Punjab University
Journal of Mathematics (ISSN 1016-2526)
Vol. 50(1)(2018) pp. 97-112

Improvement of The Hotelling's T^2 Charts Using Robust Location Winsorized One Step M-Estimator (WMOM)

Firas Haddad
College of Applied Studies and Community Service,
University of Dammam, KSA.
Email: fshaddad@uod.edu.sa

Mutasem K Alsmadi
College of Applied Studies and Community Service,
University of Dammam, KSA.
Email: mksalsmadi@gmail.com & mkalsmadi@uod.edu.sa

Received: 23 January, 2017 / Accepted: 05 April, 2017 / Published online: 10 November, 2017

Abstract. For product manufacturing, control charts are important tools. Different types of control charts are used to monitor different measures of product characteristics. Among the numerous charts, the popular one is the Hotelling's T^2 chart. It is designed while taking into account the most sensitive measures such as sample covariance matrix and sample mean vector. However, the chart becomes ineffective in the presence of outliers. This work proposes a new robust chart that overcomes the problems associated with other control charts. Our chart uses robust scale estimator, Q_n instead of the covariance matrix and Winsorized modified one step M-estimator (*WMOM*) in place of the sample mean vector. The robust chart could be removed without losing anything in the explanation. There are two main performance measurements, false alarm and probability of detection outliers. The results indicate that robust chart's performance is superior to that of the conventional control chart.

AMS (MOS) Subject Classification Codes: 35S29; 40S70; 25U09
Key Words: Robust Location Estimator *MOM* and Robust Scale Estimator Q_n Hotelling's T^2 Control Chart.

1. Introduction

The production process has many characteristics; the quality of the product heavily depends upon these characteristics. Therefore, it is important to pay close attention to monitor the complicated production process. Due to the diversity of the characteristics along with their impact on the overall product's quality, statistics are computed in different ways in order to make the production process better. One of the most famous and effective ways to monitor processes is using Control charts.

These control charts act as a statistical tool that has the ability to monitor the accomplishment and progress of the entire production process. In real time, the quality of the products does not solely depend on one factor. On the contrary, it is a contribution of different characteristics. Several characteristics should be monitored simultaneously, thus, we use Hotelling's T^2 and Shewhart-type χ^2 control charts.

The multivariate generalization of the conventional Student's *t*-statistic is known as the Hotelling's T^2. The formula is given below.

$$T^2 = n(\overline{X} - \mu_0)^T S^{-1}(\overline{X} - \mu_0) \tag{1.1}$$

In the formula above, the scatter *S*, the sample covariance matrix; it provides us with information regarding the relationship between observation vectors of different variables. The values that lie in the main diagonal of *S* matrix are the variance of each variable, whereas the other elements that are part of the *S* matrix denote the covariance between different pairs of variables. The distance between the process center and the observation is measured by Hotelling's T^2statistics observation vector's values [1].

The individual multivariate control chart will come into action and the formula for Hotelling,s T^2 will be as follows:

$$T_i^2 = (X_i - \overline{X})^T \overline{S}^{(-1)}(X_i - \overline{X}) \quad i = 1, ..., n \tag{1.2}$$

Here $\overline{X}$ and *S* represent the arithmetic mean vector and the covariance matrix estimators for the given sample.

The Hotelling's T^2 method is considered to be among the common methods in the multivariate statistical control charts [2, 3]. It might be able to produce efficient results when the data come from normal distribution, but it fails when outliers become part of the 'measurements of characteristics.' Thus, in this, we have replaced the variance covariance and arithmetic mean 5 with Q_n covariance and Winsorized modified one step M-estimator (*WMOM*).

In order to investigate and test the performance of this new robust control chart, we first generate the data from standard normal distribution and later on we put the outliers in accord with the contaminated normal model. This type of model relies on the types of variables such as dependent variables along with independent variables. Probability of detecting outliers and false alarm rate are the two measurements that judge the robustness of the control chart.

The conventional Hotelling's T^2 chart depends upon the conventional variance covariance and arithmetic mean matrix which makes them sensitive to the outliers. Therefore, in our modified control chart, we have used the robust estimators instead of the conventional means in order to overcome this problem. Consequently, many statisticians used these estimators to lessen the extreme values effect on the production process. For example, in [4, 5], the authors replaced the conventional arithmetic mean and variance co-variance with trimmed mean and trimmed covariance, respectively. In [6], the authors came up with the idea of using Hodges-Lehmann location estimator and Shamos-Bickel-Lehman's scale estimator. In 1994, robust location estimator was used as the median by [7]. An alternative for the conventional chart was proposed by [8]. According to their work, the trimmed mean vector was used instead of the arithmetic mean and the sample trimmed covariance was used in place of the sample variance covariance matrix. [9, 10] worked on the same model.

Furthermore, other statisticians used the minimum volume ellipsoid (*MVE*) and minimum covariance determinant (*MCD*) estimators in order to create alternative charts such as [11-13].

Likewise, Hanif et al. in [14] used trimmed means and deciles instead of the usual mean in the first phase of the conventional Hotelling's T^2 chart. This was done to develop the efficiency of the chart. Both the trimmed mean and decile depend upon the simulation and the results showed that the performance of the trimmed mean and decile Hotelling T^2 charts is better than that of the classical Hotelling's T^2 chart.

Winsorized covariance and Winsorized mean were used by [15]. Median absolute deviation (MAD_n) and S_n are the estimators that have the highest breakdown points. These estimators are good for the criterion in *MOM*. The results showed that these Hotelling's T^2 control charts outperformed the conventional Hotelling's T^2 chart.

