Punjab University
Journal of Mathematics (ISSN 1016-2526)
Vol. 51(10)(2019) pp. 1-14

A New Modification of the Least Squares Method with Real Life Applications

Zahid Khan[a, 1], Katrina Lane Krebs[2], Sarfraz Ahmad[3], Aamir Saghir[4], Serpil Gumusteki[5]
[1]Department of Mathematics & Statisitcs, Hazara University Mansehra Pakistan
[a,1] zahidkhan@hu.edu.pk
[2]Higher Education Division, Central Queens Lands (CQ) University, Rockhamton, Australia
[3]Department of Mathematics, Abbottabad UST, Pakistan
[4]Department of Mathematics, Mirpur University of Science and Technology, Pakistan
[5]Department of Statistics, Ondokuz Mayis University, Samsun, Turkey

Received: 06 November, 2018 / Accepted: 10 May, 2019 / Published online: 01 September, 2019

Abstract. A new regression M-estimator namely modified least squares (MLS) in the class of M-estimators is presented in this study. The proposed estimator overcomes the non-robustness property associated with traditional approach of the least square (LS) estimator. The effectiveness of the loss function used for proposed estimator has been compared with that of commonly implemented approach of the LS estimator. The influence and weight functions have been derived to analyze the robustness of the proposed estimator against the polluted measurements. Real data examples in statistical applications have been used to analyze the effectiveness of proposed estimator. The empirical results from real applications also confirm that MLS estimator substantially enhances the non-robustness property of the LS estimator.

AMS (MOS) Subject Classification Codes: 62F35; 93E10; 62J05
Key Words: Least squares estimator, gross errors, robust estimator, M-estimator.

1. Introduction

Regression analysis is one of the primary statistical techniques which is currently being applied extensively across many discipline areas for data analysis. The method of the LS is a root for most of the regression analysis used for estimating the model parameters [11, 12]. The procedure of the LS in many senses is an optimal under specific assumptions of error terms having a Gaussian distribution in a regression model [13].These assumptions are however, rarely satisfied in full when applying this analysis technique in real world situations but, they are usually taken as a best approximations to reality. The results for many non-Gaussians situations using the LS approach are markedly outlying from optimal and proving inefficiency specially, in the presence of contaminated values in the data [2]. These contaminated observations can be added from different sources such as measuring equipment failure, improper calibration, mistakes in observing data and so on [14].The occurrence of corrupted measurements obviously depends upon the quality of data but Huber [9] pointed out the prevalence of the corrupted measurements even in data which are collected in a highly careful manner. Hampel [6] emphasized that 5-10% gross errors in high quality samples (such as in geodesy and astronomy) comprising of thousands of measurements is not an exception rather, seem to be the rule. These errors are not always detectable so one initially cleans up the measurements, making it unlikely to fit a perfect model; robust methods should be devised. Statisticians in midst of the 20th century have become acutely aware about the fact that commonly used statistical methods specifically those involving normality assumptions, are not being very insensitive to smaller deviations from the assumptions taken [3]. The reaction of this probably resulted in robust procedures. Many classes of robust methods have been developed so far, however, details of some of commonly used techniques in context of regression analysis can be seen in [15, 16]. One of the important classes of robust estimators which are rooted on the principle of maximum likelihood approach are known as M-estimators. Such estimators can be primarily embodied in two categories, Hubers and Hampel estimators. The effect of outlying values is reduced by the former whereas the latter are used to nullify their effect. This effect of M-estimators against the offset values are usually evaluated in terms of their influence functions [1]. While larger errors are basically not rejected by Huber estimators however, their influence is bounded whereas the influence function for a Hampel estimators becomes zero for distorted values [17].

This work is concerned with development of a new regression M-estimator which is sufficiently to address the gross errors commonly occurring in data used for regression analysis. This new estimator has a bounded influence and yields estimates for unknown regression parameters as excellent as those that resulted by traditional approach of the LS when no contaminated value is present in data. The proposed MLS estimator was tested on variety of examples taken from relevant literature on robust regression applications.

The reminder of this work is established in the following sections as these lines: Section 2 includes the methodology of the M-estimators. The proposed estimator of MLS approach is established with its properties in Section 3. Section 4 presents a relative performance of proposed estimator with LS estimator and other conventional estimators using examples from relevant literature. Section 5 contains the application of MLS estimator in real world problem of SE of power system. Eventually, the summary of the work is highlighted in Section 6.

