International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-10, August 2019

New Robust MEWMA Control Chart for Monitoring Contaminated Data

Faridzah Jamaluddin, Hazlina Haji Ali, Sharipah Soaad Syed Yahaya

Abstract: ***Multivariate Exponential Weighted Moving Average (MEWMA), E^2 control chart is a popular multivariate control chart for monitoring the stability of time series data (non-random pattern). However, in this paper, we have shown that the existing MEWMA, E^2 control chart is sensitive in contaminated data or in the presence of outliers. To address this problem, this paper proposed an alternative MEWMA E^2 control chart using robust mean vector and covariance matrix instead of the classical mean vector and covariance matrix respectively. The classical mean vector in MEWMA E^2 control chart is replaced by Winsorized Modified One-step M-estimator (WM) while the classical covariance matrix is replaced by the Winsorized covariance matrix. The proposed MEWMA E^2 control chart known as robust MEWMA control chart, denoted as RE^2 control chart. The control limit for the RE^2 control chart was calculated based on simulated data. The performance of RE^2 and existing MEWMA E^2 control charts are based on the false alarm rate. The result revealed that the RE^2 control chart is more effective in controlling false alarm rates as compared to the existing MEWMA, E^2 control chart. The zinc-lead flotation data show that the RE^2 performs better in application.***

Keywords: ***Control chart, Contaminated data, Multivariate Exponential Weighted Moving Average (MEWMA), Robust estimator, Winsorized One-step M-estimator (WM)***

I. INTRODUCTION

Control chart is a popular statistical process control, SPC, tool used by practitioners to monitor processes with the aim of detecting unfavorable condition in process parameters. Although it was first proposed for manufacturing industry, control charts have now been applied in a wider range of disciplines, including health-care [1]–[3], environment [4]–[6] and general services [6]–[9]. Exponential weighted moving average (EWMA) and multivariate EWMA (MEWMA) are examples of time-weighted moving control charts, that use the weighted average as the location estimator [10]. EWMA control chart is efficient in detecting small and moderate process mean shifts [10]–[12]. The control chart by [13] is quite robust to normality assumption [10], [14] especially under skewed distributions such as *t*, gamma, uniform, right triangular and bimodal distributions[14], [15]. However, studies on contaminated normal data observed that the chart produces high false alarm rate [15]. Meanwhile, in multivariate aspect, several studied on the performance of MEWMA control chart under non-normal distributions observed that the standard MEWMA control chart is fairly robust against multivariate non-normal data[16][17]. For example, [17] extended the work of [14] to multivariate process for individual observations through Monte Carlo simulations. Focusing on two non-normal distributions namely multivariate *t* distribution and multivariate gamma distribution, the finding indicated that the MEWMA control chart is robust under both non-normal distributions. However, the question is whether the existing MEWMA control chart can still perform well under contaminated data?

Contaminated normal distribution can be defined as a mixture distribution which implies that majority of data are good but there are infrequent outliers. Outlier is an observation that appears to deviate markedly from remaining data [18]–[20]. In such contaminated normal distribution, a robust control chart is preferable. A lot of work has been done in the literature to design robust control charts based on robust estimators. For example, the used of robust estimators such as trimmed mean, Minimum Covariance Determinant, Minimum Volume Ellipsoid, Minimum Vector Variance and Winsorized Modified One-step M-estimator in the existing Hotelling T^2 control chart [21]–[26]. All these studies concluded that the robust control charts outperform the existing Hotelling T^2 control chart when samples deviated from normality assumption or outliers present in the data.

Thus, in this study we are interested in determining the robustness of the existing MEWMA control chart under various contaminated data. Although it has been proven by [15] that existing EWMA is not robust against contaminated data, in multivariate setting, the existing MEWMA, it is has yet to be proven. Hence, it is necessary to empirically investigate the performance of the existing MEWMA control chart under various contaminated data, since this distribution always exist in real data. Along with the existing MEWMA control chart, we also proposed a robust MEWMA control chart with an integration of robust estimator known as Winsorized Modified One-step M-estimator, WM. The ability of WM in controlling the false alarm rate has been proven in the Hotelling T^2 control chart [24] but never been tested on the existing MEWMA control chart. The performance of the investigated MEWMA control charts were assessed in the case where there were no changes in the covariance structures. The performance evaluation measured in term of false alarm rates. The organization of the remaining of the article is as follows. Section II explains on the glossary of symbols and Section III discusses about the existing MEWMA control chart.

Manuscript published on 30 August 2019.
***Correspondence Author(s)**
Faridzah Jamaluddin, School of Quantitative Sciences, College of Arts and Sciences, Universiti Utara Malaysia, 06010 UUM Sintok, Malaysia.
Hazlina Haji Ali, School of Quantitative Sciences, College of Arts and Sciences, Universiti Utara Malaysia, 06010 UUM Sintok, Malaysia.
Sharipah Soaad Syed Yahaya, School of Quantitative Sciences, College of Arts and Sciences, Universiti Utara Malaysia, 06010 UUM Sintok, Malaysia.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/

In Section IV, we formally introduce a robust MEWMA control chart based on WM estimator. While, Section V demonstrates Monte Carlo method to compute the control limit of robust MEWMA control chart. Sections VI and VII presents simulation design and result of the analysis respectively. Next, Section VIII discusses the performance of investigated MEWMA control chart on real data application. Finally, conclusion is given in the last section.

