Iraqi Journal of Statistical Sciences

www.stats.mosuljournals.com

Using Wavelet Shrinkage in the Cox Proportional Hazards Regression model (simulation study)

Taha.H. Ali and **Nasradeen Haj Salih Albarwari** **Diyar Lazgeen Ramadhan**

Department of Informatics & Statistic, College of Administration and Economics, Salahaddin University, Erbil,, Iraq

Article information	Abstract
	In this paper, a hybrid method of quantile regression and multivariate wavelet is proposed to deal with the problem of data contamination or the presence of outliers, which uses the median instead of the mean on which the linear regression model and the estimation method for ordinary least squares depend. The paper included a comparison between the proposed (for several wavelets and different threshold) and classical method based on mean absolute error, to get the best fit quantile regression model for the data. The application part dealt with two types of data representing simulation, real data, and analysis using a program designed for this purpose in the MATLAB language, as well as the statistical program SPSS-26 and EasyFit-5.5. The study concluded that the proposed method is more efficient than the classical method in estimating the parameters of a quantile regression model depending on the coefficient of determination and on the mean absolute error and mean squared error criteria.

1. Introduction

Quantile regression offers the opportunity for a more complete view of the statistical landscape and the relationships among stochastic variables. The simple expedient of replacing the familiar notions of sorting and ranking observations in the most elementary one-sample context by optimization enables us to extend these ideas to a much broader class of statistical models. Just as minimizing sums of squares permits us to estimate a wide variety of models for conditional mean functions, minimizing a simple asymmetric version of absolute errors yields estimates for conditional quantile functions. For linear parametric models, computation is greatly facilitated by the reformulation of the optimization problem as a parametric linear program. Formal duality results for linear programs yield a new approach to rank statistics and rank based inference for linear models (Koenker R. 2005).

Wavelet shrinkage is a non-parametric technique used in curve estimation. The idea is to shrink wavelet coefficients towards zero using statistical methods. More specifically, a threshold value is chosen and wavelet coefficients whose absolute values exceed that threshold are kept while others are removed. This has the effect of both reducing the noise contribution and compressing the original data while keeping a

good quality of approximation. A key step in the wavelet shrinkage procedure is the choice of the threshold (Raimondo, M. 2002).

The aim of our study is estimation of a more appropriate regression model from the classic methods for constructing a regression model by using the proposed hybrid method for multivariate wavelet shrinkage and quantile regression model, and data de-noise using wavelet, and comparison between the classical and the proposed method (multivariate wavelets quantile regression) in estimating regression model.

2. Theoretical Part (Quantile Regression Model & Wavelet Shrinkage)

The theoretical part is the most important part of academic research because it will be the viewfinder through which researchers evaluate their study problems. This part provides some information regarding statistical techniques including quantile regression models, and wavelet shrinkage. So, this part describes theories and models for classical and proposed methods that will be applied in the practical part of the paper.

2.1. Quantile Regression

Given the sample of size n, $\{(X_i, y_i), i = 1, \ldots, n\}$, with random error ε, from the following linear quantile model:

$$Qy\,(\tau) = \ X^T\,\beta_\tau + \varepsilon \tag{1}$$

The τ-specific regression parameter $\boldsymbol{\beta}_\tau$ is interpreted as the 'quantile treatment effect' of X on **y**. An estimate is obtained by solving the L_1-norm regression problem (Koenker and Bassett, 1978) is

$$\min_\beta \sum_{i=1}^n \rho_\tau(y_i - X_i^T\beta) \tag{2}$$

where $\rho_\tau(r) = r(\tau - I(r < 0))$ is the quantile loss function and *I* denoted the indicator function.

May be reformulated as a linear program by introducing 2n artificial, or "slack," variables $\{u_i, v_i : 1, \ldots, n\}$ to represent the positive and negative parts of the vector of residuals. This yields the new problem

$$\min_{u,v} = \{\tau 1_n^T.u + (1-\tau).1_n^T v \mid 1_n\,Qy(\tau)+u-v = y\} \tag{3}$$

where 1n denotes an n-vector of 1. Clearly, in (3) we are minimizing a linear function on a polyhedral constraint set consisting of the intersection of the (2n + 1)-dimensional hyperplane determined by the linear equality constraints (Bickel and Freedman, 1981). Quantile regression problem (2) may be reformulated as a linear program as in (3):

$$\min_{\beta,u,v} = \{\tau 1_n^T.u + (1-\tau).1_n^T v \mid 1_n\,X^T\,\beta+u-v = y\} \tag{4}$$

Where X now denotes the usual n by p regression design matrix. Again, we have split the residual vector y − Xβ into its positive and negative parts, and so we are minimizing a linear function on a polyhedral constraint set, and most of the important properties of the solutions, $\hat{\beta}(\tau)$, which we call "regression quantiles," again follow immediately from well-known properties of solutions of linear programs.

One of the important measures for the efficiency of the estimated regression model is the mean absolute error (MAE) criterion, which is usually used in quantile regression, as in the following formula, (Willmott, C. J., & Matsuura, K. 2005):

$$MAE = \frac{\sum_{i=1}^{n}|e_i|}{n} \qquad\qquad (5)$$

And to extract a mean (MAE) for a number of samples or iterations (k), we use the following formula:

$$Mean(MAE) = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n}e_{ij}}{n*k} \qquad\qquad (6)$$

2.2 Wavelet Shrinkage

Wavelet shrinkage is a well-established technique to data de-noise, it consists of Wavelets and a Shrinkage, as will be explained later.

2.2.1 Multivariate Wavelets

Wavelets are small waves that can be grouped together to form larger waves or different waves. A few fundamental waves were used, stretched in infinitely many ways, and moved in infinitely many ways to produce a wavelet system that could make an accurate model of any wave. Consider generating an orthogonal wavelet basis for functions $f \in L^2(R)$ (the space of square integrable real functions), starting with two parent wavelets: the scaling function ϕ (also called the farther wavelet) and the mother wavelet ψ. Other wavelets are then generated by dilations and translations of ϕ and ψ (Percival et al., 2004). The dilation and translated of the functions are defined by formulas (7) and (8).

$$\phi_{j,q}(y) = 2^{\frac{j}{2}}\phi(2^j\, y - q) \qquad j, q \in z \qquad\qquad (7)$$

$$\psi_{j,q}(y) = 2^{\frac{j}{2}}\psi(2^j\, y - q) \qquad j, q \in z \qquad\qquad (8)$$

Where j is the level, y is the observations, and q is the shift. Wavelet is a small wave that grows and decays in limited time period instead of being a wave that goes on forever, such as sinusoidal waves which are determined on a whole time domain. The wavelet Daubechies and Symlets used in the paper can be clarified:

