Journal of Physics: Conference Series

PAPER • **OPEN ACCESS**

Performance Analysis of Robust Locations Estimators

To cite this article: F. Z. Okwonu and Md. Y. Zahayu 2020 *J. Phys.: Conf. Ser.* **1529** 042095

Performance Analysis of Robust Locations Estimators

F. Z. Okwonu[1] and Md. Y. Zahayu[1]

[1]School of Quantitative Sciences, College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia

o.friday.zinzendoff@uum.edu.my

Abstract: This paper describes novel computational procedures that reduces the influence of outliers by transformational process. It transforms the influential observations to moderate data points. The procedures utilize all the information contained in the data set. These techniques are compared with the existing robust procedures, such as the weighted mean based on Mahalanobis distance, minimum covariance determinant (MCD), trimmed and winsorized mean. The performance analysis showed that the proposed techniques performed comparably with the well-established robust procedures based on the real data set.

Keywords: Mean; Trimmed mean; Winsorized mean; MCD; Mahalanobis distance

1. Introduction

1.1 The conventional mean

The conventional mean $\bar{x}$ is a well-known measure of central tendency and is often applied in different statistical methods [1,2]. Though, for normally distributed data set, the classical method for computing the mean yield better results than the robust methods. It is observed that, when the data set contains many large values or many smaller values, the classical mean overshot its true value. When the data set contain many smaller values, which are close to zero (inliers), the mean is underestimated. However, extreme smaller values or extreme larger values can be regarded as outlier because of the deviation from the clustered majority value that is closer to the mean. The break down point of $\bar{x}$ is $\frac{1}{k}$. This implies that as the sample size increases, the breakdown points tend to zero.

The uniqueness of the conventional mean irrespective of its non-robustness to contaminated data set is that it utilizes all the information contained in the data set. The usual robust procedures often expunge the influential data points thereby resulting in information loss. These robust techniques derive strength from contaminated data points. The aforementioned robust procedures do not rely on the original data set to compute the celebrated robust locations.

This paper briefly discusses few of the robust techniques for computing mean, such as trimming, winsorizing, weighted approach based on the Mahalanobis distance and the minimum covariance determinant (MCD). The techniques proposed in this paper utilize all the data point and yet preserve originality and content of the data set. These techniques reduce the influence of the larger outliers and enhances the value of the smaller outliers with reasonable percentage. The methods are very sensitive to influential observations.

The paper is arranged as follows: Section 2 briefly reviews some of the robust location estimators and their properties. Section 3 describes the different techniques and section 4 contains performance analysis of robust location estimators. The conclusions follow in section 5.

1.2 Robust location estimators

Overtime, the winsorized, trimmed means and other techniques that enable the reduction of the influence of outliers have been discussed extensively [3-5]. To enhance robustness in mean computation, several techniques have been established to transform and to trim the data set that containing influential data points [6]. The concept of robustness was coined when researchers assumed that the data set is contaminated, and the likelihood of the assumption of normality and homoscedasticity maybe violated. Though, robust techniques can handle any form of assumption violation but the price to be paid is information loss due to deleting outliers and relying heavily on inliers for mean computation.

Based on the instability in the presence of outliers, different robust techniques have been advanced to stabilize the mean through various means and given names. In some cases, the influential observations are deleted thereby losing the information contains in the data point. Other techniques also involve trimming or weighting the data set before the mean is computed for the purpose of stabilization as compared with the median. The robustification of the mean and covariance are advanced for the purpose of applying it to robustify well known classical methods that do not perform well in the presence of influential observations.

Other well-known high breakdown point and affine equivariant robust procedure such as the minimum covariance determinant (MCD) estimator which can be computed based on the Fast-minimum covariance determinant algorithm developed by Rousseeuw and Van Driessen (1999) is considered [7]. In its explicit form, the minimum covariance determinant selects h observations from the sample size k to compute the covariance matrix with the minimal determinant [8]. The MCD performs very well when the data set contains influential observations.

1.3 Properties of the robust estimators

The term affine equivalent and high breakdown estimator was coined in 1984 [9]. The concept of affine equivariant is vital since the transformation process of the data matrix does not affect the data positions. The affine equivariance nature of the sample mean is applicable to the sample variance [9-10].