Shabbak and Midi in [16] came up with a different methods to develop the performance of control charts. Their control chart used minimum co-variance determinant and minimum volume ellipsoid. There are many other researches who assessed the performance of their multivariate control charts based on the signals, and yet, they did not give much heed to whether those signals are outliers or not. These researches proposed the idea to evaluate control charts not only based on correct positions, but also on the number of outliers found. The idea was the upper control limit (*UCL*); this limit was established based on the median along with the median absolute deviation (*MAD*). This idea did improve the detection of the outliers, but it was affected by the swapping effect that occurs when outlier's positions are not taken into consideration. Nevertheless, a robust control chart is presented in order to get over the drawbacks mentioned above. This generalized potential method tends to produce better results.

AbuShawiesh et al. in [17] used the below robust alternatives to the conventional Hotelling's T^2 : T^2*MVE*, T^2*MCD* and median of Median Absolute Deviation (T^2*MedMAD*). To compare the performances of these control charts, a simulation study is conducted. To determine the strength and the performance of these charts, two sets of real data are analyzed.

The contribution of this work is the introduction of a new control chart that overcomes the problems associated with other control charts. Our chart uses the high breakdown robust scale estimator, Q_n instead of the covariance matrix and the *WMOM* in place of the sample mean vector, which is suitable when there are asymmetrical data. Furthermore, the paper suggests new and better technique to determine the control limits through an accustomed function using different dimensions of the process and different sample sizes.

2. Robust location and Scale Estimators and Robust Control Chart

Take a random sample of p variables from the distribution and let that be $X_{ij} = \{X_{1j}, \ldots, X_{nj}\}, j = 1, \ldots, p$; Wilcox and Keselman in [18] defined the *MOM* estimator as,

$$\widehat{\theta} = \sum_{i=i_1+1}^{n-i_2} \frac{X_{(i)j}}{n_j - i_1 - i_2} \tag{2.3}$$

Here,$X_{(i)j}$: $i - th$ to order statistics in $j - th$ characteristic variable.
i_1 : The number of X_{ij} that satisfies the criterion
i_2 : The number of X_{ij} that satisfies the criterion

$$(X_{ij} - \widehat{\mu}_j < -k * (Q_{nj}) \tag{2.4}$$

$$(X_{ij} - \widehat{\mu}_j > k * (Q_{nj}) \tag{2.5}$$

$n_j =$ Number of observations in each $j - th$ variable
$\widehat{\mu}_j = med\{X_{1j}, \ldots, X_{nj}\}, j = 1, \ldots, p$

$$Q_n = d\{|x_i - x_j|; i < j\} \tag{2.6}$$

Yahaya et al. in [20] reported that if different trimming criterion in *MOM* is used, it results in a robust scale estimators such as $Q_n = d\{|x_i - x_j|; i < j\}$. These estimators have the ability to improve false alarms probability of some statistics tests. Therefore, we have used Q_n instead of *MAD*$_n$ in the trimming criterion. In *MOM* estimator, the defaulting robust scale estimator for the trimming criterion *MAD*$_n$ is used.

Croux and Rousseeuw in [19] proposed the highly robust estimator Q_n which is as follows:

$$Q_n = d\{|x_i - x_j|; i < j\} \tag{2.7}$$

Here, *d* acts as a constant factor that is used to make Q_n unbiased estimator for the value of σ when the distribution is normal. $h = \dfrac{n}{2} + 1$represents half of the number of observations.

Q_n has the same properties of S_n, which is a definition that is equally good for 50% breakdown point and asymmetric distribution. In addition, we will notice that its influence function is very close and very competent when used with a Gaussian distribution, at almost 82%. For more discussion about the Q_n , see [20].

Once the elimination of outliers is done for every sample through the criteria defined in inequalities 2.4 & 2.5, the data are Winsorized as the following:
For each random variable, $X_{ij} = \{X_{1j}, \ldots, X_{nj}\}, j = 1, \ldots, p$ the Winsorized sample is made as below:

$$W_{ij} = \begin{cases} X_{(i_1+1)j}, & \text{if } X_{ij} \leq X_{(i_1+1)j} \\ X_{ij}, & \text{if } X_{(i_1+1)j} < X_{ij} < X_{(n-i_2)j} \\ X_{(n-i_2)j}, & \text{if } X_{ij} \geq X_{(n-i_2)j} \end{cases} \tag{2.8}$$

where
i_1 : Denote to the smallest outliers in the data.
i_2: Denote to the largest outliers in the data

Thus, the Winsorized *MOM* for ***j-th*** variable and the estimated Winsorized covariance matrix between W_i and W_j variables are defined as follows:

$$\overline{W}_j = \frac{1}{m_j} \sum_{i=1}^{m_j} W_{ij} \tag{2.9}$$

The vector of Winsorized *MOM* estimator is given below:
$\overline{W} = \begin{bmatrix} \overline{W}_1 \\ . \\ . \\ . \\ \overline{W}_p \end{bmatrix}$ and

$$S_{WQn}(W_i, W_j) = \frac{1}{(n-1)}[\sum_{k=1}^{n} W_{ki}W_{kj} - n\overline{W}_i\overline{W}_j] \tag{2.10}$$

To create the alternative of the conventional Hotelling's T^2 chart, we have used the Winsorized *MOM* denoted by $\overline{W}$ instead of the usual mean vector and replaced the inverse of the covariance matrix with the inverse of Winsorized covariance matrix $S^{-1}_{WQ_n}$. Thus, the new robust chart that was used in this work is written as follows:

$$T^2_{WQ_n}(X_i) = (X_i - \overline{W})^T S^{-1}_{WQ_n}(X_i - \overline{W}) \tag{2.11}$$

Where $S^{-1}_{WQ_n}$: inverse of the sample covariance matrix.