Daubechies wavelet: These wavelets are named after the researcher (Ingrid Daubechies), who is considered the pioneer researcher in the subject of wavelength, and the so-called normal orthogonal wavelets with an anchor in the year (1988), which made the discrete wavelet analysis applicable. (D) and (db) are the abbreviations of the researcher's name (daubechies), (N) is the length of the candidate or rank, while (L1) is the number of vanishing or ephemeral moments of the wavelet function, for example, (4D) and (db2). (L1) indicate the same candidate who is the second-ranked candidate in this family, and (L1) is associated with (N) in the following relations (Abd Al-Kader,2011).

$$L1 = \frac{N}{2} \qquad\qquad (9)$$

In general, (dbn) represents the family of small waves of the order n (note that the small wave haar is one of the members of this family and has the following characteristics (Taha H. A., 2009):

● The anchor of the small wave (dbn) is on the period [0, 2n-1].
● A wavelet (dbn) has n ephemeral moments, i. e:

$$\left|\frac{d^j}{dy^j}\psi(y)\right| < \infty$$

$$\Rightarrow \quad \int_{-\infty}^{\infty}y^j\psi(y)dy = 0 \ , \ \ 1 \le j \le n \qquad\qquad (10)$$

From formula (10), the moment j is obtained for the mother function
● Anomalies increase with rank, (dbn) have (rn) continuous derivatives (r is about 0.2).

Symlets wavelet: In symN, N is the order. Some authors use 2N instead of N. The symlets are nearly symmetrical, orthogonal and biorthogonal wavelets proposed by Daubechies as modifications to the (db) family. The properties of the two wavelet families are similar (Chavan et al., 2011).

2.2.2 Shrinkage

Threshold usually involves the conversion of signal values below a certain threshold to zero or equal to the threshold when values are higher than the threshold. In image processing, threshold is commonly used for image segmentation. According to the type of image, different types of threshold mechanisms are used (Ahmadi et al., 2015).

Threshold is the simplest method of non-linear wavelet de-noising, in which the wavelet coefficient is divided into two sets, one of which represents the signal while the other represents noise. There are different rules for applying the thresholds of the wavelet coefficients, and several different methods for choosing a threshold value exist, such as:

I. Universal Threshold Method

(Donoho and Johnstone, 1994) submitted the universal threshold method, which is given by formula (11).

$$\delta^U = \hat{\sigma}_{MAD} \sqrt{2 \, log \, (n)} \tag{11}$$

Where $\hat{\sigma}_{MAD}$ is the standard deviation estimator of details coefficients, and equal to $\frac{MAD}{0.6745}$. Where MAD is the median absolute deviation of the wavelet coefficients at the finest scale.

II. Minimax Threshold Method

The optimal minimax threshold method, submitted by (Donoho and Johnstone, 1994) as an improvement to the universal threshold method. Minimax is based on an estimator $\tilde{f}$ that attains to the minimax risk, as:

$$\tilde{R}(F) = \inf_{\tilde{f}} \, \sup_{f \in \tilde{R}(F)} R(\tilde{f}, f) \tag{12}$$

Where

$$R(\tilde{f}, f) = \frac{1}{n} \sum_{i=1}^{n} E[\tilde{f}, f]^2 \tag{13}$$

Where $f = f(x_i)$ and $\tilde{f} = \tilde{f}(x_i)$, denote the vectors of true and estimated sample values. The threshold minimax estimator is different from universal counter parts, in which the minimax threshold method is concentrate on reducing the overall mean square error (MSE) but the estimates are not over-smoothing.

III. SURE Threshold Method

The sure threshold proposed by (Donoho and Johnstone, 1994), which based upon the minimization of stein's risk estimator. In sure threshold method specifies a threshold estimate of δ_k at each level k for the wavelet coefficients, and then for the soft threshold estimator we have.

$$SURE(\delta, W) = n - 2 \neq \{W_k \le \delta\} - \sum_{k=0}^{d} min(|W_k|, \delta) \tag{14}$$

Where $[W_k: k = 1,2, ..., d]$ be a wavelet coefficients in the kth level, and then, select δ^s that minimizes $SURE(\delta, W)$

$$\delta^s = arg \; SUER \,(\delta, W) \tag{15}$$

On the other hand, there are many rules for threshold (Threshold Rules). The two types used in this research will be discussed.

I. Soft Threshold

The other standard technique for wavelet de-noising is Soft threshold of the wavelet coefficient, also proposed by Donoho and Johnstone, which is defined as follows (Joy et al., 2013).

$$W_n^{(s)} = sign \, \{W_n\}(|W_n| - \delta) \tag{16}$$

Where

$$\text{and} \qquad Sign\{Wn\} = \begin{bmatrix} +1 & if & Wn > 0 \\ 0 & if & Wn = 0 \\ -1 & if & Wn < 0 \end{bmatrix} \tag{17}$$

$$(|W_n| - \delta) = [(|W_n| - \delta) \; if \; (|W_n| - \delta) \geq 0 \; and \; 0 \; if \; (|W_n| - \delta) < 0 \,] \tag{18}$$

Coefficients smaller than the threshold are set to zero, and additionally, all coefficients greater than the threshold are reduced by the amount of the threshold. Thus, soft threshold is a continuous mapping.

II. Hard Threshold

Donoho and Johnstone proposed Hard threshold, it is a simplest scheme threshold interpreting the statement of (keep or kill). The Hard threshold used straightforward technique for implementing wavelet de-noising (Hagiwara, 2021). The wavelet coefficient is set to the vector $Wn^{(H)}$ with element.

$$Wn^{(H)} = \begin{bmatrix} 0 & if & |Wn| \leq \delta \\ Wn & if & |Wn| > \delta \end{bmatrix} \tag{19}$$

exceeding δ are left untouched, while smaller than or equal to δ are eliminated or set to 0. Thus, the operation of hard threshold is not continuous mapping.

2.3 Detecting data noise

To display noise detection in observations (y_i) when i=0,1,..,n, researcher (Greenblatt, 1996) suggested a way to do this work by using discrete wavelet transform (DWT), specifically, $W_1 = W_{1,0}, W_{1,1}, .., W_{1,\frac{N}{2}-1}$ which represent the discrete wavelet transformation coefficients at the first level for observation . Express the test statistics using the following formula:

$$D_2 = \frac{y_i \, |W_{1,y} - W_1|}{\hat{\delta}_y(\lambda_1)} \tag{20}$$

Where (W_1) represents the mean of the sample to $(\frac{N}{2})$ of the wavelet coefficients at the first level, and $(\hat{\delta}_y(\lambda_1))$ represents the standard deviation of the wavelet coefficients of the first level. The above test statistic is used to test for the presence of noise in observations.

