İstatistikçiler Dergisi: İstatistik & Aktüerya

Journal of Statisticians: Statistics and Actuarial Sciences

IDIA 18, 2025, 1, 1-15

Geliş/Received:17.01.2025, **Kabul**/Accepted: 30.05.2025

Araştırma Makalesi / Research Article

A genetic algorithm for robust regression in linear models

Ahmet Toy[1]

Ondokuz Mayıs University,
Terme Vocational School,
Samsun, Turkey
ahmet.toy@omu.edu.tr
ORCID: 0000-0002-2647-7259

Erol Terzi

Ondokuz Mayıs University,
Faculty of Sciences,
Samsun, Turkey
eroltrz@omu.edu.tr
ORCID: 0000-0002-2309-827X

Abstract

Outliers negatively affect the parameter estimate. Therefore, observation values can be weighted to minimize the negative impact of outliers on the parameter estimate. In this study, a robust method is proposed in which observation values are weighted with Genetic Algorithm (*GA*), which can be used both for outlier detection and parameter estimation. The proposed Genetic Algorithm for Robust Regression (*GA-RR*) method and M-estimators were compared to the root mean square error *(RMSE)* and mean absolute error *(MAE)* performance criterion using simulation study. Furthermore, the performance of the methods was evaluated using real data.

Anahtar sözcükler: Robust Regression, M-Estimators, Genetic Algorithm, Outlier.

Öz

Doğrusal Modellerde Robust Regresyon için Bir Genetik Algoritma

Aykırı değerler parametre tahminini olumsuz etkilemektedir. Bu nedenle aykırı değerlerin parametre tahmini üzerindeki olumsuz etkisini minimize etmek için gözlem değerleri ağırlıklandırılabilmektedir. Bu çalışmada, gözlem değerlerinin Genetik Algoritma (GA) ile ağırlıklandırıldığı, hem aykırı değerlerin tespiti hem de parametre tahmini için kullanılabilecek bir robust yöntem önerilmektedir. Önerilen Robust Regresyon için Genetik Algoritma (GA-RR) yöntemi ve M-tahminciler, simülasyon çalışması ile hata kareler ortalamasının karekökü (RMSE) ve ortalama mutlak hata (MAE) performans kriterince karşılaştırılmıştır. Ayrıca, gerçek veri kullanılarak yöntemlerin performansları değerlendirilmiştir.

Keywords: *Robust Regresyon, M-Tahminciler, Genetik Algoritma, Aykırı Değer.*

[1] Bu çalışma, birinci yazarın, ikinci yazar danışmanlığında hazırladığı doktora tezinden üretilmiştir.

1. Introduction

Classical statistics develops optimal methods under parametric model assumptions. Robust statistical theory deals with deviations from model assumptions and aims to develop statistical methods that are still reliable and efficient around the model. Therefore, sound statistical theory can be seen as a bridge between the Fisher parametric and fully nonparametric approaches. That is, it is a compromise between the difficulty of the parametric model and the potential difficulties of interpreting the fully nonparametric model [1].

Regression analysis is a method that models the relationship between variables. Multiple linear regression model;

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \varepsilon_i, \tag{1}$$

it is shown as. Here, i is the observation index, Y is the dependent variable, $X_1, X_2, \ldots, X_k$ are the independent variables, $\beta_0, \beta_1, \ldots, \beta_k$ are the model parameters and ε is the error term. The least squares (*LS*) method aims to minimize the sum of differences between Y_i and the predicted value $\hat{Y}_i$ for parameter estimation. When the linear regression assumptions are linearity, independence of errors (no autocorrelation), normality of errors, zero mean of errors, constant variance of errors (homoscedasticity), and no multicollinearity, the least squares method provides the best parameter estimation. However, there exist some factors that lead to the violation of model assumptions. One of these factors is that the data set contains outliers.

Outliers positioned differently from other observations in the data set cause negative parameter estimates. Different robust regression estimators have been developed for data sets containing outliers. While some estimators make parameter estimates by excluding outliers from the data set, others make parameter estimates by reducing the effects of outliers. One of the reasons for parameter estimation by reducing the effect of outliers is to ensure that the observations remain in the data set because sometimes an observation value that is thought to be an outlier can provide important information about the research topic. For this reason, it would be more accurate for a reliable parameter estimate to reduce the effect of a controversial observation value and to ensure that it remains in the data set. In some cases, it is difficult to determine which observations in the data set are outliers. Therefore, these different approaches to outliers have led to the development of different robust regression estimators. One of the most important estimators developed for parameter estimation in data sets containing outliers is the robust regression M-estimators, named after the maximum likelihood method. Since M-estimators are resistant to outliers in the dependent variable, it is considered a robust method. M-estimators attempt to minimizing sum the function ρ instead of minimizing the sum square error in the least-squares method. Here, the aim is to limit the effect of outlier observations by using the ρ function. Different M-estimators have been developed for this purpose. In this study, Huber, Cauchy, and Hampel M-estimators are considered. M-estimators have been used in different fields of study. In this context, Wu and Wu [2] used M-estimators to develop a method that provides the accuracy of orbit improvement for 20 satellites. Hu et al. [3] used M-estimators to examine absolute gravimetric data. Besides, Leoni et al. [4] used M-estimators to estimate the shaping coefficients of possible electricity prices, while Su et al. [5] used M-estimators to reject large measurement errors in studies in the field of pharmacology.

Robust regression methods attempt to prevent outliers that cause the regression model to be distorted. Therefore, the *GA* method, which can perform a global search, can be used to detect outliers and reduce the effect of outliers. *GA* is described as the genetic heuristic search algorithm that attempts to get results with random search methods [6]**.** *GA* has been used in many different fields of study. One of the main reasons for the popularity of *GA* is its robustness and reliability [7]**.** In this direction, *GA* studies have been carried out to develop robust methods. Vankeerberghen et al. [8], stated that a *GA* can be used to obtain Least Median Squares parameters in nonlinear models. Hu [9], proposed a *GA* based method to define the upper and lower bound of the data range in the robust nonlinear range regression model. Wiegand et al. [10], presented a technique to combine a robust outlier detection method with an optimized *GA* for spectral variable selection. Sykas and Karathanassi [11], worked to automatically determine the optimal regression model that allows the estimation of a parameter. Duraj and Chomatek [12], conducted *GA* studies on outliers in calibration and pharmacokinetic data. In this study, *GA-RR* [13], which is an alternative new robust method that can be used for outlier detection and parameter estimation in linear models, is introduced.

2. Methodology

2.1. M-Estimators

The M-estimator was introduced in 1964 by Peter J. Huber. The M-estimator is defined as follows.

$$\min_{\beta} \sum_{i=1}^{n} \rho(r_i) \tag{2}$$

Here, i is the observation index, ρ is a symmetric function [14]. Some properties that a suitable $\rho(r)$ function must have are as follows [15, 16].

- $\rho(r) \geq 0$
- $\rho(0) = 0$
- $\rho(r) = \rho(-r)$
- $\rho(r_i) \geq \rho(r_{i'})\ for\ |r_i| > |r_{i'}|$

Influence function $\psi(r)$ is obtained by calculating the derivative of $\rho(r)$ in M-estimators. The weight function $w(r)$ is obtained by the ratio of the $\psi(r)$ function to the residual. Estimation with M-estimators obtained from weight functions can be made using the Iteratively Reweighted Least Squares (*IRLS*) method[17]. The scale to be used for parameter estimation is $\sigma = [Median(|r_i - Median(r_i)|)]/0.6745$ and the steps to be taken are as follows [18].

i. The initial $\hat{\beta}$'s are obtained with the *LS* and the residuals *(*r_i*)* are calculated.
ii. Weights are obtained with $\sigma = [Median(|r_i - Median(r_i)|)]/0.6745$
iii. The next $\hat{\beta}$'s are obtained with Weighted Least Squares (*WLS*) $\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{W} \mathbf{y}$
iv. As a result of iterations, the difference between $\hat{\beta}$'s is iterated until convergence is reached.

Table 1 shows the $\rho(r)$, $\psi(r)$ and $w(r)$ functions of the M-estimators. Table 1 shows the tuning constants of the M-estimators for 95% asymptotic efficiency in the normal distribution [19,20].

