African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 6, Issue 3, 2023 (pp. 59-69)
www.abjournals.org

REVIEW OF SOME ROBUST ESTIMATORS IN MULTIPLE LINEAR REGRESSIONS IN THE PRESENCE OF OUTLIER(s)

Alanamu T.[1], Oyeyemi G.M.[2], Olaniran R.O.[3] and Adetunji K.O.[4]

[1]Department of Mathematics, Kwara State College of Education, Ilorin

taoheedatalanamu@yahoo.com

[2]Department of Statistics, University of Ilorin, Nigeria

gmoyeyemi@gmail.com

[3]Department of Staistics, University of Ilorin, Nigeria

rid4stat@yahoo.com

[4]Department of Mathematics, Kwara State College of Education, Ilorin

kalmarx4@yahoo.com

Cite this article:

Alanamu T., Oyeyemi G.M., Olaniran R.O., Adetunji K.O. (2023), Review of Some Robust Estimators in Multiple Linear Regressions in the Presence of Outlier(s). African Journal of Mathematics and Statistics Studies 6(3), 59-69. DOI: 10.52589/AJMSS-EC8EWXUL

Manuscript History

Received: 14 April 2023

Accepted: 5 May 2023

Published: 26 June 2023

ABSTRACT: *Linear regression has been one of the most important statistical data analysis tools. Multiple regression is a type of regression where the dependent variable shows a linear relationship with two or more independent variables. OLS estimate is extremely sensitive to unusual observations (outliers), with low breakdown point and low efficiency. This paper reviews and compares some of the existing robust methods (Least Absolute Deviation, Huber M-Estimator, Bisquare M-Estimator, MM Estimator, Least Median Square, Least Trimmed Square, S-Estimator); a simulation method is used to compare the selected existing methods. It was concluded based on the results that for y direction outlier, the best estimator in terms of high efficiency and breakdown point of at most 0.3 is MM; for x direction outlier, the best estimator in term breakdown point of at most 0.4 is S; for x, y direction outlier, the best estimator in terms of high efficiency and breakdown point of at most 0.2 is MM.*

KEYWORDS: Linear Regression, Breakdown Point, Robust Estimators, Outlier.

Article DOI: 10.52589/AJMSS-EC8EWXUL
DOI URL: https://doi.org/10.52589/AJMSS-EC8EWXUL

INTRODUCTION

Multiple linear regression is a statistical technique that is used to predict the outcome of a variable based on the value of two or more variables. Multiple regression is a type of regression where the dependent variable shows a linear relationship with two or more independent variables.

Linear regression has been one of the most important statistical data analysis tools. Given the independent and identically distributed (iid) observations $(x_i,\ y_i), i = 1,2,3,. \ . \ \ ., n$ in order to understand how the response y_i is related to the covariates x_i, we traditionally assume the linear regression model $y_i = x_i^T\beta + \varepsilon_i$, (1)

where β is an unknown $p \times 1$ vector, and the ε_i are i.i.d. and independent of x_i with $E(x_i) = 0$. The most commonly used estimate for β is the ordinary least square (OLS) estimate which minimizes the sum of squared residuals $\sum_{i=1}^{n} \ (y_i - x_i^T\beta)^2$ (2)

However, it is well known that the OLS estimate is extremely sensitive to the unusual observations (outliers). Many robust methods have been proposed to achieve high breakdown point or high efficiency or both, as the efficiency and breakdown point are two important criteria for comparing robust methods (Donoho & Huber, 1983). In this paper, we review and compare some of the existing robust methods; a simulation method is used to compare the selected existing methods.

The efficiency is used to measure the relative efficiency of the robust methods compared to the OLS estimate when the error distribution is exactly normal and there are no outliers. Breakdown point is used to measure the proportion of outliers an estimate can tolerate before it goes to infinity.

In this paper, finite sample breakdown point (Yu et al., 2014) is used and defined as follows: Let $z_i = (x_i,\ y_i)$. Given any sample $z = (z_i, \dots, z_n)$, denote $T(z)$ the estimate of the parameter β. Let z' be the corrupted sample where any m of the original points of z are replaced by arbitrary bad data. Then the finite sample breakdown point $\delta^*$is defined as:

$$\delta^*(z,T) = \left\{\frac{m}{n}: \|T(z') - T(z)\| = \infty\right\}, \qquad (3)$$

where || . || is the Euclidean norm.

