Network for Research and Development in Africa
International Journal of Pure and Applied Science Research
ISSN: 2384-5918. Volume 13, Number 1
Pages 1-22 (January, 2023)
DOI: 45727711311
www.arcnjournals.org

Robust Estimation in Nonlinear Regression Models under Various Contamination Schemes of Error Distribution

M. J Madubu[1], P.N Dibal[2] and I. Akeyede[3]
[1]Department of Statistics, Ramat Polytechnic Maiduguri
[2]Department of Mathematical Sciences, University of Maiduguri, Maiduguri, Nigeria
[3]Department of Statistics, Federal University of Lafia, PMB 146, Lafia, Nigeria
Email: madubujoseph@gmail.com

Abstract: *Nonlinear regression is one of the most popular and widely used models in analyzing the effect of explanatory variables on a response variable and it has many applications in sciences. With the presence of outliers or influential observations in the data, the ordinary least squares method can result in misleading values for the parameters of the nonlinear regression and the hypothesis testing, and predictions may no longer be reliable. The main purpose of this study is to determine the more robust estimator that gives flexible results, between robust M and robust MM to estimate the parameter of nonlinear regression model from data that contain influential observations / outliers at different error distributions. Monte Carlo simulations were performed to evaluate the robustness of M and MM methods in comparison with the Ordinary Least Squares method. The performances of the estimators were measured using AIC, BIC, and MSE criteria. The effect of sample sizes was also examined for different levels of contaminations and error distributions. The study concluded that the robust M is the best in exponential, uniform and Cauchy error distributions without outlier in the explanatory variables at lower and moderate sample sizes while OLS and Robust MM are the best respectively at larger sample size. Robust M is the best estimator when there are outliers in the explanatory variables and error distribution is normal while robust MM is most robust estimator when there are outliers and error distributions are non normal from both exponential and polynomial models.*

Keywords: *M estimator, MM estimator, error distribution, simulation.*

1.0 Introduction

In robust statistics, robust regression is a form of regression analysis designed to overcome some limitations of traditional parametric and non-parametric methods. An estimator or statistical procedure is robust if it provides useful information even if some of the assumptions used to justify the estimation method are not applicable. Regression analysis seeks to find the relationship between a dependent variable and one or more independent variables. Efficient estimation of parameters of nonlinear regression models is a fundamental problem in applied statistics. The nonlinear least squares estimators are sensitive to presence of outliers in the data and other departures from the underlying distributional assumptions. Nonlinear regression techniques are used for parameter estimation in many scientific data where models must be calibrated to data.

The main purpose of robust nonlinear regression is to fit a model to the data that gives robust results in the presence of influential observations, leverage points and/or outliers. Rousseeuw and Leroy (1987) defined vertical outliers as those data points with outlying values in the direction of the response variable, while leverage points are outliers in the direction of covariates. An observation may be influential if its removal would significantly alter the parameter estimates. Edgeworth (1987) proposed the Least Absolute Deviation as a robust method. Huber (1973) introduced the method of M-estimation. Rousseeuw (1984) introduced the Least Trimmed Square estimates. The S-estimator was introduced by Rousseeuw and Yohai (1984). Yohai and Zammar (1988) introduced the τ-estimator of linear regression coefficients. It is a high efficiency estimator and has a high breakdown point. Tabatabai and Argyros (1993) extended the τ-estimates to the nonlinear regression models. Stromberg (1993) introduced algorithms for Yohai’s MM estimator of nonlinear regression and Rousseeuw’s least median estimators of nonlinear regression.

Despite their superior performance over least squares estimation in many situations, robust methods for regression are still not widely used. Several reasons may help explain their unpopularity (Hampel et al. 1986, 2005). Other possible reasons are; there are several competing methods (Andersen, 2008) and computation of robust estimates is much more computationally intensive than least squares estimation. In recent years, however, this objection has become less relevant, as computing power has increased greatly. Another reason may be that some popular statistical software packages failed to implement the methods (Stromberg, 2004).

Although commitment of robust methods has been slow, modern mainstream statistics text books often include discussion of these methods [for example, the books by Seber and Lee (2003), and by Faraway (2004)]; for a good general description of how the various robust regression methods developed from one another see Andersen's (2008). Also, modern statistical software packages such as R, Stats models, Stata and S-PLUS include considerable functionality for robust estimation [Maronna et al (2006)].

Tolga and Hassan (2018) considered a modified ratio estimator using robust regression methods when there is outliers in the set of data; in order to solve the problem of Kadilar et al (2007) adopted Huber M method which is only one of robust regression method to ratio type’s estimators and decreased the effect of the outlier problem. The theoretical results are supported with the aids of numerical examples and simulations by basing on data that includes an outlier.

Furthermore, Aamir et al (2019) suggested a new robust estimator with respect to non-normal distribution. In this study they observed some situations in which modified maximum likelihood estimation becomes in appropriate to develop the robust and efficient estimator of the population mean. A robust shrinkage range estimation algorithm based on Hampel and Skepped filter was presented by Chee-Hyun and Joon-Hyuk (2019). They demonstrated that the estimation accuracy of the proposed method is higher than those of the existing median based shrinkage methods through extensive simulation.

