-PLUS® 8.0 for Windows.

As was mentioned earlier the $\chi_{\alpha,p=2}^2$ is used as the desired cutoff- points of the T^2 and the simulated control limits are used for the RT^2 because of its unknown exact distribution. The performance of each control chart is evaluated by using the probability of a signal for different shifts of mean and the degree of closeness of each control chart in detecting the exact number of observations with shifted process means.

Assume that μ_0 is the mean of in-control process. If μ_0 changes to a new value like μ_s , the difference between the in-control process mean and the out-off control process mean is labeled as non-centrality parameter and it can expressed as:

$$\lambda = (\mu_s - \mu_0)' \Sigma_0^{-1} (\mu_s - \mu_0) \quad (5)$$

The distance between μ_0 and μ_s is defined as the square root of λ . It is obvious that changes in the non-centrality parameter produces changes in the distribution of the T^2 statistics.

Changes in values of λ , occurs due to the change in the values of μ_s which means there is a shift in the in-control process mean. This shift can happen in the extent of δ , either on one element, more elements or on all elements of the mean vector and in different directions, although in term of non-centrality parameter, the orientation of the shifts is not considered. For example in the bivariate case, on the assumption that $\mu_0 = (0, 0)$, the following cases have the same non-centrality parameter value:

- (a) $\mu_s = (\delta, \delta)'$, (b) $\mu_s = (-\delta, \delta)'$,
- (c) $\mu_s = (-\delta, -\delta)'$, (d) $\mu_s = (\delta, -\delta)'$

The probability of a signal is defined as the probability of signalling after the first shift in the process mean. It should be noted that in this article, after changing the process mean we assume that the covariance matrix has not changed and remains constant for bivariate distributions. Here we only consider shift one element of in-control mean vector in the extent of δ . Since, the robust estimators of location and scale based on the MVE and MCD are affine equivariant estimators, without loss of generality, it is analogous to the classical sample mean and covariance, in which we assume that μ_0 is the zero vector and Σ_0 is the identity matrix. Subsequently, the in-control process is the multivariate normal standard

Table 1: Simulated empirical cutoff-points values for the UCLs of RT_1^2 control chart

| Sample size | Significance level | | | | | | | |
|-------------|--------------------|----------|-------------|----------|-------------------|----------|-------------|----------|
| | $\alpha_1 = 0.001$ | | | | $\alpha_2 = 0.01$ | | | |
| | EMPR UCL | | MAD-MED UCL | | EMPR UCL | | MAD-MED UCL | |
| | MVE | MCD | MVE | MCD | MVE | MCD | MVE | MCD |
| 20 | 53.71423 | 201.9933 | 6.289645 | 10.08569 | 20.53264 | 80.69049 | 5.231476 | 5.231476 |
| 30 | 29.35543 | 73.34435 | 6.269205 | 9.073127 | 15.24485 | 36.06677 | 5.212607 | 7.524045 |
| 40 | 23.43623 | 46.99034 | 6.290572 | 8.417557 | 13.65172 | 24.48091 | 5.229743 | 6.983040 |
| 50 | 21.03457 | 34.57649 | 6.289405 | 7.966992 | 12.71337 | 19.58086 | 5.228829 | 6.612813 |
| 100 | 17.53624 | 21.13389 | 6.314541 | 7.145443 | 11.50695 | 13.60266 | 5.249702 | 5.935296 |

distribution with μ_0 as the mean vector and Σ_0 as the covariance matrix, where:

$$\mu_0 = (0,0)', \Sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Hence, for constructing the empirical UCLs, firstly we generated a bivariate standard normal distribution for five sets of data each of size $m = 20, 30, 40, 50$ and 100 . Then the RT_1^2 values for each data set are calculated and this procedure is repeated 5000 times. Finally, the α -th upper percentile of these $5000 \times m$ values was recorded as empirical cutoff-point corresponding to each m .

The cutoff-points that are obtained by this methods hereafter, are referred as EMPR UCLs. Another approach of constructing empirical UCLs for the RT_1^2 is based on the Median $(RT_1^2) + kMAD(RT_1^2)$ values of that $5000 \times m$ data with respect to each m . The value of k corresponding to the significance levels of α . These cutoff-points are denoted by MAD-MED UCLs. In this study, we considered only two alpha levels, that is $\alpha_1 = 0.001$ and $\alpha_2 = 0.01$ their corresponding k values are 3 and 2.33, respectively. The results are presented in Table 1.

