Using Some Wavelet Shrinkage Techniques and Robust Methods to Estimate the Generalized Additive Model Parameters in Non-Linear Models

Bashar A. Majeed Al-Talib[a], Ala'a A. Hammodat[b]

[a]*Department of Statistics and Informatics, Faculty of Computer Sciences and Mathematics, University of Mosul, Iraq*
Email: bashar.altalib@uomosul.edu.iq

[b]*Faculty of Computer Sciences and Mathematics, University of Mosul, Iraq*

Abstract*—In this paper, the method of estimating the Generalized Additive Models (GAM) was highlighted, and a proposed robust weighted composition was found by combining the robust M method with the smoothing splines to estimate the Robust Generalized Additive Model and its notation is (RGAM). This estimator is used to deal with the effect of outliers' presence in the data that do not fit into the overall data pattern by relying on some of the weight functions of the robust M method. Wavelet Shrinkage technique is used as well, which has been proposed as a smoothing of data using several types of wavelet filters in calculating the discrete wavelet transformation and relying on it in estimating the wavelet generalized additive model symbolized by (WGAM). In using the simulation method, when data is contaminated with distributions ((t) Dis., Exp. Dis.) And with contamination rates (5%, 15%, 35%) and with sample sizes (50,150,300) it is noted that the smoothing method is with the Bisequare weight (BRGAM). It had a better performance compared to the rest of the methods for the simulated scenarios covered. The GCV criterion showed a marked advantage over other criteria, especially when estimating the proposed robust M (RGAM) model. Some statistical criteria have been adopted. These criteria of the Generalized Additive Model (GAM) is used to compare estimation methods, the proposed methods were tested on simulation experiments as well as on real data collected from Ibn Sina Learning Hospital on cases of short stature. The RGAM method gave the best results compared to the ordinary GAM and WGAM methods, and that by obtaining the smallest GCV value, this is because it is responsible for selecting the most suitable smoothing parameter for the smoothing spline estimator.

***Keywords*—generalized additive model; wavelet shrinkage; robust estimator; M-estimator; GCV.**

I. INTRODUCTION

The additive model (GAM) is one of the methods of Non-Parametric (or semi-parametric) regression. It is one of the practical solutions, especially when it is combined with the method of backfitting to deal with the problem of dimensionality that makes the researchers suffer using multiple Non-Parametric regression. It is restricted from the possibility of generalizing the case of univariate regression to a multivariate case [1]. When an approximate estimator of two or more variables is found, it is difficult to make a matrix of variables measured in different units and other problems related directly and indirectly. The additive method works in accumulation (sum of one-dimensional compounds) to interpret various phenomena. The GAM is a compromise between multiple regression and the matching of a surface with several dimensions, using Partial Residuals, and it can also be considered promoted Non-Parametric state of the Generalized Linear Model (GLM) and is obtained by replacing $(\sum_{i=1}^{p} x_i\beta_i)$ Linear Predictor with other Non-Parametric terms as additive Predictor $(\sum_{i=1}^{p} m_i(x_i))$; therefore, the Non-Parametric additive model [2] is shown below.

$$y_i = \alpha + \sum_{i=1}^{n} m_i(x_i) + \varepsilon_i \tag{1}$$

m_i: is a vector of estimated function vector of data that expresses the explanatory variable x_i and can be estimated by smoothing splines. There are several methods for obtaining a GAM estimate. The most important of them is the Backfitting algorithm [3].

This paper aims to make data smoothing (filtering) from outliers using some fortified weighting functions of robust M-estimators in estimating the generalized additive model to obtain an estimator symbolized as (RGAM) [4]. This paper aims as well to use some types of wavelets as a filter in calculating the discrete wavelet transformation and using the smoothed data to estimate the proposed generalized additive model (WGAM). Finally, the results' efficiency is tested using some statistical comparison criteria to determine the best estimation method among the ordinary GAM models and the proposed RGAM methods, and the proposed WGAM methods using simulation and real data analysis.

The first step to introduce the backfilling algorithm [5] depends on the iterative method. Fitting the equation (1), the conditional expectation of the response variable for each K of x ^ 's will be as follows:

$$\begin{aligned} \mathrm{E}(\mathrm{Y}/\mathrm{X} = x_1, x_2, \ldots, x_\mathrm{k}) &= m(x_1, x_2, \ldots, x_\mathrm{k}) \\ &= m_1(x_1) + m_2(x_2) + \cdots + m_\mathrm{k}(x_\mathrm{k}) \end{aligned}$$

The first step for estimating this equation is done by starting with initial values as an initial step (°) in $m_i = \mathring{m}_i$, i = 1,2, ..., p

As for the second step, it is carried out by doing the i-th iteration:

$$\widehat{m}_i = S_i\left(y - \alpha - \textstyle\sum_{k=i}^{p} m_k / x_i\right) \qquad (2)$$

S is a smoothing transformation matrix (n × n).

$S_i = (I + \lambda_i K_i)^{-1}$
$K_i = Q_i R_i^{-1} Q_i'$

K_i is recalculated for each i of the explanatory variables. To find new estimates of the functions, these operations are repeated by smoothing the errors until the partial functions converge and stopping when the smoothing functions $m_i(\mathrm{x}_i)$ remain converges [6]:

$$\Delta\left(m_i^d, m_i^{d-1}\right) = \sum_{i=1}^{p} \left(m_i^d, m_i^{d-1}\right)^2$$

Partial residuals are defined as the identical values for each function plus the total residuals from the additive model provided that:

$$E\big(m_i(x_i)\big) = 0$$

The S_i smoothing matrix is replaced by another matrix S_i^*, this is called the Centered Smoother so that:
$S_i^* = (I - 11')S_i$
Since 1 is a vector (nxn) all its elements equal one, and when generalizing equation (3), the GAM will be obtained:

$$\textstyle\sum_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{y_i} - a - m(x_\mathrm{i}))^2 + \sum_{i=1}^{n} \lambda \int_\mathrm{a}^\mathrm{b} m''(x_i)^2 \, \mathrm{dx} \qquad (3)$$

Whereas λ represents the smoothing parameter, it is noted from equation (3) that it is a penalty parameter with a separate constant λ_i for each term, and equation (3) can be written in a matrix form [7] as follows:

$$(\mathrm{y_i} - a - \textstyle\sum_{i=1}^{n} m_i)'(\mathrm{y_i} - a - \sum_{i=1}^{n} m_i) \qquad (4)$$

k_i is a penalty matrix for each x_i, and by taking the differentiation for equation (4) with respect to m_i, and make the results equals zero, we obtain:

$$\widehat{m}_i = S_k(y - a - \textstyle\sum_{i=1}^{n} \widehat{m}_i) \qquad (5)$$

Whereas $S_k = (I + \lambda_k \mathrm{K}_k)^{-1}$, the amount $(y - a - \sum_{i=1}^{n} \widehat{m}_i)$ is called the partial residuals from the smoothing, and then the smoothing process is repeated or repeated for the partial residuals until we get the required convergence.

To choose the smoothing parameter (λ), automated method is used, and one of the most known criteria for selecting the smoothing parameter is the Generalized Cross Validation criterion ((GCV_λ) Generalized Cross-Validation), and this is done by reducing this criterion, which in turn depends on the residuals, as it leaves the data to determine by itself the optimal value of the smoothing parameter [8].

II. Materials and Method

A. *The Concept of Robust Regression*

The problem of the existence of outliers in the data has received significant attention in recent years. Many researchers' awareness with extreme values in the data are often associated with a violation of the assumptions of errors that are supposed to be distributed normally. The idea of robust statistics based on statistical treatments deals with some deviations from the premises of the ideal model that are sometimes associated with outlying values in the data. Robust methods usually reduce the impact of those extreme values on the estimate, as the robust method used to diagnose, isolate, and prevent it from withdrawing (pulling) the model estimated towards it [9]–[11].

