Pak. J. Statist.
2017 Vol. 33(3), 223-236

AN EFFICIENT AND HIGH BREAKDOWN ESTIMATION PROCEDURE FOR NONLINEAR REGRESSION MODELS

Dost Muhammad Khan[1§], Shumaila Ihtesham[2], Amjad Ali[2], Umair Khalil[1], Alamgir[3], Sajjad Ahmad Khan[1] and **Sadaf Manzoor[2]**

[1] Department of Statistics, Abdul Wali Khan University Mardan,
KP, Pakistan

[2] Department of Statistics, Islamia College Peshawar, KP, Pakistan

[3] Department of Statistics, University of Peshawar, KP, Pakistan

[§] Corresponding Author Email: dost_uop@yahoo.com

ABSTRACT

In regression analysis least square (LS) estimator fails because of its sensitivity to unusual observations present in the data. Robust estimation provides alternative estimates which are insensitive and efficient, when the data are not normally distributed or polluted with distant observations usually called outliers. To cope with the problem of outliers is more challenging in nonlinear regression (NLRM) than linear regression. In this study, least trimmed absolute (LTA) estimator which is a robust and high breakdown estimator is adopted for nonlinear regression model fitting. Bias and mean square error is used to check the efficiency of the proposed estimator. The performance of the estimator is compared with LS and existing robust M-estimators using simulated data sets and real world problems. It has been observed that LTA is efficient in case of contaminated data sets as compared to LS and M estimator. Furthermore, it has been concluded that in case of 40% contamination LTA outperforms LS and M estimator, while in case of 20% outliers LTA and M estimators perform equally well. In regard to clean data sets, LS, LTA and M estimator performs equally well. Conclusion has been made on the basis of simulated data sets as well as real data sets.

KEY WORDS

High Breakdown, nonlinear least squares, least trimmed square, leverage points, outliers, M-estimator.

1. INTRODUCTION

The fundamental goal of nonlinear regression is identical to linear regression that is to regress a response variable Y to a set of predictor variable X. In nonlinear regression model, the response variable is nonlinearly related to explanatory variable or parameters. For example, the strengthening of concrete is a nonlinear process. The strength increases rapidly at the start, and then flattens out. Similarly, marginal cost of production is nonlinear function of the unit produced. Recently, the issue of analysis and model estimation in nonlinear regression is, frequently, addressed by statisticians and scholars because of its growing popularity and application.

© 2017 Pakistan Journal of Statistics

Consider the nonlinear regression model of the form

$$y_{i=}\tau(x_i,\beta)+\varepsilon_i \; i=1,2,3,4 \dots\dots n$$

where x_i is the explanatory variable and y_i is the response variable, τ is a nonlinear function, β denotes the unknown p dimensional vector of parameters, ε_i are independent and identically distributed random variable with mean *0* and unknown variance σ^2. If the model assumptions are satisfied, the parameters can be estimated through ordinary least squares (OLS), which minimizes the sum of square residuals.

In general, there are two types of nonlinear models. In the first type, models can be made linear through transformation. In this method, either independent variable, dependent variable or the combination of both can be transformed. These models are generally called intrinsically linear models or nonlinear in variable models. Usual least square method can be applied to the transformed model in order to get estimates of unknown parameters. Unluckily, this method gives an inaccurate result as the linear regression is applied to transformed data, which may falsify the error term or completely change the relationship between x and y variables. Furthermore, the confidence intervals are no more symmetric. Therefore, this method has become obsolete and should not be applied (Brown, 2001).

The second category includes intrinsically nonlinear models. Such models cannot be transformed by linearization as these models are nonlinear in parameters. For example

Logistic Model: $y=\frac{\theta 1}{1+\theta_2\exp(\theta_3 x)}+\varepsilon$

Asymptote regression model: $y=\theta_1+\theta_2\exp(\theta_3 x)$

Gompertz growth model: $y=\theta_1\exp(\mathrm{\theta_2-\theta_3 x})$

To fit nonlinear models in parameters, iterative optimization procedures are used to compute the parameter estimates. Starting value of the parameters must be provided using a guess or prior experience. Sum of Squares is computed in first iteration using starting values then in the next iteration parameter values are changed by small amount and recalculating the sum of squares, this process is repeated until converge (Brown, 2001).

For iteration, different algorithms are introduced by the researchers. The Marquardt method is used frequently by pharmacological and biochemical researchers. Some other algorithms include steepest descent, Nelder-Mead and Gauss Newton method. Another method is simplex algorithm. In this method, initial value of the parameter and its increment is provided. This method is fast, and easy but it does not provide standard errors of the parameters (Motulsky & Ransnas, 1987).