3. CONTROL LIMITS

Since we are not aware of the alternative Hotelling's T^2 distribution, we will control the upper control limit for every proposed alternative control chart. The *UCL* can be determined by the simulation [11, 12, 21]. We generated 5000 data sets from the $N_p(0, I_p)$ at the value of nominal false alarm α = 0.05. In the same phase, i.e. I, we calculate the robust and conventional estimators for every data set. In phase II, we work on generating extra observation for every data set. Then, we also calculate the robust and conventional charts for these additional observations by using the estimators that were calculated in the previous phase. We calculate the *UCL* for the conventional Hotelling's T^2 through the formula defined in equation 1.2 whereas the *UCL* for the robust control chart is calculated by 95^{th} percentile of the simulation values of robust statistics for 5000 replications.

4. SIMULATION DESIGN AND THE RESULTS

In order to highlight the merits and limitations of different control charts, they were compared and investigated for their probability of detection and false alarm under different conditions. The sample sizes were *n*=50, 100 and 150 observations and the numbers of quality characteristics were *p*= 2, 5 and 10. Due to the fact that the characteristics variables can be of both types, i.e. dependent as well as independent, this study incorporates all kinds of variables.

$$(1-\varepsilon)N_p(0, I_p) + \varepsilon N_p(\mu_1, I_p) \tag{4.12}$$

The proportion of the outlier data is represented by $\varepsilon N_p(\mu_1, I_p)$ and $N_p(0, I_p)$ as the out of control and in control distribution, respectively. If we do not lose the generality, then the value for the in control mean parameter is set to be zero and the value of the out of control parameter is set to be $\mu_1.\mu_1$ took the value zero in case of no change, 3 in case of moderate of leverage points and 5 in case of very high leverage points. Nonetheless, the identified matrix I_p, acts as the dispersion matrices, in both the distributions. This indicates that there is no correlation among the *p* variables. In cases where the variables are dependent, the data is generated according to the formula given below.

$$(1-\varepsilon)N_p(0, \sum\nolimits_0) + \varepsilon N_p(\mu_1, \sum\nolimits_0) \tag{4.13}$$

The proportion of the outliers data is represented by $\varepsilon N_p(\mu_1, I_p)$ and $N_p(0, I_p)$ as the out of control and in control distribution, respectively. If we do not lose the generality, then the value for the in control mean parameter is set to be 0 and the value of the out of control parameter is set to be $\mu_1.\sum_0$ is the covariance matrix for both out of control and in control distribution. $\sum_0$ represents the homogenous covariance matrix and its size is *pxp*. It has a high correlation between the variables, where all of the elements of the main diagonal are 1's and the 0.9's for the off diagonal [22-24].

Since this paper mainly focuses on the contamination and outlying of the observation, we have considered the outliers with Gaussian distribution to check the performances of the control charts and to see whether there is any violation of the normality assumption.

In phase I, we estimate the control parameters along with the limits that are used to enhance the charts. The process is defined as below:

(1) Characteristic p took the values 2, 5 and 10. We generate this from the formula defined in equation 4.12 and equation 4.13 in accord with the cases of independent and dependent variability.
(2) Computation of both the robustness and conventional scale and location estimators for every set of data used as an estimation of the in control parameters.

The phase II comprises the computation of the probability of detection of outliers and false alarms according to the estimations that were computed in phase I.

(1) Computation of the conventional and robust Hotelling's T^2 for every new observation that was randomly generated from the in control distribution using formula 2.11. The location and scale estimators in the formula 2.11 are computed in phase I.
(2) Computation of the conventional and robust Hotelling's T^2 for every new observation that was randomly generated from the out of control distribution using formula 2.11. The location and scale estimators in the formula 2.11 are computed in phase I.
(3) Then, the statistics in steps 1 and 2 are compared with the limits that are calculated during the simulation process.
(4) The probabilities of detecting outliers and false alarms represent the proportion of the number of the statistical values that are greater than the limits.
(5) The simulation is excecuted using MATLAB 2009a.

Tables 1-3 demonstrate the results of the investigation. They are arranged according to the number of characteristics variables *p*. Every table consists of two cases; the independent variables case, which is denoted by case (A) and the dependent variables case, which is denoted by case (B). The group sizes are displayed in the first column; the second column represents the proportion of outliers and the third column shows the shifted means vectors values. The performances of the two methods are represented in the last column. T^2 and $T^2_{WQ_n}$ are the procedures that represent the control chart of the customary Hotelling's T^2 for the charts of historical data sets. They also represent the control charts for the alternative robust chart. The false alarm in the brackets and the probability of detection outliers corresponding to each procedure are presented. When the values of the false alarms are near to the nominal value, α=0.05, the chart shows stronger performance of false alarms. If the probabilities are nearer to one, then the performance of the charts are stronger in terms of detection outliers.

Table (1) demonstrates the bivariate case with independent variables. As shown by the table, the performance of the robust chart is under control, in terms of false alarms. Whereas, the performance of the conventional chart tends to decline as the rates of outliers are increased. Meanwhile, the mean is shifted, regardless of the sample size. The

TABLE 1. The Detection of Outliers Probability and the Empirical Rates of False Alarm for the Robust Chart, when *p*=2, alpha=0.05.