The noise represents a high oscillation event or phenomenon , which is captured and well detected by (W_1) because its coefficients are related to frequencies (f) during the period $(1/4 \leq f \leq 1/2)$, noting that the use of (w) in the truth is not necessary because the wavelet coefficients have a rate equal to zero (zero mean) as

long as the process spectrum does not change in the aforementioned period, so the test statistic can be approximated to the following formula:

$$D_2 = \frac{y_i |W_{1,i}|}{\hat{\delta}_y(\lambda_1)} \qquad (21)$$

The value of (D_2) determines the upper limit that is acceptable and that values greater than them are considered undesirable noise values, while (- D_2) represents the minimum and that values less than them are also considered undesirable noise values, and on this basis, will be determined by determining the values that fall outside the limits of the specified period, i.e. [- D_2 , D_2] and trying to remove their influence and obtaining the candidate observations that fall within the period (عبدالقادر، 2011).

2.4 Proposed method

we will treat the data from contamination (de-noising of data) before creating Quantile Regression Models using multivariate wavelet shrinking. In particular, Daubechies and Symlets wavelets, and Universal, Minimax, and SURE as Threshold Method, and Soft and Hard as threshold Rules. The de-noise data (for the dependent and independent variables) is used to estimate parameters of the quantile regression model and then compare its efficiency with the classical method depending on the (MAE).

3. Practical part (Simulation and Real Data)

For the purpose of comparing the proposed and classical method, two types of data were used, namely simulation and real data.

3.1 Simulation

Simulation of the first experiment using a program designed in MATLAB language (Appendix) to generate a multiple regression model with three independent variables and a sample size equal to (100), and assuming that the parameter vector is [1.5, 3, 2, 2.5] with a random error that has an identical independent $N(0, 1)$ and noise additive. Figure (1) shows the scatter plot of the regression standardized predicted Values versus regression standardized residual (using OLS method).
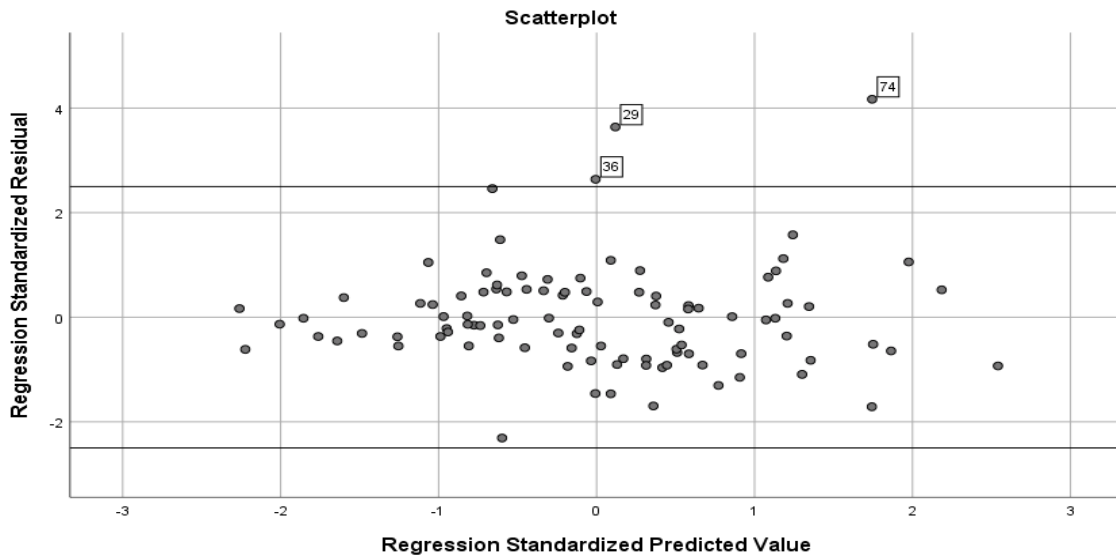

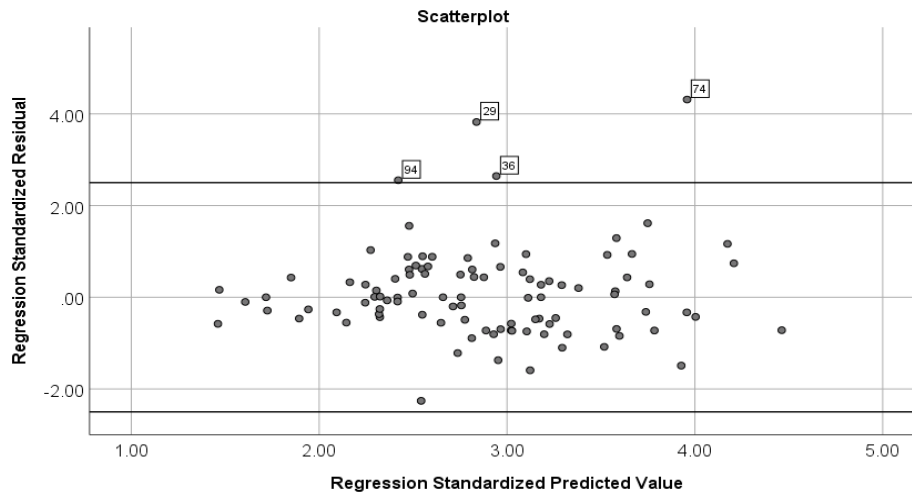Figure 1. Scatter plot for Regression Standardized Residual (OLS)

Figure (1) shows that there are three outliers (points: 29, 36, and 74) because they are greater than (2.5). The information of the linear model estimated by (OLS) is summarized in the table (1):

Table (1): Analysis of OLS Model

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	R^2	F	MSE	MAE
OLS	B	Std. Error							
(Constant)	1.776	.660		2.692	.008				
x1	3.492	.774	.381	4.513	.000	0.318	14.89	3.834	1.3684
x2	1.781	.678	.223	2.628	.010				
x3	2.794	.676	.352	4.131	.000				

Table (1) shows that the F-Statistic (14.89) supports the fit of the linear model to the data because it is greater than its tabulated value under the significant level (0.05) and degrees of freedom (3, 96) which is equal to (2.7114). All parameters are significant and the coefficient of determination is 31.8%, MSE and MAE are (3.834 and 1.3684) respectively.