B. *The Robust M-estimate with Smoothing Splines*

Huber introduced the robust M method in 1964. It is one of the robust methods of estimation, which gives less weight to extreme observations in the dependent variable to reduce its effect (effect of large residuals). The M-estimator corresponds to the estimates of the maximum likelihood because the function ρ (.) becomes a likelihood function when choosing a suitable distribution of residuals [12]. The penalized likelihood estimator can be obtained as the following criterion is maximized [13]. Penalized Likelihood is considered a generalization of the penalized squares method's concept if the probability distributions are from the exponential family. The exponential distributions are the basis upon which all derivations depend. Because of the characteristics that make it distinct in the inferential part, which makes most researchers inclined to this type of probability functions, the general form for these distributions is shown in equation 6 [14] below.

$$f(x, \theta, \varphi) = exp\left\{\frac{y\theta - b(\theta)}{\alpha\varphi} + c(y, \varphi)\right\} \qquad (6)$$

as: -θ: normal parameter. φ: Scale Parameter. c, b: functions that the shape of the distribution depends on. We find the log of the maximum likelihood as follows:

$$L(m, \varphi) = \textstyle\sum_{i=1}^{n} \left\{\frac{y_i m(x_i) - b(m(x_i))}{\alpha\varphi} + c(y_i, \varphi)\right\} \qquad (7)$$

$m(x_i)$: It represents an unknown smoothing function that requires estimation based on the sample observations and will be considered the Canonical Parameter. $\widehat{m}$ is considered as the solution that maximizes the logarithm of the likelihood function of equation (7), and then the penalized maximum likelihood estimator can be obtained, and the following criterion is maximized [15].

$$L(m,\varphi)-\frac{1}{2}\lambda\int m''(x_i)^2dx) \quad (8)$$

To maximize the log of the penalized likelihood function, we multiply equation (7) by $\alpha\varphi$:

$$\sum_{i=1}^{n}[y_i m(x_i)-b(m(x_i))] \quad (9)$$

Then maximizing equation (9) will be equivalent to minimizing equation (3). Then Minimizer of Penalized Least Squares (MPLS) will be obtained. This will be used in equation (9) to balance between the amount of smoothing for the fitted curve m which is the minimum values for non-smoothed penalty, and between the accuracy of the data, which are the higher values for the log likelihood. And when the estimation of the extremes is to be robustified, it is preferable to use the robust M method with the smoothing splines in the Non-Parametric analysis, because M-estimators method possesses the property of Scale Invariant, and by taking the standardized residuals using equation (10) as follows:

$$e_i=\frac{(y_i-a-m(x_i))}{\hat{\sigma}} \quad (10)$$

Estimates are determined by a specific objective function on all values of m, and by minimizing the criterion:

$$min\,\lambda_i\int_0^1 m''(x_i)^2dx \quad (11)$$

Or by using the matrix formula:

$$\sum_{i=1}^{n}\hat{\sigma}^2\rho\{e_i\}^2+\sum_{i=1}^{n}\lambda_i\, m_i'\, k_i m_i=0 \quad (12)$$

Let Ψ = ρ ^ 'represent the derivative of ρ, where Ψ is called the Influence curve, and to minimize function (12) we derive it partially with respect to the parameters m and make the resulted amount equal to zero:

$$-\hat{\sigma}\Psi+\sum_{i=1}^{n}\lambda_i\, m_i'\, k_i m_i=0$$

Whereas, ρ is a function concerning errors, and the M value does not have to be a fixed estimator; that is, the estimators may be affected by the size of the errors. To find σ̂, which represents the measurement parameter, it is estimated only once before starting the iteration, using the initial values, and there are several formulas for estimating $\hat{\sigma}$ including [16].

Since e_i represents the residuals and that the value of σ is approximately due to an unbiased estimator of the standard deviation of errors when n is large, and the error is a normally distributed, and that the function $\sum_{i=1}^{n}\rho(\frac{y_{i-a}-m(x_i)}{\sigma})$ is a lower bound by the first partial derivative of $\rho(.)$.

As $\Psi(x)$ represents the influence function, that is, it measures the extent of the effect of observation p, and the researchers have proposed a few functions $\rho(.)$ and their derivatives $\Psi(.)$ So that they make the estimator robust and not affected by the presence of outliers. The weight function can be defined as:

$$w(x)=\frac{\Psi(x)}{x} \quad (13)$$

$$w_i=w(e_i)=\frac{\Psi(e_i)}{e_i} \quad (14)$$

Accordingly, the new estimator will be as follows:

$$S=(W+\lambda_i K_i)^{-1}W \quad (15)$$

$$min\sum_{i=1}^{n}w\left(e_i^{(v-1)}\right)e_i^2 \quad (16)$$

Where v indicates the index of the iteration. The weight $w\left(e_i^{(v-1)}\right)$ is recalculated after each iteration in order to use it in the next iteration.

C. *Some M Robust Weighting Function*

Here are some commonly used weight weights.

1) Huber function: M-estimators are based on the Huber function with mathematical advantages [17]. However, it is sensitive to the Leverage Point, and increase linearly at the $|x|>c$ level, where a 95% approximation efficiency is obtained when errors are distributed normally with the Tuning Constant $c=1.345$.

$$\Psi_{Huber(e_i,c)}=\begin{cases}1 & if \quad |e_i|\le c\\ \frac{c}{|e_i|} & if \quad o.w\end{cases} \quad (17)$$

2) Hampel function [18]*:* as the default values for the cutoff constants a=2 , b=4 , c=8

$$\Psi_{(e_i,c)}=\begin{cases}1 & if \quad |e_i|\le a\\ \frac{a}{|e_i|} & if \quad a\le|e_i|\le b\\ \frac{a(c-|e_i|)}{|e_i|(a-b)} & if \quad b\le|e_i|\le c\\ 0 & \quad |e_i|\ge c\end{cases} \quad (18)$$

3) Bisquare or Tukey Beaton function is sometimes called a double squared weight function (Tukey, Biweight) [19], which reaches 95% efficiency when errors are distributed normally.

$$\Psi_{(e_i,c)}=\begin{cases}\left[1-\left(\frac{e_i}{c}\right)^2\right]^2 & if \quad |e_i|\le c\\ 0 & if \quad 0\end{cases} \quad (19)$$

D. *Wavelet Transform*

Wavelet transformation is one of the types of mathematical functions. It divides the original signal (data) or partitioning a given function into different frequency compounds and studying each compound with the appropriate resolution at each measurement. In other words, dividing the functions into several frequential components using different Windows sizes, and then studying these components separately, taking into account the match of the range (Scale) and the used wavelet.

Wavelet transformation analyzes the function or the time series within the range of time and frequency. Wavelet transformation is used with short time and high-frequency signals, which gives good time accuracy and weak frequency accuracy, as well as used with a long time and low frequency, which gives low time accuracy and good frequency accuracy. The wavelet can be defined mathematically as a real value function defined on an entire real axis and oscillating up and down regularly around zero. The wavelet is also considered a distinctive tool, being an effective and powerful technique for representing and analyzing data. The wavelet was developed mathematically to be wavelet for its smallness [20]. It is a signal of limited continuity with a mean equal to zero, unlike the big wave signal such as the sine wave and the cosine wave that extends (-∞ and ∞). The wavelet compounds can be described as follows:

1) The Scaling Function $\emptyset(.)$, Which is also known as the Father Function [21], which represents the dilation equation, and is considered the approximate part of the data (which is proportional to the data mean), which we obtain from the following formula:

$$\mathrm{f}(x) = \sum_{\mathrm{k}=0}^{\mathrm{N}} C(\mathrm{k})\mathrm{f}(2x - \mathrm{k}) \quad (20)$$

Whereas, C (k): represents the parameters of the Low-Pass Filter.

2) The Wavelet function Ψ (.), Which is also known as the Mother function [22], which represents the Wavelet equation, which we get from the following formula: -

$$\mathrm{w}(x) = \sum_{\mathrm{k}=0}^{\mathrm{N}} \mathrm{d}(\mathrm{k})\mathrm{f}(2x - \mathrm{k}) \quad (21)$$

Whereas, d (k) represents the High-Pass Filter parameters, as it acts as a prototype, in which all used windows to process the time series signal are generated from it.

