

## A New Robust Method for Nonlinear Regression

Tabatabai MA<sup>1</sup>, Kengwoung-Keumo JJ<sup>2</sup>, Eby WM<sup>3</sup>, Bae S<sup>4</sup>, Manne U<sup>5</sup>, Fouad M<sup>4</sup> and Singh KP<sup>\*\*</sup>

<sup>1</sup>School of Graduate Studies and Research, Meharry Medical College, Nashville, TN 37208, USA

<sup>2</sup>Department of Mathematical Sciences, Cameron University, Lawton, OK 73505, USA

<sup>3</sup>Department of Mathematics, New Jersey City University, Jersey City, NJ 07305, USA

<sup>4</sup>Department of Medicine Division of Preventive Medicine and Comprehensive Cancer Center, University of Alabama Birmingham, Birmingham, AL 35294, USA

<sup>5</sup>Department of Pathology and Comprehensive Cancer Center, University of Alabama Birmingham, Birmingham, AL 35294, USA

### Abstract

**Background:** When outliers are present, the least squares method of nonlinear regression performs poorly. The main purpose of this paper is to provide a robust alternative technique to the Ordinary Least Squares nonlinear regression method. This new robust nonlinear regression method can provide accurate parameter estimates when outliers and/or influential observations are present.

**Method:** Real and simulated data for drug concentration and tumor size-metastasis are used to assess the performance of this new estimator. Monte Carlo simulations are performed to evaluate the robustness of our new method in comparison with the Ordinary Least Squares method.

**Results:** In simulated data with outliers, this new estimator of regression parameters seems to outperform the Ordinary Least Squares with respect to bias, mean squared errors, and mean estimated parameters. Two algorithms have been proposed. Additionally and for the sake of computational ease and illustration, a Mathematica program has been provided in the Appendix.

**Conclusion:** The accuracy of our robust technique is superior to that of the Ordinary Least Squares. The robustness and simplicity of computations make this new technique more appropriate and useful tool for the analysis of nonlinear regressions.

**Keywords:** Robust nonlinear regression; Least Square estimator; Growth models; Tumor size; Metastasis; Monte-carlo simulation

### Background

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 biomedical research. 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 robust nonlinear regression is to fit a model to the data that gives resilient results in the presence of influential observations, leverage points and/or outliers. Rousseeuw and Leroy [1] 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 [2] proposed the Least Absolute Deviation as a robust method. Huber [3] introduced the method of M-estimation. Rousseeuw [4] introduced the Least Trimmed Squares-estimates. The S-estimator was introduced by Rousseeuw and Yohai [5]. Yohai and Zammar [6] introduced the  $\tau$ -estimator of linear regression coefficients. It is a high efficiency estimator and has a high breakdown point. Tabatabai and Argyros [7] extended the  $\tau$ -estimates to the nonlinear regression models. Stromberg [8] introduced algorithms for Yohai's MM estimator of nonlinear regression and Rousseeuw's least median estimators of nonlinear regression. Tabatabai et al. [9] introduced the TELBS robust linear regression method.

In Medical, biological and pharmaceutical research and development nonlinear regression analysis has been a major tool for investigating the effect of multiple explanatory variables on a response variable when the data follows a nonlinear pattern. When outliers and influential observations are present, nonlinear least squares performs poorly. In this paper we introduce a new robust nonlinear regression method capable of handling such cases. Minn et al. [10] showed that

lung tumor size can lead to metastasis. Also, aggressive tumor growth is a marker for cells destined to metastasize. They validated their statement by analyzing the lung metastasis gene-expression signature using a nonlinear model. Arisio et al. [11] study of breast cancer confirmed that the size of tumor is an important predictor of axillary lymph node metastases. Ramaswamy, et al. [12]) found that gene-expression signature is a significant factor associated with metastasis in solid tumors carrying such gene expressions. Maffuz et al. [13] showed that pure ductal carcinoma in situ is not associated with lymphatic metastasis independently of tumor size. Hense et al. [14] found that the occurrence and primary metastases in Ewing tumors is related to tumor size, pelvic site and malignant peripheral neuroectodermal tumors. Umbreit et al. [15] studied a group of patients who had undergone surgical resection for a unilateral, sporadic renal tumor. They concluded that tumor size is significantly associated with metastasis in patients suffering from renal masses. Wu et al. [16] retrospectively analyzed 666 patients with nasopharyngeal carcinoma and concluded that the tumor volume was correlated with cervical lymph node metastasis as well as distant metastasis after radiation therapy. In computer vision, robust regression methods have been used extensively to estimate surface model parameters in small image regions and imaging geometry of multiple cameras. Coras et al. [17] used nonlinear regression and

**\*Corresponding author:** Singh KP, Department of Medicine Division of Preventive Medicine and Comprehensive Cancer Center, University of Alabama Birmingham, Birmingham, AL 35294, USA, Tel: +205-934-6887; E-mail: [kpsingh@uab.edu](mailto:kpsingh@uab.edu)

**Received** June 30, 2014; **Accepted** September 10, 2014; **Published** September 18, 2014

**Citation:** Tabatabai MA, Kengwoung-Keumo JJ, Eby WM, Bae S, Manne U, et al. (2014) A New Robust Method for Nonlinear Regression. J Biomet Biostat 5: 211. doi:10.4172/2155-6180.1000211

**Copyright:** © 2014 Tabatabai MA, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

showed that micromolar doses of peroxisome proliferator-activated receptor  $\gamma$  reduce glioma cell proliferation.

Roth [18] applied nonlinear sigmoidal curves to monitor the accumulation of polymerase chain reaction products at the end of each cycle by fluorescence. In human blood samples, Kropf et al. [19] found a nonlinear binding association between transforming growth factor beta1 (TGF- $\beta_1$ ) and  $\alpha_2$ -Macroglobulin as well as TGF- $\beta_1$  and latency-associated peptide (LAP). Yang and Richmond [20] used nonlinear least squares to estimate the effective concentration of unlabeled human interferon-inducible protein 10 that yields 50% maximal binding of iodinated protein 10 to chemokine receptor CXCR3. Hao et al. [21] examined the significance of Nav1.5 protein in cellular processes by applying a nonlinear regression which relates the gene expression of Nav 1.5 protein and TGF- $\beta_1$  as well as Nav 1.5 protein and vimentin. TGF- $\beta$  families are important factors in regulation of tumor initiation, progression, and metastatic activities, Bierie, et al. [22]. Coras et al. [17] applied nonlinear regression models to show that traglitazone concentration has a tendency to inhibit 1 TGF- $\beta_1$  release in glioma cell culture.

This paper introduces a new robust nonlinear regression estimator. This new method for robust nonlinear regression has a bounded influence and high breakdown point and asymptotic efficiency under normal distribution and is able to estimate the parameters of nonlinear regression in such a way that is close to the parameter estimates we would have estimated with the absence of outliers in the data. In addition, this new robust nonlinear regression method is computationally simple enough to be used by practitioners.