II. GLOSSARY OF SYMBOLS

n_1: Number of sample sizes for Phase I control limit
n_2: Number of sample sizes for Phase II control limit
n_3: Number of sample sizes for Phase I control chart
n_4: Number of sample sizes for Phase I control chart
p_1: Number of dimensions for Phase I control limit
p_2: Number of dimensions for Phase II control limit
p_3: Number of dimensions for Phase I control chart
p_4: Number of dimensions for Phase I control chart
α: False alarm rate
ε: Percentage of outliers
r: Smoothing parameter
CL: Control Limit
$\boldsymbol{\mu}_1$: Process mean shift

III. EXISTING MEWMA CONTROL CHART

The development of MEWMA control chart is based on the MEWMA E^2 statistic [27]. Let $\boldsymbol{X}_i$ represent p number of dimensions at time i. Then, the E^2 statistic for $\boldsymbol{X}_i$ is defined as:

$$E_i^2 = \boldsymbol{Z}_i^t \boldsymbol{\Sigma}_{\boldsymbol{Z}_i}^{-1} \boldsymbol{Z}_i \quad (1)$$

where $\boldsymbol{Z}_i$ represent the MEWMA vectors and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_i}$ is the covariance matrix of $\boldsymbol{Z}_i$. The $\boldsymbol{Z}_i$ and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_i}$ are estimated as Equation (2) and Equation (3) respectively.

$$\boldsymbol{Z}_i = r\boldsymbol{X}_i + (1-r)\boldsymbol{Z}_{i-1} \quad (2)$$

and

$$\boldsymbol{\Sigma}_{\boldsymbol{Z}_i} = \frac{r}{(2-r)}\boldsymbol{\Sigma}_0 \quad (3)$$

where $\boldsymbol{Z}_{i-1}$ is a process mean vector, $\boldsymbol{\mu}_0$.

The MEWMA control chart gives an out-of-control signal when the E^2 statistic is above the control limit, CL. The control limit used to achieve the desired false alarm rate. The MEWMA control chart developed by [27] assumed known parameters; process mean vector, $\boldsymbol{\mu}_0$ and covariance matrix, $\boldsymbol{\Sigma}_0$, used in estimating the E^2 statistics. However, practically these parameters are usually unknown and are estimated by the sample mean vector, $\bar{\boldsymbol{X}}$, and sample covariance matrix, $\boldsymbol{S}$. When using these estimators, the Equation (1) can be redefined as follows:

$$E_i^2 = (\boldsymbol{Z}_i - \bar{\boldsymbol{X}})^t \boldsymbol{S}_{\boldsymbol{Z}_i}^{-1} (\boldsymbol{Z}_i - \bar{\boldsymbol{X}}), \quad (4)$$

IV. ROBUST MEWMA CONTROL CHART BASED ON WM ESTIMATOR

The classical estimators, the $\bar{\boldsymbol{X}}$ and $\boldsymbol{S}$ used in (4) are sensitive to outliers [28]. In alleviating the problem, in this study, the classical estimators are replaced with robust estimators of mean vector and covariance matrix based on Winsorized Modified One-step M-estimator, *WM*. The mean vector and covariance matrix of *WM* estimator given by [24] are estimated as follows:

$$WM_j = \frac{\Sigma_{i=1}^{m_j} w_{ij}}{m_j} \quad (5)$$

and

$$S_{WM}(w_i, w_j) = \frac{1}{m-1}\Sigma_{k=1}^{m}(w_{ki} - WM_i)(w_{kj} - WM_j) \quad (6)$$

where W_{ij} is Winsorized sample, which obtained from two steps process; First step is to identify the existence of outlier based on trimming criterion used for Modified One-step M-estimator and then the data are winsorized. Let $\boldsymbol{X}_{ij} = \{\boldsymbol{X}_{1j}, \ldots, \boldsymbol{X}_{nj}\}$, j = 1, …, p be a random sample of p dimensions. The Modified One-step M-estimator proposed by [29] is defined as

$$MOM_j = \Sigma_{i=i_1+1}^{n_j-i_2} X_{ij}/n_j - i_1 - i_2 \quad (7)$$

where

$\boldsymbol{X}_{ij}$ = i^{th} order statistic in j^{th} dimension

i_1 = Number of $\boldsymbol{X}_{ij}$ that satisfies the criterion $(\boldsymbol{X}_{ij} - \widehat{\boldsymbol{M}}_j < -K \times MAD_{n_j})$ (8)

i_2 = Number of $\boldsymbol{X}_{ij}$ that satisfies the criterion $(\boldsymbol{X}_{ij} - \widehat{\boldsymbol{M}}_j > K \times MAD_{n_j})$ (9)

n_j = Number of observations in each j^{th} dimension

$\widehat{\boldsymbol{M}}_j$ = med $\{\boldsymbol{X}_{1j}, \ldots, \boldsymbol{X}_{nj}\}$, j = 1, ..., p

MAD_{n_j} = 1.4826 × med $\{|\boldsymbol{X}_{ij} - \widehat{\boldsymbol{M}}_j|\}$