E. Haar Wavelet

Haar Wavelet [23] is an example of the Orthonormal system in the interval [1.0], and it is considered one of the simplest and oldest types of wavelets and is best for educational purposes, and it is the basis for generating other types of wavelets. The Haar wavelet consists of two functions, the wavelet function $\Psi(\mathrm{X})$ and the scaling function $\emptyset(\mathrm{X})$ (Scaling Function).

III. Results and Discussion

A. Used Generation Functions

The functions vary in the diversity of the phenomena that they represent, as these functions are characterized by being designed to display a set of phenomena that often occur in real life, and two accredited functions have been employed in most research papers, namely: -

- Linear Function of higher degrees: $\mathrm{f}_1(\mathrm{x})$ =0.2*x^11 $*(10*(1-\mathrm{x}))$^6 + 10*(10*x)^3*(1-x)^10
- Doppler function [24]

$f_2(x) = \{x(1-x)\}^{1/2} sin\{2\pi(1+\varepsilon)/(x+\varepsilon)\}$, $\varepsilon = 0.05$

B. Simulation Trials Algorithm

Several scenarios were applied to simulation experiments, as the explanatory variables were contaminated at one time and y at other times with different distributions ((t) dis., Exp.dist.), And for different sample sizes (50, 150, 300) and with contamination ratios (5, 15, 35%), then repeat each experiment once and to obtain consistent results and to give a comprehensive picture of the efficiency of the methods, different parameters were chosen for the probability distributions as follows:

- Generate four explanatory variables Standard Uniform Distribution.
- Generate random errors from a normal distribution with a mean of zero and variance σ^2.
- Generate the random variable y directly through the model used in simulation experiments, using the regression function in terms of the explanatory variables generated above and random error.
- Estimating the Generalized Additive Model GAM model and then smoothing the data with wavelet functions (Db, Haar, Least A., Coiflets) to estimate the proposed Wavelet Generalized Additive Model (WGAM) to obtain the smoothed estimators (DWGAM, HWGAM, LWGAM, CWGAM).
- Estimating the proposed Robust Generalized Additive Model (RGAM) using some weight functions of the robust M-estimators method (Huber, Hampel, Bisquare) to obtain the smoothed estimators (HRGAM, HaRGAM, BRGAM).
- Make a Comparison between GAM, WGAM, and RGAM for the smoothed estimators in points (4) and (5) through some comparison criteria (GCV, Con., BIC, AIC).

C. Simulation Results

Four examples of Tables (1,2,3,4) have been developed due to the limited space in the paper. The rest of the tables are available (ready upon request) and for the rates of contamination and samples' sizes. Different probability distributions to display and compare classical and proposed estimation methods will be discussed within the Table.

TABLE I
The Comparison Between (GAM, WGAM, RGAM) Represents the First Model, when y Contaminated with t Distribution, with Different Rates of Contamination and Different Sample Sizes

	GCV	Con.	AIC	BIC	GCV	Con.	AIC	BIC	GCV	Con.	AIC	BIC
	0.5				**0.15**				**0.35**			
GAM	2666.2	2.78E-28	533.0	540.3	2290.0	1.03E-28	529.4	536.5	1767	1.74E-27	512.8	521.1
DWGAM	2516.7	2.44E-29	266.7	271.4	2240.6	3.06E-29	264.0	268.8	1686	3.32E-29	256.9	261.4
HWGAM	2569.9	7.40E-29	267.4	272.0	2283.4	9.75E-29	264.7	269.5	1740	1.00E-28	257.4	262.0
LWGAM	2527.1	1.22E-27	266.9	271.9	2242.5	2.82E-28	264.2	268.9	1696	1.01E-27	257.0	261.5
CWGAM	2544.2	2.39E-28	266.9	271.6	2284.5	1.19E-28	265.6	268.9	1760	8.09E-28	258.2	261.0
	at 5% contamination rate and n=50 ,the best method was the DWGAM method				at 15% contamination rate and n= 50 ,the best method was the DWGAM.				at 35% contamination rate and n = 50, the best method was the DWGAM, although the BIC is smaller for a CWGAM.			
GAM	2578.5	4.48E-28	1592	1605.	3498.3	5.19E-28	1600.	1612.	1751	1.26E-27	1537.	1550.

DWGAM	2456.8	9.77E-29	799.1	807.8	2531.9	8.09E-29	801.4	810.0	1619	1.04E-27	770.6	780.2
HWGAM	2411.5	1.51E-28	797.7	806.7	2530.6	8.57E-29	801.7	810.3	1729	6.32E-28	771.2	780.6
LWGAM	2442.8	4.97E-28	798.7	807.9	2530.8	8.17E-29	801.5	810.2	1721	2.33E-28	771.6	780.3
CWGAM	2435.7	1.83E-28	798.5	807.5	2530.7	1.35E-28	801.5	810.1	1716	2.85E-29	771.3	779.9
	at 5% contamination rate and n = 150, the best method was HWGAM,although BIC is smaller for CWGAM.				at 15% contamination rate and n = 150, the best smoothing method was the (DWGAM) method, although the GCV Index had a smaller value for a (HWGAM).				at a 35% contamination rate and n = 150, the best method was the DWGAM method, although the Con. index had a smaller value for a CWGAM.			
GAM	2481.0	2.51E-27	3187.	3203	2138.2	2.50E-27	3152	3169	1672	1.02E-27	3078	3095
DWGAM	2424.5	1.27E-28	1595.0	1606	2101.0	3.86E-28	1578	1589	1675	8.75E-28	1539	1550
HWGAM	2428.8	4.86E-27	1595.1	1607	2106.5	1.95E-28	1575	1587	1683	6.49E-27	1540	1552
LWGAM	2428.5	4.65E-28	1595.6	1607	2115.3	1.58E-26	1576.1	1588	1676	1.60E-26	1589	1551
CWGAM	2426.3	3.43E-28	1595.5	1607	2143.6	9.31E-28	1576.6	1588	1676	1.60E-26	1579	1553
	at 5% contamination rate and n = 300, the best smoothing method was the (DWGAM)				at 15% contamination rate and n = 300, the best smoothing method was the (DWGAM) method, although the GCV Index had a smaller value for a (HWGAM).				at a 35% contamination rate and n = 300, the best method was the (DWGAM) method			
GAM	2579.7	3.16E-28	556.84	564.4	2245.0	5.96E-28	549.6	557.0	1641	1.82E-28	532.7	540.0
HRGAM	92.407	1.68E-28	487.28	495.2	46.062	2.10E-28	455.5	462.8	29.35	4.48E-28	382.6	389.6
HaRGAM	532.38	2.39E-29	505.12	511.8	427.53	1.71E-28	492.7	499.6	208.3	3.18E-28	421.2	428.0
BRGAM	14.69	1.30E-29	401.32	399.5	8.9252	1.24E-29	421.8	456.2	4.684	3.88E-29	324.7	356.7
	at a 5% contamination rate and n = 50, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 15% contamination rate and n = 50, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 35% contamination rate and n = 50, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)			
GAM	2378.5	4.48E-27	1592.8	1605	2498.3	5.19E-28	1600	1612	1651.	1.26E-27	1537	1550
HRGAM	48.207	1.17E-28	1367.2	1379	50.535	8.64E-28	1395	1407	26.69	8.67E-28	1076	1090
HaRGAM	479.69	1.72E-28	1470.6	1481	517.74	7.52E-28	1495	1506	192.6	7.93E-28	1199	1211
BRGAM	5.7324	1.08E-29	1321.5	1312	6.1992	1.49E-29	1325	1399	3.489	1.55E-29	1043	1021
	at a 5% contamination rate and n = 150, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 15% contamination rate and n = 150, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 35% contamination rate and n = 150, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)			
GAM	2401.0	2.51E-27	3187.1	3203	2138.2	2.50E-27	3152	3169	1648.	1.10E-27	3074	3092
HRGAM	48.116	1.77E-27	2748.8	2764	42.342	4.37E-28	2629	2645	25.85	1.32E-27	2138	2161
HaRGAM	456.53	2.13E-27	2957.2	2971	411.35	3.71E-28	2868	2883	191.1	1.22E-27	2397	2414
BRGAM	5.3927	1.00E-28	2711.6	2645	4.5370	1.98E-29	2601	2621	3.060	1.44E-29	2054	2074
	at a 5% contamination rate and n = 300, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 15% contamination rate and n = 300, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 35% contamination rate and n = 300, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)			