In order to calculate the probability of a signal, five data sets were generated from in-control process each of size c and then add 10 observation to each data set from another bivariate normal distribution with the same covariance matrix but a shift in the mean. This implies that the process mean changes from μ_0 to μ_s where, $\mu_s = (\delta, 0)$. The T_1^2 and the RT_1^2 values for each c and δ are recorded for $i = c+1$ to $c+10$ and this procedure is repeated 5000 times, where:

$$c = \{10, 20, 30, 40, 50, 90\} \text{ and } \delta = \{0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 10\}$$

The proportion of datasets in which the T_1^2 or RT_1^2 are greater than the respective UCLs after shifts have occurred are recorded as the probability of a signal. These probability versus the shift in the process mean are plotted in Fig. 1-5.

It can be observed from Fig. 1-4 that the Hotelling's T^2 control charts cannot detect shifts in the process mean for higher percentage (more than 10%) of shifted observations, irrespective of the α levels and type of cutoff-points. However, its probability of detection increases as the process in means increases, only for the empirical and the MAD-MED cutoff-points at $\alpha = 0.01$ and 10% contaminations. The plot in Fig. 1 also pointed out that the performance of the RT_1^2 control chart is reasonably close to the Hotelling T^2 for the empirical cutoff-points and in the presence of 50% shifted observations. The results are as expected because the breakdown points of the MVE and the MCD estimators is 50%. Nonetheless, both the RT_1^2 show certain percentage of detections based on the MAD-MED cutoff-points.

The plot of Fig. 2-5 suggest that by increasing the sample size and keep the number of shifted observations equal 10, has increased the probability of detection of the RT_1^2 control charts, for both types of cutoff-points.

It can be seen that the MAD-MED cutoff-points always outperform the empirical cutoff-points and their performances are fairly close for larger shifts.

We further investigate the performance of the robust control charts based on two different cutoff-points by looking at the number of exact detection of shifted observations. A good control chart is the one that can detect the exact number of shifted observations that were added to the process. In this study, we added 10 observations with shifted in means. Due to space constraint, we only include the results when $m = 20$ and $m = 100$. It can be seen from Table 2 that for $m = 20$ the T^2 chart cannot detect any shifted observation.

The MCD- RT_1^2 chart and the MVE- RT_1^2 chart based on the MAD-MED cutoff-points can only detect 2 shifted observations and 1 shifted observation, respectively. Nevertheless for larger subsamples ($m = 100$), the performance of both RT_1^2 based on the MAD-MED cutoff-points outperform the RT_1^2 based on the empirical cutoff-points.

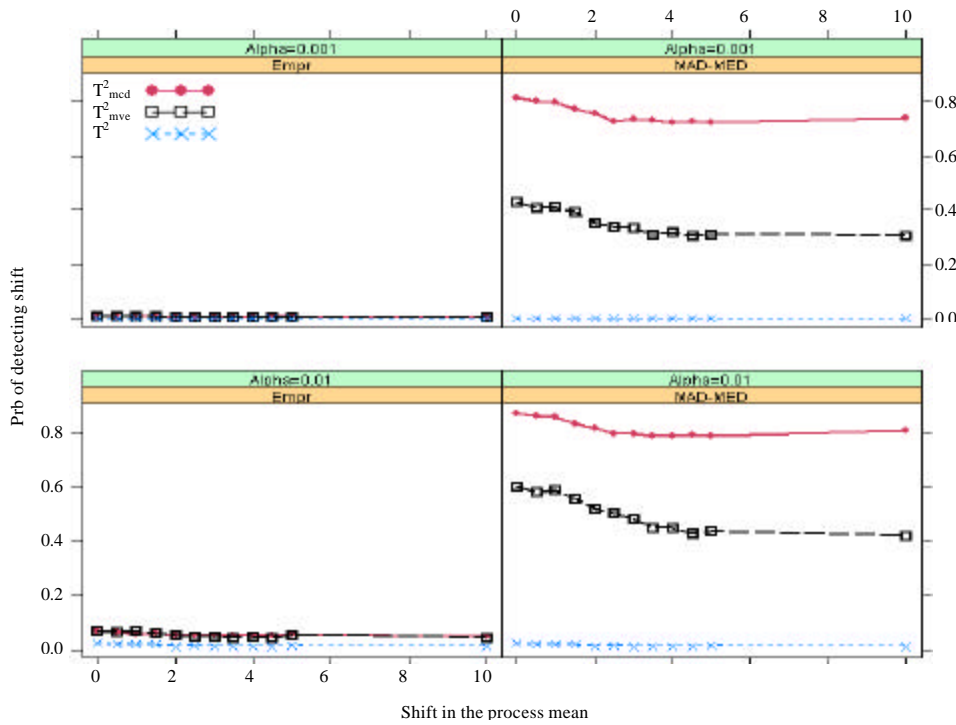


Fig. 1: Probability of a signal when $m = 20$ and 10 observations (50%) come from shifted distribution