For a reasonably smaller standard error and better efficiency under normality, the constant K was fixed to 2.24 [30]. They observed improved efficiency (i.e., 0.88 and 0.9) when the K value is equal to 2.24 for n = 10 and 20 respectively. MAD_n is a scale estimator with bounded influence function and the best possible breakdown point [31]. The simplicity of the MAD_n formula and fast computation time are among other strengths of MAD_n. After removing the outliers from each sample using criterion (8) and (9), the data are then winsorized. For each random variable $\boldsymbol{X}_{ij} = \{\boldsymbol{X}_{1j}, \ldots, \boldsymbol{X}_{nj}\}$, j = 1, ..., p dimensions, the sample is winsorized as follows:

$$W_{ij} = \begin{cases} \boldsymbol{X}_{(i_1+1)j}, & if \quad \boldsymbol{X}_{ij} \le \boldsymbol{X}_{(i_1+1)j} \\ \boldsymbol{X}_{ij}, & if \quad \boldsymbol{X}_{(i_1+1)j} < \boldsymbol{X}_{ij} < \boldsymbol{X}_{(n-i_2)j} \\ \boldsymbol{X}_{(n_j-i_2)j}, & if \quad \boldsymbol{X}_{ij} \ge \boldsymbol{X}_{(n-i_2)j} \end{cases}$$

where i_1 and i_2 represent number of the smallest and largest outliers in the data respectively. Thus, the robust MEWMA, RE^2 statistic for $\boldsymbol{X}_i$ is estimated as follows:

$$RE_i^2 = (\boldsymbol{Z}_{WM_i} - WM)^t \boldsymbol{S}_{\boldsymbol{Z}_{WM_i}}^{-1} (\boldsymbol{Z}_{WM_i} - WM) \quad (10)$$

where $\boldsymbol{Z}_{WM_i}$ is the robust MEWMA vectors and $\boldsymbol{S}_{\boldsymbol{Z}_{WM_i}}$ is the covariance matrix of $\boldsymbol{Z}_{WM_i}$.

The estimators, $\boldsymbol{Z}_{WM_i}$ and $\boldsymbol{S}_{\boldsymbol{Z}_{WM_i}}$, are defined as Equation 11 and Equation 12 respectively.

$$\boldsymbol{Z}_{WM_i} = r\boldsymbol{X}_i + (1-r)\boldsymbol{Z}_{WM_{(i-1)}} \quad (11)$$

and

$$\boldsymbol{S}_{\boldsymbol{Z}_{WM_i}} = (r/(2r)) \times \boldsymbol{S}_{WM} \quad (12)$$

V. CONTROL LIMIT OF ROBUST MEWMA CONTROL CHART

Since the distribution of the robust MEWMA, RE^2 control chart, is unknown, then the control limit of the RE^2 was calculated using Monte Carlo simulation approach.

Retrieval Number: J95880881019/19©BEIESP
DOI: 10.35940/ijitee.J9588.0881019
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)
© Copyright: All rights reserved.

The Phase I control limit involves the simulation of 5000 data sets with n_1 = 50, 200 and 400 and p_1 = 2, 5, and 10 from standard multivariate normal distribution $MVN_{p1}(0, \boldsymbol{I}_{p1})$ when false alarm rate, α is equal to 0.05. The Winsorized mean vector and covariance matrix, $\boldsymbol{WM}$ and $\boldsymbol{S_{WM}}$ for each 5000 data set were then estimated. Later, in Phase II of control limit, we generate a group of observation with sample size of n_2 = 30 for 5000 data sets and the robust MEWMA vectors, $\boldsymbol{Z_{WMi}}$ for all n_2 observations in the Phase II control limit are calculated. Then, the robust MEWMA statistics, RE_i^2 are calculated for each of the 30th observation using the corresponding estimators obtained from Phase I control limit. The control limit for the RE^2 was estimated by taking the 95th percentile of the 5000 values of RE_i^2. Table 1 reported the control limit values at α = 0.05 for several number of n_1 and p_1 with a given r = 0.05.

Table. 1 The control limits of E^2 and RE^2 control charts.

p_1	n_1	E^2	RE^2
2	50	9.539	524.018
	200	6.634	1138.75
	400	6.113	2038.82
5	50	19.200	1978.14
	200	12.445	4359.98
	400	11.396	77797.0
10	50	36.468	5860.67
	200	21.032	11971.1
	400	19.198	21395.5

VI. SIMULATION DESIGN

The performances of E^2 and RE^2 control charts were then examined in terms of false alarm rate under numerous conditions to highlight the strength and weakness of both MEWMA control charts. Sample sizes n_3 = 50, 200 and 400 observations with p_3 = 2, 5 and 10 dimensions were simulated from the mixture normal distribution. The mixture distribution suggested in [26] is as follows:

$$(1-\varepsilon)N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) + \varepsilon N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) \qquad (13)$$

where $\boldsymbol{\mu_0}$ and $\boldsymbol{\Sigma_0}$ are the in-control parameters; while $\boldsymbol{\mu_1}$ and $\boldsymbol{\Sigma_1}$ are the out-of-control parameters. The covariance matrix, $\boldsymbol{\Sigma_0}$ and $\boldsymbol{\Sigma_1}$ in equation (13) represent the identity matrix of p_3 dimensions ($\boldsymbol{I}_{p3}$) since we assume contamination with shift in the mean only but no changes in covariance structure. Table 2 presented the three different values of $\boldsymbol{\varepsilon}$ represent small, moderate and large percentage of outlier with seven values of $\boldsymbol{\mu_1}$ are also considered in creating various contaminated conditions. The manipulation of the $\boldsymbol{\varepsilon}$ and $\boldsymbol{\mu_1}$ produced 21 different types of contaminated distributions.