Quantile regression results for the same data and when $\tau = (0.5)$ were as follow:

Figure 2. Scatter plot for Regression Standardized Residual (QR)

Figure (2) shows that there are four outliers (points: 29, 36, 74, and 94) because they are greater than (2.5), because Quantile regression does not treat outliers, but is robust to them when estimating the parameters of the linear model. The information of the linear model estimated by (QR) is summarized in the table (2):

Table (2): Analysis of QR Model

Model	B	Std. Error	t	Sig.	Pseudo R^2	MSE	MAE
OLS							
(Constant)	1.751	.7071	2.476	.015			
x1	3.455	.8296	4.164	.000	0.205	3.898	1.3557
x2	1.922	.7264	2.645	.010			
x3	2.310	.7250	3.186	.002			

Table (2) shows that all parameters are significant and the pseudo R^2 is 20.5%, MSE and MAE are (3.898 and 1.3557) respectively. The Quantile regression model is more efficient than (OLS) linear regression because the value of (MAE = 1.3557) is less than (MAE = 1.3684), it is the most important in comparison

to the Quantile regression model, which depends on the median instead of the mean of (OLS) linear regression.

Proposed method (Multivariate Wavelet Quantile regression), Beginning with the use of the Multivariate Wavelet (Sym2), Minimax method, and Soft rule (threshold). We get the following results (using a program prepared for this purpose in MATLAB language in the appendix).
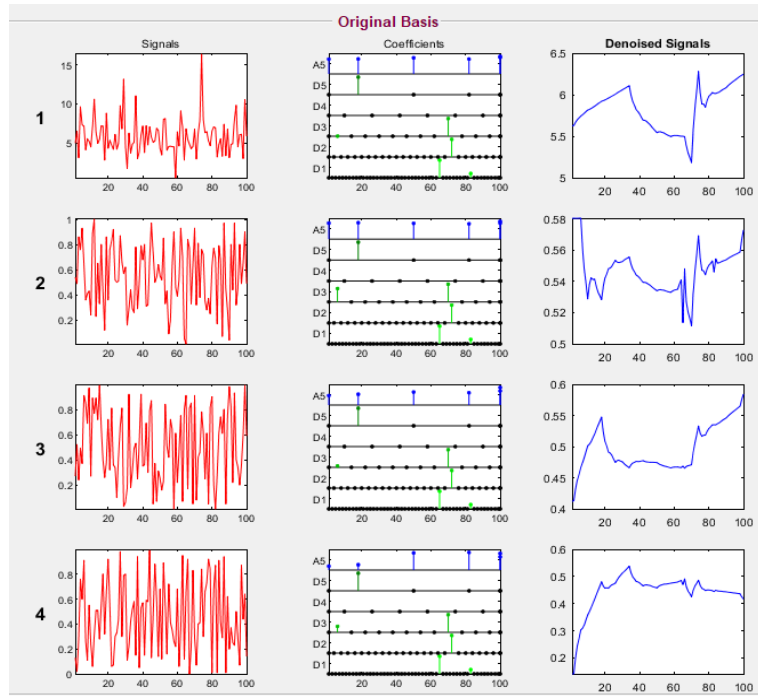
Figure 3. Original data with de-noised data for all variable

Figure (3) shows the original data (red color) and the data processed using multivariate wavelet (blue color), which confirms the presence of noise in the data that has been processed or de-noise. Multivariate Wavelet Quantile regression results for the same data and when $\tau = (0.5)$ were as follow:
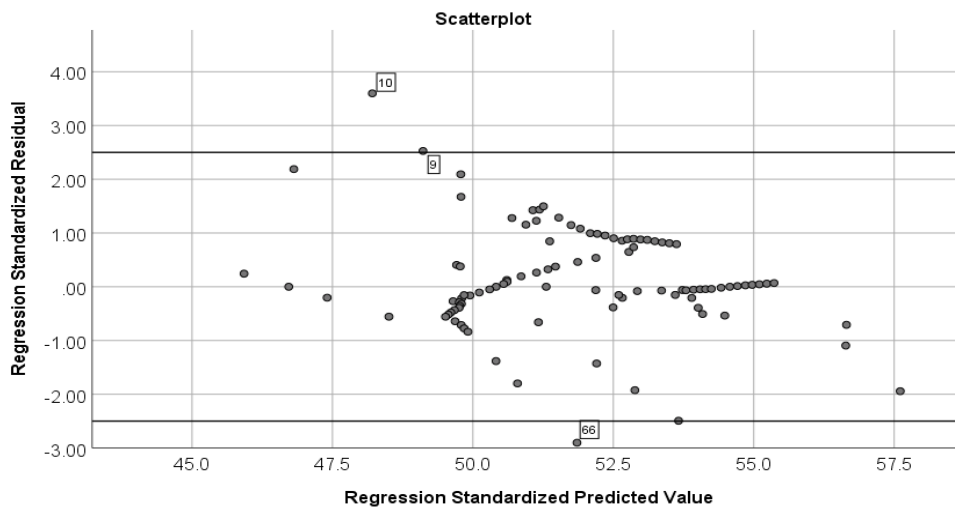
Figure 4. Scatter plot for Regression Standardized Residual (MWQR)

Figure (4) shows that there are three outliers (points: 9, 10, and 66) Because they were outside the limits of the interval (-2.5 and 2.5). The information of the linear model estimated by (MWQR) is summarized in the table (3):

Table (3): Analysis of MWQR Model

Model OLS	B	Std. Error	t	Sig.	Pseudo R^2	MSE	MAE
(Constant)	-5.093	.6433	-7.918	.000	0.631	0.0126	0.0769
x1	15.734	1.1194	14.056	.000			
x2	3.171	.3859	8.217	.000			
x3	1.659	.2286	7.256	.000			

Table (3) shows that all parameters are significant (p-value = 0.000), and It is more significant than the parameters estimated by the classical method. The pseudo R^2 is 63.1%, Which is greater than its value (20.5%) in the classical method. MSE and MAE are (0.0126 and 0.0769) respectively, they are less than their values (3.898 and 1.3557, respectively) in the classical method. Thus Multivariate Wavelet Quantile regression model is more efficient than (QR) linear model. After knowing that the multivariate wavelet quantile regression was better than the classical method in first random sample, the experiment will be repeated (1000) once and for several different sample sizes (30, 50, 100), and the results of the proposed method will be compared with the classical method, depending on the average of MAE criterion, as shown in the following table:

Table (4) Average of MAE for simulation

Sample size	Method	Wavelet	De-noising Method	MAE
30	Proposed method	Sym2	SURE	0.4242
	Proposed method	Db2	SURE	0.4272
	Proposed method	Sym2	minimax	0.3751
	Proposed method	Db2	minimax	0.3740
	Proposed method	Sym2	universal	0.2816
	Proposed method	**Db2**	**universal**	**0.2810**
	Quantile regression	-----	------	0.8523
50	Proposed method	Sym2	SURE	0.3721
	Proposed method	Db2	SURE	0.3715
	Proposed method	Sym2	minimax	0.3268
	Proposed method	Db2	minimax	0.3275
	Proposed method	**Sym2**	**universal**	**0.2608**
	Proposed method	Db2	universal	0.2614
	Quantile regression	-----	------	0.7412
100	Proposed method	Sym2	SURE	0.2479
	Proposed method	Db2	SURE	0.2480
	Proposed method	Sym2	minimax	0.1905
	Proposed method	Db2	minimax	0.1906
	Proposed method	**Sym2**	**universal**	**0.1178**
	Proposed method	Db2	universal	0.1179
	Quantile regression	-----	------	0.6711

Table (4) shows that all the proposed methods were better than the classical method because the values of MAE were less than their value in the classical method (MAE = 0.8523, 0.7412, and 0.6711 respectively) for all different sample sizes. When the sample size is equal to (30), the proposed method (Db2-universal, MAE = 0.2810) was better than the other proposed methods, while for the sample sizes (50 and 100), the proposed method (Sym2-universal, MAE = 0.2608 and 0.1178, respectively) was the best compared to the other proposed methods. For all the methods used, the values of MAE decrease when the sample size increases, so the estimated models were better.

3.2 Real Data

In this part, we will analyze real data with multiple linear regression and quantile regression and then proposed method by using some multivariate wavelets, and compare them and choose the method that gives us better results. The data in (Appendix) have been retrieved from a book located in the Directorate of Statistics in Dohuk under the name (IHSEES II Household Social and Economic Survey 2012), where the first edition was in 2014. The available copy can be obtained on the website (www.cost.gov.iq) where the census process was comprehensive in all governorates of Iraq and the sample size in each governorate was 1512 samples (family). The dependent variable represents (total household spending **y**), the independent variables represents (food and non-alcoholic beverages x_1, clothes and shoes x_2, and housing, water, gas, electricity and other fuels x_3). Figure (1) shows the scatter plot of the regression standardized predicted Values versus regression standardized residual (using OLS method).
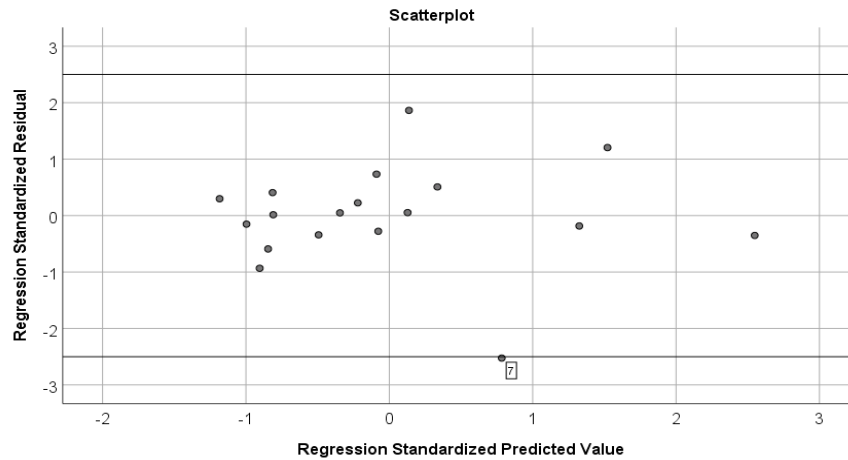
Figure 5. Scatter plot for Regression Standardized Residual (OLS) for real data

Figure (5) shows that there is one outlier (point 7) because it was less than (-2.5). The hypotheses of the linear regression model were verified and they are summarized in the following table:

Table (5) Linear model hypotheses tests

Test of Normality		Test of Homogeneity	Test of Autocorrelation		Test of Multi-collinearity		
K. S.	Chi-Squared	Levene test (Based on Median)	Durban-Watson		Model	Tolerance	VIF
0.1479	0.6937	2.649 (0.126)	1.827		constant		
(0.371)	(0.707)				X_1	.416	2.404
Critical values			d_l	1.046	X_2	.360	2.779
0.3094	5.9915		d_u	1.535	X_3	.671	1.489

Table (5) shows that to test the hypothesis of a normal distribution for random error (H_0: the random error a has normal distribution vs H_1: the random error has no normal distribution), Kolmogorov-Smirnov (K. S.) and Chi-Square tests support the null hypothesis and the random error a has normal distribution (p-values = 0.371 and 0.707 respectively, and they are greater than $\alpha = 0.05$). Test of homogeneity of random error variance (H_0: homogeneity of random error variance vs H_1: heterogeneity of random error variance), Levene test (based on median) support the null hypothesis and homogeneity of random error variance (p-value = 0.126, and it is greater than $\alpha = 0.05$). Test of autocorrelation for random error, (H_0: There is no autocorrelation problem vs H_1: There is autocorrelation problem), Durban-Watson test support the null hypothesis and there is no autocorrelation problem between random error values (D. W. = 1.827 and it is lies between (d_u = 1.535, 4-d_u = 2.465) and this means that it is lies in the acceptance region of the null

hypothesis). Test of multi-collinearity problem, (H_0: There is no multi-collinearity problem vs H_1: There is multi-collinearity problem), VIF test support the null hypothesis and there is no multi-collinearity problem between independent variables (VIF = 2.404, 2.779, and 1.489 respectively, and they are less than (3). Thus, we conclude that the estimation hypotheses of the linear regression model are available.

The estimated linear model information (OLS) is summarized in Table (6):

Table (6): Analysis of OLS Model for real data

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	R^2	F	MSE	MAE
OLS	B	Std. Error							
(Constant)	-290.4	169.45		-1.714	.109				
x1	1.406	.373	.265	3.768	.002	0.971	157.7	8267.4	54.118
x2	6.138	.752	.617	8.164	.000				
x3	1.369	.329	.230	4.159	.001				

Table (6) shows that the F-Statistic (157.7) supports the fit of the linear model to the data because it is greater than its tabulated value under the significant level (0.05) and degrees of freedom (3, 14) which is equal to (3.3439). All parameters (slops) are significant and the coefficient of determination is 97.1%, MSE and MAE are (8267.4 and 54.118) respectively.