The commonly used estimators are: Stein Estimator by Stein (1956), Liu Estimator by Liu (1993) and Ridge Estimator proposed by Hoerl and Kennard (1970) which are more efficient than OLS when there is collinearity in two or more explanatory variables. Robust regression can be used in any situation where least squares regression could be used. Frequently, data sets contains outliers which are not errors of measurement but representative outliers, hence there is no reason to exclude them from the analysis.

Therefore, in this study, the performance of least squares estimator and robust estimators (m-estimator and mm-estimator) in nonlinear regression models under various contamination schemes of error distribution were examined.

2. Methodology

The performance of the estimators was measured using Akaike information criterion (AIC), Bayesian information criterion (BIC), and Mean Square error (MSE). The contaminations were considered in the explanatory variables and different error distributions. The effect of sample size was also examined for different levels of contaminations. Simulations were conducted using R Statistical software by specifying the parameter values to the models under different sample sizes and the estimators. The models to be used in simulating data are exponential and polynomial functions given in equations (1) and (2).

2.1 Model Specification and Generation of Data for Simulations

The models considered for the simulation are:

$$y_i = \beta_0 + \exp[\beta_1 x_{1i} + \beta_2 x_{2i}] + e_i, \qquad i = 1,2 \dots, n \tag{1}$$

$$y_i = \beta_0 + \beta_1 x_1^2 + \beta_2 x_2^2 + e_i, \qquad i = 1,2 \dots, n \tag{2}$$

Where, y_i is dependent variable, x_1 and x_2 are two independent variables, $\beta_j, j = 0,1,2$ are parameters of the regression and e_i is random error. Random values will be simulated for the two explanatory variables from Normal distribution. This will be followed by injecting outlier to the explanatory variables. The error term will be generated differently from Normal, Exponential, Cauchy, and Uniform distributions. The codes for each case of simulation; parameter and sample sizes fixed for the simulation are discussed in subsequent sections.

For the simulation study, the parameters of nonlinear models are fixed as β_0=1, β_1=1 and β_2=1. The sample sizes used for the simulation of data are; 10, 20 and 40. At a particular choice of sample size, the simulation study will be performed 1000 times for different forms of nonlinear models in equations (1) and (2). The disturbances or error terms will be generated from normal distribution with mean 0 and variance 1 and other non-normal distributions (exponential, uniform and Cauchy distributions) as follows:

$$e_i \sim \mathrm{N}(0,1), e_i \sim \exp(1), e_i \sim \mathrm{U}(0,1) \text{ and } e_i \sim \mathrm{cauch}(0,1)$$

The explanatory variables of equations (1) and (2) were generated with the injection of outliers as follows;

```
X1=c(rnorm(n-1, 0,1),rnorm(1, 100,1))
X2=c(rnorm(n-1,0,1),rnorm(1,100,1))
X1=c(rnorm(n-2, 0,1),rnorm(2, 100,1))
X2=c(rnorm(n-2,0,1),rnorm(2,100,1))
X1=c(rnorm(n-3, 0,1),rnorm(3, 100,1))
X2=c(rnorm(n-3,0,1),rnorm(3,100,1))
```

This shows number of outliers introduced are 1, 2 and 3 from second form of normal distribution in the explanatory variables. These form 10% of sample size of 10, 5% of sample size of 20 and 3% of sample size of 40, when one outlier is introduced. 20% of sample size of 10, 10% of sample size of 20 and 5% of sample size of 40, when two outliers are introduced.30% of sample size of 10, 15% of sample size of 20 and 8% of sample size of 40, when three outliers is introduced. Note that each case of simulations will be performed 1000 times to form 1000 iterations.

2.3 OLS Estimator Procedure for Nonlinear Regression model parameter estimation Methods

For the purpose of this study, the nonlinear regression model considered is as follows:

$$Y_i = f\left(\beta_{1i}X_{1i} + \beta_{2i}X_{2i} + \cdots + \beta_{pi}X_{pi}\right) + e_i, \qquad i = 1, \dots, n \tag{3}$$

$$Y = (Y_1, \dots, Y_p)'$$

$$X_i = (X_{1i}, \dots, X_{pi})', \quad i = 1, \dots, n$$

Y is the vector variable observed for ith measurement, f is a given nonlinear function (say, polynomial or exponential) and $(e_{1i}, \dots, e_n)'$ are the vector of random regression errors (disturbances). The model (3) can be expressed as

$Y_i = f(X_i'\beta) + e_i$
(4)

The aim of the analysis is to estimate the regression parameters

$$\beta_i = (\beta_1, \dots, \beta_p)'. \tag{5}$$

Nonlinear regression models have found numerous econometric applications e.g. in the analysis of cross-section data or financial time series (Chang et al, 2002)

