مجلة العلوم البحتة والتطبيقية
Journal of Pure & Applied Sciences
www.Suj.sebhau.edu.ly ISSN 2521-9200
Received 25/04/2019 Revised 18/08/2019 Published online 11/12/2019

A study on the performance of five robust nonlinear regression estimators

* Abdelgadir Khalifa Alsalem[a], Alsaidi M. Altaher[b]
[a]Department of Mathematical sciences, Basic Science School/Libyan Academy, Libya
[b]Department of Statistics, Faculty of Science/University of Sebha, Libya
*Corresponding author: abdelgadirkhalifa@yahoo.com

Abstract Nonlinear regression is one of the most popular and widely used models in analysing the effect of explanatory variables on a response variable when the underling regression function is nonlinear. It has many applications in scientific research such as dose response studies conducted in agricultural sciences, toxicology and other biological sciences. Estimating the parameters of a nonlinear regression model is usually carried out by the least squares (LS. However, In the presence of outliers, even one single unusual value may have a large effect on the parameter estimates. The aim of this paper is to introduce the most commonly used methods as a better choice to the classical least squares. This includes M-estimator, MM-estimator, CM-estimator, tau-estimator and mtl-estimator. Moreover, the target is to compare their practical performance under a variety of circumstances such as sample size, percentage of outliers and model formula. Results of Monte Carlo simulations using R software, indicated that the best performance has been achieved by MM followed by CM estimator for all possible percentages of outliers (10%, 20%, 30%, 40%) as well as all sample sizes (n=50, n=100, and n=150). Moreover, results approved that the LS estimator remains the best when there is no outlier in data.

Keywords: robust estimator, nonlinear model, outliers, simulation**.**

دراسة حول اداء خمسة مقدرات متينة لنماذج الانحدار اللا خطي

*عبدالقادر خليفة السالم[1] و السعيدي المهدي الطاهر[2]
[1] قسم العلوم الرياضية–مدرسة العلوم الأساسية–الأكاديمية الليبية، ليبيا
[2] قسم الإحصاء–كلية العلوم–جامعة سبها، ليبيا
*للمراسلة: abdelgadirkhalifa@yahoo.com

الملخص الانحدار غير الخطي هو أحد النماذج الأكثر شيوعاً والأكثر استخداماً في تحليل تأثير المتغيرات التفسيرية على متغير الاستجابة عندما تكون طبيعة العلاقة غير خطية. لنماذج الانحدار الغير خطي العديد من التطبيقات في مختلف العلوم كالزراعية والبيولوجية وغيرها من العلوم الاخرى. عادةً ما يتم تقدير المعلمات الخاصة بنموذج الانحدار غير الخطي بواسطة المربعات الصغرى (LS)، الا انه في وجود القيم المتطرفة يخفق ادائها فيكون حتى لقيمة واحدة متطرفة تأثيرا كبير اعلى تقديرات المعلمات. الهدف من هذه الورقة هو التعريف بالطرق الأكثر شيوعاً كخيارات أفضل للمربعات الصغرى الكلاسيكية، ويشمل ذلكM-Estimator، MM-Estimator، CM-Estimator، tau-Estimator، mtl-Estimator. علاوة على ذلك، فإن الهدف هو مقارنة أدائها العملي ضمن ظروف مختلفة مثل حجم العينة، النسبة المئوية من القيم المتطرفة والصيغة النموذجية للدالة الانحدارية. أشارت نتائج المحاكاة باستخدام برنامج R إلى أن أفضل أداء تم تحقيقه بواسطة MM متبوعا بتقدير CM لجميع النسب المئوية المحتملة من القيم المتطرفة (10 ٪، 20 ٪، 30 ٪، 40 ٪) وكذلك جميع أحجام العينات (ن = 50، ن = 100، ن = 150). وعلاوة على ذلك، وافقت النتائج الاعتقاد النظري الذي ينص على أن مقدر LS يبقى الأفضل عندما لا يكون هناك قيم متطرفة في البيانات.

