Thailand Statistician
April 2021; 19(2): 411-419
http://statassoc.or.th
Contributed paper

A Comparison of MM-estimation and Fuzzy Robust Regression for Multiple Regression Model with Outliers

Kanittha Yimnak*[a] and Kriangsak Piampholphan [b]
[a] Department of Mathematics and Computer Science, Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi, Pathum Thani, Thailand.
[b] School of General Education, Kasem Bundit University, Bangkok, Thailand.
*Corresponding author; e-mail: kanittha_y@rmutt.ac.th

Received: 18 September 2019
Revised: 24 May 2020
Accepted: 11 August 2020

Abstract

In this paper, MM-estimation and fuzzy robust regression, developed from M-estimation, are used for modelling and are compared the model performance by simulation the data sets containing outliers in X, outliers in Y and outliers in X and Y, respectively. The robust regression is considered by the estimated mean square error (EMSE) (or the mean square of the bias values) and the mean absolute error (MAE) values. These values indicate estimation precision. The fuzzy robust regression provides the models being more robust than the MM-estimation almost all types of outliers because it has a lower EMSE and MAE, especially, outliers in both X and Y. Nevertheless, the MM-estimation is more effective than the fuzzy robust regression for case of 30% of outliers in X when n = 20 and 40, respectively and case of 10% and 20% of outliers in Y.

Keywords: Membership function, simulation for outlier.

1. Introduction

Data set, unfulfilled assumptions, affects the performance of the regression model, especially, outlier problem. Many outliers infringe the presumptions of normally distributed residuals in the least square regression. Both outliers in dependent and independent variables have a prejudicial effect on the estimate (Alma 2011). Robust regression increases the robustness of regression model because it offers an alternative to least squares regression for the natural data set that does not correspond the assumptions. The properties of robust estimators are a high finite sample breakdown point (or the confinement percentage of adulteration in the data that any test statistics first turn into swamped) and high leverage points (Donoho and Huber 1983). Robust regression models are shunted into three categories: L, R and M estimations (Alma 2011). M-estimation method, presented by Huber (1964), is a well- known method and is an estimation of maximum likelihood type. The method provides the un-bias estimator and minimum variance (Huber 1964, Susanti et al. 2014). M-estimation was developed continuity to increase the efficiency of regression such as S-estimation and MM-estimation. S-estimation, proposed by Rousseeuw and Yohai (1984), offers a high breakdown value method. The advantage of this method is minimizing the pervading of the residuals. The MM-estimation, proposed by Yohai (1987), has together the properties of high efficacy when the errors have a normal distribution and high breakdown point (BP) = 0.5 (Yohai 1987). The algorithm of the MM-estimation similar to the M-estimation, but it uses the S-estimation to minimize the scale of residual from the M-estimation (Yohai 1987, Susanti et al. 2014). These methods are often compared the efficacy of model by many researchers. For example, Alma (2011) compared the efficiency between these estimations and least trimmed squares (LTS) method for data sets containing 10%, 25% and 40% of outliers in Y. M-estimation, S-estimation and MM-estimation provide models that are higher R^2 than least trimmed squares method. For real data applying, Susanti et al. (2014) created the model for maize production data with 6 independent variables using M-estimation, S-estimation and MM-estimation. These methods provide the model with a high R^2. However, the S-estimation is the highest R^2. In addition to the described methods, there is another science that is applied extensively with statistical science, namely fuzzy logic. Presently, fuzzy logic is applied statistical research when the data set is fuzzy or not according assumptions. There are many categories of applying fuzzy logic for creating regression model. For example, linear programming approach proposed by Tanaka et al. (1982) and fuzzy least square approach proposed by Diamond (1988) (Redden and Woddall 1996, Yang and Ko 1997, Yang and Lin 2002, Nasrabadi and Nasrabadi 2004). However, the mentioned fuzzy model does not correspond the data set containing outliers. Kula et al. (2012) presented fuzzy robust regression (FRR) method. They studied multiple regression models by transforming explanation and dependent variables to triangular fuzzy number and parameter estimation to crisp number. Then, they applied this model with the insurance data set. Their results showed that the fuzzy robust regression provides the model with low residual.