Table 1. $\rho(r)$, $\psi(r)$, $w(r)$ functions and tuning constants of M-estimators

Method	$\boldsymbol{\rho(r)}$	$\boldsymbol{\psi(r)}$	$\boldsymbol{w(r)}$	Range	Tuning constant
Huber	$\frac{r^2}{2}$	r	1	$\lvert r\rvert \leq c_H$	c_H=1.345 [19]
	$c_H\lvert r\rvert - \frac{{c_H}^2}{2}$	$c_H\ sign(r)$	$\frac{c_H}{\lvert r\rvert}$	$\lvert r\rvert > c_H$	
Cauchy	$\left(\frac{{c_C}^2}{2}\right) log\left(1 + \left(\frac{r}{c_C}\right)^2\right)$	$\frac{r}{\left(1 + \left(\frac{r}{c_C}\right)^2\right)}$	$\frac{1}{\left(1 + \left(\frac{r}{c_C}\right)^2\right)}$	$r \epsilon \mathbb{R}$	c_C=2.3849

	$\frac{r^2}{2}$	r	1	$\lvert r\rvert \le a$	
	$a\lvert r\rvert - \frac{a^2}{2}$	$a\,sign(r)$	$\frac{a}{\lvert r\rvert}$	$a < \lvert r\rvert \le b$	$a = 1.35$ $b = 2.7$ $c = 5.4$
Hampel	$ab - \frac{a^2}{2} + \frac{a(c-b)}{2}\left[1 - \left(\frac{c-\lvert r\rvert}{c-b}\right)^2\right]$	$\frac{a(c-\lvert r\rvert)}{c-b} sign(r)$	$\frac{a(c-\lvert r\rvert)}{(c-b)\lvert r\rvert}$	$b < \lvert r\rvert \le c$	
	$ab - \frac{a^2}{2} + \frac{a(c-b)}{2}$	0	0	$\lvert r\rvert > c$	

2.2. Genetic Algorithm

GA was introduced by John Holland in the early 1970s and is a population genetics-based heuristic search method [21]. The *GA* is based on the principle of "survival of the fittest individuals" suggested by Charles Darwin in the 19th century [22]. *GA* deals with large search spaces and searches these spaces to search for optimal combinations of solutions [23]. The basic elements of *GA* are genes, chromosomes, and population. Gene refers to the inheritance that enables the transmission of individual characteristics from generation to generation. Chromosomes are structures that are formed by the combination of genes. The population is the set of individuals [21]. One of the issues required to be emphasized in *GA* is the fitness function. The fitness function is the function value defined by the researcher to obtain the optimal result of the problem discussed. *GA* operators are selection, crossover, and mutation.

- Selection; It is a step in which the fitness values obtained from the chromosomes are investigated. At this stage, it is intended to include the chromosomes having good fitness values d in the next generation. Therefore, the k-medoids method was used for the selection process in this study. While the k-medoids method creates the data set in the form of k clusters called medoids, it aims to keep similarities of each cluster within themselves at the maximum level and the similarities between the clusters at the minimum level [24]. The k-medoids method attempts the minimization of the total distance between the elements of each cluster. Therefore, the algorithm considers the element with the least distance from the center of the cluster as the cluster medoid instead of averaging out the elements in the cluster to detect the medoid. The outlier cluster elements are prevented from pulling the cluster center toward the boundaries through this operation. Therefore, k-medoids is considered a robust clustering method.
- Crossover; It is required to apply the crossover process for the formation of new individuals. In the crossover process, the chromosomes in the population are randomly matched. A value in the range [0,1] is randomly generated for each pair of matched chromosomes. Then, the value generated for each chromosome pair is compared with the predetermined crossover rate. If the number produced for the chromosome pair after the comparison is smaller than the crossover rate, new individuals are obtained by applying the crossover process.
- Mutation; Following the crossover process, the mutation process is applied in order not to match the same chromosomes and quickly reach different chromosomes in the search space. The mutation process is applied to genes in chromosomes. In order to apply the mutation process to a gene, the previously determined mutation rate is compared with the randomly generated value in the range [0,1]. The mutation process is applied if the number produced after the comparison is smaller than the mutation rate.

The first step in the operation of *GA* is the formation of population initialization. Then, a new population is obtained by applying crossover, mutation, and selection processes to the chromosomes in the initial population. These steps are repeated until the predetermined number of iterations is reached, and when the number of iterations is achieved, the *GA* is stopped.

2.2.1. Proposed Algorithm

Generally, the LS estimator is used for parameter estimation in linear regression analysis. The LS estimator gives the best results when making the linear regression model assumptions. However, these assumptions are generally not provided. Therefore, alternative methods have been developed that can be used if there are deviations from the assumptions. One of them is the WLS method. WLS estimators of β are obtained by Equation (3).

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}^{\top}\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^{\top}\mathbf{W}\mathbf{y} \quad (3)$$

What is important in Equation 3 is to obtain the appropriate W weights for the data set. Here, to obtain the appropriate W weights, it is necessary to detect the outlier observations in the dataset and determine the outlier rates of the observations detected as outliers. This is the purpose of the *GA-RR* approach proposed in this study. For this purpose, a hybrid approach is presented in which GA, and k-medoid methods that provide robust clustering are used together as an alternative to robust regression estimators for parameter estimation. In the *GA-RR* approach, first, the chromosomes in the initial population are created with binary encoding, and the *RMSE* values of the chromosomes are obtained. Then, the k-medoids method is applied to select the chromosomes in the cluster with low *RMSE* values for the next generation. Crossover and mutation are applied, and the *GA* is stopped when the determined number of iterations is reached. When the iteration ends, the Min-Max normalization transformation is applied to the total number of observations in the chromosomes in the clusters with low *RMSE* values, and the W weights are obtained. Parameter estimation is performed by using the W weights in the WLS method. The algorithm steps of the *GA-RR* method are as follows;

Step 1. The tuning parameters to be used in the *GA* are determined. The optimal tuning parameters for this study were determined as population size = 100, crossover rate = 0.9, mutation rate = 0.001 and number of iterations = 1000.

Step 2. The initial population is created by binary encoding. Genes are derived as 0 or 1. 0 indicates that the observed value is not included in the chromosome, 1 that it is.

Step 3. The fitness value is calculated. *RMSE* values of chromosomes are calculated. In Equation (4), e_i is the residual value of the *i.* observation.

$$RMSE = \sqrt{\left(\sum_{i=1}^{n} e_i^2\right)/n} \quad (4)$$

Step 4. The selection process is applied. The k-medoids method is applied to *RMSE* values obtained from chromosomes. Chromosomes in the cluster with lower *RMSE* values are selected for the next generation.

Step 5. The crossover and mutation operations are applied. First, new chromosomes are obtained by applying the crossover process to the randomly paired chromosomes. Then, the mutation process is applied to the genes in the chromosomes to ensure that different chromosomes are obtained quickly.

Step 6. The processes between Step 3-5 are repeated according to the number of iterations. The stopping criterion is applied.

Min-Max normalization transformation in Equation (5) is applied to the total number of genes in clusters with small *RMSE* values. In Equation (5), G_i is the total number of the *i.* observation detected as a result of iterations. G_i values are used to obtain G_{i_v} values in the range [0,1]. Here, i_v is the *i.* observation value obtained by the Min-Max normalization transformation. W weights are obtained by squaring the obtained G_{i_v} values.

$$G_{i_v} = \frac{G_i - G_{min}}{G_{max} - G_{min}} \quad (5)$$

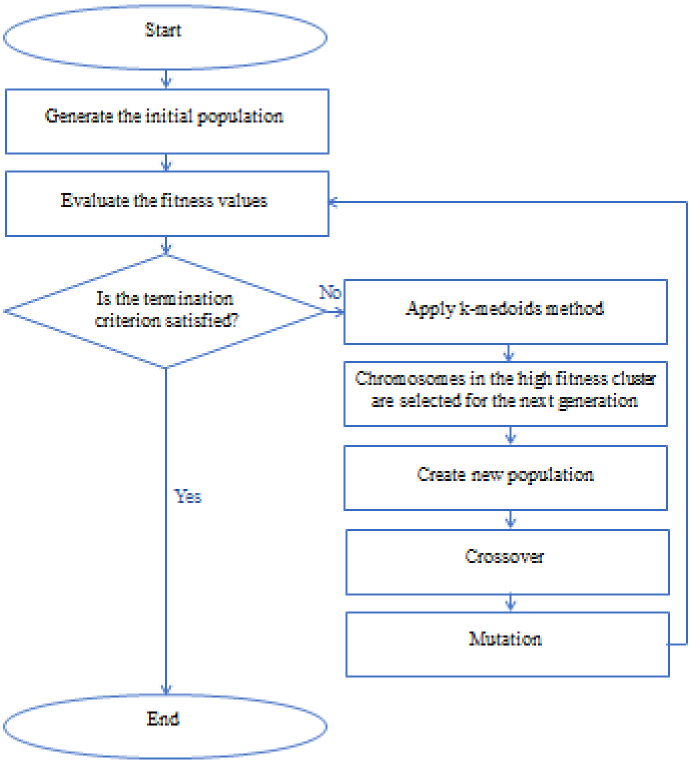

Figure 1. Flowchart of Genetic Algorithm

3. Applications

3.1. Simulation

In this part of the study, M-estimators and the *GA-RR* method are compared with the simulation study. The results in the study were obtained with the MATLAB program. To compare the methods in the simulation study, multiple linear regression analysis was performed with three independent variables. The multiple linear regression model established for the simulation study is shown in Equation (6).