1.3.1. Breakdown point. It has been observed that different techniques for estimating means and covariance often breakdown when the data set contain $\frac{k}{(p+1)}$ influential observations, where k denotes the number of observations and p is the number of variables of interest [8]. The MCD has high breakdown points [7], similarly to the trimmed mean and the winsorized mean that have $\gamma\%$ of breakdown.

2.0 Methods

2.1 The Conventional Mean

The conventional mean is known to every statistician as the most useful measure of central tendency, even with instability issues in the presence of outliers. In addition, conventional mean is highly influenced by outliers with breakdown point of $\frac{1}{k}$. This implies that as the sample size increases, the breakdown points tend to zero, that is $lim_{k\to\infty}\frac{1}{k} = 0$. The mean can be computed for univariate and multivariate data set depending on the objective of the study.

Let $X = [x_1, x_2, .., x_p]$ be $k \times p$ data matrix, then the mean is obtained as

$$\bar{x}_p = \frac{\sum_{i=1}^{k_p} x_i}{k_p}, k > p. \quad (1)$$

This formulation is based on p variables. The variance, standard deviation, standard error or coefficient of variation of the data set can also be computed. The formula for the ~~is~~ variance is:

$$S = \frac{\sum_{i=1}^{k_p}(x_i - \bar{x}_p)(x_i - \bar{x}_p)}{k_p - 1} \quad (2)$$

2.2 Weighted Mean Based on Mahalanobis Distance

This technique tends to robustify the mean and the covariance of a given data set X such that the sample size k is greater than the dimension p. Let apply equations (1-2) to define the Mahalanobis distance $M_{\alpha\beta}$ as follows;

$$M_{\alpha\beta} = \sqrt{(x_i - \bar{x}_p)S^{-1}(x_i - \bar{x}_p)}, \quad (3)$$

such that $M_{\alpha\beta} = \sqrt{\chi_p^2 \gamma}, \gamma = 0.975$, where $\chi_p^2 \gamma$ comes from the chi square distribution. From the above information, the weight of each data point is determined as follow:

$$w_l = \begin{cases} 1, if\ M_{\alpha\beta} \leq \sqrt{\chi_p^2 \gamma}, \gamma = 0.975 \\ 0, if M_{\alpha\beta} \geq \sqrt{\chi_p^2 \gamma}, \gamma = 0.975 \end{cases} \quad (4)$$

Based on the weights of the data points, the weighted mean and the weighted covariance are computed, that is;

$$\bar{x}_{w_i} = \frac{\sum_{i=1}^{k_p} w_l x_l}{\sum_{l=1}^{k_p} w_l}, \quad (5)$$

$$S_{w_l} = \frac{\sum_{l=1}^{n_p}(w_l x_i - \bar{w}_l)(w_l x_i - \bar{w}_l)^T}{\sum_{l=1}^{n_p} w_l - 1}. \quad (6)$$

This approach deletes the outliers and set the data point to zero.

2.3 The Trimmed Mean Technique (TMT)

The trimmed mean as proposed by Tukey is well established [11-13]. It is insensitive to influential observations by expunging certain percentage of the largest and the smallest data points in the data set [6,11,13-15]. The TMT then compute the mean based on the remaining data set. Trimming can be performed on the data set by applying the symmetric or the asymmetric procedures. Trimming based on equal percentage on both tails is classified, as symmetric trimming whereas trimming of unequal percentage of the data points on both tails is asymmetric trimming. The researcher needs to sort the data set before determining the percentages to be expunged from both tails. The percentage of data points to expunge when using the trimmed mean is user defined and depends on the identified influential observations. When the data set is normal the trimmed mean performance is comparable to the conventional mean [13].

Let X be as defined above. Then, the data points are ordered in ascending order such that, $x_1(n) \leq x_2(n) \leq x_3(n) \leq \cdots \leq x_k(n)$, where $x_1(n)$ and $x_k(n)$ are the smallest and largest data points. The trimmed mean is given by

$$\bar{X}_{TM} = \frac{\sum_{i=[k\gamma]+1}^{k-k\gamma} X_k}{k-2[k\gamma]} \tag{7}$$