الكلمات المفتاحية: المقدر المتين، النموذج اللاخطي، القيم المتطرفة، المحاكاة.

1. Introduction

Nonlinear models have been applied to a wide range of situations, even to finite population. Life is full of random phonemes that behave in a nonlinear manner, say for instance, growth from birth to maturity in human subjects typically is nonlinear in nature, characterized by rapid growth shortly after birth, pronounced growth during puberty, and a levelling off sometime before adulthood.; see, e.g., [1] and [2]. Estimation of the parameters of a nonlinear regression model is usually carried out by the method of least squares or the method of maximum likelihood, just as for linear regression models. Unlike linear regression, it is usually not possible to find analytical expressions for the least squares and maximum likelihood estimators for nonlinear regression models since the normal equations are not linear in this case and are, in general, difficult to solve. Instead, numerical search or iterative procedures must be used with both of these estimation procedures, requiring intensive computations. The iterative procedures usually require to starting values for the parameter estimates. The choice of the starting values is very important because a poor choice may result in a lengthy computation with many iterations. It may also lead to divergence, or to a convergence due to a local minimum. Therefore, good initial values will result in fast computations with few iterations and if multiple minima exist, it will lead to a solution that is a minimum. Hence, we conclude that nonlinear regression is a hot and complex research area in which one can contribute specially when the classical assumptions are not satisfied such as the normality assumption and outlying observations.

With the presence of outliers in the data, the ordinary least squares LS method provides misleading values for the parameters of the nonlinear regression, and predictions may no longer be reliable, see [3]. Outliers are those observations that deviate markedly from other members of the observations or data points which are unusually large or small from the majority of the observations. They are also called the abnormal data. Outliers can arise due to measurement or recording error, natural variation of the underlying distribution, or a sudden alteration in the operating system. One disadvantages of nonlinear least square with Gauss-Newton method is its sensitivity to the presence of even few outlying observations. As a result, the errors in the process of prediction and estimating as they amplify the variance of errors, leading to extended the length of confidence interval and reduced estimation efficiency. The proposed solutions for estimating the parameters of the nonlinear regression model in the presence of outliers is the use of robust estimators rather than the method of LS. In statistical literature, most robust linear regression techniques are successfully adopted for nonlinear setting, such as M-estimator, MM-estimator, Least Median of Squares (LMS), Least Trimmed Squares (LTS), See [4], [5], [6], [7] and [8]. So much comparisons had been conducted to compare the robust estimators in linear regression, see for example [9], [10], and [11]. However, Little work has been done in nonlinear regression. Hence, such simulation comparisons in nonlinear regression are required. This paper comes as an extension of a previous study by [12] in which only two robust estimators are compared. This shall contribute in presenting new simulation results to determine which robust estimator should be used when outliers are existing in data.

The rest of this paper is organized in the following manner: Section 2 gives a brief review to the Gauss- Newton Method, M-estimator, MM-estimator, CM-estimator, tau-estimator and mtl-estimator. In section 3, simulation study is conducted. Conclusions are present in section 4.

2. Methods and Models

In this section we present brief theoretical descriptions of Gauss-Newton method for classical LS, M-estimator, MM-estimator, CM estimator, tau estimator and mtl estimator.