### Methods and Models

We begin with the introduction of our new robust nonlinear regression model. The introduction of the model is followed by two algorithms describing its implementation. We then apply this new model to a real data set with an outlier present. In addition, we will analyze a problem involving tumor size and metastases with and without outliers. Monte Carlo simulations are also performed to evaluate the robustness of our method, in comparison with the ordinary least squares method.

#### Robust nonlinear regression model

Consider the general nonlinear model of the form

$$y_i = g(\theta; x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n$$

Where  $y_1, y_2, \dots, y_n$  is a sample of  $n$  observations with  $k$  predictor variables in the model and the parameter vector  $\theta = (\theta_1, \theta_2, \dots, \theta_p)$ . The errors  $\varepsilon_i$ 's are random variables. In a designed experiment,  $x_{ij}$ 's are fixed but when  $x_{ij}$ 's are observational, they are random variables. The predictor can be fixed, random, or mixed. The ordinary least squares estimate of the parameter vector  $\theta$  is given by

$$\hat{\theta} = \operatorname{argmin}_{\theta \in R^p} \sum_{i=1}^n \varepsilon_i^2,$$

and the robust estimate of the parameter vector  $\theta$  is derived by

$$\hat{\theta} = \operatorname{argmin}_{\theta \in R^p} \sum_{i=1}^n \frac{P_\omega(t_i)}{L_i}, \tag{1}$$

where the function  $\rho_\omega(x)$  is defined as

$$\rho_\omega(x) = 1 - \operatorname{Sech}(\omega x),$$

and the positive real number  $\omega$  is called the tuning constant. The function  $\operatorname{Sech}(\cdot)$  is the hyperbolic secant function and  $t_i$ 's are defined by

$$t_i = \frac{(1 - h_{ii})(y_i - g(\theta; x_i))}{\sigma}, \tag{2}$$

where  $\sigma$  is the error standard deviation and  $h_{ii}$ 's are the diagonal elements of the matrix  $H$  of the form

$$H = \operatorname{Diag}[G(G'G)^{-1}G']$$

where the matrix  $G$  is defined as

$$G = \begin{pmatrix} \frac{\partial g(\theta; x_1)}{\partial \theta_1} & \frac{\partial g(\theta; x_1)}{\partial \theta_2} & \dots & \frac{\partial g(\theta; x_1)}{\partial \theta_p} \\ \frac{\partial g(\theta; x_2)}{\partial \theta_1} & \frac{\partial g(\theta; x_2)}{\partial \theta_2} & \dots & \frac{\partial g(\theta; x_2)}{\partial \theta_p} \\ \frac{\partial g(\theta; x_n)}{\partial \theta_1} & \frac{\partial g(\theta; x_n)}{\partial \theta_2} & \dots & \frac{\partial g(\theta; x_n)}{\partial \theta_p} \end{pmatrix} = (G_1 G_2 \dots G_p).$$

For  $j = 1, 2, \dots, k$ , we define

$$M_j = \operatorname{Median}\{|x_{1j}|, |x_{2j}|, \dots, |x_{nj}|\}$$

and for  $i = 1, 2, \dots, n$ , we define

$$L_i = \sum_{j=1}^k \operatorname{Max}\{M_j, |x_{ij}|\}$$

If  $\sigma$  is unknown, one may use one of the following two estimators of  $\sigma$  which were proposed by Rousseeuw and Croux [23].

$$\hat{\sigma} = 1.1926 \operatorname{Median}(\operatorname{Median}_{\{i:1 \leq i \leq n\}} |r_i - r_j|)$$

$$\text{or } \hat{\sigma} = 2.2219 \{ |r_i - r_j|; i < j, i, j = 1, \dots, n \}_{(l)}, \tag{4}$$

where  $r_i = y_i - g(\hat{\theta}; x_i)$  and  $l = \binom{[n/2]+1}{2}$  is the binomial coefficient

and  $\{.\}_{(l)}$  is the  $l$ th order statistic.

The above estimators of  $\sigma$  have high breakdown points. Under the normality assumption for error terms, the estimators given in (3) and (4) have higher efficiency than median absolute deviation (MAD). In this paper all of our computations are performed using formula (3).

The function  $\rho_\omega: \mathbf{R} \rightarrow \mathbf{R}$  is a differentiable function satisfying the following properties:

- i.  $\rho_\omega(0) = 0$ ,
- ii.  $\rho_\omega$  is bounded,
- iii.  $\forall x \in \mathbf{R}, \rho_\omega(x) \geq 0$ ,
- iv.  $\forall x \in \mathbf{R}, \rho_\omega(x) = \rho_\omega(-x)$ ,
- v.  $\lim_{x \rightarrow \infty} \rho_\omega(x) = \lim_{x \rightarrow -\infty} \rho_\omega(x) = 1$ ,
- vi.  $\forall a, b \in \mathbf{R}, |a| > |b| \Rightarrow \rho_\omega(a) \geq \rho_\omega(b)$ ,
- vii.  $\lim_{|x| \rightarrow \infty} \frac{d\rho_\omega(x)}{dx} = 0$ .
- viii.  $\forall \kappa > 0, \lim_{x \rightarrow \infty} \frac{\rho_\omega(\kappa x)}{\rho_\omega(x)} = 1$ .

Taking the partial derivatives of (1) with respect to parameters and setting them equal to zero results in the following system of equations

$$\sum_{i=1}^n \frac{\psi_{\omega}(t_i)}{L_i} \frac{\partial t_i}{\partial \beta_j} = 0, j = 1, 2, \dots, p, \tag{5}$$

where  $\psi_{\omega}$  is the derivative of  $\rho_{\omega}$  which is equal to

$$\psi_{\omega}(x) = \omega \text{Sech}(\omega x) \text{Tanh}(\omega x)$$

Define the weights  $w_i$  as

$$w_i = \frac{\psi_{\omega}(t_i)(1-h_i)}{\sigma(y_i - g(\theta; x_i))L_i} \tag{6}$$

Then for  $j = 1, 2, \dots, p$ , the equation (5) can be written as

$$\sum_{i=1}^n w_i (y_i - g(\theta; x_i)) \frac{\partial g(\theta; x_i)}{\partial \theta_j} = 0.$$

The matrix of weights,  $W$  is a diagonal matrix whose elements on the main diagonals are  $w_1, w_2, \dots, w_n$ , and the estimator of the parameter vector  $\theta$  is given by

$$\widehat{\theta}(X, y) = (G^T W G)^{-1} G^T W y.$$

If  $g(\theta; x_i)$  is linear function of parameters, then the above model would be identical to TELBS robust linear regression model. Asymptotically,  $\hat{\theta}$  has a normal distribution with mean  $\theta$  and variance-covariance matrix of the form

$$V = \frac{\sigma^2 E(\psi_{\omega}^2(t))}{[E(\psi_{\omega}'(t))]^2} E((G^T G)^{-1}),$$

where

$$E[\psi_{\omega}'(t)] = \int_{-\infty}^{\infty} \psi_{\omega}'(t) \frac{e^{-t^2}}{\sqrt{2\pi}} dt$$

and

$$E[\psi_{\omega}^2(t)] = \int_{-\infty}^{\infty} \psi_{\omega}^2(t) \frac{e^{-t^2}}{\sqrt{2\pi}} dt.$$

The function  $\psi_{\omega}'(t)$  is defined as

$$\psi_{\omega}'(t) = \omega^2 [\text{Sech}^3(\omega t) - \text{Sech}(\omega t) \text{Tanh}^2(\omega t)].$$

Under the assumption of normality for the underlying distribution, the asymptotic efficiency,  $A_{eff}$ , is defined as

$$A_{eff} = \frac{(E[\psi_{\omega}'(t)])^2}{E[\psi_{\omega}^2(t)]} \tag{7}$$

The tuning constant  $\omega$  can be calculated by solving equation (7) for  $\omega$ .