Where $0 \leq \gamma \leq \frac{1}{2}$ and X_k denote the data points in the data set X, [16-18]. The computational time of the trimmed mean is $o(klogk)$ [19]. From equation (7), γ denotes the percentage of influential observations to be deleted from the data set. If $\gamma = 0$, the trimmed mean reduces to the arithmetic mean, that is,

$$\bar{X}_{TM} = \frac{\sum_{k=1}^{k} X_k}{k} = \bar{X} \quad . \tag{8}$$

2.4 Winsorized Mean

The winsorized mean computational procedure simply replaces the smallest and the largest data points with the nearest inliers at both end of the data set [2,20-21]. In general, the average of the winsorized data point is called the winsorized mean [22]. The formula is stated mathematically as follows;

$$\bar{X}_{wi} = \frac{(i+1)x_{(i+1)} + \sum_{j=i+2}^{n-i-1} X(j) + (i+1)x_{(n-i)}}{n} \quad , \tag{9}$$

where $\bar{X}_{wi}$ is an unbiased estimate of μ. The value of the winsorized mean lie between the value of the median and the conventional mean. For non-normal data set, $\bar{X}_{wi}$ is applied to compare groups [22]. The winsorized mean $\bar{X}_{wi}$ certainly have lower standard error than the conventional mean $\bar{x}_p$. Definitely, the advantage of winsorizing is that it retains the content of the data set without much information loss compared to trimming.

2.5 The Minimum Covariance Determinant (MCD)

The minimum covariance determinant (MCD) method does not utilize the entire data points contained in the data set rather utilizes the information provided by "finding" h observations ($h = 0.75k$) out of k observations to compute the mean which is the average of the h data point. Based on the MCD approach, covariance matrix has the smallest determinant [8].

This technique is well established and studied as a robust technique with high breakdown point and is affine equivariant [7-8]. The parameter of the MCD is based on $\frac{[k+p+1]}{2} \leq h \leq k$, where k is the sample size and p is the dimension of the data set. Then the following points are vital to study the MCD:

a. The robustness of the MCD is based on defining $h = \frac{k+p+1}{2}$.
b. The mean $\hat{\mu}_h$ is computed based on h observations, the determinant of the minimum covariance matrix is obtained based on the mean.
c. The covariance matrix $\widehat{\sum}_h$ is multiplied by a consistency factor $\beta_{hi(i=0)}$[7].
d. The mean and covariance matrix of the MCD is obtained when the h observations is greater than the data set dimension otherwise singularity problem will occur when computing the covariance.

e. To achieve consistency at the normal distribution, consistency constant is defined as $\beta_{hi} = \frac{\delta}{F_{\chi^2_{p+2}}(\chi^2_{p,\delta})}$, where $\delta = \lim_{k\to\infty} \frac{h(k)}{k}$, [23].

In order to increase efficiency, the MCD is reweighted based on the Mahalanobis distance.

$$d_j = \left(\left(x_j - \hat{\mu}_h\right)^T S^{-1}{}_h\left(x_j - \hat{\mu}_h\right)\right)^{\frac{1}{2}}$$

Applying the reweighting technique, the mean and covariance is obtained as

$$\hat{\mu}_{mcdh} = \frac{\sum_{j=1}^{k} w(d_j)x_j}{\sum_{j=1}^{k} w(d_j)} \tag{10}$$

$$\widehat{\sum}_{mcdh} = \beta_{hi}\frac{\sum_{j=1}^{k}\left(w(d_j)x_j - \hat{\mu}_{mcdh}\right)\left(w(d_j)x_j - \hat{\mu}_{mcdh}\right)}{k}$$

where w denotes the weight and defined as $w(d_j) \leq \chi^2{}_{p,0.975}$. Details of the MCD is contained in [7-10,23].

2.6 Modified Weighted Approach

The value of the information any data point provided is vital to data analyst. To prevent information loss due to assigning one to inliers and zero to outlier, the weighting procedure in equation (4) is modified such that all the data points in the data matrix can contribute meaningful. This procedure ensures that all the data point contribute to the computation of the mean and covariance. The modified weighting formula is given as

$$\overleftarrow{w_l} = \begin{cases} 1, \; inliers \\ 0 + \left(1 - \frac{\alpha p}{k_p + p}\right), outliers \end{cases} \tag{11}$$

Hence the mean and the covariance can be obtained with all data points position intact. The mean and the variance based on this description is stated as follows:

$$\bar{\bar{x}}_{w_i} = \frac{\sum_{i=1}^{k_p} \overleftarrow{w_l} x_l}{\sum_{l=1}^{n_p} \overleftarrow{w_l}}, \; \left(k_p = \sum_{l=1}^{n_p} \overleftarrow{w_l}\right) \tag{12}$$

$$S_{w_l} = \frac{\sum_{l=1}^{k_p}(\overleftarrow{w_l}x_l - \bar{\bar{w}}_{wl})(\overleftarrow{w_l}x_l - \bar{\bar{w}}_{wl})^T}{\sum_{l=1}^{k_p} \overleftarrow{w}_l - 1}. \tag{13}$$