In this study, the MM-estimation and the fuzzy robust regression were applied to increase the robustness of the regression model. The membership function of the residuals by this method defines the weighted matrix. Thus, the weighted fuzzy least squares are constructed by the weighted matrix. This process is similar M-estimation. The performance of the MM-estimation and the fuzzy robust regression are compared by the simulation the data sets containing outliers in X, outliers in Y and outliers in both X and Y. The efficacy of the both methods is compared using mean absolute error (MAE) and the estimated mean square error (EMSE). Multiple regression model is described in Section 2 and MM-estimation is described in Section 3. Fuzzy robust regression model is explained in Section 4. Evaluations of the model are indicated in Section 5. Section 6 represents numerical examples. Finally, conclusions and discussions are revealed in Section 7.

2. Multiple Regression Model

The data set, consisting of n observations and p regressors $(y_i, x_{ij}); i = 1,2,...,n; j = 1,2,...,p,$ can be written the multiple regression model as follows (Wei et al. 2009):

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

or

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{bmatrix}_{n\times(p+1)} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}_{(p+1)\times 1} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix},$$

where $\mathbf{Y}$ is the vector of dependent variable, $\mathbf{X}$ is the matrix of independent variables, $\boldsymbol{\beta}$ is the vector of regression coefficients and $\boldsymbol{\varepsilon}$ is the vector of error. The common assumption is as (Mahaboob et al. 2018):

$$E(\varepsilon)=0 \text{ and } E(\varepsilon\varepsilon')=\sigma^2 I_n,$$

where I_n is a unit matrix of order n. The residual sum of square can be written as

$$\begin{aligned} e'e &= (Y-X\hat{\beta})'(Y-X\hat{\beta}), \\ &= Y'Y-\hat{\beta}'X'Y-Y'X\hat{\beta}+\hat{\beta}'X'X\hat{\beta},\ \left(\hat{\beta}'X'Y=Y'X\hat{\beta}\right), \\ &= Y'Y-2\hat{\beta}'X'Y+\hat{\beta}'X'X\hat{\beta}, \end{aligned}$$

where $\hat{\beta}$ is the least squares estimator of β. $\hat{\beta}$ can be solved by minimizing the residual sum of squares using the first order condition:

$$\frac{\partial}{\partial\hat{\beta}}(e'e)=0,$$

$$-2X'Y+2X'X\hat{\beta}=0,$$

$$X'X\hat{\beta}=XY,$$

$$\hat{\beta}=(X'X)^{-1}XY.$$

3. MM-estimation

The algorithm of MM-estimation is developed from M-estimation. This method used S-estimation to minimize the residual scale from M-estimation, so the estimation by this method have high breakdown value and more effective. The proportion of outliers, breakdown value, are addressed, these don't affect the model (Chen 2002). S-estimation, solves problem frailness of M-estimation in part of the default discretion in the distribution of a data set and not function of all data set because it usually applies the median like the weighted value. The S-estimation can be written as

$$\hat{\beta}_S=\min_\beta \hat{\sigma}_S(e_1,e_2,\dots,e_n),$$

where $\hat{\sigma}_s$ is determined minimum robust scale estimator and realizing $\min\sum_{i=1}^{n}\rho\left(\frac{y_i-\sum_{j=1}^{n}x_{ij}\beta}{\hat{\sigma}_S}\right)$ (Salibian-Barrera and Yohai 2006),

$$\hat{\sigma}_S=\sqrt{\frac{1}{nk}\sum_{i=1}^{n}w_ie_i^2}, \tag{1}$$

k is turning constant and $w_i=w_\sigma(u_i)=\frac{\rho(u_i)}{u_i^2}$. The initial estimate $(\hat{\sigma}_s^0)$ can be written as

$$\hat{\sigma}_s^0=\frac{\text{median}\left|e_i-\text{median}(e_i)\right|}{0.6745}.$$