An estimate for the variance-covariance matrix is derived and given as follows

$$\hat{V} = \frac{n^2 \hat{\sigma}^2 \sum_{i=1}^n \psi_{\omega}^2(t_i)}{(n-p) \left( \sum_{i=1}^n \psi_{\omega}'(t_i) \right)^2} (G^T G)^{-1}.$$

The robust deviance is defined as

$$D = 2\hat{\sigma}^2 \sum_{i=1}^n \frac{1 - \text{Sech}(\omega t_i)}{L_i}.$$

The deviance plays a major role in model fitting. A smaller value of deviance is preferred over larger values. Following Akaike Criterion [24] and Ronchetti [25], the robust equivalence of AIC is denoted by AICR, and is given by

$$AICR = \frac{D}{\hat{\sigma}^2} + 2p \frac{E[\psi_{\omega}^2(t)]}{E[\psi_{\omega}'(t)]},$$

and the Robust Schwarz Information Criterion BICR is given by

$$BICR = \frac{D}{\hat{\sigma}^2} + 2p \ln(n).$$

A robust coefficient of determination is given by

$$R^2 = 1 - \left( \frac{\text{Median}_{\{i:1 \leq i \leq n\}} |r_i|}{\text{Median}_{\{i:1 \leq i \leq n\}} |y_i|} \right)^2.$$

For more details, see Rosseeuw and Leroy [1].

There are numerous variable selection techniques available in the literature. One may use the stepwise procedure that may involve in forward selection or backward elimination. For each set  $S \subseteq \{x_1, x_2, \dots, x_p\}$  of explanatory variables, the robust final predicted error of Maronna et al. [26] is denoted by  $RFPE(S)$  and is defined as

$$RFPE(s) = \frac{\sum_{i=1}^n \rho_{\omega}(t_i)}{n} + \frac{\#(s) \sum_{i=1}^n \psi_{\omega}^2(t_i)}{n \sum_{i=1}^n \psi_{\omega}'(t_i)},$$

where  $\#(S)$  is the number of elements in the set  $S$ . In the forward selection or backward elimination, choose the one whose inclusion or deletion results in the smallest value of  $RFPE$ . To perform hypothesis testing, we let  $\Omega \subseteq \mathbb{R}^p$  be the parameter space and  $\{\theta_1, \theta_2, \dots, \theta_p\}$  be a subset of  $\{\theta_1, \theta_2, \dots, \theta_p\}$ .

$$\text{Define } \Omega_0 = \{\theta \in \Omega : \theta_1 = \theta_2 = \dots = \theta_p = 0\},$$

and the function  $f(\theta)$  as

$$f(\theta) = \sum_{i=1}^n \frac{\rho_{\omega}(t_i)}{L_i}.$$

Then a robust likelihood ratio type test statistic for testing the null hypothesis  $H_0: \theta \in \Omega_0$  against the alternative  $H_1: \theta \in \Omega_0^c$  is

$$S_n^2 = 2 \left( \text{Sup}_{\theta \in \Omega_0} f(\theta) - \text{Sup}_{\theta \in \Omega} f(\theta) \right) / q$$

For more information, the reader is referred to Hampel et al.

[27]. Asymptotically under the null hypothesis  $\frac{E[\psi_{\omega}'(t)]}{E[\psi_{\omega}^2(t)]} S_n^2$  has a chi-square distribution with  $q$  degrees of freedom, where the Wald type test statistic is defined as  $W_n^2 = n(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_q) V_q^{-1} (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_q)^t$ , and  $\frac{1}{n} V_q$  is the asymptotic variance-covariance matrix for the given vector  $(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_q)$ . the null distribution of the statistic  $W_n^2$  is asymptotically a chi-square distribution with  $q$  degrees of freedom.

Any of the following two robust algorithms can be used to estimate the parameter vector  $\theta$  and standard deviation  $\sigma$  of a nonlinear

regression model:

**Algorithm I:**

1. Set  $j=0$  and  $\hat{\sigma}^{(0)}=1$ . Calculate the initial estimate  $\hat{\theta}^{(0)}$  of parameter vector  $\theta$  by minimizing  $\sum_{i=1}^n \frac{\rho_w(y_i - g(\theta; x_i))}{L_i}$ .
2. Set  $j=j+1$  and calculate the followings :
  - i.  $\hat{\sigma}^{(j)}$  using formula (3) or (4).
  - ii. Use  $\hat{\theta}^{(j-1)}$  to evaluate matrix  $\hat{G}^{(j)}$ .
  - iii. Evaluate matrix  $H^{(j)}=Diagonal [(\hat{G}^{(j)}((\hat{G}^{(j)})^t)^{-1}\hat{G}^{(j)})]$ .
  - iv. The  $i$ th diagonal element of matrix  $H^{(j)}$  is denoted by  $h_i^{(j)}$ .
3. Calculate  $\hat{\theta}^{(j)}$  by minimizing  $f(\theta) = \sum_{i=1}^n \frac{\rho_w(t_i^{(j)})}{L_i}$ , where  $t_i^{(j)} = \frac{(y_i - g(\theta; x_i))(1 - h_i^{(j)})}{\hat{\sigma}^{(j)}}$ .
4. If convergence occurs, stop. Otherwise go to step 2 and continue the process.

The convergence occurs when  $\|\hat{\theta}^{(j)} - \hat{\theta}^{(j-1)}\| \cong 0$ . The convergence criterion is to stop the algorithm when  $\|\hat{\theta}^{(j)} - \hat{\theta}^{(j-1)}\| \leq \epsilon$  or  $\frac{\|\hat{\theta}^{(j)} - \hat{\theta}^{(j-1)}\|}{\|\hat{\theta}^{(j)}\|} \leq \epsilon$ . One may choose the value of  $\epsilon = 0.00001$ .