Quantile regression results for real data, when $\tau = (0.5)$ were as follow:
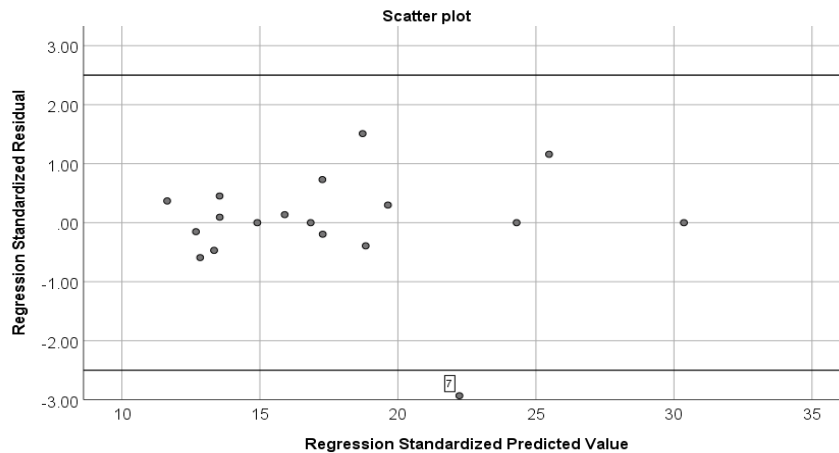
Figure 6. Scatter plot for Regression Standardized Residual (QR) for real data

Figure (6) shows that also there is one outlier (point 7) because it was less than (-2.5). The hypotheses of the linear regression model were verified and they are summarized in the following table:

Table (7): Analysis of QR Model

Model	B	Std. Error	t	Sig.	Pseudo R^2	MSE	MAE
OLS							
(Constant)	-429.304	145.1331	-2.958	.010			
x1	1.787	.3194	5.593	.000	0.863	8904.3	49.7139
x2	5.512	.6439	8.560	.000			
x3	1.378	.2820	4.889	.000			

Table (7) shows that all parameters are significant and the pseudo R^2 is 86.3%, MSE and MAE are (8904.3 and 49.7139) respectively. The Quantile regression model is more efficient than (OLS) linear regression because the value of (MAE = 49.7139) is less than (MAE = 54.118), it is the most important in comparison

to the Quantile regression model, which depends on the median instead of the mean of (OLS) linear regression.

Proposed method (Multivariate Wavelet Quantile regression), Beginning with the use of the Multivariate Wavelet (Sym2), Minimax method, and Soft rule (threshold). We get the following results (using a program prepared for this purpose in MATLAB language in the appendix).
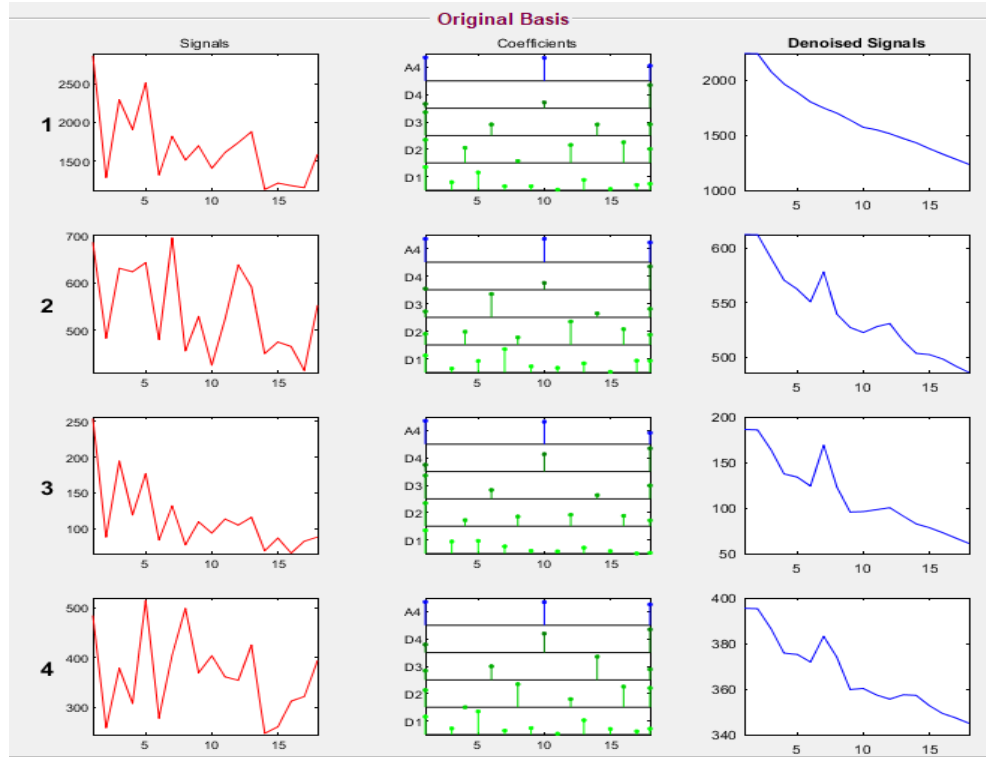
Figure 7. Original data with de-noised data for all variable (real data)

Figure (7) shows the original data (red color) and the data processed using multivariate wavelet (blue color), which confirms the presence of noise in the data that has been processed or de-noise. Multivariate Wavelet Quantile regression results for the real data and when $\tau = (0.5)$ were as follow:
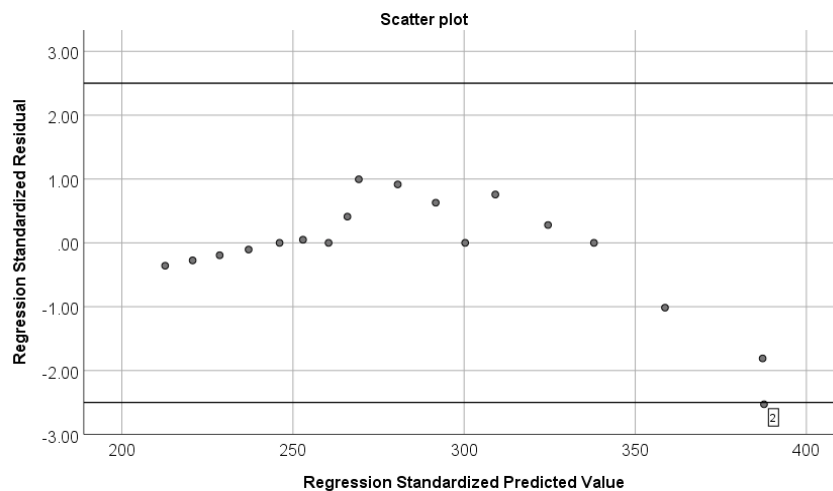
Figure 8. Scatter plot for Regression Standardized Residual (MWQR) for real data

Figure (8) shows that also there is one outlier (point 2) because it was less than (-2.5)., The information of the linear model estimated by (MWQR) is summarized in the table (8):

Table (8): Analysis of MWQR Model for real data

Model OLS	B	Std. Error	t	Sig.	Pseudo R^2	MSE	MAE
(Constant)	-12588.26	267.7654	-47.012	.000			
x1	14.551	.3104	46.881	.000	0.987	33.8086	3.3344
x2	-15.672	.4187	-37.429	.000			
x3	22.360	.7294	30.654	.000			

Table (8) shows that all parameters are significant (p-value = 0.000), and It is more significant than the parameters estimated by the classical method. The pseudo R^2 is 98.7%, Which is greater than its value (86.3%) in the classical method. MSE and MAE are (33.8086 and 3.3344) respectively, they are less than their values (8310.68 and 49.7139, respectively) in the classical method. Thus Multivariate Wavelet Quantile regression model is more efficient than (QR) linear model.

