



OPEN

A redescending M-estimator approach for outlier-resilient modeling

Aamir Raza¹, Muhammad Noor-ul-Amin^{2✉}, Amel Ayari-Akkari³, Muhammad Nabi⁴ & Muhammad Usman Aslam⁵

The OLS model is built on the assumption of normality in the distribution of error terms. However, this assumption can be easily violated, especially when there are outliers in the data. A single outlier can disrupt the normality assumption of error terms, making the OLS model less effective. In such situations, M-estimators (MEs) come into play to obtain reliable estimates. We introduce a redescending M-estimators (RME) for robust regression to handle datasets with outliers. The proposed RME produces more robust estimates by effectively managing the influence of outliers, even at lower values of the tuning constant. We compared the performance of this estimator with existing RMEs using real-life data examples and an extensive simulation study. The results show that our suggested RME is more efficient than the compared ME in various situations.

Keywords Ordinary least squares, Outliers, Redescending, Robust regression

The ordinary least square (OLS) method has many useful applications in real life, but sometimes real data do not fulfill the conditions of OLS in terms of normality of error terms due to the presence of outliers. Even a single value of the outlier can disturb the performance of the least square estimates by producing inefficient and unreliable results. To overcome this drawback of OLS, researchers developed a technique called robust regression, which modifies the OLS in the presence of outliers. In heavy tailed distributions, the OLS estimates are badly influenced by the presence of outliers and MEs provide a possible alternative to the classical OLS estimates to compensate the sensitivity of the estimates towards the outliers. A large number of robust estimators are presented in the literature: M-estimators (Huber¹), least median square (Roisseeuw²), least trimmed square (Roisseeuw³), MM estimators (Yohai⁴) and S-estimators (Sakata and White⁵) are very famous robust estimators.

Huber¹ introduced the procedure of the ME to deal with the data having outliers. This estimator provided weights approximately equal to one for the central observations and zero weights for outliers. The maximum likelihood formulation is used in ME when the condition of normality in error terms is violated in OLS. In robust MEs, the symmetric loss function is used in place of the square error term in OLS, i.e.

$$\text{Minimize } \sum_{k=1}^n \rho(r_k) \quad (1)$$

where $\rho(\cdot)$ provides the role of error terms in the objective function. The technique of the ME is usually based on the iterative reweighted least squares technique (IRLS) to obtain the optimum estimates of objective function, Birch⁶. The objective function is chosen in such a way that it reduces the weight of outliers and hence produces maximum efficiency.

The psi function is obtained by differentiating Eq. (1) w.r.t. residual, the $\psi(r_k)$ function is given as.

$$\sum_{k=1}^n \psi(r_k) X_k = 0 \quad (2)$$

Weight function $w(r_k)$ is obtained by dividing $\psi(r_k)$ function by “r” residuals, given as

¹Govt. College Women University Sialkot, Sialkot, Pakistan. ²Lahore Campus, COMSATS University Islamabad, Lahore, Pakistan. ³Biology Department, College of Sciences in Abha, King Khalid University, P.O. Box 960, Abha, Saudi Arabia. ⁴Khost Mechanics Institute, Khost, Afghanistan. ⁵Xi'an Jiaotong University, Xi'an, China. ✉email: nooramain.stats@gmail.com

$$\sum_{k=1}^n w(r_k)X_k = 0 \quad (3)$$

If the deviation of $\rho(\cdot)$ is redescended, the ME is known as the RME. Many researchers have done useful work on ME; well-known names are Hampel (1986), Andrew⁷, Tukey⁸, Qadir⁹, Ali (2005), Insha¹⁰, Alamgir (2015), Khalil (2016), Noor-ul-Amin¹¹, Anekwe¹², Luo¹³, and Mukhtar¹⁴. These RME perform well to reduce the effects of outliers. We have also proposed a RME to deal with the data having outliers given in Eq. (13). The proposed estimator is a generalized RME, as the weights of outliers can be controlled by changing the value of the generalized tuning constant 'a'.

The design of the upcoming sections is as follows: Section "Redescending M-estimator" describes the existing RME; Section "Redescending M-estimator" discusses the proposed RME with its graphical presentations. In Section "Proposed redescending M-estimator", the performance of the proposed ME relative to existing ME is presented. In Section "Graphical comparison of the proposed RME with existing RME", practical applications and a simulation study of the proposed RME are discussed in Session 6. The concluding remarks of the proposed ME are cited in Section "Simulation study".

