FUZZY REGRESSION ANALYSIS AND AN APPLICATION

BULANIK REGRESYON ANALİZİ VE BİR UYGULAMA

SADIK ÖZKAN GÜNDÜZ

PROF. DR. ÖZLEM MÜGE TESTİK
Supervisor

Submitted to Graduate School of Science and Engineering of Hacettepe University as a Partial Fulfillment to the Requirements for the Award of the Degree of Master of Science in Industrial Engineering

2021

ABSTRACT

FUZZY REGRESSION ANALYSIS AND AN APPLICATION

Sadık Özkan GÜNDÜZ

Master of Science, Department of Industrial Engineering

Supervisor: Prof. Dr. Özlem Müge TESTİK

June 2021, 106 pages.

Wind energy, one of the renewable energy sources, is immensely popular today due to the increase in environmental awareness, the decrease in the number of fossil fuels, and the increase in the cost of these fuels. Although wind energy is a clean and nature-friendly energy source, the wind is not a continuous energy source. In addition, the establishment of farms that convert wind energy into electrical energy is expensive and requires technical capacity. Determining the locations where wind turbine farms will be established, which will provide long-term profit to its investors and require considerable amounts of financing at the beginning, is significantly vital in terms of the economic use of resources. It is necessary to collect many meteorological data such as wind speed, wind direction, air density, temperature, pressure, and relative humidity from at least one year ago at the stage of determining the wind turbine construction locations. The wind turbine manufacturer creates theoretical information about how much electrical energy the turbine will generate at what wind speed. Following the collection of meteorological data, various numerical and statistical models are made with the help of the theoretical electricity generation data, and the suitability of the construction location is evaluated. However, when similar wind turbines are examined, it will be seen that there are differences between the theoretical production amount given by the manufacturer and the actual amount of electricity produced at the same wind speed. In this condition, it is clear that there is a fuzzy relationship between wind speed and the electrical energy produced.

For this thesis, the amount of electrical energy to be produced by a wind turbine is estimated by using only wind speed or wind speed and wind direction data with fuzzy linear regression methods. In addition, the amount of produced electrical energy and the wind speed in the data set are fuzzified. Succeeding, crisp input crisp output, crisp input fuzzy output, and fuzzy input fuzzy output situations were estimated with four different fuzzy regression methods and the results were compared.

This application is intended to determine the general framework for the locations where the wind turbine is planned to be installed before the complex calculations and modeling, or when seasonal observations are made rather than annual, or in cases where the observed values are not dependable or there are many site alternatives but there is not enough time to decide on site selection. It has been determined that it will be beneficial in situations. Therefore, it will bring a different approach to the literature.

Finally, this study will open a new window to the methods by establishing the basis for the Fuzzy Partial Regression Method and Fuzzy Nonlinear regression methods that are expected to be used in the future in estimating the energy produced by wind turbines.

Keywords: Fuzzy Logic, Fuzzy Regression, Wind Energy.

ÖZET

BULANIK REGRESYON ANALİZİ SVE BİR UYGULAMA

Sadık Özkan GÜNDÜZ

Yüksek Lisans, Endüstri Mühendisliği Bölümü
Tez Danışmanı: Prof. Dr. Özlem Müge TESTİK
Haziran 2021, 106 sayfa

Günümüzde çevre duyarlılığının artması, fosil yakıtların miktarının azalması ve dolayısıyla bu yakıtların maliyetinin artmasından dolayı, yenilenebilir enerji kaynaklarından rüzgâr enerjisi çok rağbet görmektedir. Ancak, rüzgâr enerjisinin temiz ve doğa ile barışık bir enerji kaynağı olmasına rağmen, rüzgâr sürekliliği olan bir enerji kaynağı değildir. Ayrıca, rüzgâr enerjisini elektrik enerjisine dönüştüren çiftliklerin kurulması oldukça pahalı ve teknik kapasite gerektirmektedir. Uzun vadede yatırımcılarına kazanç sağlayacak ve başlangıçta yüksek meblağlarda finansman gerektiren rüzgâr türbini çiftliklerinin, kurulacağı mevkilerin belirlenmesi ekonomik kaynakların tasarruflu kullanılması açısından büyük önem arz etmektedir. Rüzgâr türbini çiftliklerinin, kurulması planlanan mevkilerin belirlenmesi aşamasında, en az bir yıl öncesine ait rüzgâr hızı, rüzgâr yönü, hava yoğunluğu, sıcaklık, basınç ve bağıl nem gibi birçok meteorolojik verilerin toplanması gerekmektedir. Rüzgâr türbini üreticisi tarafından oluşturulan, türbinin hangi rüzgâr hızında ne kadar elektrik enerjisi üreteceğine dair teorik bilgiler mevcuttur. Meteorolojik verilerin toplanmasına müteakip, üreticinin verdiği teorik elektrik üretim miktarları verisi yardımıyla çok çeşitli nümerik ve istatistiksel modellemeler yapılarak, seçilecek yerlerin uygunluğu tespit edilmeye çalışılır. Ancak, benzer rüzgâr türbinleri incelendiğinde, üreticinin verdiği teorik üretim miktarı ile aynı rüzgâr hızında gerçekte üretilen elektrik miktarları arasında

farklılıklar olduğu görülecektir. Bu koşulda, rüzgâr hızı ile üretilen elektrik enerjisi arasında aslında bulanık bir ilişki olduğu açıktır.

Bu tezin amacı kapsamında, bulanık doğrusal regresyon metotları ile sadece rüzgâr hızı ya da rüzgâr hızı ve rüzgâr yönü verilerini kullanarak bir rüzgâr türbini tarafından üretilecek elektrik enerjisi miktarı tahmin edilmiştir. Ayrıca, çeşitli modellerde veri setinde tahmin edilmeye çalışılan üretilecek elektrik enerjisi miktarı ile rüzgâr hızı da bulanıklaştırılarak, kesin girdi kesin çıktı, kesin girdi bulanık çıktı ve bulanık girdi ve bulanık çıktı durumları da dört farklı regresyon metodu ile tahmin edilmiş ve sonuçlar karşılaştırılmıştır.

Bu uygulama, güç üretimini tahmin etmeye yönelik bulanık regresyon yöntemlerinin, rüzgar türbininin kurulması planlanan yerler için karmaşık hesaplamalar ve modellemelerden önce genel çerçevesinin belirlenmesinin istendiğinde, ya da yıllık değil mevsimsel gözlemlerin yapıldığı durumlarda, ya da gözlemlenen değerlerin güvenilir bulunmadığı durumlarda ya da çok yer alternatifinin olduğu ancak yer seçimi kararının verilmesi için yeterli zamanın bulunmaması durumlarında faydalı olacağı tespit ediliş olup literatüre farklı bir yaklaşım getirecektir.

Son olarak, bu çalışma rüzgâr türbinlerinin üreteceği enerjinin tahmin edilmesinde gelecekte kullanılması beklenen Bulanık Parçalı Regresyon Metodu ile Bulanık Doğrusal olmayan regresyon metotlarına temel oluşturarak, bu metotlara yeni bir pencere açacaktır.

Anahtar Kelimeler: Bulanık mantık, Bulanık Regresyon, Rüzgâr Enerjisi.

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere appreciation to my advisor, Prof.Dr. Ozlem Müge TESTIK, for her unwavering support of my M.Sc. studies, as well as her patience, encouragement, and vast knowledge. Her advice was invaluable to me during the research and writing of this thesis. I attribute her encouragement and effort to the degree of my M.Sc., and I feel incredibly lucky to have an advisor like her.

I am also thankful to Prof.Dr. Murat Caner TESTIK for giving me the opportunity to enroll the Graduate program at Hacettepe University.

I would like to thank to esteemed instructors Assoc.Prof.Kumru Didem ATALAY and Assistant Professors; Banu Y. OZKAYA, Diclehan T. OZTURK, Volkan SONMEZ, who are in the Thesis Monitoring Committee and shared their valuable contributions.

On this difficult road we move forward together; I would like to express my endless thanks and gratitude to my love, Ipek, who always sheds light on me and always supports me with the understanding, respect and endless love. I dedicate this work to her with all my heart.

S. Ozkan GUNDUZ

June 2021, Ankara

TABLE OF CONTENTS

TABLE OF FIGURES

TABLES

ABBREVIATIONS

CICO	Crisp Input Crisp Output
CIFO	Crisp Input Fuzzy Output
FIFO	Fuzzy Input Fuzzy Output
FLR	Fuzzy Linear Regression
FLS	Fuzzy Least Squares
GOF	Goodness of Fit
GPA	Goal Programming Approach
LP	Linear Programming
MF	Membership Function
MSE	Mean Square Error
nsTFN	Non-symmetric Triangular Fuzzy Number
sTFN	Symmetric Triangular Fuzzy Number
TEF	Total Error Fit
TFN	Triangular Fuzzy Number
WTPC	Wind Turbine Power Curve

1 INTRODUCTION

Since the industrial revolution, the usage of energy and the need for energy has increased with high consumption unceasingly [1]. Human beings have benefited from conventional energy resources such as biomass and fossil resources in order to encounter the increasing energy need with the increasing human population. In the last century, researchers have turned to different energy resources such as wind and sun, due to the decrease in the number of fossil resources, increase in prices of them, and especially the increase in the amount of carbon dioxide in the atmosphere [2].

Usage of wind energy have increased speedily in the last 20 years. Renewable energy makes wind energy attractive because it is an environmental-friendly and efficient energy source when applied properly [3].

The installed power generated by wind energy was only 24 GW in 2001, it reached 743 GW all over the world by 2021. In twenty years, wind energy installation increased thirty times. The situation proves that wind energy is quite important for the future of humanity and wind energy investments will continue to increase in the future. Europe is the world leader in wind energy with a total installed power of 220 GW [4]. Turkey ranks seventh in this list with an installed power value of 9 GW [5]. So, wind energy earns popularity not only in Europe but also in Turkey. Correspondingly to the situation, many studies have been done to improve the capacity and the efficiency of wind energy tools due to the potential of wind energy.

Statistical analysis and big data applications have permeated many sectors, including renewable energy, storage, in tandem with the speedy growth of information science and the fast convergence of conventional industries and intelligent technology [6].

In essence, finding an effective method that can provide precise wind power projections is crucial, as it can assist the efficiency of wind turbine plants, reduce unfavorable effects in the wind energy installation scenario, and improve profits of the wind energy earnings through the optimization of bidding strategies [7].

Wind energy is naturally intermittent due to the high correlation of fluctuating and varying wind speeds and other meteorological parameters. The feature makes it difficult to predict accurately and results in a relatively poor generation result for large-scale wind energy input into the system. For this reason, incorrect wind energy estimation can cause many problems and ineffective usage of economic resources [8].

To maximize the generated electrical energy, it is necessary to estimate the power generation of a wind turbine. In recent years, various models have been developed. Usually, a similar format is used when conducting wind power studies.

Primarily, meteorological models and historical power generation data of wind plants are combined for a place where is evaluated as a potential construction area. Then, seasonal or characteristic correlated data of wind speed and load are collected for the same period. The forecasts are provided by the preprocessed results that are the adoption of the power system models and statistical analysis models [9]. The mentioned models are physical models and conventional statistical models [10].

Proposed methods for predicting power generation have been revealed to be frequently complex. Since estimated parameters are affected by a variety of factors, the condition necessitates the use of either parametric or non-parametric techniques. The main parametric methods are the Linearized Model, Polynomial Model, Probabilistic Model, and Logistic function model. The main non-parametric methods are Neural Networks, Data Mining Algorithms, and Fuzzy Clustering Methods [11].

It is assumed that there is a relation between wind speed and generated power, but there are also unexplainable factors that affect wind speed, power generation, and the relation of them. The thesis was driven by the unknown associations in the dataset's context, which led to the use of fuzzy regression. Because the application of fuzzy set theory is popular for the kind of subjects, and it is capable to generate not only crisp decisions but also corresponding degrees of membership.

In this study, the power that would be generated by a wind turbine with wind speed and wind direction data is used to predict with the help of the various fuzzy linear regression (FLR) methods. Fuzzy regression methods do not consider distribution and there is a fuzzy relation between wind speed, wind direction, and power generation. Thus, fuzzy logic is implemented to the regression analysis. The methods to predict the power generation are especially influential and advantageous when a general frame of the wind turbine models is required before complex calculations, when there are a small number of observations that are not trustworthy, or when there are more optional places to construct a wind farm for decision-maker.

In addition to the goals, it is believed that the study will be the basis and open a new door to the unused application of fuzzy piecewise regression and fuzzy non-linear regression for the estimation of wind turbine energy estimation.

2 REGRESSION ANALYSIS

The systems which have not randomness for succeeding observations are called Deterministic Systems. In these systems, the same outputs are always created by the same inputs at the identified beginning. Non-deterministic or stochastic systems, conversely include randomness, are the systems that cannot produce equivalent outputs with equivalent inputs. Nevertheless, it does not mean that calculations or predictions related to results cannot be made.

The complications in the science and engineering era can be resolved by examination or analysis of the relationship between two or more variables. The statistical tools are generally used to explore and model the non-deterministic connection between variables. The mentioned methods are termed Regression Analysis [12]. Francis Galton, one of the leading scientists of the 18th century, was the first to use and develop the concepts of correlation and regression. Karl Pearson, a colleague, and researcher of Galton continued Galton's work after Galton's death. Although the correlation coefficient, which is widely used today, is known as the Pearson correlation coefficient, this concept is essentially based on Galton's studies [13].

In regression models, the one or more variables that affect or produce the output are defined as independent variables, outputs are stated as a dependent variable in general. However, the definition can generate confusion. Therefore, it is possible to call the independent variable a regressor or predictor variable and a dependent variable as a response variable [14].

2.1 Regression Types

In literature, the regression analysis can be classified with the number of regressors the shape of the regression line or the type of the dependent variable, or whether the regression has parametric or non-parametric structures. Not only these main attributes but also subsidiary attributes are used to detail the type of regression analysis.

Montgomery et al. [14], classified the types of regression analysis methods shown as in Figure 2.1.

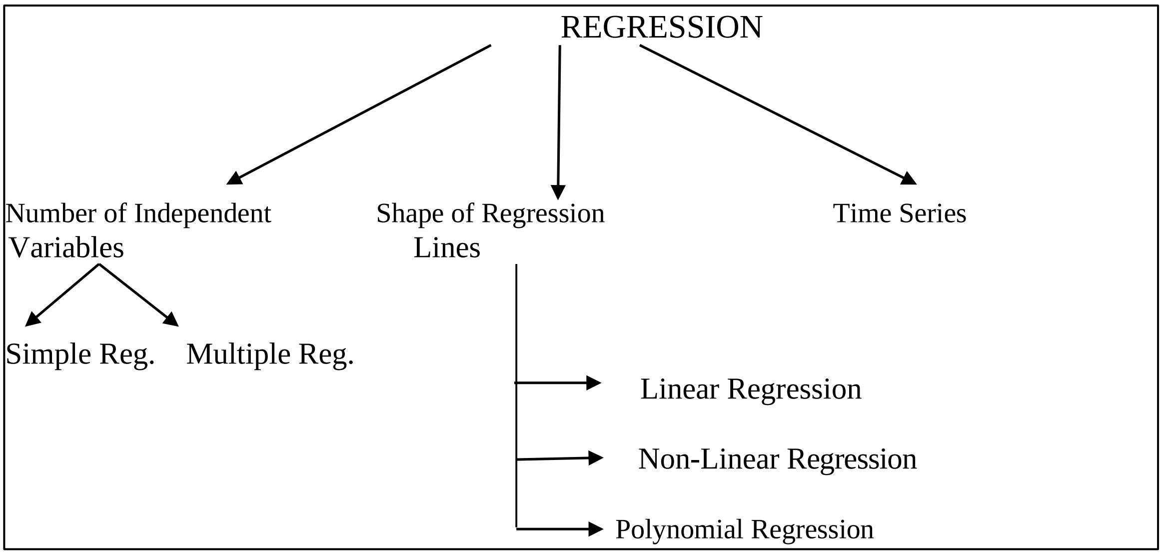

Figure 2. 1 Classification of the Regression Models

Although there are many subordinate types for regression models, only principal types for regression analysis are indicated above.

Regression analysis is a statistical model used to define or predict the causal relationship between a dependent variable and one or more independent variables. Within the framework of this definition, a regression model, *Y* dependent variable, *k* number of independent variables, and *X*'s independent variables can be defined as below:

$$\boldsymbol{Y = f(X_1, X_2, X_3, \dots, X_k)} + \varepsilon \qquad (2.1)$$

The last term in Equation 2.1 is called the error term and this term indicates the mismatch between the estimated value obtained by the model and the actual value. The expression $f(X_1, X_2, X_3, \dots, X_k)$ seen in the equation allows the model to be defined as linear or nonlinear. The mathematical model of Equation 2.1 above is shown below:

$$Y = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \varepsilon \qquad (2.2)$$

Written in Equation 2.2 format, it is a multiple linear regression model in terms of both variables and parameters which is the most general form:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \varepsilon \tag{2. 3}$$

However, the model in Equation 2.3 is an example of a non-linear relation between variables. Also, some models are non-linear according to their parameters like in Equation 2.4 below:

$$Y = \beta_0 X^{(\beta_1)} e^{\varepsilon} \tag{2. 4}$$

However, when the expression "X_1^2" in Equation 2.3 is defined as "X_2" as a new variable, a model in Equation 2.5 is obtained and the model transforms into a linear model with this new form.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon \tag{2. 5}$$

After the logarithms of both sides are taken in Equation 2.4, a linear model is reached as in Equation 2.6 which is written below:

$$log(Y) = log(\beta_0) + \beta_1 log(X_1) + \varepsilon \tag{2. 6}$$

As described above, some models that are not linear in terms of variables and parameters can be transformed into linear models with the help of appropriate transformations. Thus, these models generally can be defined as linear regression models. However, there are also models in which linearity cannot be achieved as a result of any transformation and these models are also classified as nonlinear regression models.