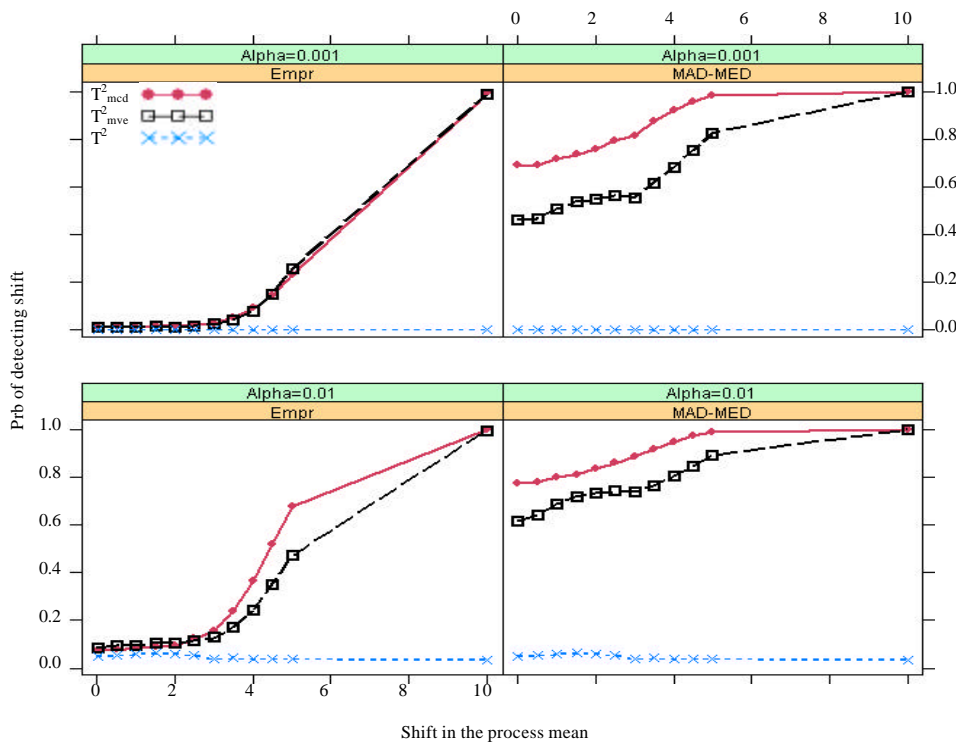


Fig. 2: Probability of a signal when $m = 30$ and 10 observations (33%) come from shifted distribution