Redescending M-estimator

A detailed discussion about the existing ME is necessary to highlight the superiority of the proposed estimator. So we have done a detailed study on the existing ME in this section. The known MEs are given below.

Huber¹ proposed well known ME with ψ function defined as

$$\psi(r) = \begin{cases} r & |r| < k \\ k & |r| \geq k \end{cases} \quad (4)$$

where k is the tuning constant and r is the residual obtained using the OLS. The value of k is 1.345 to obtain 95% efficient result for the data having Gaussian distribution. The efficiency of Huber proposed model is found to be low when data contain larger residuals as its ME has not redescending characteristics.

Hampel (1986) introduced a piece wise ME with-function given by

$$\psi(r) = \begin{cases} r & |r| \leq k \\ k \cdot \text{sign}(r) & k < |r| \leq l \\ k \frac{m-|r|}{m-l} \text{sign}(r) & l < |r| \leq m \\ 0 & |r| > m \end{cases} \quad (5)$$

where k , l and m are constant and $0 < k \leq l < m < \infty$. Differentiability of $\psi(r)$ is not ideal here and a smooth $\psi(r)$ function should be preferred. This suggested robust estimator performed well for the Princeton Robustness Study. The Hampel psi function has lack of differentiability, and it uses three tuning constants, which are undesirable.

Andrew⁷ developed a very popular and commonly used robust ME called the Andrew sine function, ψ -function for Princeton Robustness Study is written as

$$\psi(r) = \begin{cases} k \sin(r/k) & |r| < k\pi \\ 0 & |r| > \pi k \end{cases} \quad (6)$$

where k is a tuning constant and the value of $k=3.2$ to obtain maximum efficiency. This function is smooth and differentiable.

Another very useful and popular redescending robust estimator is given by Tukey⁸ with ψ -function defined as

$$\psi(r) = \begin{cases} r \left[1 - \left(\frac{r}{k} \right)^2 \right]^2 & |r| < k \\ 0 & |r| > k \end{cases} \quad (7)$$

where k is the tuning constant. For Tukey's ME, value of $k=4.865$ to obtain 95% efficiency for the data have Gaussian distribution. Tukey's bi-weighted and Andrew's sine functions covers the drawbacks of Hampel ψ function to some extent, but they give less weight to some good observations.

Qadir⁹ developed another RME known as the Qadir Beta function. The ψ -function is given as

$$\psi(r) = \begin{cases} \frac{r}{16k^4} (k+r)^2 (k-r)^2 & |r| \leq k \\ 0 & |r| > k \end{cases} \quad (8)$$

where k is an arbitrary constant called the tuning constant, the value of k is 4.0 to attain maximum efficiency.

Ali (2005) also derived a modified Tukey's RME with the psi function.

$$\psi(r) = \begin{cases} \frac{2r}{3} \left[1 - \left(\frac{r}{k} \right)^4 \right]^2 & |r| \leq k \\ 0 & |r| > k \end{cases} \quad (9)$$

This ME gives most efficient results if $k=4$.

Insha¹⁰ developed a RME for the detection of outliers, whose ψ -function is given as

$$\psi(r) = \left\{ \left[1 + \left(\frac{r}{k} \right)^4 \right]^{-2} \mid r \geq 0 \right\} \quad (10)$$

Tuning constant $k=4$ provides the most efficient results for Insha's robust estimator. Alamgir (2015) suggested another RME for robust regression; his proposed psi-function is given as

$$\psi(r) = \begin{cases} \frac{16r \cdot e^{-2(r/k)^2}}{\left(1 + e^{-(r/k)^2}\right)^2} & |r| \leq k \\ 0 & |r| > k \end{cases} \quad (11)$$