2.1.1 Linear Regression Analysis

Linear regression models are shaped with the assumption that there is a linear relationship between the regressor(s) and the dependent variable. If there is a single independent variable in the model, the model is called a **Simple Linear Regression Model**; on the other hand, if there are more than two independent variables in the model, the model is called a **Multiple Linear Regression Model** [15].

The researchers have two general purposes in multiple linear regression. The first is to estimate the value of the dependent variable via the independent variables assumed to affect the dependent variable. The second is to identify which of the independent variables affect the dependent variable more and capable to define the relationship between them [16]. If the relationship between *Y* dependent variable and *p* number of independent variables is linear and if there are observation values of *Y* and ***X***, the multiple linear regression model is expressed as follows:

$$Y = \boldsymbol{X}\beta + \varepsilon \tag{2. 7}$$

In the formula, *Y* designates an *n*-dimension vector for *n*- number of observations, ***X*** labels a matrix which is formed by the number of independent variables and observations. *β* shows an *n*-dimension vector like *Y*. ε, which is an *n*-dimension vector that defines the error between observation and prediction. The error terms are considered that they are normally distributed, the error is assumed that its mean equals to zero (*E(ε)=0*) and its variance is constant (*var(ε)=* $\sigma^2\boldsymbol{I}$) [17].

2.2 Least Square Estimation Method

The primary objective of regression analysis is to model the relationship between variables accurately by predicting the unknown parameters. Numerous methods are applied to estimate the parameters of regression models. The main purpose of the estimation methods is to minimize the total errors which are calculated as the distance between the estimated regression line and the observations. One of the methods which use the minimization of total sums calculated by the negative or positive observations located above an under-regression line is the Least Absolute Deviation Regression [18]. This method is expressed as below:

$$\boldsymbol{Min}\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}\left|\left(\boldsymbol{Y}_{\mathrm{i}} - \widehat{\boldsymbol{Y}}_{\mathrm{i}}\right)\right| \tag{2. 8}$$

The method occasionally cannot reach common solutions; moreover, it permits obtaining many regression lines that have the same total absolute error. Contrarily, in the least Squares Estimation method, using the squares of errors vanishes the problem mentioned above because of the negative and positive of errors, also it emphasizes the effects of the observation with a bigger error. Because of this reason, the Least Squares method is mostly used to estimate the parameters of a regression model today. The mathematical expression of the Least Squares Method is given below:

$$\text{Min} \sum_{i=1}^{n} \left(\boldsymbol{Y_i} - \boldsymbol{\widehat{Y}_i}\right)^{2} \tag{2. 9}$$

If Equation 2.9 is detailed:

$$\text{Min}\sum_{i=1}^{n} \left(\boldsymbol{Y_i} - \left(\boldsymbol{\beta_0} + \boldsymbol{\beta_1 x_{1i}} + \boldsymbol{\beta_2 x_{2i}} + \cdots + \boldsymbol{\beta_p x_{pi}}\right)\right)^{2} \tag{2. 10 }$$

is attained.

Here, the Multiple Regression Model can be written again as:

$$Y = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{ip} + \varepsilon_i \qquad \text{(i=1,2, … , n)}$$

$$= \boldsymbol{\beta_0} + \sum_{j=1}^{k} \boldsymbol{\beta_j x_{ij}} + \boldsymbol{\varepsilon_i}, \boldsymbol{i = 1, 2, \ldots, n} \tag{2. 11}$$

After, the least square function is shown below:

$$\boldsymbol{S(\beta_0, \beta_1, \ldots, \beta_k)} = \sum_{i=1}^{n} \boldsymbol{\varepsilon_i}^{2} = \sum_{i=1}^{n} \left(\boldsymbol{y_i} - \boldsymbol{\beta_0} - \sum_{j=1}^{k} \boldsymbol{\beta_j xij}\right)^{2} \tag{2. 12}$$

According to normal procedure, the function ***S*** should be minimized concerning $\beta_0, \beta_1, \ldots, \beta_k$. In other words, the derivation of the function ***S*** with respect to β_0 will be calculated. Thus, the least-square estimators of $\beta_0, \beta_1, \ldots, \beta_k$ should fulfill the equation below:

$$\frac{\partial S}{\partial \beta_0} \Big|_{\widehat{\beta_0}, \widehat{\beta_1}, \ldots, \widehat{\beta_k}} = -2 \sum_{i=1}^{n} \left(y_i - \widehat{\beta_0} - \sum_{j=1}^{n} \beta_j x_{ij} \right) = 0 \tag{2. 13}$$

and

$$\frac{\partial S}{\partial \beta_j}\Big|_{\widehat{\beta_0},\widehat{\beta_1},\ldots,\widehat{\beta_k}} = -2\sum_{i=1}^{n}\left(y_i - \widehat{\beta_0} - \sum_{j=1}^{n}\beta_j x_{ij}\right)x_{ij} = 0,\ j = 1, 2, \ldots, k \tag{2. 14}$$

After expanding Equation 2.14, The least-squares normal equation can be found :

$$n\hat{\beta}_0 + \hat{\beta}_1\sum_{i=1}^{n} x_{i1} + \hat{\beta}_2\sum_{i=1}^{n} x_{i2} + \cdots + \hat{\beta}_k\sum_{i=1}^{n} x_{ik} = \sum_{i=1}^{n} y_i$$

$$\begin{array}{l}\hat{\beta}_0\sum_{i=1}^{n} x_{i1} + \hat{\beta}_1\sum_{i=1}^{n} x_{i1}^2 + \hat{\beta}_2\sum_{i=1}^{n} x_{i1}x_{i2} + \cdots + \hat{\beta}_k\sum_{i=1}^{n} x_{i1}x_{ik} = \sum_{i=1}^{n} x_{i1}y_i \\ \quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad\qquad \vdots \qquad\qquad \vdots \\ \hat{\beta}_0\sum_{i=1}^{n} x_{ik} + \hat{\beta}_1\sum_{i=1}^{n} x_{ik}\, x_{i1} + \hat{\beta}_2\sum_{i=1}^{n} x_{i1}x_{i2} + \cdots + \hat{\beta}_k\sum_{i=1}^{n} x_{ik}^{\,2} = \sum_{i=1}^{n} x_{ik}y_i\end{array} \tag{2. 15}$$

When the least-square estimators are calculating, it will be significant that there is one more equation from the number of the estimators, in other words, when least-square estimators are $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$, the number of the regression equation is *p*, $p = k + 1$.

The researchers who want to see whole variables compactly generally prefer to use matrix notation to manage multiple regression easier [17]. The matrix representation of the model is shown with Eq. (2.11) below:

$$Y = X\beta + \epsilon$$

Where:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{21} & \ldots & x_{1k} \\ 1 & x_{21} & x_{22} & \ldots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

As understood from the matrix notation; the dimension of y vector called observations is $n \times 1$, the dimension of $\boldsymbol{X}$ called regressor variables is $n \times p$, the dimension of the vector β called regression coefficients is $n \times 1$ and finally, the dimension of the vector ε called random errors is $n \times 1$.

$$S(\beta_0, \beta_1, \dots, \beta_k) = \sum_{i=1}^{n} \varepsilon_i{}^2 = \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j xij \right)^2$$

$$= \varepsilon' \varepsilon = (y - X\beta)'(y - X\beta)$$

If the least-square function *S(β)* improved as below:

$$S(\beta) = y'y - \beta'X'y - y'X\beta + \beta'X'X\beta \qquad (2.16)$$

Due to $\beta'X'y$ and its transpose $y'X\beta$ are scalar or in other words, 1×1 matrix, the least-square function above will be transformed to function below :

$$S(\beta) = y'y - 2\beta'X'y + \beta'X'X\beta \qquad (2.17)$$

Also, the derivation of function according to estimators must equal to zero as below:

$$\frac{\partial S}{\partial \beta}|_{\hat{\beta}} = -2X'y + 2X'X\hat{\beta} = 0 \qquad (2.18)$$

After simplifying the derivation above, the **least-squares normal equations** below will be found as:

$$X'X\hat{\beta} = X'y \qquad (2.19)$$

The least-square estimators ($\hat{\beta}$) can be reached by multiplying both sides of (2.16) by $\left(X'X\right)^{-1}$. Thus, the estimator is:

$$\hat{\beta} = (X'X)^{-1}X'y$$

When the estimator $\hat{\beta}$ put in place in the equation below:

$$\hat{y} = X\hat{\beta} = X(X'X)^{-1}X'y = Hy \qquad (2.20)$$

$H = X(X'X)^{-1}X'$ is generally named hat matrix and its dimension is $n \times n$. The assets of the hat matrix are quite important, and it is also used to calculate the residuals which are the difference between observed and fitted values.

$$e = y - \hat{y}$$

After the H matrix is added to the equation above, the residual will be:

$$e = y - X\widehat{\beta} = y - Hy = (I - H)y \qquad (2.21)$$

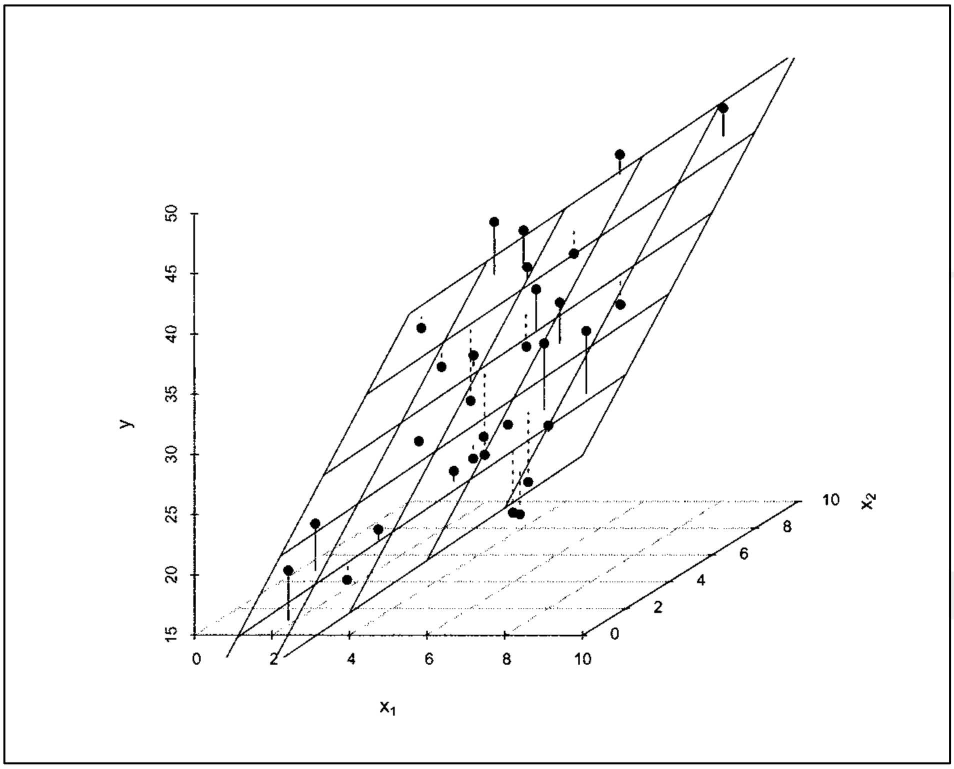

Figure 2. 2 The Least Square Method with two independent variables [19].

The plane in Figure 2.2 indicates the regression model that is obtained by the least square estimation with two independent variables, after minimizing the squares of errors. Grey lines between the plane and dots show the error values. The least Square Estimation method is used not only in linear regression models but also in non-linear and fuzzy regression models. A detailed explanation of the method will be given in the following chapters.

The theoretical part of the least-square method is emphasized above superficially, however, some certain assumptions are obliged to satisfy for the usage of the least-square method. Otherwise, the method will not give appropriate results to analyze data. The assumptions are listed below [19]:

- Firstly, the error terms are considered to have normal distribution, otherwise, other assumptions would be invalid, and the least-square method would be an unnecessary application for the data set wanted to practice.
- The expected value of the error terms is assumed to equal to zero ($E(\varepsilon)=0$). The assumption avoids that the regression line cannot be violated, in other words for a group of observations the model is not systematically low or high.
- The variance of the error terms is stable ($var(\varepsilon)= \sigma^2 \boldsymbol{I}$). It means the variance of the errors cannot be different for some part of the data set. The assumption for error terms that have constant variance is called **homoscedasticity.**
- There cannot be a correlation between each error term. The correlation between errors happens in time series data and the situation called **autocorrelation**.

The violations of the assumption mentioned above lead the model to be misled and they weaken the strength and capacity of estimation. For prevention from the violation, some plots and tests are advised to check the model.

2.3 Circular Regression

The researchers prefer to use not only linear observations but also directional observations to reach a proper and efficient solution according to the structure of the dataset. When the dataset includes directional and circular data, it leads the researchers to use the directional statistics for appropriate prediction. In science, the Directional Statistic generally is used in Meteorology, Biology, Geology, Geophysics, Geography, and Psychology, etc.[20]

2.3.1 Literature Review

When linear statistical methods are applied to directional data causes errors in subjects such as parameter estimation and regression analysis. Therefore, many methods have been developed differently from known statistical techniques. The development of directional data analysis that has its statistical attributes has begun with the study of Gumbel et al. [21]. They examined the theoretical background of circular normal distribution.

Gould [22] proposed the first regression model with circular variables in 1969. The model uses the dependent variable as a circular variable and independent variables are linear variables also the dependent variable has von Mises distribution.

Examining that the probability function proposed by Gould led to erroneous Maximum Likelihood Errors, Johnson and Wehrly proposed models for direct estimation of the mean direction and concentration parameter in the case of a single covariate [23].

Mardia calculated the correlation coefficient for bivariate circular distributions [24]. Lund implemented the least-squares errors method to circular regression [25]. Jammalamadaka and Sarma proposed a regression model with two circular random variables that explain the relationship between these variables and demonstrating the conditional expectation of the given vector [26].

Probability density functions that are defined on the timeline can be wrapped around a unit circle. The application of this concept encouraged the use of time-dependent variables as circular variables. Mardia and Jupp describe the properties of wrapped distributions as quite inclusive in their book [27].

Downs and Mardia (2002) offered a circular regression model based on a one-to-one match between the independent angle and the mean of the dependent angle [28].

Hussin et al. developed Mardia's model [29] by addressing the situation in which both response and explanatory variables are circular.

2.3.2 Circular Descriptive Statistic

When the classic statistical methods that are formed for the linear dataset are applied to the directional dataset causes errors for parameter estimation, regression analysis, etc. [27]. The directional data cannot be mentioned in any size; therefore, the observed value can be described by points over a unit circle whose center is the origin or a vector that joins the points with the origin of the circle. Directional data can be defined with an interval between 0 and 2π. $\boldsymbol{\theta}^{\circ}$ that is used to define vector shows the angle between the vector and positive x-axis with clockwise or counterclockwise. The Cartesian coordinates of the vector are: $\boldsymbol{cos(\theta)}$ $\boldsymbol{and}$ $\boldsymbol{sin(\theta)}$.

The radian and degree units can be used for directional data, the angle in the degree unit is symbolized with $\boldsymbol{\theta}^{\circ}$ and angle in radian unit is symbolized with $\boldsymbol{\theta}$. Transformations from radian to degree unit and from degree to radian unit are shown below:

$$\boldsymbol{\theta}^{\circ} = \boldsymbol{180\,\theta}/\boldsymbol{\pi}\ , \boldsymbol{0^{\circ} < \theta^{\circ} < 360^{\circ}} \quad (2.\ 22)$$

$$\boldsymbol{\theta} = \boldsymbol{\pi\theta}^{\circ}/\boldsymbol{180}\ , \boldsymbol{0 < \theta < 2\pi} \quad (2.\ 23)$$

Also, the relation between Cartesian coordinates and polar coordinates is depicted in Figure 2.3.

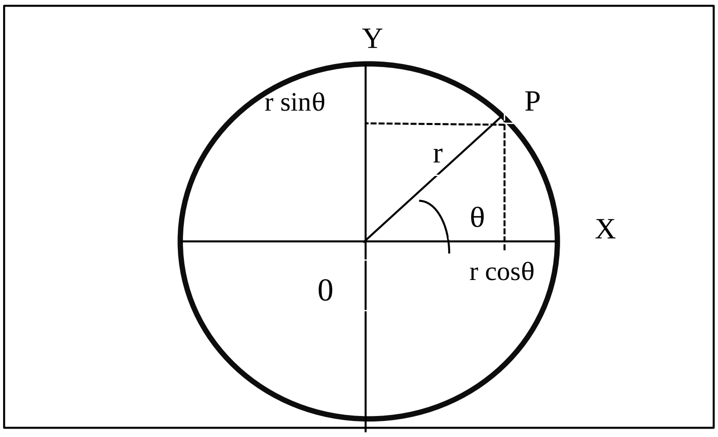

Figure 2. 3 The relation between cartesian coordinates and polar coordinates.

The polar coordinates of the P are defined with radius (r) and angle (θ), the Cartesian coordinates of the point can be reached with trigonometric transformation below [30]:

$$\boldsymbol{x} = \boldsymbol{r}\sin\boldsymbol{\theta} \ \boldsymbol{and}\ \boldsymbol{y} = \mathbf{r}\cos\boldsymbol{\theta} \tag{2. 24}$$

2.3.3 Mean Direction

When directional data is used in research, the mean of the data set cannot be calculated as the arithmetic mean. When the arithmetic mean of 001 and 359 is 180, on the other side, the directional mean of the 001° and 359° will be 360°. The basic example explains the importance of the mean directional. The same difference between linear and directional data can be seen in variance calculations. The directional mean is calculated as a resultant vector of the directions that are accepted as directional vectors. If All P_i are points on a unit circle with an angle $\theta_i (i = 1,2, \dots, n)$, the mean of points $\bar{\theta}$ is the resultant of the unit vectors $\overline{OP_1}, \dots, \overline{OP_n}$.