M-estimates (Huber, 1981) are solutions of the normal equation with appropriate weight functions. They are resistant to unusual y observations but sensitive to high leverage points on x; the breakdown point of an M-estimate is $1/n$. R-estimates (Jaeckel, 1972) minimize the sum of scores of the ranked residuals; they have relatively high efficiency but with breakdown points as low as those of OLS estimates. Least Median of Squares (LMS) estimates (Siegel, 1982) minimize the median of squared residuals, Least Trimmed Squares (LTS) estimates (Rousseeuw, 1983) minimize the trimmed sum of squared residuals, and S-estimates (Rousseeuw & Yohai, 1984) minimize the variance of the residuals; all have high breakdown point but low efficiency. Generalized S-estimates (GS-estimates) (Croux et al., 1994) maintain high breakdown point as S-estimates and have slightly higher efficiency. MM-estimates proposed by Yohai (1987) simultaneously can attain high breakdown point and efficiencies. Mallows Generalized M-estimates (Mallows, 1975) and Schweppe

Article DOI: 10.52589/AJMSS-EC8EWXUL
DOI URL: https://doi.org/10.52589/AJMSS-EC8EWXUL

Generalized M-estimates (Handschin et al., 1975) downweight the high leverage points on x but cannot distinguish "good" and "bad" leverage points, thus resulting in a loss of efficiencies.

In addition, Mallows Generalized M-estimates (Mallows, 1975) and Schweppe Generalized M-estimates (Handschin et al., 1975) have low breakdown points when the number of explanatory variables is large. Schweppe one-step (S1S) Generalized M-estimates (Coakley & Hettmansperger, 1993) overcome the problems of Schweppe Generalized M-estimates, and are calculated in one step. They both have high breakdown points and high efficiencies. Gervini and Yohai (2002) proposed a new class of high breakdown point and high efficiency robust estimate called robust and efficient weighted least squares estimator (REWLSE). Lee et al. (2011) and She and Owen (2011) proposed a new class of robust methods based on the regularization of case specific parameters for each response.

ROBUST REGRESSION METHODS

Least Absolute Deviation (LAD)

Least absolute deviation (LAD, also called median regression) estimates is achieved by minimizing the sum of the absolute values of the residuals:

$$\hat{\beta} = arg \sum_{i=1}^{n} \left| y_i - x_i^T \beta \right| . \qquad (4)$$

The LAD is also called L_1 estimate due to the L_1 norm used. Although LAD is more resistant than OLS to unusual y values, it is sensitive to high leverage outliers, and thus has a breakdown point of $BP = 1/n \to 0$ (Rousseeuw & Yohai, 1984). Moreover, LAD estimates have a low efficiency of 0.64 when the errors are normally distributed.

Least Median Squares (LMS) Estimates

The LMS estimates (Siegel, 1982) are found by minimizing the median of the squared residuals:

$$\hat{\beta} = argMed \left| y_i - x_i^T \beta \right| . \qquad (5)$$

One good property of the LMS estimate is that it possesses a high breakdown point of nearly 0.5. However, the LMS estimate has at best an efficiency of 0.37 when the assumption of normal errors is met (see Rousseeuw & Croux, 1993). Moreover, LMS estimates do not have a well-defined influence function because of its convergence rate of $n^{-\frac{1}{3}}$ (Rousseeuw, 1982). Despite these limitations, the LMS estimate can be used as the initial estimate for some other high breakdown points and high efficiency robust methods.

Article DOI: 10.52589/AJMSS-EC8EWXUL
DOI URL: https://doi.org/10.52589/AJMSS-EC8EWXUL

Least Trimmed Squares (LTS) Estimates

The LTS estimate (Rousseeuw, 1983) is defined as

$$\hat{\beta} = arg \sum_{i=1}^{q} \quad r_{(i)}(\beta)^2 \, . \qquad (6)$$

where $r_{(i)}(\beta) = y_{(i)} - x_{(i)}^T \beta$, $r_{(1)}(\beta)^2 \leq . \; . \; . \leq r_{(q)}(\beta)^2$are ordered squared residuals, $q = [n\,(1 - \alpha) + 1]$, and α is the proportion of trimming. Using $q = \left(\frac{n}{2}\right) + 1$ ensures that the estimator has a breakdown point of BP = 0.5, and the convergence rate of $n^{-\frac{1}{2}}$ (Rousseeuw, 1983). Although highly resistant to outliers, LTS suffers badly in terms of very low efficiency, which is about 0.08, relative to OLS estimates (Stromberg et al., 2000). The reason that LTS estimates call our attention is that it is traditionally used as the initial estimate for some other high breakdown point and high efficiency robust methods.