**Algorithm II:**

1. Set  $j=0$  and  $\hat{\sigma}^{(0)}=1$ . Calculate the initial estimate  $\hat{\theta}^{(0)}$  of parameter vector  $\theta$  by minimizing  $\sum_{i=1}^n \frac{\rho_w(y_i - g(\theta; x_i))}{L_i}$ .
2. Set  $j=j+1$ . For  $1 \leq i \leq n$ , calculate  $\hat{\sigma}^{(j)}$ ,  $t_i^{(j)}$  and weights  $w_i^{(j)}$  where  $w_i^{(j)} = \frac{\psi(t_i^{(j)})(1 - h_i^{(j)})}{\hat{\sigma}^{(j)}(y_i - g(\hat{\theta}^{(j)}; x_i))L_i}$  and  $t_i^{(j)} = \frac{(y_i - g(\theta; x_i))(1 - h_i^{(j)})}{\hat{\sigma}^{(j)}}$ , use them to calculate the weights matrix  $W^{(j)}$ .
3. Use information from step 2 to calculate.  $\hat{\theta}^{(j)} = (G^{(j)})^t W^{(j)} G^{(j)-1} (G^{(j)})^t W^{(j)} y$
4. If convergence occurs, stop. Otherwise go to step 2 and continue the process.

The convergence occurs when  $\|\hat{\theta}^{(j)} - \hat{\theta}^{(j-1)}\| \cong 0$ . The convergence criterion is to stop the algorithm when  $\|\hat{\theta}^{(j)} - \hat{\theta}^{(j-1)}\| \leq \epsilon$  or  $\frac{\|\hat{\theta}^{(j)} - \hat{\theta}^{(j-1)}\|}{\|\hat{\theta}^{(j)}\|} \leq \epsilon$ . One may choose the value of  $\epsilon = 0.00001$ .

**Drug concentration data**

Kenakin [28] used a set of responses to the concentration of an agonist in a functional assay. They fit the following model to their data. In this data, observation 5 has an outlier in the response direction,

$$\text{Response} = \text{Basel} + \frac{\text{Max} - \text{Basel}}{1 + 10^{n(\log(EC_{50} - \log(A))}} \tag{8}$$

Table 1 shows their actual and predicted concentrations as well as our results for fitted Hyperbolic model of type III (H3). For this example, the new robust technique is an effective regression tool in estimating model parameters in the presence of outliers. Figure 1 shows the fitted curve using hyperbolic model of type III (H3). Figure 2 uses formula (8) and the least squares fitted curve for the concentration data.

**Tumor metastasis**

The data in Table 2 consist of 12 observations. The response variable is the fraction of breast cancer patients with metastases and the predictor variable is the tumor size. Table 2 is from Michaleson et al. [29]. This data was originally collected by Tabar et al. [30-32] and Tubiana et al. [33,34]. To assess the robustness of our new method with regard to a special class of nonlinear growth models, we utilize this tumor metastasis data that is free of outliers. We first fit a model to the data using the robust method as well as least squares when there is no outlier present. Then we plant outliers in X direction, Y direction and both X and Y direction. In the X direction we change the X value in observation 12 from 90 to 2. In the Y direction, we change the y value in observation 6 from .55 to 3 and in both X and Y direction we change observation 12 in X direction from 90 to 2 and observation 7 in Y from .56 to 3.

For illustrative purposes, we have fitted hyperbolic of type II, Gompertz and logistic models. In the past, these models have been used to monitor cancer progression and regression. Each model has three parameters  $\theta_1, \theta_2,$  and  $\theta_3$  with  $\theta_1$  and  $\theta_2$  being positive and  $\epsilon_i$  are random errors. The response  $y_i$  is the fraction metastasized and  $x_i$  is the tumor size for individual  $i$ . The left graphs in Figures 3-5 are fitted curves using our proposed robust nonlinear regression technique and the graphs on the right sides of Figures 3-5 have been drawn using the nonlinear least square regression technique by planting outliers in X direction, Y direction and both X and Y direction. As you can see, when there is no outlier in the data all models perform well regardless of using the robust method or Least Squares. But when we plant outliers in the X, Y, and/or XY directions the fits become unacceptable for Least Squares whereas the robust method performs well for all models.

The hyperbolic model of type II or simply H2 has the form

$$y_i = \frac{\theta_1}{1 + \theta_2 \text{Arcsinh}[\exp[-\theta_3 x_i]]} + \epsilon_i .$$

The Gompertz model is of the form:  $y_i = \theta_1 \exp[-\theta_2 \exp(-\theta_3 x_i)] + \epsilon_i .$

The logistic model is of the form  $y_i = \frac{\theta_1}{1 + \theta_2 \exp[-\theta_3 x_i]} + \epsilon_i .$

**Simulation**

We perform simulation experiments to evaluate the robustness of our new nonlinear regression method compared to the least squares

OBS	Concentration	Response	Least Square	Calculating using 1	Hyperbolic H3
1	0.01	2	2	0.8	2
2	0.03	8	8.7	5.2	8
3	0.1	28	27.8	28	27.9
4	0.3	59	59.0	62.6	61.4
5	1	95	84.3	77.9	78.1
6	3	78	84.3	80.4	78.4
7	10	80	84.3	80.8	78.4

**Table 1:** Parameter Estimates for the Concentration Data.

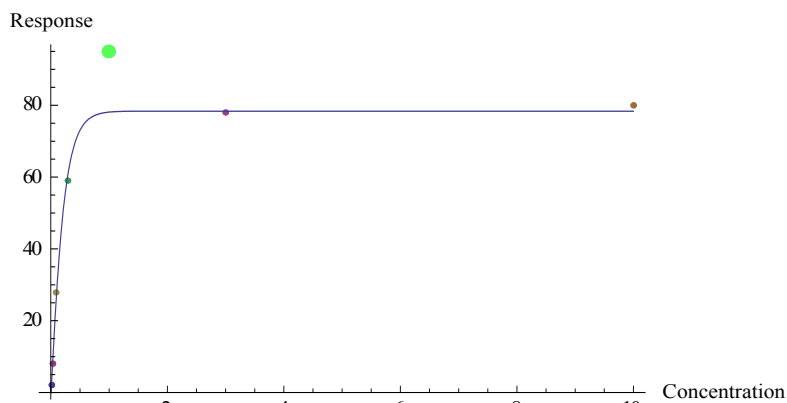


Figure 1: Fitted curve using hyperbolic model of type III (H3).

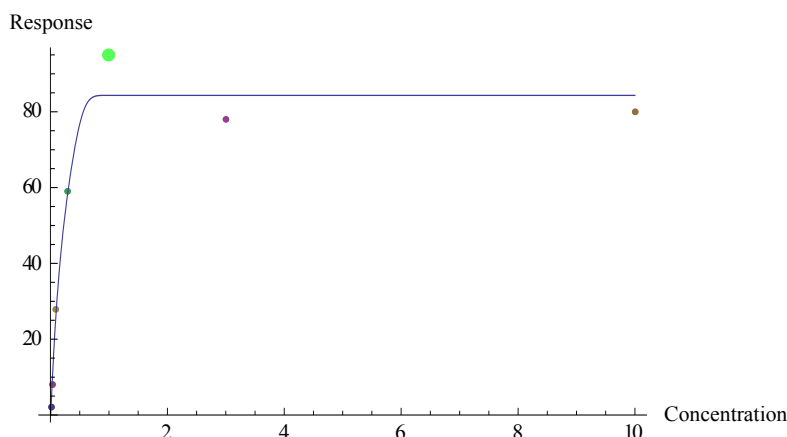


Figure 2: The least squares fitted curve for the concentration data.