2. M-ESTIMATORS

The innovative idea of M-estimators in context of regression parameters was first introduced by Huber [8]. These estimators utilize the approach of ML in formulations providing optimal weights to data values under non-normal environments. The approach of M-estimator is actually a generalization of most commonly used criterion of the LS by substituting the quadratic objective function; with the following less rapidly growing function $\rho\left(\widehat{\varepsilon}_i\right)$ of residuals:

$$\min_{\widehat{\beta}} \sum_{i=1}^{n} \rho\left(\widehat{\varepsilon}_i\right) \tag{2. 1}$$

where the chosen function $\rho\left(\widehat{\varepsilon}_i\right)$ holds at least the properties of symmetric, non-negative, and increasing monotonically in either direction of $\widehat{\varepsilon}_i$ and having unique minimum at zero. In respect to other regression robust estimators, the approach of M-estimators is theoretically and computationally simplified [18]. Despite of non-robustness of M-estimators against the leverage points, they are remained extensively employed in regression applications where it assumes that corrupted measurements are mainly in direction of response variable. Obviously, selection of an appropriate form of objective function should be found on information about the actual errors behaviour which is generally not known in advance. Thereby, the function is chosen instead with the view of how the resulting approach reduces the effects of larger residuals by assigning less weight to them. The workout of this approach can be helpful to detect the outliers which can be identified as values acquiring zero weights in the estimation procedure [5]. Instead of minimizing (2. 1) with respect to unknown parameter β, the following expression is usually used due to scale invariant issue of the M-estimators [10].

$$\min_{\widehat{\beta}} \sum_{i=1}^{n} \rho\left(\frac{\widehat{\varepsilon}_i}{\sigma}\right) \tag{2. 2}$$

The unknown value of sigma is typically substituted by one of its most extensively used preliminary estimates namely; mean absolute value deviation (MAD) which can be expressed as below:

$$\widehat{\sigma} = \frac{\text{median}\left|\widehat{\varepsilon}_i\right|}{0.674} \tag{2. 3}$$

This estimator of σ is highly robust to contaminated values with a break down point of 50% [10]. Solving (2. 2) for unknown values of β is weighted least squares (WLS) problem with an iterative scheme because here, the weights depend upon estimated residuals, the residuals are found on the coefficient estimated and the coefficient estimation relies on the weights. In order to compute the M-estimators, different iterative techniques are available but iterative re-weighted least square (IRLS) approach is more numerical stable and efficient technique [4].

3. The proposed M-estimator

In this section the proposal of the new objective function $\rho\left(\widehat{\varepsilon}_i\right)$ with its associated influence and weight functions are discussed. The proposed objective function belongs to the group of the soft re-descending M-estimators according to classification of these estimators on the basis of corresponding behaviour of their influence functions.
Beginning with the following general regression model:

$$y_i = f\left(x_{ij}, \theta\right) + \varepsilon_i, \quad i = 1,\ 2, \ldots m; \quad j = 1, 2, 3 \ldots k \tag{3. 4}$$

where $y_1,\ y_2,\ \ldots, y_m$ be a sample of m measurements with k response variables in defined model, the vector of the unknown parameters $\theta = [\theta_1, \theta_2, \ldots \theta_p]$ and ε_i represents a Gaussian random term.

To find the unknown values of regression parameters, the following estimator in the most familiar criterion of the LS approach is employed in the optimization process:

$$\rho\left(\varepsilon_i\right) = \frac{1}{2}\varepsilon_i^2, \tag{3. 5}$$

where $\varepsilon_i = y_i - \ f\left(x_{ij}, \theta\right)$.

The larger deviations of the residuals can distort the results obtained from the expression (3. 5). This justifies the non-robustness of the LS criterion against outlying observations. The influence of these troublesome values on estimator performance can be measured using its influence function. Thereby a way of observing the effect of outliers on estimation results obtained by the LS estimator, can be seen by the following its influence function:

$$\Psi\left(\varepsilon_i\right) = \varepsilon_i. \tag{3. 6}$$

Equation (3. 6) suffices to conclude that the influence function of the LS criterion is not bound and the effect of error is proportionally related to estimated quantity. The consequence of this non-robustness signifies that a single outlier value can have an overriding impact on estimation results. In contrast to the LS, our suggested estimator based on the modified form, utilizes the following function in optimization criteria:

$$\rho\left(\varepsilon_i\right) = \frac{\varepsilon_i^2}{\sqrt{4 + \alpha\varepsilon_i^2}}, \quad \alpha \geq 0 \tag{3. 7}$$

where α signifies the tune constant. The function defined in (3. 7) is referred as the modified least squares (MLS) estimator and satisfies the same properties as mentioned above for the function defined in (2. 1). The MLS estimator is constructed by robustizing the LS estimator in such a way, that the contaminated values cannot have significant influence on the estimated quantities. This tolerance of MLS estimator is achieved by adding the descending term $\alpha\varepsilon_i^2$ to diminish the influence of the larger residuals on the estimated parameters and its inclusion robustizes the LS approach. It should be noticed that zero value of the tune constant provides the equivalent results as those obtain from the routinely LS estimator. Likewise, the other existing M-estimators, the proposed estimator is resistant to outlier in Y observations and highly sensitivity to leverage points in x-direction. Thus, the MLS estimator holds the same property of breakdown point equal to $1/n$, where n is number of observations in the sample. In order to assess the breakdown property of the proposed estimator for outlying in y-direction, the following sensitivity curve (SC) for the proposed estimator has been computed for the chem data:

$$SC = n\left[T_n\left(X_1, X_2 \ldots X_{n-1}, Y\right) - T_{n-1}\left(X_1, X_2 \ldots X_{n-1}\right)\right]$$

where $T_n\left(X\right)$ denotes the value of the estimator based on data with contaminated observation Y. To view the effect of arbitrarily changes in Y on the proposed estimator, an outlier has been introduced in chem data that varies within the range $\{-10 < Y < 10\}$. The chem data set can be accessed from the MASS package in R. Thus, the sensitivity of curves of the proposed estimator for the location model can be viewed below in Figure 1.