The weight assigned to the outliers depends on the sample size and the dimension of the data set with the constant $\alpha = 1.5$. For instance, if the sample size of the data set is $k = 21$ and dimension $p = 3$, then the weight assigned to an influential observation is 0.8125. Suppose the data point 80 is considered outlier, the procedure in equations(4,7) will delete the data point and assign zero value but this procedure will simply reduce the influential observation to inlier such that the data point becomes inlier.

2.7 Reweighted Delta Bar (RDB)
Let X be $k \times p$ data matrix, where k denotes the number of observations and p denotes the number of variable of interest, then the data matrix can be expressed as $X = [x_1, x_2, \ldots, x_p], k > p$. The data set X is sorted automatically by the computational process leading to computing £=med (X) and $\beta = \sqrt{\delta/£}$, δ is user define, then

$$\Delta = (\beta X)^2 \tag{14}$$

be the transformed data matrix, then the mean can be computed as follows;

$$\bar{\Delta} = \frac{\sum_{i=1}^{n_i} \Delta}{k_i} \tag{15}$$

$$d_\Delta = \sqrt{(\Delta_i - \bar{\Delta})\nabla^{-1}(\Delta_i - \bar{\Delta})},$$

$$\text{where} \quad \nabla = \frac{\sum_{i=1}^{k_p}(\Delta_i - \bar{\Delta})(\Delta_i - \bar{\Delta})}{k_p - 1}.$$

We invoke equation (11) to reweight the data set. After applying equation (11), the reweighted mean can be computed as follows:

$$\bar{\Delta}_{RW} = \frac{\sum_{i=1}^{k_p} \bar{w}_R \Delta_i}{k_p}. \tag{16}$$

$$\nabla = \frac{\sum_{l=1}^{k_p}(\bar{w}_R\Delta_i - \bar{\Delta}_{RW})(\bar{w}_R\Delta_i - \bar{\Delta}_{RW})^T}{\sum_{l=1}^{k_p} \bar{w}_R\Delta_i - 1}. \tag{17}$$

The denominator of equation (16) is retained since the entire data set was applied to compute the mean, no data point was deleted, the value of the identified outliers was reduced. However, equation (14) allows the would-be outliers to be identified. Based on (11,14), this method is very sensitive to outliers, because (14) exposes the masking outliers and (11) transforms it to inlier.

3.0 Performance Analysis based on Real Data
We investigate the performance of the aforementioned-techniques on the premise that real data set often violate the assumption of normality and homoscedasticity [24-25]. To investigate the performance of the new techniques proposed, we compared the well-established robust location techniques using real data set from the stockloss applied in the minimum volume ellipsoid subroutine [26,28] and the red wine data [27]. The stockloss data set consist of $K = 21$ and $p = 3$. The performance of the different methods is reported in Table 1.

Table 1: comparison of mean based on real data set [26]

Methods	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$
$\bar{X}_p$ (eq.1)	60.428571	21.095238	86.285714
$\bar{X}_w$ (10%)	59.952381	20.952381	86.714286
$\bar{X}_{TM}$ (10%)	58.933333	20.733333	86.8
$\bar{X}_w$ (20%)	57.714286	20.809524	86.619048

$\bar{X}_{TM}$ (20%)	59.272727	20.636364	87.181818
MCD	59.50	20.83	87.33
$\bar{X}_{Wi}$(eq.5)	49.47619	20.285714	82.857143
$\bar{\bar{X}}_{wi}$(eq.12)	59.915179	21.057292	86.125
$\bar{\Delta}$(eq.15)	57.905172	20.453571	77.302463
$\bar{\Delta}_{RW}$(eq.16)	54.134968	20.318103	76.743842

In Table 2 below, we extracted the first fifty-three $(K = 53)$ data points with seven variables $(p = 7)$. The mean based on the above methods is contained in Table 2. Based on the principle behind the formulation of $\bar{X}_{TM}$ and $\bar{X}_{Wi}$, these techniques expunge the influential observations while the MCD uses a fraction based on the h observations [7]. While $\bar{\bar{X}}_{wi}$, $\bar{\Delta}$ and $\bar{\Delta}_{RW}$ reduces the influence of influential observations and compute the average with the entire data set, which is almost similar to winsorization approach.