<b>Tumor size x</b>	12	17	17	25	30	39	40	50	60	70	80	90
<b>Fraction Metastasized y</b>	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 2: Tumor Size Versus Fraction Metastasized Data.

n=20		0%		10%		20%		30%		40%	
		$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$
Least-Squares	Bias	0.0375	0.1083	25343	47179	8583	6953	2407	5185	1604	4015
	MSE	0.0479	0.4842	5.5E+9	1.5E+10	4.6E+9	2.1E+9	3.2E+8	2.2E+9	1.0E+8	1.4E+9
	MEP	5.0375	1.1083	25348	45180	9324	7329	3456	7733	1934	5234
Robust Method	Bias	0.0901	0.1525	0.0056	0.2242	0.034	0.2421	0.0199	0.2265	0.0176	0.3269
	MSE	0.0534	0.5646	0.2111	0.9423	0.2502	1.2292	0.2578	0.8812	0.2767	1.6324
	MEP	5.0701	1.1327	5.0056	1.2201	5.0134	1.2440	5.0199	1.2263	4.9824	1.3269
n=50		0%		10%		20%		30%		40%	
		$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$
Least-Squares	Bias	0.0131	0.1430	5214	4108	7326	5423	2113	4927	1476	3104
	MSE	0.0396	0.2566	6.1E+9	1.7E+9	3.2E+9	1.8E+9	2.5E+8	1.8E+9	7.9E+7	1.2E+9
	MEP	5.0131	1.0414	30123	58972	8589	6954	2412	6186	1609	4016
Robust Method	Bias	0.0701	0.1625	0.0328	0.0914	0.02925	0.1285	0.0494	0.1843	0.0050	0.2876
	MSE	0.2167	0.5221	0.2333	0.6257	0.2718	0.7224	0.3150	0.8543	0.2937	0.9745
	MEP	5.0701	1.1131	5.0328	1.0914	5.0292	1.1285	5.0494	1.1843	4.9734	1.1567

Table 3: Bias, Mean Square Errors (MSE) and Mean Estimated Parameter (MEP) with Percentage Contamination in the X Direction.

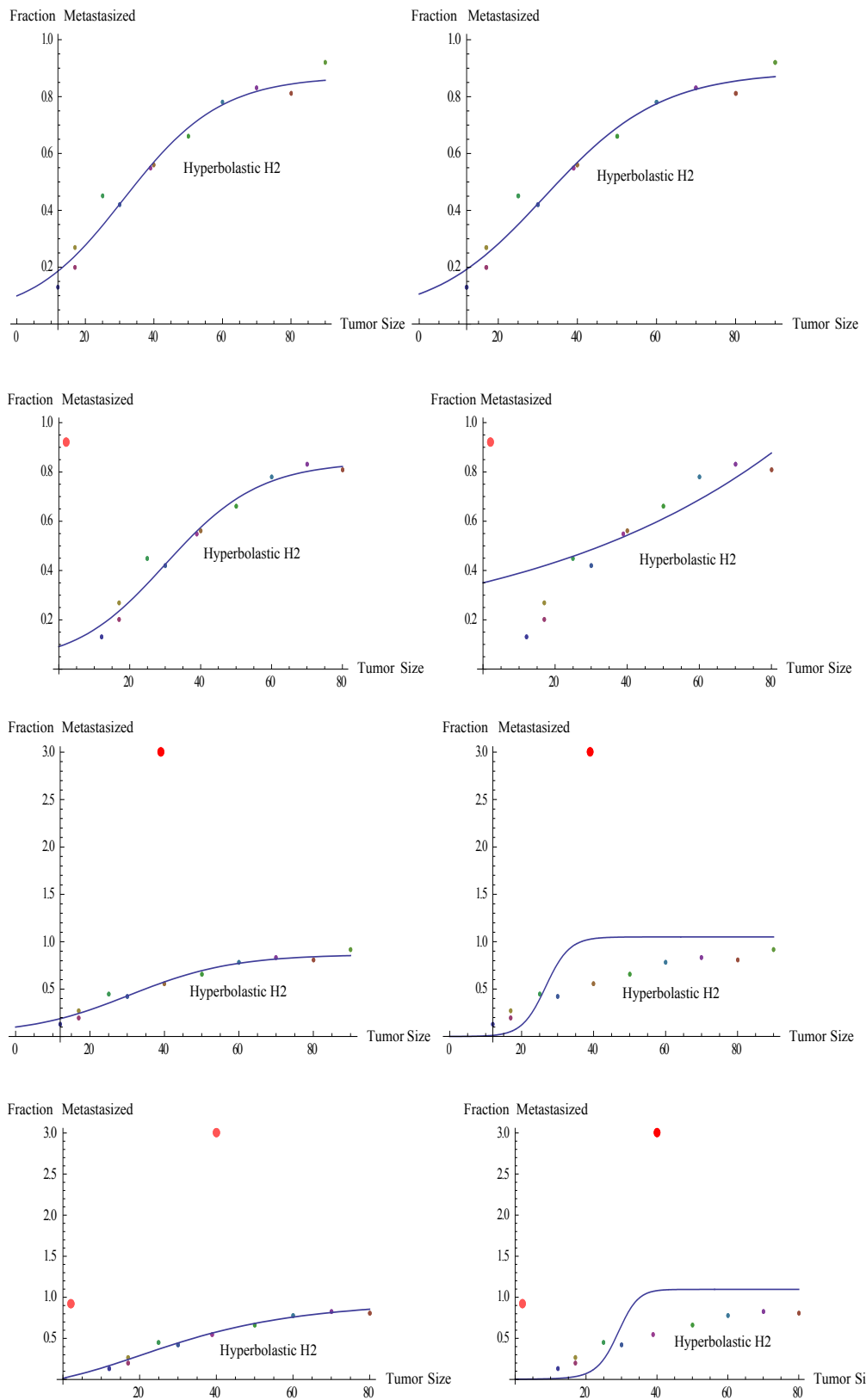


Figure 3: Hyperbolic model of type II.