Let ρ be a Tukey's bi-square objective function as (Susanti et al. 2014),

$$\rho(u_i)=\begin{cases}\dfrac{u_i^2}{2}-\dfrac{u_i^4}{2c^2}+\dfrac{u_i^6}{6c^4}; & -c\le u_i\le c\\ \dfrac{c^2}{6}; & u_i<-c \text{ or } u_i>c.\end{cases} \tag{2}$$

where c is the turning constant. The frequent choice for the constant c (see in (2)) and k (see in (1)) are 1.548 and 0.1995, respectively, for 50% breakdown and about 28% asymptotic efficiency (Rousseeuw and Leroy 1987). MM-estimation provides the equation as (Susanti et al. 2014):

$$\sum_{i=1}^{n}\rho_i(u_i)X_{ij}=0 \quad\text{or}\quad \sum_{i=1}^{n}\rho_i\left(\frac{Y_i-\sum_{j=0}^{k}X_{ij}\hat{\beta}_j}{S_{MM}}\right)X_{ij}=0,$$

where S_{MM}, the standard deviation, obtained from the residual of S-estimation. The algorithm of M-estimation adjusted from Susanti et al. (2014) is

Step 1 Approximate regression coefficients on the data set using OLS.

Step 2 Compute residual value $e_i=y_i-\hat{y}_i$.

Step 3 Compute value $\hat{\sigma}_{sn}$ by (1).

Step 4 Compute value $u_i=\dfrac{e_i}{\hat{\sigma}_{sn}}$.

Step 5 Compute weighted value w_i,

$$w_i=\begin{cases}\left[1-\left(\dfrac{u_i}{4.685}\right)^2\right]^2, & -4.685\le u_i\le 4.685\\ 0, & u_i<-4.685 \text{ or } u_i>4.685.\end{cases}$$

Step 6 Compute $\hat{\beta}_{MM}$ by using a weighted least squares method with weighted w_i.

Step 7 Repeat Steps 3-7 to obtain a convergent values of $\hat{\beta}_{MM}$ (or $\left|\hat{\beta}_{MM}^{k+1}-\hat{\beta}_{MM}^{k}\right|<\varepsilon,\ \varepsilon>0.$ is a very small number.) In this study, $\varepsilon=0.0005$ is used.

4. Fuzzy Robust Regression Model

Let $X_1,X_2,\ldots,X_p$ be explanatory variables or independent variables and Y be dependent variable. A triangular fuzzy number of them can be written as $\mathbf{X}=(x,\underline{x},\overline{x})$, where x is the element value of $\mathbf{X}$, $\underline{x}$ and $\overline{x}$ are left spreads and right spreads, respectively, $\mathbf{Y}=(y,\underline{y},\overline{y})$, where y is the element value of $\mathbf{Y}$, $\underline{y}$ and $\overline{y}$ are left spreads and right spreads, respectively, the steps of the proposed method are as the following (Kula et al. 2012):

Step 1 Define fuzzy regression model, $Y_i=b_0+b_1X_i$, where $X_i=(x_i,\underline{x}_i,\overline{x}_i)$ and $Y_i=(y_i,\underline{y}_i,\overline{y}_i)$; $i=1,2,\ldots,n$ are triangular fuzzy number, b_0 and b_1 are crisp number. The optimization of the fuzzy least squares is shown as (3).

$$\min r(b_0,b_1)=\sum d\left(b_0+b_1X_i,Y_i\right)^2, \tag{3}$$

where

$$d\left(b_0+b_1X_i,Y_i\right)^2=\left[b_0+b_1x_i-y_i-\left(b\underline{x}_i-\underline{y}_i\right)\right]^2+\left[b_0+b_1x_i-y_i-\left(b\overline{x}_i-\overline{y}_i\right)\right]^2+\left(b_0+b_1x_i-y_i\right)^2. \tag{4}$$

In this study, the fuzzy least squares model is applied for multiple regression model. Thus, (4) can be written as

$$\min r\left(b_0,b_1,b_2,\ldots,b_p\right)=\sum d\left(b_0+b_1X_{i1}+b_2X_{i2}+\ldots+b_pX_{ip},Y_i\right)^2.$$