Where C and S are the components of the resultant vector R,

$$R = \left(\sum_{i=1}^{n} cos\theta_i , \sum_{i=1}^{n} sin\theta_i \right) = (C, S) \tag{2. 25}$$

Subsequently, the mean of the direction is stated as below [27]:

$$\bar{\theta} = \begin{cases} arctan\left(\frac{S}{C}\right), & if\ S \geq 0, C > 0 \\ {}^{\pi}/_{2}, & if\ C = 0, S > 0 \\ arctan\left(\frac{S}{C}\right) + \pi, & if\ C < 0 \\ arctan\left(\frac{S}{C}\right) + 2\pi, & if\ C \geq 0, S < 0 \\ undefined, & if\ C = 0, S = 0 \end{cases}$$

The mod and the median of the directional data can be calculated like the method that is used to calculate the directional mean, nevertheless, the terms are not detailed in the study.

2.3.4 Circular Regression Types

Although regression analysis is a frequently used method in statistics, the usage percent of it with circular data is quite low. Studies about circular regression have begun approximately 50 years ago [31].

A regression method is named circular regression whether it has a circular variable as a dependent or independent variable in a regression model. Comprehensively, in a circular regression when there is/are circular independent variable(s), a linear or circular dependent variable can exist or when there is a circular dependent variable, just linear independent variable(s) can exist. The directional descriptive statistics are practiced in circular regression. The methods and approaches can change according to circular variables used as dependent or independent variables. Three Circular Regression Types are detailed below [32], [33]:

1. Circular – Linear Regression: When one of the independent variables in the model is a circular variable, if the dependent variable is linear, the model is called the Circular-Linear Regression model.
2. Linear – Circular Regression: When independent variable(s) is/are linear variable(s) if the dependent variable is a circular variable, the model is named Linear – Circular Regression model.
3. Circular – Circular Regression: When both independent and dependent variables are circular variables, the model is called the Circular – Circular Regression Model.

The Linear – Circular Regression and Circular- Circular Regression models are not dilated since the models are not used in the study.

2.3.4.1 Circular-Linear Regression

When Y is the linear dependent, the a_1 is the circular independent variable with a single period. The simple regression model for these variables is:

$$\boldsymbol{y = A_0 + A_1 cos\omega(a_1 - a_0)} \quad (2.\ 26)$$

The above model has T period, A_0 is named the mean level, A_1 is named amplitude, ω is named the angular frequency with $2\pi\ /\ T$, and lastly, a_0 is named acrophase that points to the highest peak and is symbolized with φ. So, the model can be written as below [30]:

$$\boldsymbol{When\ \varphi = \omega a_0, Y = A_0 + A_1 cos(\omega a - \varphi)} \quad (2.\ 27)$$

After determination of the constants via the Least Square Method, the model in Equation 2.28 can be generalized as in Equation 2.29 :

$$y = A_0 + A_1 cos(\omega a - \varphi) + A_2 cos(2\omega a - \varphi_2) + \cdots + A_k cos(k\omega a - \varphi_k) \quad (2.\ 28)$$

Researchers report the distortion of the oscillation, in other words, the sinusoidal curve of the variable does not follow peak and trough points with the same period and skewed oscillations occur [30], [34]. To solve the problem, different methods or terms are generated to avoid the mentioned deviations.

Johnson and Wehrly developed a new approach for angular linear distributions and the regression models were relatively generated with these distributions in 1978 [23]. They tried to forecast the Air Quality Index as the dependent variable with the temperature that is a linear independent variable and wind direction that is an angular variable. When x_1 is the air pollution index, x_2 is the temperature and θ is the wind direction, the model offered by Johnson and Wehrly is seen below:

$$\widehat{x_1} = A_0 + A_1 x_2 + A_3 cos\theta + A_4 sin\theta \quad (2.\ 29)$$

3 FUZZY SET THEORY

Firstly, propounded by Zadeh in 1965 [35], the fuzzy set theory is an approximation logic system that uses imprecise or linguistic data or human experiences to compute based on mathematical models, in other words, it allows humanity to utilize fuzzy data for decision-making [36].

Aristotle's logic uses only “0” and only”1” to explain events, it means an event is white or black; the event exists or not exists. Contrarily fuzzy logic uses a membership degree which is a value between “0” and “1” [37].

According to Zadeh, fuzzy logic is calculating with words instead of numbers [38]. For instance, the word “ice cream” represents different tastes and preferences to anyone. Indeed, an expression of a human reflects his or her thoughts which have uncertainty. The Fuzzy logic skillfully breaks the classical two-valued (0 or 1) approach and gives a capability to a thought, an event, or a variable to be denoted by an infinite number of values between 0 and 1.

Fuzzy logic is a revolution against classical logic. It also led to new applications not only in mathematical science but also in engineering. In the first decade after fuzzy logic was proposed, some researchers approached the theory with suspicion. Due to not having an application, the theory stayed as a philosophical debate among scientists [39]. However, the fuzzy set theory and fuzzy systems took the attention of academia, afterward, Mamdani and Assilian applied the fuzzy set theory to a steam engine controller in 1975 [40]. Following years, the fuzzy set theory was practiced in hoovers, washing machines, elevators, metro systems, business administrations, and many fields in the economy and engineering, etc. So far, the fuzzy set theory has continued to boost its popularity.

3.1 Literature Review

The fuzzy set theory gained popularity in Japan in the 1970s, after Zadeh's proposal for fuzzy set theory. And many new objections have been proposed [41].

Bezdek developed the fuzzy clustering method to cluster and analyze plants in a well-known botanical dataset. It is another example of the implementation of the fuzzy set theory to a crisp method [42].

Dubois and Prade presented information about algebraic operations on fuzzy numbers and examined L-R type fuzzy numbers [43] Lowen studied convex fuzzy sets and put forwards the preliminaries of the convexity of the fuzzy sets [44].

Pedrycz worked on fuzzy logic operators to be used in fuzzy membership degrees, essentially incorporating statistical properties [45].

Some issues related to choosing suitable operators for combination and intersection of fuzzy subsets have been criticized by Yager and new operator suggestions were made on the topic [46].

Laarhoven and Pedrycz [47] implemented fuzzy triangular numbers to one of the most popular decision-making methods that are Saaty's [48] Analytic Hierarchy Process. Buckley [49] developed the Fuzzy AHP method by using trapezoidal fuzzy numbers.

Pehlivan and Apaydın analyzed the fuzzy linear programming problem using the simplex method and artificial neural networks approach in their studies and compared the results. As a result of the research, they reported that the artificial neural network approach was quite useful and could be an alternative to the other method examined [50].

Karwowski and Evans emphasized that the application of fuzzy set theory to production management whose sub-areas are new product advancement, facility location, production planning, inventory and stock controlling , cost-benefit analysis would produce effective results [51]. Many studies are contributed to the production management area and over two hundred studies between 1994-2001 are reviewed detailly by Bansal [52].

Bellman and Zadeh firstly dealt with the implementation of the fuzzy set to decision-making methods [53]. After, Baldwin and Guild applied fuzzy set theory to decision-making methods by comparing fuzzy sets in the space [54]. Since fuzzy multi-criteria decision-making methods have emerged, many contributions are made to the issue. The last two decades' studies are examined systematically by Mardani et al. [55]. Also, Kahraman et al. comprehensively reviewed the studies until 2015 [56].

Mamdani and Assilian developed the first fuzzy logic controller and put the fuzzy set theory into practice in 1975 [40]. Hereby, the Fuzzy Inference System established and began to use in the industry. After, Takagi-Sugeno-Kang [57] proposed a system with fuzzy inputs and a crisp output (linear combination of inputs) to be computationally efficient and suitable for working with optimization and adaptive techniques different from the method of Mamdani.

Examples of the application of fuzzy logic in the industry are the control of a cement kiln created by Smith & Co. in Denmark in 1980 and the design of the Sendai metro by the Hitachi company in Japan in 1987. Later, a Japanese government-industry joint activity, LIFE (Laboratory for Industrial Fuzzy Engineering) was established as a consortium of about 50 members [58]. After the industrial applications of the fuzzy set theory were proven, the usage and trial of the fuzzy set-in sub-areas of the industry have begun to attract the attention of industrial companies. Although just fifty years passed after the fuzzy set theory is propounded, the theory is developed quite speedily and implemented to not only social sciences but also engineering and industry. Consequently, it is significant to prove that the importance of the theory and application of fuzzy logic will continue to appear in every area of human life due to existing vagueness in every part of human life [59].

3.2 Fuzzy Sets

The fuzzy set is generally described with its membership functions which sign different belonging degrees to a set or a phenomenon. In classic set theory, a candidate for membership is either an element of a set or not. It can be stated that if the candidate is an element of the set, so it is 1 or if the candidate is not a member of the set so it is 0, like in Aristo Logic. However, the fuzzy set theory transforms the certain membership term in classical logic into a generalized partial membership concept.

An element of a fuzzy set could be a member of another fuzzy set. Consequently, the fuzzy sets are vague and have imprecise boundaries comparing to classic sets [60], [61]. For instance, when it is considered that X is possible temperature value in Celsius; if it is used to define the boiling temperature of the water, it is a crisp or classic set:

A(x)= Boiling Temperature of water (Precise definition: a classic set),

$$\boldsymbol{A} = \{\boldsymbol{x} \in \boldsymbol{X} \mid \boldsymbol{x} \geq \mathbf{100}\}, \tag{3.1}$$

Or detailed formulation is:

$$\mu_A(x) = \begin{cases} 1 \text{ for } x \geq 100 \\ 0 \text{ for } x < 100 \end{cases}, \boldsymbol{x} \in \boldsymbol{X} \tag{3.2}$$

on the other side, if it is used to define hot temperature, it is absolutely a fuzzy set:
$\tilde{A}$(x)= Extremely hot temperature (According to whom? a fuzzy set)

$$\tilde{A} = \{(x,\ \mu_{\tilde{A}}(x)) | \ x \epsilon X,\ \ \mu_{\tilde{A}}(x) \epsilon [0,1] \tag{3.3}$$

3.2.1 Basic Fuzzy Set Operations

Basic fuzzy set operations are intersection, union, complementation, and inclusion similar to the classical set operations, however, the properties of the fuzzy sets like membership functions, etc. must primarily be applied to the classic set operation and all fuzzy sets must be in the same universe as the first axiom in the operations [61]. Owing to different applications of the feature of the fuzzy set theory; Sugeno [62], Yager [63], Dubois, and Prade [64], [65] enhanced their set approaches in the fuzzy set operations. Zadeh's standard operations are described below, but the researchers who want detailed knowledge about fuzzy set operations can see the approaches mentioned above.

Intersection

Fuzzy sets $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ are a subset of X and their membership functions are $\boldsymbol{\mu}_{\tilde{A}}(\boldsymbol{x})$ and $\boldsymbol{\mu}_{\tilde{B}}(\boldsymbol{x})$ respectively, the membership function of the intersection of the two fuzzy sets can be stated as in Equation 3.4:

$$\boldsymbol{\mu_{(\tilde{A}\cap\tilde{B})}(x) = min[\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)]\forall x \in X} \quad (3.4)$$

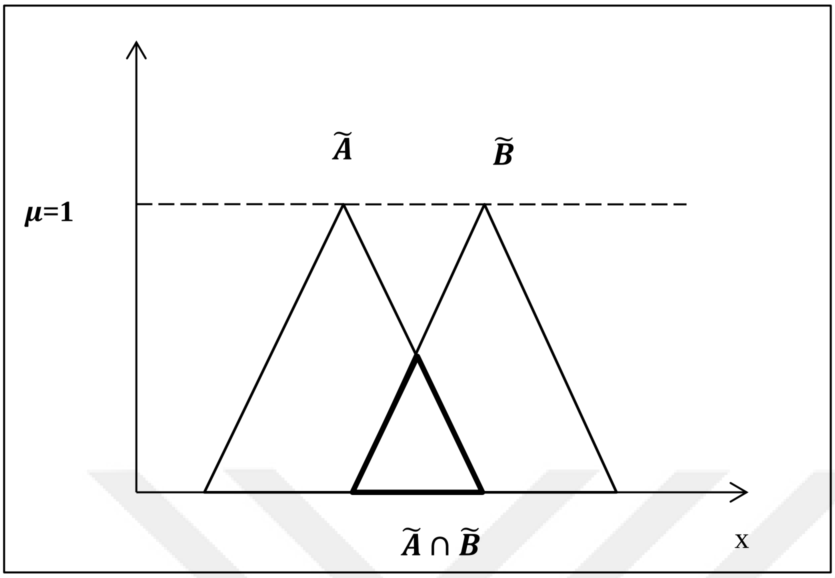

Figure 3. 1 Intersection of the fuzzy sets.

Union

Fuzzy sets $\tilde{A}$ and $\tilde{B}$ are a subset of X and their membership functions are $\mu_{\tilde{A}}(x)$ and $\mu_{\tilde{B}}(x)$ respectively. The union operation is also referred to as MAX-union or standard fuzzy union. The membership function of the union of the two fuzzy sets can be stated as in Equation 3.5:

$$\mu_{(\tilde{A}\cup\tilde{B})}(x) = max[\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)]\forall x \in X \ \mathbf{z} \quad (3.5)$$

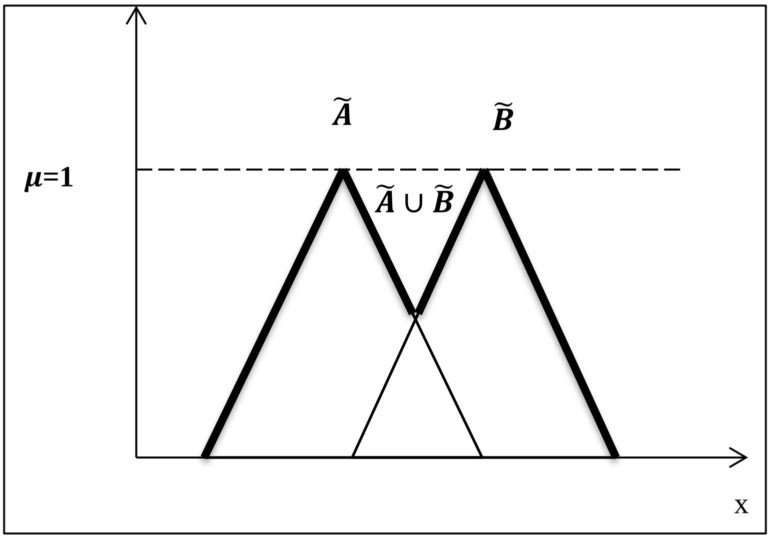

Figure 3. 2 Union of the two fuzzy sets.

Complementation

The complement of a fuzzy set Ã can be symbolized as A^c, $\overline{A}$ or A', the membership functions of fuzzy set Ã and fuzzy set $\tilde{A}^c$ are $\mu_{\tilde{A}}(x)$ and $\mu_{A^c}(x)$ correspondingly and formulated as:

$$\mu_{A^c}(x)=1-\mu_A(x) \quad \forall x \in X \tag{3. 6}$$

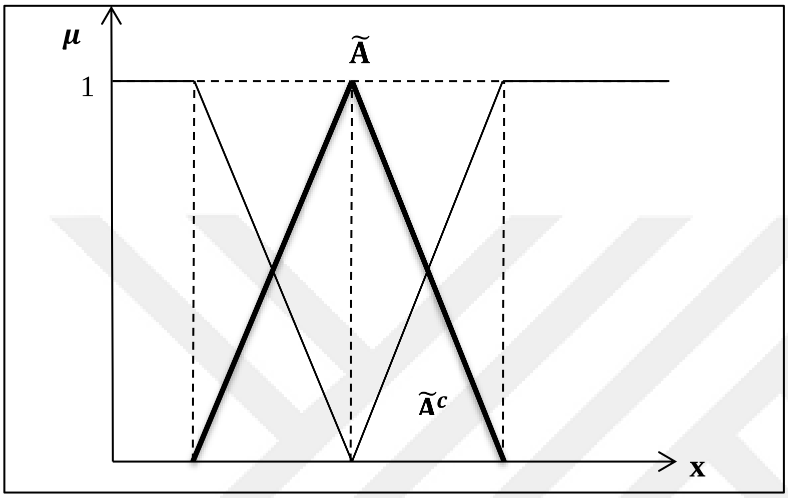

Figure 3. 3 Complementation of Fuzzy set **Ã**.

The bold membership function is the fuzzy set **Ã** and the other membership function is the complementary fuzzy set $\tilde{\mathbf{A}}^{\mathbf{c}}$.

Inclusion (Containment)

If all elements of a fuzzy set $\tilde{A}$ are also elements of a fuzzy set $\tilde{B}$, the fuzzy set $\tilde{A}$ is said that it is included by the fuzzy set $\tilde{B}$. In this case, the fuzzy set $\tilde{A}$ is a subset of the fuzzy set $\tilde{B}$; in other words, the fuzzy set $\tilde{B}$ is the superset of the fuzzy set $\tilde{A}$.

$$\tilde{A} \subseteq \tilde{B} \Leftrightarrow \forall x \in X\left[x \in \tilde{A} \Rightarrow x \in \tilde{B}\right] \quad \text{or} \quad \tilde{B} \supseteq \tilde{A} \tag{3.7}$$

$$\tilde{A} \subseteq \tilde{B} \Leftrightarrow \mu_A(x) \leq \mu_B(x) \; \forall x \in X \tag{3.8}$$

The summary of basic fuzzy set operations is shown in Table 3.1 below [66]:

Table 3.1 Summary of the basic fuzzy set operations.