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \varepsilon_i \quad (6)$$

Here, i is the observation index, and β_k (*k*=0, 1, 2, 3), is taken as 1. The independent variables are $X_1 \sim N(5,\sigma), X_2 \sim N(10,\sigma), X_3 \sim N(20,\sigma)$, $\sigma=1, \sigma=3, \sigma=5$, and $X_1 \sim N(5,5)$, $X_2 \sim N(10,3)$, $X_3 \sim N(20,1)$ derived in four different forms, and the error term is derived as $\varepsilon \sim N(0,1)$. In the study, the sample size was taken as $n = [30, 50, 100, 300]$, and the number of outliers as 1, 2 and 3. Outliers are formed by changing the value of any dependent variable i. as $Y_{i_v} = Y_i * 10$. For the simulation study, *t*=1000 different data sets were derived. The comparison of the methods was made using $RMSE_k$ and MAE_k with $\hat{\beta}_{kh}$ being the estimated value of β_k for the h. iteration.

$$RMSE_k = \sqrt{\left(\sum_{h=1}^{t} (\beta_k - \hat{\beta}_{kh})^2\right) \Big/ t} \quad (7)$$

$$MAE_k = \left(\sum_{h=1}^{t} |\beta_k - \hat{\beta}_{kh}|\right) \Big/ t \quad (8)$$

The following tables were obtained by examining the methods according to the simulation results obtained.

Table 2. Comparison of *GA-RR* method with M-estimators for n=30

Outlier	σ	Method	Average of coefficients				*RMSE*				*MAE*			
			β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
1	1	Huber	1.0931	0.9918	0.9914	1.0047	5.3762	0.2454	0.2251	0.2337	4.0548	0.1786	0.1746	0.1786
		Cauchy	1.0478	0.9934	0.9910	1.0039	4.9142	0.2137	0.2129	0.2137	3.8584	0.1677	0.1651	0.1678
		Hampel	1.0233	0.9930	0.9916	1.0048	4.9967	0.2144	0.2144	0.2181	3.9199	0.1682	0.1667	0.1705
		GA-RR	**1.0743**	**0.9938**	**0.9922**	**1.0019**	**4.7372**	**0.2066**	**0.2062**	**0.2054**	**3.7576**	**0.1628**	**0.1587**	**0.1618**
	3	Huber	1.0809	1.0034	1.0001	0.9983	1.7271	0.0767	0.0757	0.0755	1.3766	0.0586	0.0577	0.0601
		Cauchy	1.0432	1.0029	0.9991	0.9978	1.6343	0.0699	0.0700	0.0713	1.3120	0.0552	0.0553	0.0566
		Hampel	1.0363	1.0029	0.9994	0.9978	1.6485	0.0699	0.0697	0.0718	1.3237	0.0554	0.0550	0.0569
		GA-RR	**1.0379**	**1.0031**	**0.9995**	**0.9976**	**1.6096**	**0.0687**	**0.0681**	**0.0699**	**1.2877**	**0.0543**	**0.0538**	**0.0555**
	5	Huber	1.0843	0.9997	0.9990	0.9996	1.0358	0.0439	0.0451	0.0432	0.8052	0.0345	0.0354	0.0331
		Cauchy	1.0159	0.9987	0.9992	1.0002	0.9729	0.0418	0.0420	0.0407	0.7634	0.0327	0.0335	0.0321
		Hampel	1.0095	0.9988	0.9991	1.0003	0.9734	0.0416	0.0421	0.0410	0.7617	0.0329	0.0340	0.0322
		GA-RR	**1.0096**	**0.9990**	**0.9991**	**1.0003**	**0.9461**	**0.0409**	**0.0408**	**0.0398**	**0.7415**	**0.0322**	**0.0324**	**0.0314**
2	1	Huber	0.9155	0.9958	1.0021	1.0101	5.4544	0.2500	0.2491	0.2338	4.2379	0.1965	0.1917	0.1839
		Cauchy	0.7756	0.9958	1.0047	1.0094	4.9471	0.2200	0.2129	0.2066	3.8416	0.1741	0.1671	0.1641
		Hampel	0.7486	0.9951	1.0061	1.0101	4.9675	0.2208	0.2147	0.2074	3.8617	0.1744	0.1681	0.1643
		GA-RR	**0.8146**	**0.9955**	**1.0034**	**1.0080**	**4.9016**	**0.2196**	**0.2116**	**0.2051**	**3.8088**	**0.1737**	**0.1663**	**0.1632**
	3	Huber	1.1095	0.9949	0.9988	1.0039	1.9412	0.0991	0.0811	0.0818	1.4782	0.0670	0.0626	0.0629
		Cauchy	0.9723	0.9969	0.9994	1.0032	1.6445	0.0732	0.0695	0.0710	1.2953	0.0569	0.0551	0.0560
		Hampel	0.9631	0.9969	0.9998	1.0031	1.6543	0.0736	0.0704	0.0713	1.3019	0.0567	0.0555	0.0561
		GA-RR	**0.9829**	**0.9963**	**0.9998**	**1.0024**	**1.6369**	**0.0705**	**0.0685**	**0.0708**	**1.2782**	**0.0558**	**0.0546**	**0.0556**
	5	Huber	1.1250	1.0025	0.9996	0.9998	1.1280	0.0554	0.0514	0.0485	0.8786	0.0398	0.0399	0.0383
		Cauchy	1.0133	1.0028	0.9991	0.9991	0.9722	0.0483	0.0451	0.0416	0.7672	0.0344	0.0352	0.0329
		Hampel	0.9989	1.0030	0.9989	0.9993	0.9754	0.0492	0.0445	0.0411	0.7650	0.0345	0.0351	0.0326
		GA-RR	**1.0317**	**1.0008**	**0.9980**	**0.9988**	**0.9521**	**0.0434**	**0.0444**	**0.0403**	**0.7563**	**0.0339**	**0.0352**	**0.0320**
3	1	Huber	1.1097	0.9926	0.9895	1.0124	6.5681	0.2784	0.3736	0.2635	4.9097	0.2095	0.2283	0.2050
		Cauchy	0.9442	0.9983	0.9929	1.0067	4.9419	0.2158	0.2321	0.2144	3.8789	0.1676	0.1822	0.1687
		Hampel	0.9591	0.9980	0.9931	1.0057	4.9635	0.2173	0.2329	0.2158	3.8780	0.1685	0.1827	0.1692
		GA-RR	**0.9919**	**0.9984**	**0.9923**	**1.0043**	**4.7879**	**0.2112**	**0.2233**	**0.2103**	**3.7523**	**0.1652**	**0.1763**	**0.1662**
	3	Huber	1.1486	1.0006	1.0030	1.0026	2.6755	0.0992	0.0955	0.1166	1.7744	0.0734	0.0732	0.0758
		Cauchy	0.9365	1.0013	1.0038	1.0013	1.7912	0.0796	0.0739	0.0749	1.3765	0.0565	0.0579	0.0579
		Hampel	0.9232	1.0017	1.0038	1.0015	1.7908	0.0800	0.0739	0.0747	1.3715	0.0565	0.0578	0.0578
		GA-RR	**0.9733**	**0.9985**	**1.0028**	**1.0003**	**1.7246**	**0.0688**	**0.0724**	**0.0731**	**1.3547**	**0.0541**	**0.0571**	**0.0570**
	5	Huber	1.2393	0.9990	1.0013	0.9992	1.3623	0.0694	0.0589	0.0572	1.0162	0.0450	0.0453	0.0426
		Cauchy	1.0087	1.0009	1.0015	0.9993	1.0065	0.0457	0.0459	0.0436	0.7908	0.0345	0.0362	0.0342
		Hampel	0.9891	1.0014	1.0018	0.9996	1.0042	0.0449	0.0459	0.0436	0.7885	0.0343	0.0363	0.0342
		GA-RR	**1.0263**	**1.0006**	**1.0008**	**0.9986**	**0.9785**	**0.0418**	**0.0443**	**0.0421**	**0.7743**	**0.0333**	**0.0356**	**0.0334**

Table 2 shows that the *GA-RR* method has the smallest *RMSE* and *MAE* values. The methods with the smallest *RMSE* and *MAE* values after the *GA-RR* method are Cauchy, Hampel and Huber M-estimator, respectively. Additionally, it is seen that when the outlier is 2 and $\sigma = 5$, the *RMSE* and *MAE* values of the Hampel M-estimator are smaller than those of the Cauchy M-estimator.