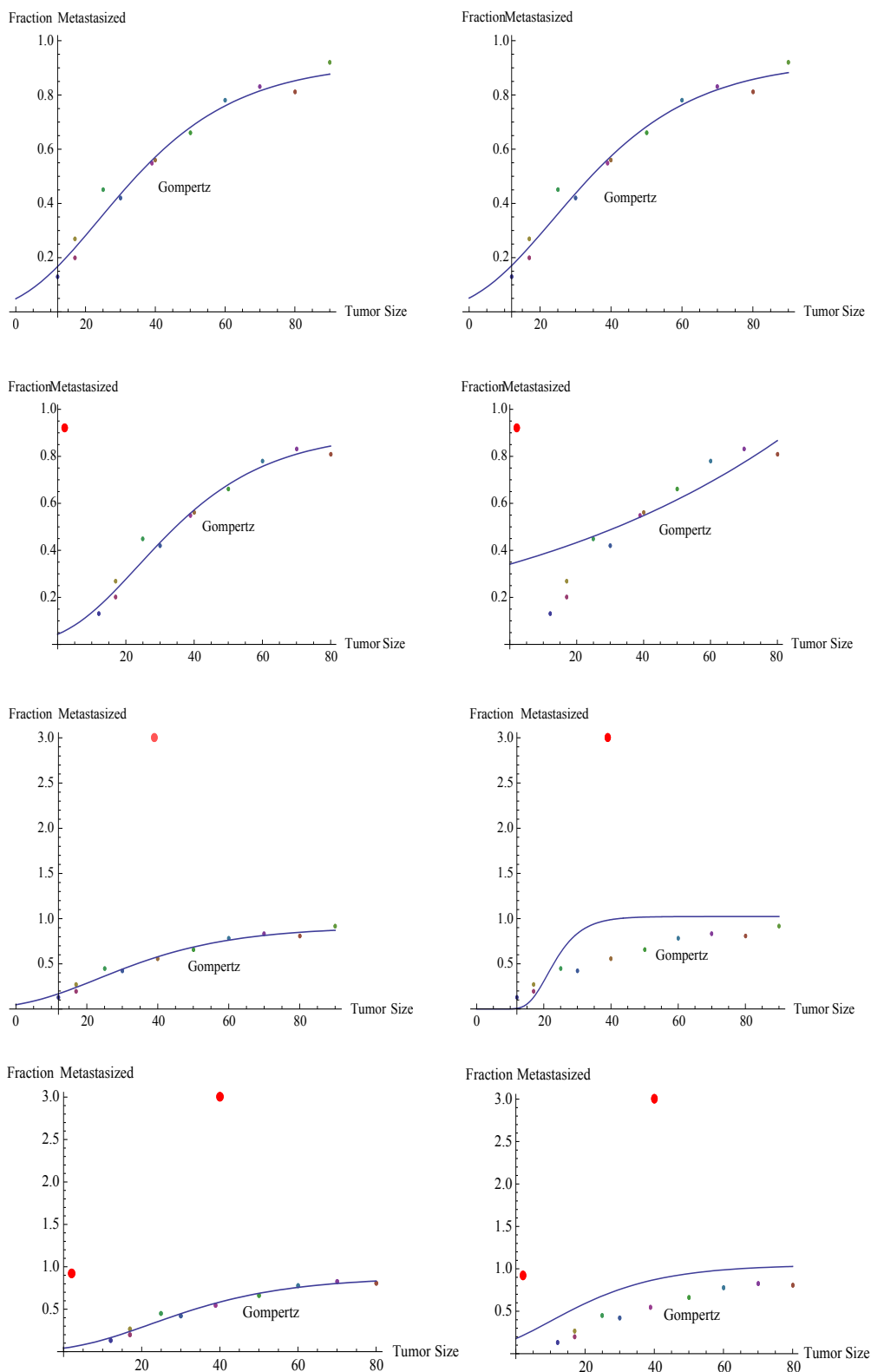


Figure 4: The Gompertz mode.



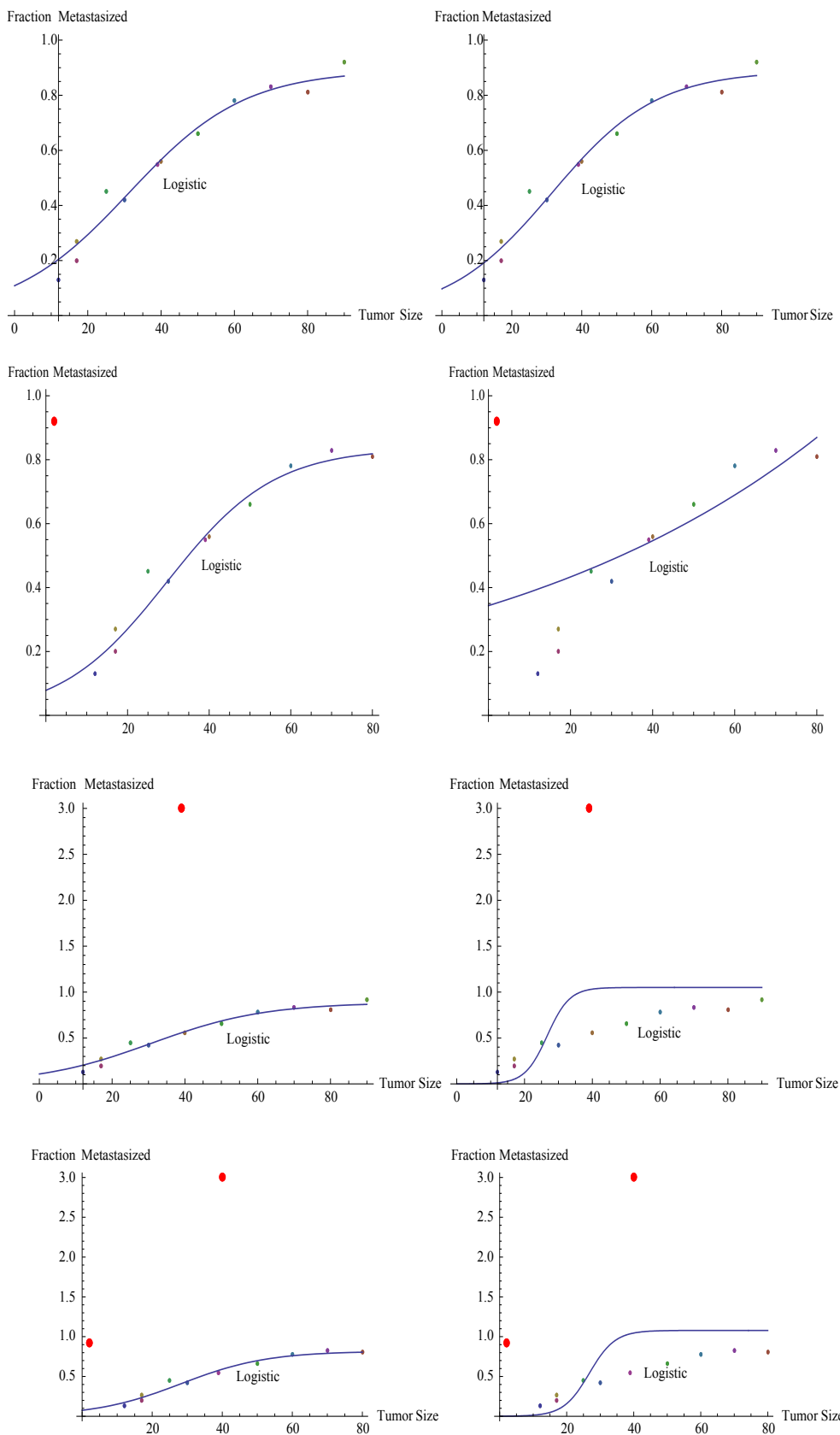


Figure 5: The logistic model.