The following procedures represent the simulation method used in Phase I and Phase II of in establishing the RE^2 control chart. The in-control parameters which are used together with control limit to develop the control chart are estimated in Phase I control chart. The simulation method of Phase I control chart is as follows:

1. Stimulate 1000 group sizes n_3 = 50, 200 and 400 observations with dimensions, p_3 = 2, 5 and 10 from the models defined in Equation (13).
2. Compute the robust location, $\boldsymbol{WM}$ and scale estimators, $\boldsymbol{S_{WM}}$ and for each data set.

While in Phase II of control chart, the false alarm rate based on the estimations in Phase I of control chart are determined through the following procedures:

1. Randomly stimulate 1000 data sets of n_4 = 30 from in-control distribution (formula 13).
2. Calculate the robust MEWMA vectors, $\boldsymbol{Z_{WM}}$ of all 30 observations for 1000 data sets;
3. Compute the RE_i^2 for each of the $\boldsymbol{Z_{WM30}}$ (step 2) using $\boldsymbol{WM}$ and $\boldsymbol{S_{WM}}$ obtained in Phase I of control chart.
4. Compared the RE_i^2 in Step 3 with the control limits obtained in simulation process of control limit discussed in Section V.
5. The estimated proportion of RE_i^2 values in step 4 that are higher than the control limit, CL in 1000 data sets represents false alarm rates.

Table. 2 The ε and μ_1 used in Phase I control chart.

Percentage of outliers, $\boldsymbol{\varepsilon}$	Small	10%
	Moderate	15%
	Large	20%
Process mean shift, $\boldsymbol{\mu_1}$	No shift	0
	Small	0.1, 0.25, 0.5, 1.0
	Moderate	1.5, 2.0, 3.0

VII. RESULT

The comparison results of false alarm rates for the existing MEWMA, E^2 and robust MEWMA, RE^2 control charts under numerous conditions are reported in Table 3-5. According to the Bradley's liberal criterion, a control chart is considered robust if its false alarm rate, α is between the robust interval of 0.5α to 1.5α [32]. Thus, when the α is equal to 0.05, the MEWMA control chart is considered robust if its α is between 0.025 to 0.075. Taking into the consideration of the Bradley's liberal criterion, the bolded values in Table 3 to Table 5 indicate that the false alarm rate are between the robust interval under specified conditions. For bivariate (p_3 = 2) case reported in Table 3, the overall results on false alarm rates show that the RE^2 control chart outperforms the E^2 control chart. The RE^2 control chart have the ability in controlling false alarm rates for almost all of the conditions investigated which is about 80% (53 out of 66) of the conditions as compared to E^2 control chart, which is only effective for 56% (37 out of 66) of the conditions.

The E^2 control chart is badly affected with moderate and high percentage of outliers, $\boldsymbol{\varepsilon}$ = 15% and 20% especially for moderate process mean shifts, which are verified by the rates of false alarm far above the significance value, α = 0.05. Thus, the performance of the RE^2 control chart is considered superior than the E^2 for bivariate case.

When the dimensions increased to multivariate data, p_3 = 5, the RE^2 control chart is successfully maintain it good performance as compared to the E^2 control chart (refer to Table 4). However, the performance of RE^2 control chart decrease in case of multivariate data as compared to bivariate data, where it capable in controlling false alarm rates for only 44 simulated condition as compared to 53 simulated condition (refer to Table 3). In addition to that, we also notice some improvements in the E^2 control chart for multivariate data, p_3 = 5 as compared to the bivariate data, p_3 = 2 especially for small percentage of outlier, $\boldsymbol{\varepsilon}$ = 10%.

© Copyright: All rights reserved.

Table 5 demonstrate the false alarm rates for the multivariate case of p_3 = 10. The RE^2 control chart continue to be the best performer in controlling false alarm rates, robust under 40 simulated conditions. While some significant improvement could be
detected in E^2 control chart for the case of small sample size (n_3 = 30) as compared with result obtained when number of dimensions, p_3 = 5 as shown in Table 4.

VIII. ILLUSTRATIVE EXAMPLE

To evaluate the performance of E^2 and RE^2 control charts on real data, zinc-lead flotation data set was considered in this study. The flotation data set consists of five random variables, feed rate, copper II sulfate ($CuSO_4$), air flow rate, pulp level and upstream *pH*. The dataset comprises 200 vectors. Figure 1a - Figure 1e demonstrate the time series plots of the floatation data. The first 170 vectors of flotation data were used as Phase I control chart or historical data while the rest 30 vectors were for Phase II future data. Both the normality tests, the Shapiro-Wilk and Kolmogorov-Smirnov conclude that all the five variables in the flotation dataset are not normally distributed when all of the *p*-values are less than $1x10^{-4}$. In addition to this, the result on Mahalanobis distance indicates that the presence of multivariate outliers is close to 5% for this flotation data. All of this information suggest that the multivariate normality assumption is not valid for the flotation dataset and thus we could expect that the RE^2 control chart would be more robust and powerful than the E^2.