The problem of model estimation becomes challenging when the data set is contaminated with outliers or an influential observations. An outlier is a data point that is not consistent with the mass of the observations. In a regression model, outlying observations are often recognized as those data points whose residuals are much larger or smaller than the remaining residuals (Gilbert, 2007). In other words, any observation which is not according to the pattern of remaining data set is called an outlier. Such data sets may also be called as contaminated data sets. A standard data set may contain the following three types of outliers; i.e., vertical outliers (outliers in y direction), leverage point (outliers in x direction) and good leverage point (outliers in xy direction) (Rousseeuw & Driessen, 2006; Tabatabai et al., 2014)

Least square criteria is not robust to outliers, hence the results of estimation and hypothesis testing would be ambiguous and unreliable in nonlinear regression (Tabatabai et al., 2014). Robust regression is a statistical method which tries to decrease or eliminate the consequence of outliers so as to attain more reliable results from the bulk of data. In case of multivariate contaminated datasets, robust estimation provides the most appropriate substitute to the classical estimation procedures to balance its sensitivity to outliers (Khalil et al., 2013). Some of the most general robust regression methods for linear regression include M-estimator, S-estimator, redescending M estimator, MM estimator, L estimator which include Least absolute deviation (LAD), least trimmed square (LTS), least median of square (LMS) and least trimmed sum of absolute (LTA), etc.

Most of the robust linear regression techniques are successfully adopted for nonlinear setting. Ekblom and Madsen (1989) proposed Marquardt algorithm for Huber estimator in nonlinear regression. Stromberg and Ruppert (1992) studied the breakdown properties of LTS and LMS in nonlinear regression problem. Verboon (1993) applied M estimators, particularly Huber and Biweight function, for nonlinear model fitting. In the same year Stromberg extended LMS and MM estimator from linear to nonlinear setting. Tabatabai and Argyros (1993) incorporated τ-estimators, originally proposed by Yohai and Zamar for linear regression, to nonlinear regression analysis. Neugebauer (1996) has analyzed the robustness properties of M estimator in nonlinear regression.

Sakata and White (2001) extended S-estimators of linear regression to nonlinear cross section and time series regression. Cizek (2002) worked on LTS and its asymptotic property in nonlinear fitting. Similarly, Hawkins and Khan (2009) proposed an algorithm for LTS estimator in this regard. Abebe and Mckean (2013) used rank based estimator for the said purpose. Rank based estimation was first proposed by Jureckova (1971) and Jaeckel (1972) for linear regression problem. Most recently, a new M estimator based on secant hyperbolic function was proposed for nonlinear case (Tabatabai et al., 2014). This article considers the development of new algorithm for the use and estimation of the parameters in nonlinear regression based on LTA approach.

2. PROPOSED ALGORITHMS IN THE LITERATURE

Least square method is generally adopted if the data is normally distributed. However in the presence of outliers, it totally fails. To remedy this problem, alternative methods have been developed in linear and nonlinear regression, which are more robust to outliers. Among these, least trimmed sum of absolute residuals (LTA) estimator is rarely been used in nonlinear regression despite the fact that LTA estimator is relatively easy as compare to other high breakdown estimators.

Hawkins and Olive (1999) have found a connection between regression outliers and case leverage. Least square estimator is affected by all types of outliers, irrespective of its direction in x space or y space. As compare to this, $L1$ estimator is robust to regression outliers on low leverage cases, but sensitive to outliers on high leverage cases. This makes it a 0 breakdown estimator as OLS. For LTS, one has to choose suitably low values of h so that all outliers can be trimmed. While using LTA, it is usually enough to trim outliers on high leverage case only. So, high values of h can be used for LTA as compared to LTS.

Few algorithms for the estimation of high breakdown estimators are discussed in the next section.

2.1 *Progress Algorithm*

Rosseeuw and Leroy (1987) proposed PROGRESS (Program for Robust reGRESSion) algorithm for the computation of LMS in linear regression case. This algorithm takes into account a trial subset of p values, where p stands for number of parameters, and fits a linear line passing through them. This method is continued many times and the fit for which median of squared residual is minimum is retained. The method is repeated $\binom{n}{p}$ times. For small datasets it is feasible to consider all p subsets but for large datasets it may become tedious. The steps for LMS estimator are given below:

- Calculate exact fit to p points, where p stands for number of parameters, and call it $\tilde{\beta}_{ex}$
- Compute median of squared residuals at $\tilde{\beta}_{ex}$
- Repeat these steps $\binom{n}{p}$ times.
- The value of $\tilde{\beta}_{ex}$ for which median of the residual is lowest, is LMS estimate.

2.2 *Feasible Solution Algorithm*

Hawkins and Olive (1999) derived a feasible solution algorithm (FSA) for LTA in linear regression. It produces perfect approximations to the exact LTA fit. For LTA estimation, $L1$ is fitted to suitable half of the cases. One of its characteristic is that the absolute residual of all cases that it does not include is greater than or equal to the absolute residuals that it includes. The algorithm is as follows:

- Generate random elemental sets of size h, and calculate the residuals.
- If the present elemental set provides the L1 fit to the h cases with the smallest absolute residuals, then it is a feasible solution.
- If not, replace one of the cases in the elemental set with a better one.
- Continue until a feasible solution is obtained.
- Repeat it with the large number of random starts, say t times.
- Use the feasible solution with the smallest sum of absolute residuals.