In Figure 1, the sensitivity curves of the MLS estimator have been established for the tuning constant equal to 0 and 0.75. Notice that, the zero of the tuning constant in the MLS estimator provides result equal to the LS estimator. Thus, the proposed estimator treats the offset values differently with different values of the tuning constant. Figure 1, shows that

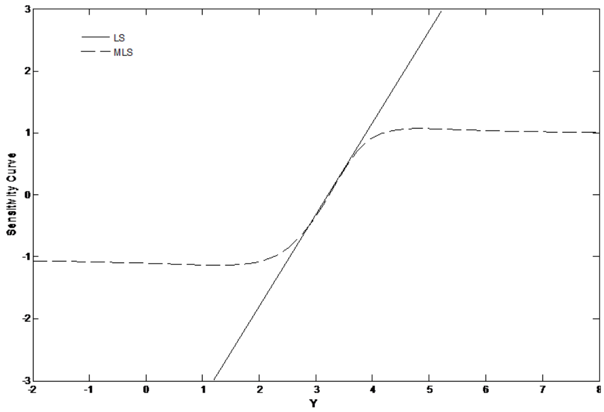

FIGURE 1. Sensitivity curve of the MLS and LS estimators

arbitrarily changes to Y do not alternate the location estimate of the proposed estimator and thus indicates its bounded influence. Whereas, for the LS estimator this effect can be considered as large by large changes to Y.
The objective and influence functions of the MLS are viewed in the following equations (3. 8) and (3. 9) respectively:

$$\Psi\left(\varepsilon_i\right)=\frac{8\varepsilon_i+\alpha\varepsilon_i^3}{\left(4+\alpha\varepsilon_i^2\right)^{\frac{3}{2}}}\qquad\forall\,\varepsilon_i\in R \tag{3. 8}$$

$$w\left(\varepsilon_i\right)=\frac{8+\alpha\varepsilon_i^2}{\left(4+\alpha\varepsilon_i^2\right)^{\frac{3}{2}}}\qquad\forall\,\varepsilon_i\in R \tag{3. 9}$$

For the different values of tune constant, graphically comparisons between the MLS approach and to that the LS are depicted in the following respective graphs of the loss, weight and influence functions.
Figure 2, describes that the large value of tune constant yields more reduction and less rapid increase of the objective function, indicating that the resultant estimator becomes progressively robust to outlying observations. Figure 3 illustrates that the behaviour of the MLS estimator is similar to the LS for the smaller errors whereas for moderately to larger errors, the effect gradually decreases and subsequently approaches to a constant quantity which confirms the bound influence of the MLS estimator, i.e.,

$$\Psi\left(\varepsilon_i\right)\rightarrow\frac{1}{\sqrt{\alpha}}\ \text{ as }\ \left|\varepsilon_i\right|\rightarrow\infty$$

Figure 4, describes that the weight pattern assigned by the MLS estimator during the estimation process. This weighting behaviour illustrates that reduced weights are attached to outlier values and thus, the estimated quantities keep on less affected by such measurement errors.
The proposed estimator follows the asymptotic normal distribution with variance equal to $\gamma\sigma^2$, where γ is the correction factor. The asymptotic variance-covariance matrix can be

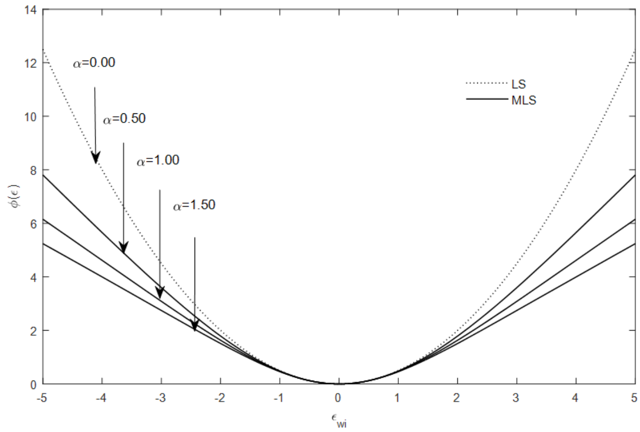

FIGURE 2. The objective function of the MLS estimator

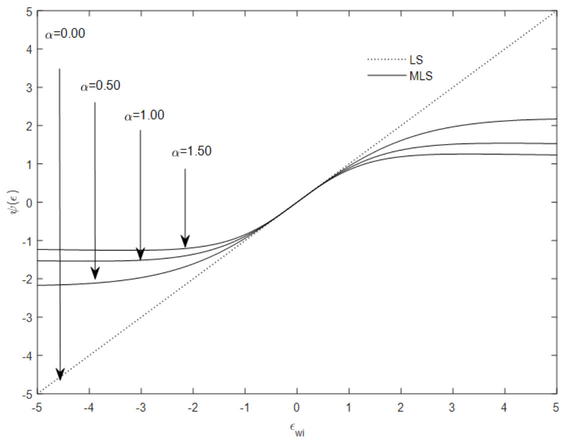

FIGURE 3. The influence function of the MLS estimator

estimated from the observed data as follows:

$$\hat{V} = \frac{1/n \sum_{i=1}^{n} \Psi\left(\widehat{\varepsilon}_i\right)^2}{\left(1/n \sum_{i=1}^{n} \Psi\left(\acute{\widehat{\varepsilon}_i}\right)\right)^2} \left(\acute{X}X\right)^{-1},$$

where $\Psi\left(\left(\widehat{\varepsilon}_i\right)\right)$ for the proposed estimator is based on the scaled residuals in computational algorithm.