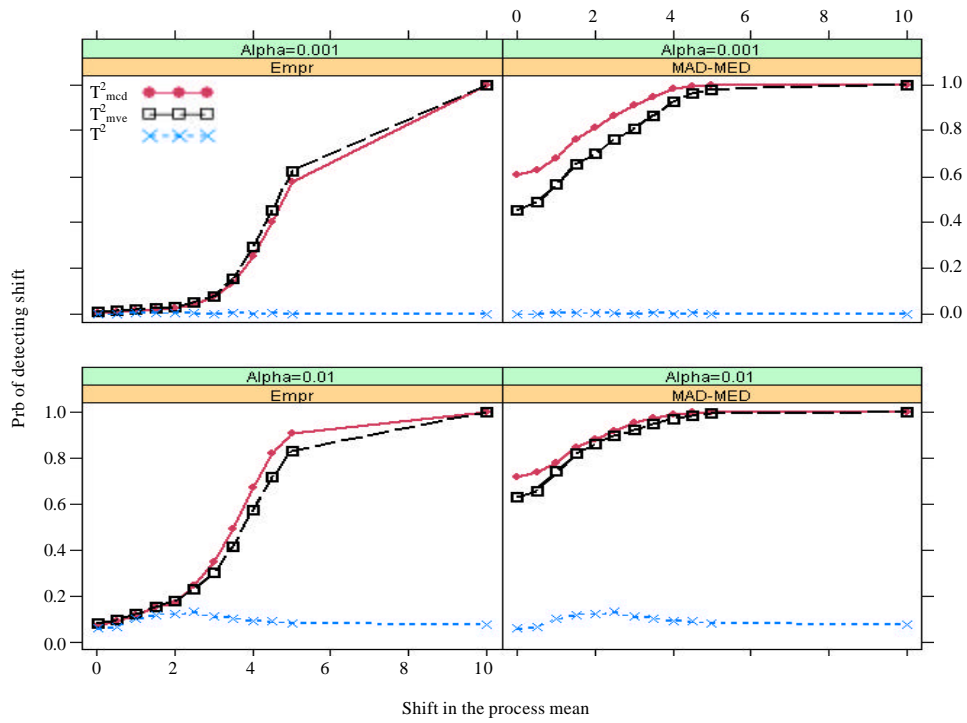


Fig. 3: Probability of a signal when $m = 40$ and 10 observations (25%) come from shifted distribution

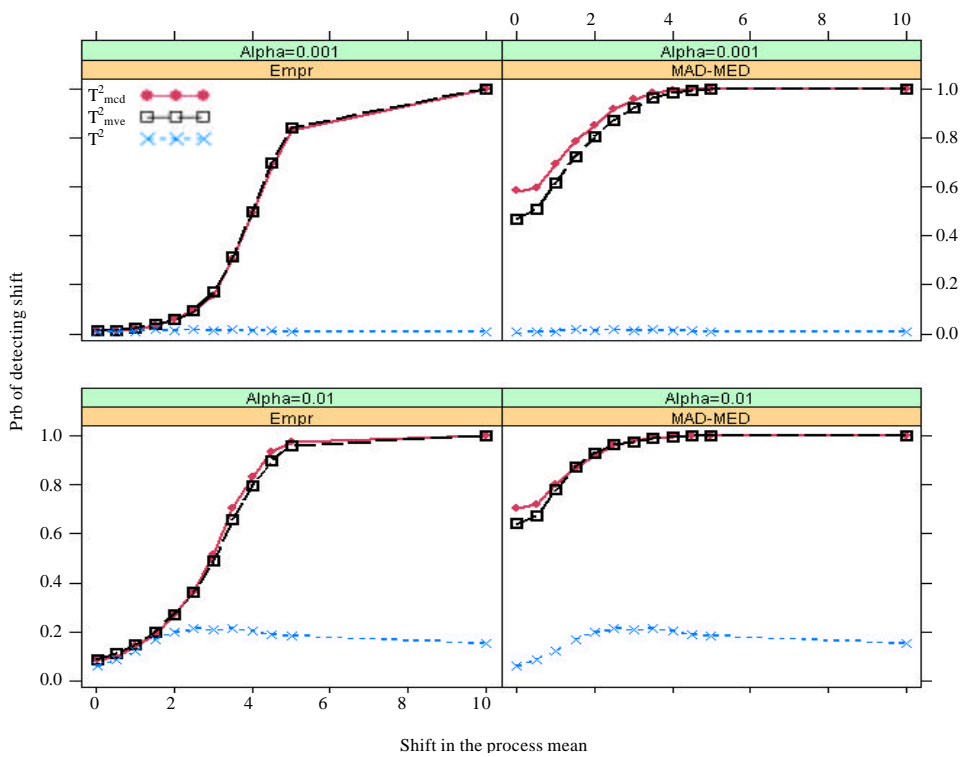


Fig. 4: Probability of a signal when $m = 50$ and 10 observations (20%) come from shifted distribution