All proposed and classical methods were applied, then the mean absolute error was calculated, and the results were summarized in the following table:

Table (9) MAE for real data

Method	Wavelet	De-noising Method	MAE
Proposed method	Sym2	SURE	3.3344
Proposed method	Db2	SURE	3.3344
Proposed method	Sym2	Minimax	3.3344
Proposed method	Db2	Minimax	3.3344
Proposed method	**Sym2**	**Universal**	**3.2841**
Proposed method	**Db2**	**Universal**	**3.2841**
Quantile regression	-----	------	49.7139

Table (9) shows that all the proposed methods were better than the classical method because the values of (MAE = 3.3344, and 3.2841, respectively) were less than their value in the classical method (MAE = 49.7139). The proposed methods (Sym2 and Db2-universal) performed better than the other proposed methods.

4. Conclusion & Recommendations

Through the study of simulation and real data, the following main conclusions and recommendations were summarized:

4.1 Conclusions

1. Models estimated using the proposed methods (MWQR) are more efficient than the classical method (QR) for all simulation and real data, depending on the MAE.

2. The classical method (QR) is better than the classical method (OLS) depending on the criterion MAE, while the proposed methods were the best of them based on the pseudo R^2, MSE, and MAE, for all simulation and real data.

3. The wavelets (Sym2 and Db2) with the Universal threshold (De-noising Method) were the best compared to the other types of threshold methods.

4. The proposed methods (MWQR) treat contamination while not treating outliers in the data, but are robust to it.

5. The parameters of the models estimated using the proposed methods (MWQR) are more significant than the classical method (QR), and the F-value was larger.

6. For real data, there is a significant effect of the independent variables (food and non-alcoholic beverages, clothes and shoes, and housing, water, gas, electricity and other fuels) on the dependent variable (total household spending), according to the classical and proposed methods.

7. For the real data, the proposed methods interpretation ratio (MWQR) was greater than that of the classical method (QR).

4.2 Recommendations

1. Using the proposed methods (MWQR), because it is more efficient than classical method.
2. Using wavelet quantile regression with a generalized regression model.
3. Using wavelet quantile regression with gamma regression.
4. Use statistical software such as MATLAB and R. To create a package that automatically selects the best wavelet according to the data entered when using wavelet quantile regression.

5. References

1. AHMADI, K., JAVAID, A. Y. & SALARI, E. (2015). An efficient compression scheme based on adaptive thresholding in wavelet domain using particle swarm optimization. *signal processing: image communication,* 32**,** 33-39.

2. Bickel, P. J., & Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *The annals of statistics*, *9*(6), 1196-1217.

3. Chavan, M. S., Mastorakis, N., Chavan, M. N., & Gaikwad, M. S. (2011, February). Implementation of SYMLET wavelets to removal of Gaussian additive noise from speech signal. In *Proceedings of Recent Researches in Communications, Automation, Signal Processing, Nanotechnology, Astronomy and Nuclear Physics: 10th WSEAS International Conference on Electronics, Hardware, Wireless and Optical Communications (EHAC'11), Cambridge* (p. 37).

4. Donoho, D. L., Johnstone, I. M., Kerkyacharian, G., & Picard, D. (1994). Wavelet shrinkage: asymptopia?. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(2), 301-337.

5. HAGIWARA, K. 2021. Bridging between soft and hard thresholding by scaling. *arXiv preprint arXiv:2104.09703*.

6. JOY, J., PETER, S. & JOHN, N. 2013. Denoising using soft thresholding. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering,* 2**,** 1027-1032.

7. Koenker R. 2005. Quantile regression. New York: Cambridge University Press.

8. Percival, D. B., Wang, M., & Overland, J. E. (2004). An introduction to wavelet analysis with applications to vegetation time series. *Community Ecology*, *5*(1), 19-30.

9. Raimondo, M. (2002). Wavelet shrinkage via peaks over threshold. *InterStat*, 1-19.

10. Roger Koenker and Gilbert Bassett, 1978. Regression Quantiles, Econometrica, Vol. 46, No. 1 (Jan., 1978), pp. 33-50.

11. Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate research, 30(1), 79-82.

12. TAHA.H.A.AL-ZUBAYDI 2009. Using Discrete Wavelet filter in Analysis of time series AR(1) and comparison with other filters. *Baghdad university*.

13. Abd Al-Kader, Qias Mustafa ,2011," Comparison of Box-Jenkins models before and after wavelet filtration in terms of demotion with applications, Erbil, Salahaddin University - Erbil (Master's thesis).

Appendix
Table: Real data

series	total family spending	Food items and non-spiritual drinks	clothes and shoes	Housing, water, gas, electricity and other fuels
	y	**x1**	**x2**	**x3**
1	2864.7	685.7	254.1	484.8
2	1286.9	483.5	88.2	259.3
3	2292.9	630.7	194.7	378.8
4	1909.1	623.3	119.3	308.1
5	2513.5	642.6	177	514.8
6	1320.3	480.8	84.2	278.4
7	1820.6	694.6	131.9	405.2
8	1513.2	456.7	77.3	498.8
9	1697.9	529.2	109.5	369.4
10	1406.3	426.7	93.7	403.9
11	1611.8	523.8	113.5	361.3
12	1740	637.7	104.9	354.6
13	1881.2	591.4	116	425.3
14	1133.3	451.2	68.9	248
15	1214.5	475.5	86.8	261.2
16	1182.3	466.2	65.7	312.6
17	1155.3	415.5	82.4	322
18	1588.9	553.2	88.2	394.4