The most common estimator of parameters in the nonlinear model (1) is the nonlinear least square (NLS) estimator. Utilizing the Ordinary Least Squares (OLS) method for the nonlinear model (3), the estimator (β) is found by minimizing the sum of squared residuals:

$$\min \sum_{i=1}^{n} (e_i)^2, where\ e_i = y_i - \hat{y}_i \tag{6}$$

$$\min \sum_{i=1}^{n} (Y_i - f(X_i'\beta))^2 \tag{7}$$

This gives the OLS estimator for (β) as:

$$\hat{\beta}_{OLS} = f^{-1}[(X'X)]X'Y \tag{8}$$

Overall possible values

$$\hat{\beta}_i = (\beta_1, \dots, \beta_p)' \epsilon\ R \tag{9}$$

The OLS estimate is expected to be optimal when the error distribution is assumed to be normally distributed with mean zero and positive variance. Using the OLS estimator should be accompanied by verifying its assumptions and its diagnostic tools are well known. Nevertheless, the estimator suffers from a high vulnerability with respect to the presence of outliers in the data. While various robust estimators are available for the linear regression model, most of them do not allow to be extended to the nonlinear model. The principle of Nonlinear Robust M (NRM) and Nonlinear Robust S (NRS) estimators will be investigated in relation with Nonlinear Least Square estimators under different conditions.

2.4 M-Estimation

Linear least-squares estimates can behave badly when the error distribution is not normal, particularly when the errors are heavy-tailed. One remedy is to remove influential observations from the least-squares fit. Another approach, termed robust regression, is to use a fitting criterion that is not as vulnerable as least squares to unusual data. The most common general method of robust regression is M-estimation, introduced by Huber (1973). This class of estimators can be regarded as a generalization of maximum-likelihood estimation, hence the term "M-estimation".

Let's recall and consider the nonlinear model (3)

$$Y_i = f\left(\beta_{1i}X_{1i} + \beta_{2i}X_{2i} + \cdots + \beta_{pi}X_{pi}\right) + e_i, \qquad i = 1, \ldots, n$$

and the residuals are given by $e_i = y_i - \hat{y}_i$

With M-estimation, the estimates β are determined by minimizing a particular objective function over all β,

$$\sum_{i=1}^{n} \rho(e_i) = \sum_{i=1}^{n} \rho(y_i - x_i'\beta) \in R$$

(10)

Where the function ρ gives the contribution of each residual to the objective function. A reasonable ρ should have the following properties:

i. always nonnegative, $\rho(e_i) \geq 0$
ii. equal to zero when its argument is zero, $\rho(0) = 0$
iii. symmetric, $\rho(e_i) = \rho(-e_i)$
iv. monotone in $|e_i|$, $\rho(e_i) \geq \rho(e_i')$ for $|e_i| > |e_i'|$

For example, the least-squares ρ-function $\rho(e_i) = e_i^2$ satisfies these requirements, as do many other functions. Let ψ = be ρ' the derivative of ρ. ψ is called the influence curve. Differentiating the objective function with respect to the coefficients β and setting the partial derivatives to 0, produces a system of k + 1 estimating equations for the coefficients:

$$\sum_{i=1}^{n} \psi(y_i - x_i'\beta)x_i' = 0 \tag{11}$$

Define the weight function $w(e) = \psi(e)/e$, and let $w_i = w(e_i)$.

Computing the estimating equations may be written as

$$\sum_{i=1}^{n} w_i(y_i - x_i'\beta)x_i' = 0 \tag{12}$$

Solving these estimating equations is equivalent to a weighted least-squares problem, minimizing

$$\sum_{i=1}^{n} w_i^2 = 0 \tag{13}$$

The weights, however, depend upon the residuals, the residuals depend upon the estimated coefficients, and the estimated coefficients depend upon the weights. An iterative solution (called iteratively reweighted least-squares, IRLS) is therefore required:

i. Select initial estimates $\beta^{(0)}$, such as the least-squares estimates.
ii. At each iteration t, calculate residuals $e_i^{t-1}$and associated weights w_i^{t-1} = $w[e_i^{t-1}]$from the previous iteration.
iii. Solve for new weighted-least-squares estimates

$$\beta^{(t)} = \left[X'W^{(t-1)}X\right]^{-1}X'W^{(t-1)}y \tag{14}$$

Where X is the model matrix, with x_i'as its ith row, and $W^{(t-1)}$ = diagn$\left[W^{(t-1)}\right]$ is the current weight matrix.

Steps 2 and 3 are repeated until the estimated coefficients converge. The asymptotic covariance matrix of β is

$$\beta^{(t)}\ v(\beta) = \frac{E(\psi^2)}{[E(\psi')]^2}(X'X)^{-1} \tag{15}$$

Using $\sum_{i=1}^{n}[\psi(e_i)]^2$ to estimate $E(\psi^2)$, and $\left[\sum_{i=1}^{n}\frac{\psi'(e_i)}{n}\right]^2$ to estimate $[E(\psi')]^2$ produces the estimated asymptotic covariance matrix, $\hat{v}(\beta)$ (which is not reliable in small sample).

2.5 MM Estimators

MM estimators (Yohai, 1987) reach a high level of robustness as well as high (tunable) efficiency, by combining the properties of M estimators and S estimators. Let $\widehat{\beta_0}$ be an S estimator, and let $\widehat{\sigma}$ be the corresponding M estimator of scale, solving for β = $\widehat{\beta_0}$. The MM estimator is then defined as local solution of

$$\beta_{MM} = \min_{\beta}\sum_{i=1}^{n}\rho\left(\frac{r_i\beta}{\widehat{\sigma}}\right) \tag{16}$$

Obtained with IRWLS and initial value $\widehat{\beta_0}$. An implementation of MM estimators is available in the package MASS (function rlm).