2.1 The Gauss - Newton Method

A nonlinear regression model can be written as

$$y_i = f(x_i, \theta) + \varepsilon_i \quad , i = 1, \dots, n \qquad (1)$$

To estimate the parameters of (1) using the Gauss – Newton method, a Taylor series expansion is used to approximate the nonlinear regression model with linear terms and then employs ordinary least squares to estimate the unknown parameters. Once the starting values for the parameters have been chosen, the mean responses $f(X;\theta)$ is approximated for the n cases by the linear terms in the Taylor series expansion around the starting values $\theta_k^{(0)}$, such as:

$f(x_i, \theta) \approx$

$$f\left(x_i, \theta^{(0)}\right) + \sum_{k=0}^{p} \left[\frac{\partial f(x_i, \theta)}{\partial \theta_k}\right]_{\theta=\theta^{(0)}} \left(\theta_k - \theta_k^{(0)}\right) \qquad (2)$$

The Taylor approximation (2) becomes in this notation:

$$f(x_i, \theta) \approx f_i^{(0)} + \sum_{k=0}^{p} F_{ik}^{(0)} \beta_k^{(0)} \qquad (3)$$

Therefore, by ordinary least squares we can estimate the parameters $\beta^{(0)}$:

$$b^{(0)} = (F^{(0)\prime} F^{(0)})^{-1} F^{(0)\prime} Y^{(0)} \qquad (4)$$

$$\theta^{(1)} = \theta^{(0)} + b^{(0)}$$

In general $\theta^{(a)} = \theta^{(a-1)} + b^{(a-1)}$ (5)

2.2 The M-Estimator

The "M" in the term "M-estimator" stands for maximum-likelihood-type estimator [3]. This name stems from the fact that an M-estimator can be loosely interpreted as a maximum-likelihood estimator, albeit for an unknown, non-Gaussian model. To cope with the problem of outliers, Huber [3] introduced a class of the so-called M-estimators, in which the sum of function ρ of the residuals is minimized. The estimated vector of the parameters $\hat{\beta}_M$ estimated by an M-estimator is given by

$$\hat{\beta}_M = \arg min_{\beta} \sum_{i=1}^{n} \rho\left(\frac{r_i}{\sigma}\right) \qquad (6)$$

A popular choice for σ is the median absolute deviation

$$\sigma = median|r_i - median(r_i)|/0.6745$$

Function $\rho(.)$ must be even, nondecreasing for positive values, and less increasing than the square. However, it is simpler to differentiate ρ with respect to β and solve for the root of the derivative. When this differentiation is applicable, the M-estimator is said to be of ψ-type. Let $\psi = \rho'$ be the derivative of ρ, and define weights such as $w_i = \psi(\frac{r_i}{\sigma})/r_i$, the estimates $\hat{\beta}_M$ is obtained by solving the system of equations:

$$\sum_{i=1}^{n} w_i^2 r_i^2 = 0 \qquad (7)$$

2.3 The MM-Estimator

The "MM" in the term "MM-estimator" stands for the two stage maximum-likelihood estimator. The MM- estimator by [6] introduces the multi-stage estimator (MM-estimator). It is a combination of high breakdown and high efficiency. It can be obtained using a three-stage procedure. At first stage, an initial consistent estimate $\hat{\beta}_0$ with high breakdown point with possibly low normal efficiency is obtained. Yohai [6] suggests using the S-estimator for this stage. In the second stage, a robust M-estimator of scale parameter $\hat{\sigma}$ of the residuals based on the initial value is calculated. In the third stage, an M-estimator $\hat{\beta}$ starting at $\hat{\beta}_0$ is obtained. Huber or bi-square functions is typically used as the initial estimate $\hat{\beta}_0$. Let $\rho_0(r) = \rho_1\left(\frac{r}{k_0}\right), \rho(r) = \rho_1\left(\frac{r}{k_1}\right)$, and assume that each of the

ρ_i functions is bounded, $i = 0$ and 1. The scale estimate $\hat{\sigma}$ satisfies the following equation:

$$\frac{1}{n}\sum_{i=1}^{n} \rho_0\left(\frac{y_i - f(x_i,\theta)}{\hat{\sigma}}\right) = 0.5 \quad (8)$$

where $k_0 = 1.56$ when the ρ function is biweight. In this case the estimator has the asymptotic breakdown point ($BP = 0.5$).