2.3 *FAST-LTS Algorithm*

This algorithm was proposed by Rousseew and Van Driessen (2006), for the computation of LTS estimator in linear regression estimation problem. Basically, it involves two steps: Initial step and concentration C step. The fundamental part of this algorithm is C-step. It states that starting from initial fit a new subset is taken for which absolute residuals are the smallest, then applying least square to new subset to obtain new fit, which guaranteed lower objective function. It can be summarized as follows:

Initial step:

- A starting value of $\tilde{\beta}$ is generated.
- Using starting values, OLS is applied to all n cases and finding corresponding residuals.
- Order absolute residuals; identify those h cases as best cases to cover for which absolute residuals are the smallest.

C step:

- Apply OLS to the set of covered cases, getting fresh $\tilde{\beta}$. Calculate residuals and discover h smallest absolute residuals and calculate sum of square residuals.
- If the sum of squared residuals has reduced from preceding step, then take these h cases and repeat OLS fit.
- If the residual sum of square is unaffected from the preceding step, then the concentration step ends, and LTS solution is obtained.

2.4 *Direct Conversion to Nonlinear Regression*

The aforementioned algorithms have shown their characteristics in linear regression case and they can be adapted in the case of nonlinear regression. The main dilemma is that the algorithms needed to be solved iteratively for a large number of optimization problems for randomly selected subsets of data. Despite the fact, it does not matter at all in linear regression case as the minimum of criterion function can be found easily and uniquely in some cases. The circumstances are differing in nonlinear regression setting as minimizing the objective function is time consuming, and convergence speed of different algorithms might be different considerably with respect to the behavior of data. Therefore, these algorithms can be used for LTA estimation in nonlinear regression, but it may be comparatively slow and has most likely lower accuracy as compared to linear regression (Cizek, 2001).

PROGRESS algorithm provides exact fit in case of simple linear regression with slope and intercept when all possible subsets $\binom{n}{p}$ are examined. But it fails when line passes through origin as it does not provide exact LMS fit when intercept is zero. The reason is that in PROGRESS algorithm intercept is adjusted given the best slope for exact fit, but as the intercept is zero, this adjustment is not possible, and the resultant fit is local minima rather than global minima (Barreto & Maharry, 2006). This failure of PROGRESS algorithm in linear case can be encountered in nonlinear setting as well since most of the nonlinear models may not include intercept term.

Furthermore, FSA algorithm can be adopted for LTA estimation in nonlinear regression after some adjustment, but FSA in nonlinear case is not as favorable as in linear regression. The weakness is that the number of random starts t depends upon three factors; the proportion of outliers, complexity of model and sample size. The value of t will be increased as long as any one of the factors increases, but the amount of increase is not obvious (Chen et al., 1997). According to Hawkins and Olive (1999), the performance of FSA is superb for text book size problems, but for big data sets, its performance is not satisfactory on efficiency grounds. Additionally, FAST algorithm may take a lot of time in nonlinear setting investigating poor elemental fits as it requires significantly more computation in nonlinear than linear setting (Hawkins & Khan, 2009).

2.5 *A Hybrid FAST Algorithm*

As discussed above, there are many drawbacks of directly adapting these algorithms to nonlinear case. For fitting LTS in nonlinear regression, Hawkins and Khan(2009) proposed a Hybrid FAST algorithm. In this paper, the best properties of FAST-LTS and PROGRESS algorithm are combined to get an efficient and fast algorithm. Their algorithm is as follows:

- Outer loop, executed I times
  - Inner loop, executed J times
  - Create an elemental set and compute elemental fit to get starting values.
  - Use these initial values to compute residuals on all n cases and find the sum of squares of h smallest absolute residuals. Keeping track of such smallest sum, the cases giving the smallest absolute residuals, and corresponding elemental fit.
- Take this best-of-J elemental fits. Use its covered cases and elemental fit as starting values for a concentration step using a usual nonlinear regression procedure.
- The final reported solution is the best of the resulting I outer loops.

3. PROPOSED ALGORITHM FOR ESTIMATION OF NLRM

The proposed algorithm is based on the idea of Hybrid FAST algorithm where the best features of the PROGRESS and FAST-LTS are implemented in the nonlinear regression setting. In the proposed algorithm, FAST-LTS approach is substituted with FAST-LTA keeping in view the high efficiency of LTA as compared to FAST-LTS approach. The steps for the proposed Hybrid FAST-LTA are as follows:

- Generating M elemental sets and its elemental fits.
- Using elemental fits for the calculation of residuals on all n cases and find sum of h smallest absolute residuals.
- Out of M times, identify h smallest absolute residuals, its corresponding cases and elemental fits.
- Using this elemental fit as starting values for concentration step.
- Repeating step 1 to 4, N times.
- Out of N, best reported candidate is obtained as Hybrid FAST-LTA solution.