Table. 3 False alarm rates for E^2 and RE^2 control charts when p_3 = 2

n_3	ε	μ_1	E^2	RE^2	n_3	ε	μ_1	E^2	RE^2	n_3	ε	μ_1	E^2	RE^2
50	0%	0.00	**0.045**	**0.052**	200	0%	0.00	**0.037**	**0.051**	400	0%	0.00	**0.040**	**0.051**
	10%	0.10	**0.056**	**0.052**		10%	0.10	**0.049**	**0.046**		10%	0.10	**0.062**	**0.055**
		0.25	**0.057**	**0.051**			0.25	**0.055**	**0.046**			0.25	**0.058**	**0.056**
		0.50	**0.058**	**0.052**			0.50	**0.056**	**0.047**			0.50	**0.062**	**0.056**
		1.00	**0.055**	**0.047**			1.00	**0.070**	**0.042**			1.00	**0.067**	**0.041**
		1.50	**0.059**	**0.042**			1.50	0.080	**0.036**			1.50	**0.073**	**0.036**
		2.00	**0.059**	**0.042**			2.00	0.086	**0.029**			2.00	0.080	**0.031**
		3.00	**0.055**	**0.030**			3.00	0.090	0.018			3.00	0.090	**0.025**
	15%	0.10	**0.048**	**0.035**		15%	0.10	**0.051**	**0.055**		15%	0.10	**0.053**	**0.051**
		0.25	**0.052**	**0.038**			0.25	**0.053**	**0.053**			0.25	**0.057**	**0.050**
		0.50	**0.060**	**0.042**			0.50	**0.060**	**0.048**			0.50	**0.060**	**0.049**
		1.00	0.078	0.024			1.00	0.088	**0.041**			1.00	0.089	**0.041**
		1.50	0.093	**0.026**			1.50	0.115	**0.033**			1.50	0.126	**0.035**
		2.00	0.109	0.019			2.00	0.138	**0.025**			2.00	0.155	0.023
		3.00	0.113	0.009			3.00	0.172	0.021			3.00	0.198	0.012
	20%	0.10	**0.053**	**0.040**		20%	0.10	**0.055**	**0.061**		20%	0.10	**0.056**	**0.054**
		0.25	**0.055**	**0.040**			0.25	**0.059**	**0.062**			0.25	**0.061**	**0.050**
		0.50	**0.072**	**0.048**			0.50	**0.070**	**0.054**			0.50	0.080	**0.043**
		1.00	0.102	**0.037**			1.00	0.121	**0.040**			1.00	0.135	**0.041**
		1.50	0.141	**0.026**			1.50	0.184	**0.029**			1.50	0.217	**0.025**
		2.00	0.165	0.020			2.00	0.248	0.018			2.00	0.273	0.020
		3.00	0.186	0.010			3.00	0.336	0.014			3.00	0.391	0.009

© Copyright: All rights reserved.

Table. 4 False alarm rates for E^2 and RE^2 control charts when p_3 = 5

n_3	ε	μ_1	E^2	RE^2	n_3	ε	μ_1	E^2	RE^2	n_3	ε	μ_1	E^2	RE^2
	0	0.00	**0.044**	**0.046**		0%	0.00	**0.045**	**0.056**		0%	0.00	**0.043**	**0.054**
	10%	0.10	**0.050**	**0.041**		10%	0.10	**0.049**	**0.048**		10%	0.10	**0.051**	**0.041**
		0.25	**0.051**	**0.041**			0.25	**0.045**	**0.058**			0.25	**0.049**	**0.043**
		0.50	**0.048**	**0.038**			0.50	**0.055**	**0.056**			0.50	**0.050**	**0.038**
		1.00	**0.055**	**0.048**			1.00	**0.064**	**0.035**			1.00	**0.054**	**0.043**
		1.50	**0.053**	**0.026**			1.50	**0.073**	**0.029**			1.50	**0.067**	**0.032**
		2.00	**0.054**	0.022			2.00	**0.074**	**0.025**			2.00	**0.070**	**0.027**
		3.00	**0.056**	0.019			3.00	0.076	0.015			3.00	**0.075**	0.016
	15%	0.10	**0.055**	**0.055**		15%	0.10	**0.049**	**0.048**		15%	0.10	**0.050**	**0.053**
		0.25	**0.059**	**0.044**			0.25	**0.049**	**0.044**			0.25	**0.055**	**0.050**
		0.50	**0.059**	**0.044**			0.50	**0.059**	**0.047**			0.50	**0.064**	**0.044**
50		1.00	**0.073**	**0.035**	200		1.00	0.078	**0.039**	400		1.00	0.083	**0.043**
		1.50	0.080	**0.025**			1.50	0.096	0.022			1.50	0.101	0.023
		2.00	0.081	0.019			2.00	0.105	0.014			2.00	0.111	0.012
		3.00	0.079	0.009			3.00	0.109	0.008			3.00	0.125	0.012
	20%	0.10	**0.047**	**0.052**		20%	0.10	**0.045**	**0.055**		20%	0.10	**0.048**	**0.045**
		0.25	**0.055**	**0.056**			0.25	**0.047**	**0.048**			0.25	**0.057**	**0.045**
		0.50	**0.068**	**0.043**			0.50	**0.066**	**0.044**			0.50	0.080	**0.043**
		1.00	0.100	**0.036**			1.00	0.129	0.290			1.00	0.015	**0.027**
		1.50	0.116	0.018			1.50	0.176	0.016			1.50	0.191	0.018
		2.00	0.123	0.015			2.00	0.197	0.011			2.00	0.208	0.005
		3.00	0.118	0.007			3.00	0.221	0.004			3.00	0.247	0.002

Retrieval Number: J95880881019/19©BEIESP
DOI: 10.35940/ijitee.J9588.0881019
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)
© Copyright: All rights reserved.