The tune value α is directly connected with the efficiency performance of the proposed estimator. The higher value of tuning constant increases down weighting and make the estimator more robustized but at the cost of efficiency loss. Thereby, we cannot assume every value of tune parameter from its defined domain due to the tradeoff between efficiency

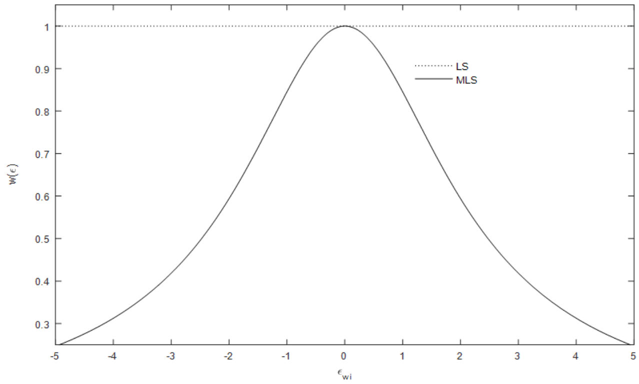

FIGURE 4. The weight function of the MLS estimator

and robustness. In general the following formula is usually applied to find the asymptotic efficiency of a M-estimator [19]:

$$E_{ff} = \frac{\left[E\left(\Psi'_{\varepsilon}\left(Z\right)\right)\right]^2}{E\left[\Psi^2_{\varepsilon}\left(Z\right)\right]}, \tag{3. 10}$$

where $\Psi_{\varepsilon}\left(Z\right)$ signifies the influence function at specific value of tuning constant α and $\Psi'_{\varepsilon}\left(Z\right)$ denotes its derivate under an ideal model of standard normal distribution. The efficiency solution using (3. 10) for some M-estimators is somewhat straightforward while the others involve numerical integration. The efficiency of MLS estimator according to (3. 10) at any specific value of the tune constant α can be calculated and its computation also necessitates the numerical integration. The 95% asymptotic efficiency of the MLS estimator can be achieved by letting the tune constant equal to 0.75. In addition, the asymptotic relative efficiency of the MLS estimator for other values of the tune constant is also computed which is displayed in Table 1.

TABLE 1. The relative efficiency computation of the MLS estimator at different values α

α	0.25	0.5	0.75	1.5	2
Efficiency	99%	97%	95%	92%	90%

The procedure for finding the unknown vector of interest $\theta = [\theta_1, \theta_2, \ldots \theta_p]$ is rooted on the IRLS approach which can be outlined below:
1: Start with initial estimates $\widehat{\theta}$
2: Obtain the initial residuals $\widehat{\varepsilon}_i = \frac{\left(y_i - x_i^T \widehat{\theta}\right)}{\widehat{\sigma}}$, $i = 1, 2, ..m$ where $\widehat{\sigma}$ is scale estimate as defined in (2. 3)

3: Construct the preliminary weights $w(\varepsilon_i) = \frac{8+\alpha\varepsilon_i^{*3}}{\left(4+\alpha\varepsilon_i^{*2}\right)^{\frac{3}{2}}}$ where $\varepsilon_i^* = \frac{\widehat{\varepsilon}_i}{\alpha*\sqrt{1-h_{ii}}}$ is standardized residual and h_{ii} is corresponding i^{th} leverage point from the vector of the LS leverages.
4: Attain the updated values of $\widehat{\theta}$ by applying MLS approach with weights as defined in step 3.
5: Iterate up to convergence.
The aforementioned algorithm starts with initial estimates of $\widehat{\theta}$ which can be assumed from the LS fit. Moreover, the termination tolerance is defined on the $\widehat{\theta}$ which means that iterations will continue until the norm of change of estimated coefficients is less than a fixed small positive value τ which may be chosen as e^{-8} i.e., $\left\|\widehat{\theta}^i - \widehat{\theta}^{i+1}\right\| \leq e^{-8}$.