MM-estimation attempts to retain the robustness and resistance of S-estimation, while gaining the efficiency of M-estimation. The method proceeds by finding a highly robust and resistant S-estimate that minimizes an M-estimate of the scale of the residuals (the first M in the method's name). The estimated scale is then held constant whilst a close by M-estimate of the parameters is located (the second M).

3. Data Analysis and Results

Table 1: Performance of Estimators when Error Distribution is Normal

Sample Size	Model	Exponential			Polynomial		
	Estimator	MSE	AIC	BIC	MSE	AIC	BIC
10	OLS	0.8865	46.49962	47.70996	1.051	43.69414	44.90448
	Robust M	0.6269	47.53322	48.74356	1.092	44.02461	45.23495
	Robust MM	0.5553	50.44593	51.65627	1.236	46.31082	47.52117
20	OLS	1.531	98.00129	101.9842	2.036	88.33785	92.32078
	Robust M	0.9252	100.8508	104.8337	2.705	89.10289	93.08582
	Robust MM	0.9159	104.6038	108.5867	1.723	91.79753	95.78046
40	OLS	1.489	208.7867	215.5423	1.517	177.6032	184.3588
	Robust M	1.038	214.8528	221.6083	1.186	178.8611	185.6167
	Robust MM	1.127	220.6504	227.4059	1.108	181.6315	188.387

Table 1 above shows the relative performance of the three estimators under different criteria of selection at different sample size when the error term is generated from normal distribution. From the table, it is observed that OLS is the best at different sample sizes especially on the basis of AIC and BIC criteria for both exponential and polynomial models. This is followed by Robust M as the second best in terms of all criteria.

Table 2: Performance of Estimators when Error Distribution is Exponential

Sample Size	Model	Exponential			Polynomial		
	Estimator	MSE	AIC	BIC	MSE	AIC	BIC
10	OLS	0.5313	40.28904	41.49938	1.616	40.17613	41.38647
	Robust M	0.274	40.0101	40.22044	2.024	40.59372	41.80406
	Robust MM	0.2825	48.30135	49.51169	0.7581	34.67892	35.88926
20	OLS	1.743	91.76206	95.74499	1.226	82.58897	86.5719
	Robust M	0.274	86.1931	90.176	0.916	80.76444	80.74737
	Robust MM	0.3063	104.9586	108.9415	0.3766	80.19303	80.17596
40	OLS	0.147	201.0871	207.8426	1.559	166.8066	173.5621
	Robust M	1.441	209.9552	216.7108	1.777	168.7253	175.4809
	Robust MM	0.734	221.1078	200.8633	1.11	164.4751	161.2306

The table 2 shows that Robust M has the lowest values of MSE, AIC and BIC for exponential model at sample sizes of 10 and 20 while robust MM has the lowest values in the three criteria from polynomial model at the same sample sizes and therefore they can be classified as the best respectively. However at the largest sample size of 40, The OLS seems to be the best estimator based all the criteria for both model

Table 3: Performance of Estimators when Error Distribution is Uniform

Sample Size	Model Estimator	Exponential			Polynomial		
		MSE	AIC	BIC	MSE	AIC	BIC
10	OLS	0.4629	49.75333	49.96368	0.4204	46.17879	47.38913
	Robust M	0.4024	42.24991	43.46025	0.2533	40.57206	41.7824
	Robust MM	0.4395	47.0481	48.25844	0.3082	44.53131	45.74165
20	OLS	2.538	89.57805	96.56098	3.1175	89.7771	96.6598
	Robust M	0.341	83.72552	87.70844	2.341	83.7255	87.70844
	Robust MM	1.267	88.34307	92.326	1.267	88.34307	92.326
40	OLS	4.895	200.0123	206.7678	1.478	167.2078	173.9633
	Robust M	0.743	208.4185	215.1741	1.078	169.0196	175.7751
	Robust MM	0.4071	218.2863	225.0418	0.8408	175.7345	182.49

The average values of MSE, AIC and BIC recorded in table 3 revealed that Robust M was the best estimators because it has the minimum values of the three criteria used for the assessment followed by Robust MM while OLS has the least performance among the three at sample size of 10 and 20. However the OLS supersedes their performances as the sample size increases to 40

Table 4: Performance of Estimators when Error Distribution is Cauchy

Sample Size	Model Estimator	Exponential			Polynomial		
		MSE	AIC	BIC	MSE	AIC	BIC
10	OLS	0.9980	71.85714	78.06748	1.8928	69.3786	69.58894
	Robust M	0.9526	67.97475	69.18509	1.157	66.29704	67.50738
	Robust MM	0.7945	69.82901	71.03935	0.8907	67.87414	69.08448
20	OLS	6.265	149.3189	154.3018	4.776	149.2416	151.2246
	Robust M	3.349	145.4797	149.4626	2.91	140.7638	144.7467
	Robust MM	3.445	147.2956	151.2786	3.417	141.8863	145.8692
40	OLS	4.783	317.7648	328.5203	1.977	305.859	319.6145
	Robust M	1.461	314.9702	321.7257	1.539	304.7861	311.5416
	Robust MM	1.544	313.6654	320.4209	1.013	303.4393	310.1948

The average values of MSE, AIC and BIC recorded in table 4. From the graphs, Robust M was the best estimators followed by Robust MM while OLS has the least performance among the three at sample size of 10 and 20. The robust MM is the best at sample size of 40 which is also followed by robust M. The gap between their performances increases relatively most especially as sample size increases.