TABLE II
REPRESENTS THE COMPARISON BETWEEN (GAM, WGAM, RGAM) OF THE FIRST MODEL AND WHEN CONTAMINATING X WITH EXP. DISTRIBUTION. CONTAMINATION RATES AND SIZES OF SAMPLES ARE DIFFERENT

		GCV	**Con.**	**AIC**	**BIC**	**GCV**	**Con.**	**AIC**	**BIC**	**GCV**	**Con.**	**AIC**	**BIC**
n	**Cont. Per.**	**0.5**				**0.15**				**0.35**			
50	**GAM**	2534.7	2.89E-28	534.83	541.96	2879.1	9.00E-27	535.3	542.91	2912.0	3.93E-27	534.0	541.3
	DWGAM	2286.6	3.92E-29	269.50	273.39	2722.1	7.13E-29	269.8	274.23	2575.6	1.29E-27	267.6	272.3
	HWGAM	2239.6	3.35E-28	268.43	273.02	2707.8	7.95E-27	268.9	273.87	2405.8	1.03E-28	266.8	271.6
	LWGAM	2290.7	8.48E-29	269.00	273.87	2774.7	1.43E-28	269.9	274.33	2556.7	2.79E-28	267.1	271.8
	CWGAM	2241.2	1.98E-29	269.03	273.55	2752.9	6.72E-28	269.9	274.84	2563.2	1.82E-28	267.2	271.8
		at 5% contamination rate and n = 50, the best smoothing method was the (HWGAM), although the Con. criterion had a smaller value for a (CWGAM).				at 15% contamination rate and n = 50, the best smoothing method was the (HWGAM), although the Con. criterion had a smaller value for a (DWGAM).				at 35% contamination rate and n = 50, the best smoothing method was the (HWGAM).			
150	**GAM**	2597.0	2.56E-27	1600.0	1612.7	2568.6	5.41E-26	1604.	1616.5	2599.3	8.23E-28	1603	1616
	DWGAM	2538.0	6.43E-28	801.66	810.44	2563.7	1.12E-28	803.1	811.3	2577.1	2.86E-27	802.9	811.8
	HWGAM	2494.8	3.61E-29	800.31	800.50	2512.3	1.03E-28	802.8	810.2	2568.6	1.51E-28	800.0	811.8
	LWGAM	2548.6	3.91E-28	801.32	811.36	2582.6	6.12E-27	803.0	811.7	2577.6	8.05E-28	803.0	811.9
	CWGAM	2539.8	1.25E-	801.12	802.88	2584.6	3.65E-27	803.0	811.6	2577.5	2.02E-	802.9	811.9

			28								28		
		at 5% contamination rate and n = 150, the best smoothing method was the (HWGAM)				at 15% contamination rate and n = 150, the best smoothing method was the (HWGAM)				at 35% contamination rate and n = 150, the best smoothing method was the (HWGAM).			
300	**GAM**	422.54	9.57E-24	9853.5	9885.3	199.04	8.17E-24	5798.7	5832	199.39	2.51E-25	7320	7357
	DWGAM	419.41	1.75E-25	9727.1	9759.0	195.03	9.76E-27	3294.7	3312.9	196.14	2.24E-26	4015	4035
	HWGAM	413.58	3.76E-25	9696.8	9728.7	174.92	4.97E-27	3209.4	3229.6	181.76	6.86E-27	3889	3912
	LWGAM	422.54	9.57E-26	9853.5	9885.3	201.21	2.04E-26	3255.7	3273.3	191.59	7.07E-26	3961	3982
	CWGAM	463.02	9.08E-24	9840.2	9872.5	202.02	3.29E-26	3230.9	3249.7	191.82	8.26E-26	3931	3952
		at 5% contamination rate and n = 300, the best smoothing method was the (HWGAM), although the Con. criterion had a smaller value for a (LWGAM).				at 15% contamination rate and n = 300, the best smoothing method was the (HWGAM)				at 35% contamination rate and n = 300, the best smoothing method was the (HWGAM).			
50	**GAM**	2611.1	1.83E-27	557.41	565.01	2497.6	3.04E-28	533.9	541.10	2491.5	1.05E-28	533.8	540.8
	HRGAM	53.217	2.65E-28	477.48	484.91	59.328	0.64048	459.1	472.20	53.07	1.11E-28	418.9	425.5
	HaRGAM	536.08	4.44E-28	511.59	518.54	515.16	4.44E-28	486.1	492.77	533.68	1.27E-28	449.6	455.9
	BRGAM	11.432	5.06E-29	454.12	468.54	10.76	1.89E-29	398.5	402.65	18.614	7.50E-29	387.4	402.8
		at a 5% contamination rate and n = 50, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 15% contamination rate and n = 50, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 35% contamination rate and n = 50, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM).			
150	**GAM**	2497	2.56E-27	1600.0	1612.7	2499.5	1.66E-26	1600	1613.1	2494.3	5.05E-27	1599	1612
	HRGAM	50.811	2.35E-27	1395.1	1407.1	50.982	1.01E-27	1383	1395.6	61.895	1.77E-27	1302	1315
	HaRGAM	523.21	1.83E-27	1494.9	1506.4	522.48	7.24E-28	1482	1493.5	538.04	1.56E-27	1378	1389
	BRGAM	5.939	1.53E-29	1359.4	1364.5	7.688	4.31E-29	1356	1376.8	7.899	1.35E-28	1287	1296
		at a 5% contamination rate and n = 150, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 15% contamination rate and n = 150, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 35% contamination rate and n = 150, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)			
300	**GAM**	2554.1	1.43E-26	3205.8	3223	2553.1	1.57E-27	3205	3223.4	2553.7	9.08E-27	3205	3223
	HRGAM	51.216	1.28E-27	2814.8	2832	51.27	4.48E-28	2796	2813.9	51.029	2.29E-27	2554	2570
	HaRGAM	536.43	1.07E-27	3013.3	3028.6	538.57	4.91E-28	2994	3010.9	541.06	3.07E-27	274	2756
	BRGAM	5.412	1.79E-28	2794.4	2805.4	5.579	1.79E-28	2711	2687.4	5.414	7.02E-28	2516	2572
		at a 5% contamination rate and n = 300, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 15% contamination rate and n = 300, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 35% contamination rate and n = 300, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)			

TABLE III

THE COMPARISON BETWEEN (GAM, WGAM, RGAM) REPRESENTS THE SECOND MODEL, WHEN Y CONTAMINATED WITH T DISTRIBUTION, WITH THE DIFFERENT RATES OF CONTAMINATION AND DIFFERENT SAMPLE SIZES

	GCV	**Con.**	**AIC**	**BIC**	**GCV**	**Con.**	**AIC**	**BIC**	**GCV**	**Con.**	**AIC**	**BIC**
	0.5				**0.15**				**0.35**			
GAM	2666.2	2.78E-28	533.0	540.3	2278	7.34E-27	529.2	536.1	1757	1.92E-27	512.5	520.48
DWGAM	2516.7	2.44E-29	266.1	271.0	2222.9	3.02E-29	263.8	268.1	1660	1.22E-28	256.4	261.03
HWGAM	2569.9	7.40E-29	267.4	272.0	2278.0	9.00E-29	264.7	269.4	1716	9.05E-28	257.2	261.74
LWGAM	2527.1	1.22E-28	266.9	271.4	2227.7	3.00E-28	264.1	268.5	1666	1.05E-27	256.5	261.07
CWGAM	2544.2	2.39E-28	266.9	271.4	2283.1	1.60E-28	264.4	268.9	1680	9.35E-28	256.7	262.57
	at 5% contamination rate and n = 50, the best smoothing method was the (DWGAM).				at 15% contamination rate and n = 50, the best smoothing method was the (DWGAM)				at 35% contamination rate and n = 50, the best smoothing method was the (DWGAM)			
GAM	2564.0	9.33E-28	1592	1602	2584.9	4.59E-27	1599	1610	1740	1.62E-28	1536	1547.8
DWGAM	2434.9	5.90E-29	798.4	807.1	2511.0	8.06E-28	800.8	809.4	1696	1.73E-27	770.5	779.30
HWGAM	2386.7	1.63E-29	796.9	805.9	2501.0	1.06E-28	800.3	809.1	1709	5.20E-28	770.9	779.13