4. Simulations

This section demonstrates the effectiveness of the MLS estimator through Monte Carlo simulation results. For better understanding the performance, several other useful estimators including Andrews (hard redescender), Welsch and Cauchy (soft redescender) and Huber (Monotone) are also readily considered. Each estimator has its own tune constant related to the relative efficiency (RE) of the corresponding estimator under the ideal model of the normal distribution. Therefore, in the comparison procedure, this tune constant value is necessary to be adopted in such a way that one desires to expect almost equal efficiency performance. This means that the Andrews estimator with 90% (for example) efficiency can be rigorously compared with the same efficiency of other chosen estimator. For this purpose, the corresponding tune constant values have been selected for each estimator which provides an equivalent efficiency of approximately 95%. The following independent linear model with known parameters $(\beta_0,\ \beta_1)$ have been used in the simulation procedure:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \tag{4. 11}$$

where $\beta_0 = 2$, $\beta_1 = 5$, $\varepsilon_i \sim N(0,\ 1\ )$ and the independent variable x_i is considered to be normally distributed with the parameters of zero mean and unit variance.
Across the various sample sizes, three cases were considered. In case-I, it had been assumed that the data set was not contaminated and the error terms were exactly normally distributed. This case is considered because the proposed estimator in the absence of outlying values provides results more or less similar to those as obtained by the LS estimator. The case- II had been considered to validate the robustness of the proposed estimator by introducing 25% contamination in the y-direction. To this end, 25% of the values in the y vector were replaced with influential values y_i, which followed $N(100,\ 400)$. In the same line 25% of the usual values in the x-direction were also included considering the same contaminated values distribution and the results were reported under case-III. The M-estimators are sensitive to bad leverage points, however; case-III had been retained due to the implementation of standardized residuals (see step 3 in the given algorithm) in the proposed method. For uniformity in results, the residuals used by each robust estimator were also standardized in the same manner in the simulation procedure. To restore the same random number generator setting amongst the various estimators, a seed command was

applied in the program written in the MATLAB environment. The results obtained in all three cases are described in Table 2 for the sample sizes (n) of 20, 50 and 100 respectively.

TABLE 2. Comparisons of different robust estimators

Methods	Case No-I		Case No-II		Case No-III	
	$\lvert\widehat{\beta}_0 - \beta_0\rvert$	$\lvert\widehat{\beta}_1 - \beta_1\rvert$	$\lvert\widehat{\beta}_0 - \beta_0\rvert$	$\lvert\widehat{\beta}_1 - \beta_1\rvert$	$\lvert\widehat{\beta}_0 - \beta_0\rvert$	$\lvert\widehat{\beta}_1 - \beta_1\rvert$
n=20						
LS	0.0342	0.0640	21.7744	3.3717	402.621	0.3354
MLS	0.0368	0.0738	0.0019	0.1905	0.0599	0.0012
Andrews	0.0441	0.0885	0.015	0.1331	0.2338	0.0016
Welsh	0.0470	0.0893	0.0142	0.1301	0.2320	0.0016
Huber	0.0467	0.0829	0.0300	0.3582	0.0871	0.0010
Cauchy	0.0436	0.0804	0.0131	0.2290	0.2346	0.0016
n=50						
LS	0.0374	0.0109	427.80	28.69	62.89	0.6105
MLS	0.0390	0.0175	0.0421	0.0059	0.0180	7.12×10^{-4}
Andrews	0.0546	0.0667	0.0724	0.1532	0.0135	2.29×10^{-4}
Welsh	0.0577	0.0582	0.0447	0.1522	0.0073	2.36×10^{-4}
Huber	0.0509	0.1919	0.1810	0.4726	0.0129	7.73×10^{-4}
Cauchy	0.0599	0.0449	0.0293	0.0920	0.0035	2.46×10^{-4}
n=100						
LS	0.0110	0.0915	37.893	16.546	80.43	1.3123
MLS	0.0167	0.1059	0.0399	0.1741	0.0807	3.52×10^{-4}
Andrews	0.0280	0.1022	0.0481	0.1416	0.0262	2.22×10^{-4}
Welsh	0.0268	0.1233	0.0435	0.1635	0.0864	2.22×10^{-4}
Huber	0.0250	0.1603	0.0813	0.1910	0.0841	4.59×10^{-4}
Cauchy	0.0274	0.1553	0.0419	0.1698	0.0291	2.23×10^{-4}

Results in Table 2, are reported for the absolute bias on the average of 10,000 simulation runs. It is obvious from the results obtained in all three cases that the proposed estimator was in close proximity to those as obtained from the LS estimator in the absence of any gross error irrespective of the various sample sizes. In the presence of usual values both in the x and y directions, it caused profound deterioration in the values of the estimated coefficients using the LS estimator. The proposed MLS estimator resulted in values closer to the benchmark values of β_0 and β_1, justifying its non-robustness property. In summary, the proposed estimator likewise with the other conventional M-estimators provides results which are insensitive to the substantial effect of the contaminated values. Also, notice that there are no markedly differences held in the results across the various sample sizes.

In addition, the computation of the RE of the MLS was also part of the simulation study which is the computation of the ratio of the minimum variance of an estimator to its real variance. In a regression context, the variance of the regression coefficients from the robust estimator was compared to that of the LS estimator since it is the most efficient when assumptions are suitably fulfilled. In order to compute the RE of the estimated coefficients defined following in (4. 12):

$$RE = \frac{MSE\left(\widehat{\beta}_{benchmark}\right)}{MSE\ \left(\widehat{\beta}\right)} \tag{4. 12}$$

where $MSE\left(\widehat{\beta}_{benchmark}\right)$ is average mean square error of the LS estimated coefficient computed over 10000 simulations when there was no contamination in the processing measurements. whereas $MSE\ \left(\widehat{\beta}\right)$ is mean square error of estimated coefficient using either the LS or MLS approach under the contaminated data. The error terms were allowed to contaminate the values from various distributions such as $N\,(0,\ 25)$, $t\,(1)$ and χ^2 (2. 3). The t-distribution with one and the chi-square distribution with 3 degrees of freedom are intuitively considered as cases of heavy tailed distributions. The RE of the MLS computed at various levels of contamination for the sample size of 50 is reported in Table 3.