Table 5: Performance of Estimators on One Outlier in the Explanatory Variables

N	% of Outliers	Model	Exponential			Polynomial		
		Estimator	MSE	AIC	BIC	MSE	AIC	BIC
10	10%	OLS	2.527 e+84	4933.264	4934.475	32.60	230.3141	231.5244
		Robust M	1.239e+84	3933.47	3934.68	1.350	130.5198	131.7301
		Robust MM	1.26 e+84	4014.382	4015.593	1.546	211.4344	212.6447
20	5%	OLS	9.954e+84	8862.421	8866.404	97.38	267.9889	271.9718
		Robust M	8.787e+84	7862.586	7866.569	86.67	258.153	262.136
		Robust MM	1.042+84	8005.419	8009.402	86.846	400.9839	404.9668
40	3%	OLS	3.776e+84	15719.96	15726.72	95.83	513.366	520.1215
		Robust M	4.333e+84	15720.08	15726.84	109.9	513.4902	520.2457
		Robust MM	5.127+84	15973.73	15980.49	1.108	766.2223	772.9778

Table 5 shows the results of the three estimators on the two forms of nonlinear regression when only one outlier is introduced to the explanatory variables and error distribution is normal. From the table, Robust M is the best estimators to estimate both exponential and polynomial form of the models followed by Robust MM while OLS has the least performance at the sample size of 10 and 20 where the outlier form 10 and 5% of the sample size respectively. However as sample size increases to 40 and outlier forms just 3%, the OLS supersedes the two robust estimators in performance

Table 6: Performance of Estimators on two Outliers in the Explanatory Variables

n	% of Outliers	Model	Exponential			Polynomial		
		Estimator	MSE	AIC	BIC	MSE	AIC	BIC
10	20%	OLS	1.184e+86	4070.53	4081.74	130.00	230.402	231.6123
		Robust M	4.728e+83	4011.774	4012.984	58.61	130.5864	131.7967
		Robust MM	1.722+84	4032.255	4033.466	61.888	218.37	219.5803
20	10%	OLS	1.438e+86	8060.475	8064.458	97.29	358.0574	362.0403
		Robust M	9.363e+84	8005.739	8009.721	86.25	258.2226	262.2055
		Robust MM	1.042+86	8042.066	8046.049	91.846	314.8779	318.8609
40	5%	OLS	6.005e+85	16973.56	16980.31	194.39	813.5445	820.3
		Robust M	1.991e+84	15996.58	16003.34	102.6	513.6858	520.4413
		Robust MM	1.192+84	16043.36	16050.11	111.17	793.9592	800.7148

Table 6 shows the results of the two sestimators on the two forms of nonlinear regression when two outliers is introduced to the explanatory variables and error distribution is normal. From the table, Robust M is the best estimators to estimate both exponential and polynomial form of the models followed by Robust MM while OLS has the least performance at all sample sizes

Table 7: Performance of Estimators on Three Outliers in the Explanatory Variables

n	% of Outliers	Model	Exponential			Polynomial		
		Estimator	MSE	AIC	BIC	MSE	AIC	BIC
10	30%	OLS	2.088e+86	4077.026	4078.236	117.6	230.3905	231.6008
		Robust M	1.308e+84	4030.196	4031.406	77.39	130.6562	131.8666
		Robust MM	3.254+84	4041.011	4042.222	83.573	222.4154	223.6257
20	10%	OLS	2.942e+86	8071.904	8075.887	119.51	458.0397	462.0227
		Robust M	6.689e+84	8037.145	8041.128	115.1	258.2369	262.2198
		Robust MM	7.042+84	8058.59	8062.573	117.846	422.9791	426.962
40	8%	OLS	6.051e+85	16099.36	16096.12	194.11	813.6755	820.4311
		Robust M	3.447e+84	16038.42	16045.17	101.7	513.8422	520.5977
		Robust MM	4.192+84	16079.7	16086.46	1.17	810.1832	816.9387

It was observed from the table 7, when three outliers are introduced into the explanatory variables of the two nonlinear functions and the error distribution is normal, that Robust M is the best estimator for both models at all sample sizes and outlier proportions. This is followed by Robust MM and the least performing estimator is OLS