Table 3. Comparison of *GA-RR* method with M-estimators for n=50

Outlier	σ	Method	Average of coefficients				*RMSE*				*MAE*			
			β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
1	1	Huber	0.9881	0.9972	1.0035	1.0008	3.6233	0.1629	0.1575	0.1553	2.8914	0.1284	0.1248	0.1234
		Cauchy	0.9627	0.9980	1.0036	1.0002	3.5161	0.1561	0.1524	0.1521	2.8149	0.1234	0.1209	0.1212
		Hampel	0.9633	0.9991	1.0033	1.0000	3.5379	0.1567	0.1528	0.1534	2.8282	0.1239	0.1210	0.1217
		GA-RR	**0.9592**	**0.9977**	**1.0024**	**1.0010**	**3.4552**	**0.1522**	**0.1489**	**0.1487**	**2.7614**	**0.1205**	**0.1182**	**0.1191**
	3	Huber	0.9985	1.0025	1.0030	0.9999	1.2751	0.0541	0.0551	0.0554	1.0312	0.0422	0.0436	0.0445
		Cauchy	0.9674	1.0022	1.0030	0.9998	1.2502	0.0519	0.0530	0.0540	1.0007	0.0406	0.0420	0.0432
		Hampel	0.9591	1.0022	1.0029	1.0002	1.2590	0.0526	0.0533	0.0542	1.0097	0.0412	0.0423	0.0434
		GA-RR	**0.9820**	**1.0017**	**1.0027**	**0.9991**	**1.2191**	**0.0508**	**0.0513**	**0.0526**	**0.9743**	**0.0398**	**0.0409**	**0.0416**
	5	Huber	1.0385	0.9999	1.0015	0.9996	0.7630	0.0318	0.0318	0.0326	0.6039	0.0250	0.0254	0.0259
		Cauchy	1.0085	0.9996	1.0012	0.9996	0.7427	0.0304	0.0307	0.0317	0.5864	0.0239	0.0244	0.0251
		Hampel	1.0053	0.9995	1.0014	0.9995	0.7454	0.0303	0.0307	0.0318	0.5882	0.0238	0.0245	0.0252
		GA-RR	**1.0101**	**0.9992**	**1.0011**	**0.9994**	**0.7276**	**0.0297**	**0.0302**	**0.0308**	**0.5710**	**0.0233**	**0.0240**	**0.0243**
2	1	Huber	1.1485	0.9973	0.9964	0.9990	3.8684	0.1624	0.1611	0.1714	3.0428	0.1289	0.1278	0.1357
		Cauchy	1.0207	0.9984	0.9993	1.0001	3.6112	0.1553	0.1524	0.1593	2.8881	0.1228	0.1199	0.1268
		Hampel	1.0262	0.9979	0.9987	1.0001	3.6320	0.1557	0.1529	0.1603	2.9097	0.1233	0.1205	0.1279
		GA-RR	**1.0131**	**0.9994**	**0.9995**	**1.0000**	**3.5392**	**0.1529**	**0.1505**	**0.1563**	**2.8294**	**0.1201**	**0.1183**	**0.1251**
	3	Huber	1.0403	1.0002	1.0011	1.0012	1.2883	0.0595	0.0562	0.0569	1.0097	0.0454	0.0444	0.0448
		Cauchy	0.9820	1.0003	1.0013	1.0003	1.1804	0.0548	0.0515	0.0513	0.9327	0.0426	0.0404	0.0409
		Hampel	0.9778	1.0005	1.0013	1.0003	1.1720	0.0544	0.0509	0.0513	0.9270	0.0425	0.0403	0.0407
		GA-RR	**0.9751**	**1.0006**	**1.0010**	**1.0004**	**1.1651**	**0.0520**	**0.0500**	**0.0508**	**0.9214**	**0.0414**	**0.0395**	**0.0404**
	5	Huber	1.0673	1.0004	1.0004	1.0003	0.7759	0.0338	0.0325	0.0343	0.6161	0.0268	0.0258	0.0272
		Cauchy	0.9922	0.9998	1.0011	1.0004	0.7295	0.0314	0.0306	0.0322	0.5766	0.0250	0.0246	0.0255
		Hampel	0.9772	1.0000	1.0014	1.0006	0.7266	0.0316	0.0307	0.0322	0.5741	0.0251	0.0246	0.0255
		GA-RR	**0.9861**	**0.9993**	**1.0013**	**1.0003**	**0.7213**	**0.0313**	**0.0300**	**0.0317**	**0.5699**	**0.0249**	**0.0241**	**0.0250**
3	1	Huber	1.2161	1.0083	0.9985	0.9938	4.7637	0.2557	0.2191	0.1803	3.4387	0.1473	0.1527	0.1431
		Cauchy	1.1052	0.9993	0.9978	0.9962	3.7280	0.1619	0.1666	0.1602	2.9811	0.1272	0.1343	0.1275
		Hampel	1.0968	0.9993	0.9976	0.9966	3.7398	0.1615	0.1670	0.1605	2.9895	0.1267	0.1345	0.1276
		GA-RR	**1.1160**	**0.9992**	**0.9973**	**0.9959**	**3.6370**	**0.1582**	**0.1637**	**0.1569**	**2.8977**	**0.1244**	**0.1318**	**0.1245**
	3	Huber	1.1206	0.9985	1.0003	1.0002	1.2683	0.0619	0.0589	0.0555	0.9984	0.0480	0.0463	0.0439
		Cauchy	1.0020	0.9990	1.0000	1.0006	1.1293	0.0526	0.0519	0.0502	0.8935	0.0418	0.0411	0.0398
		Hampel	0.9941	0.9993	1.0000	1.0005	1.1297	0.0527	0.0520	0.0504	0.8957	0.0421	0.0411	0.0400
		GA-RR	**0.9877**	**0.9990**	**0.9996**	**1.0011**	**1.1046**	**0.0514**	**0.0506**	**0.0493**	**0.8736**	**0.0410**	**0.0401**	**0.0394**
	5	Huber	1.1082	1.0001	1.0010	0.9998	0.8311	0.0377	0.0353	0.0358	0.6635	0.0299	0.0280	0.0279
		Cauchy	1.0272	0.9988	1.0007	0.9989	0.7355	0.0330	0.0309	0.0318	0.5875	0.0258	0.0245	0.0249
		Hampel	1.0134	0.9989	1.0005	0.9992	0.7353	0.0327	0.0310	0.0319	0.5859	0.0258	0.0245	0.0249
		GA-RR	**1.0230**	**0.9985**	**1.0005**	**0.9988**	**0.7174**	**0.0320**	**0.0301**	**0.0308**	**0.5721**	**0.0254**	**0.0239**	**0.0241**

Table 3 shows that the *GA-RR* method has the smallest *RMSE* and *MAE* values. The methods with the smallest *RMSE* and *MAE* values after the *GA-RR* method are Cauchy, Hampel and Huber M-estimator, respectively. Additionally, it is seen that when the outlier is 2 and σ =3, the *RMSE* and *MAE* values of the Hampel M-estimator are smaller than those of the Cauchy M-estimator.