LWGAM	2419.7	2.11E-28	798.0	806.9	2512.8	1.29E-28	800.9	809.7	1697	1.86E-28	770.5	779.33
CWGAM	2413.0	8.84E-28	797.8	806.7	2513.4	1.58E-28	800.9	809.7	1693	3.11E-29	770.3	779.78
	at 5% contamination rate and n = 150, the best smoothing method was the (HWGAM).				at 15% contamination rate and n = 150, the best smoothing method was the (HWGAM)				at 35% contamination rate and n = 150, the best smoothing method was the (CWGAM), although the BIC criterion had a smaller value for a (HWGAM).			
GAM	2464.0	1.33E-27	1592	1602	2128.4	1.20E-27	3151	3165	1663	1.17E-27	3076	3090.4
DWGAM	2434.9	5.90E-29	796.4	804.1	2144.9	3.24E-28	1577	1588	1656.3	9.73E-27	1537	1548.9
HWGAM	2386.7	1.63E-28	796.9	805.9	2133.6	1.96E-28	1576	1587.8	1661.3	6.10E-27	1538	1550.7
LWGAM	2419.7	2.11E-28	798.0	806.9	2131.4	8.60E-27	1576	1587.9	1656.1	1.68E-27	1537	1549.5
CWGAM	2413.0	8.84E-28	797.8	806.7	2128.2	1.15E-28	1575	1587.4	1656.0	4.44E-28	1537	1549.2
	at 5% contamination rate and n = 300, the best smoothing method was the (DWGAM), although the GCV criterion had a smaller value for a (HWGAM).				at 15% contamination rate and n = 300, the best smoothing method was the (CWGAM)				at 35% contamination rate and n = 300, the best smoothing method was the (CWGAM)			
GAM	2512.9	2.31E-28	534.1	541.0	2233.2	3.56E-28	528.1	535.2	1705	3.22E-28	514.4	521.59
HRGAM	51.598	9.65E-29	450.4	457.4	44.227	1.31E-28	427.1	433.7	28.38	2.51E-29	360.3	366.64
HaRGAM	494.26	3.02E-29	482.4	489.0	397.34	7.21E-29	461.9	468.5	199.6	3.08E-29	396.8	403.07
BRGAM	10.772	9.08E-28	440.1	442.3	9.303	3.07E-28	402.8	423.4	4.366	3.55E-28	337.1	346.47
	at a 5% contamination rate and n = 50, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM), although the Con. The index had a smaller value for a (HRGAM).				at a 15% contamination rate and n = 50, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM), although the Con. Index had a smaller value for a (HRGAM).				at a 35% contamination rate and n = 50, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM), although the Con. Index had a smaller value for a (HRGAM).			
GAM	2364.0	1.33E-28	1592	1602	2124.4	9.55E-28	1575	1586	1625.	9.48E-29	1535.	1546.4
HRGAM	47.657	3.78E-29	1362	1372	41.624	4.30E-28	1288	1299	24.17	2.34E-29	1027.	1037.2
HaRGAM	380.49	2.85E-29	1459	1470	390.88	2.63E-28	1402	1412	170.3	3.19E-29	1152	1162.6
BRGAM	5.376	1.87E-29	1314	1337	4.146	7.65E-29	1182	1205	2.730	2.14E-30	1004	1014.8
	at a 5% contamination rate and n = 150, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 15% contamination rate and n = 150, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 35% contamination rate and n = 150, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)			
GAM	2388.9	3.09E-26	3185	3199	2128.2	2.77E-27	3150	3164	1618.4	1.23E-27	3068	3081.9
HRGAM	47.627	3.66E-27	2740	2754	40.655	1.35E-28	2569	2582	24.081	9.95E-30	2056	2068.2
HaRGAM	481.66	2.89E-27	2949	2963	384.37	1.07E-28	2805	2817	176.96	3.24E-30	2321	2332.7
BRGAM	4.854	3.91E-29	2601	4568	3.906	4.29E-29	2465	2487	2.348	1.08E-30	1987	1998.4
	at a 5% contamination rate and n = 300, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 15% contamination rate and n = 300, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 35% contamination rate and n = 300, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)			

TABLE IV

REPRESENTS THE COMPARISON BETWEEN (GAM, WGAM, RGAM) OF THE SECOND MODEL AND WHEN CONTAMINATING X WITH EXP. DISTRIBUTION. CONTAMINATION RATES AND SIZES OF SAMPLES ARE DIFFERENT

	GCV	**Con.**	**AIC**	**BIC**	**GCV**	**Con.**	**AIC**	**BIC**	**GCV**	**Con.**	**AIC**	**BIC**
	0.5				**0.15**				**0.35**			
GAM	2862.0	6.94E-26	535.1	542.1	2793.2	3.53E-28	533.8	540.6	2690	3.57E-27	536.0	542.8
DWGAM	2714.1	1.38E-28	269.0	273.51	2669.7	2.46E-29	268.6	273.0	2636	5.73E-28	268.4	272.9
HWGAM	2668.2	5.79E-28	268.5	273.50	2721.97	3.48E-28 2721.97	268.3	272.9	3246E-28	5.97E-28	268.5	273.0
LWGAM	2725.9	2.90E-28	269.0	273.52	2680.1	6.89E-29	268.7	272.9	2657	6.12E-28	268.5	273.9
CWGAM	2725.9	3.28E-27	269.0	273.56	2672.6	2.50E-29	268.8	272.5	2637	5.94E-28	268.8	273.8
	at 5% contamination rate and n = 50, the best smoothing method was the (HWGAM), although the Con. Index had a smaller value for a (DWGAM).				at 15% contamination rate and n = 50, the best smoothing method was the (DWGAM) for Gam and Con. and LWGAM for AIC Criterion and is smaller value for a (DWGAM) in BIC Criterion				at 35% contamination rate and n = 50, the best smoothing method was the (HWGAM)			
GAM	2591	1.75E-28	1599	1609.	2654.3	3.75E-28	1603.	1614	2631	7.71E-26	1602.	1613
DWGAM	2499	8.99E-28	800.4	809.3	2558.2	1.21E-29	802.3	810.6	2545	1.92E-28	802.0	810.4
HWGAM	2463	4.13E-29	799.3	808.3	2595.2	1.55E-28	803.3	811.8	2556	1.27E-27	802.4	810.5
LWGAM	2499	6.38E-29	800.2	808.9	2561.8	4.04E-29	802.6	810.9	2551	9.90E-28	802.2	810.8
CWGAM	2488	1.11E-28	800.0	808.7	2565.8	1.72E-28	802.5	810.9	2551	1.83E-28	802.2	810.8
	at 5% contamination rate and n = 150, the best smoothing method was the (HWGAM).				at 15% contamination rate and n = 150, the best smoothing method was the (DWGAM)				at 35% contamination rate and n = 150, the best smoothing method was the (DWGAM)			
	2637	3.39E-27	3203	3217.4	2518	6.50E-27	3201	3215	2583	3.43E-27	3197	3211
	2567	6.37E-28	1603	1615.0	2491	1.31E-28	1599.0	1610	2477	2.59E-28	1598.2	1609.0
	2546	1.14E-28	1602	1614.1	2495	1.83E-28	1599.9	1613	2479	1.47E-27	1598.9	1610
	2562	2.03E-27	1603	1614.4	2496	4.59E-28	1599.4	1612	2609.8	3.94E-28 1598.677	1598.4	1609.7 3.94E-28
	2556	1.17E-27	1603	1614.5	2495	2.00E-28	1599.5	1612	2479	8.25E-28	1598.8	1609.2