Description	Notation	Definition
A is a subset of B	$\tilde{B} \supseteq \tilde{A}$	$\mu_{\tilde{A}}(x) \leq \mu_{\tilde{B}}(x)$
A is equal to B	$\tilde{A} = \tilde{B}$	$\mu_{\tilde{A}}(x) = \mu_{\tilde{B}}(x)$
Complement of A	A^c, $\bar{A}$ or A'	$\mu_{A^c}(x)=1-\mu_A(x)$
A intersect B	$\tilde{A} \cap \tilde{B}$	$\mu_{\tilde{A}\cap\tilde{B}}(x) = \min\left[\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\right] \forall x \in X$
A union B	$\tilde{A} \cup \tilde{B}$	$\mu_{\tilde{A}\cup\tilde{B}}(x) = \max\left[\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\right] \forall x \in X$

3.3 Extension Principle

The extension principle permits general mathematical notions and theories used in fuzzy situations. The operation was firstly offered by Zadeh [67], and also known as Zadeh's Extension Principle. The main goal of the principle is to reach the reflection of a fuzzy set $\tilde{A}$ of the universe X, after a function ($f: X \rightarrow Y$) applied to the fuzzy set $\tilde{A}$. The fuzzy set $\tilde{A}$ is defined in X and the fuzzy set $\tilde{B}$ is defined in Y, mathematical presentation of the extension principle is as in Equations 3.9 - 3.12 [68]:

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) | x \in X\} ; \; \tilde{B} = \{(y, \mu_B(y)) | y = f(x_1, x_2, \dots, x_n)\} \quad (3.9)$$

$$\tilde{B} = f(A_1, A_2, \dots, A_n)$$

$$\mu_{\tilde{B}}(y) = \begin{cases} \sup \mu_A(x), if\ f^{-1}(y) \neq \emptyset; when\ x \in f^{-1}(y) \\ 0\ otherwise \end{cases} \quad (3.10)$$

$$\tilde{B} = \int_{X=X_1\times X_2\times\ldots\times X_n} \min\left((\mu_{A_1}(x_1), \mu_{A_2}(x_2), \mu_{A_3}(x_3), \dots, \mu_{A_n}(x_n)/f(x_1, x_2, x_3)\right) \quad (3.11)$$

The detailed definition is:

$$\tilde{A} = \frac{\mu_A(x_1)}{x_1} + \frac{\mu_A(x_2)}{x_2} + \cdots + \frac{\mu_A(x_n)}{x_n} = \sum_{i=1}^{\infty} \frac{\mu_A(x_i)}{x_i} \quad (3.)$$

$$\tilde{B} = f\left(\sum_{i=1}^{\infty} \frac{\mu_A(x_i)}{x_i}\right) = \sum_{i=1}^{\infty} \frac{\mu_A(x_i)}{f(x_i)} \quad (3.12)$$

3.4 Membership Functions

In fuzzy logic, every input has a belonging degree to a fuzzy function. The belonging degree is identified with a membership function. The membership function has a value between [0, 1] and is indicated by $\mu(x)$. For instance, $\mu_{\tilde{A}}(x) = 0.7$ means that membership degree of element x to fuzzy set $\tilde{A}$ is 0.7, in other words, x is a member of the fuzzy set $\tilde{A}$ with a 70 % possibility. According to the statement, an element may not be completely the member of a fuzzy set [60].

3.4.1 Basic Definitions of Fuzzy Membership Functions

When objects signified by x generate the collection of X, a fuzzy set $\tilde{A}$ can be defined as:

$$\tilde{A} = \{(\mathrm{x}, \boldsymbol{\mu}_{\tilde{A}}(x)) | x \epsilon X\}$$

Fuzzy membership functions also have parts; the subset which involves all interval of a fuzzy set is called the Support. The support of the fuzzy set $\tilde{A}$, $S(\tilde{A})$, is the crisp set of all $x \epsilon X$, when $\mu_{\tilde{A}}(x)$>0. Similarly, more than one value in the fuzzy set $\tilde{A}$ can be equal to 1. These values are doubtlessly subset of fuzzy set Ã. These values are generally accepted to stay at the center area of the fuzzy set Ã. Thus, the subsets whose membership degrees are equal to 1 named the Core of the fuzzy setÃ. Correspondingly, other subsets not mentioned which are not equal to 0 and 1 called the boundary of fuzzy set Ã. All parts of the fuzzy set $\tilde{A}$ can be seen in Figure 3.4.

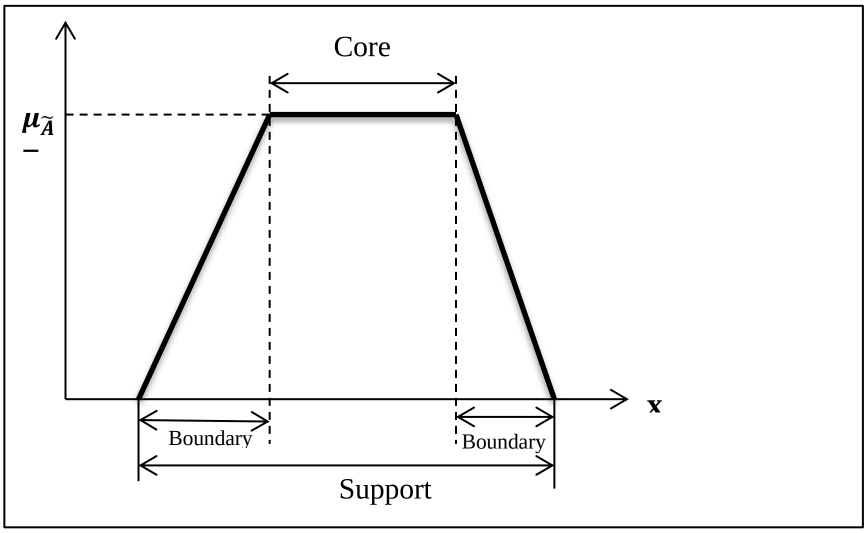

Figure 3. 4 Parts of a Membership Function.

3.4.2 Convexity of Fuzzy Sets

Convexity is also an important attribute for Fuzzy Membership Functions. A fuzzy set $\tilde{A}$ is convex if:

$$\boldsymbol{\mu}_{\tilde{A}}(\lambda x_1 + (1-\lambda)x_2) \geq \min\{\boldsymbol{\mu}_{\tilde{A}}(x_1), \boldsymbol{\mu}_{\tilde{A}}(x_2)\}, x_1, x_2 \in X, \lambda \in [0,1] \quad (3.13)$$

However, the property above is the conclusion of inequalities that are found by basic set operations [35]. For fuzzy sets, A, B, ∧ the basic concept of convex combination is:

$$A \cap B \subset (A, B; \wedge) \subset A \cup B \text{ for all } \wedge . \quad (3.14)$$

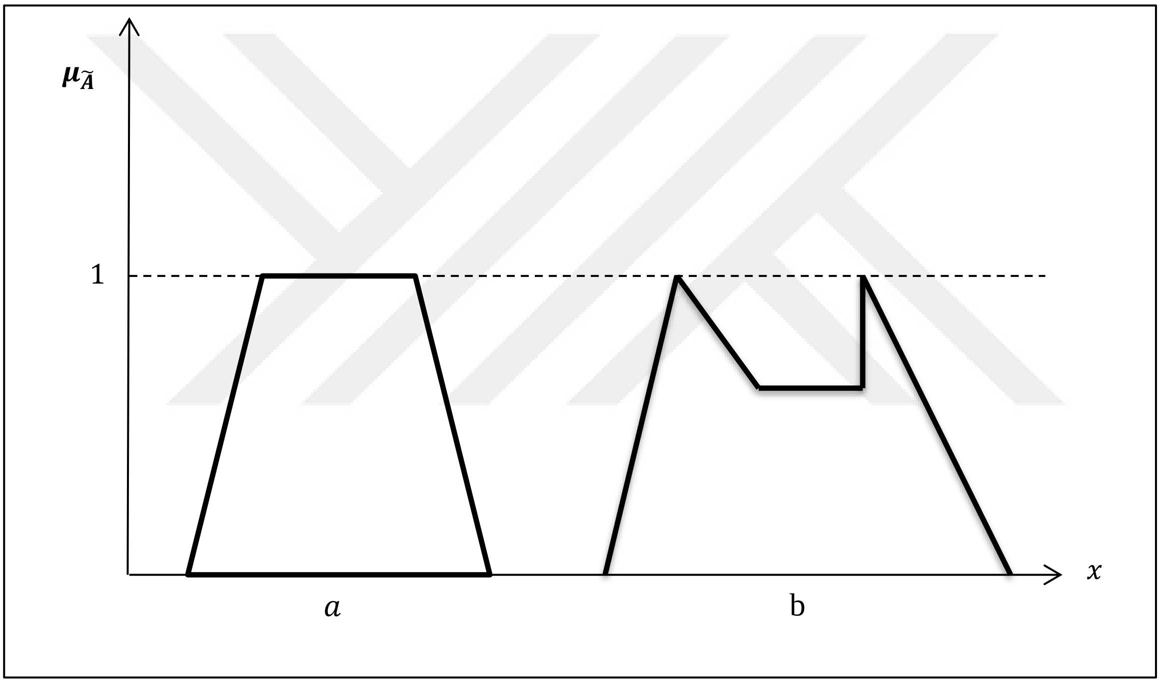

Figure 3. 5 Convex (a) and Nonconvex (b) Fuzzy Sets.

In Figure 3.5, when the shape at left is an example of a convex fuzzy set, the other shape symbolizes a nonconvex fuzzy set.

3.4.3 Normality of Fuzzy Sets

In this case, there is at least one point that equals one; the fuzzy set $\tilde{A}$ is called Normal, otherwise the fuzzy set $\tilde{A}$ is a non-normal fuzzy set. In other words, if the fuzzy set $\tilde{A}$ is normal, the maximum value of $\mu_{\tilde{A}}(x)$ that is occasionally stated as the height of the fuzzy membership function must equal 1 [39]. The examples of Normal and Non-normal Fuzzy Sets are depicted in Figure 3.6.

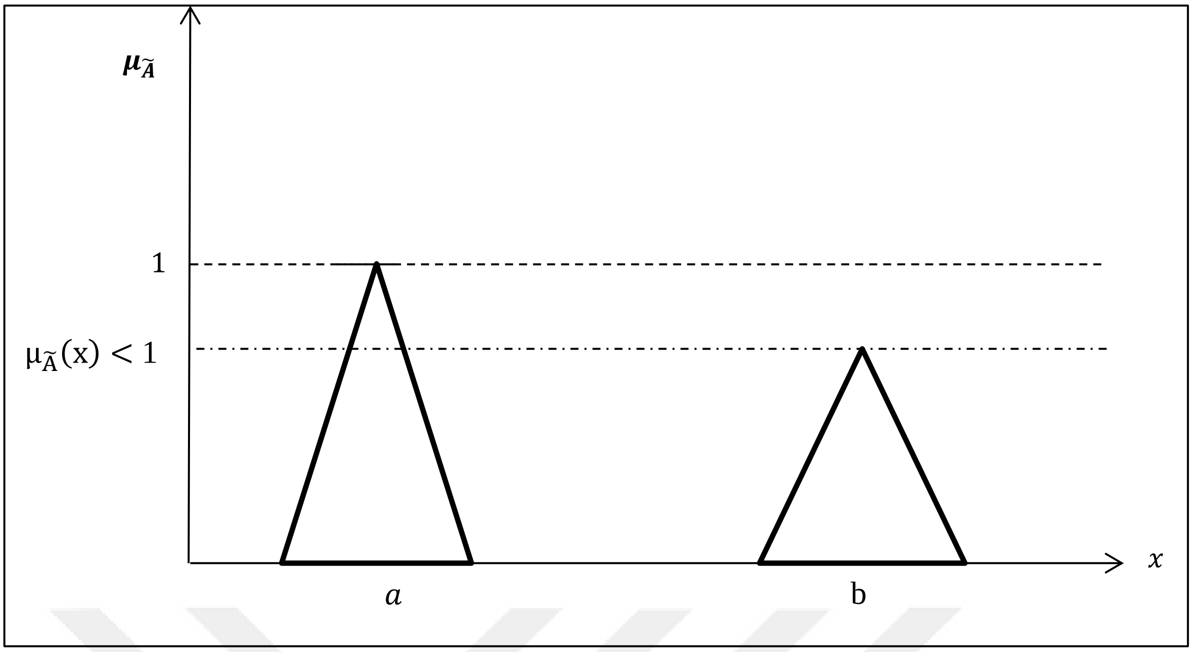

Figure 3. 6 Normal (a) and non-normal (b) Fuzzy Sets.

3.5 Types of Fuzzy Membership Functions

There is various type of membership function has different shapes. In literature, mostly used membership functions are Triangular, Trapezoidal, Gaussian, Sigmoidal membership functions [69].

3.5.1 Triangular Membership Functions

The triangular membership function is defined by three parameters a, b, c which draw the borders of a fuzzy set.

$$\boldsymbol{\mu_{\tilde{A}}(x; a, b, c)} = \begin{cases} \frac{(x-a)}{b-a} & \boldsymbol{if\ a \leq x \leq b} \\ \frac{(c-x)}{c-b} & \boldsymbol{if\ b \leq x \leq c} \\ \boldsymbol{0\ if\ x > c\ or\ x < a} \end{cases} \tag{3.15}$$

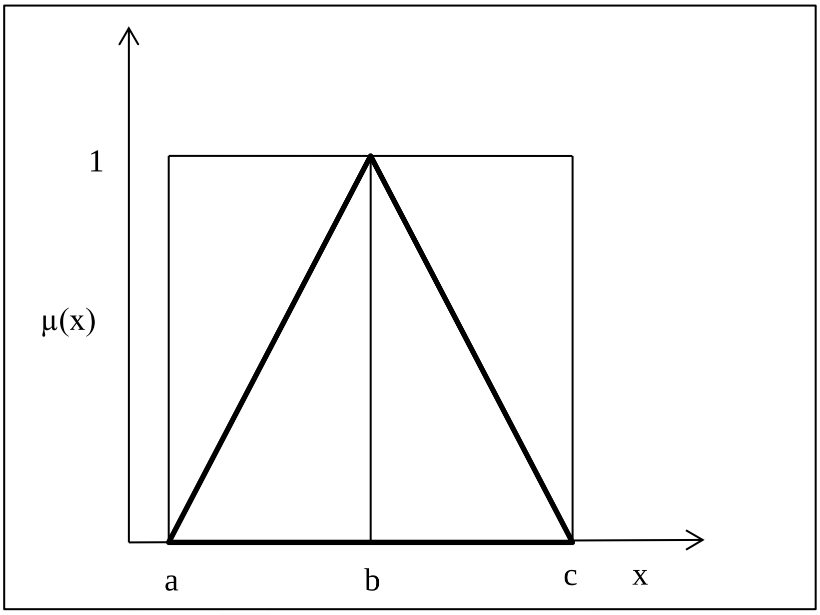

Figure 3. 7 Triangular Fuzzy Numbers

3.5.2 Trapezoidal membership function

The trapezoidal membership function is expressed by four parameters a, b, c, and d, like the triangular fuzzy number in a fuzzy set.

$$\mu_{\tilde{A}}(x; a, b, c, d) = \begin{cases} \frac{(x-a)}{b-a} \quad if\ a \leq x \leq b \\ 1\ \ if\ b \leq x \leq c \\ \frac{(c-x)}{c-b} \quad if\ c \leq x \leq d \\ 0\ if\ x > d\ or\ x < a \end{cases} \tag{3.16}$$

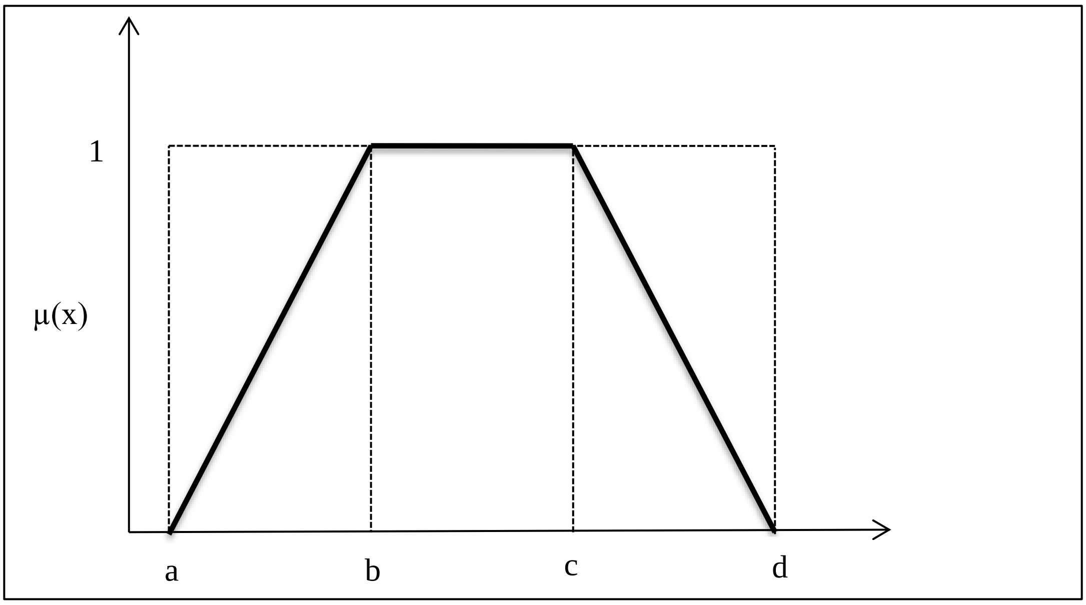

Figure 3. 8 Trapezoidal Fuzzy Numbers.