Table 4. Comparison of *GA-RR* method with M-estimators for n=100

Outlier	σ	Method	Average of coefficients				*RMSE*				*MAE*			
			β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
1	1	Huber	1.0760	0.9988	0.9996	0.9977	2.4292	0.1052	0.1060	0.1051	1.9676	0.0847	0.0847	0.0838
		Cauchy	1.0739	0.9988	0.9993	0.9971	2.3925	0.1031	0.1048	0.1035	1.9376	0.0827	0.0837	0.0824
		Hampel	1.0750	0.9987	0.9995	0.9970	2.3963	0.1034	0.1051	0.1039	1.9436	0.0830	0.0837	0.0829
		GA-RR	**1.0616**	**0.9990**	**0.9996**	**0.9975**	**2.3580**	**0.1014**	**0.1034**	**0.1019**	**1.9117**	**0.0816**	**0.0827**	**0.0813**
	3	Huber	0.9919	1.0015	1.0004	1.0008	0.8181	0.0361	0.0345	0.0357	0.6488	0.0287	0.0273	0.0284
		Cauchy	0.9743	1.0017	1.0005	1.0008	0.7974	0.0359	0.0337	0.0350	0.6285	0.0286	0.0267	0.0278
		Hampel	0.9746	1.0016	1.0005	1.0007	0.7966	0.0360	0.0339	0.0349	0.6274	0.0286	0.0268	0.0278
		GA-RR	**0.9713**	**1.0016**	**1.0005**	**1.0008**	**0.7871**	**0.0354**	**0.0334**	**0.0345**	**0.6210**	**0.0281**	**0.0263**	**0.0273**
	5	Huber	1.0223	1.0000	0.9994	1.0002	0.4997	0.0215	0.0213	0.0216	0.4002	0.0172	0.0171	0.0172
		Cauchy	0.9992	1.0001	0.9996	1.0004	0.4952	0.0210	0.0206	0.0215	0.3956	0.0167	0.0167	0.0171
		Hampel	0.9961	1.0001	0.9997	1.0004	0.4987	0.0210	0.0208	0.0216	0.3991	0.0168	0.0168	0.0172
		GA-RR	**0.9888**	**1.0001**	**0.9998**	**1.0006**	**0.4894**	**0.0207**	**0.0204**	**0.0213**	**0.3926**	**0.0166**	**0.0164**	**0.0169**
2	1	Huber	0.9559	1.0005	0.9990	1.0045	2.5558	0.1062	0.1058	0.1107	2.0558	0.0848	0.0832	0.0899
		Cauchy	0.9445	0.9993	0.9981	1.0042	2.5066	0.1039	0.1025	0.1089	2.0076	0.0840	0.0811	0.0880
		Hampel	0.9414	0.9994	0.9985	1.0041	2.5163	0.1039	0.1025	0.1094	2.0188	0.0842	0.0810	0.0884
		GA-RR	**0.9347**	**0.9997**	**0.9983**	**1.0043**	**2.4690**	**0.1016**	**0.1011**	**0.1068**	**1.9687**	**0.0822**	**0.0796**	**0.0863**
	3	Huber	1.0186	1.0008	0.9984	1.0015	0.8774	0.0357	0.0365	0.0382	0.6956	0.0281	0.0291	0.0305
		Cauchy	0.9803	1.0009	0.9985	1.0017	0.8546	0.0345	0.0352	0.0369	0.6776	0.0273	0.0281	0.0296
		Hampel	0.9770	1.0009	0.9985	1.0017	0.8574	0.0345	0.0352	0.0371	0.6805	0.0274	0.0281	0.0297
		GA-RR	**0.9812**	**1.0008**	**0.9983**	**1.0015**	**0.8340**	**0.0339**	**0.0344**	**0.0361**	**0.6632**	**0.0269**	**0.0274**	**0.0292**
	5	Huber	1.0341	1.0004	1.0003	0.9998	0.5152	0.0217	0.0216	0.0220	0.4082	0.0175	0.0173	0.0175
		Cauchy	1.0066	1.0003	1.0002	0.9996	0.4951	0.0209	0.0207	0.0211	0.3887	0.0167	0.0165	0.0168
		Hampel	1.0025	1.0003	1.0002	0.9996	0.4958	0.0210	0.0208	0.0211	0.3879	0.0167	0.0166	0.0167
		GA-RR	**1.0039**	**1.0003**	**1.0001**	**0.9996**	**0.4849**	**0.0205**	**0.0203**	**0.0206**	**0.3798**	**0.0162**	**0.0163**	**0.0163**
3	1	Huber	0.9378	1.0028	1.0057	1.0025	2.6403	0.1103	0.1128	0.1156	2.0968	0.0870	0.0896	0.0922
		Cauchy	0.8834	1.0023	1.0045	1.0034	2.5049	0.1071	0.1065	0.1099	2.0064	0.0851	0.0844	0.0886
		Hampel	0.8809	1.0025	1.0040	1.0036	2.5093	0.1078	0.1069	0.1100	2.0107	0.0860	0.0846	0.0886
		GA-RR	**0.8636**	**1.0033**	**1.0041**	**1.0041**	**2.4405**	**0.1030**	**0.1040**	**0.1078**	**1.9468**	**0.0821**	**0.0822**	**0.0867**
	3	Huber	1.0437	1.0009	1.0005	1.0001	0.8553	0.0374	0.0381	0.0366	0.6792	0.0299	0.0299	0.0287
		Cauchy	0.9866	1.0006	1.0006	1.0004	0.8095	0.0356	0.0366	0.0349	0.6443	0.0284	0.0288	0.0275
		Hampel	0.9802	1.0007	1.0006	1.0004	0.8083	0.0354	0.0367	0.0349	0.6430	0.0284	0.0288	0.0275
		GA-RR	**0.9817**	**1.0006**	**1.0007**	**1.0003**	**0.7889**	**0.0347**	**0.0357**	**0.0341**	**0.6283**	**0.0277**	**0.0282**	**0.0270**
	5	Huber	1.0565	1.0006	0.9998	1.0002	0.5403	0.0232	0.0227	0.0225	0.4346	0.0185	0.0180	0.0179
		Cauchy	1.0088	1.0002	0.9995	1.0003	0.5065	0.0225	0.0216	0.0213	0.4134	0.0180	0.0171	0.0172
		Hampel	1.0023	1.0003	0.9996	1.0003	0.5069	0.0225	0.0215	0.0213	0.4131	0.0180	0.0172	0.0172
		GA-RR	**1.0071**	**1.0000**	**0.9991**	**1.0003**	**0.4943**	**0.0218**	**0.0209**	**0.0208**	**0.4023**	**0.0173**	**0.0166**	**0.0168**

Table 4 shows that the *GA-RR* method has the smallest *RMSE* and *MAE* values. The methods with the smallest *RMSE* and *MAE* values after the *GA-RR* method are Cauchy, Hampel and Huber M-estimator, respectively. Additionally, it is seen that when the outlier is 3, the *RMSE* and *MAE* values of the Hampel and Cauchy M-estimator are close.