M-Estimates

By replacing the least squares criterion (2) with a robust criterion, M-estimate (Huber, 1964) of β is

$$\hat{\beta} = arg \sum_{i=1}^{n} \quad \rho\left(\frac{y_i - x_i^T \beta}{\hat{\sigma}}\right) \qquad (7)$$

where $\rho(\cdot)$ is a robust loss function and $\hat{\sigma}$ is an error scale estimate. The derivative of ρ, denoted by $\psi(\cdot) = \rho'(\cdot)$, is called the influence function. In particular, if $\rho(t) = \frac{1}{2}t^2$, then the solution is the OLS estimate. The OLS estimate is very sensitive to outliers. Rousseeuw and Yohai (1984) indicated that OLS estimates have a breakdown point $BP = 1/n$, which tends to zero when the sample size n is getting large. Therefore, one single unusual observation can have a large impact on the OLS estimate. One of the commonly used robust loss functions is Huber's ψ function (Huber, 1981), where $\psi_c(t) = \rho'(t) = max\{-c, min(c, t)\}$. Huber (1981) recommends using c = 1.345 in practice. This choice produces a relative efficiency of approximately 95% when the error density is normal. Another possibility for $\psi(\cdot)$ is Tukey's bisquare function $\psi_c(t) = t\{1 - (t/c)^2\}_+^2$. The use of c = 4.685 produces 95% efficiency.

S-Estimates

S-estimates (Rousseeuw & Yohai, 1984) are defined by

$$\hat{\beta} = arg \; arg \; \hat{\sigma} \; (r_1(\beta), . \; . \; . , r_n(\beta)), \qquad (8)$$

where $r_i(\beta) = y_i - x_i^T \beta$ and $\hat{\sigma}(r_1(\beta), . \; . \; . , r_n(\beta))$ is the scale M-estimate which is defined as the solution of

$$\frac{1}{n} \sum_{i=1}^{n} \quad \left(\frac{r_i(\beta)}{\hat{\sigma}}\right) = \delta \qquad (9)$$

for any given β, where δ is taken to be $E_{\phi}[\rho\,(r)]$. For the biweight scale, S-estimates can attain a high breakdown point of BP = 0.5 and has an asymptotic efficiency of 0.29 under the assumption of normally distributed errors (Maronna, Martin & Yahai, 2006).

Article DOI: 10.52589/AJMSS-EC8EWXUL
DOI URL: https://doi.org/10.52589/AJMSS-EC8EWXUL

MM-Estimates

First proposed by Yohai (1987), MM-estimates have become increasingly popular and are one of the most commonly employed robust regression techniques. The MM-estimates can be found by a three-stage procedure. In the first stage, compute an initial consistent estimate $\hat{\beta}_0$with high breakdown point but possibly low normal efficiency. In the second stage, compute a robust M-estimate of scale $\hat{\sigma}$ of the residuals based on the initial estimate. In the third stage, find an M-estimate $\hat{\beta}$ starting at $\hat{\beta}_0$. In practice, LMS or S-estimate with Huber or bisquare functions is typically used as the initial estimate $\hat{\beta}_0$. Let $\rho_0(r) = \rho_1(r/k_0)$, $\rho(r) = \rho_1(r/k_1)$, and assume that each of the ρ-functions is bounded. The scale estimate $\hat{\sigma}$ satisfies

$$\frac{1}{n}\sum_{i=1}^{n} \rho_0\left(\frac{r_i(\hat{\beta})}{\hat{\sigma}}\right) = 0.5 \qquad (10)$$

If the ρ-function is biweight, then $k_0 = 1.56$ ensures that the estimator has the asymptotic $BP = 0.5$. Note that an M-estimate minimizes

$$L(\beta) = \sum_{i=1}^{n} \rho\left(\frac{r_i(\hat{\beta})}{\hat{\sigma}}\right) \qquad (11)$$

Let ρ satisfy $\rho \leq \rho_0$. Yohai (1987) showed that if $\hat{\beta}$ satisfies $L(\beta) \leq (\hat{\beta}_0)$, then $\hat{\beta}$'s BP is not less than that of $\hat{\beta}_0$. Furthermore, the breakdown point of the MM-estimate depends only on k_0 and the asymptotic variance of the MM-estimate depends only on k_1. We can choose $k_1$in order to attain the desired normal efficiency without affecting its breakdown point. In order to let $\rho \leq \rho_0$, we must have $k_1 \geq k_0$; the larger the k_1 is, the higher efficiency the MM-estimate can attain at the normal distribution.

Maronna, Martin, and Yahai (2006) provide the values of k_1 with the corresponding efficiencies of the biweight ρ-function. However, Yohai (1987) indicates that MM-estimates with larger values of k_1 are more sensitive to outliers than the estimates corresponding to smaller values of k_1. In practice, an MM-estimate with bisquare function and efficiency 0.85 ($k_1 = 3.44$) starting from a bisquare S-estimate is recommended.

Simulation Study

In this part, we compare different robust methods and report the mean squared errors (MSE) and relative efficiency of the parameter estimates for each estimation method. We compare the OLS estimate with seven other commonly used robust regression estimates: the M estimate using Huber's ψ function (M-H), the M estimate using Tukey's bi-square function (M-T), the S estimate, the LTS estimate, the LMS estimate, the MM estimate (using bi-square weights and k_1 = 4.68), and the LAD.