Case	*n*	ε	μ	$T^2_{\overline{X}-S}$	$T^2_{\overline{X}_{mom-Q_n}}$
A	50	0	(0,0)	0.059	0.052
		0.1	(3,3)	(0.029) 0.533	(0.019) 0.742
			(5,5)	(0.023) 0.794	(0.0.02) 1
		0.2	(3,3)	(0.028) 0.145	(0.017) 0.326
			(5,5)	(0.024) 0.135	(0.019) 0.964
	100	0	(0,0)	0.056	0.05
		0.1	(3,3)	(0.02) 0.491	(0.025) 0.80
			(5,5)	(0.016) 0.747	(0.024) 1
		0.2	(3,3)	(0.021) 0.176	(0.013) 0.35
			(5,5)	(0.016) 0.171	(0.024) 0.979
	150	0	(0,0)	0.055	0.045
		0.1	(3,3)	(0.021) 0.522	(0.024) 0.78
			(5,5)	(0.016) 0.806	(0.024) 1
		0.2	(3,3)	(0.021) 0.175	(0.02) 0.32
			(5,5)	(0.016) 0.175	(0.02) 0.968
B	50	0	(0,0)	0.059	0.047
		0.1	(5,5)	(0.029) 0.459	(0.029) 0.738
		0.2	(5,5)	(0.028) 0.147	(0.019) 0.488
	100	0	(0,0)	0.056	0.041
		0.1	(5,5)	(0.023) 0.429	(0.029) 0.774
		0.2	(5,5)	(0.023) 0.173	(0.029) 0.534
	150	0	(0,0)	0.055	0.038
		0.1	(5,5)	(0.024) 0.457	(0.025) 0.746
		0.2	(5,5)	(0.022) 0.174	(0.019) 0.482

probability of detecting outlier for the conventional chart is less than the probability of detecting outliers for the robust chart. In some cases, the probability of the detection outlier for the robust chart is up to 1. Thus, we conclude that in bivariate case and independent variables, the performance of robust chart is better when compared with the performance of the conventional charts. Likewise, in terms of the detection outliers, the robust chart's performance is superior to the conventional chart's performance except when both the μ and ε have smaller values.

Despite the fact that the performance of the robust chart declines in case of dependent variables, they are still able to control the false alarm rates. The rates of the conventional chart detection outliers are smaller when compared to the rates of the detection outliers of the robust chart. This implies that the robust chart outperformed the conventional control chart.

TABLE 2. The Detection of Outliers Probability and the Empirical Rates of False Alarm for the Robust Chart, when *p*=5, alpha=0.05.

Case	*n*	ε	μ	$T^2_{\overline{X}-S}$	$T^2_{\overline{X}_{mom-Q_n}}$
A	50	0	(0,0,0,0,0)	0.051	0.046
		0.1	(3,3,3,3,3)	(0.03) 0.405	(0.026) 0.81
			(5,5,5,5,5)	(0.031) 0.468	(0.025) 1
		0.2	(3,3,3,3,3)	(0.032) 0.099	(0.035) 0.255
			(5,5,5,5,5)	(0.033) 0.101	(0.037) 0.986
	100	0	(0,0,0,0,0)	0.053	0.05
		0.1	(3,3,3,3,3)	(0.027) 0.364	(0.027) 0.865
			(5,5,5,5,5)	(0.026) 0.442	(0.025) 1
		0.2	(3,3,3,3,3)	(0.026) 0.119	(0.038) 0.23
			(5,5,5,5,5)	(0.025) 0.109	(0.036) 0.981
	150	0	(0,0,0,0,0)	0.054	0.039
		0.1	(3,3,3,3,3)	(0.029) 0.417	(0.021) 0.857
			(5,5,5,5,5)	(0.028) 0.507	(0.021) 1
		0.2	(3,3,3,3,3)	(0.030) 0.118	(0.021) 0.337
			(5,5,5,5,5)	(0.029) 0.120	(0.021) 0.971
B	50	0	(0,0,0,0,0)	0.051	0.038
		0.1	(5,5,5,5,5)	(0.039) 0.176	(0.018) 0.247
		0.2	(5,5,5,5,5)	(0.03) 0.086	(0.02) 0.15
	100	0	(0,0,0,0,0)	0.053	
		0.1	(5,5,5,5,5)	(0.038) 0.153	(0.016) 0.258
		0.2	(5,5,5,5,5)	(0.036) 0.091	(0.021) 0.147
	150	0	(0,0,0,0,0)	0.054	0.027
		0.1	(5,5,5,5,5)	(0.038) 0.182	(0.018) 0.229
		0.2	(5,5,5,5,5)	(0.037) 0.094	(0.016) 0.124

Table (2) displays the results when the value of *p* =5. The robust chart has stronger performance than the performance of the conventional chart in terms of the detection of outliers and false alarms in most cases of variables. However, in the case of the dependent variables, and despite that it is good in terms of false alarms, the robust along with the conventional control chart fail to detect outliers. The reason behind this failure is that the rates of detection of outliers are very low.