2.4 The CM-Estimator

Constrained M-estimators (CM-estimators) for the regression parameters β and the scale parameter σ were introduced by Mendes and Tyler [13] The aim is to have robust regression estimators with high breakdown point and high asymptotic relative efficiency. The CM-estimates for β and σ are defined as the global minimum of the objective function

$$L(\beta,\sigma) = ave\{p(r_i/\sigma)\} + log\sigma \quad (9)$$

over all $\beta \in R^p$ and $\sigma > 0$ subject to

$$ave\{p(r_i/\sigma)\} \leq \varepsilon\, p(\infty) \quad (10)$$

where ε is a fixed number between 0 and 1, and ave stands for the arithmetic average.

As shown in [13], the CM-estimators are regression and affine equivariant, and possess at the same time, the good local properties of the M-estimators for regression and good global robustness properties of the regression S-estimators. The breakdown point of the CM-estimates is approximately **min (ε,1- ε)** or approximately 0.5 when $\varepsilon = 0.5$. Also, when p is properly tuned the CM-estimates can have good local robustness properties. They are consistent, asymptotically normal and very efficient estimators. See [13] for more details

2.5 The tau-Estimator

A new class of robust estimators; called the τ-estimator is introduced by Yohai and Zamar [14]. The τ-estimator possess simultaneously the following properties: (i) they are qualitatively robust, (ii) their breakdown point is 0.5, and (iii) they are highly efficient for regression models with normal errors.

Let p_1 and p_2 be two function satisfying assumption of ρ which defined in MM-estimator and let s_n be the M-estimate of scale based on p_1. Then given a sample $u = (u_1, u_2, \dots, u_n)$, the scale estimate τ_n is defined by

$$\tau_n^2(u) = s_n^2(u)\frac{1}{n}\sum_{i=1}^{n} p_2\left(\frac{u_i}{s_n(u)}\right) \quad (11)$$

Clearly τ_n is scale equivariant, i.e.,

$$\tau(\lambda u) = |\lambda|\tau(u) \qquad \forall \quad \lambda \in \mathbb{R}.$$

If $p_1 = p_2$ we get $\tau_n = \sqrt{p}\, s_n$. If $p_2(u) = u^2$ we get the sample standard deviation. The τ-estimates for regression are defined by the value $\hat{\theta}$ such that

$$\tau_n\left(r(\hat{\theta})\right) = min_\theta\ \tau_n(r(\theta)) \quad (12)$$

2.6 The mtl-Estimator

Let X be a random variable having a probability density $p(x;\theta)$ which depends on an unknown parameter θ, where both X and θ may be vector valued, and $x_1, x_2, \dots, x_n$ be n independent realizations of X.

The maximum likelihood (ML) estimator of θ is the value of θ that maximizes the likelihood function

$$L(\theta;\ x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} p(x_i;\theta) \quad (13)$$

or, equivalently, the logarithm of the likelihood function

$$\sum_{i=1}^{n} \ell(\theta; x_i), \quad (14)$$

where

$$\ell(\theta; x_i) = \ln p(x_i;\theta) \quad (15)$$

is a function of θ which represents the contribution of the ith observation to the log likelihood function (14). This method can be used irrespective of whether X and θ are scalar-or vector-valued. We shall refer to $L(\theta; x_i) = p(x_i;\theta)$ and $\ell(\theta; x_i)$ as the likelihood and the log likelihood curves, respectively, when considered as functions of θ for every x_i. The above robust methods have one aspect in common: each is based on some sort of trimming the observations (note that down weighting is a type of trimming).

