11. BÖLÜM / CHAPTER 11

USE OF WEIGHTED MEDIAN IN THEIL-SEN REGRESSION ANALYSIS

Necati Alp ERİLLİ*

*Assoc. Prof. Dr., Sivas Cumhuriyet University, Faculty of Economics and Administrative Science, Econometrics Department, Sivas, Turkey
E-mail: aerilli@cumhuriyet.edu.tr

DOI: 10.26650/B/SS10.2021.013.11

ABSTRACT

Regression analysis; is one of the most used analytical techniques in statistical prediction studies. The aim of regression analysis; to investigate the relationship between the dependent and independent variable or variables, which is estimated to have a cause-effect relationship between them, to explain the assumed relationship between variables functionally and to define this relationship with a model. Non-parametric regression analysis is a method used in cases where some assumptions that are valid for parametric regression methods cannot be met and give successful results.

Nonparametric methods make calculations with the median parameter instead of arithmetic mean. The median parameter does not reflect the effect of outliers on the calculations. Weighted Median calculates the contribution of outlier values to the model even though it gives weight to each observation value. In this study, the Weighted Median parameter was used in the Theil-Sen method which is frequently used in nonparametric regression analysis. Optimum and Hodges-Lehmann calculations which are used in the computation of constant parameters in the Theil-Sen method are also used in this study. The results obtained from Theil-Sen, Hodges-Lehmann and Optimum method used with Weighted Median were tested in different data structures. Thus, the advantages or weaknesses of the use of the weighted median parameter in the Theil-Sen method over the classical methods were determined.

Keywords: Theil-Sen Regression, Weighted Median, Non-parametric Regression Analysis, Hodges-Lehmann, Optimum

1. Introduction

Statistics is the science of making deductions and interpretation from data. In order to determine the meaning of the data, to extract the desired information from the data and to make it ready for use; a great number of statistical techniques have been developed for different types of data. All these statistical methods cannot explain or eliminate uncertainties in the world. However, we can use statistics to make all these uncertainties useful, to explain and interpret them, even if partly. One of the most well-known and widely used of the aforementioned statistical methods is regression analysis.

1.1. Regression Analysis

Regression analysis is one of the most widely used statistical techniques. Regression analysis is used in many fields of academic and applied sciences such as social sciences, medical research, economic studies, agriculture and biology, meteorology, physics, and chemistry. In addition to being an easily understandable method, regression analysis finds wide usage and application area through all kinds of statistical software packages.

Regression analysis is a statistical technique in which we use the observed data to correlate a variable called a dependent variable and one or more independent variables. The aim is to create a regression model or estimation equation that can be used to define, predict and control the dependent variable based on independent variables (Bowerman et al., 2012:491). According to another definition, regression analysis is concerned with the study of dependence of one variable, the dependent variable, on one or more other variables, the explanatory variables, with a view to estimating and/or predicting the (population) mean or average value of the former in terms of the known or fixed values of the latter (Gujarati, 1999:16).

Statistically, the concept of the relationship between two variables is the examination of whether there is a correlation between the exchange of values of these variables. This relationship can also be defined as a cause-effect relationship. The most important problem here is to determine the structure of this relationship. In real-life problems, this structure can often be predicted to a sufficient degree of accuracy, although it cannot be determined most of the time.

The concept of prediction here includes the functional form of the relationship, the direction, and the degree of the relationship. The functional form of the relationship between the variables indicates the mathematical function between the variables. Thus, it can be determined whether the relationship is linear or curved. The purpose in the direction of the relationship is to reveal whether the variables change in the same direction or in different directions. The degree of the relationship gives us information about the strength of the relationship between the two variables. In general, the functional form of the relationship is analyzed by regression analysis and the degree of the relationship is analyzed by correlation analysis (Erilli, 2018:179).

Regression analysis is based on some assumptions. The most important of these assumptions is that the form of the relationship between dependent and independent variables is known. When assumptions are not met, the estimates will not be a good estimate. In this case, regression methods which allow the linearity assumption in the parametric regression to be stretched in order to make a better estimate are needed. These methods are regression models known as non-parametric and semi-parametric regression methods.