The data generation processes that follow were adapted from Yu, Yao and Bai (2014). We generated n samples from the model

$$Y = X\beta + \varepsilon \qquad (12)$$

where $X = [X_1, X_2, X_3] \sim N_3(0, I)$, $\beta = [1,\ 1,\ 1]$. In order to compare the performance of different methods, we consider the following three cases for the error density of ε and independent variables X:

Article DOI: 10.52589/AJMSS-EC8EWXUL
DOI URL: https://doi.org/10.52589/AJMSS-EC8EWXUL

Case I (with y direction outlier): $\varepsilon \sim (1-\pi)N(0,1) + \pi N(0,10^2)$ - contaminated normal mixture with $\pi = 0.1, 0.2, 0.3, 0.4, 0.5$ proportions of contamination.

Case II (with **x** direction outlier): $X \sim (1-\pi)N_3(0,I) + \pi N_3(0,10^2 I)$ - contaminated multivariate normal mixture with $\pi = 0.1, 0.2, 0.3, 0.4, 0.5$ proportions of contamination.

Case III (with **x**, y direction outlier):$X \sim (1-\pi)N_3(0,I) + \pi N_3(0,10^2 I)$ – contaminated multivariate normal mixture and $\varepsilon \sim (1-\pi)N(0,1) + \pi N(0,10^2)$ with $\pi = 0.1, 0.2, 0.3, 0.4, 0.5$ proportions of contamination, and in the overall cases, the sample size $n = 10, 30, 100,\ 200\ and\ 1000$ were used. The replication size was fixed at 1000.

Criteria for Assessing the Estimators Performance

Mean Square Error of Parameter

$$MSE_i = \frac{\sum_{j=1}^{p} (\hat{\beta}-\beta)^2}{p} \quad (13)$$

$$AMSE = \frac{\sum_{i=1}^{I} MSE_i}{I} \quad (14)$$

Relative Efficiency of Robust Estimators

$$RE = \frac{AMSE_{OLS}}{AMSE_{RE}} \quad (15)$$

where I is the number of iteration and p is the number of parameters in the model.

Table 1: Average MSE of various methods over all sample sizes for case I at $\pi = 0.1, 0.2, 0.3, 0.4, 0.5$

π	LS	LA	H-M	B-M	MM	LMS	LTS	S
0	0.032	0.0489	0.034	0.0369	0.0441	**0.5468**	**0.1506**	**0.0955**
0.1	0.3504	0.0904	0.0982	0.0902	0.0677	**1.9829**	**0.2115**	**0.1105**
0.2	0.6947	0.1918	0.2403	0.2092	0.1182	**7.6956**	**0.2907**	**0.1365**
0.3	1.0874	0.4212	0.5151	0.4752	0.4158	**11.284**	**0.5122**	**0.3925**
0.4	1.4611	0.7532	0.845	0.8359	0.9152	**13.392**	**1.4485**	**0.9707**
0.5	1.8004	1.1028	1.2276	1.2752	1.4132	**17.4091**	**2.6971**	**1.7331**

Table 2: Average RE of various methods over all sample sizes for case I at $\pi = 0.1, 0.2, 0.3, 0.4, 0.5$

π	LS	LA	H-M	B-M	MM	LMS	LTS	S
0	100	64.3	94.02	90.91	87.52	**11.94**	**15.71**	**24.34**
0.1	7.64	46.79	55.63	69.59	72.34	**11.14**	**13.91**	**22.95**
0.2	3.68	32.88	31.18	51.98	53.27	**10.72**	**12.98**	**20.94**
0.3	2.28	22.62	15.58	34.04	32.74	**10.76**	**12.63**	**16.64**
0.4	1.62	14.07	6.77	16.48	14.59	**9.48**	**11.74**	**12.9**
0.5	1.27	8.41	2.99	5.58	4.73	**7.14**	**9.71**	**7.82**

Article DOI: 10.52589/AJMSS-EC8EWXUL
DOI URL: https://doi.org/10.52589/AJMSS-EC8EWXUL

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 6, Issue 3, 2023 (pp. 59-69)