The method proposed here is to replace the log likelihood function by the trimmed log likelihood function

$$\sum_{i=a}^{b} w_i \ell(\theta; x_i), \quad (16)$$

where $a \leq b$, $(a,b) \in [1,2,\dots,n\}$, and $w_i \geq 0$ are weights. By maximizing (3.37), an estimator of θ can be obtained. This estimator is denoted by $\hat{\theta}(a,b,w)$ indicating that it depends on a,b, and w, where w is a vector containing the weights w_i. Unless otherwise stated, we investigate the case $w_i = 1, a \leq i \leq b$. In this case, the estimators are denoted by $\hat{\theta}(a,b)$, for simplicity of notation. When a = 1 and b = n, (1, n) is clearly the maximum likelihood estimator (MTL) of θ. When $a > 1$ and $b < n$, the estimator is based on trimming the log likelihood function. We therefore refer to the method as the maximum trimmed likelihood (MTL) method and to $\hat{\theta}(a,b)$, as the maximum trimmed likelihood estimator (MTLE).

3. *Results and Discussion*

In this section we summarize and discus the numerical results from simulation study.

3.1 Simulations setup

Simulation is a technique for guiding the experiments of a model. Compared with analytical methods, simulation is easily understandable and highly realistic. In this study, all computations and graphics were carried out using the software package R.

3.2 Models used in simulation

3.2.1 Michaelis-Menten Model

Michaelis-Menten model, used by [15] and [16], expresses the reaction velocity as a function of concentration of substrate as

$$y_i = \frac{\beta_0 x_i}{\beta_1 + x_i} + \varepsilon_i$$

Where response variable y_i is velocity and predictor variable x_i is substrate; the parameter is β_0 the maximum reaction velocity and β_1 denotes concentration of substrate. In this simulation the true parameter values are chosen to be β_0 = 5 and β_1 = 1 as in [17]. However, different true parameters are possible as long as convergence occurs in optimization process.

3.2.2. Exponential Model

It is a two parameter model given by the following relationship

$$y_i = \beta_0\, e^{\beta_1 x_i} + \varepsilon_i$$

Where β_0, β_1 are parameters, x is independent predictor, y response predictor, ε is random variable. Where β_0 = 0.2, β_1 = 0.3 are the true parameters (these values have been chosen arbitrary with the advantage of fast computations and the convergence occurs very quickly).

3.2.3 Logistic Model

It is a three parameter model given by the following relationship

$$y_i = \frac{\beta_0}{1 + e^{(\beta_1 - \log(x_i)/\beta_2)}} + \varepsilon_i$$

with β_0 = 5, β_1 = 1, β_2 = 0.6 . These values have been chosen as in *nlr* package.

3.2.4 Logarithm Model

From the R package and using *nlr* Library, with the command *nlrobj3 [[6]]* we get the following model:

$$y_i = log\ (\beta_0 + \beta_1 x_i) + \varepsilon_i$$

with β_0 = 6.64, β_1 = 0.369. These values have been chosen as in *nlr* package.

3.3 Data generation

The contaminated normal distribution a simple useful distribution that can be used to simulate outliers [16]. For each model the explanatory x is chosen uniformly within the range (1,10). We control the outlier percentage through the outlier generating model by [16] such as:

$$\varepsilon :\ (1-\tau)N(\mu_1, \sigma_1^2) + (\tau)N(\mu_2, \sigma_2^2)$$

Where the proportion $(1-\tau)\%$ refers to the percentage of non-outlier data, while the proportion $\tau\%$ refers to percentage of outliers. For each sample size, 500 random data are generated with $\mu_1 = 0, \sigma_1^2 = 10, \mu_2 = 0, \sigma_2^2 = 0.2$.

3.4 Comparison criterion

The mean squared error (MSE) is used as a comparison criterion. The mean squared error is estimated by

$$MSE = \frac{\sum_{i=1}^{m}(\hat{f}_i - f)^2}{m}$$

where $\hat{f}_i$ is the fitted values and f the true function with m=500 (replication number for each model).

3.5 Sample sizes and outliers percentages

We have used three sample sizes (n=50,100,150), and four outlier percentages (10%,20%,30%,40%).