1.2. Parametric - Nonparametric Regression

The most important difference between parametric and nonparametric regression methods is based on the confidence in the information obtained from the researcher and data about the regression function. In parametric regression, the researcher selects a possible family of curves from all curves and needs very specific quantitative information about the form of the regression function. Nonparametric regression techniques rely on data more than parametric techniques to obtain information about regression function. Therefore, they are suitable for inference problems. Non-parametric estimators are more appropriate to use when a suitable parametric form cannot be obtained for the regression function because when the parametric model is valid, non-parametric models will have less effectiveness. In addition, non-parametric models can be used to test the validity of the parametric model (Eubank, 1988).

1.2.1. Parametric Regression Analysis

Parametric regression is the expression of dependent and independent variables and the average relationship between these variables with a mathematical function. Assumptions such as normal distribution, equal variance, autocorrelation should be met for the successful implementation of parametric regression analysis. They are the most powerful regression methods in case of assumptions being met.

1.2.2. Non-Parametric Regression Analysis

They are the methods used when some assumptions valid for parametric regression methods are not met. They are effective methods for data with very low sample size or with outliers. In statistical studies, there are robust parametric methods that handle the effects of outlier observations in different ways. However, even these powerful methods may not be able to produce appropriate solutions because the parameters are corrupted due to outlier observations, and the actual structure of the data may not be reflected in the model. In this case, non-parametric regression provides preliminary information (Hardle, 1994). Non-parametric regression method has some drawbacks, although it does not have restrictive assumptions when making predictions. When the number of independent variables is large, it is difficult to make predictions and the obtained graphs are complex. In addition, it is difficult to take into consideration discrete independent variables with a nonparametric method and to interpret individual effects of dependent variables depending on the increase in the number of independent variables. The drawbacks of non-parametric methods can be solved by using a semi-parametric regression model (Horowitz, 1993).

1.2.3. Semi Parametric Regression Analysis

The semi-parametric regression model uses both the parametric and non-parametric regression models together. Therefore, the semi-parametric regression model is not affected by the restrictive assumptions of parametric models, but it combines the attractive features of non-parametric methods (such as Cox, Kernel Regression). Semi-parametric regression models can be defined as a combination of parametric and non-parametric regression models. Semi-parametric regression models are used when nonparametric regression methods cannot make good predictions or when the distribution of errors is not known although the researcher wants to use parametric methods. Normality assumption is not required when parameter estimation is performed with these models (Takezawa, 2006). While it is possible to work with up to two explanatory variables in order to obtain interpretable results in the estimation of the non-parametric model, it is possible to examine the dependent effect of k explanatory variables in the semi-parametric method. In addition, it is suggested that this approach should be used in applied studies since not as many assumptions are made as in the parametric model (Horowitz, 1993).

2. Literature Review

Theil-Sen regression method has been known and used for many years. Different applications in the literature are being tried and the effectiveness of the method in different structures is being investigated. Scholz (1978) studied with the weighted median in regression analysis. Fernandes ve Leblanc (2005), investigated for Parametric (Modified Least Squares) and non-parametric (Theil-Sen) consistent predictors are given for linear regression in the presence of measurement errors together with analytical approximations of their prediction confidence intervals. Lavagnin et al. (2011) reports the combined use of the nonparametric Theil–Sen (TS) regression technique and of the statistics of Lancaster–Quade (LQ) concerning the linear regression parameters to solve typical analytical problems, like method comparison, calculation of the uncertainty in the inverse regression, determination of the detection limit. Wilcox (1998) compares the small-sample efficiency of the extended Theil-Sen estimator to the modified Buckley-James estimator when the predictor is random. Zhou and Serfling (2008) introduce multivariate spatial U-quantiles and develop a corresponding Bahadur–Kiefer representation extending the classical Theil–Sen nonparametric simple linear regression slope estimator, and for robust estimation of multivariate dispersion. Peng et al. (2008) obtain the strong consistency and asymptotic distribution of the Theil–Sen estimator in simple linear regression models with arbitrary error distributions. Siegel (1982) used the repeated median algorithm as a robustified U-statistic in which nested medians replace the single mean. Zaman and Alakus (2019) are examined and discussed with various weights of Theil-Sen method and estimators. In an attempt to show the need for non-parametric methods, results are evaluated based on real-life data. Erilli and Alakus (2014) offer a new method for the estimation of regression parameters under data which have many equal values. The Proposed method and other nonparametric methods such as Theil, Mood-Brown, Hodges-Lehmann methods, and OLS method were compared with the sample data and the results were evaluated. Dang et al. (2008) introduced Theil-Sen estimators in multiple linear regression analysis.