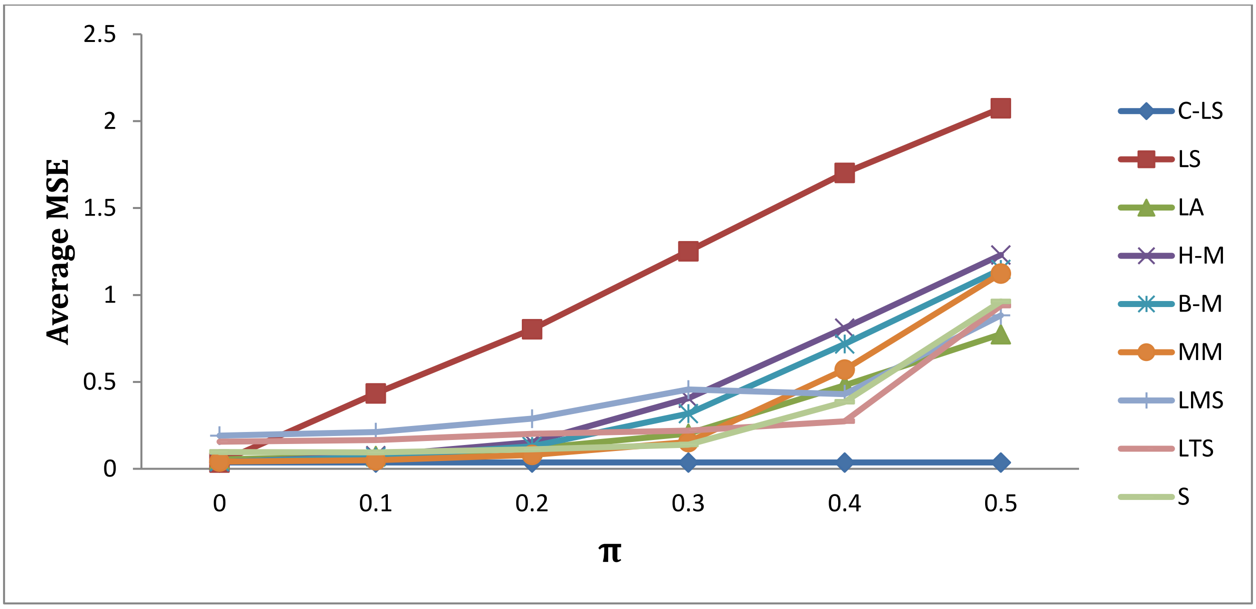

Fig 1: Plot of average MSE against proportion of outlier in the y direction for the various methods

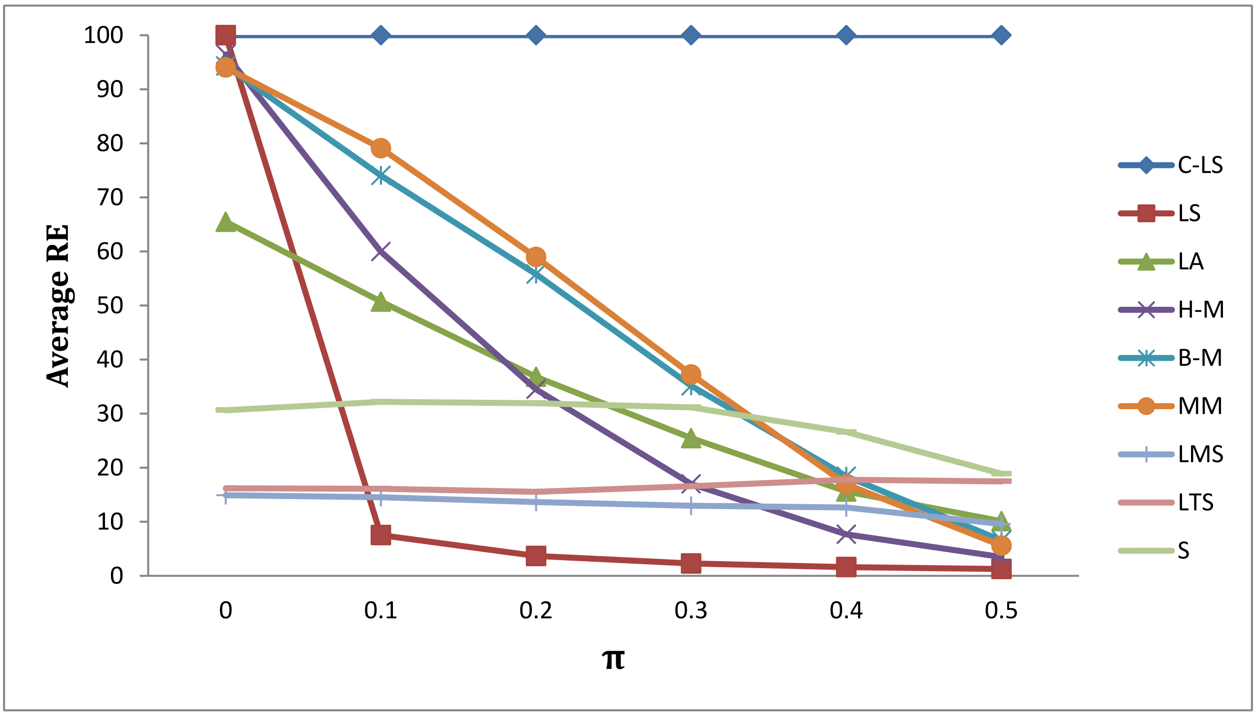

Fig 2: Plot of average RE against proportion of outlier in the y direction for the various methods