3.5.3 Gaussian Membership Function

Another fuzzy membership function mentioned before is the Gaussian Membership Function which is generally used to signify impreciseness and fuzzy occurrences. Gaussian Membership Function is defined with c which represents the center and σ which represents the width of fuzzy set $\tilde{A}$ [70].

$$\mu_{\tilde{A}}(x; c, \sigma) = exp\left(-\frac{(c_i - x)^2}{2\sigma^2{}_i}\right) \tag{3.17}$$

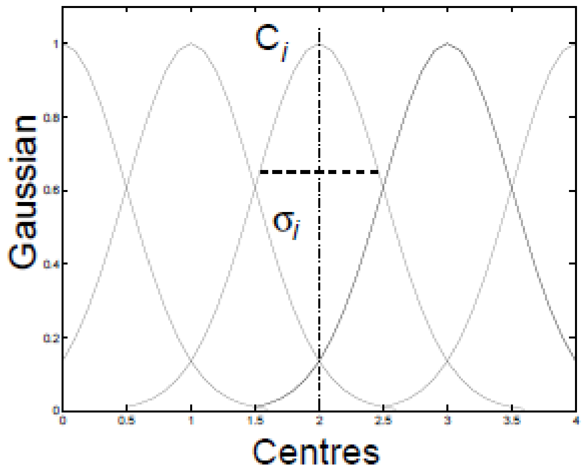

Figure 3. 9 Gaussian Fuzzy Membership Function.

3.5.4 Sigmoidal Membership Function

One of the other frequently used membership functions is the sigmoidal membership function which has two parameters. The first parameter '*c*' the crossover of the S-shaped curve and the second parameter '*a*' is the value of the slope of the S wave. A characteristic of the Sigmoidal MF is having open right and left sides which enable researchers to describe exceptionally large or negative fuzzy conceptions [71].

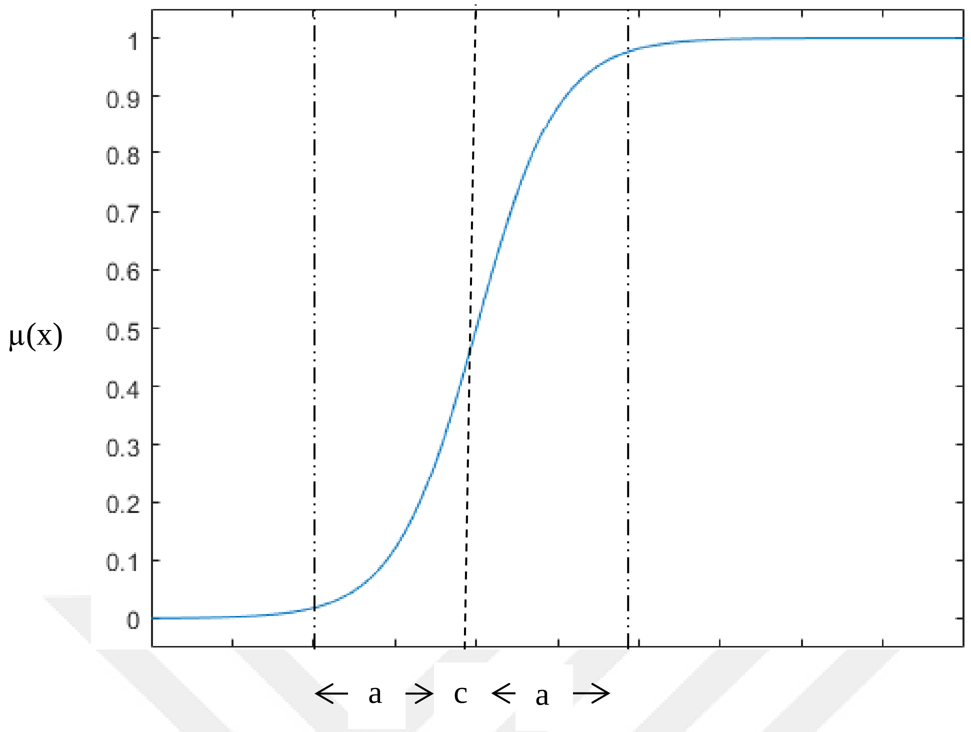

Figure 3. 10 Sigmoidal Fuzzy Membership Functions.

The mathematical formula of Sigmoidal Membership Function is described below:

$$\mu_{\tilde{A}}(x; a, c) = \left(\frac{1}{1+ e^{-a(x-c)}}\right) \tag{3.18}$$

3.6 A α-Levels Concept in Fuzzy Sets

Due to avoid complex calculations with the fuzzy sets, producing less complex arithmetic operations, and using as a defuzzification or ranking method, the α-level concept is generated by leading scientists. The concept is also used to produce a subset of the fuzzy set $\tilde{A}$ whose membership degree is greater than the given α. In other words, the subset is a crisp interval and continuous function [67]. The fuzzy set $\tilde{A}$ is in a set of real numbers and its α-cut is denoted by $\widetilde{A_\alpha}$. Then, $\widetilde{A_a} = \mu_{\tilde{A}}^{-1}[(a, 1)]$, is a crisp set and written as:

$$A_a = \{x \in \mathcal{R} | A(x) \geq a\} \quad (3.19)$$

The strong α-level which is denoted by A_{a^+} is also crisp set and written as:

$$A_{a^+} = \{x \in \mathcal{R} | A(x) > a\} \quad (3.20)$$

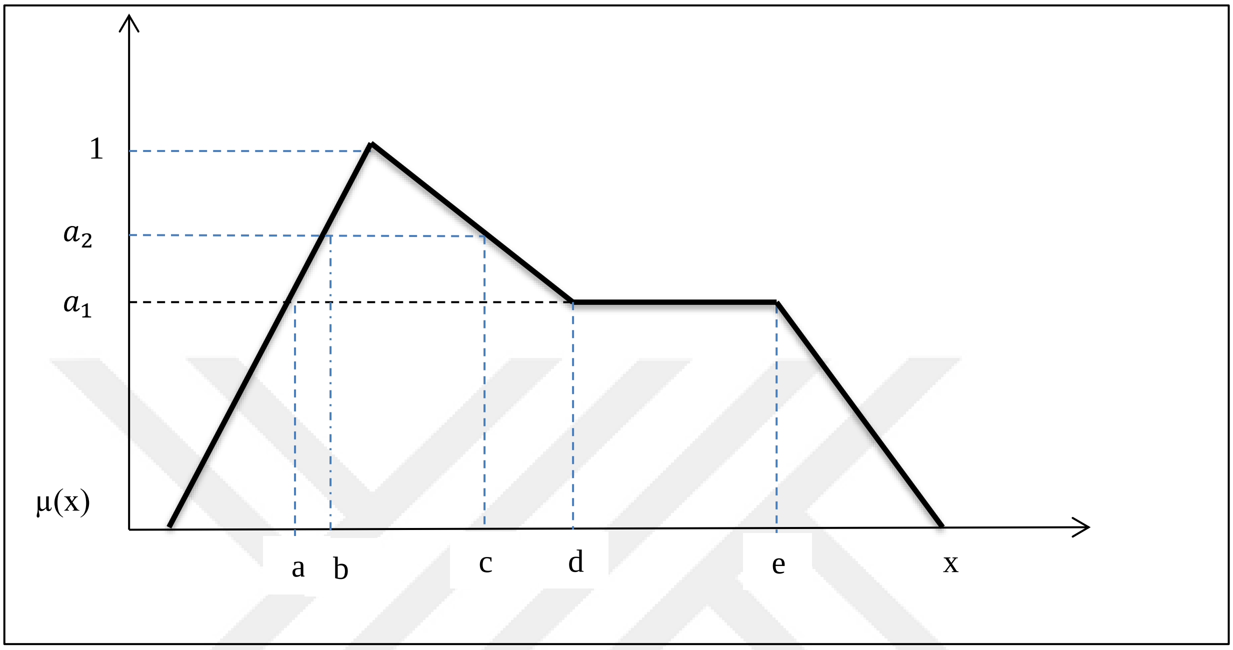

Figure 3. 11 α-levels in a Fuzzy Number.

According to figure 3.11, α-level for α_1 is $\mu^{-1}{}_A(a_1) = [a, e]$,

strong α-level for α_1 is $\mu^{-1}{}_A(a^+{}_1) = [a, d]$ and α-level for α_2 is $\mu^{-1}{}_A(a_2) = [b, c]$.

3.7 Fuzzy Arithmetical Operations

The fuzzy set operators mentioned in the previous section are insufficient for calculations with fuzzy numbers. When the fuzzy numbers are in an equation, the arithmetical operations that are addition, subtraction, multiplication, and division should be activated for solving the problem. Since the equations that include fuzzy numbers are important components of mathematical programming and the other scientific areas for modeling real-life problems. Like in basic fuzzy set operations, there are different approaches for fuzzy arithmetical operations. However, α-cut and max-min convolution methods are defined in the next section [72].

3.7.1 Addition of Fuzzy Numbers

If $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{B}}$ are fuzzy numbers, the addition of the two fuzzy numbers can be calculated by two methods as mentioned:

I. α-level cut Method: The method uses the upper and lower values that are generated after the application of the α-level cut method to the fuzzy numbers. The fuzzy numbers' lower and upper values are $\widetilde{\mathbf{A}}_{\boldsymbol{a}} = [\, \boldsymbol{A}_{\boldsymbol{a}}^{\boldsymbol{L}}, \boldsymbol{A}_{\boldsymbol{a}}^{\boldsymbol{U}}]$ and $\widetilde{\mathbf{B}}_{\boldsymbol{a}} = [\, \boldsymbol{B}_{\boldsymbol{a}}^{\boldsymbol{L}}, \boldsymbol{B}_{\boldsymbol{a}}^{\boldsymbol{U}}]$. If the fuzzy number $\widetilde{\boldsymbol{C}}$ is the summation of the two fuzzy number, the fuzzy number $\widetilde{\boldsymbol{C}}$:

$$\widetilde{\boldsymbol{C}_{\boldsymbol{a}}} = \widetilde{\mathbf{A}}_{\boldsymbol{a}} + \widetilde{\mathbf{B}}_{\boldsymbol{a}} = [\boldsymbol{A}_{\boldsymbol{a}}^{\boldsymbol{L}} + \boldsymbol{B}_{\boldsymbol{a}}^{\boldsymbol{L}}, \boldsymbol{A}_{\boldsymbol{a}}^{\boldsymbol{U}} + \boldsymbol{B}_{\boldsymbol{a}}^{\boldsymbol{U}}] \qquad \text{for every } \alpha \in [0,1] \qquad (3.21)$$

As a result, a lower value of $\widetilde{\boldsymbol{C}}$ is visibly summation of the lower values of the fuzzy numbers and the upper value of $\widetilde{\boldsymbol{C}}$ is similarly the summation of the fuzzy numbers.

II. Max-Min Convolution: The method especially uses Zadeh's extension principle. If the fuzzy number $\widetilde{\boldsymbol{C}}$ is the summation of the two fuzzy number, the fuzzy number $\widetilde{\boldsymbol{C}}$:

$$\boldsymbol{\mu}_{\widetilde{C}}(\boldsymbol{z}) = \boldsymbol{max}_{\boldsymbol{z}=\boldsymbol{x}+\boldsymbol{y}}\, \{\boldsymbol{min}[\boldsymbol{\mu}_{\widetilde{A}}(\boldsymbol{x}) + \boldsymbol{\mu}_{\widetilde{B}}(\boldsymbol{y})]\} \qquad (3.22)$$

where x, y, and $z \in R$

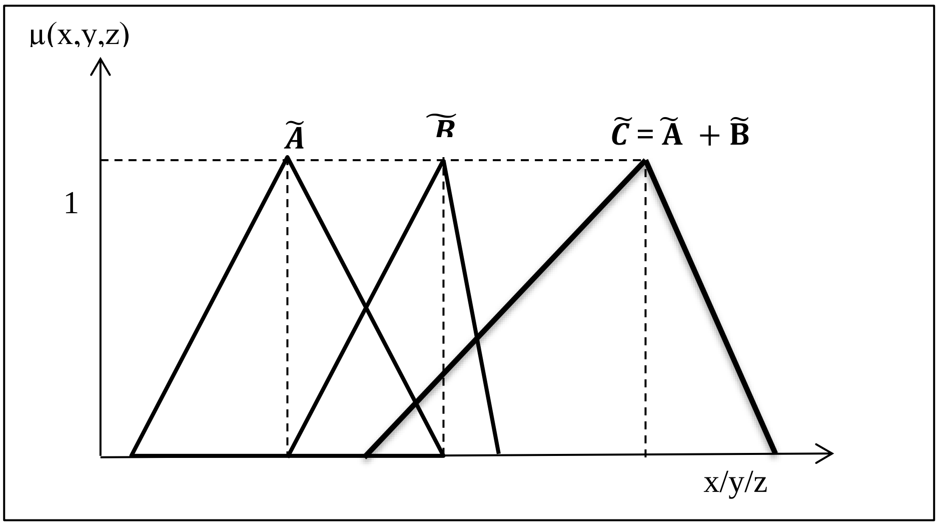

Figure 3. 12 Addition of Two Fuzzy Number in Max-Min Convolution Method.

3.7.2 Subtraction of Fuzzy Numbers

The two methods practiced besides operation are also applied to subtraction operations. The defined fuzzy numbers are used similarly:

I. α-level cut Method: The fuzzy numbers' lower and upper values are $\widetilde{\mathbf{A}}_{\boldsymbol{a}} = [\boldsymbol{A}_{\boldsymbol{a}}^{\boldsymbol{L}}, \boldsymbol{A}_{\boldsymbol{a}}^{\boldsymbol{U}}]$ and $\widetilde{\mathbf{B}}_{\boldsymbol{a}} = [\boldsymbol{B}_{\boldsymbol{a}}^{\boldsymbol{L}}, \boldsymbol{B}_{\boldsymbol{a}}^{\boldsymbol{U}}]$ again. If the fuzzy number $\widetilde{\boldsymbol{C}}$ is the subtraction of the two fuzzy number, the fuzzy number $\widetilde{\boldsymbol{C}}$:

$$\widetilde{\boldsymbol{C}_{\boldsymbol{a}}} = \widetilde{\mathbf{A}}_{\boldsymbol{a}} - \widetilde{\mathbf{B}}_{\boldsymbol{a}} = [\boldsymbol{A}_{\boldsymbol{a}}^{\boldsymbol{L}} - \boldsymbol{B}_{\boldsymbol{a}}^{\boldsymbol{U}},\ \boldsymbol{A}_{\boldsymbol{a}}^{\boldsymbol{U}} - \boldsymbol{B}_{\boldsymbol{a}}^{\boldsymbol{L}}] \quad \text{for every } \alpha \in [0,1] \tag{3.23}$$

II. Max-Min Convolution: The method uses the extension principle too. If the fuzzy number $\widetilde{\boldsymbol{C}}$ is the subtraction of the two fuzzy number, the fuzzy number $\widetilde{\boldsymbol{C}}$:

$$\boldsymbol{\mu}_{\widetilde{\boldsymbol{C}}}(\boldsymbol{z}) = \boldsymbol{max}_{\boldsymbol{z}=\boldsymbol{x}-\boldsymbol{y}}\ \{\boldsymbol{min}[\boldsymbol{\mu}_{\widetilde{\boldsymbol{A}}}(\boldsymbol{x}),\ \boldsymbol{\mu}_{\widetilde{\boldsymbol{B}}}(\boldsymbol{y})]\}$$

$$\boldsymbol{\mu}_{\widetilde{\boldsymbol{C}}}(\boldsymbol{z}) = \boldsymbol{max}_{\boldsymbol{z}=\boldsymbol{x}+\boldsymbol{y}}\ \{\boldsymbol{min}[\boldsymbol{\mu}_{\widetilde{\boldsymbol{A}}}(\boldsymbol{x}), \boldsymbol{\mu}_{\widetilde{\boldsymbol{B}}}(-\boldsymbol{y})]\} \tag{3.24}$$

$$\boldsymbol{\mu}_{\widetilde{\boldsymbol{C}}}(\boldsymbol{z}) = \boldsymbol{max}_{\boldsymbol{z}=\boldsymbol{x}+\boldsymbol{y}}\ \{\boldsymbol{min}[\boldsymbol{\mu}_{\widetilde{\boldsymbol{A}}}(\boldsymbol{x}), \boldsymbol{\mu}_{\widetilde{-\boldsymbol{B}}}(\boldsymbol{y})]\}$$

where x, y, and z $\in$ R.

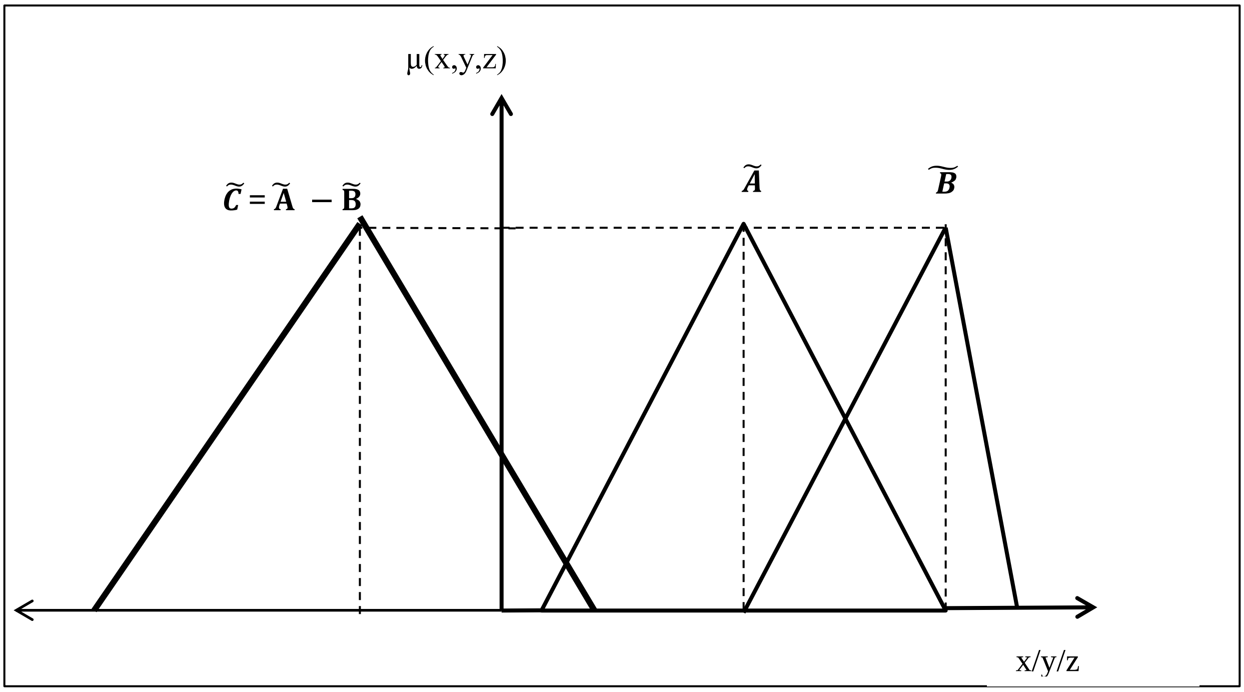

Figure 3. 13 Subtraction of Two Fuzzy Number in Max-Min Convolution Method.