Table 5. Comparison of *GA-RR* method with M-estimators for n=300

Outlier	σ	Method	Average of coefficients				*RMSE*				*MAE*			
			β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
1	1	Huber	0.9461	1.0014	1.0016	1.0017	1.4277	0.0608	0.0574	0.0619	1.1519	0.0484	0.0456	0.0494
		Cauchy	0.9491	1.0014	1.0010	1.0016	1.4093	0.0605	0.0571	0.0612	1.1344	0.0481	0.0455	0.0488
		Hampel	0.9410	1.0014	1.0012	1.0019	1.4209	0.0605	0.0572	0.0618	1.1448	0.0480	0.0456	0.0493
		GA-RR	**0.9356**	**1.0016**	**1.0008**	**1.0023**	**1.3840**	**0.0592**	**0.0559**	**0.0601**	**1.1139**	**0.0471**	**0.0447**	**0.0479**
	3	Huber	1.0365	0.9996	0.9996	0.9989	0.4879	0.0206	0.0205	0.0206	0.3942	0.0164	0.0164	0.0165
		Cauchy	1.0335	0.9996	0.9996	0.9989	0.4856	0.0205	0.0204	0.0206	0.3918	0.0164	0.0164	0.0165
		Hampel	1.0336	0.9995	0.9996	0.9988	0.4857	0.0205	0.0204	0.0206	0.3929	0.0164	0.0163	0.0165
		GA-RR	**1.0293**	**0.9996**	**0.9996**	**0.9989**	**0.4733**	**0.0201**	**0.0199**	**0.0200**	**0.3815**	**0.0161**	**0.0160**	**0.0159**
	5	Huber	1.0183	0.9992	0.9993	1.0001	0.2822	0.0116	0.0120	0.0123	0.2250	0.0093	0.0096	0.0098
		Cauchy	1.0192	0.9990	0.9991	1.0000	0.2821	0.0115	0.0120	0.0122	0.2247	0.0092	0.0096	0.0098
		Hampel	1.0134	0.9991	0.9993	1.0001	0.2785	0.0116	0.0120	0.0122	0.2217	0.0092	0.0096	0.0097
		GA-RR	**1.0117**	**0.9990**	**0.9993**	**1.0001**	**0.2710**	**0.0113**	**0.0116**	**0.0118**	**0.2161**	**0.0090**	**0.0092**	**0.0095**
2	1	Huber	1.0485	0.9989	1.0004	0.9983	1.3974	0.0592	0.0594	0.0619	1.1217	0.0475	0.0472	0.0497
		Cauchy	1.0511	0.9989	0.9995	0.9981	1.3826	0.0588	0.0588	0.0612	1.1173	0.0471	0.0467	0.0492
		Hampel	1.0479	0.9988	0.9996	0.9982	1.3918	0.0590	0.0587	0.0618	1.1234	0.0473	0.0466	0.0495
		GA-RR	**1.0443**	**0.9989**	**0.9999**	**0.9981**	**1.3474**	**0.0568**	**0.0573**	**0.0600**	**1.0884**	**0.0456**	**0.0457**	**0.0481**
	3	Huber	1.0382	0.9997	1.0000	0.9990	0.4687	0.0196	0.0203	0.0201	0.3735	0.0154	0.0166	0.0161
		Cauchy	1.0299	0.9998	1.0001	0.9989	0.4612	0.0193	0.0202	0.0198	0.3711	0.0151	0.0164	0.0159
		Hampel	1.0270	0.9998	1.0000	0.9989	0.4629	0.0193	0.0203	0.0199	0.3711	0.0152	0.0165	0.0159
		GA-RR	**1.0270**	**0.9995**	**1.0000**	**0.9988**	**0.4524**	**0.0191**	**0.0195**	**0.0194**	**0.3631**	**0.0150**	**0.0159**	**0.0156**
	5	Huber	1.0215	0.9999	0.9999	0.9998	0.2977	0.0130	0.0129	0.0123	0.2329	0.0102	0.0103	0.0098
		Cauchy	1.0122	0.9999	0.9998	0.9998	0.2932	0.0129	0.0126	0.0121	0.2298	0.0102	0.0100	0.0098
		Hampel	1.0109	0.9998	0.9998	0.9998	0.2930	0.0130	0.0126	0.0121	0.2297	0.0102	0.0101	0.0097
		GA-RR	**1.0029**	**0.9998**	**0.9998**	**0.9999**	**0.2852**	**0.0125**	**0.0122**	**0.0118**	**0.2257**	**0.0099**	**0.0097**	**0.0095**
3	1	Huber	0.9506	0.9986	1.0000	1.0038	1.4205	0.0607	0.0621	0.0602	1.1331	0.0484	0.0490	0.0481
		Cauchy	0.9387	0.9984	1.0001	1.0037	1.3938	0.0605	0.0610	0.0592	1.1109	0.0479	0.0487	0.0470
		Hampel	0.9366	0.9986	0.9997	1.0038	1.4038	0.0605	0.0614	0.0596	1.1212	0.0479	0.0489	0.0474
		GA-RR	**0.9434**	**0.9984**	**0.9996**	**1.0035**	**1.3633**	**0.0594**	**0.0604**	**0.0581**	**1.0792**	**0.0474**	**0.0483**	**0.0457**
	3	Huber	1.0093	0.9991	1.0000	1.0008	0.4507	0.0202	0.0202	0.0196	0.3655	0.0160	0.0159	0.0159
		Cauchy	0.9913	0.9992	1.0000	1.0010	0.4433	0.0199	0.0199	0.0191	0.3600	0.0158	0.0157	0.0155
		Hampel	0.9895	0.9992	1.0001	1.0009	0.4451	0.0199	0.0199	0.0193	0.3610	0.0158	0.0157	0.0156
		GA-RR	**0.9921**	**0.9992**	**1.0000**	**1.0009**	**0.4301**	**0.0193**	**0.0190**	**0.0188**	**0.3496**	**0.0154**	**0.0150**	**0.0153**
	5	Huber	1.0280	0.9999	1.0001	0.9998	0.2975	0.0124	0.0119	0.0124	0.2387	0.0098	0.0094	0.0101
		Cauchy	1.0143	0.9999	1.0000	0.9999	0.2882	0.0122	0.0117	0.0121	0.2308	0.0097	0.0093	0.0098
		Hampel	1.0102	0.9998	0.9999	0.9999	0.2908	0.0122	0.0118	0.0123	0.2317	0.0098	0.0093	0.0099
		GA-RR	**1.0078**	**0.9998**	**0.9997**	**0.9998**	**0.2799**	**0.0119**	**0.0114**	**0.0118**	**0.2242**	**0.0095**	**0.0091**	**0.0096**

Table 5 shows that the *GA-RR* method has the smallest *RMSE* and *MAE* values, while the M estimators have very close values to each other.

Table 6. Comparison of *GA-RR* method with M-estimators for $\sigma_1 = 5$, $\sigma_2 = 3$, $\sigma_3 = 1$