	at 5% contamination rate and n = 300, the best smoothing method was the (HWGAM).				at 15% contamination rate and n = 300, the best smoothing method was the (DWGAM)				at 35% contamination rate and n = 300, the best smoothing method was the (HWGAM)			
GAM	2510	1.50E-28	534.1	540.9	2493.2	3.53E-28	533.8	540.6	2584	3.58E-28	557.0	564.1
HRGAM	53.5	1.12E-28	462.0	468.6	52.607	2.23E-29	455.6	462.2	53.73	1.97E-27	439.6	446.2
HaRGAM	539	7.04E-29	494.5	500.9	516.43	2.12E-29	487.0	493.4	553.0	8.94E-28	471.5	477.7
BRGAM	9.422	1.36E-28	227.1	235.7	13.914	1.04E-29	228.7	249.4	17.11	1.95E-28	405.0	424.8
	at a 5% contamination rate and n = 50, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 15% contamination rate and n = 50, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 35% contamination rate and n = 50, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)			
GAM	2481	1.75E-26	1599	1609	2483.9	3.93E-26	1599	1610	2478	1.48E-26	1598	1609
HRGAM	50.37	2.63E-27	1395	1405	50.388	3.50E-27	1381	1392	50.5	7.02E-27	1284	1294
HaRGAM	519.4	1.32E-27	1495	1505	518.99	1.64E-27	1480	1490	527.9	4.55E-27	1378	1388
BRGAM	5.302	2.73E-28	1063	1278	6.963	1.13E-27	1173	1333	5.775	4.27E-27	1197	1254
	at a 5% contamination rate and n = 150, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 15% contamination rate and n = 150, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)				at a 35% contamination rate and n = 150, the best method was the smoothing method with weighted function of Robust M estimator (BRGAM)			
GAM	2537	7.07E-28	3203	3217.1	2534	6.19E-28	3203.5	3216	2536	9.17E-29	3203	3216
HRGAM	50.76	7.50E-29	2813	2825.9	50.78	2.52E-28	2794.8	2807	50.59	8.58E-29	2548	2561
HaRGAM	532.5	4.43E-29	3012	3024.5	534.2	2.30E-28	2993.5	3005	537.2	1.20E-28	2735	2747
BRGAM	5.168	5.47E-31	2467	2578.1	5.268	1.93E-30	2367.4	2413	5.269	8.17E-32	2341	2484
	at a 5% contamination rate and n = 300, the best method was the smoothing method with the weighted function of Robust M estimator (BRGAM)				at a 15% contamination rate and n = 300, the best method was the smoothing method with the weighted function of Robust M estimator (BRGAM)				at a 35% contamination rate and n = 300, the best method was the smoothing method with the weighted function of Robust M estimator (BRGAM)			

As a result of simulation experiments through tables (1,2,3,4) for the results of non-parametric analysis when X and y are contaminated and at sample sizes (50, 150, 300) and contamination rates (5%, 15%, 35%), the two proposed methods WGAM. And the proposed RGAM outperforms the ordinary GAM method by obvious decreasing the values of the comparison criteria (Concurvity, BIC, AIC GCV,), and the proposed robust M method (RGAM) showed an advantage over the proposed WGAM method by lowering the values of the criteria at Weight function (BRGAM) in all sample sizes and contamination ratios. Table (5) and as an overall result of simulation experiments, and for cases whose tables did not appear, we notice that the Bisequar (BRGAM) weighting method has had a better performance than the rest of the methods for the simulated scenarios that were addressed.

We notice from the Final Table of 216 different trials for studied simulation scenarios, for three other contamination distributions and three sample sizes (50,150,300) and different contamination scenarios, that 50% of the trials recommended the BRGAM method as the most efficient method in the first and second simulation functions. The rest of the HWGAM and DWGAM have ratios more considerable than 27% and 22%, respectively. The second function was 22% and 16%, respectively. Therefore, the estimation of the Generalized Additive Model with robust methods was superior to all other methods.

TABLE V
REPRESENTS THE SUMMARY OF SIMULATION EXPERIMENTS FOR WAVELET AND FORTIFIED METHODS FOR 216 SIMULATION ATTEMPTS

Y (t dist.)	**Wavelet Method**	**50**	DWGAM	DWGAM	DWGAM	**DWGAM**	Probability Distributions	**Wavelet Method**	**50**	DWGAM	DWGAM	DWGAM	**DWGAM**
		150	HWGAM	DWGAM	DWGAM	**DWGAM**			**150**	HWGAM	HWGAM	CWGAM	**HWGAM**
		300	DWGAM	HWGAM	DWGAM	**DWGAM**			**300**	DWGAM	CWGAM	CWGAM	**CWGAM**
	Robust Method	**50**	BRGAM	BRGAM	BRGAM	**BRGAM**		**Robust Method**	**50**	BRGAM	BRGAM	BRGAM	**BRGAM**
		150	BRGAM	BRGAM	BRGAM	**BRGAM**			**150**	BRGAM	BRGAM	BRGAM	**BRGAM**
		300	BRGAM	BRGAM	BRGAM	**BRGAM**			**300**	BRGAM	BRGAM	BRGAM	**BRGAM**
Y (Exp. Dist.)	**Wavelet Method**	**50**	CWGAM	HWGAM	HWGAM	**HWGAM**		**Wavelet Method**	**50**	DWGAM	HWGAM	DWGAM	**DWGAM**
		150	DWGAM	DWGAM	DWGAM	**DWGAM**			**150**	CWGAM	HWGAM	HWGAM	**HWGAM**
		300	HWGAM	DWGAM	DWGAM	**DWGAM**			**300**	HWGAM	DWGAM	HWGAM	**HWGAM**
	Robust Method	**50**	BRGAM	BRGAM	BRGAM	**BRGAM**		**Robust Method**	**50**	BRGAM	BRGAM	BRGAM	**BRGAM**
		150	BRGAM	BRGAM	BRGAM	**BRGAM**			**150**	BRGAM	BRGAM	BRGAM	**BRGAM**
		300	BRGAM	BRGAM	BRGAM	**BRGAM**			**300**	BRGAM	BRGAM	BRGAM	**BRGAM**
Y (Laplace dist.)	**Wavelet Method**	**50**	DWGAM	DWGAM	DWGAM	**DWGAM**		**Wavelet Method**	**50**	HWGAM	HWGAM	DWGAM	**HWGAM**
		150	CWGAM	DWGAM	DWGAM	**DWGAM**			**150**	DWGAM	HWGAM	HWGAM	**HWGAM**
		300	DWGAM	DWGAM	DWGAM	**DWGAM**			**300**	DWGAM	HWGAM	HWGAM	**HWGAM**
	Robust Method	**50**	BRGAM	BRGAM	BRGAM	**BRGAM**		**Robust Method**	**50**	BRGAM	BRGAM	BRGAM	**BRGAM**
		150	BRGAM	BRGAM	BRGAM	**BRGAM**			**150**	BRGAM	BRGAM	BRGAM	**BRGAM**
		300	BRGAM	BRGAM	BRGAM	**BRGAM**			**300**	BRGAM	BRGAM	BRGAM	**BRGAM**