The results when the value of *p*=10 is shown in Table (3). In case of independent variables, the false alarm probabilities are in control for the robust charts, but the performance deteriorates when the proportion of outliers ε and μ increase in terms of probability of detection. Nonetheless, conventional chart control on the false alarm anyway the change in the proportion of outliers ε. Despite the fact that the probabilities of false alarm values increase far from the nominal α, the modified charts have quite high values of probabilities for detecting outliers. Lastly, the robust charts are able to control false alarm probability

TABLE 3. The Detection of Outliers Probability and the Empirical Rates of False Alarm for the Robust Chart, when *p*=10, alpha=0.05.

Case	*n*	ε	μ	$T^2_{\overline{X}-S}$	$T^2_{\overline{X}_{mom-Q_n}}$
A	50	0	(0,0,0,0,0,0,0,0,0,0)	0.049	0.048
		0.1	(3,3,3,3,3,3,3,3,3,3)	(0.031) 0.244	(0.04) 0.655
			(5,5,5,5,5,5,5,5,5,5)	(0.032) 0.259	(0.038) 1
		0.2	(3,3,3,3,3,3,3,3,3,3)	(0.033) 0.082	(0.048) 0.21
			(5,5,5,5,5,5,5,5,5,5)	(0.032) 0.082	(0.055) 0.949
	100	0	(0,0,0,0,0,0,0,0,0,0)	0.057	0.057
		0.1	(3,3,3,3,3,3,3,3,3,3)	(0.041) 0.247	(0.035) 0.808
			(5,5,5,5,5,5,5,5,5,5)	(0.038) 0.265	(0.032) 1
		0.2	(3,3,3,3,3,3,3,3,3,3)	(0.042) 0.106	(0.043) 0.211
			(5,5,5,5,5,5,5,5,5,5	(0.040) 0.106	(0.0541) 0.96
	150	0	(0,0,0,0,0,0,0,0,0,0)	0.055	0.044
		0.1	(3,3,3,3,3,3,3,3,3,3)	(0.033) 0.294	(0.036) 0.837
			(5,5,5,5,5,5,5,5,5,5)	(0.034) 0.314	(0.036) 1
		0.2	(3,3,3,3,3,3,3,3,3,3)	(0.035) 0.107	(0.043) 0.203
			(5,5,5,5,5,5,5,5,5,5)	(0.034) 0.108	(0.048) 0.942
B	50	0	(0,0,0,0,0,0,0,0,0,0)	0.049	0.023
		0.1	(5,5,5,5,5,5,5,5,5,5)	(0.034) 0.077	(0.02) 0.091
		0.2	(5,5,5,5,5,5,5,5,5,5)	(0.032) 0.063	(0.014) 0.056
	100	0	(0,0,0,0,0,0,0,0,0,0)	0.047	0.03
		0.1	(5,5,5,5,5,5,5,5,5,5)	(0.038) 0.106	(0.025) 0.102
		0.2	(5,5,5,5,5,5,5,5,5,5)	(0.04) 0.08	(0.025) 0.068
	150	0	(0,0,0,0,0,0,0,0,0,0)	0.045	0.025
		0.1	(5,5,5,5,5,5,5,5,5,5)	(0.036) 0.099	(0.019) 0.098
		0.2	(5,5,5,5,5,5,5,5,5,5)	(0.034) 0.078	(0.019) 0.062

under independent case, but it fails in the case of dependent variables.

Figures 1-3 represents the case of independent variables (case A) when the number of characteristics *p*=2, 5 and 10. We note that the new modified chart has stronger performance than the conventional chart in detecting the outliers, particularly when μ=5, regardless of the rate of the contaminated data. Whereas, there is a slight decline in the performance of the new modified chart when μ=3, and ε=0.1, and heavy decline when μ=3, and ε=0.2.

In the case of dependent variables (case B), through Figures 4 to 6, there was a significant decline in the performance of the detection of the outliers data. This applies to the traditional and the new robust charts although there is a simple superiority in the performance of the new robust charts especially when *p*=2, 5. While in the case of *p*=10, we note that the new chart was not superior to the traditional chart, which indicates that the new chart is not suitable in this case.

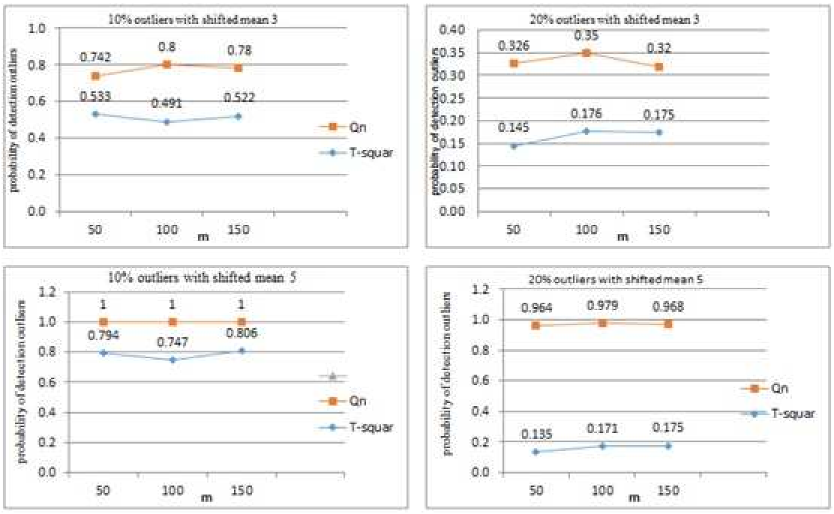

FIGURE 1. The Outliers Detection Rates in Case A (p=2)