n=20		0%		10%		20%		30%		40%	
		$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$
Least-Squares	Bias	0.0375	0.1083	68395	12511	115870	12323	115212	9858	136703	8915
	MSE	0.0479	0.4842	1.50E+10	1.50E+8	1.2E+11	1.2E+9	1.4E+11	9.9E+8	2.4E+11	1.0E+9
	MEP	5.0375	1.1083	68400	12512	115875	12324	115217	9859	136708	8916
Robust Method	Bias	0.0901	0.1525	0.0358	0.0523	0.0362	0.0250	0.0899	0.1605	0.2034	0.3292
	MSE	0.0534	0.5646	0.2296	0.5710	0.2555	0.6212	0.3534	0.8233	0.3928	1.095
	MEP	5.0701	1.1327	4.9641	1.0523	4.9638	0.9751	5.0899	1.1601	5.2034	1.3292
n=50		0%		10%		20%		30%		40%	
		$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$
Least-Squares	Bias	0.0131	0.1430	23650	11692	18843	5605	25855	5462	33790	5170
	MSE	0.0396	0.2566	4.6E+9	1.1E+9	7.6E+9	6.3E+9	1.2E+10	5.1E+08	3.1E+10	7.1E+08
	MEP	5.0131	1.0414	23655	11693	18848	5606	25860	5463	33795	5171
Robust Method	Bias	0.0701	0.1625	0.0365	0.0646	0.0262	0.0143	0.0504	0.0640	0.0022	0.0510
	MSE	0.2167	0.5221	0.0555	0.2067	0.0660	0.3303	0.0673	0.3861	0.0710	0.6139
	MEP	5.0701	1.1131	5.0365	1.0646	4.974	0.9857	5.0505	1.0650	5.0022	1.0510

Table 4: Bias, Mean Square Errors (MSE) and Mean Estimated Parameter (MEP) with Percentage Contamination in the Y Direction.

n=20		0%		10%		20%		30%		40%	
		$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$
Least-Squares	Bias	0.0375	0.1083	25342.0	45178	8582	6952	2406	5184	1603	4015
	MSE	0.0479	0.4842	5.5E+9	1.5E+10	4.6E+9	2.1E+9	3.2E+8	2.2E+9	1.0E+8	1.4E+9
	MEP	5.0375	1.1083	25347	45179	8587	6953	2411	5185	1608	4016
Robust Method	Bias	0.0901	0.1525	0.0762	0.0895	0.0081	0.0599	0.1191	0.1685	0.0176	0.0417
	MSE	0.0534	0.5646	0.1676	0.3112	0.1484	0.6847	0.2901	0.7813	0.2539	0.7448
	MEP	5.0701	1.1327	5.0762	1.0895	4.9919	1.0599	5.1191	1.1685	5.0176	0.0417
n=50		0%		10%		20%		30%		40%	
		$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$	$\theta_1$	$\theta_2$
Least-Squares	Bias	0.0131	0.1430	660	892	583	582	566	478	559	404
	MSE	0.0396	0.2566	455959	1.0E+6	345138	437874	323227	252951	314270	176481
	MEP	5.0131	1.0414	665	893	588	583	571	479	564	405
Robust Method	Bias	0.0701	0.1625	0.0173	0.0736	0.0892	0.1767	0.0399	0.0059	0.01445	0.0256
	MSE	0.2167	0.5221	0.0526	0.3141	0.0746	0.4530	0.0474	0.2435	0.0857	0.4877
	MEP	5.0701	1.1131	4.9827	1.0736	5.0892	1.1767	4.9601	1.0060	4.9856	1.0256

Table 5: Bias, Mean Square Errors (MSE) and Mean Estimated Parameter (MEP) with Percentage Contamination in the X-Y-Direction.

method. We have simulated the biochemistry model known as Michaelis-Menten kinetics. In biochemistry this model expresses the reaction velocity  $V$  as a function of concentration of substrate  $C$  as

where the parameter  $\alpha$  denotes the maximum reaction velocity and  $\beta$  is the substrate concentration at which the initial velocity  $V_0$  is 50% of the maximum reaction velocity. The larger the parameter  $\beta$ , the lower is the efficiency between the substrate and enzyme. This model has also been used in many biological systems such as gene regulatory system. In order to investigate the robustness of our new method relative to the method of least squares, we considered the nonlinear Michaelis-Menten equation of the form

$$y_i = \frac{\theta_1 x_i}{\theta_2 + x_i} + \varepsilon_i, \quad i = 1, 2, \dots, n$$

Where the response variable is  $y_i$  and  $x_i$  is fixed. In our simulations we set  $x_i=i$  and  $\varepsilon_i$  as the standard normal distribution with mean 0 and standard deviation 1. We performed 1000 repetitions using two sample sizes  $n=20$  and  $n=50$ . The outliers were randomly chosen in the direction of  $X$ ,  $Y$  and both  $X$  and  $Y$ . We used contamination levels of 0%, 10%, 20%, 30%, and 40%. In this simulation the parameter values are  $\theta_1=5$  and  $\theta_2=1$ . The software Mathematica is used in the simulation

process. To evaluate the robustness of these estimators, we randomly choose 10%, 20%, 30% and 40% of the simulated observations and contaminate the selected data by magnifying their size by a factor of 100 in the direction of explanatory variable  $X$ , response variable  $Y$ , and both response  $Y$  and explanatory  $X$  variables. Finally, we estimate both bias and mean squared errors using the following equations

$$bias = \left| \frac{\sum_{l=1}^m \hat{\theta}_l}{m} - \theta \right|$$

Where  $m$  is the number of iterations in the simulation. The mean squared error is estimated by

$$MSE = \frac{\sum_{l=1}^m (\hat{\theta}_l - \theta)^2}{m}$$

Tables 3-5 give the summary of our simulation outcome for both small and large sample sizes. The asymptotic efficiency for our simulation studies has been set to 95% level. By examining the simulation tables, we find out that in the absence of contamination in the simulated data, both the least square and the proposed robust method perform well with respect to bias, mean square error and mean estimated parameter

values. However, when contamination enters into our simulated data in the direction of explanatory variable  $X$ , or response variable  $Y$ , or both  $X$  and  $Y$ , then the new method outperforms the least squares method for both small and large samples. We also observe that the estimated values of parameters  $\theta_1$  and  $\theta_2$  are in close proximity of the true values of the parameters  $\theta_1$  and  $\theta_2$ . The simulation results clearly indicate the robustness of our new nonlinear regression technique relative to least squares method when outliers or influential observations are present.