Outlier	n	Method	Average of coefficients				*RMSE*				*MAE*			
			β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
1	30	Huber	0.9270	1.0010	0.9990	1.0069	4.7805	0.0442	0.0860	0.2402	3.6524	0.0351	0.0590	0.1785
		Cauchy	0.9352	1.0009	1.0010	1.0025	4.2844	0.0412	0.0680	0.2102	3.4135	0.0326	0.0534	0.1678
		Hampel	0.9067	1.0007	1.0011	1.0038	4.3136	0.0416	0.0686	0.2117	3.4384	0.0330	0.0537	0.1690
		GA-RR	**0.9425**	**1.0013**	**1.0010**	**1.0019**	**4.1653**	**0.0398**	**0.0664**	**0.2044**	**3.3178**	**0.0318**	**0.0520**	**0.1633**
	50	Huber	0.9746	0.9994	1.0012	1.0027	3.1510	0.0318	0.0497	0.1554	2.4589	0.0250	0.0393	0.1212
		Cauchy	0.9466	0.9991	1.0012	1.0025	3.0760	0.0311	0.0482	0.1510	2.4030	0.0247	0.0381	0.1180
		Hampel	0.9278	0.9991	1.0012	1.0033	3.0740	0.0312	0.0484	0.1510	2.4042	0.0248	0.0383	0.1179
		GA-RR	**0.9257**	**0.9993**	**1.0014**	**1.0032**	**3.0415**	**0.0307**	**0.0478**	**0.1493**	**2.3807**	**0.0247**	**0.0382**	**0.1166**
	100	Huber	1.0640	0.9993	1.0002	0.9981	2.2094	0.0215	0.0357	0.1091	1.7562	0.0170	0.0286	0.0863
		Cauchy	1.0549	0.9994	1.0004	0.9976	2.1612	0.0211	0.0350	0.1067	1.7303	0.0168	0.0279	0.0855
		Hampel	1.0575	0.9993	1.0003	0.9975	2.1703	0.0211	0.0353	0.1071	1.7329	0.0168	0.0281	0.0855
		GA-RR	**1.0472**	**0.9994**	**1.0005**	**0.9978**	**2.1273**	**0.0209**	**0.0348**	**0.1050**	**1.7063**	**0.0166**	**0.0277**	**0.0844**
	300	Huber	0.9090	1.0000	1.0002	1.0049	1.2019	0.0125	0.0199	0.0593	0.9481	0.0099	0.0159	0.0469
		Cauchy	0.9116	0.9999	1.0002	1.0046	1.1979	0.0125	0.0198	0.0591	0.9438	0.0099	0.0158	0.0467
		Hampel	0.9122	0.9999	1.0002	1.0045	1.2014	0.0125	0.0198	0.0592	0.9459	0.0099	0.0158	0.0468
		GA-RR	**0.9184**	**0.9999**	**1.0001**	**1.0042**	**1.1845**	**0.0122**	**0.0193**	**0.0585**	**0.9352**	**0.0096**	**0.0155**	**0.0462**
2	30	Huber	1.2516	1.0027	0.9956	0.9963	7.1320	0.0655	0.0814	0.3528	4.2185	0.0413	0.0626	0.2056
		Cauchy	1.1873	1.0013	0.9955	0.9930	4.4217	0.0438	0.0719	0.2145	3.4262	0.0343	0.0560	0.1669
		Hampel	1.1478	1.0016	0.9961	0.9943	4.3763	0.0441	0.0711	0.2133	3.4385	0.0345	0.0561	0.1656
		GA-RR	**1.1470**	**1.0008**	**0.9956**	**0.9947**	**4.3568**	**0.0435**	**0.0709**	**0.2122**	**3.4001**	**0.0340**	**0.0560**	**0.1652**
	50	Huber	0.9576	1.0005	1.0023	1.0049	3.7464	0.0338	0.0595	0.1861	2.8270	0.0267	0.0445	0.1394
		Cauchy	0.9361	1.0000	1.0027	1.0023	3.2375	0.0316	0.0527	0.1603	2.5910	0.0251	0.0415	0.1279
		Hampel	0.9451	1.0001	1.0029	1.0015	3.2413	0.0314	0.0525	0.1604	2.5828	0.0252	0.0414	0.1275
		GA-RR	**0.9467**	**1.0000**	**1.0022**	**1.0017**	**3.1881**	**0.0311**	**0.0517**	**0.1581**	**2.5522**	**0.0248**	**0.0408**	**0.1262**
	100	Huber	1.0000	0.9998	0.9993	1.0027	2.1777	0.0220	0.0371	0.1073	1.7255	0.0174	0.0293	0.0851
		Cauchy	0.9640	0.9995	0.9991	1.0030	2.0894	0.0211	0.0367	0.1028	1.6525	0.0167	0.0290	0.0813
		Hampel	0.9688	0.9996	0.9993	1.0024	2.0883	0.0211	0.0367	0.1027	1.6503	0.0167	0.0290	0.0811
		GA-RR	**0.9546**	**0.9995**	**0.9989**	**1.0033**	**2.0396**	**0.0205**	**0.0360**	**0.1005**	**1.6168**	**0.0163**	**0.0284**	**0.0794**
	300	Huber	0.9976	1.0007	0.9999	1.0008	1.2493	0.0117	0.0196	0.0611	0.9977	0.0093	0.0156	0.0485
		Cauchy	0.9962	1.0008	0.9998	1.0004	1.2396	0.0116	0.0192	0.0604	0.9837	0.0092	0.0153	0.0477
		Hampel	0.9919	1.0007	0.9999	1.0005	1.2337	0.0117	0.0193	0.0601	0.9821	0.0093	0.0153	0.0477
		GA-RR	**0.9938**	**1.0007**	**0.9998**	**1.0003**	**1.1936**	**0.0113**	**0.0186**	**0.0581**	**0.9523**	**0.0089**	**0.0149**	**0.0461**
3	30	Huber	1.2340	1.0020	0.9953	1.0013	5.6421	0.0567	0.0924	0.2761	4.2880	0.0437	0.0736	0.2112
		Cauchy	1.0890	1.0020	0.9956	0.9979	4.4446	0.0425	0.0747	0.2175	3.4409	0.0339	0.0593	0.1687
		Hampel	0.9352	0.9999	0.9949	1.0065	6.1556	0.0837	0.0757	0.3308	3.5859	0.0364	0.0596	0.1769
		GA-RR	**1.0931**	**1.0016**	**0.9951**	**0.9976**	**4.4043**	**0.0415**	**0.0730**	**0.2157**	**3.4253**	**0.0332**	**0.0576**	**0.1678**
	50	Huber	1.2336	0.9992	0.9988	0.9951	3.6599	0.0355	0.0597	0.1806	2.8371	0.0282	0.0469	0.1395
		Cauchy	1.1634	0.9985	0.9983	0.9934	3.2377	0.0320	0.0521	0.1592	2.5680	0.0251	0.0410	0.1260
		Hampel	1.1555	0.9987	0.9984	0.9935	3.1920	0.0318	0.0522	0.1572	2.5424	0.0250	0.0415	0.1251
		GA-RR	**1.1177**	**0.9984**	**0.9983**	**0.9953**	**3.1261**	**0.0313**	**0.0513**	**0.1531**	**2.5366**	**0.0245**	**0.0407**	**0.1243**
	100	Huber	1.1027	0.9999	0.9993	0.9982	2.2260	0.0232	0.0380	0.1098	1.7924	0.0185	0.0301	0.0878
		Cauchy	1.0137	1.0000	0.9994	1.0000	2.0983	0.0221	0.0363	0.1037	1.6832	0.0177	0.0289	0.0830
		Hampel	1.0061	1.0000	0.9996	1.0001	2.1110	0.0222	0.0363	0.1042	1.6941	0.0178	0.0289	0.0835
		GA-RR	**1.0093**	**0.9998**	**0.9996**	**1.0000**	**2.0652**	**0.0217**	**0.0355**	**0.1017**	**1.6518**	**0.0173**	**0.0281**	**0.0813**
	300	Huber	1.0162	0.9991	1.0000	1.0005	1.2690	0.0122	0.0200	0.0625	1.0134	0.0097	0.0163	0.0500
		Cauchy	0.9902	0.9992	1.0002	1.0010	1.2491	0.0120	0.0196	0.0616	1.0019	0.0096	0.0159	0.0494
		Hampel	0.9914	0.9991	1.0002	1.0008	1.2483	0.0120	0.0197	0.0615	1.0018	0.0095	0.0159	0.0496
		GA-RR	**0.9841**	**0.9990**	**1.0001**	**1.0011**	**1.2028**	**0.0117**	**0.0191**	**0.0593**	**0.9616**	**0.0093**	**0.0155**	**0.0477**

Table 6 shows that the *GA-RR* method has the smallest *RMSE* and *MAE* values. After the *GA-RR* method, the methods with the smallest *RMSE* and *MAE* values are seen to be the Cauchy and Hampel M-estimators, which have values close to each other. The Huber M-estimator follows this.

3.2. Real Data

The stack loss data is obtained from 21 days of operation of a plant. In the Stack loss data Table 7, i is the observation index (days), Y is the stack loss, X_1 is the air flow, X_2 is the cooling water inlet temperature and X_3 is the acid concentration [25].

Table 7. Stack loss data

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Y	42	37	37	28	18	18	19	20	15	14	14	13	11	12	8	7	8	8	9	15	15
X_1	80	80	75	62	62	62	62	62	58	58	58	58	58	58	50	50	50	50	50	56	70
X_2	27	27	25	24	22	23	24	24	23	18	18	17	18	19	18	18	19	19	20	20	20
X_3	89	88	90	87	87	87	93	93	87	80	89	88	82	93	89	86	72	79	80	82	91

Stack loss data has been used in many studies [26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36]. In some of these studies, it is stated that 1, 2, 3, 4 and 21 observations may be outliers, while in others it is stated that some of these observations may be outliers. The results regarding the outliers detected by the methods used in this study are given in Figure 2. In Figure 2, i is the observation index, and W is the weights in the range [0,1].

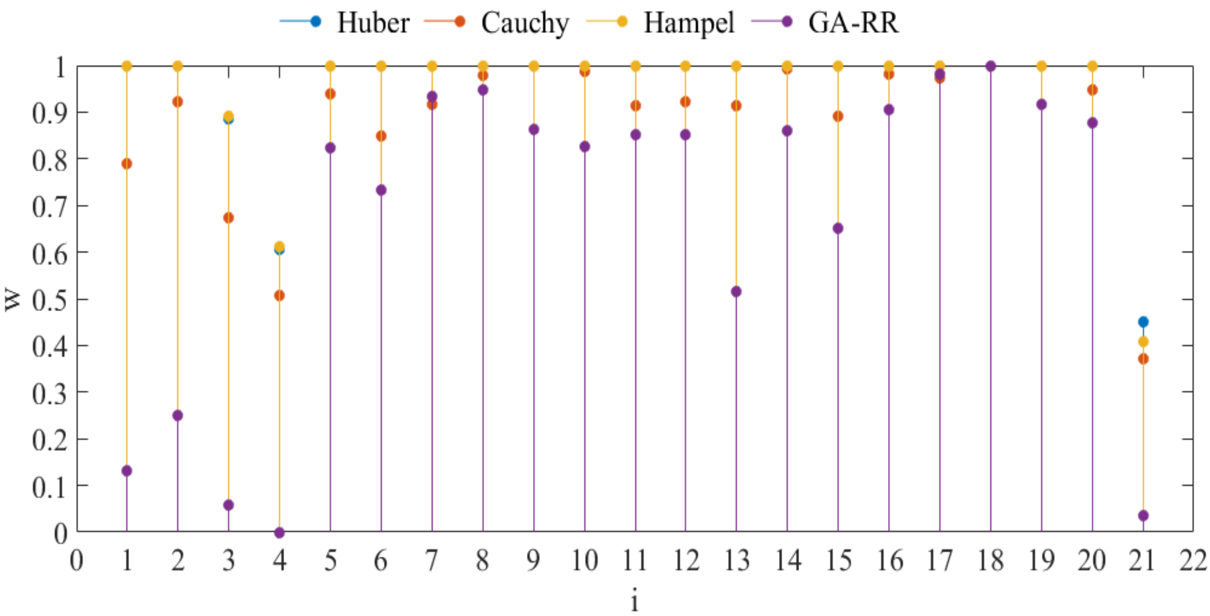

Figure 2. Observation weights of the methods

Figure 2 shows that the Huber and Hampel M-estimator detects the 3rd, 4th and 21st observations as outliers, while the Cauchy M-estimator and *GA-RR* method detects the 1st, 3rd, 4th and 21st observations as outliers. It is seen that the *GA-RR* method gives low weight to the 2nd observation, which is said to be an outlier in the studies. The results of the parameter estimates of the *GA-RR* method with M-estimators for the stack loss data are given in Table 8.