Table 2. Red Wine data [27]

Methods	Fixed acidity	Volatile acidity	Citric acid	Residual sugar	Chlorides	Free sulfur dioxide	Total sulfur dioxide
$\bar{X}_p$ (eq.1)	7.50	0.55	0.19	2.61	0.12	15.02	52.25
$\bar{X}_w$ (10%)	7.539622	0.539339	0.179622	2.488679	0.094490	13.54717	50.35849
$\bar{X}_{TM}$ (10%)	7.536585	0.547926	0.169268	2.178048	0.087024	13.12195	47.97561
$\bar{X}_w$ (20%)	7.484905	0.544434	0.175283	2.075471	0.085830	12.69811	48.07547
$\bar{X}_{TM}$ (20%)	7.545161	0.554677	0.164838	2.058064	0.083935	12.83871	46
MCD	7.36	0.57	0.13	1.95	0.08	12.16	35.78
$\bar{X}_{Wi}$(eq.5)	7.411839	0.545634	0.183957	2.554056	0.113596	14.45094	51.87217
$\bar{\bar{X}}_{wi}$(eq.12)	7.477959	0.550748	0.187498	2.594174	0.116785	14.87688	52.15200
$\bar{\Delta}$(eq.15)	6.797569	0.5192655	0.3506816	4.3408868	0.3347271	22.942816	88.229292
$\bar{\Delta}_{RW}$(eq.16)	6.704052	0.501778	0.3233921	4.1424195	0.2989411	21.101315	82.294081

The results in Table 3 and Table 4 consist of the measurements of Bull and Cow horn length, horn width and base length. The sample size consists of $K = 100$ and $p = 3$, respectively.

Table 3. Bull horn measurement (cm)

Methods	Horn-length (cm)	Horn Width (cm)	Base length (cm)
$\bar{X}_p$ (eq.1)	19.178	7.4402	1.556
$\bar{X}_w$ (10%)	19.24	7.4422	1.561
$\bar{X}_{TM}$ (10%)	19.185	7.33275	1.545
$\bar{X}_w$ (20%)	19.19	7.316	1.543
$\bar{X}_{TM}$ (20%)	19.083333	7.2933333	1.5383333
MCD	19.96	7.25	1.46
$\bar{X}_{Wi}$(eq.5)	17.68	7.107	1.415
$\bar{\bar{X}}_{wi}$(eq.12)	19.158034	7.4365607	1.55446
$\bar{\Delta}$(eq.15)	17.606939	7.3142258	1.452375
$\bar{\Delta}_{RW}$(eq.16)	17.520783	7.2582564	1.4429795

Table 4. Cow horn measurement (cm)

Methods	Horn-length (cm)	Horn Width (cm)	Base length (cm)
$\bar{X}_p$ (eq.1)	27.32	11.24	2.38
$\bar{X}_w$ (10%)	27.354	11.05	2.4
$\bar{X}_{TM}$ (10%)	27.4675	11.075	2.3875
$\bar{X}_w$ (20%)	27.464	11.082	2.38
$\bar{X}_{TM}$ (20%)	27.573333	11.07	2.3666667
MCD	28.19	10.99	2.35
$\bar{X}_{Wi}$(eq.5)	27.212359	11.194107	2.3749126
$\bar{\bar{X}}_{wi}$(eq.12)	27.29159	11.231527	2.3817282
$\bar{\Delta}$(eq.15)	24.708938	10.530505	2.2814609
$\bar{\Delta}_{RW}$(eq.16)	24.607091	10.465276	2.2699007

Based on the different methods, the performance analysis for the robust methods and the proposed methods are comparable. Though, performance differs based on the data set as in the case of MCD in Table 1-2 and Table 3-4, respectively.