4. Main findings

Having examined Tables 1 to 4 carefully, we have noted many important features: The best performance has been achieved by the OLS-estimator when there are no outliers in the simulated data (0%) for all possible of sample sizes (n=50, n=100, n=150) as well as for the four different models. An interesting feature to note is that the performance of the OLS estimator is the worst one for all possible percentages of outliers (10%, 20%, 30%, 40%) as well as all possible of sample sizes (n=50, n=100, and n=150) with the four different models. For all possible of percentages of outliers (10%, 20%, 30%, 40%) the robust MM-estimator outperforms the CM-estimator, and they are both are superior to LS and other robust estimator in all cases.

Table 1: Simulation results for first model with sample sizes (n=50,100,150) and outlier percentages (0%, 10%, 20%, 30% ,40%).

N		0%	10%	20%	30%	40%
	OLS	0.00073	0.11074	4.11536	3.27919	0.39553
	M	0.00163	0.01101	0.01088	0.05132	0.00185
	MM	0.00265	0.00563	0.00056	0.00142	0.00189
50	CM	0.00372	0.00635	0.00062	0.00685	0.01434
	tau	0.003	0.01849	0.00208	0.00858	0.01379
	mtl	0.01145	0.00991	0.00034	0.0129	0.0024
	OLS	0.00242	0.35308	0.01755	0.30446	0.33077
	M	0.0028	0.00051	0.00117	0.0087	0.01975
	MM	0.00282	0.00091	0.00034	0.00164	0.00095
100	CM	0.00549	0.0012	0.00029	0.00252	0.00181
	tau	0.00269	0.00101	0.00051	0.00455	0.00181
	mtl	0.03214	0.00184	0.00113	0.00422	0.00056
	OLS	0.00179	0.16295	0.09501	0.04144	0.17088
	M	0.00223	0.00023	0.0016	0.00358	0.01206
	MM	0.0023	0.00051	0.00065	0.00053	0.00046
150	CM	0.00448	0.00034	0.00093	0.00109	0.00107
	tau	0.00278	0.00051	0.00182	0.00059	0.00139
	mtl	0.00675	0.00067	0.00126	0.00174	0.00501

Table 2: Simulation results for second model with sample sizes (n=50,100,150) and outlier percentages (0%, 10%, 20%, 30% ,40%).

N		0%	10%	20%	30%	40%
50	OLS	3.32E-03	4.58E-01	8.23E-01	3.94E+00	2.06E+00
	M	4.51E-03	1.72E-03	2.32E-01	3.47E-01	1.34E+00
	MM	4.62E-03	2.60E-03	1.90E-02	8.06E-02	3.44E-02
	CM	7.95E-03	2.34E-03	1.85E-02	5.78E-02	3.35E-02
	tau	5.02E-03	2.78E-03	2.65E-02	6.01E-02	2.88E-01
	mtl	4.51E-03	5.13E-03	7.01E-03	6.73E-02	2.36E-01
100	OLS	2.95E-03	3.05E-01	1.36E-01	8.97E-01	1.56E+00
	M	4.23E-03	3.46E-03	2.48E-03	1.79E-01	1.33E-01
	MM	4.00E-03	2.50E-04	1.82E-02	6.48E-02	1.65E-01
	CM	6.59E-03	1.33E-03	1.70E-02	3.66E-03	1.51E-01
	tau	3.86E-03	1.64E-03	2.75E-03	5.91E-03	1.60E-01
	mtl	1.98E-02	7.95E-03	4.47E-03	2.54E-02	1.24E-01
150	OLS	1.42E-03	1.08E+00	2.27E-01	7.59E-01	1.31E+00
	M	1.64E-03	4.15E-03	4.59E-02	2.23E-01	4.12E-02
	MM	1.70E-03	1.75E-03	7.10E-03	4.14E-01	4.55E-01
	CM	2.27E-03	3.16E-03	6.07E-03	1.47E-02	5.82E-01
	tau	1.58E-03	1.03E-03	2.28E-03	2.30E-02	5.81E-01
	mtl	2.78E-03	2.59E-03	1.92E-02	1.35E-02	0.06191

Table 3: Simulation results for third model with sample sizes (n=50,100,150) and outlier percentages (0%, 10%, 20%, 30% ,40%).