Article DOI: 10.52589/AJMSS-EC8EWXUL
DOI URL: https://doi.org/10.52589/AJMSS-EC8EWXUL

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 6, Issue 3, 2023 (pp. 59-69)
www.abjournals.org

Table 3: Average MSE of various methods over all sample sizes for case II at $\pi = \mathbf{0.1, 0.2, 0.3, 0.4, 0.5}$

π	C-LS	LS	LA	H-M	B-M	MM	LMS	LTS	S
0	0.0365	0.0365	0.0540	0.0385	0.0404	0.0414	**0.1909**	**0.1559**	**0.0967**
0.1	0.0365	0.5335	0.5013	0.5125	0.4515	0.0714	**0.6750**	**0.1911**	**0.1153**
0.2	0.0365	0.6551	0.6913	0.6610	0.6675	0.1176	**0.9939**	**0.2165**	**0.1446**
0.3	0.0365	0.7118	0.7494	0.7199	0.729	0.4781	**0.7520**	**0.2431**	**0.1993**
0.4	0.0365	0.7368	0.7665	0.7448	0.7535	0.6983	**0.8745**	**0.3689**	**0.6320**
0.5	0.0365	0.7498	0.776	0.7552	0.7611	0.7517	**1.8074**	**0.7781**	**0.8003**

Table 4: Average RE of various methods over all sample sizes for case II at $\pi = \mathbf{0.1, 0.2, 0.3, 0.4, 0.5}$

π	C-LS	LS	LA	H-M	B-M	MM	LMS	LTS	S
0	100	100	65.49	96.47	94.31	94.02	**14.91**	**16.19**	**30.63**
0.1	100	8.04	6.65	7.77	7.48	40.6	**11.35**	**14.03**	**21.49**
0.2	100	5.44	4.50	5.22	4.99	19.7	**10.55**	**13.00**	**18.55**
0.3	100	4.52	3.92	4.36	4.18	7.41	**9.42**	**11.82**	**12.85**
0.4	100	4.16	3.76	4.03	3.89	4.61	**3.52**	**6.49**	**4.44**
0.5	100	4.06	3.73	3.98	3.89	4.02	**1.72**	**3.48**	**3.57**

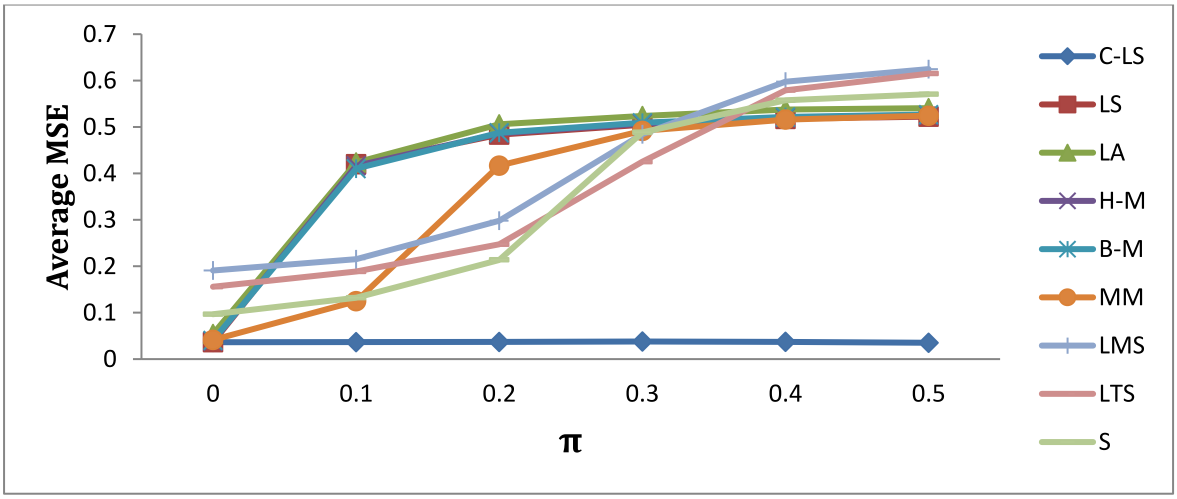

Fig 3: Plot of average MSE against proportion of outlier in the x direction for the various methods

Article DOI: 10.52589/AJMSS-EC8EWXUL
DOI URL: https://doi.org/10.52589/AJMSS-EC8EWXUL