The initial estimation parameter can be calculated as

$$\hat{\boldsymbol{\beta}}_F^0=\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}+\underline{\mathbf{X}}^{\mathrm{T}}\underline{\mathbf{X}}+\overline{\mathbf{X}}^{\mathrm{T}}\overline{\mathbf{X}}\right)^{-1}\left(\mathbf{X}^{\mathrm{T}}\mathbf{Y}+\underline{\mathbf{X}}^{\mathrm{T}}\underline{\mathbf{Y}}+\overline{\mathbf{X}}^{\mathrm{T}}\overline{\mathbf{Y}}\right),$$

where $\mathbf{X}=\begin{bmatrix}1 & x_{11} & \cdots & x_{1p}\\ 1 & x_{21} & \cdots & x_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & x_{n1} & \cdots & x_{np}\end{bmatrix}$, $\mathbf{Y}=\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{bmatrix}$, $\underline{X}=\begin{bmatrix}1 & x_{11}-\underline{x}_{11} & \cdots & x_{1p}-\underline{x}_{1p}\\ 1 & x_{21}-\underline{x}_{21} & \cdots & x_{2p}-\underline{x}_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & x_{n1}-\underline{x}_{n1} & \cdots & x_{np}-\underline{x}_{np}\end{bmatrix}$,

$\overline{\mathbf{X}}=\begin{bmatrix}1 & x_{11}+\overline{x}_{11} & \cdots & x_{1p}+\overline{x}_{1p}\\ 1 & x_{21}+\overline{x}_{21} & \cdots & x_{2p}+\overline{x}_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & x_{n1}+\overline{x}_{n1} & \cdots & x_{np}+\overline{x}_{np}\end{bmatrix}$, $\underline{\mathbf{Y}}=\begin{bmatrix}y_1-\underline{y}_1\\ y_2-\underline{y}_2\\ \vdots\\ y_n-\underline{y}_n\end{bmatrix}$ and $\overline{\mathbf{Y}}=\begin{bmatrix}y_1+\overline{y}_1\\ y_2+\overline{y}_2\\ \vdots\\ y_n+\overline{y}_n\end{bmatrix}$.

The $(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}$ is exist.

Step 2 Compute the residual $e_i=y_i-\hat{y}_i$.

Step 3 Determine median with defer to the absolute residual values. Then, the distance is solved by:

$$D=\left\|\mathrm{abs}(e_i)-\mathrm{median}(\mathrm{abs}(e_i))\right\|,\quad i=1,2,\ldots,n\ ,$$

where $\|\bullet\|$ is the Euclidean distance (Sanli and Apaydin 2004).

Step 4 Determine the membership function $\mu(e)$,

$$\mu(e)=\begin{cases}1, & |e|\le p\\ \dfrac{q-|e|}{q-p}, & p<|e|<q\\ 0, & \text{otherwise,}\end{cases}\tag{5}$$

where $p=\mathrm{median}(D_i)$, $q=\max(D_i)+r$ and $r=\dfrac{\mathrm{median}|\ e_i-\mathrm{median}(e_i)|}{0.6745}$. Figure 1 reveals the membership function.

Step 5 Generate fuzzy weighted least square regression (Chang and Lee 1996). Weighted function, used to fit the model, is included the degree of membership. The weight matrix is used for procreation the weighted fuzzy least squares (Sanli and Apaydin 2004). Equation (5) provides the membership function and the values of the membership function. Thus, the weight matrix is composed. The weight matrix used for solving parameter estimation is a diagonal matrix, where $W=diag(\mu_1,\mu_2,\ldots,\mu_n)$ and $\mu_1,\mu_2,\ldots,\mu_n$ is the elements of the degree of membership. The weighted fuzzy least square function can be calculated as

$$\hat{\boldsymbol{\beta}}_F=(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X}+\underline{\mathbf{X}}^{\mathrm{T}}\mathbf{W}\underline{\mathbf{X}}+\overline{\mathbf{X}}^{\mathrm{T}}\mathbf{W}\overline{\mathbf{X}})^{-1}(\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{Y}+\underline{\mathbf{X}}^{\mathrm{T}}\mathbf{W}\underline{\mathbf{Y}}+\overline{\mathbf{X}}^{\mathrm{T}}\mathbf{W}\overline{\mathbf{Y}}).$$