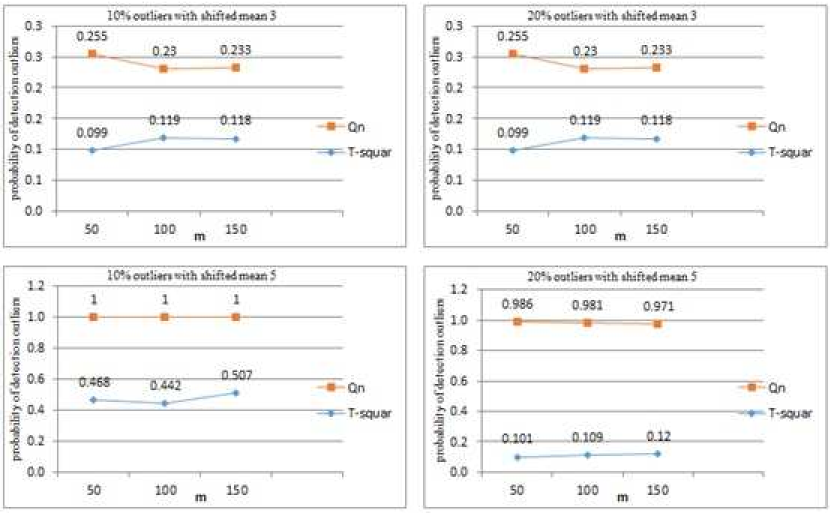

FIGURE 2. The Outliers Detection Rates in Case A (p=5)

5. PERFORMANCE OF THE MODIFIED CHART

Tables 1-3 show that in terms of false alarms, the robust Hotelling's T^2 charts outperformed the conventional control chart even when the number of variables is increasing and outliers are presented. The values of probabilities of detection outliers in the conventional

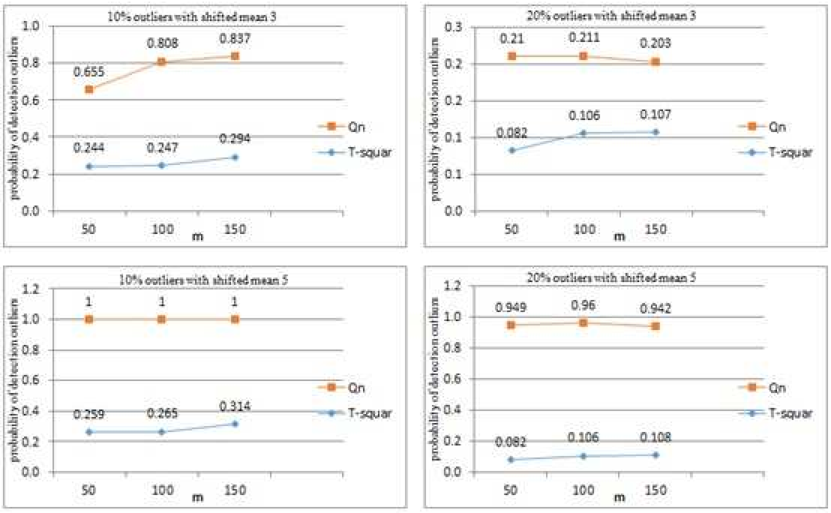

FIGURE 3. The Outliers Detection Rates in Case A (p=10)

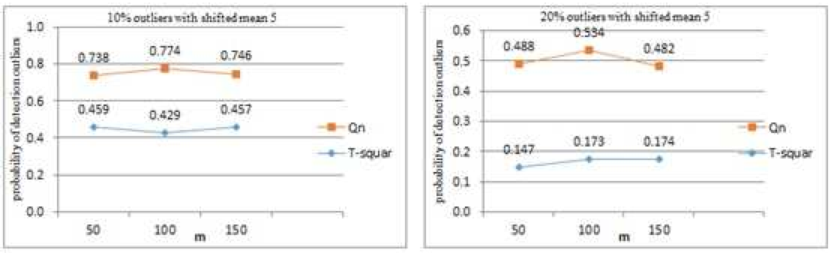

FIGURE 4. The outliers detection rates in case B (p=2)

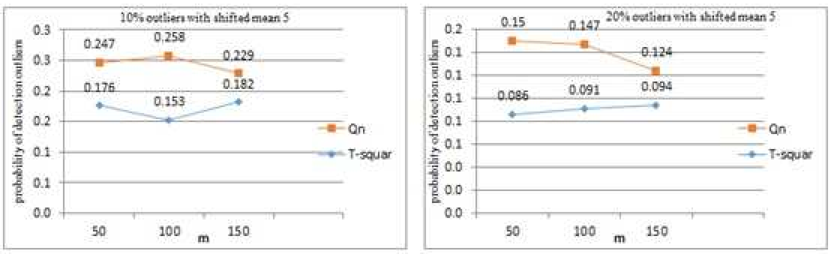

FIGURE 5. The outliers detection rates in case B (p=5)

control charts were not better than the probabilities values of detection outliers for the robust charts. Therefore, we can say that the robust chart outperformed the conventional

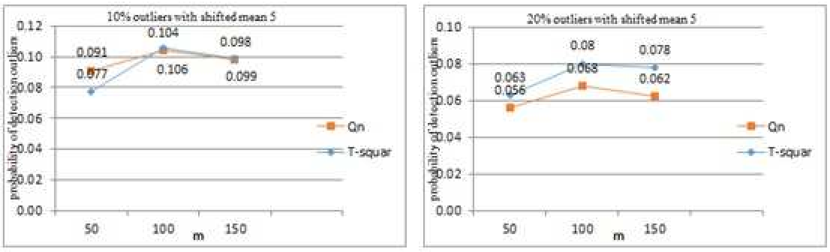

FIGURE 6. The outliers detection rates in case B (*p*=10)

Hotelling's T^2 chart. In addition, the modified Hotelling's T^2 charts' performance improves as the sample sizes increase because there are enhancements for the robust charts in terms of probability of detection outliers when sample size n increases. On the contrary, the performance of the conventional control chart in the outliers' detection declines once the samples size n starts to increase. In addition, with the increase of the sample sizes n, the rates of false alarm start to decrease and this implies that the modified Hoteling's T^2 performance starts to decline. For the rates of false alarm, increasing the number of variables is important, but it is not imperative for the probability of detection outliers, while the performance and results of the conventional control chart are bad in terms of probability detection outliers and false alarm.