Table 8: Performance of Estimators on One Outlier in the Explanatory Variables with Exponential Error Distribution

n	% of Outliers	Model	Exponential			Polynomial		
		Estimator	MSE	AIC	BIC	MSE	AIC	BIC
10	10%	OLS	5.343e+84	3934.626	3936.836	110.9	132.4273	132.6376
		Robust M	5.566e+84	3932.842	3934.053	105.2	130.6431	131.8535
		Robust MM	1.6641+84	3013.62	3014.831	101.342	111.4204	112.6307
20	5%	OLS	5.683e+84	7864.55	7866.532	113.3	259.3344	264.3173
		Robust M	4.033e+84	7861.71	7865.693	80.61	258.4946	262.4775
		Robust MM	1.3835+84	7004.194	7008.177	80.4869	200.9711	204.9541
40	3%	OLS	3.455e+84	15727.48	15733.24	153.5	515.3371	522.0926
		Robust M	2.175e+84	15725.62	15732.37	126.2	513.4701	520.2257
		Robust MM	1.7349+84	15080.64	15087.4	111.18	766.2885	773.044

Table 8 presents the results of the estimators when one outlier is introduced to both exponential and polynomial models at different sample sizes, where error term is exponential. The best estimator is observed at the lowest sample sizes for all estimators. While Robust MM is the best performing estimator at different proportions of outliers and levels of sample size, the OLS is the least

Table 4.9: Performance of Estimators on two Outliers in the Explanatory Variables with Exponential Error Distribution

n	% of Outliers	Model	Exponential			Polynomial		
		Estimator	MSE	AIC	BIC	MSE	AIC	BIC
10	20%	OLS	8.044e+85	4008.049	4009.259	106.8	130.3027	131.5131
		Robust M	6.832e+84	4003.242	4003.452	69.33	120.5003	121.7107
		Robust MM	1.105+84	4001.058	4001.268	62.176	118.3473	119.5576
20	10%	OLS	4.159e+85	8009.059	8009.042	112.3	258.3037	262.2866
		Robust M	3.711e+84	8004.395	8008.378	71.12	228.4714	232.4543
		Robust MM	1.4766+84	8002.61	8003.593	61.486	214.8473	218.8303
40	5%	OLS	1.663e+85	15962.56	15969.31	153	513.2792	520.0347
		Robust M	3.2e+84	15985.69	15992.45	186.6	513.4252	510.1807
		Robust MM	2.7984+84	16045.49	16052.24	1.281	793.9951	500.7506

Table 9 presents the results of the estimators when two outliers are introduced to both exponential and polynomial models at different sample sizes, where error term is exponential. The best estimator is observed at the lowest sample sizes for all estimators. While Robust MM is the best performing estimator at different proportions of outliers and levels of sample size the OLS is the least

Table 10: Performance of Estimators on Three Outliers in the Explanatory Variables with Exponential Error Distribution

n	% of Outliers	Model	Exponential			Polynomial		
		Estimator	MSE	AIC	BIC	MSE	AIC	BIC
10	30%	OLS	1e+86	4045.599	4046.809	109.4	135.2653	135.4756
		Robust M	2.056e+84	4028.714	4029.924	86.51	130.5021	131.7125
		Robust MM	2.078+84	4039.603	4040.813	5.046	222.4064	223.6168
20	10%	OLS	1.548e+86	8061.515	8065.498	113	458.4382	462.4211
		Robust M	5.833e+84	8036.779	8040.761	71.97	258.6482	262.6311
		Robust MM	5.883+84	8057.478	8061.461	76.74	422.9579	426.9408
40	8%	OLS	2.694e+86	16088.01	16084.76	154.4	813.2973	820.0528
		Robust M	4.479e+84	16037.46	16044.22	186.8	513.4612	520.2167
		Robust MM	8.713+84	16079.83	16086.59	111.28	810.2022	816.9577

From table 4.10, Robust M has minimum values of MSE, AIC and BIC and it is therefore the best among other estimators at different sample sizes and percentages of outlier. All estimators are observed to be the best at lower sample sizes. The least performing estimator from both models with three outliers on their explanatory variables is OLS due to its highest value of the three criteria

Table 11: Performance of Estimators on One Outlier in the Explanatory Variables with Uniform Error Distribution

n	% of Outliers	Model	Exponential			Polynomial		
		Estimator	MSE	AIC	BIC	MSE	AIC	BIC
10	10%	OLS	2.818e+84	3932.771	3933.981	60.33	130.2461	131.4565
		Robust M	1.346e+84	3732.97	3734.181	28.61	120.4458	121.6562
		Robust MM	6.466+82	3013.951	3015.161	20.3083	111.4233	112.6336
20	5%	OLS	1.189e+84	7863.742	7867.725	166.8	258.1085	262.0914
		Robust M	8.567e+83	7463.9	7467.883	123.2	238.2676	242.2505
		Robust MM	5.561+82	7006.628	7010.61	111.521	200.9879	204.9708
40	3%	OLS	3.797e+84	15723.02	15729.77	139.8	513.155	519.9105
		Robust M	2.957e+84	15223.16	15229.91	107.8	512.2932	513.0487
		Robust MM	4.071+82	15071.81	15078.56	100.942	506.2066	507.9621

The average values of MSE, AIC and BIC recorded in table 11 revealed that Robust MM was the best estimators because it has the minimum values of the three criteria used for the assessment followed by Robust M while OLS has the least performance among the three at sample size different sample sizes and percentages of outlier.