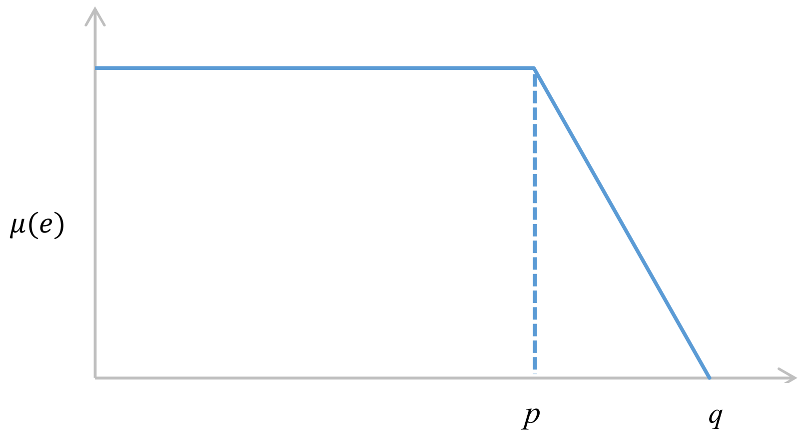

Figure 1 Picture of the membership function

Step 6 Repeat Steps 2-5 to convergent values of $\hat{\beta}_F$ (or $\left|\hat{\beta}_F^{k+1} - \hat{\beta}_F^{k}\right| < \varepsilon,\ \varepsilon > 0$ is a very small number). In this study, $\varepsilon = 0.0005$ is used.

5. Evaluations of the Model

The performance of the both methods is compared by the mean absolute error (MAE) and the estimated mean square error (EMSE) as (6) and (7), respectively,

$$MAE = n^{-1}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|, \quad \text{(Willmott and Matsuura 2005)} \tag{6}$$

$$EMSE = \frac{1}{1{,}000}\sum_{t=1}^{1{,}000}\sum_{j=0}^{p}\left(\hat{\beta}_{j(t)} - \beta_j\right)^2, \quad \text{(Ozkale and Arican 2015)} \tag{7}$$

6. Numerical Examples

The model $y_0 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon_i$ is generated 1,000 data sets by using MATLAB 9.7.0 (R2019b) software with the fixed regression coefficients $\beta_0 = 1, \beta_1 = 1$ and $\beta_2 = 1,$ respectively. The error term (ε_i) is assumed to be independent (Ampanthong and Suwattee 2009) and is generated from the normal distribution N(0,1). In this study, the independent variables, $x_{ij}, i = 1,\ldots,n, j = 1,2$ are generated from the uniform distribution (−5, 5) with sample size 20, 40, 60, 80 and 100, respectively. The percentage of outliers 10%, 20% and 30% are designed for comparing. The generated 50% of outliers are less than $Q_1 - 3IQR$ and 50% of outlier are more than $Q_3 + 3IQR,$ where Q_1 is the first quartile, Q_3 is the third quartile and IQR is the interquartile range. In case of outliers in X and $Y,$ the chosen center, left and right spreads of X and Y are $(x_{ij}, y_i), \left(\frac{x_{ij}}{1.5}, \frac{y_i}{1.5}\right)$ and $\left(\frac{x_{ij}}{2}, \frac{y_i}{2}\right),$ respectively. In case of outliers in $X,$ the chosen center, left and right spreads of X

and Y are $(x_{ij}, y_i), \left(\frac{x_{ij}}{1.5}, \frac{y_i}{1.5}\right)$ and $\left(\frac{x_{ij}}{1}, \frac{y_i}{1}\right)$, respectively. In case of outliers in Y, the chosen center, left and right spreads of X and Y are $(x_{ij}, y_i), \left(\frac{x_{ij}}{0.75}, \frac{y_i}{0.75}\right)$ and $\left(\frac{x_{ij}}{0.5}, \frac{y_i}{0.5}\right)$, respectively.