Table. 5 False alarm rates for E^2 and RE^2 control charts when $p_3 = 10$

The mean vector and covariance matrix estimator for constructing the two statistics, E^2 and RE^2 are summarized in Table 6. While, the last column of Table 6 reported the *UCL* values for E^2 and RE^2 control chart for $\alpha = 0.05$ with 170 observations.

Meanwhile, Figure 2 shows the E^2 and RE^2 control charts along with their *UCL* values. From the plot, it can be seen that the RE^2 control chart gives an early signal as compared to the E^2 control chart. The RE^2 control chart exceeds its control limit at the 25th observations, while the E^2 control chart does not give any signal until the 28th observation.

Table. 6 Estimates of mean vector, covariance matrix and *UCL*

Control Chart	Mean Vector	Covariance Matrix	*UCL*
E^2	$\begin{bmatrix} 324.36 & 10.79 & 6.82 & 26.01 & 2.80 \end{bmatrix}$	$\begin{bmatrix} 222.636 & -0.050 & 0.510 & 7.973 & 0.011 \\ -0.050 & 0.005 & 0.005 & 0.178 & 0.0002 \\ 0.510 & 0.005 & 0.155 & 0.033 & 0.0008 \\ 7.973 & 0.178 & 0.033 & 20.138 & 0.006 \\ -0.001 & 0.0002 & 0.0008 & 0.006 & 4.27E-05 \end{bmatrix}$	12.576
RE^2	$\begin{bmatrix} 324.361 & 10.788 & 6.755 & 25.888 & 2.801 \end{bmatrix}$	$\begin{bmatrix} 216.083 & 1.039 & 3.456 & 61.585 & 0.080 \\ 1.039 & 0.005 & 0.017 & 0.292 & 0.0004 \\ 3.456 & 0.017 & 0.056 & 0.989 & 0.001 \\ 61.585 & 0.292 & 0.989 & 17.920 & 0.022 \\ 0.080 & 0.0004 & 0.001 & 0.022 & 3.10E-05 \end{bmatrix}$	3835.023

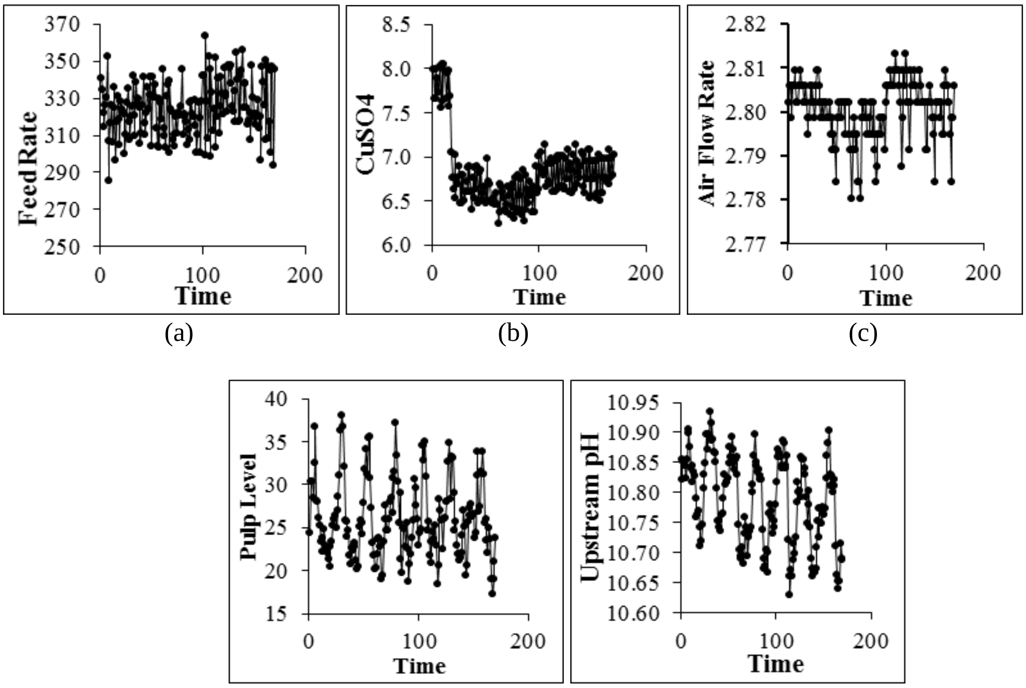

Fig. 1 The time series plot of the flotation process

Retrieval Number: J95880881019/19©BEIESP
DOI: 10.35940/ijitee.J9588.0881019
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)
© Copyright: All rights reserved.