Alamgir suggested $k=3$ to obtain 95% efficient estimates for the data having a normal distribution. In the series of RME developments, Khalil (2016) also suggested his RME, whose psi-function is given as

$$\psi(r) = \begin{cases} r \left(\frac{3}{2}\right) \left\{ 1 - \left(\frac{r}{k}\right)^4 \right\}^2 \sin \left[\left(\frac{2}{3}\right) \left\{ 1 - \left(\frac{r}{k}\right)^4 \right\}^2 \right] & |r| \leq k \\ 0 & 0 \end{cases} \quad (12)$$

where k is the tuning constant, and a value of $k=4$ is suggested by Khalil to obtain 95% efficiency in the normal case. The ME proposed by Ali, Alamgir, and Khalil rejected the observations completely with a larger residual. Insha provided an estimator that covers this drawback, but it lacks generalization. We have produced a RME that covers the drawbacks of previously proposed estimators and has redescending characteristics. If the outliers are completely rejected by the ME, then it is called a RME. Hence, a ME is redescending if the derivative of the objective function, i.e. $\rho(r_k)$ is redescending that satisfies $\rho'(r_k) = 0$.

Proposed redescending M-estimator

We proposed a new RME called Aamir's RME, which contains the properties of a redescending estimator for the detection of outliers in robust regression. The proposed RME is more robust even at low value of tuning constant, it produces more efficient and reliable estimate even at small sample sizes, the mathematical formulation of the suggested RME is simple and easy to apply in real life situations, the proposed RME is generalized ME as we can adjust the weights of extreme values of residuals by changing the value of the generalized tuning constant 'a'. The psi function of proposed estimator is more linear than the existing RMEs. The proposed estimator is more flexible to control the weights of outliers. The validity and usefulness of the proposed estimator are confirmed by extensive simulation analysis. The characteristics and shape of $\rho(r_i)$ function, $\psi(r_i)$ and $w(r_i)$ weight function of proposed RME are discussed below.

The objective function of proposed ME is given as

$$\rho(r_i) = \left\{ \frac{k^2}{2a} \left[1 - \left\{ 1 + \left(\frac{r}{k} \right)^2 \right\}^{-a} \right] \mid r \geq 0 \right\} \quad (13)$$

where k is any arbitrary tuning constant and 'a' is a generalized tuning constant, and we have taken values of 'a' is 6 and 8 for our present study, and 'r' is the residual associated with i^{th} observation. The proposed objective function fulfills the following properties of $\rho(\cdot)$ function.

- (i) $\rho(r_i) \geq 0$
- (ii) $\rho(0) = 0$
- (iii) $\rho(-r_i) = \rho(r_i)$
- (iv) $\rho(r_i) \geq \rho(r_j)$ for $|r_i| \geq |r_j|$, i.e. $\rho(\cdot)$ is an increasing function
- (v) $\rho(\cdot)$ is continuous and differentiable

To demonstrate the performance of the proposed objective function, a sequence of residuals is generated using the R program. Utilizing the generated data, the graph of the objective function given in Eq. (13) is constructed and presented in Fig. 1.

Figure 1, shows that the proposed function is a positive, symmetrical, differentiable, and increasing function which satisfies the conditions of the objection function for robust regression.

Differentiating $\rho(\cdot)$ function with respect to residual, we obtained ψ -function, which is given below

$$\psi(r_i) = \left\{ r \left[1 + \left(\frac{r}{k} \right)^2 \right]^{-a-1} \mid r \geq 0 \right\} \quad (14)$$

Data used in Fig. 1, the graph of psi function is constructed and shown in Fig. 2.

Figure 2 shows that the proposed function is differentiable and more linear at the center than the other existing ψ -function of robust estimators. The proposed ψ -function follows the pattern needed to construct the RME. Moreover, the proposed ψ -function provides more weightage to the central values and vice versa.

Dividing the ψ -function by residual "r", we obtain the corresponding weight function, which is given as

$$w(r_i) = \left\{ \left[1 + \left(\frac{r}{k} \right)^2 \right]^{-a-1} \mid r \geq 0 \right\} \quad (15)$$

Graph of weight function using the data in Fig. 1, is represented in Fig. 3.