The performance of the robust control chart is better when compared with the performance of conventional T^2 chart in case (B). The probability of detection outliers is smaller in the conventional control charts when compared to the robust Hotelling's T^2 control charts. Moreover, an increase in the sample size *n* results in increase in the probability of detection outlier and decrease in the values of false alarms. The increase in the number of variables *p* and subsequently the decrease in the probability of detection outliers imply that the performance of the modified Hotelling's T^2 is not affected by the increasing number of variables.

Lastly, in the capability for outliers' detection, we can say that the modified robust Hotelling's T^2 performance is better when compared to the conventional Hotelling's T^2 chart.

The obtained results were compared with the results for $T^2_{wMAD_n}$, $T^2_{wS_n}$ and $T^2_{wT_n}$ in [9], where the procedure of the simulation is approximately the same. The comparison illustrates the improvement of the proposed control chart when the number of characteristics variables *p* increases especially when *p*=5 and 10, where we noticed that case A has stronger performance in terms of false alarms when the sample sizes *n*= 50, 100 and *p*=10 whereas the mentioned robust charts are out of control. Seemingly, when *p* equals 5, the chart displays the same behavior. This implies that the performance of the robust control chart is so well. On the other hand, the proposed chart has stronger performance than the performance of the three mentioned charts in detecting outliers regardless of the values of *p, n* and the shifted mean μ.

To investigate the proposed chart, another comparison has been performed in [12]. They used other types of robust estimators in their proposed charts such as the minimum volume ellipsoid (*MVE*) estimators, minimum covariance determinant (*MCD*) estimators and reweighted *RMCD* estimators. The obtained results showed inconsistency between the percentages of outliers' detection and the overall false alarm rates. For example, for the values of μ, ε and the sample sizes n, the false alarm rates dropped from the nominal value 0.05 mean while the probability of detection increased. However, the proposed chart $T^2_{\overline{X}_{mom-Q_n}}$ is consistent in terms of detecting outlier's probability and controlling false alarm rates. Even though the conventional Hoteling T^2 chart performs well in terms of controlling false alarm rates, it fails to achieve good probability of detection, especially in case of large number of quality characteristics.

6. EMPIRICAL CASE

We used the example from [13] in order to compare and evaluate results of the performance of both the conventional and modified control charts. Their data comprises of two characteristics random variables, namely and on 30 different products taken from the production process. In Vargas, Queensberry datasets two variables were used and we choose the first characteristic values of the first twenty-five products. The observations of both random variables are shown in Table (4). The table also shows the values of the new Hotelling's T^2 along with the conventional T^2 chart statistics.

The scale and the location estimators are also shown in Table (4) presented in the appendix.

We calculated the *UCL* using the simulation for the robust charts, while the formula 2.4 to compute the *UCL* for the conventional T^2 chart is used, because the vector X is required for calculating the estimators. We set the value of all *UCLs* to be 10.81512 for α=0.05. This case has false alarm probability with 25 observations. The final results show that in the case of conventional chart, the production process is not in control at two observations namely the second and twentieth observations, whereas the process is out of control only on the 22^{th} observation in case of robust charts.

7. CONCLUSION

This work proposed an effective Hotelling's T^2 chart using the Winsorized covariance matrix and Winsorized *MOM* as the scale covariance matrix and the location vector, respectively. In the *MOM*, the default-trimming criterion MAD_n was exchanged with another maximum breakdown point scale estimators (Q_n). The modified charts' performance was compared with the performance of conventional chart, in terms of probabilities of detecting outliers and false alarms rates. Studies on the performance involve two cases namely the independent case (A) and the dependent variables case (B). The outcomes of the simulation indicated that the performance of the modified charts is under control for the false alarm probabilities under most of conditions of the study, and starts to fail control once the shifted mean vector μ and outliers proportion increases. Moreover, this robust chart is able to produce detection probability of around 1. The conventional T^2 chart hardly reaches strong performance in detecting the outliers only under the condition bivariate p=2 with μ=2 and ε=0.1. The highest rate in this case reaches the value of 0.806 based on the three conditions of the sample sizes 50, 100 and 150. In other conditions, the conventional charts failed to produce strong performance. On the other hand, increasing p value enhanced the detection probability outliers for the proposed chart. In this case, the maximum value of probability of detecting outliers for the conventional chart reaches 0.468. This means that, the new robust charts has strong performance as the number of characteristics is increasing.

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] R. L. Mason and J. C. Young, *Multivariate Statistical Process Control with Industrial Applications* American Statistical Association-Society for Industrial and Applied Mathematics(ASA-SIAM), 2002.

[2] F. B. Alt, *Multivariate quality control*, Encyclopedia of Statistical Sciences **6**, (1985) 110-122.