Table 5. Number of influential observations identified and transformed

Methods	Stockloss	Wine	Bull	Cow
$\bar{X}_{Wi}$(eq.5)	**5**	**18**	**20**	**26**
$\bar{\bar{X}}_{wi}$(eq.12)	5	18	20	26
$\bar{\Delta}_{RW}$(eq.16)	6	22	24	24

The method $\bar{X}_{Wi}$ equation (5) delete the outliers while the other two techniques in Table 5 transform the data set into inliers. The other techniques not listed in Table 5 are well known procedures on handling influential observations hence we are silent on discussing them. The performance analysis of the proposed techniques is benchmarked on the performance of the well-established techniques such as the winsorized, trimmed mean and the robust mean based on the minimum covariance determinant by comparing the computed mean from the conventional approach.

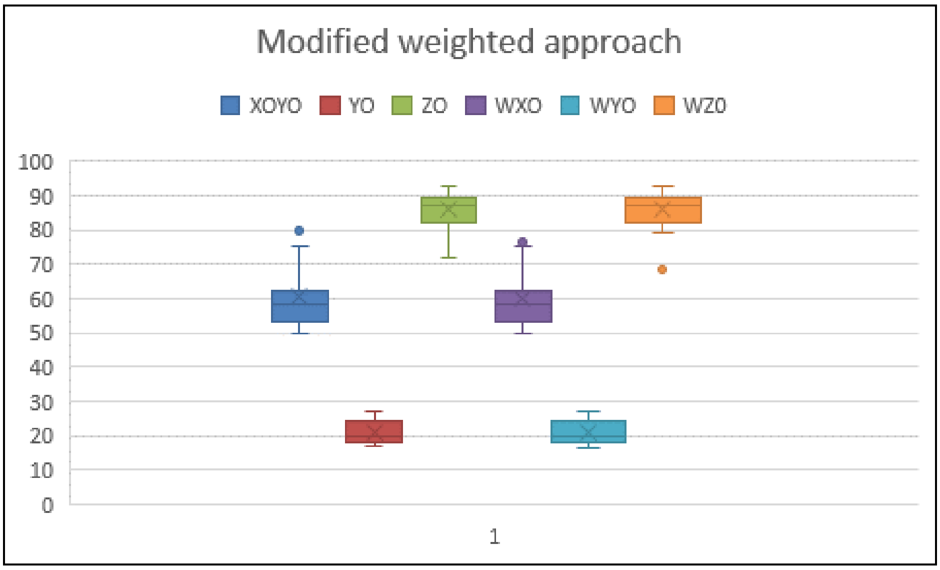

Figure 1. Modified weighted approach.

In Figure 1 above, XOYO, YO and ZO are the origin data set while WXO, WYO and WZO are the transformed data based on the method in equation (12). Based on the method, the outlier influence in XOYO was drastically reduced and WZO identified an outlier.

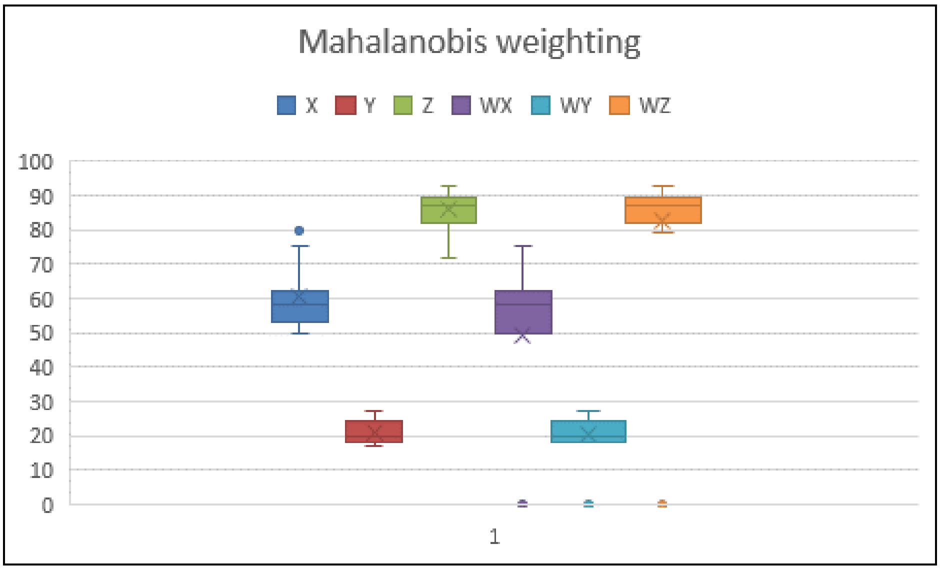

Figure 2. Mahalanobis distance weighting approach.