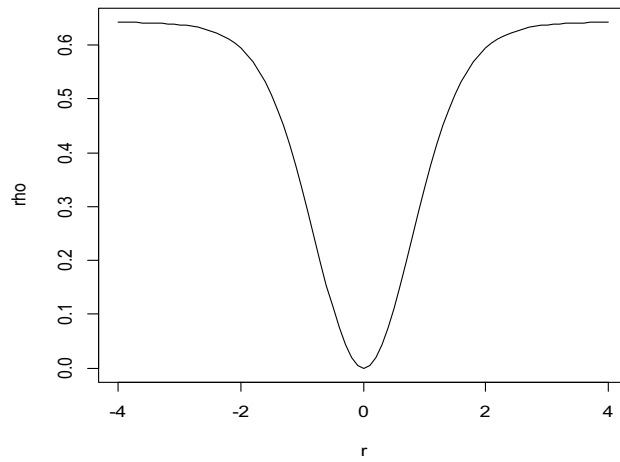


Figure 1. The Graph of $\rho(r_i)$ Function of Proposed RME w.r.t. Residual (r).

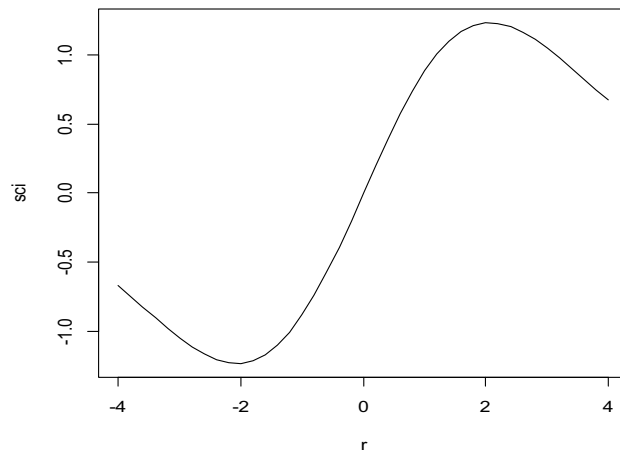


Figure 2. Graph of ψ -function for proposed RME w.r.t. residual (r).

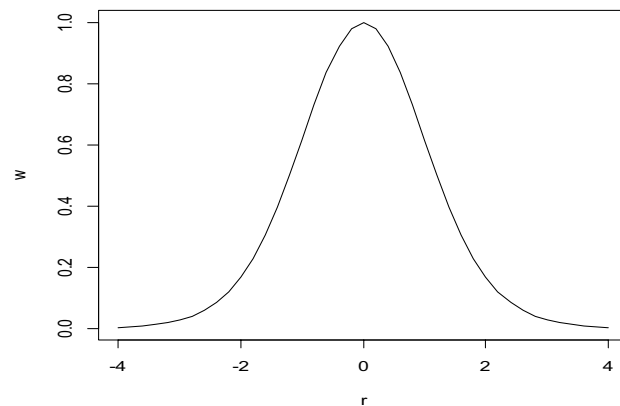


Figure 3. Graph of Weight Function of Proposed RME w.r.t. Residual (r).

Figure 3, shows that the proposed weight function has the unique property of robustness with a generalized tuning constant ' a '. By increasing the value of ' a ', we can obtain more redescending estimates. We can control the weights of outliers by selecting the value of ' a '. It is analyzed that if the value of the generalized constant is higher, the proposed weight function provides lower the weights of the outliers, and vice versa.

Graphical comparison of the proposed RME with existing RME

The comparison between the existing RME and the proposed estimator is done by using a graphical representation of ψ -functions. The proposed function is continuous differentiable at every point, and its ψ -function has more linearity at the center of the data than the other existing redescending ψ -functions which enhances the efficiency of suggested ME. The graphical comparison of existing and proposed ψ -functions is represented in Fig. 4.

Figure 4 indicates that the suggested psi function is much more linear at the center and more continuous than all previously suggested RME. For example, psi function given by Asad Ali gives a weight of approximately one to the central values, but weighted values rapidly tend to zero for the larger residuals. The process of weighting the outliers approaches zero more rapidly for the estimators proposed by Tukey, Insha, and Qadir, respectively, as compared to the proposed psi function. Hence, the proposed RME is more useful and efficient than all existing ME.

Practical applications

Two real-life data examples and extensive simulation studies are included in this section to demonstrate the performance of the proposed ME. The proposed RME is found to be more efficient than considered estimators.