The performance of the proposed Hybrid FAST-LTA will be assessed through simulation study to be conducted in next section.

4. SIMULATION STUDY

Simulations study in this section is performed using R software. Michealis-Menten model is used for three different sample sizes $(n = 30, n = 50, n = 100)$ and different proportions of outliers (0%, 20%, and 40%) were considered. Each model is repeated 500 times, and mean square error and bias are calculated for LTA, LS and M estimate.

4.1 Michaelis-Menten Model

In biochemistry, Michaelis-Menten model is used to study the relationship between reaction velocity V and concentration of substrate S as

$$V = \frac{V_{max}[S]}{K_M + [S]}$$

In this model, V_{max} is a constant representing maximum velocity at saturation of substrate concentration. Michaelis constant K_M shows concentration of substrate at which reaction time is 50% of V_{max}. It also offers substrate concentration measure for significant catalysis. Furthermore, the parameter K_M increases as the efficiency between substrate and enzymes decreases (Berg et al., 2002).

Consider two parameter's Michaelis-Menten model

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

The data has been simulated by taking a sample from Puromycin data set used in Batts and Watts (1988), where $\beta_0 = 112.48$, $\beta_1 = 0.0096$ are the parameter values; using these values a sample size of 50, 100 and 500 were generated with x values, uniformly spaced over the range (0.02,1.1). Normal random errors with mean 10 and standard deviation 5 were added to the regression model. The data generation process has been adopted from Hawkins and Khan (2009). Then 0%, 20% and 40% outliers were incorporated in y-direction. The results of least square LS, M estimate and proposed least trimmed absolute residuals LTA are exhibited in Table 1.

Table 1
Mean Square Errors (MSE) and Bias for Michaelis-Menten Model with Percentage Contamination

Sample size	Procedure	MSE/BISE	0%		20%		40%	
			β_0	β_1	β_0	β_1	β_0	β_1
n=30	Least Squares	MSE	25.09	0.000016	993.77	0.0124	4800.79	0.068
		BIAS	4.68	0.0023	31.26	0.11	69.02	0.26
	M Estimator	MSE	24.94	0.000016	54.62	0.00017	6744.67	0.097
		BIAS	4.65	0.0024	6.92	0.0114	81.78	0.31
	Hybrid LTA	MSE	36.51	0.000055	36.38	0.000048	466.27	0.0077
		BIAS	4.65	0.0024	4.83	0.0021	7.12	0.0085
n=50	Least Squares	MSE	22.92	0.0000108	966.51	0.012	4736.42	0.067
		BIAS	4.61	0.0023	30.92	0.109	68.66	0.2588
	M Estimator	MSE	22.93	0.000013	51.57	0.00015	6672.05	0.094
		BIAS	4.54	0.0023	6.91	0.0114	81.47	0.307
	Hybrid LTA	MSE	30.77	0.000042	28.66	0.0000312	25.92	0.000025
		BIAS	4.46	0.0025	4.54	0.0025	4.55	0.0025
n=100	Least Squares	MSE	22.48	0.00001	962.97	0.0119	4739.75	0.06
		BIAS	4.65	0.0023	30.96	0.108	68.75	0.25
	M Estimator	MSE	22.48	0.00001	50.875	0.00014	6698.97	0.093
		BIAS	4.64	0.0023	7.011	0.115	81.72	0.31
	Hybrid LTA	MSE	28.32	0.000025	26.56	0.000022	24.19	0.000015
		BIAS	4.65	0.0024	4.63	0.0023	4.59	0.024

It can be seen clearly from Table 1 that with clean data LS, M and LTA estimator give similar results for small as well as for large samples. As 20% outliers are incorporated in x direction, proposed method outperforms the least square estimator, but produced the same results as that of existing robust methods. While increasing the level of contamination up to 40%, our proposed estimator improves on its other counterpart.

4.2 Gompertz Model

Gompertz model was initially developed by Benjamin Gompertz in 1832 to fit human mortality data (Missov et al., 2015). Later on, it was successfully used in predicting cancer tumor growth. It is a three parameter model given by the following relationship

$$y_i = \beta_0 + \exp[-\beta_1(-\beta_2 x_i)] + \varepsilon_i$$

In this model, x_i is the tumor size and y_i is the fraction of breast cancer people with metastases. The data has been simulated by the pattern of Tumor growth data used in Tabatabai et al. (2014). Where $\beta_0 = 25.67$,$\beta_1 = 2.58$ and $\beta_2 = 0.036$ are the parameter values. Using these values a sample size of 30, 50, and 100 were generated with x values uniformly spaced over the range (20,160). Normal random errors with mean 1 and standard deviation 0.5 were added to the regression model. Then 0%, 20% and 40% outliers were incorporated in xy-direction. The results of least square LS, M estimator and proposed Hybrid least trimmed absolute residuals LTA are exhibited in Table 2.