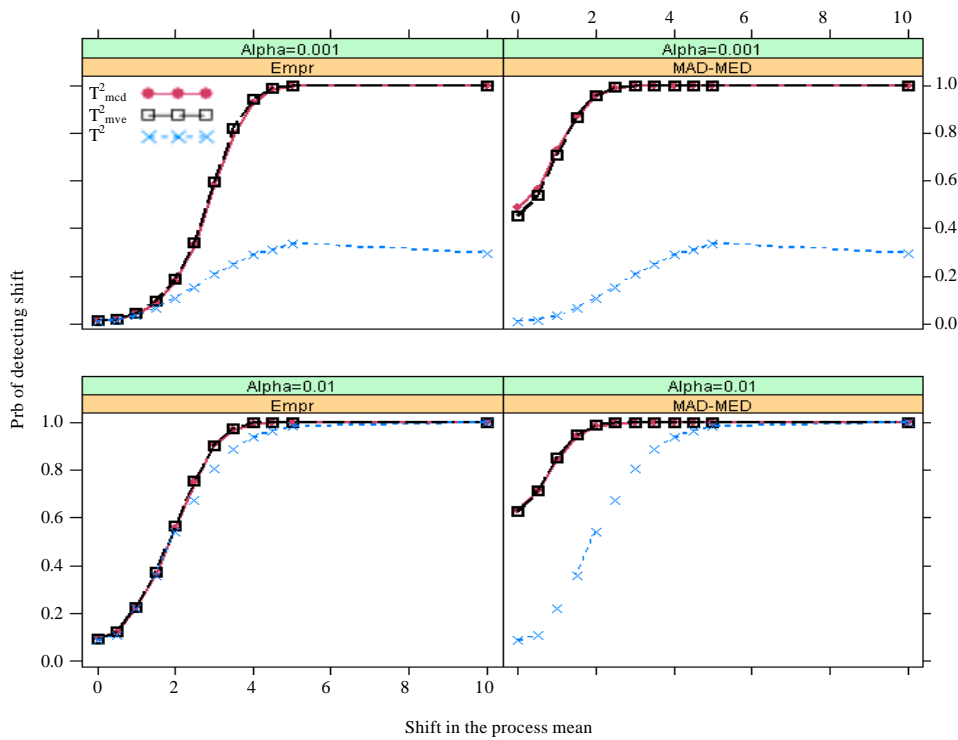


Fig. 5: Probability of a signal when $m = 100$ and 10 observations come (10%) from shifted distribution

Table 2: Number of exact detection out of 10 additional shifted observations

| $\mu_s (\delta, 0)$ | Upper Control Limits (UCLs) ($m = 20$) | | | | | Upper Control Limits (UCLs) ($m = 100$) | | | | |
|---------------------|--|------------|-------------|------------|-------------|---|------------|-------------|------------|-------------|
| | T^2 | EMPR | | MAD-MED | | T^2 | EMPR | | MAD-MED | |
| | | T^2_{mv} | T^2_{mcd} | T^2_{mv} | T^2_{mcd} | | T^2_{mv} | T^2_{mcd} | T^2_{mv} | T^2_{mcd} |
| 00.0 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 1 | 1 |
| 00.5 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 1 | 1 |
| 01.0 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 1 | 1 |
| 01.5 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 2 | 2 |
| 02.0 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 3 | 3 |
| 02.5 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 1 | 5 | 5 |
| 03.0 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 1 | 6 | 7 |
| 03.5 | 0 | 0 | 0 | 0 | 2 | 0 | 3 | 3 | 8 | 8 |
| 04.0 | 0 | 0 | 0 | 0 | 2 | 0 | 5 | 4 | 9 | 9 |
| 04.5 | 0 | 0 | 0 | 0 | 2 | 0 | 7 | 6 | 10 | 10 |
| 5.0 | 0 | 0 | 0 | 0 | 2 | 0 | 8 | 8 | 10 | 10 |
| 10.0 | 0 | 0 | 0 | 0 | 2 | 0 | 10 | 10 | 10 | 10 |

Both robust charts (based on both empirical methods) are able to detect the exact number of shifted observations in the data for large shifts ($\delta = 10$). The performance of T^2 chart is very poor because it cannot detect any shifted observations even at a larger shifted values.

NUMERICAL EXAMPLE

Here, we wanted to show the application of the robust and the classical quality control charts and

different cutoff- points to a Malaysian Palm Oil data set taken from Shi (2006).

It presents two quality characteristics of 54 observations namely the free fatty acid and dirt. We confine our study only on sample size equals to one. Hence, we only refer to the first column of the respective data. Firstly, the distribution of the Hotelling's T^2 statistics is checked by using the Q-Q plot to confirm whether it really follows a chi square distribution. The characteristics of the Q-Q plot corresponding to the Chi-square distribution should follow a linear trend.