Example-1: number of phone calls from Belgium 1950–1973

The first example is taken from the article by Rousseeuw³, which represents the year-wise number of international calls (in tens of millions) from Belgium 1950–1973 (Belgium Statistical Survey) given in Table 1. This data has some outliers in the response variable. The years represent the independent variable X, and the numbers of annual calls represent the dependent variable Y. Some of the authors, Qadir⁹, Ali (2005), and Khalil (2016), have used this in their research. The performance of the proposed RME with the considered estimators is presented in Table 2.

Table 2 depicts the performance of OLS and other robust estimators along with our proposed robust estimator in terms of estimates of regression coefficient and sum of squares of errors (SSE). The OLS estimates are very poor and misleading everywhere within the data due to the presence of outliers with the highest SSE, i.e., 659.44. All the remaining ME perform efficiently to minimize the effect of outliers, but our proposed robust estimator

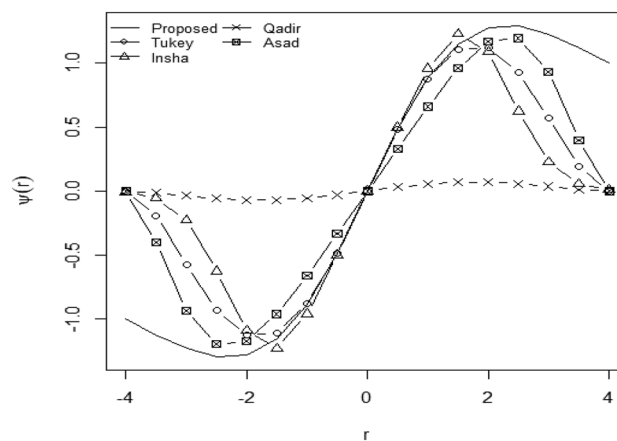


Figure 4. Graphical Comparison of Proposed ψ -Function with Well-known ψ -Functions w.r.t. Residual (r).

Years	Calls	Years	Calls
1950	0.44	1962	1.61
1951	0.47	1963	2.12
1952	0.47	1964	11.9
1953	0.59	1965	12.4
1954	0.66	1966	14.2
1955	0.73	1967	15.9
1956	0.81	1968	18.2
1957	0.88	1969	21.2
1958	1.06	1970	4.3
1959	1.2	1971	2.4
1960	1.35	1972	2.7
1961	1.49	1973	2.9

Table 1. Data Set of Number of International Calls (in tens of millions) from Belgium.

Method	Coefficients			
	Constant	X	Values used	SSE
OLS	-26.01	0.504	24	695.44
LMS	-5.164	0.108	16	0.1313
Andrew (1.5)	-4.9072	0.1033	16	0.17308
Tukey (3.8)	-5.2439	0.1102	16	0.13665
Qadir (1.0)	-5.505	0.1122	16	0.36726
Ali (3.0)	-5.2092	0.1094	16	0.13404
Insha (4)	-5.243	0.1102	16	0.1406
Alamgir (3.0)	-5.2454	0.1102	16	0.14127
Khalil (4.0)	-5.2343	0.1099	16	0.13863
Proposed (a=6, k=3.5)	-5.5109	0.1094	16	0.13256

Table 2. Fitting OLS, Existing M-Robust Methods and Proposed RME on Telephone Call Data.

has the least SSE among all except the LMS estimator. So it is analyzed that the suggested estimator outfits the model by giving the least SSE, like the LMS method.

Example-2: annual average price growth rate in China 1940 to 1948

The data used in this example is taken from Rousseeuw³. Insha¹⁰ also used this data in his research work on robust regression. This data consists of nine values of the annual average growth rates of prices from 1940 to 1948. The response variable is represented by the annual average growth rate, where years represent the predictor. The growth rates are 1.62, 1.63, 1.90, 2.64, 2.05, 2.13, 1.94, 15.50, and 364.00. It means growth rates increased by 1.62% in 1940 as compared with 1939, 1.63% in 1941 as compared with 1940, and so on. An exponential growth in prices was seen in 1948 due to a budget deficit, the war, and massive government expenses. The last two values of the response variable are outliers, as they are very different from other values. The performance of the proposed RME with considered estimators is presented in Table 3.