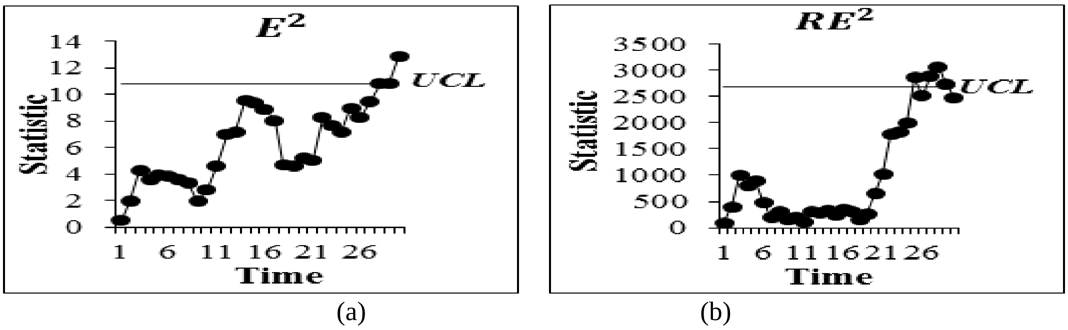

Fig. 2 The E^2 and RE^2 control charts for monitoring the flotation process

IX. CONCLUSION

In this article, we propose a robust MEWMA control chart using Winsorized One-step M-estimator, WM. Instead of using the classical mean vector and covariance matrix, as in the construction of the existing MEWMA control chart, E^2, we propose to use the robust estimators of the mean vector and covariance matrix using WM. The proposed MEWMA control chart, RE^2 is more robust than E^2 control chart. The E^2 control chart is less robust especially for moderate and high percentage of outliers with moderate process mean shift, where its false alarm rates inflated above the 0.075 robust limit regardless of the sample sizes. This is in line with past study on the EWMA control chart, the univariate version of MEWMA, E^2 control chart. The study by [15] conclude that the EWMA control chart is not robust against contaminated data, producing high false alarm rates.

Although the performance of the RE^2 control chart is better than E^2 control chart, its performance decreases as number of dimensions increases from bivariate data to multivariate data. While this study uses a single value of smoothing parameter, r = 0.05 as suggested by [33] it is recommended that future studies to consider other values of r in between 0.05 to 0.2 as well to achieve the robustness against contaminated data[10]. Apart from that, the future research might also consider other robust estimators such as minimum vector variance, minimum covariance determinant or modified one-step M-estimator, which is believed to be more capable in controlling the false alarm rates especially for high percentage of outliers regardless of number of dimensions investigated.

ACKNOWLEDGMENT

The authors would lie to acknowledge the work that led to this paper, which was fully funded by Fundamental Research Grant Scheme (S/O 14216) of Ministry of Higher Education, Malaysia.

REFERENCES

1. J. C. Benneyan, R. C. Lloyd, and P. E. Plsek, "Statistical process control as a tool for research and healthcare improvement," *Qual. Saf. Heal. Care*, vol. 12, no. 6, pp. 458–464, 2003.
2. F. Correia, R. Nêveda, and P. Oliveira, "Chronic respiratory patient control by multivariate charts," *Int. J. Health Care Qual. Assur.*, vol. 24, no. 8, pp. 621–643, 2011.
3. M. Waterhouse, I. Smith, H. Assareh, and K. Mengersen, "Implementation of multivariate control charts in a clinical setting," *Int. J. Qual. Heal. Care*, vol. 22, no. 5, pp. 408–414, 2010.
4. L. Burgas, J. Melendez, J. Colomer, J. Massana, and C. Pous, "Multivariate statistical monitoring of buildings. Case study: Energy monitoring of a social housing building," *Energy Build.*, vol. 103, pp. 338–351, 2015.
5. J. P. George, Z. Chen, and P. Shaw, "Fault Detection of Drinking Water Treatment Process Using PCA and Hotelling's T2 Chart," *Int. J. Comput. Electr. Autom. Control Inf. Eng.*, vol. 3, no. 2, pp. 430–435, 2009.
6. M. Shaban, "Drainage water reuse: State of control and process capability evaluation," *Water. Air. Soil Pollut.*, vol. 225, no. 11, 2014.
7. A. Faraz, C. Heuchenne, E. Saniga, and E. Foster, "Monitoring delivery chains using multivariate control charts," *Eur. J. Oper. Res.*, vol. 228, no. 1, pp. 282–289, 2013.
8. I. G. Guardiola, T. Leon, and F. Mallor, "A functional approach to monitor and recognize patterns of daily traffic profiles," *Transp. Res. Part B Methodol.*, vol. 65, pp. 119–136, 2014.
9. Y. Samimi and A. Aghaie, "Monitoring usage behavior in subscription-based services using control charts for multivariate attribute characteristics," in *2008 IEEE International Conference on Industrial Engineering and Engineering Management, IEEM 2008*, 2008, pp. 1469–1474.
10. D. C. Montgomery, *Introduction to Statistical Quality Control*, 6th ed. New York: John Wiley & Sons, 2009.
11. G. Capizzi and G. Masarotto, "Evaluation of the run-length distribution for a combined Shewhart-EWMA control chart," *Stat. Comput.*, vol. 20, no. 1, pp. 23–33, 2010.
12. N. Khan, M. Aslam, and C. H. Jun, "Design of a Control Chart Using a Modified EWMA Statistic," *Qual. Reliab. Eng. Int.*, vol. 33, no. 5, pp. 1095–1104, 2017.
13. S. W. Roberts, "Control Chart Tests Based on Geometric Moving Averages," *Technometrics*, vol. 1, no. 3, pp. 239–250, 1959.
14. C. M. Borror, D. C. Montgomery, and G. C. Runger, "Robustness of the EWMA Control Chart to Non-Normality," *J. Qual. Technol.*, vol. 31, no. 3, pp. 309–316, 1999.
15. S. W. Human, P. Kritzinger, and S. Chakraborti, "Robustness of the EWMA control chart for individual observations," *J. Appl. Stat.*, vol. 38, no. 10, pp. 2071–2087, 2011.
16. Z. G. Stoumbos and J. H. Sullivan, "Robustness to Non-normality of the Multivariate EWMA Control Chart," *J. Qual. Technol.*, vol. 34, no. 3, pp. 260–276, 2002.
17. M. C. Testik, G. C. Runger, and C. M. Borror, "Robustness properties of multivariate EWMA control charts," *Qual. Reliab. Eng. Int.*, vol. 19, no. 1, pp. 31–38, 2003.
18. P. . Rousseeuw and B. . Van Zomeren, "Unmasking Multivariate Outliers and Leverage Points," *J. Am. Stat. Assoc.*, vol. 85, no. 411, pp. 633–639, 1990.
19. V. S. Aelst, E. Vandervieren, and G. Willems, "A Stahel-Donoho estimator based on huberized outlyingness," *Comput. Stat. Data Anal.*, vol. 56, no. 3, pp. 531–542, 2012.
20. P. J. Rousseeuw and M. Hubert, "Robust statistics for outlier detection," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 1, no. 1, pp. 73–79, 2011.
21. J. L. Alfaro and J. F. Ortega, "A robust alternative to Hotelling's T2 control chart using trimmed estimators," *Qual. Reliab. Eng. Int.*, vol. 24, no. 5, pp. 601–611, 2008.
22. J. L. Alfaro and J. F. Ortega, "A comparison of robust alternatives to Hotelling's T2 control chart," *J. Appl. Stat.*, vol. 36, no. 12, pp. 1385–1396, 2009.
23. H. Ali, S. S. Syed Yahaya, and Z. Omar, "Robust hotelling T2 control chart with consistent minimum vector variance," *Math. Probl. Eng.*, vol. 2013, 2013.