In Figure 2 X,Y and Z are the original data set while WX,WY and WZ are the transformed data based on equation (11). Three outliers were identified and deleted.

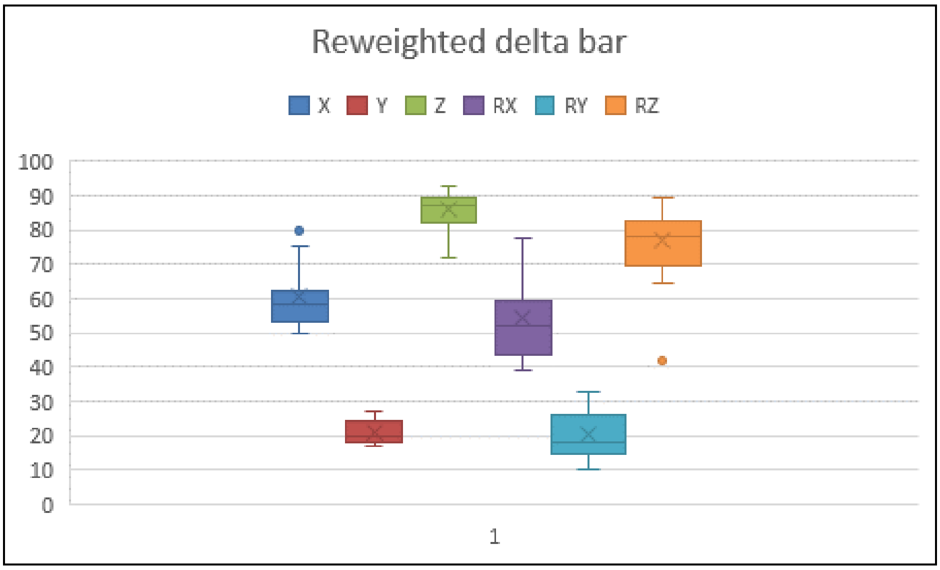

Figure 3. Reweighted delta bar.

From the figure 3 above, X,Y and Z are the original data set while RX,RY and RZ are based on equation(14). From above we observed that the outlier in X was transformed to inlier in WX and RZ identified an outlier in Z.

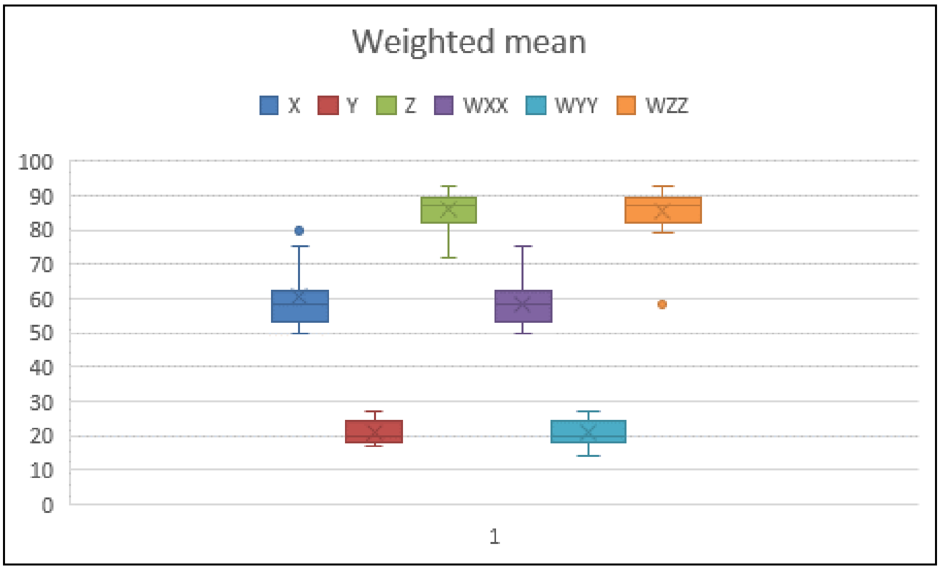

Figure 4. Weighted mean approach.

This procedure transforms the original data set X, Y and Z into WXX, WYY and WZZ. It transforms the outlier in x into inlier in WXX and the masking outlier in Z was identified in WZZ.

4.0 Conclusion

This study revealed that the proposed methods and the well-known robust procedures such as the minimum covariance determinant (MCD), the trimmed mean technique (TMT), winsorized mean and the weighted approach based on Mahalanobis distance performed comparably based on the real data set used. The level of outlier identification and weighting process is unique. The uniqueness of these techniques lies on the transformation of outliers by percentage reduction to inliers.