Table 2
Mean Square Errors (MSE) and Bias for Gompertz Model with Percentage Contamination

N	Sample Size		0%			20%			40%		
			β_0	β_1	β_2	β_0	β_1	β_2	β_0	β_1	β_2
n=30	LS	MSE	1.12	0.0535	2.45e-6	32.09	0.43	0.00033	105.59	0.0749	0.00046
		BIAS	1.042	0.196	0.00066	5.63	0.65	0.018	10.23	0.266	0.0214
	M Estimator	MSE	1.12	0.055	2.62e-6	3.49	0.29	6.49e-5	105.59	0.0749	0.00046
		BIAS	1.04	0.198	0.00697	1.82	0.53	0.0077	10.23	0.266	0.0214
	Hybrid LTA	MSE	1.99	0.297	2.3e-5	1.71	0.22	1.45e-5	1.38	0.078	5.41e-6
		BIAS	1.34	0.489	0.0038	1.26	0.40	0.0029	1.14	0.24	0.0013
n=50	LS	MSE	1.09	0.05	1.88e-6	97.496	0.22	0.00047	108.39	0.078	0.00046
		BIAS	1.03	0.203	0.007	9.84	0.46	0.022	10.38	0.28	0.02
	M Estimator	MSE	1.10	0.05	1.93e-6	3.44	0.325	6.67e-5	108.39	0.078	0.00046
		BIAS	1.03	0.203	0.0007	1.83	0.56	0.008	10.38	0.28	0.02
	Hybrid LTA	MSE	1.87	0.24	1.73e-5	1.38	0.106	6.96e-6	1.26	0.084	4.76e-6
		BIAS	1.298	0.456	0.0034	1.13	0.284	0.0016	1.09	0.252	0.0013
n=100	LS	MSE	1.088	0.043	1.12e-6	33.71	0.45	0.00034	110.05	0.075	0.0005
		BIAS	1.037	0.196	0.0006	5.79	0.667	0.018	10.47	0.27	0.02
	M Estimator	MSE	1.086	0.044	1.16e-6	3.45	0.317	6.55e-5	110.05	0.075	0.0005
		BIAS	1.036	0.196	0.00065	1.84	0.56	0.008	10.47	0.27	0.02
	Hybrid LTA	MSE	1.65	0.204	1.27e-5	1.387	0.145	7.59e-6	1.175	0.069	3.16e-6
		BIAS	1.25	0.425	0.0031	1.15	0.346	0.0022	1.065	0.233	0.001

It is apparent from the results displayed in Table 2 that with clean data set LS, M estimate gives minimum mean square error (MSE) and bias, while Hybrid LTA estimator's MSE and bias is a bit higher than both for small sample as well as for increasing sample size. It is concluded that LS and M estimate are efficient in case of clean data. Perturbing the data with 20% outliers, M estimate and Hybrid LTA shows better performance than LS. While in case of 40% contamination, our proposed estimator beats its other counterparts.

5. APPLICATIONS ON REAL DATASETS

The performance of Hybrid LTA is illustrated using nonlinear examples from the field of biochemistry and medicines.

5.1 Michaelis-Menten Model

Michaelis-Menten model, used by Stromberg (1993) and Tabatabai et al. (2014), 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 β_0 is the maximum reaction velocity and β_1 denotes concentration of substrate. The Puromycin data, taken from Bates and Watts (1988) which was primarily used by Treloar (1974), consists of 12 observations given below in Table 3.

Table 3
Concentration of Substrate versus Reaction Velocity

Concentration Ppm	**0.02**	**0.02**	**0.06**	**0.06**	**0.11**	**0.11**	**0.22**	**0.22**	**0.56**	**0.56**	**1.10**	**1.10**
Velocity Counts/min	**76**	**47**	**97**	**107**	**123**	**139**	**159**	**152**	**191**	**201**	**207**	**200**

The speed of an enzymatic reaction depends on the concentration of a substrate. As outlined in Bates and Watts (1988), an experiment was conducted to inspect how a treatment of the enzyme with an additional substance called Puromycin affects the reaction speed. To check the performance of our proposed estimator, with least square and M estimator, model is fitted on clean data as well as contaminated data. Twenty percent outliers are incorporated in *X* space, *Y* space and both *XY* space. In *X* space, shifting the *X* value in observation 5,6,7 from 0.11, 0.11, 0.22 to 2, 2.01, 2.05. In the *Y* space, changing the *Y* value in observation 10, 11,12 from 201, 207, 200 to 65, 50, 60. And in both *XY* space, transforming *Y* value in observation 6, 8, 11 from 139, 191, 207 to 85, 100, 80 and *X* value in observation 6,8,11 from 0.11, 0.56, 1.10 to 0.65, 0.61, 0.6. Table 3 shows standard error of estimate (SE) and residuals standard errors (RSE) for LS, M and Hybrid LTA estimators. By examining the SE and RSE, it is clear that Hybrid LTA is more efficient than its counterpart. The results are also presented through graphs in figure 1 in which (a) presents clean data (b) shows the effect of outliers in *X* space (c) demonstrates the pattern of contamination in *XY* space and (d) displays the consequences of outliers in both *Y* space. From these graphs, it is apparent that in clean data set all three estimators perform well but when outliers enter in X,Y and XY space, the least square becomes unacceptable whereas M estimator and Hybrid LTA execute well.