Figure 2 represents the generated data sets with sample size 100 and 10%, 20% and 30% of outliers. The results of a performance comparison of the both methods are shown in Table 1.

Table 1 The values of EMSE and MAE by MM-estimation and FRR method by sample sizes with percentage of both X 's and Y 's outliers, percentage of X 's outliers and percentage of Y 's outliers

Sample Size	% of outliers	Outliers in X and Y				Outliers in X				Outliers in Y			
		MM-estimation		FRR method		MM-estimation		FRR method		MM-estimation		FRR method	
		EMSE	MAE	EMSE	MAE	EMSE	MAE	EMSE	MAE	EMSE	MAE	EMSE	MAE
20	10	6.719	3.901	2.066	3.482	3.270	2.525	1.271	2.328	**0.162**	3.065	**0.217**	2.365
	20	15.198	4.519	9.807	4.344	5.696	2.784	5.650	2.744	5.155	4.919	2.041	4.074
	30	18.819	4.697	14.835	4.607	**7.346**	2.982	**7.777**	2.934	12.708	5.911	6.139	5.144
40	10	7.326	4.011	2.837	3.645	3.226	2.596	1.339	2.407	**0.031**	3.027	**0.334**	2.403
	20	14.279	4.552	10.666	4.482	5.072	2.903	4.487	2.871	**0.866**	4.884	**1.395**	3.925
	30	18.344	4.761	14.704	4.714	**6.490**	3.005	**6.610**	2.973	8.133	6.307	4.052	5.324
60	10	7.247	4.034	3.320	3.676	3.066	2.629	1.542	2.441	**0.021**	3.109	**0.401**	2.483
	20	14.285	4.593	11.125	4.545	5.031	2.980	4.425	2.952	**0.781**	1.874	**5.250**	4.245
	30	17.872	4.812	14.626	4.788	6.801	3.061	6.732	3.045	5.204	6.612	3.839	5.508
80	10	7.589	4.039	3.634	3.706	3.049	1.785	2.634	2.451	**0.022**	3.145	**0.491**	2.536
	20	13.673	4.627	10.685	4.581	5.397	2.977	4.901	2.946	**0.216**	5.445	**1.767**	4.343
	30	17.783	4.844	14.443	4.824	6.700	3.051	6.600	3.035	9.794	7.099	4.877	6.007
100	10	7.357	4.022	3.812	3.711	2.778	2.675	1.612	2.508	**0.013**	3.199	**0.480**	2.574
	20	13.798	4.618	11.006	4.580	5.368	2.961	4.863	2.938	**0.270**	5.461	**1.893**	4.361
	30	17.654	4.820	14.456	4.809	6.645	3.107	6.526	3.093	9.689	7.164	5.646	6.105

Note: These values are calculated via MATLAB 9.7.0 (R2019b).

For Table 1, it is found that the MAEs by the FRR method are lower than the MM-estimation for almost all types of outliers. The estimated mean square values (or EMSE) by the FRR method are clearly lower than the MM-estimation when outliers containing in both X and Y. In case of outliers in X, the EMSEs by the FRR method are lower than the MM-estimation except in case of 30% of outliers when n = 20 and 40. In case of 10% and 20% of outliers in Y, the EMSEs by the MM-estimation are lower than the FRR method.

7. Conclusions and Discussions

In this study, the efficacy of the MM-estimation method and the FRR method are compared using generated thousand data with three types of outliers. The FRR method provides the model with lower EMSE and MAE than the MM-estimation method almost all types of outliers, especially, the data sets containing outliers in both X and Y. However, there are some cases that the EMSE by the FRR method are higher than the MM-estimation such as 30% of outliers in X when n = 20 and 40 and 10%, 20% of outliers in Y. From the results, it is said that the MM-estimation is effective for modeling when outliers containing in Y are not over 20%. On the other hand, the FRR method is more effective than the MM-estimation when outliers in Y are greater than 20%. If the data set contains excessive outliers, the values of EMSE and MAE tend to be higher that means the parameter estimations are bias estimation, so the both methods are not suitable for modeling.