Reference

[1] Wang, T., Li, Y. and Cui, H. (2007). On weighted random trimmed means. Journal of System Science & Complexity, **20**, 47-65.
[2] Dodonov, Y. S. and Dodonova, Y. A. (2011). Robust measures of central tendency: weighting as a possible alternative to trimming in response time data analysis. Psikologicheskie Issledovaniya, **15**, 1-11.
[3] Huber, P. J. (1964). Robust estimation of a location parameter. Ann. Math. Statistic, **35**,73-101.
[4] Tukey, J.W. and Mclaughlin, D. H. (1963). Less vulnerable confidence and significance procedures for location based on a single sample: Trimming/Winsorization I, Sankhya. A, **3**, 331-352.
[5] Adam L. and Bejda, P. (2017). Robust estimators based on generalization of trimmed mean. Communications in statistics-simulation and computation, 1-13.
[6] Muhammad DI, N. F., Yahaya, S.S.S. and Abdullah, S. (2014). Comparing groups using robust H-statistic with adaptive trimmed mean. Sains Malaysia, **43**, 634-648.
[7] Hubert, M. and Debruyne, M. (2010). Minimum covariance determinant. Computational Statistics, **2**,36-43.
[8] Rousseeuw, P. J., and Van Driessen, K. (1999). A fast algorithm for the minimum covariance determinant estimator. Technometrics, **41**,212-223.
[9] Rousseeuw, P. J. (1984). Least median of squares regression. Journal of the American Statistical Association, **79**,871-880.
[10] Pison, G., Van Aelst, S. and Willems, G. (2002). Small sample corrections for LTS and MCD. Metrika, **55**,111-123.
[11] Dhar, S.S. and Chaudhuri, P. (2012). On the derivatives of the trimmed mean. Statistica Sinica, **22**,655-679.
[12] Tukey, J. W. (2003) Some elementary problems of importance to small practice. Human Biology, 205-214.
[13] Stigler, S.M. (1973). The asymptotic distribution of the trimmed mean. The Annals of Statistics, **1**, 472-477.
[14] Wilcox, R.R. (2005). Trimmed means. Encyclopedia of Statistics in Behavioral science, 4,2066-2067.
[15] Alkhazaleh, A. M.H. and Razali, A.M. (2010). New techniques to estimate the asymmetric mean. Journal of Probability and Statistics, 1-9.
[16] Midi, H., et al, (2018). Robust trimmed mean direction to estimate circular location parameter in the presence of outliers. Journal of Engineering and Applied Science, **13**,6612- 6615.
[17] Maronna, R.A. Martin, R. D. and Yohai, V. J. (2006). Robust statistics, theory and methods, (3rd ed.), John Wiley and Sons Ltd, Hobokon, New Jersey, USA.
[18] Oosterhoff, J. (1994). Trimmed mean or sample median? Statistics and probability letters, 20, 401-409.
[19] Beliakov, G. (2011). Fast computation of trimmed means. Journal of Statistical Software, **39**, 1-6.

[20] Wilcox, R.R. (2003). Applying contemporary statistical techniques, (1st ed.), CA academic press, San Diego.
[21] Courvoisier, D.S. and Renaud, O. (2010). Robust analysis of the central tendency, simple and multiple regression and ANOVA: A step by step tutorial. International journal of Psychological research, **3**, 78-87.
[22] Wilcox, R.R. (2011). Winsorized robust measures. Wiley StatsRef. Statistics Reference Online,1-2 https://doi.org/10.1002/9781118445112.stat06339.pub2.
[23] Croux, C. and Haesbroeck,G. (1999). Influence function and efficiency of the minimum covariance determinant scatter matrix estimator. Journal of Multivariate Analysis,**71**,161-190.
[24] Wilcox, R.R. (2010). Fundamental of modern statistical method, (2nd ed.), Springer, New York.
[25] Thompson, B. (2006). Foundation of behavioral statistics: An insight-based approach. New York. The Guilford Press.
[26] Brownlee, K. A. (1965). Statistical Theory and Methodology in Science and Engineering. John Wiley & Sons, New York.
[27] Https://archive.ics.uci.edu>datasets>wine
[28] https://documentation.sas.com