Table 4
Standard Errors of Estimate and Residuals Standard Errors

Procedure	**Clean Data**			**Outliers in X space**			**Outliers in Y space**			**Outliers in XY space**		
	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	**RSE**	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	**RSE**	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	**RSE**	$\boldsymbol{\beta_0}$	$\boldsymbol{\beta_1}$	**RSE**
Least Squares	6.94	0.008	10.93	12.2	0.015	29.2	20.0	0.014	46.7	19.20	0.018	41.3
M Estimator	6.45	0.008	8.500	13.8	0.017	30.03	24.2	0.017	51.9	24.35	0.022	34.9
Hybrid LTA	2.65	0.004	3.390	3.9	0.005	5.324	7.55	0.005	6.31	6.32	0.006	6.09

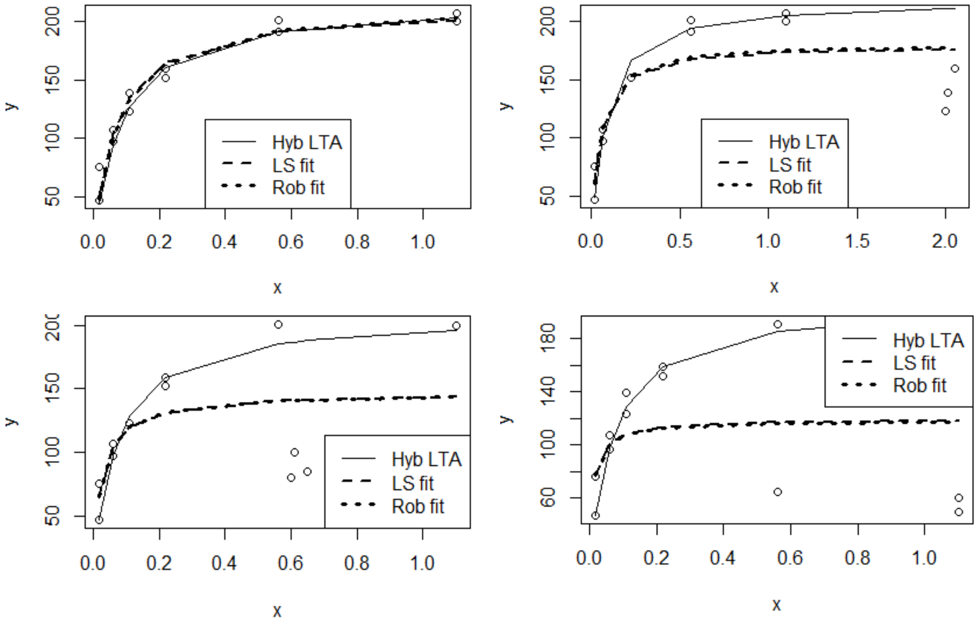

Figure 1: Michaelis-Menten Model; (a, b, c, d)

5.2 Gompertz Model

Gompertz model is a sigmoid function that is the slowest at the ends and fastest at the middle. For the first time, it was successfully used by Laird (1964) to analyze the growth of cancer tumor. Gompertz Model is of the form:

$$y_i = \theta_1 + \exp[-\theta_2(-\theta_3 x_i)] + \varepsilon_i$$

It is frequently used for screening cancer regression and progression. The data in the Table 5 comprise of 12 observations. This data is taken from Tabatabai et al. (2014) which was primarily collected by Tubiana & Koscielny (1990). The given data is clean as there is no outlier in it. Model is fitted on this data using Hybrid LTA estimator, M estimator and least square. Then outliers are inserted in *X* direction, both *XY* direction and in *Y* direction. In *X* space shifting the *X* value in observation 12 from 90 to 9. In the *Y* space changing the *Y* value in observation 6 from 0.55 to 3. And in both *XY* space transforming *Y* value in observation 7 from 0.56 to 3 and *X* value in observation 12 from 90 to 2. Table 6 shows standard error of estimate (SE) and residuals standard errors (RSE) for LS, M and Hybrid LTA estimators. By examining the SE and RSE, it is clear that Hybrid LTA is more efficient than its other counterparts. The results are also presented through graphs in figure 2 in which (a) presents clean data (b) shows the effect of outliers in *X* space (c) demonstrates the pattern of contamination in *XY* space and (d) displays the consequences of outliers in both *Y* space. From these graphs it is apparent that in clean data set all three estimators perform well, but when outliers enter in *X*,*XY* and *Y* space, the least square becomes unacceptable whereas M estimator and Hybrid LTA executes well.