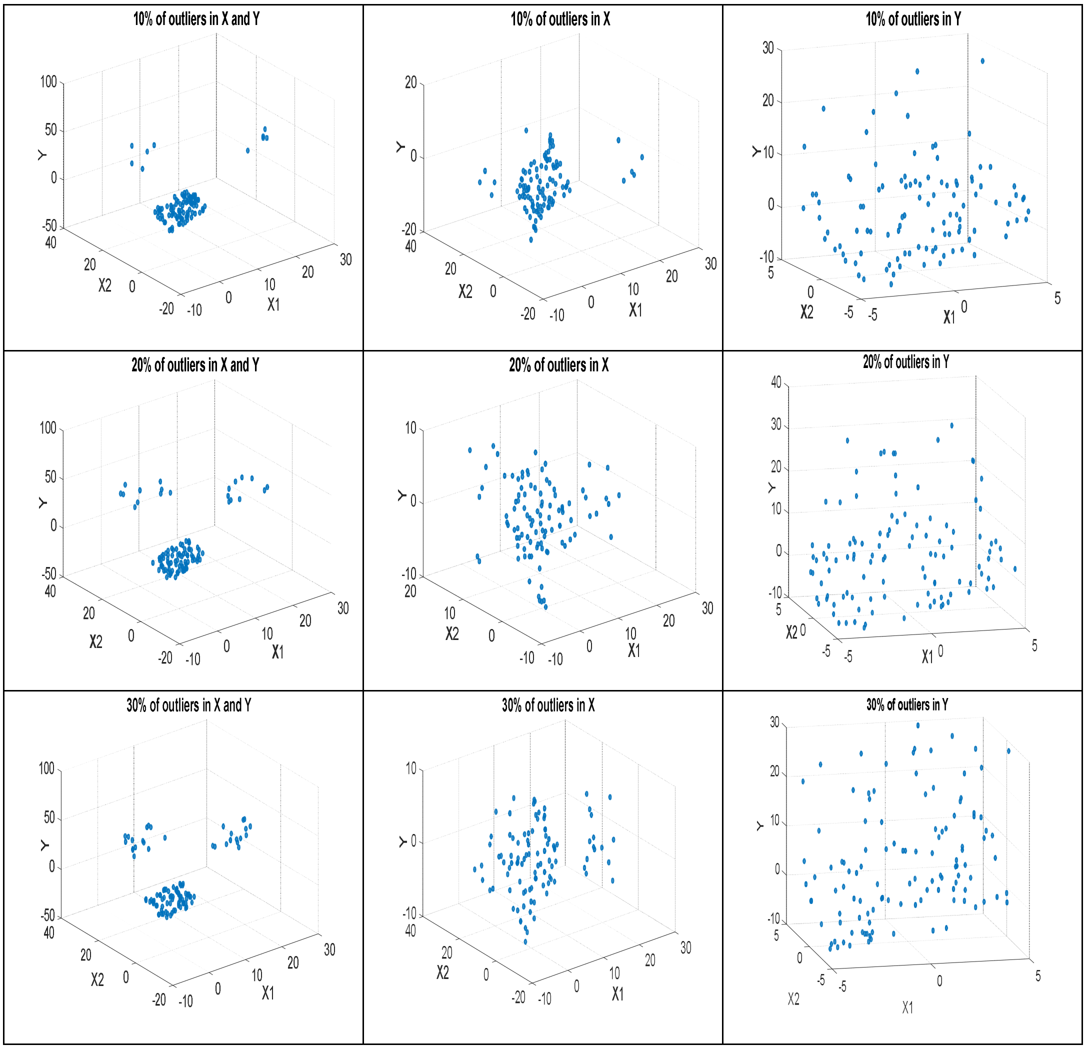

Figure 2 Scatter plot for the generated data sets (n =100) containing 10%, 20% and 30% of outliers in X, outliers in Y and outliers in X and Y

Acknowledgements

The authors are thankful for anonymous reviewers and the editor for their fruitful suggestion which improved the quality of the manuscript.