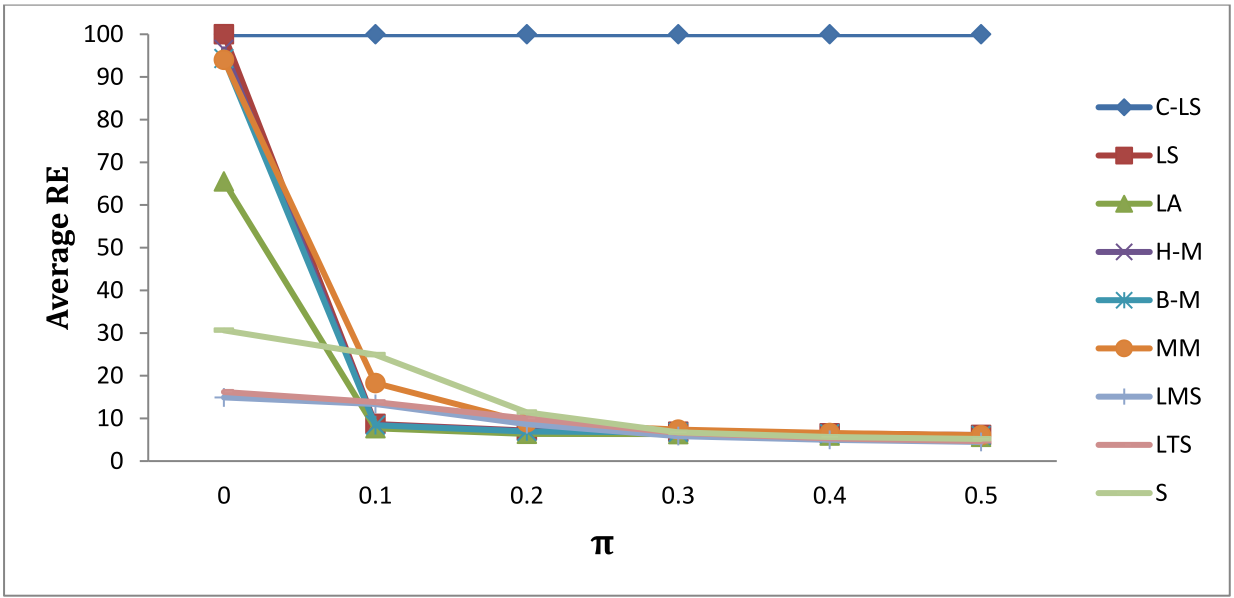

Fig 4: Plot of average RE against proportion of outlier in the x direction for the various methods

Table 5: Average MSE of various methods over all sample sizes for case III at $\pi = \mathbf{0.1, 0.2, 0.3, 0.4, 0.5}$***.l***

π	C-LS	LS	LA	H-M	B-M	MM	LMS	LTS	S
0	0.0365	0.0365	0.0540	0.0385	0.0404	0.0414	**0.1909**	**0.1559**	**0.0967**
0.1	0.0366	0.6544	0.3897	0.4382	0.2511	0.0692	**0.9763**	**0.197**	**0.1126**
0.2	0.0370	0.8387	0.6837	0.7027	0.5091	0.1156	**0.9915**	**0.2114**	**0.1471**
0.3	0.0379	1.0251	0.8726	0.8819	0.7996	0.1974	**0.8775**	**0.2697**	**0.2039**
0.4	0.0372	1.1071	0.9169	0.9427	0.9087	0.4021	**1.2288**	**0.3191**	**0.2874**
0.5	0.0352	1.205	0.9334	1.0092	0.9832	0.7527	**1.7151**	**0.3995**	**0.4420**

Table 6: Average RE of various methods over all sample sizes for case III at $\pi = \mathbf{0.1, 0.2, 0.3, 0.4, 0.5}$

π	C-LS	LS	LA	H-M	B-M	MM	LMS	LTS	S
0	100	100	65.49	96.47	94.31	94.02	**14.91**	**16.19**	**30.63**
0.1	100	5.78	5.89	6.31	14.51	41.87	**11.44**	**14.26**	**21.94**
0.2	100	3.24	3.37	3.53	4.11	22.43	**10.82**	**13.31**	**18.69**
0.3	100	2.33	2.65	2.69	2.78	12.38	**10.79**	**12.90**	**15.48**
0.4	100	2.08	2.5	2.48	2.49	5.79	**9.02**	**11.53**	**10.39**
0.5	100	1.9	2.52	2.33	2.36	3.34	**6.26**	**9.33**	**5.03**

Article DOI: 10.52589/AJMSS-EC8EWXUL
DOI URL: https://doi.org/10.52589/AJMSS-EC8EWXUL

DISCUSSION OF RESULTS AND CONCLUSION

The results presented cover mean squared errors (MSE) and relative efficiency (RE) of the parameter estimates for each estimation method with sample sizes n = 10, 30, 100, 200 and 1000 respectively. The number of replicates is 1000. For Case I presented in Table 1, H-M , B-M and MM work better than other estimates. Similar results were observed when the proportion of outlier is 0.2. Increasing the proportion of outlier in Table 1 reduces the efficiency of all the estimators with LS having much larger MSE than other robust estimators; H-M, B-M and MM have similar MSE to S. Tables 1 and 2 give the average results of all the estimators at various proportions of outlier. The tables serve as a standard check for assessing the breakdown point of the estimators. The empirical breakdown point of LS and LA is 0, H-M, B-M and MM is 0.3 while LTS, LMS and S is 0.5. The graphical plots in Fig 1 and 2 confirm the results.