TABLE 3. RE of the estimated coefficients when the contamination in errors is made from different distributions with different rates

	$\widehat{\boldsymbol{\beta_0}}$					$\widehat{\boldsymbol{\beta_1}}$				
Method	1%	5%	10%	20%	30%	1%	5%	10%	20%	30%
(a) $\mathbf{N(0,1)}$ with contamination $\mathbf{N(0,25)}$										
LS	1.0121	0.7936	0.6357	0.4118	0.3283	1.0121	0.7936	0.6357	0.4118	0.3283
MLS	1.0664	0.9274	0.8508	0.7193	0.6092	1.0664	0.9274	0.8508	0.7193	0.6092
(b) $\mathbf{N(0,1)}$ with contamination $\mathbf{t(1)}$-distribution (Cauchy distribution)										
LS	1.0123	0.9019	0.7614	0.5694	0.4433	1.0123	0.9019	0.7614	0.5694	0.4433
MLS	1.0661	0.9478	0.8837	0.7777	0.6889	1.0661	0.9478	0.8837	0.7777	0.6889
(c) $\mathbf{N(0,1)}$ with contamination $\boldsymbol{\chi^2}\ \mathbf{(3)}$-distribution										
LS	1.0105	0.8770	0.7582	0.6149	0.5364	1.0105	0.8770	0.7582	0.6149	0.5364
MLS	1.0664	0.9431	0.8821	0.7691	0.6747	1.0664	0.9431	0.8821	0.7691	0.6747

Since most of the conventional estimators in the variety of the statistical package manuals are described at the tuning value which provides approximately 95% efficiency; thereby, the MSE of the estimated coefficients has also been evaluated using the MLS estimator at tuning constant of equivalent efficiency. The larger size of the RE fairly indicates that the estimator with respect to the benchmark value is more efficient. It can be seen clearly from the Table 3 that the results of the LS estimator became unfavourable as the contamination level increased. Whereas, the MLS, in the absence of or only a small amount of contamination had a performance equivalently efficient to that of the LS estimator. For a moderate to higher rate of contamination, the MLS is more efficient than the LS estimator with respect to the benchmark value. Additionally, the MLS apparently outperforms to some extent in the presence of contaminated errors from the normal distribution.

5. APPLICATIONS

In order to investigate the effectiveness of the proposed estimator, we have taken some realistic data examples from the relevant literature. The first example is taken from Rousseauw and Leroy [15] about data originally presented by Hampel et al. [7] for the usefulness of the robust estimation procedures. The measurements in given data set concern the water flow rates at two different located cities (Libby and Newgate) for the month of January on the Kootnenary river in time span of 1931 to 1943. Hampel internationally introduced a bad leverage point by replacing original value of 44.9 to 15.7 for the Newgate city for the year 1934. Consequently, the contaminated value introduced in this way attracted the LS line towards itself and resulted in a drastic changed in the estimated coefficients. We applied proposed estimator on the actual and contaminated data set to verify its effectiveness. The standardized data as used by Hampel et al. [7], along with estimated values from the LS approach and proposed method are shown in Table 4.

TABLE 4. The original data on water flow rates and estimated values using different estimators

Libby (x_i)	Newgate(y_i)	LS	MLS	Huber	Cauchy	Welsch
0.6575	0.5678	0.6569	0.6454	0.6434	0.6436	0.6432
0.5071	0.5188	0.6593	0.556	0.5548	0.5336	0.5319
0.8104	0.7523	0.6543	0.7363	0.7335	0.7555	0.7564
1.8828	1.292	0.6366	1.3738	1.3653	1.5401	1.5501
0.8977	0.7523	0.6529	0.7882	0.7849	0.8194	0.821
0.5241	0.5736	0.6591	0.5661	0.5648	0.5460	0.5445
0.427	0.4525	0.6607	0.5084	0.5076	0.4750	0.4726
0.8516	0.7956	0.6536	0.7608	0.7578	0.7856	0.7869
0.7901	0.7177	0.6547	0.7242	0.7215	0.7406	0.7414
0.6308	0.6745	0.6573	0.6295	0.6277	0.6241	0.6235
0.6696	0.6658	0.6567	0.6526	0.6506	0.6525	0.6522
0.9389	0.9022	0.6522	0.8127	0.8092	0.8495	0.8515
0.6745	0.686	0.6566	0.6555	0.6534	0.6561	0.6558

The results given in Table 4, show that the proposed estimator provides reliable results like the other well-known robust procedures in the presence of a bad leverage value. The procedure of the LS is highly sensitive to even a single outlier value whereas; the MLS estimator is relatively non-robust in sense of providing fewer weights to outlying values. Note here, the results given in Table 4 are based on different tuning constant values which give equal efficiency of 95% for all robust methods. The coefficient of determination for the LS was only 0.002 whereas its value for the proposed estimator was 0.9335 which was close to benchmark value of 0.9453 for the original data without contamination. The scatter plots with fitted lines from each method in Figure 5 further depict the usefulness of the proposed method.