## Conclusion

In this paper we introduced a new robust estimator of nonlinear regression parameters. In addition, robust testing for hypothesis about model parameters was introduced. Moreover, two algorithms were developed to perform the robust nonlinear estimation of model parameters. The computer simulation revealed the robustness of our new estimator. This robust method provides a powerful alternative to least squares method. The robust method presented in this paper has influence functions bounded in both the response and the explanatory variable direction. It has high asymptotic breakdown point and efficiency. A Mathematica program is also provided to ease in computations. This program does the necessary calculations to perform the robust nonlinear regression analysis of the drug concentration example given in this paper.

## Acknowledgement

Research reported in this paper was partially supported by the Center grant of the National Cancer Institute of the National Institutes of Health to the University of Alabama at Birmingham Comprehensive Cancer Center (P30 CA013148), the Cervical SPORE grant (P50CA098252), the Morehouse/Tuskegee University/UAB Comprehensive Cancer Center Partnership grant (2U54-CA118948), and the Mid-South Transdisciplinary Collaborative Center for Health Disparities Research (U54MD008176). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## References

- Rousseeuw PJ, Leroy AM (1987) Robust Regression and Outlier Detection. New York Wiley.
- Edgeworth FY (1987) On observations relating to several quantities. *Hermathena* 6: 279-285.
- Huber PJ (1973) Robust regression: asymptotics, conjectures, and Monte Carlo. *Ann Stat* 1: 799-821.
- Rousseeuw PJ (1984) Least median of squares regression. *J Am Stat Assoc* 79: 871-880.
- Rousseeuw PJ, Yohai VJ (1984) Robust regression by means of S estimators. In *Robust and nonlineartime series analysis*. Springer-Verlag 26: 256-274.
- 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.
- Tabatabai MA, Argyros IK (1993) Robust estimation and testing for general nonlinear regression models. *Appl Math Comput* 58: 85-101.
- Stromberg AJ (1993) Computation of high breakdown nonlinear regression parameters. *Annals of Statistics* 15: 642-656.
- Tabatabai MA, Eby WM, Li H, Bae S, Singh KP (2012) TELBS robust linear regression method. *Open Access Medical Statistics* 2: 65-84.
- Minn AJ, Gupta GP, Padua D, Bos P, Nguyen DX, et al. (2007) Lung metastasis genes couple breast tumor size and metastatic spread. *PNAS* 104: 6740-6745.
- Arisio R, Sapino A, Cassoni P, Accinelli G, Cuccorese MC, et al. (2000) What modifies the relation between tumor size and lymph node metastases in T1 breast carcinomas? *J Clin Pathol* 53: 846-850.
- Ramaswamy S, Ross KN, Lander ES, Golub TR (2003) A molecular signature of metastasis in primary solid tumors. *Nature Genetics* 33: 49-54.
- Maffuz A, Barroso-Bravo S, Nájera I, Zarco G, Alvarado-Cabrero I, et al. (2006) Tumor size as predictor of microinvasion, invasion, and axillary metastasis in ductal carcinoma in situ. *J Exp Clin Cancer Res* 25(2): 223-227.
- Hense HW, Ahrens S, Paulussen M, Lehnert M, Jürgens H (1999) Factors associated with tumor volume and primary metastases in Ewing tumors: Results from the (E1) CESS studies. *Annals of Oncology* 10: 1073-1077.
- Umbreit EC, Shimko MS, Childs MA, Lohse CM, Chevillie JC, et al. (2011) Metastatic potential of a renal mass according to original tumor size at presentation. *BJU International* 109: 190-194.
- Wu Z, Mo-Fa Gu, Zeng, Shao-Min H, Yong S (2013) Correlation between nasopharyngeal carcinoma tumor volume and the 2002 International Union Against Cancer tumor classification system. *Radiation Oncology* 8:87.
- Coras R, Hölsken A, Seufert S, Hauke J, Eyüpoglu IY, et al. (2007) The peroxisome proliferator-activated receptor- $\gamma$  agonist roglitazone inhibits transforming growth factor- $\beta$ -mediated glioma cell migration and brain invasion. *Mol Cancer Ther* 6: 1745-1754.
- Roth CM (2002) Quantifying Gene Expression. *Curr Issues Mol Biol* 4: 93-100.
- Kropf J, Schurek JO, Wollner A, Gressner A (1997) Immunological measurement of Transforming growth factor-beta 1 (TGF- $\beta$ 1) in blood; assay development and comparison. *Clinical Chemistry* 43: 10 1965-1974.
- Yang J, Richmond A (2006) The Angiostatic Activity of Interferon-Inducible Protein-10/CXCL10 in Human Melanoma Depends on Binding to CXCR3 but Not to Glycosaminoglycan. *Mol Ther* 9 (6): 846-855.
- Hao X, Silva EA, Månsson-Broberg A, Grinnemo KH, Siddiqui AJ, et al. (2007) Angiogenic effects of sequential release of VEGF-A165 and PDGF-BB with alginate hydrogels after myocardial infarction. *Cardiovasc Res* 75: 178-185.
- Bierie B, Stover DG, Abel TW, Chytil A, Gorska AE, et al. (2008) Transforming Growth Factor- $\beta$  Regulates Mammary Carcinoma Cell Survival and Interaction with the Adjacent Microenvironment. *Cancer Res* 68: 1809-1819.
- Rousseeuw PJ, Croux C (1993) Alternatives to the Median Absolute Deviation. *J Am Stat Assoc* 88: 1273-1283.
- Akaike H (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6): 716-723.
- Ronchetti E (1985) Robust model selection in regression. *Stat and Prob Letters* 3(1): 21-23.
- Maronna R, Martin D, Yohai VJ (2006) Robust Statistics: Theory and Methods. New York: Wiley.
- Hampel FR, Ronchetti EM, Rousseeuw PJ, Stahel WA (1986) Robust Statistics: The Approach Based on Influence Functions. New York: Wiley.
- Kenakin TP (2009) A Pharmacology Primer: Theory, Applications, and Methods. Third Edition, Academic Press: 286- 287.
- Michaelson JS, Halpern E, Kopans D (1999) Breast cancer: Computer simulation method for estimating optimal intervals for screening. *Radiology* 21: 551-560.
- Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A, et al. (1992) Update in the Swedish two-county program of mammographic screening for breast cancer. *Radiol Clin North Am* 30: 187-210.
- Tabar L, Fagerberg G, Chen HS, Duffy SW, Smart CR, et al. (1995) Efficacy of breast cancer screening by age: New results from the Swedish two-county trial. *Cancer* 75: 2507-2517.
- Tabar L (1996) Breast cancer screening with mammography in women aged 40-49 years. *Int J Cancer* 68: 693-699.
- Tubiana M, Koscielny S (1991) Natural history of human breast cancer: recent data and clinical implications. *Breast Cancer Res Treat* 18: 125-140.
- Tubiana M, Koscielny S (1990) The natural history of human breast cancer: implications for a screening strategy. *Int J Radiat Oncol Biol Phys* 19: 1117-112.