Moving to Case II (x direction outlier), LS, LA, H-M, B-M and MM estimators could not withstand at most 0.1 proportion of outlier. The results presented in Tables 3 and 4 reveal that a single leverage outlier will significantly increase the MSE of the estimators. Although LTS, LMS and S still withstand up to 0.1 proportions of leverage outlier but their MSE are relatively far from the true scenario when there is no outlying point.

Tables 5 and 6 present the results of Case III (x,y direction outlier); the results are similar to Case II with LTS, LMS and S having high breakdown points but extremely low efficiency.

Based on the results the following conclusions were drawn:

1. For the y direction outlier, the best estimator in terms of high efficiency and breakdown point of at most 0.3 is MM,
2. For the x direction outlier, the best estimator in terms of high breakdown point of at most 0.4 is S,
3. For the x,y direction outlier, the best estimator in terms of high efficiency and breakdown point of at most 0.2 is MM.

REFERENCES

Coakley, C. W. and Hettmansperger, T. P. (1993), A Bounded Influence, High Breakdown, Efficient Regression Estimator. Journal of American Statistical Association, 88, 872-880.

Croux, C., Rousseeuw, P. J., and H¨ossjer O. (1994), Generalized S-estimators. Journal of American Statistical Association, 89, 1271-1281.

Donoho, D. L. and Huber, P. J. (1983), The Notation of Break-down Point, in A Festschrift for E. L. Lehmann, Wadsworth 18Freedman, W. L., Wilson, C. D., and Madore, B. F. (1991), New Cepheid Distances to Nearby Galaxies Based on BVRI CCD Photometry. Astrophysical Journal, 372,455-470.

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. Journal of the American Statistical Association, 96, 1348-1360.

Article DOI: 10.52589/AJMSS-EC8EWXUL
DOI URL: https://doi.org/10.52589/AJMSS-EC8EWXUL

Gervini, D. and Yohai, V. J. (2002), A Class of Robust and Fully Efficient Regression Estimators. The Annals of Statistics, 30, 583-616.

Handschin, E., Kohlas, J., Fiechter, A., and Schweppe, F. (1975), Bad Data Analysis for Power System State Estimation. IEEE Transactions on Power Apparatus and Systems, 2, 329-337.

Huber, P.J. (1981), Robust Statistics. New York: John Wiley and Sons.

Jackel, L.A. (1972), Estimating Regression Coefficients by Minimizing the Dispersion of the Residuals. Annals of Mathematical Statistics, 5, 1449-1458.

Lee, Y., MacEachern, S. N., and Jung, Y. (2011), Regularization of Case-Specific Parameters for Robustness and Efficiency. Submitted to the Statistical Science.

Mallows, C.L. (1975), On Some Topics in Robustness. unpublished memorandum, Bell Tel. Laboratories, Murray Hill.

Maronna, R. A., Martin, R. D. and Yohai, V. J. (2006), Robust Statistics. John Wiley.

Naranjo, J.D., Hettmansperger, T. P. (1994), Bounded Influence Rank Regression. Journal of the Royal Statistical Society B, 56, 209-220.

Rousseeuw, P.J.(1982), Least Median of Squares regression. Research Report No. 178, Centre for Statistics and Operations research, VUB Brussels.

Rousseeuw, P.J.(1983), Multivariate Estimation with High Breakdown Point. Research Report No. 192, Center for Statistics and Operations research, VUB Brussels.

Rousseeuw, P.J. and Croux, C.(1993), Alternatives to the Median Absolute Deviation. Journal of American Statistical Association, 94, 388-402.

Rousseeuw, P.J. and Yohai, V. J. (1984). Robust Regression by Means of S-estimators. She, Y. (2009). Thresholding-Based Iterative Selection Procedures for Model Selection and Shrinkage. Electronic Journal of Statistics, 3, 384C415.

She, Y. and Owen, A. B. (2011), Outlier Detection Using Nonconvex Penalized regression. Journal of American Statistical Association, 106, 626-639.

Siegel, A.F. (1982), Robust Regression Using Repeated Medians. Biometrika, 69, 242-244. Stromberg, A. J., Hawkins, D. M., and H¨ossjer, O. (2000), The Least Trimmed Differences

Regression Estimator and Alternatives. Journal of American Statistical Association, 95, 853-864.

Yohai, V. J. (1987), High Breakdown-point and High Efficiency Robust Estimates for Regression. The Annals of Statistics, 15, 642-656.

Article DOI: 10.52589/AJMSS-EC8EWXUL
DOI URL: https://doi.org/10.52589/AJMSS-EC8EWXUL