N		0%	10%	20%	30%	40%
50	OLS	5.30E-03	7.42E-01	6.52E-02	1.17E+00	4.38E-01
	M	6.00E-03	2.87E-03	3.46E-03	2.81E-02	1.39E-02
	MM	5.71E-03	4.88E-03	1.54E-03	1.90E-03	6.89E-03
	CM	1.53E-02	5.98E-03	1.10E-03	3.04E-03	4.48E-03
	tau	5.46E-03	6.44E-03	3.70E-03	3.27E-03	4.39E-03
	mtl	4.63E-02	1.04E-02	9.00E-03	6.75E-03	7.32E-03
100	OLS	6.20E-04	0.467033	2.16E-01	4.74E-01	9.04E-01
	M	1.05E-03	0.004108	6.41E-03	5.37E-02	2.11E-02
	MM	1.14E-03	0.001235	2.78E-03	1.10E-02	1.19E-02
	CM	2.29E-03	0.000928	2.25E-03	1.71E-03	1.36E-02
	tau	1.09E-03	0.000792	2.24E-03	3.92E-03	1.38E-02
	mtl	3.69E-03	0.002359	2.36E-03	6.15E-03	2.39E-02
150	OLS	0.001301	1.20E-01	3.84E-01	4.06E-02	8.12E-02
	M	0.001427	7.70E-04	2.41E-03	3.50E-04	1.19E-03
	MM	0.0016	1.10E-03	5.60E-04	9.30E-04	3.02E-02
	CM	0.002561	1.54E-03	6.80E-04	8.60E-04	2.22E-03
	tau	0.001593	2.80E-03	2.58E-03	1.17E-03	1.70E-03
	mtl	0.0085	3.04E-03	1.02E-02	4.14E-03	0.00678

Table 4: Simulation results for fourth model with sample sizes (n=50,100,150) and outlier percentages (0%, 10%, 20%, 30% ,40%).

N		0%	10%	20%	30%	40%
50	OLS	2.10E-03	1.49E-01	7.36E-02	1.84E-02	2.32E-01
	M	2.30E-03	1.45E-03	8.60E-04	3.47E-03	1.93E-03
	MM	2.14E-03	1.16E-03	1.78E-03	2.81E-03	1.21E-03
	CM	2.35E-03	1.49E-03	1.30E-03	1.71E-03	1.67E-03
	tau	2.30E-03	3.48E-03	6.18E-03	2.18E-03	1.42E-03
	mtl	1.34E-02	3.67E-03	6.11E-03	1.70E-03	6.58E-03
100	OLS	3.50E-04	7.87E-02	9.31E-02	6.12E-02	7.79E-02
	M	4.60E-04	7.90E-04	3.47E-03	2.00E-03	1.45E-02
	MM	4.40E-04	3.80E-04	1.78E-03	1.65E-03	7.25E-03
	CM	5.30E-04	3.50E-04	1.71E-03	5.30E-04	1.90E-03
	tau	5.40E-04	2.90E-04	2.47E-03	4.20E-04	1.84E-03
	mtl	2.25E-03	5.60E-04	2.72E-03	1.12E-03	4.16E-03
150	OLS	3.00E-05	8.28E-02	2.00E-01	8.65E-02	4.05E-02
	M	8.00E-05	1.11E-03	5.60E-04	2.13E-03	1.46E-02
	MM	5.00E-05	5.10E-04	3.00E-05	1.06E-03	1.79E-03
	CM	6.00E-05	5.50E-04	7.00E-05	9.90E-04	1.97E-03
	tau	4.00E-05	1.00E-03	2.00E-04	1.23E-03	1.98E-03
	mtl	8.50E-04	1.00E-03	1.90E-04	4.90E-04	0.0029