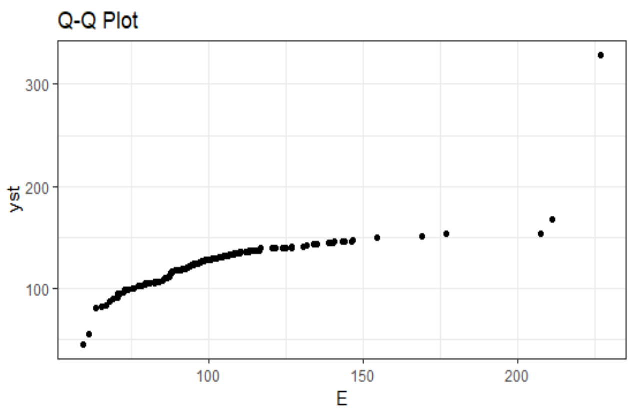
X (t dist.)	Wavelet Method	**50**	CWGAM	HWGAM	HWGAM	**HWGAM**		Wavelet Method	**50**	DWGAM	DWGAM	HWGAM	**DWGAM**
		150	HWGAM	HWGAM	HWGAM	**HWGAM**			**150**	CWGAM	DWGAM	DWGAM	**DWGAM**
		300	HWGAM	HWGAM	HWGAM	**HWGAM**			**300**	HWGAM	HWGAM	HWGAM	**HWGAM**
	Robust Method	**50**	BRGAM	BRGAM	BRGAM	**BRGAM**		Robust Method	**50**	BRGAM	BRGAM	BRGAM	**BRGAM**
		150	BRGAM	BRGAM	BRGAM	**BRGAM**			**150**	BRGAM	BRGAM	BRGAM	**BRGAM**
		300	BRGAM	BRGAM	BRGAM	**BRGAM**			**300**	BRGAM	BRGAM	BRGAM	**BRGAM**
X (Exp. Dist.)	Wavelet Method	**50**	HWGAM	HWGAM	HWGAM	**HWGAM**		Wavelet Method	**50**	HWGAM	NA	DWGAM	**NA**
		150	HWGAM	HWGAM	HWGAM	**HWGAM**			**150**	HWGAM	DWGAM	DWGAM	**DWGAM**
		300	HWGAM	HWGAM	HWGAM	**HWGAM**			**300**	HWGAM	DWGAM	DWGAM	**DWGAM**
	Robust Method	**50**	BRGAM	BRGAM	BRGAM	**BRGAM**		Robust Method	**50**	BRGAM	BRGAM	BRGAM	**BRGAM**
		150	BRGAM	BRGAM	BRGAM	**BRGAM**			**150**	BRGAM	BRGAM	BRGAM	**BRGAM**
		300	BRGAM	BRGAM	BRGAM	**BRGAM**			**300**	BRGAM	BRGAM	BRGAM	**BRGAM**
X (Laplace dist.)	Wavelet Method	**50**	HWGAM	HWGAM	HWGAM	**HWGAM**		Wavelet Method	**50**	HWGAM	HWGAM	HWGAM	**HWGAM**
		150	HWGAM	HWGAM	HWGAM	**HWGAM**			**150**	**CWGAM**	LWGAM	DWGAM	**BRGA**
		300	HWGAM	HWGAM	HWGAM	**HWGAM**			**300**	DWGAM	DWGAM	DWGAM	**DWGAM**
	Robust Method	**50**	BRGAM	BRGAM	BRGAM	**BRGAM**		Robust Method	**50**	BRGAM	BRGAM	BRGAM	**BRGAM**
		150	BRGAM	BRGAM	BRGAM	**BRGAM**			**150**	BRGAM	BRGAM	BRGAM	**BRGAM**
		300	BRGAM	BRGAM	BRGAM	**BRGAM**			**300**	BRGAM	BRGAM	BRGAM	**BRGAM**

D. Collection

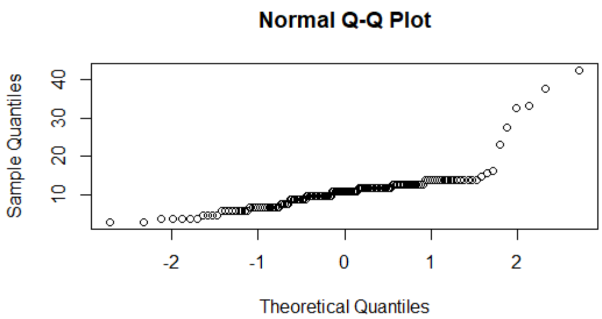
This study was applied to real data collected from Ibn Sina Teaching Hospital (Al-Wafa Specialist Center for Diabetes and Endocrinology Consultant of Short Stature) for Nineveh Governorate, 2019. On the cases with short stature, this research data was collected for 150 people with this disease. It is a very suitable sample for a model with nine explanatory variables (most research brings together that the appropriate sample size for estimating the regression models is to be ten times the number of explanatory variables at least). One response variable (height) was after reviewing a group of specialist doctors who were consulted. They demonstrated that they are the main factors that affect the incidence of this disease.

E. Normality Test

The normal distribution was tested using the normal probability plot (Q-Q plot), the response variable, and one of the explanatory variables tested (the rest were the same). Figure 1 shows that our data are not distributed as normal.



(a) Probability graph of response y



(b) Probability plot of one of the explanatory variables (x1)

Fig. 1 Q-Q plot illustrates the scheme

F. Outliers Detection

In this step, the extreme values are to be detected, where the box plot was used to detect the extreme values in the response variable. On the other hand, a Cook distance method was used to detect the explanatory variables' extreme values. As we note in figures (2), (3) there are extreme values (some of them are outliers) in response and explanatory variables, respectively.

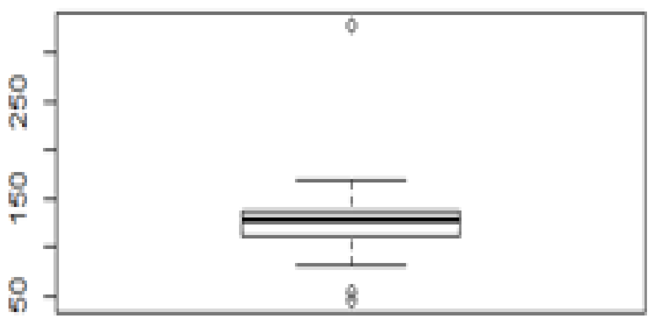


Fig. 2 shows the Box Plot

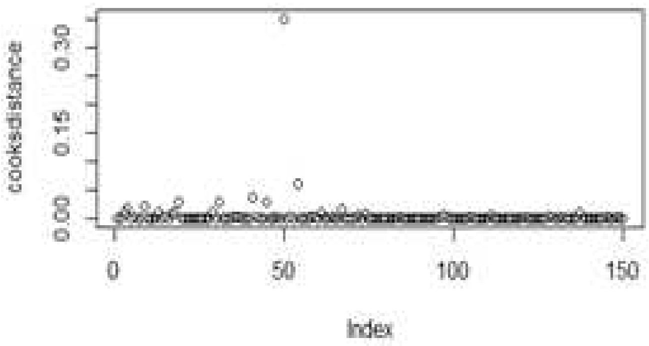


Fig. 3 shows Cook's Distance

The data is prepared by relying on three methods. Firstly, estimate the Generalized Additive Model GAM based on the smoothing splines. Secondly, to filter the data using the wavelet shrinkage method and estimating the proposed Weighted Generalized Additive Model estimation of WGAM, based on four types of the most common wavelet functions (Daubechies), Haar, Least Asymmetric, Coiflets) using smooth thresholds. Thirdly, to estimate the Generalized Additive Model based on the proposed robust M estimator RGAM and three weights of the hippocampus M amount (Huber, Hampel, Bisquare), as shown in Table (7).

TABLE VI
SHOWS THE RESULTS OF ESTIMATING THE GAM USING WAVELET FUNCTIONS

Comparative		**GCV**	**Concavity**	**AIC**	**BIC**
GAM		0.8318	0.026306	399.9243	416.9655
Wavelet function (WGAM)	**DWGAM**	0.9076	2.33E-05	207.4834	215.7273
	HWGAM	1.0710	0.00119	219.8305	230.0873
	LWGAM	0.975	1.29E-26	212.9603	219.9128
	CWGAM	0.7772	1.18E-08	204.9526	212.2963
Robust function (RGAM)	**HRGAM**	0.367951	0.003224	210.4578	223.7547
	HaRGAM	0.441533	0.004095	229.9464	242.903
	BRGAM	0.287627	0.004178	200.015	209.447

From observing the results in Table (7) and using the real data, the proposed WGAM and RGAM methods recorded a clear superiority over the ordinary GAM method through a clear decrease in the comparison criteria' values (Concurvity, BIC, AIC GCV,). The estimated RGAM showed progress on WGAM through the decline in the comparison criteria' values (BIC, AIC, GCV) at the Bisequare weighting function (RBGAM) to get. On the other hand, the wavelet shrinkage technique (WGAM) recorded a decrease in the non-linear multicollinearity index (Concurvity) at the LWGAM filter (wavelet). The GCV criterion is considered one of the most prominent comparison criteria for the Generalized Additive Model (GAM) that works to choose the smoothing parameter's value.