3.7.3 Multiplication of Fuzzy Numbers

There is an assumption to avoid sign effect of the operation, it is $\boldsymbol{\mu_{\tilde{A}}(x) = 0\ for\ x < 0}$ **and** $\boldsymbol{\mu_{\tilde{B}}(y) = 0\ for\ y < 0}$.

I. α-level cut Method: The fuzzy numbers' lower and upper values are $\tilde{\mathbf{A}}_{\boldsymbol{a}} = [\, \boldsymbol{A_a^L}, \boldsymbol{A_a^U}]$ and $\tilde{\mathbf{B}}_{\boldsymbol{a}} = [\, \boldsymbol{B_a^L}, \boldsymbol{B_a^U}]$ again. The fuzzy number $\tilde{\boldsymbol{C}}$ is the multiplication of the fuzzy number$\tilde{\boldsymbol{A}}$ and the fuzzy number$\tilde{\boldsymbol{B}}$, the multiplication is:

$$\widetilde{\boldsymbol{C_a}} = \tilde{\mathbf{A}}_{\boldsymbol{a}} \cdot \tilde{\mathbf{B}}_{\boldsymbol{a}} = [\boldsymbol{A_a^L} \cdot \boldsymbol{B_a^L}, \boldsymbol{A_a^U} \cdot \boldsymbol{B_a^U}] \text{for every } \alpha\ \boldsymbol{\epsilon}\ [0,1] \tag{3.25}$$

II. Max-Min Convolution: The method also uses the extension principle; however, the principle makes the multiplication operation more complicated. So, Kaufmann and Gupta suggest a new procedure [73]. The procedure is defined as:

1. If the fuzzy number is normal, the point where membership value equals 1 is found, if else the fuzzy number is not normal, the peak value is found. Next, the peak value of the fuzzy number $\tilde{\boldsymbol{C}}$ is determined after the left and right sides of the fuzzy number are defined.
2. The left side of the fuzzy number $\tilde{\boldsymbol{C}}$:

$$\boldsymbol{\mu_{\tilde{c}}(z) = \max_{xy \le z} \ \{\min[\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(y)]\}} \tag{3.26}$$

3. The right side of the fuzzy number $\tilde{\boldsymbol{C}}$:

$$\boldsymbol{\mu_{\tilde{c}}(z) = \max_{xy \ge z} \ \{\min[\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(y)]\}} \tag{3.27}$$

3.7.4 Division of Fuzzy Numbers

The same approaches in multiplication operation can be applied to division operation. Therefore, the division operation is defined as follows:

I. α-level cut Method: The fuzzy number $\tilde{C}$ is the multiplication of the fuzzy numbers and the result is:

$$\widetilde{C_a} = \widetilde{A}_a \div \widetilde{B}_a = [A_a^L \div B_a^U, A_a^U \div B_a^L] \qquad \text{for every } \alpha \in [0,1] \tag{3.28}$$

II. Max-Min Convolution: The procedure used in multiplication operation is also applied to division operation. Firstly, the peak value of the fuzzy number $\widetilde{C}$ is determined and the left side of the division is defined as:

$$\mu_{\tilde{C}}(z) = \max_{\frac{x}{y} \le z} \{\min[\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(y)]\} \tag{3.29}$$

$$\mu_{\tilde{C}}(z) = \max_{xy \le z} \left\{\min\left[\mu_{\tilde{A}}(x), \mu_{\tilde{B}}\left(\frac{1}{y}\right)\right]\right\} \tag{3.30}$$

The right side of the fuzzy number $\widetilde{C}$ is calculated as:

$$\mu_{\tilde{C}}(z) = \max_{\frac{x}{y} \ge z} \{\min[\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(y)]\} \tag{3.31}$$

$$\mu_{\tilde{C}}(z) = \max_{xy \ge z} \left\{\min\left[\mu_{\tilde{A}}(x), \mu_{\tilde{B}}\left(\frac{1}{y}\right)\right]\right\} \tag{3.32}$$

The division operation can be obtained by continuing as the multiplication operation from these equations.

3.8 L-R Representation of Fuzzy Set

Arithmetic calculations between fuzzy numbers cannot be easy as expected. Thus, a new and easier definition not only for academicians and but also for computers is needed. Moreover, Due to reach the limit of effectiveness when a vague real-world model is explained by computers, fuzzy numbers are preferred to depict with its left and right parameters. The main theme in the L-R form of fuzzy number is first determining the center of the fuzzy number after diving it into two pieces called to left and right sides [69].

L-R representation of the fuzzy number M is shown in formula 3.5 below.

$$\mu_M(x) = \begin{cases} L\left(\frac{m-x}{a}\right), a > 0, x \le m; \\ R\left(\frac{x-m}{\beta}\right), \beta > 0, x \ge m \end{cases} \tag{3. 33}$$

In formula 3.7, L (for Left) and R (for Right) are the reference functions of the fuzzy number $\widetilde{M}$. α and β are scalar that called left and right spread of the fuzzy number $\widetilde{M}$ respectively and they are bigger than zero, also m denotes the mean of the fuzzy number $\widetilde{M}$. When the spreads (α and β) of the fuzzy number $\widetilde{M}$are equal to zero, which means that M is a crisp number. One of the most important advantages of the L-R representation of a fuzzy membership function is to define the left and right parts of the function individually [74]. In other words, two distinct functions can be combined in a fuzzy membership function. The other advantage of the representation is straightforward computational calculations in a complex model, in case that fuzzy membership functions are symmetric.

Mathematical notation of the fuzzy number $\widetilde{M}$ is:

$$\widetilde{M} = (m, a, \beta)_{LR} \tag{3.34}$$

An illustration of the fuzzy number $\widetilde{M}$ is given in figure 3.14 below.

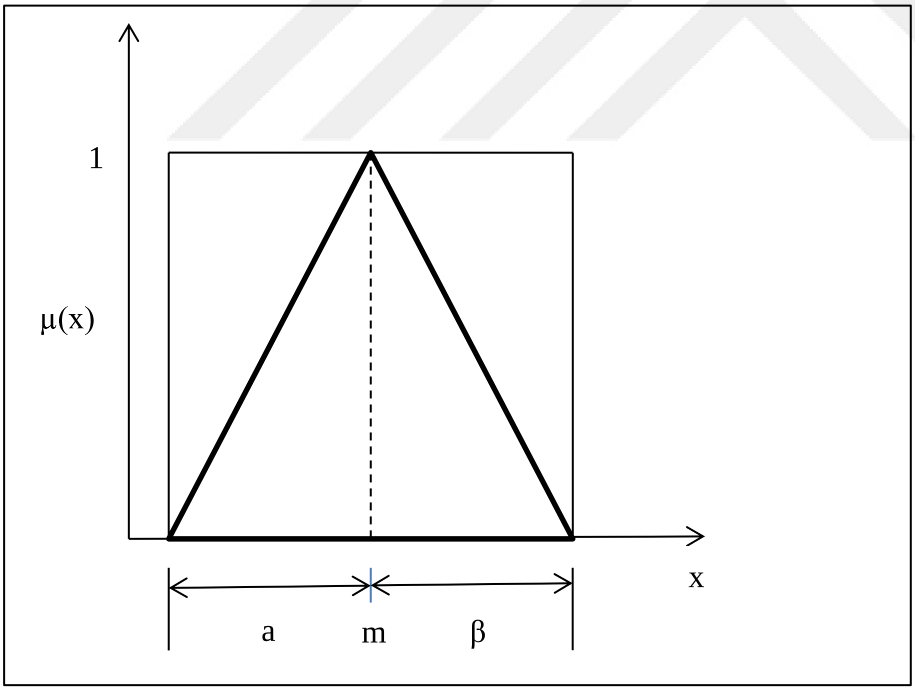

Figure 3. 14 L-R Representation of Fuzzy Sets

4 FUZZY REGRESSION

After the dissertation and improvement of the fuzzy set theory approach by Zadeh [35], it allowed researchers to assess vague variables as numeric variables. In 1982, Tanaka et al. developed the Linear Regression Analysis with a fuzzy model by taking advantage of the fuzzy set theory [75]. Every qualitative variable or observation is injected into the model with its fuzzy membership degrees.

Although the crisp regression models are quite prevailing to detect the relationship between variables, and huge varieties of crisp models are generated by the researcher; the models cannot determine the relationship between variables as expected due to some reasons and obscures. The situations led researchers to implement fuzzy set theory to crisp regression models and improve the fuzzy regression models. For instance, if the number of observations is considerable small, the type of the distribution of the observations are not perceived, the distribution of the errors is not normal, the relationships between dependent and independent variables are ambiguous or the structure of the data set is spoiled after linearization process; the fuzzy regression methods are advised to use for modeling dataset [76].

The distribution of the error terms is disregarded in fuzzy regression models. When the errors in crisp regression models are the result of mismeasurement or selecting not appropriate model, the errors in the fuzzy regression model are the result of the fuzziness of the model parameters. Thus, the error in the fuzzy regression model is equal to the total spread of the fuzzy parameters. When a small number of observations are used in a crisp regression model, the distribution of the observation cannot be determined, and the main assumptions of the crisp regression cannot be provided. Moreover, the reliability of the model with a small number of observations would not be at the desired level.

Lots of researchers have tried to enhance the fuzzy regression model by criticizing the prior models. Thus, many fuzzy regression models have been generated. Subsequently, the classification of fuzzy regression models has been complicated. Chukhrova and Johannssen classified the fuzzy regression models with more updated and inclusive techniques. They also added fuzzy application of Machine Learning Techniques, which are currently popular, to their classification method, as is seen in Figure 4.1 below [77]:

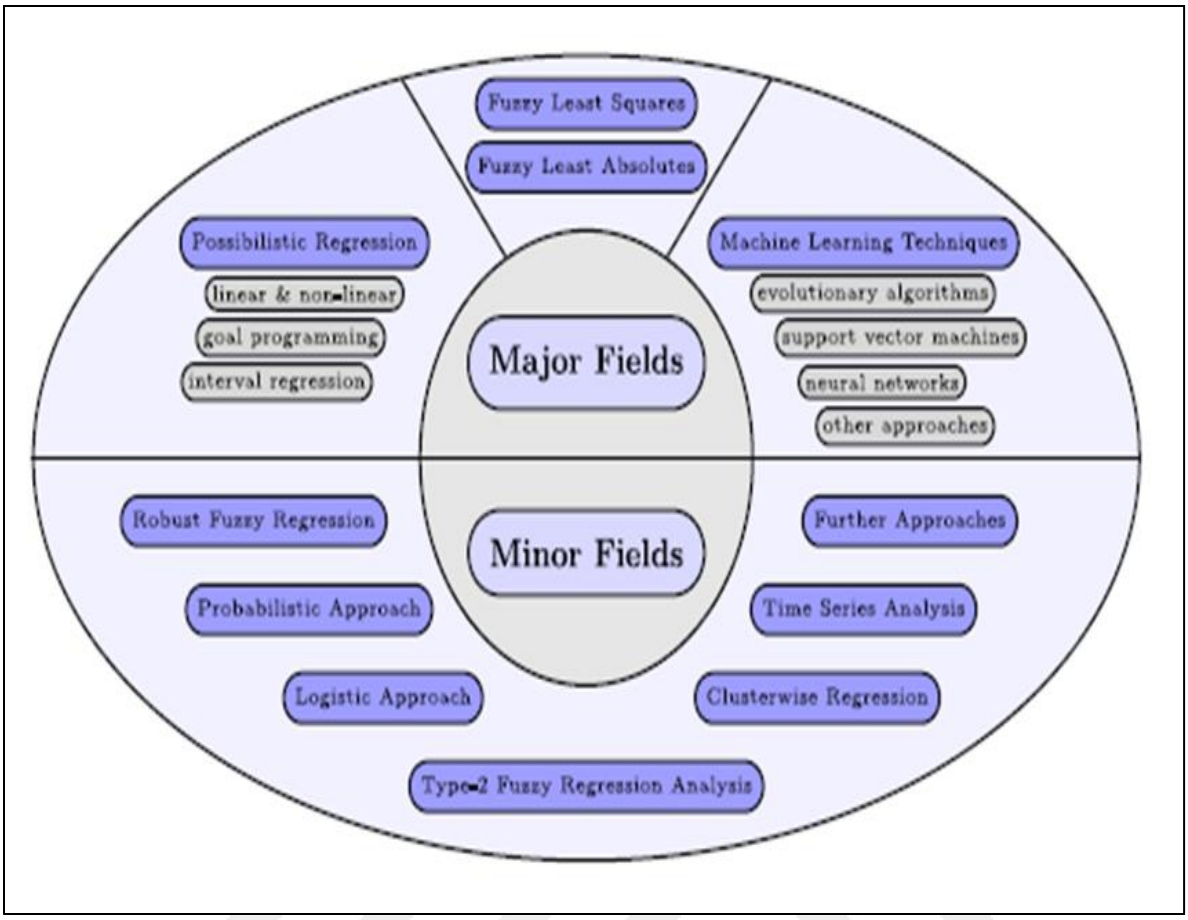

Figure 4. 1 Detailed Classification of Fuzzy Regression Method [77].

Although Chukhrova and Johannssen's study is contemporary and comprehensive, it gives general facts about fuzzy regression models to researchers.

As explained above, the classification concept is applied just for fuzzy linear regression models, so the generated scheme is shown in Figure 4.2.

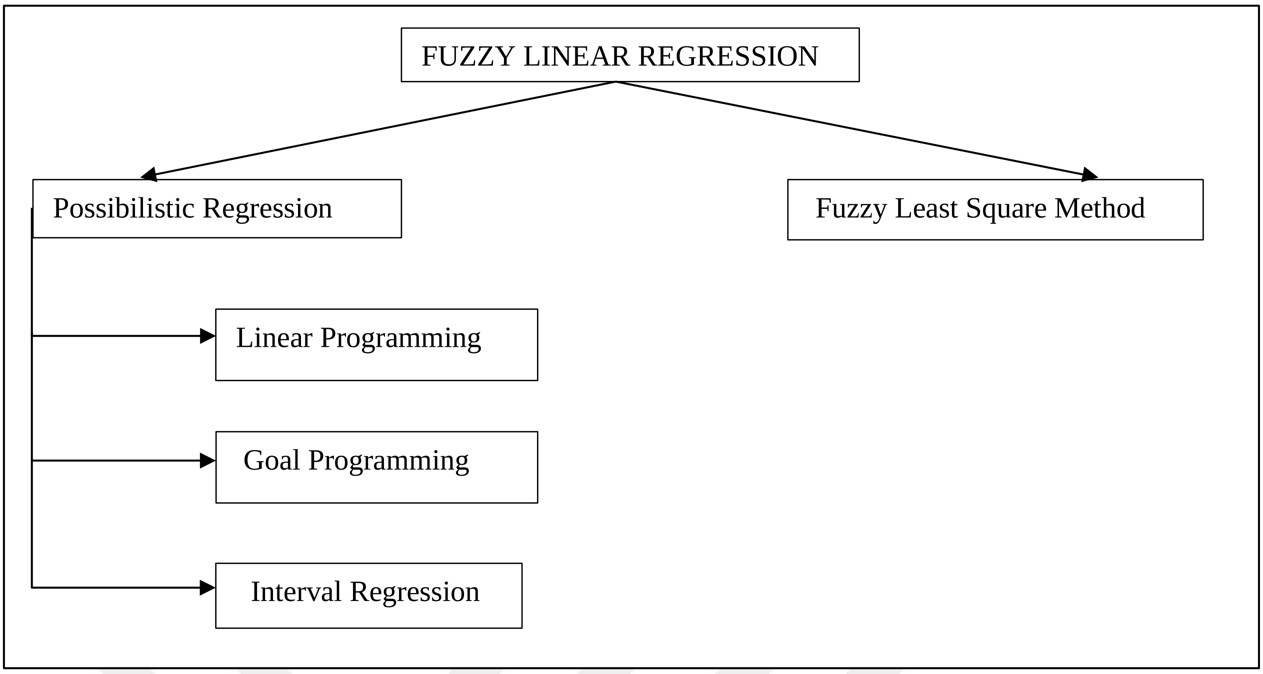

Figure 4. 2 Classification of the FLR Models.

It is also possible to classify the fuzzy regression models according to the type of independent or dependent variables. So, there are three types of fuzzy regression models with diverse types of input and output. The classes are:

- crisp input and crisp output (CICO),
- crisp input and fuzzy output (CIFO),
- fuzzy input and fuzzy output (FIFO).