Table 5
Tumor Size Versus Fraction Metastasized Data

Tumor Size x	**12**	**17**	**17**	**25**	**30**	**39**	**40**	**50**	**60**	**70**	**80**	**90**
Fraction Metastasized y	**0.13**	**0.20**	**0.27**	**0.45**	**0.42**	**0.55**	**0.56**	**0.66**	**0.78**	**0.83**	**0.81**	**0.92**

Table 6
Standard Errors of Estimate and Residuals Standard Errors

	Clean data				Outliers in X space				Outliers in XY space				Outliers in Y space			
	θ_1	θ_2	θ_3	RSE	θ_1	θ_2	θ_3	RSE	θ_1	θ_2	θ_3	RSE	θ_1	θ_2	θ_3	RSE
LS	0.05	0.40	0.007	0.044	2.3e6	3.03e2	4.1e-2	0.224	0.65	3.3	0.13	0.8	0.3	279	0.3	0.73
M	0.05	0.40	0.007	0.040	0.06	0.408	0.008	0.040	0.08	0.7	0.01	0.06	0.6	0.7	0.01	0.06
LTA	0.01	0.11	0.001	0.008	0.03	0.174	0.003	0.012	0.03	0.4	0.004	0.02	0.02	0.16	0.002	0.01

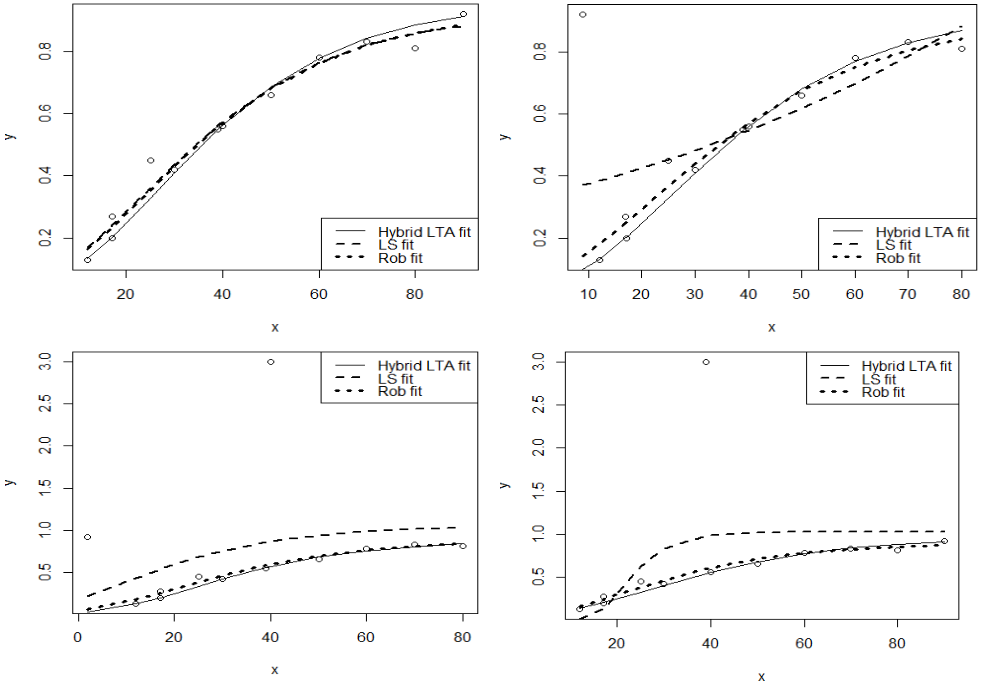

Figure 2: Gompertz Model (a,b,c,d)

6. CONCLUSIONS

In this study, Hybrid FAST LTA estimator has been presented for univariate nonlinear models. It is hoped that this study offers an introduction to Hybrid LTA estimation in nonlinear regression. It makes the nonlinear robust fitting possible in usual practice. It is expected that this algorithm may be adopted for other nonlinear robust fitting procedures, for instance S estimators, M estimators and rank based estimators.

It has been concluded that Hybrid LTA estimator performs well in contrast to LS and M estimator in case of contaminated data set, with level of contamination upto 40 percent. Through simulation study, it has been shown that Mean Square Error and Bias of Hybrid LTA is less than LS and M estimator when the data is polluted with 40% outliers, and for 20% outliers Hybrid LTA and M estimators are equally efficient to LS. In regard to data sets, with 0% contamination LS performs well as compared to Hybrid LTA estimator.

The introduction of robust techniques and its applications to real data has made high breakdown estimation procedures more attractive to researchers rather than sensitive procedures. Applications to real world data such as Puromycin data and Tumor growth data show that Hybrid LTA estimator can also be used in diverse fields of medical sciences.