3. Method

Non-parametric regression methods can be grouped under two headings as parameter estimation and Smoothing techniques according to the median parameter. Prediction methods according to median parameter can be listed as Mood-Brown, Theil-Sen, Adichie, Maritz, Kildea methods; while examples of Smoothing techniques can be listed as Kernel Smoothing, Priestley and Chao method, Nadaraya-Watson estimator, K-Nearest neighbor method (Topal, 1999:12-22).

The most commonly used method among these is probably the Theil-Sen method. Its ease of application and success compared to other methods are seen as its most important advantages. Although the increase in the number of observations creates computational difficulties, with the support of package programs, this issue is no longer a problem. Therefore, non-parametric regression studies have increased in recent years.

3.1. Theil-Sen Regression Method

The method was first proposed by Theil (1950) and the procedure is firstly known as Theil's Method. After Sen (1968) highlighted the method with the relationship to Kendall's tau, it is named as the Theil-Sen method. In literature, it's also named as Theil–Kendall as well.

Theil proposed estimating the slope of a regression line as the median of the slopes of all lines joining pairs of points with different x values. For a pair (x_i, y_i) and (x_j, y_j) the relevant slope can be calculated as $S_{ij} = (y_j - y_i) / (x_j - x_i)$. So, there must be $n(n+1) / 2$ slopes for any data.

The $\hat{\beta}_1$ statistic, which is the estimator of the parameter β_1 in simple regression analysis, is calculated as the median of the slope values: $\hat{\beta}_1 = Median(S_{ij})$.

Theil suggested for the estimation of the intercept as $\hat{\beta}_1 = Median(y_i) - \hat{\beta}_1 Median(x_i)$ (Theil, 1950).

3.2. Intercept Parameter Calculations with Theil-Sen Slopes

$\hat{\beta}_1$; To show the slope parameter estimated from the Theil-Sen method, the following two different methods are developed methods for estimating the β_0 parameter (Lehmann and Dabrera, 1975; Hodges and Lehmann, 1963).

3.2.1. Optimum Method

Let's define $d_i = y_i - \hat{\beta}_1 x_i$ calculated for all observations where $\hat{\beta}_1$ is calculated with the Theil-Sen method. $\hat{\beta}_0$ is calculated as the median parameter of all d_i; d_i $\left(\hat{\beta}_0 = Median\{d_i\}\right)$. The optimum approach does not require the assumption of symmetrically distributed d_i. It is better suited especially for data with outliers.

3.2.2. Hodges-Lehmann Method

Hodges-Lehmann method to predict $\hat{\beta}_0$ is defined as the mean value of d_i $\left(\hat{\beta}_0 = Mean\{d_i\}\right)$. The Hodges-Lehmann method can not give power against data with outliers.

3.3. Test of significance of slope parameter

To test $H_0 : \hat{\beta}_1 = 0$, we can use the test statistics given in Equation.1 and Equation.2 (Birkes and Dodge, 1993: 119).

$$|t| = \frac{|U|}{SD(U)} \tag{1}$$

where

$$U = \sum\left[rank(y_i) - \frac{n+1}{2}\right]x_i \text{ and } SD(U) = \sqrt{\frac{n(n+1)}{12}\sum(x_i - \bar{x})^2} \tag{2}$$

The approximate p-value of the test is calculated to be Prob $\left[|Z| \geq |t|\right]$, where Z is a random variable having a standard normal distribution.

3.4. Weighted Median

Median of a list of numbers is obtained by putting the numbers in increasing order and selecting the number in the middle of the ordered list. The weighted median of a list of numbers x_i with heights w_i is obtained as follows:

First, put the numbers x_i in increasing order. By changing the indices, we can arrange so that $x_1 \leq x_2 \leq \cdots \leq x_n$. The weights w_i should be nonnegative and should add to 1. Find the index k such that;

$$w_1 + w_2 + \cdots + w_{k-1} < 0.5$$
$$w_1 + w_2 + \cdots + w_{k-1} + w_k > 0.5 \quad (3)$$

Then x_k is the weighted median. Sometimes it happens that there is an index k such that $w_1 + w_2 + \cdots + w_{k-1} = 0.5$. Then $(w_{k-1} + w_k) / 2$ is the weighted median.