Table 8. Parameter estimates of the methods

Method	β_0	β_1	β_2	β_3
Huber	-41.012	0.813	0.990	-0.132
Cauchy	-40.650	0.810	0.961	-0.127
Hampel	-41.193	0.820	0.970	-0.130
GA-RR	**-37.360**	**0.830**	**0.563**	**-0.088**

Besides, since the *GA-RR* method gives much less weight to outliers than M-estimators, the β_i values differ. The results obtained without the 1st, 3rd, 4th and 21st observations, which were identified as outliers, are given in Table 9.

Table 9. Stack loss data results of the methods

Method	*RMSE*	*MAE*
Huber	1,469	1,158
Cauchy	1,419	1,119
Hampel	1,456	1,150
GA-RR	**1,002**	**0,722**

Table 9 shows that the *GA-RR* method has lower *RMSE* and *MAE* values than the M-estimators.

4. Conclusion

M-estimators and *GA-RR* method were compared with the simulation study. As a result of the comparison, it is seen that the parameter estimation *RMSE* and *MAE* values obtained with the *GA-RR* method are smaller than the parameter estimation *RMSE* and *MAE* values obtained with M-estimators. As the sample size increases the *RMSE* values of the *GA-RR* method and M-estimators become closer to each other. This is due to the gradual decrease in the percentage of outliers in sample sizes. When M-estimators are compared among themselves, the Cauchy M-estimator usually has the smallest *RMSE* value, and then the Hampel M-estimator.

It was compared with M-estimators and the *GA-RR* method using stack loss data, which is the real data set. As a result of the comparison, it is seen that the *GA-RR* method can detect outliers detected by M-estimators and can also detect other observations that may be outliers. The *GA-RR* method makes a sound parameter estimate by giving weights to the observations in the data set in proportion to their contradictions. In this context, it can be said that the *GA-RR* method can also be used as an outlier detection method.

In conclusion, an alternative new method for robust regression estimators has been developed to make robust parameter estimations in this study. Since the *GA-RR* method is resistant to outliers, it can make reliable parameter estimation. Therefore, the *GA-RR* method can contribute to the reliable examination of the variables that are the subject of research in several different fields of study.

References

[1] E. Ronchetti, 2006, The Historical Development of Robust Statistics. ICOTS-7.

[2] H. Y. Wu, L. D. Wu 2005, A robust estimation method in orbit improvement. *Chinese Astronomy and Astrophysics*, 29(4), 430-437.

[3] M. Hu, W. M. Zhang, M. Zhong, 2017, Robust regression and its application in absolute gravimeters. *Review of Scientific Instruments*, 88(5), 054501.

[4] P. Leoni, P. Segaert, S. Serneels, T. Verdonck, 2018, Multivariate constrained robust M-regression for shaping forward curves in electricity markets. *Journal of Futures Markets,* 38(11), 1391-1406.

[5] Q. Su, Y. Bommireddy, Y. Shah, S. Ganesh, M. Moreno, J. Liu, ... Z. K. Nagy, 2019, Data reconciliation in the Quality-by-Design (QbD) implementation of pharmaceutical continuous tablet manufacturing. *International journal of pharmaceutics,* 563, 259-272.

[6] D. E. Goldberg, 1989, Genetic algorithms in search, optimization, and machine learning. Addison. Reading.

[7] A. Hussain, S. Riaz, M. S. Amjad, E. U. Haq, 2022, Genetic algorithm with a new round-robin based tournament selection: Statistical properties analysis. *PloS one,* 17(9), e0274456.

[8] P. Vankeerberghen, J. Smeyers-Verbeke, R. Leardi, C. L. Karr, D. L. Massart, 1995, Robust regression and outlier detection for non-linear models using genetic algorithms. *Chemometrics and Intelligent Laboratory Systems,* 28(1), 73-87.

[9] Y. C. Hu, 2009, Functional-link nets with genetic-algorithm-based learning for robust nonlinear interval regression analysis. *Neurocomputing,* 72(7-9), 1808-1816.

[10] P. Wiegand, R. Pell, E. Comas, 2009, Simultaneous variable selection and outlier detection using a robust genetic algorithm. *Chemometrics and Intelligent Laboratory Systems,* 98(2), 108-114.

[11] D. Sykas, V. Karathanassi, 2015, An automatic method for producing robust regression models from hyperspectral data using multiple simple genetic algorithms. *In Third International Conference on Remote Sensing and Geoinformation of the Environment* (RSCy2015) (Vol. 9535, p. 953502). SPIE.

[12] A. Duraj, Ł. Chomątek, 2017, Outlier detection using the multiobjective genetic algorithm. *Journal of Applied Computer Science, 25*(2), 29-42.

[13] A. Toy, 2022, Robust regresyon tahmin edicilerine yönelik yeni bir yaklaşım. *Doktora Tezi.* Ondokuz Mayıs Üniversitesi, Lisansüstü Eğitim Enstitüsü.

[14] P. J. Rousseeuw, A. M. Leroy, 1987, Robust regression and outlier detection John Wiley and Sons. Inc., New York.

[15] R. A. Maronna, R. D. Martin, V. J. Yohai, 2006, Robust Statistics Theory and Methods John Wiley and Sons. Inc., USA.

[16] P. Pennacchi, 2008, Robust estimate of excitations in mechanical systems using M-estimators—Theoretical background and numerical applications. *Journal of Sound and Vibration,* 310(4-5), 923-946.

[17] P. W. Holland, R. E. Welsch, 1977, Robust regression using iteratively reweighted least-squares. *Communications in Statistics-theory and Methods,* 6(9), 813-827.

[18] J. Fox, 2002, Robust regression: Appendix to an R and S-PLUS companion to applied regression.

[19] Z. Zhang, 1997, Parameter estimation techniques: A tutorial with application to conic fitting. *Image and vision Computing,* 15(1), 59-76.

[20] D. B. Özyurt, R. W. Pike, 2004, Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes. *Computers and chemical engineering,* 28(3), 381-402.

[21] M. Kumar, D. Husain, N. Upreti, D. Gupta, 2010, Genetic algorithm: Review and application. Available at SSRN 3529843.

[22] Z. Michalewicz, M. Schoenauer, 1996, Evolutionary algorithms for constrained parameter optimization problems. Evolutionary computation, 4(1), 1-32.

[23] X. S. Yang, 2020, Nature-inspired optimization algorithms. Academic Press.

[24] H. Jiawei, K. Micheline, P. Jian, 2016, Data mining concepts and techniques.

[25] K. A. Brownlee, 1965, Statistical Theory and Methodology in Science and Engineering. 2nd ed., John Wiley and Sons, New York.

[26] N. R. Draper, H. Smith, 1966, Applied Regression Analysis. John Wiley and Sons, New York.

[27] D. F. Andrews, D. Pregibon, 1978, Finding the outliers that matter, J. R. Stat. SOC. Series B.

[28] A. P. Dempster, M. Gasko-Green, 1981, New tools for residual analysis, Ann. Stat., 9, 945-959.

[29] W. J. J. Rey, 1983, Introduction to Robust and Quasi-Robust StatisticalMethods. Berlin, Springer.

[30] F. R. Hampel, E. M. Ronchetti, P.J. Rousseeuw, W. A. Stahel, 1986, Robust Statistics. The Approach Based on Influence Functions. Wiley, New York.

[31] S. Chatterjee, M. Machler, 1997, Robust regression : a weighted least squares approach, *Communications in Statistics - Theory and Methods*, 26:6, 1381-1394.

[32] D. Birkes, Y. Dodge, 1990, Alternative Methods of Regression. John Wiley& Sons, New York.

[33] P. Cizek, J. A. Visek, 2000, Least trimmed squares, Interdisciplinary Research Project 373: No. 2000, 53, Humboldt University of Berlin, *Quantification and Simulation of Economic Processes,* Berlin.

[34] J. Jureckova, J. Picek, 2006, Robust Statistical Methods with R, Chapman & Hall/CRC, Boca Rotan, FL.

[35] P. J. Rousseeuw, M. Hubert, 2018, Anomaly detection by robust statistics. WIREs *Data Mining and Knowledge Discovery,* 8:e1236.

[36] H. S. Seo, 2019, Unified methods for variable selection and outlier detection in a linear regression. *Communications for Statistical Applications and Methods,* Vol. 26, No. 6, 575–582.