To further describe the foregoing, the second example is taken from Street et al. [20] work which has also been used by other investigators for illustration of their robust procedures [21].The data set comprise of US population (in millions) recorded at census interval of ten years from 1790 to 1990.The scatter plot of the given data set is shown in Figure 6(a). According to Figure 6(a), second degree polynomial seems to be appropriate for the given data. Therefore the LS and MLS quadratic fits are calculated and shown in Figure 6(b).

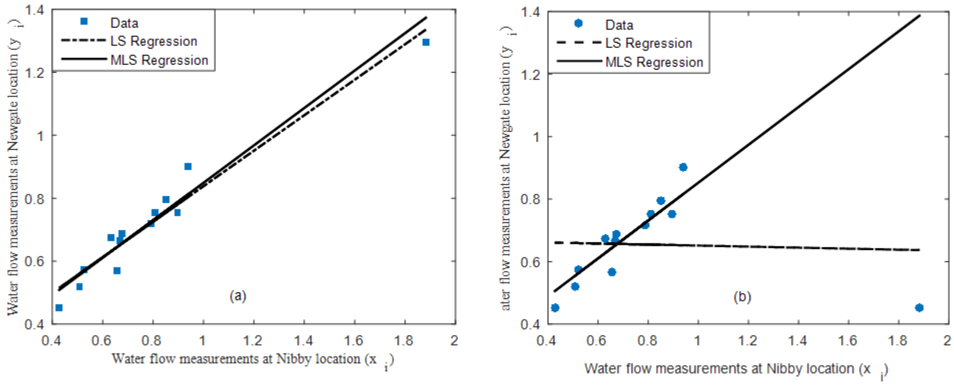

FIGURE 5. The LS and MLS fits on data (a) with original measurements (b) with contamination measurements

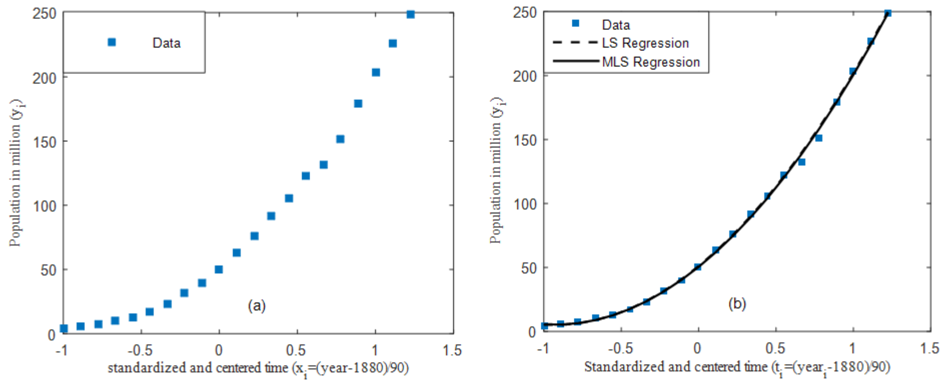

FIGURE 6. The scatter plot of the US population growth data vs standardized years (a) without fitted lines (b) with fitted lines

In absence of any outlier value the estimated coefficients from our proposed estimator and the LS are in proximity such that one cannot distinguish between the fitted lines without looking critically on given data.

Now we intentionally corrupted some certain values by adding identical amount of (1.2, 50) in given data $(x_i \;\; y_i)$ in order to produce gross errors in some measurements. The LS method based on the estimated values of the regression coefficients from uncontaminated data were taken as benchmarks. The relative mean absolute error (RMAE) value for the regression coefficients were calculated according to (5. 13).

$$RMAE = \frac{1}{k}\sum_{j=1}^{k}\left|\frac{\widehat{\beta}_j - \beta_j}{\beta_j}\right| \tag{5. 13}$$

where $\widehat{\beta}_j$ is j^{th} coefficient estimated value for contaminated data, β_j is a benchmark value for the coefficient and k is total number of the estimated coefficients which is three for our quadratic fit including the intercept term.
By taking (5. 13) into consideration, the results summary for the relative performance of the MLS is described in Table 5.

TABLE 5. The estimation results of proposed method and others conventional methods on the US population growth (in million) in time span of 1790 to 1990

	RMAE				
Number of Outliers	OLS	MLS	Huber	Cauchy	Welsch
1	0.2233	0.0126	0.0127	0.0108	0.0132
2	0.1826	0.0100	0.0097	0.0104	0.0131
3	0.4529	0.0515	0.0711	0.0105	0.0136
4	0.4227	0.0541	0.0725	0.0107	0.0145
5	0.6186	0.4540	0.4539	0.4544	0.4539

To better understand the effectiveness of the procedure, we have presented the results of comparative performance of the proposed estimator to that of the LS estimator and other conventional robust procedures. All other robust procedures were evaluated at respective tuning constant values of equal efficiency of 95%. The results reported in Table 5, indicates that the proposed procedure is not very sensitive to outlying values. In, the presence of different number of deliberated outliers the performance of the MLS is excellent in comparison to that of the LS method. In the occurrence of extreme outlying values the performance of the MLS was a slightly affected like the other M-estimators. However, such extreme offset values usually do not occur and can readily be diagnosed. In addition, its performance is almost identical to the other conventional robust procedures.