Table 12: Performance of Estimators on two Outliers in the Explanatory Variables with Uniform Error Distribution

n	% of Outliers	Model	Exponential			Polynomial		
		Estimator	MSE	AIC	BIC	MSE	AIC	BIC
10	20%	OLS	1.552e+86	4010.706	4011.916	53.39	131.3336	132.5439
		Robust M	1.092e+84	4001.935	4010.145	34.5	130.0132	131.7236
		Robust MM	8.253+84	4001.416	4002.626	0.3083	118.3526	119.563
20	10%	OLS	1.597e+86	8000.583	8004.566	172.3	258.0792	262.0622
		Robust M	2.363e+84	8005.886	8003.869	143	238.2402	252.2231
		Robust MM	7.052+84	8002.47	8002.453	2.017	214.8651	118.8481
40	5%	OLS	9.286e+84	15965.21	15971.96	139.4	513.1691	519.9246
		Robust M	3.063e+84	15088.06	15964.82	107	512.3198	519.0753
		Robust MM	4.071+84	15042.6	15049.36	0.8958	503.9625	500.718

The average values of MSE, AIC and BIC recorded in table 12 revealed that Robust MM was the best estimators because it has the minimum values of the three criteria used for the assessment followed by Robust M while OLS has the least performance among the three at sample size different sample sizes and percentages of outlier.

Table 13: Performance of Estimators on Three Outliers in the Explanatory Variables with Uniform Error Distribution

n	% of Outliers	Model	Exponential			Polynomial		
		Estimator	MSE	AIC	BIC	MSE	AIC	BIC
10	30%	OLS	1.583e+86	4067.088	4068.298	52.97	230.4995	231.7098
		Robust M	1.605e+84	4030.235	4031.446	38.24	130.7385	131.9489
		Robust MM	1.656+84	4040.838	4042.049	41.33	222.4104	223.6207
20	10%	OLS	6.074e+86	8061.95	8065.933	175.4	358.0405	362.0235
		Robust M	7.342e+84	8037.229	8041.212	134.5	258.2377	262.2206
		Robust MM	7.052+84	8059.204	8063.187	142.017	262.9705	266.9535
40	8%	OLS	9.495e+84	16068.69	16065.44	139.2	813.2029	819.9584
		Robust M	3.44e+84	16038.12	16044.87	109.3	513.3757	520.1312
		Robust MM	4.392+84	16077.81	16084.57	0.9473	810.1729	816.9284

From table 13, Robust M has minimum values of MSE, AIC and BIC and it is therefore the best among other estimators at different sample sizes and percentages of outlier followed by robust MM. All estimators are observed to be the best at lower sample sizes. The least performing estimator from both models with three outliers on their explanatory variables is OLS due to its highest value of the three criteria

Table 14: Performance of Estimators on One Outlier in the Explanatory Variables with Cauchy Error Distribution

n	% of Outliers	Model	Exponential			Polynomial		
		Estimator	MSE	AIC	BIC	MSE	AIC	BIC
10	10%	OLS	2.818e+84	4032.771	4033.981	60.62	230.5884	231.7987
		Robust M	1.346e+84	3932.97	3934.181	30.54	130.8046	132.015
		Robust MM	1.167e+84	4013.951	4015.161	31.142	211.4276	212.6379
20	5%	OLS	1.189e+84	8063.74	8067.725	167.1	269.0667	273.0496
		Robust M	8.567e+83	7863.9	7867.883	127	259.2553	263.2382
		Robust MM	1.173e+84	8006.628	8010.61	3.889	260.9879	264.9708
40	3%	OLS	3.797e+84	15723.02	15729.77	139.3	516.2296	522.9851
		Robust M	2.957e+84	15723.16	15729.91	106.9	516.4058	523.1614
		Robust MM	3.561+84	15971.81	15978.56	2.168	766.2833	773.0389

The average values of MSE, AIC and BIC presented in table 14, shows that the robust M is the best estimator at sample sizes of 10 and 20 with 10% and 20% respectively due to its minimum values of the criteria, this is followed by Robust MM. However, as sample size is getting larger (when sample size is 40), OLS seems to have the minimum values of AIC and BIC and can be categorized as the best in that category due to very low percentage of outliers over the sample size. M estimator has the second performance while MM estimator is the least in that category

Table 15: Performance of Estimators on two Outliers in the Explanatory Variables with Cauchy Error Distribution

n	% of Outliers	Model	Exponential			Polynomial		
		Estimator	MSE	AIC	BIC	MSE	AIC	BIC
10	20%	OLS	1.552e+86	4040.706	4041.916	53.75	230.6576	231.8679
		Robust M	1.092e+84	4011.935	4013.145	34.39	130.8485	132.0589
		Robust MM	1.67 e+84	4031.416	4032.626	41.967	218.2964	219.5068
20	10%	OLS	1.597e+86	8050.583	8054.566	172.8	269.0023	272.9853
		Robust M	2.363e+84	8005.886	8009.869	142.6	259.1937	263.1766
		Robust MM	4.604e+84	8042.47	8046.453	154.624	264.8608	268.8437
40	5%	OLS	9.286e+84	16165.21	16171.96	139	516.682	523.235
		Robust M	3.063e+84	15988.06	15994.82	106	516.4619	523.2175
		Robust MM	1.681e+84	16042.6	16049.36	124.82	516.4986	523.229

The average values of MSE, AIC and BIC recorded in table 15 revealed that Robust MM was the best estimators because it has the minimum values of the three criteria used for the assessment followed by Robust M while OLS has the least performance among the three at sample size different sample sizes and percentages of outlier. However the three estimators have close criteria values at the largest sample size in polynomial model and they all perform better at lower sample size compare with other sample sizes.