The slope of the line passing through the point (x_0, y_0) that minimizes the sum of absolute deviations from from the data points, can be described as the weighted median of the slopes $b_i = (y_i - y_0) / (x_i - x_0)$ of the lines between data points (x_i, y_i) and the given point (x_0, y_0), with each weight proportional to the x-distance $| x_i - x_0|$ between the two points (Birkes and Dodge, 1993:113-114).

3.4.1. Numerical Example

To understand the use of weighted median in Theil-Sen regression, let's calculate step by step with data with 5 observations. Let our observations be $Y = 100,120,138,145,162$ for the dependent variable and $X = 40,47,45,50,65$ for the corresponding argument. Table.1 shows Theil Slopes and Weighted Median slopes.

Table 1

Theil and weighted median slopes calculations for the numerical example

$Y_i - Y_j$	$X_i - X_j$	$\|X_i - X_j\|$	Theil Slope	Weighted Median Slope
-20	-7	7	-20/-7=2.857142857	7/110=0.0636363
-38	-5	5	-38/-5=7.6	5/110=0.0454545
-45	-10	10	-45/-10=4.5	10/110=0.0909090
-62	-25	25	-62/-25=2.48	25/110=0.22727272
-18	2	2	-18/2=-9	2/110=0.0181818
-25	-3	3	-25/-3=8.33333	3/110=0.0272727
-42	-18	18	-42/-18=2.33333	18/110=0.16363636
-7	-5	5	-7/-5=1.4	5/110=0.0454545
-24	-20	20	-24/-20=1.2	20/110=0.181818
-17	-15	15	-17/-15=1.133333	15/110=0.136363
		$\Sigma = 110$		

The Theil slopes values calculated in Table 1 are sorted from small to large. The corresponding weighted median slopes values are also written with them and their cumulative totals are taken in another column. This total must be equal to 1. All these values are given in Table 2.

Table 2
Ordered and corresponding weighted median values for the numerical example

Ordered Theil Slope	Corresponding weighted median	Ordered weighted median
-9	0.018181818	0.018181818
1.133333333	0.136363636	0.154545455
1.2	0.181818182	0.336363636
1.4	0.045454545	0.381818182
2.333333333	0.163636364	0.545454545
2.48	0.227272727	0.772727273
2.857142857	0.063636364	0.836363636
4.5	0.090909091	0.927272727
7.6	0.045454545	0.972727273
8.333333333	0.027272727	1

In the ordered weighted median column given in Table 2, the first ordered Theil slope value exceeding 0.50 is taken as the weighted median value. In this example, the first ordered weighted median value exceeding 0.50 is 0.5454 and the corresponding Theil slope value is 2.333. This value is our β_1 value calculated by the weighted median (The median value of the ordered Theil slope values is calculated as 2.4066667 for comparison).

We can calculate the intercept parameter with this formula: $\hat{\beta}_0 = Median(Y) - \hat{\beta}_1 \times Median(X)$. It can be found as; $\hat{\beta}_0 = 138 - 2.333 \times 47 = 28.333$. Thus, the Theil-Sen regression equation calculated by the weighted median will be $\hat{Y} = 28.333 + 2.333X$.

4. Application

In the application part, the Theil-Sen regression method is performed with the weighted median parameter. The advantages and weaknesses of the Theil-Sen method with weighted median methods are determined and interpreted according to OLS and Theil-Sen regression methods with Mean Absolute Error (MAE) given in Equation.4:

$$MAE = \frac{1}{n}\sum_{j=1}^{n}\left|y_i - \hat{y}_j\right| \tag{4}$$

This is because MAE is more robust to outliers since it does not make use of the square. On the other hand, MSE is more useful if we are concerned about large errors whose consequences are much bigger than equivalent smaller ones.

The estimation model for the sample data was obtained by 7 different methods. These methods are; Ordinary Least Squares, Theil-Sen, Theil-Sen with Hodges-Lehmann, Theil-Sen with Optimum, Weighted Median Theil-Sen, Weighted Median Theil-Sen with Hodges-Lehmann and Weighted Median Theil-Sen with Optimum.