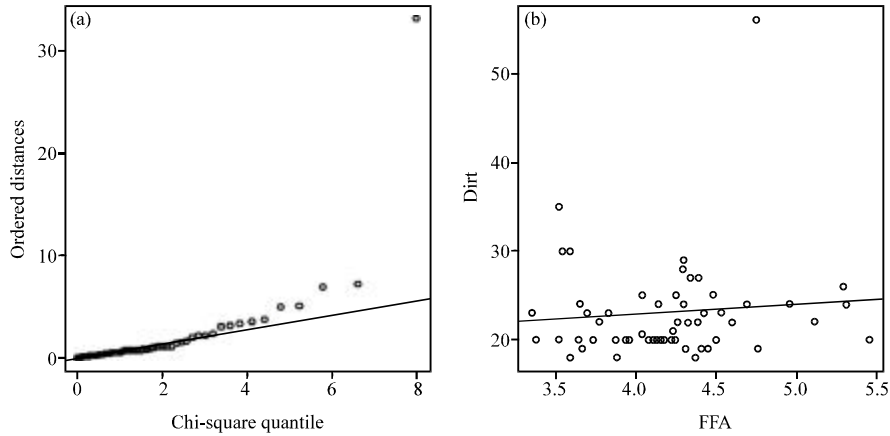


Fig. 6: (a) Q-Q plot of T^2 values of palm oil data and (b) scatter plot of palm oil data variables

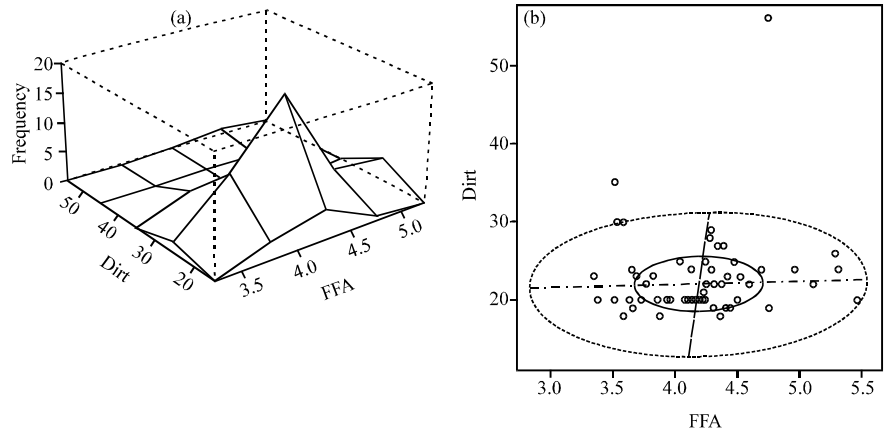


Fig. 7: (a) Density plot and (b) Robust bivariate box plot of palm oil data

A little deviation from linear trend at the end point can be acceptable to say that the Hotelling's T^2 is approximately distributed as chi square. By looking at Fig. 6a, we can see that the bulk of the data follows a linear trend but there is a severe departure in the upper tails of the T_1^2 plot. Hence, it is very obvious that the Hotelling's T^2 does not follow a Chi-square distribution. Moreover, there is an indication of four outliers present in the data. The data is further explored by plotting a scatter plot of free fatty acid against dirt and displayed in Fig. 6b. The plot suggests that there are four outliers in the data set. A further helpful enhancement to the scatter plot is often provided by the two-dimensional analogue of the box plot for univariate data, known as the bivariate box plot (Goldberg and Iglewicz, 1992). This type of box plot may be useful in indicating the distributional properties of the data and in identifying possible outliers as a primary investigation.

The bivariate box plot is based on calculating robust measures of location, scale and correlation. It can be

Table 3: The empirical UCLs for numerical example

| UCLs | $\alpha_1 = 0.001$ | | $\alpha_2 = 0.01$ | |
|---------|--------------------|-------|-------------------|-------|
| | MVE | MCD | MVE | MCD |
| EMPR | 20.75 | 33.50 | 12.62 | 19.10 |
| MAD-MED | 5.23 | 6.56 | 5.62 | 7.05 |

observed from the density plot and robust bivariate box-plot presented in Fig. 7a and b, that the distribution of the data is not normal and several apparent outliers can be observed.

The exploratory data analysis on the palm oil data suggested that there are approximately four or more outliers in the data. Now we would like assess the performance of the Hotelling's T^2 and robust charts on the palm oil data based on three cutoff-points, namely the chi-square, empirical and MAD-MED cutoff-points. We wanted to know whether these charts are capable of detecting the outliers. Here, the simulated UCLs in Table 1 are used. By using linear interpolation, the UCLs for $m = 54$ are presented in Table 3.