The OLS method is applied along with other robust estimators, including the proposed robust method, on the average growth rate increase in China since 1940–1948 to check the relative performance of the proposed method with existing robust methods. We have calculated the regression coefficients and SSEs for all methods in Table 3 using R-program. A critical analysis of the results obtained shows that the performance of OLS is very poor and goes astray throughout the data. It also depicts the sensitivity of OLS towards the outliers, and the SSE is highest for the OLS method, which is 78,532.88. All remaining robust estimators perform well to tackle the effects of outliers, but among all methods, our proposed robust estimator has a minimum SSE. This means our proposed estimator is superior among all the estimators to deal with the data having outliers. One can also use real life data set included by Yasin (2021) from the Economic Survey of Pakistan 2017–18.

Simulation study

The validation and reliability of the proposed ME are done by using an extensive simulation study. The efficiency of the proposed RME is compared with well-known RME. We have used the following linear model for simulation

$$y_i = \alpha + \beta x_i + e_i$$

Method	Coefficients			
	Constant	X	Values used	SSE
OLS	-1049	24.85	9	78,532.88
LMS	-2.47	0.102	7	0.69534
Andrew (0.58)	-2.7542	0.10898	7	0.616687
Tukey (2)	-2.7535	0.10896	7	0.616652
Qadir (1.0)	-2.7779	0.1095	7	0.617785
Ali (3.0)	-2.7792	0.10956	7	0.617851
Insha (1.5)	-2.6392	0.10618	7	0.615962
Alamgir (3.0)	2.7853	0.1097	7	0.61815
Khalil (2.0)	-2.7851	0.1097	7	0.618139
Proposed (a=8, k=2)	-2.6486	0.10656	7	0.611112

Table 3. Fitting OLS, Existing M-Robust Methods and Proposed RME on Annul Average Price Growth Rate in China 1940–1948.

Method	Case-I		Case-II	
	a	b	a	b
OLS	1.99842	1.00002	7.0426	0.99794
Hampel (4)	1.99742	1.00008	1.99586	1.00018
Tukey (4.69)	1.99742	1.00011	1.99465	1.00023
Andrew(3.2)	1.99801	1.00021	1.99484	1.00045
Huber (4)	1.99834	1.00002	2.1976	0.99764
Insha (4)	2.00924	0.99964	1.99743	1.00004
Ali (4)	2.00303	0.99984	1.99923	1.00011
Khalil (5)	2.02813	0.99846	2.00358	0.999761
Proposed (2)	1.99736	1.00009	1.99331	1.00047

Table 4. Coefficient Estimates using Different RMEs for $n = 100$.

where $\alpha = 2, \beta = 1, X_i \sim N(20, 10)$ and $e_i \sim N(0, 1)$. We have generated a population of 10,000 values by using the above model. A sample of 100 values is chosen from the population using the R-Program, and estimates of parameters are obtained using the proposed and considered MEs, results are shown in Table 4.

We have studied two cases; in the first case, we estimated parameters by using data with no outliers. In the second case, parameters are estimated for the data having 10% contamination of the observations as outliers in the y-direction. The ME unfortunately gives poor results when outliers are in the X direction (Norazan,¹⁵). The results are obtained by averaging the 50,000 iterations of simulated data using a sample size of 100. This table shows that for case 1, all methods performed equally well when the data has no outlier. But for case II, the estimates provided by the proposed ME are very close to the values of the parameters from which the simulation was carried out. The results obtained from the considered estimator are also efficient, except for the OLS method. The OLS method again failed to provide reliable estimates when the data had outliers, which validated the results obtained in Section “Graphical comparison of the proposed RME with existing RME”. The proposed ME is also efficient for small sample sizes and can be used to save money, which is a major purpose of sampling and estimation.

Conclusion

The results obtained in previous sections showed that the proposed RME is more general and efficient than the considered RME. The proposed ME showed that its behavior is exactly similar to other well-known RMEs, and the ψ -function given by the ME is more continuous before it redescends. The proposed estimator is very simple, more general, and flexible, and it converges very quickly as compared to previous MEs. The real-life data applications showed that the proposed RME is more efficient and has minimum SSE in the presence of outliers. It is also revealed from simulation studies that the values of coefficients obtained from the proposed robust estimator are very close to parameters and very similar to those of famous robust estimators such as Huber, Hampel, Andrew, and Tukey. The proposed estimator can also be used to estimate the population mean using different sampling techniques for the data with outliers.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request. Further, no experiments on humans and/or the use of human tissue samples involved in this study.