IV. CONCLUSION

The use of GAM model based on smoothing splines represents a very flexible method of the data problem. It does not need a preliminary determination of the form of the relationship between the explanatory and response variables. In using the simulation method, when data is contaminated with distributions ((t) Dis., Exp. Dis.) And with contamination rates (5%, 15%, 35%) and with sample sizes (50,150,300) it is noted that the smoothing method is with the Bisequare weight (BRGAM). It had a better performance compared to the rest of the methods for the simulated scenarios covered. The GCV criterion showed a marked advantage over other criteria, especially when estimating the model in the proposed robust M (RGAM) model.

It has a better performance compared to other methods of simulation scenarios that have been addressed. The GCV criterion showed a marked advantage over other criteria, especially when estimating the model using the proposed robust M (RGAM) model. When estimating the generalized additive model according to the proposed wavelet shrinkage GAM method (WGAM) and robust M (RGAM) method using the real data, it was noted that the two methods performed better than the usual GAM method. It works through a clear decrease in the comparison criteria' values (Concurvity, BIC, AIC, GCV) as the proposed robust GAM using M-estimator (RGAM) progress on the WGAM wavelet functions. It leads to a decrease in the values of the two comparison criteria at the Bisequare weight function (BRGAM).

On the other hand, the wavelet recorded a decrease in the comparison criteria (BIC, AIC). The two methods helped to smooth the data from the extreme values. This is done by obtaining the smallest values for the comparison criteria. It was noted that the GCV criterion decreases with the increase of the sample size in general, as the GCV is the efficiency criterion for the GAM model, which is responsible for choosing the best smoothing parameter. Accordingly, the GCV criterion is considered as the most crucial efficiency criterion used in the research, and accordingly, the proposed robust method can be considered better than the proposed wavelet method. As for the AIC and BIC standard, there has been an increase with an increase in the sample size, and we find that the non-linear multicollinearity index (Concurvity) fluctuates up and down. Its results are close in all ways, so it is less sensitive to outliers than other criteria.

REFERENCES

[1] V. B. Khyade and M. Eigen, "Key Role of Statistics for the Fortification of Concepts in Agricultural Studies," *Int. Acad. J. Psychol. Educ. Stud.*, vol. 6, no. 1, pp. 20–34, 2019.

[2] H. K. Sharaf, M. R. Ishak, S. M. Sapuan, N. Yidris, and A. Fattahi, "Experimental and numerical investigation of the mechanical behavior of full-scale wooden cross arm in the transmission towers in terms of load-deflection test," *J. Mater. Res. Technol.*, vol. 9, no. 4, pp. 7937–7946, 2020.

[3] H. K. Sharaf, M. R. Ishak, S. M. Sapuan, and N. Yidris, "Conceptual design of the cross-arm for the application in the transmission towers by using TRIZ–morphological chart–ANP methods," *J. Mater. Res. Technol.*, vol. 9, no. 4, pp. 9182–9188, 2020.

[4] A. N. Johari *et al.*, "Fabrication and cut-in speed enhancement of savonius vertical axis wind turbine (SVAWT) with hinged blade using fiberglass composites," in *Proceedings of the Seminar Enau Kebangsaan*, 2019, pp. 978–983.

[5] M. R. M. Asyraf *et al.*, "Creep test rig for full-scale composite cr ossarm: simulation modelling and analysis," in *Seminar Enau Kebangsaan*, 2019, pp. 34–38.

[6] A. Buniello *et al.*, "The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D1005–D1012, 2019.

[7] Y. Zhang, G. Qi, J.-H. Park, and N. Chatterjee, "Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits," *Nat. Genet.*, vol. 50, no. 9, pp. 1318–1326, 2018.

[8] O. Weissbrod, J. Flint, and S. Rosset, "Estimating SNP-based heritability and genetic correlation in case-control studies directly and with summary statistics," *Am. J. Hum. Genet.*, vol. 103, no. 1, pp. 89–99, 2018.

[9] R. Andersen, *Modern methods for robust regression*, no. 152. Sage, 2008.

[10] G. Li, "Robust regression," *Explor. data tables, trends, shapes*, vol. 281, p. U340, 1985.

[11] J. Fox and S. Weisberg, "Robust regression," *An R S-Plus companion to Appl. Regres.*, vol. 91, 2002.

[12] X. Zhou, "A unified framework for variance component estimation with summary statistics in genome-wide association studies," *Ann. Appl. Stat.*, vol. 11, no. 4, pp. 2027–2051, Dec. 2017, doi: 10.1214/17-AOAS1052.

[13] A. Cichonska *et al.*, "metaCCA: summary statistics-based multivariate meta-analysis of genome-wide association studies using canonical correlation analysis," *Bioinformatics*, vol. 32, no. 13, pp. 1981–1989, Jul. 2016, doi: 10.1093/bioinformatics/btw052.

[14] M. R. Mahmoudi and A. Abbasalizadeh, "How statistics and text mining can be applied to literary studies?," *Digit. Scholarsh. Humanit.*, vol. 34, no. 3, pp. 536–541, Sep. 2019, doi: 10.1093/llc/fqy069.

[15] M. Shelke, S. S. Deshpande, and S. Sharma, "Quinquennial Review of Progress in Degradation Studies and Impurity Profiling: An Instrumental Perspective Statistics," *Crit. Rev. Anal. Chem.*, vol. 50, no. 3, pp. 226–253, May 2020, doi: 10.1080/10408347.2019.1615863.

[16] A. Worster and T. Haines, "Advanced statistics: Understanding Medical Record Review (MRR) Studies," *Acad. Emerg. Med.*, vol. 11, no. 2, pp. 187–192, Feb. 2004, doi: 10.1111/j.1553-2712.2004.tb01433.x.

[17] Y.-G. Wang, X. Lin, M. Zhu, and Z. Bai, "Robust Estimation Using the Huber Function With a Data-Dependent Tuning Constant," *J. Comput. Graph. Stat.*, vol. 16, no. 2, pp. 468–481, Jun. 2007, doi: 10.1198/106186007X180156.

[18] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust statistics: the approach based on influence functions*, vol. 196. John Wiley & Sons, 2011.

[19] J. Ji, "Robust inversion using Biweight norm," in *SEG Technical Program Expanded Abstracts 2011*, Jan. 2011, pp. 2717–2721, doi: 10.1190/1.3627758.

[20] S. Jaffard, Y. Meyer, and R. D. Ryan, *Wavelets: tools for science and technology*. SIAM, 2001.

[21] R. C. Guido, "A note on a practical relationship between filter coefficients and scaling and wavelet functions of Discrete Wavelet Transforms," *Appl. Math. Lett.*, vol. 24, no. 7, pp. 1257–1259, Jul. 2011, doi: 10.1016/j.aml.2011.02.018.

[22] J. Rafiee, M. A. Rafiee, N. Prause, and M. P. Schoen, "Wavelet basis functions in biomedical signal processing," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 6190–6201, May 2011, doi: 10.1016/j.eswa.2010.11.050.

[23] R. S. Stanković and B. J. Falkowski, "The Haar wavelet transform: its status and achievements," *Comput. Electr. Eng.*, vol. 29, no. 1, pp. 25–44, Jan. 2003, doi: 10.1016/S0045-7906(01)00011-8.

[24] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, Sep. 1994, doi: 10.1093/biomet/81.3.425..