REFERENCES

1. Abebe, A. and McKean, J.W. (2013). Weighted Wilcoxon Estimators in Nonlinear Regression. *Australian & New Zealand Journal of Statistics*, 55(4), 401-420.
2. Barreto, H. and Maharry, D. (2006). Least median of squares and regression through the origin. *Computational Statistics & Data Analysis*, 50(6), 1391-1397.
3. Batts, D.M. and Watts, D.G. (1988). *Nonlinear Regression Analysis and its Application*. John Wiley & Sons, New York.
4. Berg, J.M., Tymoczko, J.L., Stryer, L. and Stryer, L. (2002). *Biochemistry*. New York: W.H. Freeman.
5. Brown, A. (2001). A step-by-step Guide to Non-linear Regression Analysis of Experimental Data using a Microsoft Excel Spreadsheet. *Computer Methods and Programs in Biomedicine*, 65(3), 191-200.
6. Chen, Y., Stromberg, A.J. and Zhou, M. (1997). *The least trimmed squares estimate in nonlinear regression*. Technical Report Department of Statistics, University of Kentucky, Lexington, KY, 40506.
7. Cizek, P. (2001). Nonlinear least trimmed squares. *SFB Discussion Paper, Humboldt University*, 78-86.
8. Cizek, P. (2002). Robust estimation in nonlinear regression and limited dependent variable models. *CERGE-EI Working Paper*, 189.
9. Ekblom, H. and Madsen, K. (1989). Algorithms for non-linear Huber estimation. *BIT Numerical Mathematics*, 29(1), 60-76.
10. Gilbert, S. (2007). Using SAS Proc NLMIXED for Robust Regression. *Statistics and Data Analysis*, 181, 1-9.
11. Hawkins, D.M. and Olive, D.J. (1999). Improved feasible solution algorithms for high breakdown estimation. Computational Statistics and Data Analysis, 30(1), 1-11.
12. Hawkins, D. and Khan, D. (2009). A procedure for robust fitting in nonlinear regression. *Computational Statistics and Data Analysis*, 53(12), 4500-4507.
13. Hawkins, D. and Olive, D. (1999). Applications and Algorithms for Least Trimmed Sum of Absolute Deviation Regression. *Computational Statistics and Data Analysis*, 32(2), 119-134.
14. Jaeckel, L. (1972). Estimating Regression Coefficients by Minimizing the Dispersion of the Residuals. *The Annals of Mathematical Statistics*, 43, 1449-1458.
15. Jureckova, J. (1971). Nonparametric estimate of regression coefficients. *Annals of Mathematical Statistics*, 42(4), 1328-1338.
16. Khalil, A., Ali, A., Khan, S., Khan, D. M. and Khalil, U. (2013). A New Efficient Redescending M-Estimator: Alamgir Redescending M-Estimator. *Research Journal of Recent Sciences*, 2(8), 79-91.
17. Laird, A.K. (1964). Dynamics of Tumor Growth. *British Journal of Cancer*, 19, 278-291.
18. Missov, T.I., Lenart, A., Nemeth, L., Canudas-Romo, V. and Vaupel, J.W. (2015). The Gompertz force of mortality in terms of the modal age at death. *Demographic Research*, 32, 1031-1048.
19. Motulsky, H. and Ransas, L.A. (1987). Fitting Curves to Data using Nonlinear Regression: A practical and Nonmathematical Review. *Official Publication of the Federation of American Societies for Experimenal Biology*, 1(5), 365-374.

20. Neugebauer, S.P. (1996). *Robust analysis of M-estimators of nonlinear models.* Unpublished Doctoral dissertation, Virginia Tech.
21. Rousseeuw, P. and Driessen, V. (2006). Computing LTS Regression for Large Data Sets. *Data Mining and Knowledge Discovery*, 12(1), 29-45.
22. Rousseeuw, P. and Leroy, A. (1987). *Robust Regression and Outlier Regression.* New York: Wiley.
23. Sakata, S. and white, H. (2001). S-Estimation of Nonlinear Regression Models with Dependent and Heterogeneous Observation. *Journal of Econometrics*, 103(1), 5-72.
24. Stromberg, A.J. (1993). Computation of high breakdown nonlinear regression parameters. *Journal of the American Statistical Association*, 88(421), 237-244.
25. Stromberg, A. and Rupert, D. (1992). Breakdown in Nonlinear Regression. *Journal of the American Statistical Association*, 87(420), 991-997.
26. Tabatabai, M. and Argyros, I. (1993). Robust Estimation and Testing for General Nonlinear Regression Methods. *Applied Mathematics and Computation*, 57(1), 85-101.
27. Tabatabai, M., Kengwoung-Keumo, J., Eby, W., Manne, U., Fouad, M. and Singh, K. (2014). A New Robust Method for Nonlinear Regression. *Journal of Biometrics & Biostatistics*, 5(5), 211.
28. Treloar, M.A. (1974). *Effects of Puromycin on Galactosyltransterase in Golgi Membranes*. M.Sc. Thesis, University of Toronto.
29. Tubiana, M. and Koscielny, S. (1990). The Natural History of Human Breast Cancer: Implications Fora Screening Strategy. *International Journal of Radiation Oncology* Biology* Physics*, 19(5), 1117-1120.
30. Verboon, P. (1993). Robust nonlinear regression analysis. *British Journal of Mathematical and Statistical Psychology*, 46(1), 77-94.