6. CONCLUSION

In this study, a new robust MLS estimator has been introduced. The MLS estimator provides a more general approach that can be converted to the LS estimator for some specific value of the tuning parameter. In addition a robust approach for estimating the regression parameters has been described which showed the bounded influence in the directions of response and explanatory variable. The simulation studies exhibited that estimated results from the MLS approach are accurate as those produced from the LS estimator in absence of outlying values. Whereas, contaminated measurements have less influence on the estimated results from MLS estimator. Moreover, the real data examples also revealed the robustness property of the proposed estimator.

REFERENCES

[1] W. H. Aeberhard and E. Cantoni and S. Heritier, *Robust Inference in the Negative Binomial Regression Model With an Application to Falls Data*, Biometrics **70**(4), (2014) 920-931.

[2] C. Anderson and R. E. Schumacker, *A Comparison of Five Robust Regression Methods With Ordinary Least Squares Regression: Relative Efficiency, Bias, and Test of the Null Hypothesis*, Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences **2**(2), (2003) 79-103.

[3] M. Avella Medina and E. Ronchetti, *Robust Statistics: A Selective Overview and New Directions*, Wiley Interdisciplinary Reviews: Computational Statistics **7**(6), (2015) 372-393.

[4] P. Bergstrm and O. Edlund, *Robust Registration of Point Sets Using Iteratively Reweighted Least Squares*, Computational Optimization and Applications **58**(3), (2014) 543-561.

[5] R. Couillet and F. Pascal and J. W. Silverstein, *The Random Matrix Regime of Maronna?s M-estimator with Elliptically Distributed Samples*, Journal of Multivariate Analysis **139**, (2015) 56-78.

[6] R. F. Hampel, *The Breakdown Points of the Mean Combined with Some Rejection Rules*, Technometrics,**27**(2), (1985) 95-107.

[7] R.F. Hampel and E. M. Roncheti and P. J. Rousseeuw, *Robust Statistics: The Approach Based on Influence Functions*, New York, John Wiley & Sons 2005.

[8] P. J. Huber, *"Robust Statistics," in International Encyclopedia of Statistical Science*, Springer, 1248-1251, 2011.

[9] P. J. Huber, *Robust Statistics: A Review*, The Annals of Mathematical Statistics **43**(4), (1972) 1041-1067.

[10] G. W. Imbens and M. Kolesr, *Robust Standard Errors in Small Samples: Some Practical Advice*, Review of Economics and Statistics, **98**(4), (2016) 701-712.

[11] Z. Khan and R. B. Razali and H. Daud and N. M. Nor and M. Fotuhi-Firuzabad and K. L. Krebs, *Bad Data Detection in Power System State Estimation Based on Generalized Likelihood Ratio Test*, International Journal of Energy and Statistics **4**(04), (2016) 1650016.

[12] P.-L. Loh, *Statistical Consistency and Asymptotic Normality for High-Dimensional Robust M-Estimators*, The Annals of Statistics **45**(2), (2017) 866-896.

[13] I. Ullah and M. A. Qadir and A. Ali, *Insha's Redescending M-Estimator for Robust Regression: A Comparative Study*, Pakistan Journal of Statistics and Operation Research **2**(2),(2006) 135-144.

[14] R.A. Maronna and R.D. Martin and V. Yohai, *Robust Statistics*, John Wiley & Sons, 2006.

[15] A. Muhammad and G. Mustafa, *Ridge Regression Based Subdivision Schemes for Noisy Data*, Punjab Univ. j. math. **51**(3), (2019) 61-69.

[16] S. Negahban and M. J. Wainwright, *Estimation of Near Low-rank Matrices with Noise and High-dimensional Scaling*, The Annals of Statistics (2011) 1069-1097.

[17] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*, John Wiley & Sons, 2005.

[18] Y. She and A. B. Owen, *Outlier Detection Using Nonconvex Penalized Regression*, Journal of the American Statistical Association **106**(494), (2011) 626-639.

[19] J. O. Street and R. J. Carroll, and D. Ruppert, *"A Note on Computing Robust Regression Estimates Via Iteratively Reweighted Least Squares*, The American Statistician **42**(2), (1988) 152-154.

[20] L. Wang,*Estimation of Exponential Population With Nonconstant Parameters Under Constant-Stress Model*, Journal of Computational and Applied Mathematics **342**, (2018) 478-494.

[21] Y.-G. Wang and X. Lin and M. Zhu and Z. Bai, *Robust Estimation Using the Huber Function with a Data-dependent Tuning Constant*, Journal of Computational and Graphical Statistics **16**(2), (2007) 468-481.