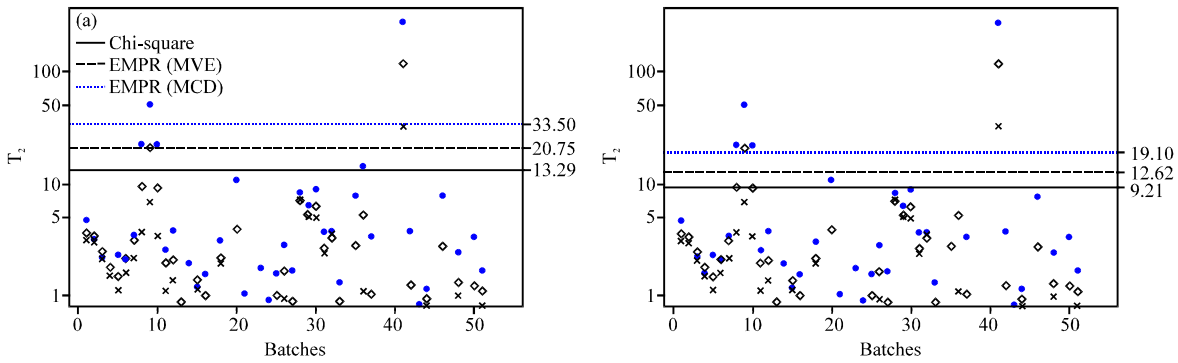


Fig. 8: T^2 control charts for palm oil data by using the EMPR UCLs, T^2 Hotelling is crosses, T^2_{mve} is empty diamond and T^2_{mcd} is solid circle, (a) $\alpha = 0.001$ and (b) $\alpha = 0.01$

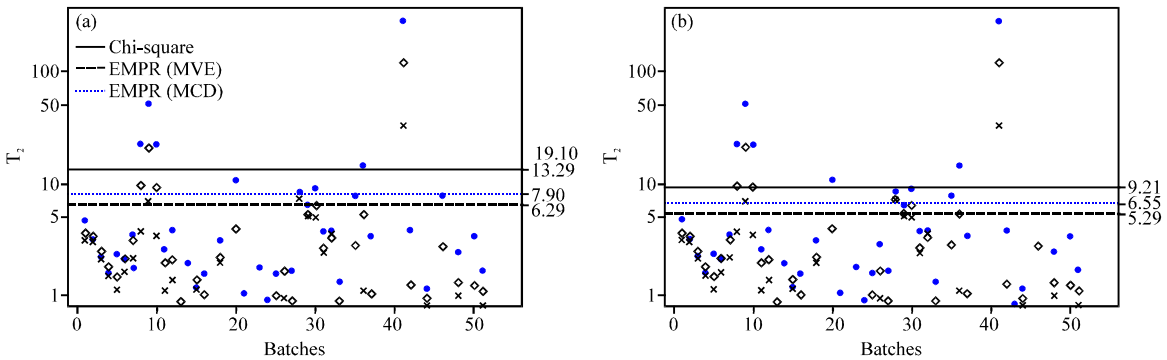


Fig. 9: T^2 Control charts for palm oil data by using the MAD-MED UCLs, T^2 Hotelling is cross, T^2_{mve} is empty diamond and T^2_{mcd} is solid circle, (a) $\alpha = 0.001$ and (b) $\alpha = 0.01$

The respective control charts are presented in Fig. 8a, b and 9a, b. The Hotelling's T^2 , the MCD- RT_1^2 chart and the MVE- RT_1^2 based on the EMPR UCLs detected 1, 2, 2 outliers respectively at the first alpha level and 1, 2, 4 at the second alpha level. On the other hands, at the first significant level the MVE- RT_1^2 based on the MAD-MED UCLs detected 5 outliers whereas, the MCD- RT_1^2 detected 8 outliers.

The results also suggested that the MVE- RT_1^2 and MCE- RT_1^2 based on the MAD-MED cutoff-points identified 8 and 10 outliers respectively at $\alpha = 0.01$. These results suggested that UCLs which were determined by the MAD-MED methods detect more outliers than when using the EMPR cutoff-points. It is clear that the Hotelling's T^2 detection capability is very poor.

CONCLUSIONS

The performance of the Hotelling's T^2 control chart based on the chi-square cutoff -point cannot be

recommended when outliers or sustained shifts in mean are present in the data. In this situation, the Hotelling's T^2 is no longer distributed as Chi-squares.