Program

```
clc
clear all
n = 30;
for i=1:1000
%rng('default'); % For reproducibility
t1 = rand(n,1);t2 = rand(n,1);t3 = rand(n,1); epsilon = randn(n,1).*sqrt((t1+t2+t3+0.01));
y = 10 + 3*t1 + t2.*sin(2*t3) + epsilon; Tbl = table(t1,t2,t3,y); numOut = 5;
[~,idx] = datasample(Tbl,numOut);
Tbl.y(idx) = Tbl.y(idx) + randsample([-1 1],numOut,true)'.*(0.9*Tbl.y(idx));
Mdl = TreeBagger(n,Tbl,'y','Method','regression');
tau =.25:.25:0.75; predT = [t1,t2,t3]; quartiles = quantilePredict(Mdl,predT,'Quantile',tau);
meanY = predict(Mdl,predT); eols=abs(meanY-y);d=y.*ones(n,3); eqr=abs(quartiles-d);
MADols(i)=sum(eols/n); MADqr(i)=min(sum(eqr/n)); data=[t1,t2,t3,y];
dataw = wdenoise(data,'Wavelet','sym2', 'DenoisingMethod','sure','ThresholdRule','soft');
t1w=dataw(:,1);t2w=dataw(:,2);t3w=dataw(:,3);yw=dataw(:,4);
Tblw=table(t1w,t2w,t3w,yw); Mdlw = TreeBagger(n,Tblw,'yw','Method','regression');
predT = [t1w,t2w,t3w]; quartilesw = quantilePredict(Mdlw,predT,'Quantile',tau);
d=yw.*ones(n,3); eqrw=abs(quartilesw-d); MADqrw1(i)=min(sum(eqrw)/n);
dataw = wdenoise(data,'Wavelet','db2', 'DenoisingMethod','sure','ThresholdRule','soft');
t1w=dataw(:,1);t2w=dataw(:,2);t3w=dataw(:,3);yw=dataw(:,4); Tblw=table(t1w,t2w,t3w,yw);
Mdlw = TreeBagger(n,Tblw,'yw','Method','regression'); predT = [t1w,t2w,t3w];
quartilesw = quantilePredict(Mdlw,predT,'Quantile',tau); d=yw.*ones(n,3);
eqrw=abs(quartilesw-d); MADqrw2(i)=min(sum(eqrw)/n);
dataw = wdenoise(data,'Wavelet','sym2', 'DenoisingMethod','minimax','ThresholdRule','soft');
t1w=dataw(:,1);t2w=dataw(:,2);t3w=dataw(:,3);yw=dataw(:,4);
Tblw=table(t1w,t2w,t3w,yw); Mdlw = TreeBagger(n,Tblw,'yw','Method','regression');
predT = [t1w,t2w,t3w]; quartilesw = quantilePredict(Mdlw,predT,'Quantile',tau);
```

```
d=yw.*ones(n,3); eqrw=abs(quartilesw-d); MADqrw3(i)=min(sum(eqrw)/n);
dataw = wdenoise(data,'Wavelet','db2', 'DenoisingMethod','minimax','ThresholdRule','soft');
t1w=dataw(:,1);t2w=dataw(:,2);t3w=dataw(:,3);yw=dataw(:,4);
Tblw=table(t1w,t2w,t3w,yw); Mdlw = TreeBagger(n,Tblw,'yw','Method','regression');
predT = [t1w,t2w,t3w]; quartilesw = quantilePredict(Mdlw,predT,'Quantile',tau);
d=yw.*ones(n,3); eqrw=abs(quartilesw-d); MADqrw4(i)=min(sum(eqrw)/n);
dataw = wdenoise(data,'Wavelet','sym2', 'DenoisingMethod','universal','ThresholdRule','soft');
t1w=dataw(:,1);t2w=dataw(:,2);t3w=dataw(:,3);yw=dataw(:,4);
Tblw=table(t1w,t2w,t3w,yw); Mdlw = TreeBagger(n,Tblw,'yw','Method','regression');
predT = [t1w,t2w,t3w]; quartilesw = quantilePredict(Mdlw,predT,'Quantile',tau);
d=yw.*ones(n,3); eqrw=abs(quartilesw-d); MADqrw5(i)=min(sum(eqrw)/n);
dataw = wdenoise(data,'Wavelet','db2', 'DenoisingMethod','universal','ThresholdRule','soft');
t1w=dataw(:,1);t2w=dataw(:,2);t3w=dataw(:,3);yw=dataw(:,4);
Tblw=table(t1w,t2w,t3w,yw); Mdlw = TreeBagger(n,Tblw,'yw','Method','regression');
predT = [t1w,t2w,t3w]; quartilesw = quantilePredict(Mdlw,predT,'Quantile',tau);
d=yw.*ones(n,3); eqrw=abs(quartilesw-d); MADqrw6(i)=min(sum(eqrw)/n);
end
MMADols=mean(MADols), MMADqr=mean(MADqr), MMADqrw1=mean(MADqrw1)
MMADqrw2=mean(MADqrw2), MMADqrw3=mean(MADqrw3), MMADqrw4=mean(MADqrw4),
MMADqrw5=mean(MADqrw5)
MMADqrw6=mean(MADqrw6)
```

استخدام طريقة مقترحة هجينة للانحدار الكمي والموجبة متعدد المتغيرات في تقدير معلمات الأنموذج الخطي

طه حسين علي   و نصر الدين حاج صلاح البرواري   و   ديار لازكين رمضان

الخلاصة

تم في هذا البحث إقتراح طريقة هجينة للإنحدار الكمي والموجبات متعدد المتغيرات لمعالجة مشكلة تلوث البيانات أو وجود القيم الشاذة والتي تستخدم الوسيط بدلاً من المتوسط الذي يعتمد عليه أنموذج الإنحدار الخطي وطريقة التقدير للمربعات الصغرى الإعتيادية. تضمن البحث المقارنة بين الطريقة المقترحة (لعدة مويجات وطرائق تقدير قطع عتبة مختلف) والتقليدية بالإعتماد على متوسط الخطأ المطلق للحصول على أفضل أنموذج إنحدار كمي ملائم للبيانات. تناول الجانب التطبيقي نوعين من البيانات تمثل المحاكاة والبيانات الحقيقية والتحليل بإستخدام برنامج مصمم لهذا الغرض بلغة ماتلاب فضلاً عن البرنامج الإحصائي SPSS-26 و EasyFit-5.5. توصلت الدراسة الى أن الطريقة المقترحة أكثر كفاءة من الطريقة التقليدية في تقدير معلمات أنموذج الإنحدار الكمي إعتماداً على معامل التحديد وعلى معيار متوسط الخطأ المطلق ومتوسط الخطأ التربيعي.

الكلمات الدالة :المويجة متعددة المتغيرات، الانحدار الكمي ، ازالة الضوضاء والعتبة.