In the study, the classification method according to inputs and outputs is not detailed, however, when the fuzzy regression models are mentioned, the input and output details will be given in section 6.

4.1 Literature Review

The first fuzzified linear regression analysis is suggested by Tanaka et al. [75]. It is assumed that the input and output variables are crisp numbers, but the system parameters are fuzzy, and the objective function is based on the minimization of the spread of the predicted value of the dependent variable.

Since the appearance of fuzzy regression, scientists have made many contributions to the method. Thus, fuzzy regression has been developed rapidly until today. Tanaka [78] proposed the possibilistic linear model for processing fuzzy data. Diamond [79] developed the fuzzy least-squares method in which the independent variable is definite, and the dependent variable is a triangular fuzzy number.

Moskowitz and Kim [80] determined the relationship between the spreads of fuzzy parameters, the forms of membership functions, and the *h* value in fuzzy linear regression.

Peters [81] introduced a new linear fuzzy regression model to address the shortcomings of the method proposed by Tanaka et al [75]. The method is referred to as interval regression analysis in the literature.

Kim et al. [82] and Kim and Chen [83] compared fuzzy linear regression with crisp non-parametric regression methods and reported that fuzzy regression analysis could be preferred to classical regression analysis when working with a small data set in their study.

Ishibuchi and Nii [84] brought a new perspective to fuzzy regression analysis by using asymmetric triangular and trapezoidal fuzzy coefficients in their studies.

Chang [85] proposed a hybrid method using weighted fuzzy arithmetic based on the fuzzy least-squares method that is capable of adapting to diverse types of data.

D'urso and Gastaldi [86] worked on a linear sub-model called "Doubly linear adaptive fuzzy regression model" based on the fuzzy regression model. They expressed the interaction between center and fuzzy spreads, also discussed the model with numerical estimates.

Because linear programming and least squares-based approaches are overly sensitive to outlines in fuzzy regression, more robust methods are needed. The least absolute deviation method based on the medians has been developed by Dielman [87]. The pioneers of this method are Chang and Lee [88] and Kim et al. [89].

Lee and Chen [90] presented a generalized fuzzy linear regression model, and they proposed a nonlinear programming model to determine fuzzy parameters.

Nasrabadi et al. criticized that fuzzy regression models are sensitive to outliers; it is not effective to predict parameters of the whole dataset. They also mentioned that the predicted values are distributed more widely when the dataset has a large number of observations in the model. In order to eliminate these deficiencies, a multi-objective fuzzy linear regression model has been developed [91].

Watada was the first to suggest methods that combined time series analysis with fuzzy regression analysis. Watada used the intersection concept of fuzzy numbers in the fuzzy time series model in this study [92].

The chronological development and various usage of the fuzzy regression at similar subjects are enlightened so far. Besides the theoretical contributions, there are diverse applications of fuzzy regression in primary areas. Examples of the applications can be seen in the automotive industry, business administration, economics, engineering, energy research, finance, hydrology, information technology, insurance, manufacturing, etc.[77].

4.2 The Components of The Fuzzy Regression Models

It is assumed that the difference between observed and predicted values in classical regression analysis is due to observational errors or incorrect selection of the model. In fuzzy regression, it is considered that the difference between the observed and predicted values naturally arises from the uncertainty or fuzziness of the system structure. The output variable that is defined for the specified inputs in the system structure has a possible value within a specified range and the output can take any value within this range. Fuzzy functions are expressed with the fuzzy coefficients in fuzzy regression models [93]. Tanaka's general fuzzy linear regression method is shown below to explain the main components of the model [75]:

$$\widetilde{Y_i} = \widetilde{A_o} + \widetilde{A_1}X_1 + \cdots + \widetilde{A_p}X_{ip} \qquad (4.\ 1)$$

In Equation 4.1, the parameter $\widetilde{A_j}$ is a fuzzy number and j signifies the parameter that belongs to which independent variable. The membership functions which belong to independent variables can be generated in different forms. In the condition of preferring symmetric triangular fuzzy numbers for the membership functions, the model parameters would be $\widetilde{A_j} = (a_j, c_j), j = 1, \dots, p; a = (a_0, a_1, \dots, a_p)\ and\ c = (c_0, c_1, \dots, c_p)$. Here, a_j designates the center of the fuzzy number and c_j designates left and right spread from the center. So, the model above would be transformed to the equation below:

$$\widetilde{Y_i} = (a_0, c_0) + (a_1, c_1)X_1 + \cdots + (a_p, c_p)X_i p \tag{4. 2}$$

Moreover, the membership function of the expected values of the $\widetilde{Y_j}$ is defined as below with help of the Zadeh's extension principle [94]:

$$\mu_{(Y_j)}(y) = \begin{cases} 1 - \dfrac{\left|y_j - \sum_{j=1}^{p} a_j x_j\right|}{\sum_{j_1}^{p} c_j \left|x_j\right|} \\ 1, x = 0, \quad y = 0 \\ 0, \quad x = 0, y \neq 0 \end{cases} \tag{4. 3}$$

Under circumstances $\widetilde{Y_j}$ is a symmetric triangular fuzzy number, it can be defined with center y and spread e_j, the membership function is:

$$\mu_{(Y_j)}(y) = 1 - \frac{|y_j - y|}{e_j} \tag{4. 4}$$

4.2.1 "*h*" Value in The Fuzzy Linear Regression Models

The "*h*" value refers to the degree of compliance of the fuzzy outcomes that are estimated in the fuzzy linear regression analysis according to the observed values of the dependent variable, in other words, it defines the desired reliability level by obtaining the width or narrowness of the fuzzy spread of parameters in fuzzy regression model [93].

The error term (ε) is used to define randomness in the classical linear regression model. When there is ε value in every observation, contrary to the classic model, the error is distributed over all in fuzzy linear regression models' coefficients. In this case, each parameter is estimated at a certain fuzzy level. This fuzzy level is called "*h term*" or *"h value"* and takes a value in the range between zero and one [0, 1] [80].

In literature, many researchers have conducted several studies to suggest what *h* value should be. Tanaka and Watada reported that values of *h* can change according to the size of the data set. If the data set is large enough, the *h* term should be "0.0"; otherwise, researchers should increase the *h* term [95].

Moskowitz and Kim aimed to determine the relationship between the "*h value*", the spread of fuzzy parameters, and the shape of membership functions in their study [80].

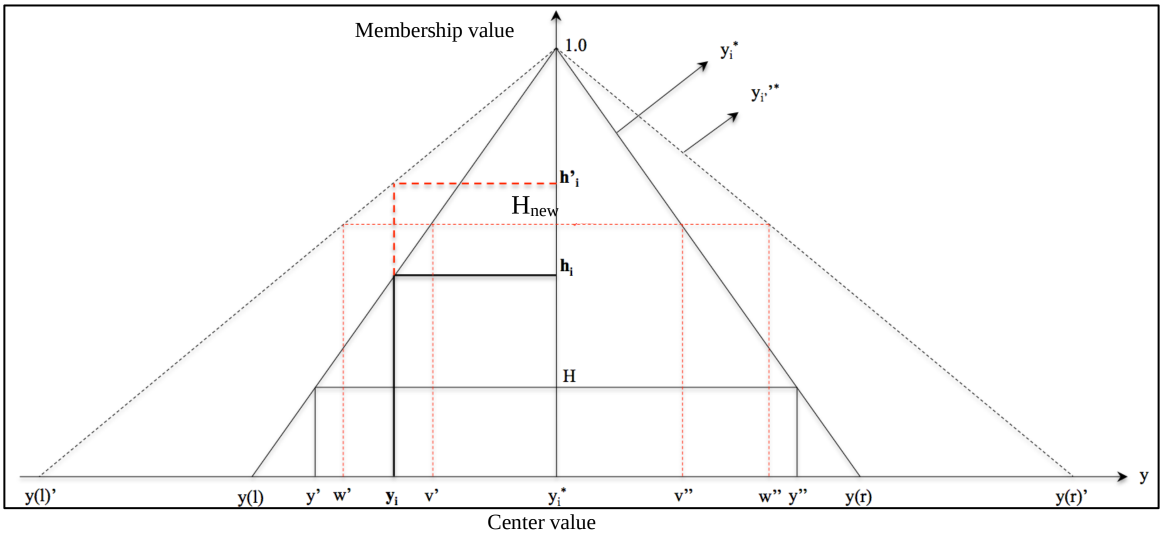

Figure 4. 3 h_i value of $\boldsymbol{y_i^*}$ to the observation y_i [80].

$\tilde{Y}_i$, the fuzzy estimates are obtained by multiplying each crisp explanatory variables Xj with fuzzy $\tilde{A}j$, although the dependent variable Y_i are crisp values. The condition that $\tilde{Y}i$ fuzzy intervals include observations Y_i is provided by the following two constraints:

$$\sum_j a_j X_i j + L^{-1}(h) \sum{}_j c_j \left| X_i j \right| \geq y_i \tag{4. 5}$$

$$\sum_j a_j X_i j \; - \; L^{-1}(h) \sum{}_j c_j \left| X_i j \right| \leq y_i \tag{4. 6}$$

$i = 1, 2 \ldots, n \qquad j = 0, 1, \ldots, k \qquad c_j \geq 0$

"n" is the number of observations for the dependent variable in the model. *k* is the number of explanatory variables. The number of constraints is determined by the number of observations, that is, *n*. Because a range is estimated by approaching Y_i from the left and right. Therefore, two constraints should be written for each Y_i observation value. In this case, the constraint number will be two times the observations. The increase or decrease in the number of explanatory variables does not change the constraints.

4.3 Possibilistic Regression Methods

If the parameters of a linear system are explained with possibility distributions, the term "possibilistic linear system" is used for such systems. Due to a fuzzy number takes a value between zero and one [0,1], a fuzzy number can signify a possibility distribution. Moreover, if a fuzzy number used as a parameter of an independent variable or observed value in a regression model, the model called Possibilistic Regression Model [96].

4.3.1 Linear Programming Approaches

Tanaka et al. first proposed the Linear programming approach for fuzzy linear regression (FLR) [75]. It has fuzzy output, crisp input, and fuzzy parameters. The model is established in a mathematical programming problem. This model aims to minimize the total spread of fuzzy parameters, depending on the support of the predicted values, and to ensure that the spread is created by the observed values for a given h-level.

As criticized by Redden and Woodall, the approaches are overly sensitive to outliers and can generate endless solutions [97]. Also, the distribution of predicted values becomes wider as more data is included in the model.

Based on this criticism, Tanaka [78], Tanaka and Watada [95], and Tanaka et al. [96] tried to develop the early fuzzy regression models.

4.3.1.1 Tanaka's Method

It is assumed that the spread between the observed and predicted data in fuzzy regression is due to the system vagueness or the fuzziness of the regression coefficients. The goal of fuzzy regression is to find a suitable regression model that covers all observed fuzzy data. Different fuzzy regression models can be produced depending on the use of appropriate criteria. The regression coefficients in the method are fuzzy numbers. Since the regression coefficients are fuzzy numbers, the predicted dependent variable Y value is also a fuzzy number. The fuzzy regression model with independent variables X_i is summarized below. A_0 is the fuzzy coefficient and A_i's are the fuzzy slope coefficients [75].

$$\widetilde{Y_\iota} = \widetilde{A_o} + \widetilde{A_1}X_1 + \cdots + \widetilde{A_p}X_{ip}$$

Each fuzzy parameter $A_i = (c_i, s_i)$ is expressed as symmetric triangular membership functions with central value c_i and spread value s_i.

According to the approach, the fuzzy output $\tilde{Y}$ is predicted when the fuzzy coefficients provide minimum spread to output and objective *"h term"* is also ensured . *"h term"* is referred to as "degree of fit", which measures the fit between data and the regression model. A basic fuzzy linear regression model is written as follows:

$$\widetilde{(Y)} = \tilde{A}_0 X_0 + \tilde{A}_1 X_1 + \cdots + \widetilde{A_N} X_N = \tilde{A} X$$

$X = [X_0, X_1, \ldots, X_N]^T$ is the vector of the independent variables,

$\tilde{A} = \left[\tilde{A}_0, \tilde{A}_1, \ldots, \tilde{A}_N\right]^T$ is the vector of the parameters that are defined below:

$\widetilde{A}_i$ parameter is a symmetric triangular fuzzy number whose structure is generated by $A_j = (c_j, s_j)$ fuzzy coefficient vectors. Here, c_j is the central value and s_j is the spread value.

$$\mu_{x_j}(a_j) = \begin{cases} 1 - \frac{|c_j - a_j|}{s_j}, c_j - s_j \leq c_j + s_j, \forall j = 1,2,\ldots,N \\ 0, \quad other\ cond. \end{cases} \quad (4.7)$$

Thus, the fuzzy regression model can be written as below :

$$\widetilde{Y_i} = (c_0, s_0) + (c_1, s_1)X_{1i} + (c_2, s_2)X_{2i} + \cdots + (c_N, s_N)X_{Ni} \quad (4.8)$$

The fuzzy regression model in Equation 4.8 predicts fuzzy outputs and fuzzy parameters which signify the fuzzy relationship between crisp input and fuzzy output data. Applying the extension principle, the membership function of the fuzzy number Y_i is calculated as follows:

$$\boldsymbol{\mu(Y_i)} = \begin{cases} \mathbf{1} - \dfrac{|\boldsymbol{Y_i} - \boldsymbol{X'c}|}{\boldsymbol{s'}|\boldsymbol{X}|}, & \boldsymbol{X} \neq \mathbf{0}, \\ \mathbf{1}, \quad \boldsymbol{X} = \mathbf{0}, & \boldsymbol{Y} \neq \mathbf{0}, \forall \boldsymbol{i} = \mathbf{1}, \mathbf{2}, \ldots, \boldsymbol{M} \\ \mathbf{0}, & \boldsymbol{X} = \mathbf{0}, \boldsymbol{Y} = \mathbf{0} \end{cases} \quad (4.9)$$

$$s^t = (s_0, s_1, \ldots, s_N), c = (c_0, c_1, \ldots, c_N)$$

Every dependent fuzzy variable can be calculated as $\widetilde{Y_\iota} = \left(Y_i^L, Y_i^{h=1}, Y_i^U\right)$, *i=1,2,3…, M.* The lower bound of the fuzzy number $\widetilde{Y_\iota}$ is $Y_i^L = \sum_{j=0}^{N}\left(c_j - s_j\right)X_{ij}$, the center of $\widetilde{Y_\iota}$ is $Y_i^{h=1} = \sum_{j=0}^{N}\left(c_j\right)X_{ij}$ and lastly the upper bound of $\widetilde{Y_\iota}$ is $Y_i^U = \sum_{j=0}^{N}\left(c_j + s_j\right)X_{ij}$.

To obtain an effective fuzzy regression model by reducing fuzziness, the objective function is adapted to minimize the total spread of the fuzzy number $\widetilde{Y_\iota}$;

$$MIN\ (s^T\,|X|) = MIN\sum_{j=0}^{N}\left(s_j\sum_{j=1}^{N}\left|x_{ij}\right|\right) \tag{4. 10}$$

The constraints require that each observation value of Y_i be related to $\widetilde{Y_\iota}$ with a minimum of *h* value. So, $\mu Yi \geq h\ (i = 1,2, \ldots, M)$

$$\mathbf{1} - \frac{|\boldsymbol{Y_i} - \boldsymbol{X'c}|}{\boldsymbol{s'}|\boldsymbol{X}|} \geq \boldsymbol{h}, \forall \mathbf{i} = \mathbf{1}, \mathbf{2}, \ldots, \mathbf{M} \tag{4. 11}$$

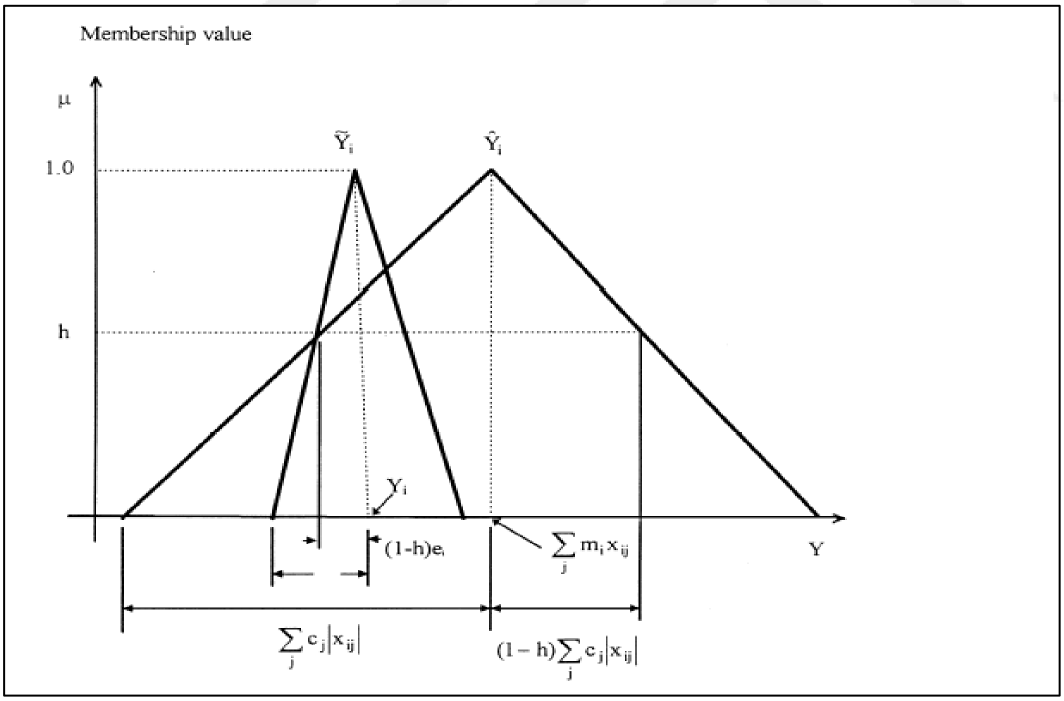

Figure 4. 4 The *h*-value, Y required to obtain fuzzy data Y_i [98].