© Copyright: All rights reserved.

24. F. S. Haddad, S. S. Syed Yahaya, and J. L. Alfaro, "Alternative Hotelling's T2 charts using winsorized modified one-step M-estimator," *Qual. Reliab. Eng. Int.*, vol. 29, no. 4, pp. 583–593, 2013.
25. W. A. Jensen, J. B. Birch, and W. H. Woodall, "High breakdown estimation methods for phase I multivariate control charts," *Qual. Reliab. Eng. Int.*, vol. 23, no. 5, pp. 615–629, 2007.
26. S. S. Syed Yahaya, H. Ali, and Z. Omar, "An alternative hotelling T2 control chart based on minimum vector variance (MVV)," *Mod. Appl. Sci.*, vol. 5, no. 4, pp. 132–151, 2011.
27. C. A. Lowry, W. H. Woodall, C. W. Champ, and S. E. Rigdon, "A multivariate exponentially weighted moving average control chart," *Technometrics*, vol. 34, no. 1, pp. 46–53, 1992.
28. H. Ali and S. S. Syed Yahaya, "On robust mahalanobis distance issued from minimum vector variance," *Far East J. Math. Sci.*, vol. 74, no. 2, 2013.
29. R. R. Wilcox and H. J. Keselman, "Repeated measures one-way ANOVA based on a modified one-step M-estimator," *Br. J. Math. Stat. Psychol.*, vol. 56, no. 1, pp. 15–25, 2003.
30. R. Wilcox, "Multiple comparisons based on a modified one-step Multiple comparisons based on a modified," *J. Appl. Stat.*, vol. 30, no. 10, pp. 1231–1241, 2003.
31. P. J. Rousseeuw and C. Croux, "Alternative to the Median Absolute Deviation," *Am. Stat. Assoc.*, vol. 88, no. 424, pp. 1273–1283, 1993.
32. J. V. Bradley, "Robustness?," *Br. J. Math. Stat. Psychol.*, vol. 31, no. 2, pp. 144–152, 1978.
33. J. H. Sullivan and W. H. Woodall, "Adapting control charts for the preliminary analysis of multivariate observations," *Commun. Stat. - Simul. Comput.*, vol. 27, no. 4, pp. 953–979, 1998.

AUTHORS PROFILE

Faridzah Jamaluddin received her Bachelor (Hons) in Industrial Statistics and Master of Science (Statistics) from Universiti Utara Malaysia since 2011 and 2015 respectively. She was currently a PhD student in Universiti Utara Malaysia. Her research interests are in Quality Control, Multivariate Control Charts, Robust Estimators.

Hazlina Ali is a senior lecturer at School of Quantitative Sciences, Universiti Utara Malaysia, UUM. She obtained her PhD in Statistics from UUM in 2013. Her research interests are developed new multivariate statistics methods based on robust estimators. The new robust methods will be applied in various industrial data.

A professor of Statistics at School of Quantitative Sciences, Universiti Utara Malaysia, Sharipah Soaad Syed Yahaya has been in academic field for almost 33 years. She holds a Bachelor and a Master degree in Mathematics from Indiana State University and Ph.D from Universiti Sains Malaysia specializing in Robust Statistics. Her works encompass the robustification of various statistical procedures such as statistical process control, group classification as well as groups comparison. Currently she is working on the robustification of mixed control charts for the detection of small shifts.

Retrieval Number: J95880881019/19©BEIESP
DOI: 10.35940/ijitee.J9588.0881019
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)
© Copyright: All rights reserved.