Received: 19 January 2024; Accepted: 22 March 2024

Published online: 26 March 2024

References

- Huber, P. J. Robust estimation of a location parameter. *Ann. Math. Stat.* **35**(1), 73–101 (1964).
- Rousseeuw, P. J. Least median of squares regression. *J. Am. Stat. Assoc.* **79**(388), 871–880 (1984).
- Rousseeuw, P. J. & Leroy, A. M. *Robust Regression and Outlier Detection* (John Wiley & Sons, New York, 1987).
- Yohai, V. J. High breakdown-point and high efficiency robust estimates for regression. *Annals Stat.* **15**, 642–656 (1987).
- Sakata, S. & White, H. S-estimation of nonlinear regression models with dependent and heterogeneous observations. *J. Econom.* **103**(1–2), 5–72 (2001).
- Birch, J. B. Exploratory and robust data analysis. (Unpublished lecture notes; 1965).
- Andrews, D. F. A robust method for multiple linear regression. *Technometrics* **16**(4), 523–531 (1974).
- Beaton, A. E. & Tukey, J. W. The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. *Technometrics* **16**(2), 147–185 (1974).
- Qadir, M. F. Robust method for detection of single and multiple outliers. *Sci. Khyber* **9**(2), 135–144 (1996).
- Ullah, I., Qadir, M. F. & Ali, A. Insha's RME for robust regression: A comparative study. *Pak. J. Stat. Oper. Res.* **1**, 135–144 (2006).
- Noor-Ul-Amin, M., Asghar, S. U., Sanaullah, A. & Shehzad, M. A. RME for robust regression. *J. Reliab. Stat. Stud.* **5**, 69–80 (2018).
- Anekwe, S. & Onyeagu, S. The Redescending M estimator for detection and deletion of outliers in regression analysis. *Pak. J. Stat. Oper. Res.* **2**, 997–1014 (2021).
- Luo, R., Chen, Y. & Song, S. On the ME under third moment condition. *Mathematics* <https://doi.org/10.3390/math10101713> (2022).

14. Mukhtar, M., Ali, N. & Shahzad, U. An improved regression type mean estimator using RME. *UW J. Sci. Technol.* **7**(1), 11–18 (2023).
15. Norazan, M. R. Weighted maximum median likelihood estimation for parameters in multiple regression model. Unpublished PhD. Thesis, University Putra Malaysia, Malaysia; (2008).
16. Alamgir, A. A., Khan, S. A., Khan, D. M. & Khalil, U. A new efficient RME: Alamgir RME. *Res. J. Recent. Sci.* **2**(8), 79–91 (2015).
17. Ali, A. & Qadir, M. F. A modified ME for the detection of outliers. *Pak. J. Stat. Oper. Res.* **1**, 49–64 (2005).
18. Andrews, D. F., Hampel, F. R. Robust estimates of location: Survey and advances. (Princeton University Press; 2015).
19. Khalil, U., Ali, A., Khan, D. M., Khan, S. A. & Qadir, F. Efficient UK's RME for robust regression. *Pak. J. Stat.* **32**(2), 125–138 (2016).
20. Yasin, S., Kamal, S. & Suhail, M. Performance of some new ridge parameters in two-parameter ridge regression model. *Iran. J. Sci. Technol. Transact. A: Sci.* **45**(1), 327–341 (2021).
21. Hampel, F. R., Ronchetti, E.M., Rousseeuw, P.J. & Stahel, W.A. Robust Statistics: The Approach Based on Influence Functions (John Wiley & Sons, New York, 1986).

Acknowledgements

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University, Saudi Arabia for funding this work through Large Groups Project under grant number R.G.P2/186/44.

Author contributions

A.R. and M.N.A. conceived and designed the research. A.R. and M.N.A. conducted the literature review and theoretical framework development. A.A.-A. refining the mathematical formulations and provided critical insights into the methodology. M.N. contributed to performed the statistical analysis. M.U.A. contributed to the interpretation of the results and provided expertise on outlier-resilient modeling techniques. All authors actively participated in drafting the manuscript and revising it critically for important intellectual content.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.N.-u.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024