5. Conclusion

In this paper the issue of robust nonlinear regression is considered. Five robust estimators are evaluated, namely **M**-estimator, **MM**-estimator, **CM**-estimator, **tau**-estimator and **mtl**-estimator using simulation study. It has been deduced that all these robust estimators are superior to classical least squares in the presence of outliers. The best performance is achieved by MM-estimator followed by CM-estimator. For further investigation one can take into consideration the issue of outliers in X-direction, XY direction, in addition to apply these methods for real data.

References

[1]- G. A. F. Seber and C. J. Wild, *Nonlinear Regression*. Wiley, 2003.

[2]- D. M. Bates and D. G. Watts, *Nonlinear Regression Analysis and Its Applications*. Wiley, 2007.

[3]- P. J. Huber, "Robust estimation of a location parameter," in *Breakthroughs in statistics*, Springer, 1992, pp. 492–518.

[4]- H. Barreto and D. Maharry, "Least median of squares and regression through the origin," *Comput. Stat. Data Anal.*, vol. 50, no. 6, pp. 1391–1397, 2006.

[5]- Y. Chen, A. J. Stromberg, and M. Zhou, "The least trimmed squares estimate in nonlinear regression," *Rel. téc., Dep. Stat. Univ. Kentucky, Lexington, KY*, vol. 40506, 1997.

[6]- V. J. Yohai, "High breakdown-point and high efficiency robust estimates for regression," *Ann. Stat.*, pp. 642–656, 1987.

[7]- D. M. Hawkins and D. M. Khan, "A procedure for robust fitting in nonlinear regression," *Comput. Stat. Data Anal.*, vol. 53, no. 12, pp. 4500–4507, 2009.

[8]- A. A. Alamgir, S. A. Khan, D. M. Khan, and U. Khalil, "A new efficient redescending M-estimator: Alamgir redescending M-estimator," vol. 2277, p. 2502, 2013.

[9]- R. Maronna, R. D. Martin, and V. Yohai, *Robust Statistics Theory and Methods*. John Wiley & Sons, 2006.

[10]- P. J. Rousseeuw, "Least median of squares regression," *J. Am. Stat. Assoc.*, vol. 79, no. 388, pp. 871–880, 1984.

[11]- P. Rousseeuw and V. Yohai, "Robust regression by means of S-estimators," in *Robust and nonlinear time series analysis*, Springer, 1984, pp. 256–272.

[12]- A. Alsalem and A. Altaher, "A comparative study of some robust nonlinear regression methods," in *PSC*, 2019.

[13]- B. Mendes and D. E. Tyler, "Constrained M-estimation for regression," in *Robust statistics, data analysis, and computer intensive methods*, Springer, 1996, pp. 299–320.

[14]- V. J. Yohai and R. H. Zamar, "High breakdown-point estimates of regression by means of the minimization of an efficient scale," *J. Am. Stat. Assoc.*, vol. 83, no. 402, pp. 406–413, 1988.

[15]- A. J. Stromberg and D. Ruppert, "Breakdown in nonlinear regression," *J. Am. Stat. Assoc.*, vol. 87, no. 420, pp. 991–997, 1992.

[16]- D. E. Herwindiati, M. A. Djauhari, and M. Mashuri, "Robust multivariate outlier labeling," *Commun. Stat. Comput.*, vol. 36, no. 6, pp. 1287–1294, 2007.

[17]- M. A. Tabatabai, J. J. Kengwoung-Keumo, W. M. Eby, S. Bae, U. Manne, M. Fouad, and K. P. Singh, "A New Robust Method for Nonlinear Regression," *J. Biom. Biostat.*, vol. 5, no. 5, p. 211, 2014.