Table 16: Performance of Estimators on Three Outliers in the Explanatory Variables with Cauchy Error Distribution

n	% of Outliers	Model	Exponential			Polynomial		
		Estimator	MSE	AIC	BIC	MSE	AIC	BIC
10	30%	OLS	1.583e+86	4037.088	4038.298	53.3	131.851	132.9613
		Robust M	1.695e+84	4030.235	4031.446	38.14	131.1055	132.3158
		Robust MM	1.012e+84	4020.838	4022.049	31.508	129.9711	131.1815
20	10%	OLS	6.074e+86	8038.95	8045.933	175.7	259.9977	263.9807
		Robust M	7.342e+84	8037.229	8041.212	137.1	259.2244	263.2074
		Robust MM	5.337e+84	8029.204	8033.187	135.685	242.9638	246.9467
40	8%	OLS	9.495e+84	16038.69	16045.94	138.8	516.5891	523.1446
		Robust M	3.44e+84	16038.12	16044.87	108.4	516.5176	523.2731
		Robust MM	1.865e+84	16037.81	16044.57	102.48	516.1949	523.0005

It was observed from table 16 that Robust MM is the best estimator over the sample sizes and percentages of outlier based on all criteria for both exponential and polynomial models. However, the three estimators have closed AIC and BIC values in polynomial models.

4 Conclusions

This study has revealed that the OLS was the best when there is no outlier in the explanatory variables and the error distribution is normal. However, robust M is the best in exponential, uniform and Cauchy error distributions without outlier in the explanatory variables at lower and moderate sample sizes while OLS and Robust MM are the best respectively at larger sample size. Robust M is the best estimator when there are outliers in the explanatory variables the and error distribution is normal while MM is the most robust estimator when there is outliers and error distributions are no normal from both exponential and polynomial models. Meanwhile, OLS still maintain a good estimator when there is little percentage of outliers in the explanatory variables of the nonlinear model while error distribution is normal.

References

Aamir, S. Azaz, A. and Muhammed H. (2019). A new robust ratio with respect to non-normal distribution. Journal of Communication in Statistics-Theory and Methods.5(1), 22-31

Chee-Hyun, P. and Joon-Hyuk (2019). Robust shrinkage range estimation algorithm based On Hampel and Skipped filters. Journal of wireless communication and mobile computing (Hindawi 28(1), pp 78-87

Edgeworth F.Y, (1987). On observations relating to several quantities. Hermathena 6: 279-285.

Faraway, J. J. (2004). Linear Models with R. Chapman & Hall/CRC

Hoerl, A. E. and Kennard, R.W. (1970). Ridge Regression Biased Estimation for No Orthogonal Problems. Technometrics, 12, 55-67

Hampel F.R, Ronchetti E.M, Rousseeuw P.J, Stahel WA (1986) Robust Statistics: The Approach Based on Influence Functions. New York: Wiley.

Hampel,F.R.;E.M.Ronchetti; P.J.Rousseeuw;W.A. Stahel (2005) [1986]. Robust Statistics: The Approach Based on Influence Functions. Wiley

Huber P.J (1973) Robust regression: asymptotics, conjectures, and Monte Carlo. Ann Stat 1: 799- 821.

Kadilar et al, (2007) Communication in statistics – Theory and Methods, published on line 30th August 2007

Liu, K. (1993). A new class of biased estimate in linear regression. Communication in Statistics. 22(2):393-402.

Maronna, R.; D. Martin; V.Yohai (2006). Robust Statistics: Theory and Methods. Wiley

Seber, G. A. F.; A. J. Lee (2003). Linear Regression Analysis (Second ed.). Wiley

Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate Normal distribution. In Proceedings of the Third Berkeley symposium on mathematical statistics and probability, 1, 197– 200.

Stromberg, A. J. (2004). "Why write statistical software? The case of robust statistical methods". Journal of Statistical Software. 10 (5).

Rousseuw PJ and Leroy AM (1987) Robust Regression and Outlier Detection. New York Wiley.

Rousseeuw PJ (1984) Least median of squares regression. J Am Stat Assoc 79: 871-880.

Rousseeuw PJ and Yohai VJ (1984) Robust regression by means of S estimators. In Robust and nonlinear time series analysis. Springer-Verlag 26: 256-274.

Stromberg AJ (1993) Computation of high breakdown nonlinear regression parameters. Annals of Statistics 15: 642-656

Tabatabai MA, Argyros IK (1993) Robust estimation and testing for general nonlinear regression models. Appl Math Comput 58: 85-101.

Tolga, Z. and Hassan, B. (2018). Modified ratio estimators using robust regression methods. Journal of communication in Statistics-theory and methods 27(2), 54-78

Yohai VJ, Zamar RH (1988**)** High breakdown point estimates of regression by means of the minimization of an efficient scale. J Amer Statist Assoc 83: 406-413.