References

Alma OG. Comparison of robust regression methods in linear regression. Int J Contemp Math Sci. 2011; 6(9): 409-421.

Ampanthong P, Suwattee P. Comparative study of outlier detection procedures in multiple linear regression. In: Proceedings of the International Multi Conference of Engineers and Computer Scientists 2009 Vol I; 2009 March 18-20; Hong Kong. 2009.

Chang PT, Lee ES. A generalized fuzzy weighted least-squares regression. Fuzzy Set Syst. 1996; 82: 289-298.

Chen C. Robust regression and outlier detection with the ROBUSTREG Procedure. North Carolina: SAS Institute Inc; 2002.

Diamond P. Fuzzy least squares. Inform Sci. 1988; 46: 141-157.

Donoho DL, Huber PJ. The notion of breakdown point. In: Bickel PJ, Doksum KA, Hodges JL, eds. A Festschrift for Erich L. Lehmann. Belmont: 1983. p. 157-184.

Huber PJ. Robust estimation of a location parameter. Ann Math Statist. 1964; 35(1): 73-101.

Kula KS, Tank F, Dalkılıç TE. A study on fuzzy robust regression and its application to insurance. Math Comput Appl. 2012; 17(3): 223-234.

Mahaboob B, Venkateswarlu B, Narayana C, Ravi Sankar J, Balasiddamuni P. A treatise on ordinary least squares estimation of parameters of linear model. Int J Engineering Tech. 2018; 7(4.10): 518-522.

Nasrabadi MM, Nasrabadi E. A mathematical-programming approach to fuzzy linear regression analysis. Math Comput Appl. 2004; 155: 873-881.

Ozkale MR, Arican E. First-order r-d class estimator in binary logistic regression model. Stat Prob Lett. 2015; 106: 19-29.

Redden DT, Woddall WH. Futher examination of fuzzy linear regression. Fuzzy Set Syst. 1996; 79(2): 203-211.

Rousseeuw PJ, Yohai V. Robust regression by means of S-estimators in robust and nonlinear time series analysis. In: Franke J, Härdle W, Martin D, editors. Robust and nonlinear time series analysis, New York: Springer-Verlag; 1984.

Rousseeuw PJ, Leroy AM. Robust regression and outlier detection. New York. John Wiley & Sons; 1987.

Salibian-Barrera M, Yohai VJ. A fast algorithm for S-regression estimates. J Comput Graph Stat. 2006; 15(2): 414-427.

Sanli K, Apaydin A. The fuzzy robust regression analysis, the case of fuzzy data set has outlier Gazi Univ J Sci. 2004; 17: 71-84.

Susanti Y, Pratiwi H, Sulistijowati HS, Liana T. M-estimation, S-estimation and MM-estimation in robust regression. Int J Pure Appl Math. 2014; 91(3): 349-360.

Tanaka H, Uegima S, Asai K. Linear regression analysis with fuzzy model. IEEE Transactions on Systems Man Cybernet. 1982; 12: 903-907.

Yang MS, Ko CH. On cluster-wise fuzzy regression analysis. IEEE Transactions on Systems Man and Cybernetics. Part B: Cybernetics. 1997; 27: 1-13.

Yang MS, Lin TS. Fuzzy least-squares linear regression analysis for fuzzy input-output data. Fuzzy Set Syst. 2002; 126: 389-399.

Yohai JV. High breakdown-point and high efficiency robust estimates for regression. Ann Statist. 1987; 17(4): 1662-1683.

Wei BC, Lin JG, Xie FC. Diagnostic statistics. Beijing: Higher Education Press; 2009.

Willmott CJ, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Res. 2005; 30(1): 79-82.