The first application is data with 20 observations. Observation values for the dependent variable (*Y*) and the independent variable (*X*) are given in Table 3.

Table 3
Values of Data.1

Y	8	12	13	15	50	15	16	48	19	20
X	36	58	61	45	63	76	41	70	69	50
Y	11	22	23	33	20	22	26	37	34	43
X	75	74	79	61	81	100	102	116	59	115

The scatter plot of these values is also given in Figure.1.

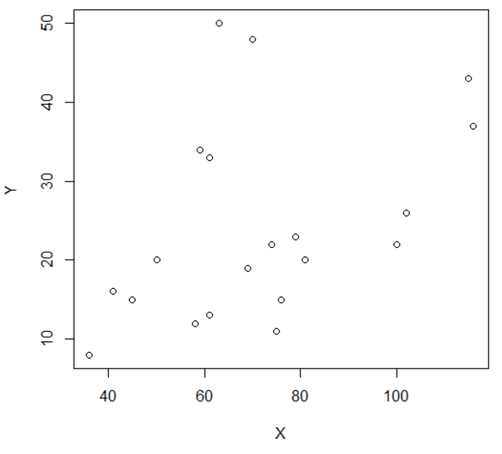

Figure 1. Scatterplot for Data.1

Parameters obtained from the 7 methods mentioned above and MAE values are given in Table 4.

Table 4

Intercept, slope, and MAE results obtained from 7 methods for Data.1

Method	Intercept	Slope	MAE
OLS	8.117673	0.2268669	8.7029361
Theil-Sen	6.257575	0.212121	7.9863622
Theil-Sen Hodges-Lehmann	9.172742	0.212121	8.7518188
Theil-Sen Optimum	5.848498	0.212121	7.9848488
Theil-Sen W. Median	4.64705	0.235294	7.9852907
Theil-Sen W. Median - Hodges-Lehmann	7.514714	0.235294	8.6750003
Theil-Sen W. Median - Optimum	4.411772	0.235294	7.9558825

When we look at the results in the table, the minimum MAE value belongs to Theil-Sen / Optimum obtained with the weighted median. The remarkable detail in the table is that the values of Theil-Sen, Hodges-Lehmann, and Optimum which are calculated by weighted median have lower MAE than the values calculated by the median.

The second data set was obtained by adding two outlier values to the data given in Sen (1968). Observation values of the data are given in Table 5.

Table 5

Values of Data.2

Y	9	15	19	20	45	55	78	30	50
X	1	2	3	4	10	12	18	12.5	4.5

The scatter plot of these values is also given in Figure 2.

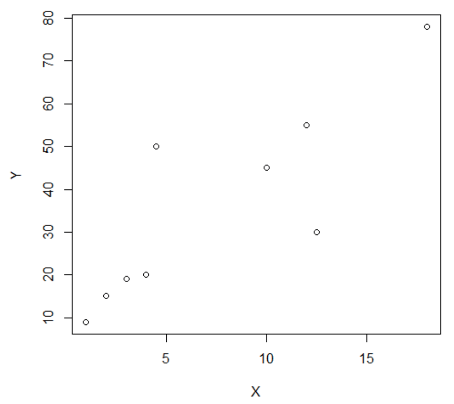

Figure 2. Scatter Plot for Data.2

Parameters obtained from the 7 methods mentioned above and MAE values are given in Table 6.

Table 6
Intercept, slope, and MAE results obtained from 7 methods for Data.2

Method	**Intercept**	**Slope**	**MAE**
OLS	11.35785	3.265364	74.5198
Theil-Sen	12.140625	3.96875	94.1719
Theil-Sen Hodges-Lehmann	6.121527	3.96875	59.2535
Theil-Sen Optimum	6.5625	3.96875	58.8125
Theil-Sen W. Median	12.28125	3.9375	93.3438
Theil-Sen W. Median - Hodges-Lehmann	6.354166	3.9375	59.3958
Theil-Sen W. Median - Optimum	7.125	3.9375	58.625

When we look at the results in Table 6, it is seen that the method of giving minimum MAE value is Theil-Sen / Optimum obtained by the weighted median. The result of OLS is better for both the Theil-Sen method with weighted median and classical method.