[3] D. C. Montgomery, *Introduction to Statistical Quality Control,* 2005.

[4] J. J. Alloway and M. Raghavachari, *Multivariate control charts based on trimmed means*, In ASQC Quality Congress TransactionsSan Francisco, 449-453.

[5] J. J. Alloway and M. Raghavachari, *An introduction to multivariate control charts*, In American Soc. Quality Control Ann. Quality Congress Trans, 781.

[6] M. O. Abu-Shawiesh and M. B. Abdullah, *A new robust bivariate control chart for location*, Communications in Statistics-Simulation and Computation **30,** No. 3 (2001) 513-529.

[7] J. Surtihadi, *Multivariate and robust control charts for location and dispersion*, 1994.

[8] J. L. Alfaro and J. F. Ortega, *A robust alternative to Hotelling's T^2 control chart using trimmed estimators*, Quality and Reliability Engineering International **24**, No. 5 (2008) 601-611.

[9] F. Haddad, Syed S. Yahaya and J. L. Alfaro, *Alternative Hotelling's T^2 Charts using Winsorized Modified OneStep Mestimator*, Quality and Reliability Engineering International **29**, No. 4 (2013) 583-593.

[10] F. Haddad, Syed S. Yahaya and J. L. Alfaro, *Robust Hotelling's Charts based on High Breakdown Points Estimators*, In Icoqsia, 2010, 2-4 November, Penang Malaysia.

[11] S. H. Steiner and A. M. Variyath, *A multivariate robust control chart for individual observations*, Journal of Quality Technology **41**, No. 3 (2009) 259.

[12] J. Alfaro and J. F. Ortega, *A comparison of robust alternatives to Hotelling's T^2 control chart*, Journal of Applied Statistics **36**, No. 12 (2009) 1385-1396.

[13] N. J. A. Vargas, *Robust estimation in multivariate control charts for individual observations*, Journal of Quality Technology, **35**, No. 4 (2003) 367-376.

[14] H. M. Hanif, R N M, Rasmani K A and G N A M. *Enhancing Hotelling's T^2 Control Performance with Decile and Trimmed Mean,* In Proceedings of the 2013 International Conference on Information, Operations Management and Statistics (ICIOMS2013), Kuala Lumpur, Malaysia, September 1-3.

[15] F. Haddad, J. Alfaro and M. K. Alsmadi, *Hotelling's T^2 Charts Using Winsorized Modified One Step M-Estimator For Individual Non Normal Data*, Journal of Theoretical & Applied Information Technology, **72**, No. 2 (2015).

[16] A. Shabbak and H. Midi, *An Improvement of the Hotelling Statistic in Monitoring Multivariate Quality Characteristics*, Mathematical Problems in Engineering (2012) 1-15.

[17] M. O. A. AbuShawiesh, F. George and B. M. G. Kibria, *A comparison of some robust bivariate control charts for individual observations*, International Journal for Quality Research **8**, No. 2 (2014) 183-196.

[18] R. R. Wilcox and H. Keselman, *Repeated measures oneway ANOVA based on a modified onestep Mestimator*, British Journal of Mathematical and Statistical Psychology **56**, No. 1 (2003) 15-25.
[19] C. Croux and P. J. Rousseeuw, *Time-efficient algorithms for two highly robust estimators of scale*, 1992.
[20] Y. Ma and M. G. Genton, *Highly robust estimation of dispersion matrices*, Journal of Multivariate Analysis **78**, No. 1 (2001) 11-36.
[21] 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 **23**, No. 5 (2007) 615-629.
[22] W. A. Jensen, J. B. Birch and W. H. Woodall, *High Breakdown Estimation Methods for Phase I Multivariate Control Charts Quality Reliability Engineering International*, 2006.
[23] N. Johnson, *A comparitive simulation study of robust estimators of standard error*. Brigham Young University 2007.
[24] J. L. Alfaro and J. F. Ortega, *A comparison of robust alternative to Hotelling T^2 control chart*, Applied Statistics 36, No. 12 (2009) 1385-1396.

APPENDIX

TABLE 4. The two variables X_1 and X_2 of Vargas data set with the values of T^2 statistics using the conventional and the Winsorized MOM estimator.

Product No	X_1	X_2	$T^2_{\overline{X}-S}$	$T^2_{\overline{X}_{mom-Q_n}}$
1	0.567	60.558	0.650	0.98
2	0.538	56.303	13.004	54.57
3	0.530	59.524	0.169	1.246
4	0.562	61.102	1.539	3.63
5	0.483	59.834	1.264	1.713
6	0.525	60.228	0.215	0.183
7	0.556	60.756	0.783	1.556
8	0.586	59.823	0.842	1.744
9	0.547	60.153	0.076	0.0385
10	0.531	60.640	0.607	1.1648
11	0.581	59.785	0.684	1.5915
12	0.585	59.675	0.876	2.2038
13	0.540	60.489	0.348	0.5381
14	0.458	61.067	4.361	8.2798
15	0.554	59.788	0.094	0.5959
16	0.469	58.640	3.239	9.0668
17	0.471	59.574	1.889	2.9619
18	0.457	59.718	2.674	3.6645
19	0.565	60.901	1.154	2.3952
20	0.664	60.180	5.941	8.3775
21	0.600	60.493	1.619	2.2008
22	0.586	58.370	3.446	13.977
23	0.567	60.216	0.350	0.4265
24	0.496	60.214	0.911	1.10567
25	0.485	59.500	1.265	2.41786