By erroneously assumed that the distribution of Hotelling's T^2 is chi-squared, the chi-square cutoff-points is used for the upper control limit. Consequently this control chart produces misleading conclusion. To rectify this problem, we proposed a robust control charts which are based on the empirical cutoff-points and the MAD-MED cutoff- points. On the basis of the findings of this research, the performance of the RT_1^2 charts based on MAD-MED cutoff-points generally are better than the RT_1^2 charts based on EMPR cutoff-points. Nonetheless their performance are reasonably close to the MAD-MED cutoff-points for large shift magnitudes in the mean vector.

The MCD- RT_1^2 control chart is preferred over the MVE- RT_1^2 based on the MAD-MED UCLs as it has better detection capability. However, their performances become closer as the percentage of shifted observations decreases (approximately up to 10%).

REFERENCES

- Abu-Shawiesh, M.O., F.M. Al-Athari and H.F. Kittani, 2009. Confidence interval for the mean of a contaminated normal distribution. *J. Applied Sci.*, 9: 2835-2840.
- Alt, F.B., 1985. *Multivariate Quality Control in Encyclopedia of Statistical Sciences*. John Wiley and Sons, New York.
- Alt, F.B. and N.D. Smith, 1988. *Multivariate Process Control*. In: *Handbook of Statistics*, Krishnaiah, P.R. and C.R. Rao (Eds.). Vol. 7. Elsevier, North-Holland, Amsterdam, pp: 333.
- Deming, W.E., 2000. *Out of the Crisis*. The MIT Press, Cambridge.
- Goegebeur, Y., V. Planchon, J. Beirlant and R. Oger, 2005. Quality assessment of pedochemical data using extreme value methodology. *J. Applied Sci.*, 5: 1092-1102.
- Goldberg, K.M. and B. Iglewicz, 1992. Bivariate extensions of the boxplot. *Technometrics*, 34: 307-320.
- Habshah, M., M.R. Norazan and A.H.M.R. Imon, 2009. The performance of diagnostic-robust generalized potentials for the identification of multiple high leverage points in linear regression. *J. Applied Statist.*, 36: 507-520.
- Hadi, A.S., 1992. A new measure of overall potential influence in linear regression. *Comput. Statist. Data Anal.*, 14: 1-27.
- Hawkins, D.M., 1980. *Identification of Outliers*. 1st Edn., Chapman and Hall, London, New York, ISBN-13: 9780412219009.
- Hotelling, H., 1931. The generalization of student's ratio. *Ann. Mathe. Statist.*, 2: 360-378.
- Jensen, W.A., J.B. Birch and W.H. Woodall, 2007. High breakdown estimation methods for phase I multivariate control charts. *Qual. Reliabil. Eng. Int.*, 23: 615-629.
- Johnson, R.A. and D.W. Wichern, 2002. *Applied Multivariate Statistical Analysis*. 6th Edn., University of Wisconsin Press, Madison.
- Montgomery, D.C., 2006. *Introduction to Statistical Quality Control 2005*. John Wiley and Sons, New York.
- Rousseeuw, P. and B. van Zomeren, 1990. Unmasking multivariate outliers and leverage points. *J. Am. Stat. Assoc.*, 85: 633-639.
- Rousseeuw, P. and K. Van Driessen, 1999. A fast algorithm for the minimum variance determinant estimator. *Technometrics*, 41: 212-223.
- Rousseeuw, P.J. and A.M. Leroy, 2003. *Robust Regression and Outlier Detection*. 1st Edn., John Wiley, New York, ISBN: 978-0471488552, pp: 360.
- Shewart, W.A., 1931. *Economic Control of Quality of Manufactured Product*. D. Van Nostrand Company, New York.
- Shi, T.H., 2006. *The comparison between classical and robust methods on process capability indices and multivariate control charts*. M.Sc. Thesis, University Putra Malaysia
- Sullivan, J.H. and W.H. Woodall, 1996. A comparison of multivariate control charts for individual observations. *J. Qual. Technol.*, 28: 398-408.
- Tracy, N.D., J.C. Young and R.L. Mason, 1992. Multivariate control charts for individual observations. *J. Qual. Technol.*, 24: 88-95.
- Vargas, J.A. and C.J. Lagos, 2007. Comparison of multivariate control charts for process dispersion. *Qual. Eng.*, 19: 191-196.
- Wilcox, R.R. and H.J. Keselman, 2003. Modern robust data analysis methods: measures of central tendency. *Psychol. Methods*, 8: 254-274.