Third data set is given in Table 7 which has 20 samples with many equal values (Erilli and Alakus, 2014).

Table 7
Values of Data.3

Y	30	40	45	50	50	55	60	70	60	75	80	90	85	80	90	92	95	98	96	100
X	4	5	5	5	6	8	9	9	10	10	10	10	10	11	11	11	12	13	14	15

The scatter plot of these values is also given in Figure 3.

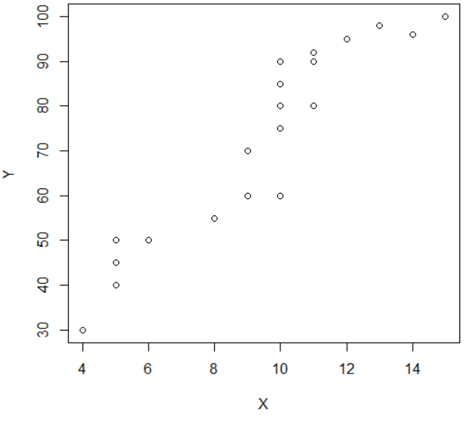

Figure 3. Scatter Plot for Data.3

Parameters obtained from the 7 methods mentioned above and MAE values are given in Table 8.

Table 8
Intercept, slope, and MAE results obtained from 7 methods for Data.3

Method	**Intercept**	**Slope**	**MAE**
OLS	10.36379	5.652363	9.721839
Theil-Sen	10.84	6.666	6.2002
Theil-Sen Hodges-Lehmann	9.3896	6.666	6.17827
Theil-Sen Optimum	10.005	6.666	6.1168
Theil-Sen W. Median	11.25	6.625	6.2125
Theil-Sen W. Median - Hodges-Lehmann	10.3125	6.625	6.125
Theil-Sen W. Median - Optimum	9.775	6.625	6.1725

4.1 Test of significance of slope parameters

In this subsection of the study, the significance of the slope parameters of the models of the three different data predicted above was tested. Table.9 shows the *t* values for the slope parameters in the models obtained according to OLS and Theil methods for three data sets. Significance level was taken as 0.05.

Table 9
Slope parameter t-values for all data sets

<table>
<tr><th></th><th>Method</th><th>t value</th><th>critic value</th></tr>
<tr><td rowspan="2">Data.1</td><td>OLS</td><td>1,93</td><td rowspan="2">1,734</td></tr>
<tr><td>Theil-Sen</td><td>2,176</td></tr>
<tr><td rowspan="2">Data.2</td><td>OLS</td><td>1,71</td><td rowspan="2">1,894</td></tr>
<tr><td>Theil-Sen</td><td>3,078</td></tr>
<tr><td rowspan="2">Data.3</td><td>OLS</td><td>11,28</td><td rowspan="2">1,734</td></tr>
<tr><td>Theil-Sen</td><td>4,138</td></tr>
</table>

When we look at the results in Table 9, it is seen that the slope parameters obtained by the Theil-Sen method are significant in all data sets ($t_{value} > t_{critical}$). According to the OLS method, the slope parameter was found to be significant only for the model obtained from the 3rd data set ($t_{value} > t_{critical}$), whereas it was not significant in the other two models ($t_{value} < t_{critical}$).

For comparison, Theil-sen significant values calculated with Equation.1 and 2 are given in Table.10:

Table 10
Test of significance values for all data sets

	Data.1	Data.2	Data.3
U	1274,5	140	331
SD(U)	585,5111	45,474	79,9875

5. Conclusion

Model prediction studies have always been one of the most interesting areas for sciences such as statistics, economics, and econometrics. There are many different methods with different model structures and almost all of them have the same object: To make strong predictions with significant parameters. The OLS method is the most powerful estimation method when assumptions are achieved. Assuming OLS method assumptions in regression analysis is a very difficult issue. In particular, working in a small data structure makes these assumptions more difficult to achieve. Since the Theil - Sen regression method does not require heavy assumptions such as OLS regression analysis, it can be applied successfully even in very small data sets. For simple regression analysis, Theil - Sen regression analysis can be considered as an alternative method to the OLS method.

In this study, the use of the weighted median instead of the median is introduced in the Theil-Sen method. The median parameter does not reflect the effect of outliers on the average calculation. Weighted median adds some effect to each observation value because it gives a certain weight. The weighted median in Theil-Sen results was tested in three different data structures. The method which was tested in data with outliers and with many equal observations was compared with OLS and classical Theil-Sen methods. According to the results, it can be said that the use of weighted median is quite successful. Although the weighted median cannot be calculated very easily, such as the Median parameter, it can be used successfully in non-parametric methods.

References

Birkes, D., & Dodge Y. (1993). *Alternative Methods of Regression*. NY, USA: John Wiley & Sons Inc.

Bowerman, B. L., O'Connell, R. T., Murphree, E. S., & Orris, J. B. (2013). *İşletme İstatistiğinin Temelleri.* (N. Orhunbilge, M. Can & Ş. Er, Çev.) Ankara: Nobel Yayınları.

Dang, X., Peng, H., Wang, X., & H. Zhang (2008). Theil-Sen Estimators in a Multiple Linear Regression Model, Submitted paper. Available on http://home.olemiss.edu/~xdang/papers.

Erilli, N. A. (2018). *İstatistik-2*. Ankara: Seçkin Yayınları.

Erilli, N. A., & Alakus K. (2014). Non-Parametric Regression Estimation for Data with Equal Values. *European Scientific Journal, 10*(4), 70–82.

Eubank, R. L. (1988). *Spline Smoothing and Nonparametric regression*. New York, USA: Marcel Dekker Inc.

Fernandes, R., & Leblanc, S. G. (2005). Parametric (modified least squares) and non-parametric (Theil–Sen) linear regressions for predicting biophysical parameters in the presence of measurement errors. *Remote Sensing of Environment, 95*, 303–316.

Gujarati, D. (1999). *Temel Ekonometri*. (Ü. Şenesen & G. Günlük Şenesen, Çev.) İstanbul: Literatür Yayıncılık.

Hardle, W. (1994). *Applied Nonparametric Regression*. London, UK: Cambridge University Press.

Hodges, J. L., & Lehmann, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statistics, 34*, 598–611.

Horowitz, J. L. (1993). Semiparametric Estimation of a Work-Trip Mode Choice Model. *Journal of Econometrics*, *58*, 49–70.

Lavagnini, I., Badocco, D., Pastore, P., & Magno F. (2011). Theil–Sen nonparametric regression technique on univariate calibration, inverse regression and detection limits. *Talanta, 87*, 180–188.

Lehmann, E. L., & Dabrera H. J. M. (1975). *Nonparametrics: Statistical Methods Based on Ranks*. SF, USA: Holden-Day Inc.

Peng, H., Wang, S., & Wang X. (2008). Consistency and asymptotic distribution of the Theil–Sen estimator. *Journal of Statistical Planning and Inference*, *138*, 1836–1850.

Scholz, F. W. (1978). Weighted median regression estimates. The annals of statistics, *6*(3), 603–609.

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall’s tau. *Journal of the American Statistical Association*, *63*, 1379–1389.

Siegel, A. F. (1982). Robust regression using repeated medians. *Biometrika*, *69*(1), 242–244.

Takezawa, K. (2006). *Introduction to Nonparametric Regression.* Canada: Wiley-Interscience.

Theil, H. A. (1950). Rank invariant method of linear and polynomial regression analysis. *III. Nederl.Akad. Wetensch.Proc., Series A*, *53*, 1397–1412.

Topal, M. (1999). *Nonparametrik regresyon metodlarının incelenmesi* (Yüksek lisans tezi, Atatürk Üniversitesi, Fen Bilimleri Enstitüsü, Erzurum). https://tez.yok.gov.tr/UlusalTezMerkezi/ adresinden edinilmiştir.

Wilcox, R. R. (1998). Simulations on the Theil-Sen regression estimator with right-censored data. *Statistics & Probability Letters*, *39*, 43–47.

Zaman, T., & Alakus, K. (2019). Comparison of Parametric and Non-Parametric Estimation Methods in Linear Regression Model. *alphanumeric journal*, *7*(1), 13–24.

Zhou, W., & Serfling, R. (2008). Multivariate spatial U-quantiles: A Bahadur–Kiefer representation, a Theil–Sen estimator for multiple regression, and a robust dispersion estimator. *Journal of Statistical Planning and Inference*, *138*, 1660–1678.