To calculate the fuzzy coefficient $A_i = (c_i, s_i)$, the following linear fuzzy regression model developed by Tanaka et al. is formulated as follows:

$$min\, S = ns_0 + s_1 \sum_{i=1}^{n} |X_i| \tag{4. 12}$$

Subject to :

$$\boldsymbol{s_0 \geq 0, s_1 \geq 0} \tag{4. 13}$$

$$\sum_{j=0}^{I} c_j X_{ij} + (1-h) \sum_{j=0}^{I} s_j |X_{ij}| \geq Y_i + (1-h)\, e_i \qquad for\ i = 1,2 \ldots, n$$

$$\sum_{j=0}^{I} c_j X_{ij} - (1-h) \sum_{j=0}^{I} s_j |X_{ij}| \leq Y_i - (1-h)\, e_i \qquad for\ i = 1,2 \ldots, n$$

The objective function in the model minimizes the total spread, in other words, fuzziness. It is supposed to determine a value for *h term* in constraints. By the way, the centers and spread values of the fuzzy parameters can be estimated. Tanaka et.al targeted to cover all observations in the model and assumed that observations are sure and possible. They do not want to label some observations as outliers as in conventional regression models. However, they recommended to users to take the advice of experts when users select the data set and *h* value [96].

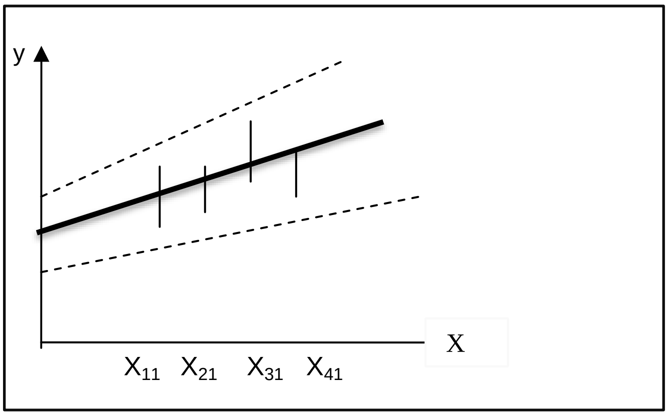

Figure 4. 5 The estimated view of the Tanaka Method.

The central tendency with bold line at the middle, lower and upper boundary with dashed lines are shown in Figure 4.5. As seen in the figure, Tanaka Method prefers to cover all observations into spreads that are at the same distance from the central line.

4.3.2 Goal Programming Approaches

When more than one objective function is considered in fuzzy regression models, the approaches are called Goal Programming Approaches (GPA) [77]. In essence, linear programming approaches can be accepted as a special case of Goal Programming. Sakawa and Yano formerly applied multi-objective programming techniques [99] and used FIFO data in FLR models [100]. However, the approach of Sakawa and Yano is criticized that the model is quite sensitive to outliers by Redden and Woodall [101]. Nasrabadi et al. suggested a GPA (The approaches boundaries are softer than previous models) to smooth over the obstacles that are emphasized by Redden and Woodall [91]. Thereafter, they enhanced the model by introducing FIFO data [102]. However, the model proposed by Nasrabadi et al. [91] would be utilized in the study.

4.3.2.1 Nasrabadi Method

When linear programming based FLR methods that minimize the fuzziness are easy for programming and calculation, on the other side FLS based methods that use least-squares of errors as a constraint reach minimum fuzziness level. The first group of methods have wider spreads and sensitive to outliers, FLS methods are quite tough for calculation [94]. Nasrabadi et al. formerly purposed a new arithmetic operation on symmetric triangular fuzzy numbers to decrease the spreads [103]. Nasrabadi et al. latterly revisited their model [103] to overwhelm the deficiencies of fuzzy regression methods by implementing the approach of Özelkan and Duckstein into a multi-objective model [91], [104]. The model can use fuzzy outputs and fuzzy inputs as in Equation 4.23 below:

$$\boldsymbol{\tilde{Y} = \widetilde{A_0}\widetilde{X_0} + \widetilde{A_1}\widetilde{X_1} + \cdots + \widetilde{A_N}\widetilde{X_N}} \quad (4.\ 14)$$

$\widetilde{A_j} = (c_j, s_j)$; are the parameters that centers are symbolized with c_j and the radius are symbolized with s_j. c $=(c_0, c_1, \ldots, c_n)^t$ and s $=(s_0, s_1, \ldots, s)^t$ are the vector of the fuzzy parameters $\widetilde{A_j}$'s. The fuzzy independent variables are defined as $\widetilde{X_{ij}} = (x_{ij}, r_{ij})$. Here x_{ij} symbolizes the center and r_{ij} symbolizes the radius of the spread of the fuzzy independent variable $\tilde{X}_{ij}$. The given outputs are $\widetilde{\widetilde{Y_{ij}}} = (\overline{y_{ij}}, \overline{e_{ij}})$ and finally, the fuzzy estimated values are defined as $\widetilde{Y_{ij}} = (y_{ij}, e_{ij})$. Here y_{ij} symbolizes the center and e_{ij} symbolizes the radius of the spread of the fuzzy estimated values $\tilde{Y}_{ij}$. The Nasrabadi method works after solving the quadratic programming problem below [103] :

$$Min: D^2(h) = \sum_{i=1}^{m} (sr_i^T - \overline{e_i})^2$$

Subject to:

$$\sum_{j=0}^{n} (c_j x_{ij}) + |(1-h)| \sum_{j=0}^{n} (s_j r_{ij}) \geq \overline{y_i} - |(1-h)| e_i, \quad (4.15)$$

$$-\sum_{j=0}^{n} (c_j x_{ij}) + |(1-h)| \sum_{j=0}^{n} (s_j r_{ij}) \geq -\overline{y_i} - |(1-h)| e_i, \quad (4.16)$$

c_j and $s_j \in \Re$, (j=0,1,2, …,n) (i=1,2, …,m) $0 \leq h \leq 1$

The brief format of Equation 4.15 is given below:

$$\boldsymbol{y_i} \geq \boldsymbol{\overline{y_i}} \quad (4.17)$$

And the brief format of Equation 4.16 is given below:

$$\boldsymbol{y_i} \leq \boldsymbol{\overline{y_i}} \quad (4.18)$$

The improved multi-objective fuzzy linear regression model is obtained by adding an extra objective function that ensures the soft border to the quadratic programming method above:

$$Min: D^2(h) = \sum_{i=1}^{m} (e_i - \overline{e_i})^2$$

$$Min: E^2(h) = \sum_{i=1}^{m} (\varepsilon_{i,L}^2 + \varepsilon_{i,R}^2)$$

s.t. $y_i - \overline{y_i} \leq \varepsilon_{i,L}$, $i = 1, \ldots, m,$

$\overline{y_i} - y_i \leq \varepsilon_{i,R}$, $i = 1, \ldots, m,$

$e_i \geq 0,$ $\varepsilon_{i,L}^2, \varepsilon_{i,R}^2 \geq 0$ $i = 1, \ldots, m,$

Here, $E^2(h) = \sum_{i=1}^{m} (\varepsilon_{i,L}^2 + \varepsilon_{i,R}^2)$ is a deviation from extreme values and ε values are the relaxation variables. In the MOFLR method, all constraints influence the solution, and herewith all observations join to estimation.

Hence, Nasrabadi Method keeps the spread in the model constant, inversely to other methods, Nasrabadi Method also uses symmetric triangular fuzzy numbers either as an independent variable or as a dependent variable.

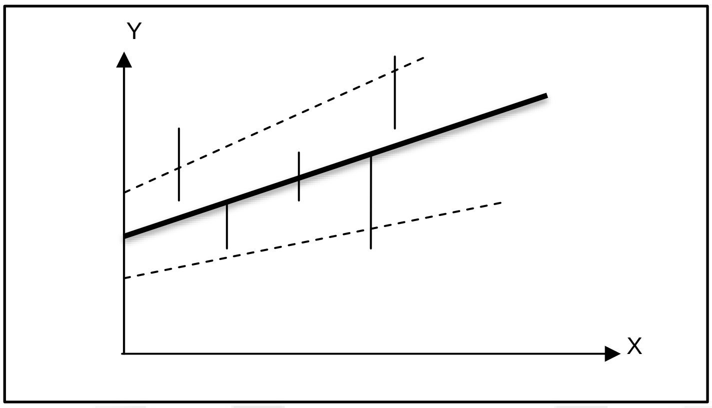

Figure 4. 6 The estimated view of the Nasrabadi Method

The central tendency with bold line at the middle, lower and upper boundary with dashed lines are shown in Figure 4.6. As seen in the figure, Nasrabadi Method models the data with soft boundaries without covering all observations. However, studies that use a vast number of variables is reported that the borders' spreads get narrower.

4.3.3 Interval Regression Approaches

The interval regression is a basic type of the possibilistic regression model that is stated by Tanaka [105]. The system parameters, in other words, fuzzy regression coefficients, are defined in intervals. Various interval regression methods are considered. When some of the proposed models use symmetric triangular fuzzy numbers, others proposed asymmetric triangular or trapezoidal fuzzy numbers [75]. Tanaka and Lee developed the basic interval model by implementing quadratic programming [105]. Lee and Tanaka also studied estimating lower and upper boundaries with their interval model [106].

4.3.3.1 Lee and Tanaka Method

When an interval regression method uses symmetric triangular fuzzy numbers, the upper and lower lines are pointlessly wide. To cope with the drawback, Lee and Tanaka suggested an interval model that estimates parameters as a non-symmetric triangular fuzzy number. Additionally, they combined the least square method for central tendency and LP techniques for lower and upper boundaries [107]. An FLR model is assumed as below:

$$\widetilde{Y_\iota} = \widetilde{A_o} + \widetilde{A_1}X_1 + \cdots + \widetilde{A_n}X_n \quad (4.19)$$

Where *X* is a crisp input vector, *Y(x)* is estimated fuzzy output $\tilde{A} = (A_0, \ldots, A_n)$ is the fuzzy coefficient vector with non-symmetric triangular fuzzy numbers, $\tilde{A}_i$ stated as $A_i = (a_i, c_i, d_i)_T$. Here, a_i is a center, c_i is a left spread, and d_i is a right spread. The estimated output data can be written as below:

$$Y(x_j) = \left(\theta_c(x_j), \theta_L(x_j), \theta_R(x_j)\right) \quad (4.20)$$

$$Y(x_j) = \left(\sum_{i=0}^{n} a_i x_{ji}\right)_c, \left(\sum_{x_{ji}\geq 0}^{n} c_i x_{ji} - \sum_{x_{ji}<0} d_i x_{ji}\right)_L, \left(\sum_{x_{ji}\geq 0}^{n} d_i x_{ji} - \sum_{x_{ji}<0} c_i x_{ji}\right)_R \quad (4.21)$$

The *h*-level set of *Y*(x) can be denoted as below:

$$[Y(x)]_h = \{y | \mu_{Y(x)}(y) \leq h\} = [y_h^-, y_h^+] \quad (4.22)$$

where

$$y_h^- = \theta_c(x_j) - (1-h)\theta_L(x_j),$$

$$y_h^+ = \theta_c(x_j) + (1-h)\theta_R(x_j)$$

y_h^- *and* y_h^+ represent the lower and upper boundaries. The minimization function for the sum of squared distance between estimated and observed outputs can be formulized as below:

$$Min: J_{LS} = \sum_{i=1}^{m} (y_j - a^t x_j)^2 \quad (4.23)$$

And the minimization function of the sum of spreads of estimated output is written below:

$$Min: J_{LP} = (1-h) \sum_{i=1}^{m} \left(c^t |x_j| + d^t |x_j|\right) \quad (4.24)$$

The newly proposed method combines the objection functions in Equation 4.23 and 4.24 above as a new joint objection function. After the specification of the objection function and constraints, Lee and Tanaka method can be designated as below:

$$Min_{a,c,d}: J = k_1 J_{LS} + k_2 J_{LP} + \xi(c^t c + d^t d)$$

Subject to: $y_h^{+} \geq y_j$,

$y_h^{-} \leq y_j, j = 1, \dots, m$,

$c_i \geq 0, d_i \geq 0, i = 0, \dots, n$,

In the QP model, k_1 and k_2 are the weight coefficients, ξ is a small positive number that is $k_1, k_2 \gg \xi$.

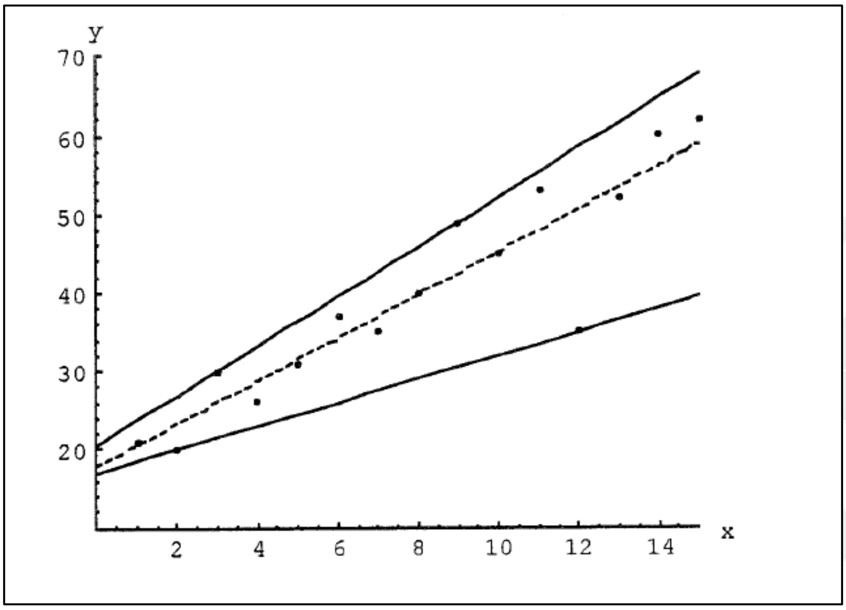

Figure 4. 7 Fuzzy regression model with non-symmetric TFN coefficient by QP [107].

The central tendency with dashed line at the middle, lower and upper boundary with bold lines are shown in Figure 4.7. As seen in the figure, Lee and Tanaka Method also prefers to cover all observations into spreads. Inputs and observed outputs are crisp numbers in the model. However, the central tendency is calculated with Least Square Method and the spreads are calculated with the LP technique. Due to the combination of different techniques and usage of non-symmetric TFN as system parameters, the lower and upper boundaries have not same distance from the central line.

4.4 Fuzzy Least Square Method

Diamond and Tanaka reported that two conditions in a fuzzy least square method are studied [108]. These are:

- Fuzzy input and fuzzy output (FIFO).
- Crisp input and fuzzy output (CIFO).

Fuzzy the Least Square method is put forward by diamond in 1988 with the aim of parameter estimation [79]. The fuzzy least square method can be interpreted as the classical linear regression method's fuzzy expansion. The process of obtaining the parameters in the fuzzy least square method has similarities in the classical regression method. The main purpose of the model proposed by Diamond is to minimize the distance between the estimated value and observed value. The inputs and outputs can be used as a fuzzy number or crisp number in the method. Two different models are offered as written below:

$$\boldsymbol{Y = a + bX, a, b \in R,} \tag{4. 25}$$

$$\boldsymbol{Y = \widetilde{E} + bX, b \in R,\ \widetilde{E} \in F(R)} \tag{4. 26}$$

It is assumed that the observations involved with X_i and Y_i couples (i=1,2, …. N). So, the model is detailed, and the version is shown below:

$$\boldsymbol{X_i = \left(x_i, \underline{\xi}_i, \overline{\xi}_i\right), Y_i = \left(y_i, \underline{\eta}_i, \overline{\eta}_i\right)} \text{ and } \boldsymbol{x_i - \underline{\xi}_i \geq 0} \tag{4. 27}$$

When the Eq.4.38 the least square optimization problem is solved:

$$\textit{Minimum } r(a, b) = \Sigma\, d(a + b X_i\,, Y_i)^2$$

In the minimization problem, two situations occur; the first is b ≥ 0 and the second one is b<0. If b ≥ 0, the distance between fuzzy outputs and the observed values of the dependent variables is calculated in Equation 4.41 below :

$$\boldsymbol{d(a + bX_i, Y_i) = \left[a + bx_i - y_i - \left(b\underline{\xi_i} - \underline{\eta_i}\right)\right]^2 + \left[a + bx_i - y_i - \left(b\overline{\xi_i} - \overline{\eta_i}\right)\right]^2 + (a + bx_i - y_i)^2} \tag{4. 28}$$

When b<0, The equation of the distance is shown in Equation 4.42:

$$\boldsymbol{d(a + bXi, Yi) = \left[a + bx_i - y_i + b\overline{\xi_i} - \underline{\eta_i}\right]^2 + \left[a + bx_i - y_i - b\underline{\xi_i} + \overline{\eta_i}\right]^2 +}$$

$$(\boldsymbol{a} + \boldsymbol{b}\boldsymbol{x}_i - \boldsymbol{y}_i)^2 \quad (4.\ 29)$$

The next step at the least square method is the derivation of the equation 4.41 according to a and b. After solving the 4.43 equations system below, a and b parameters can be calculated[69], [108].

$$\frac{\partial \boldsymbol{d}}{\partial \boldsymbol{a}} = \boldsymbol{3Na} + \boldsymbol{b\Sigma}[\boldsymbol{3x_i} + \boldsymbol{\delta}(\boldsymbol{X_i})] = \boldsymbol{\Sigma}[\boldsymbol{3y_i} + \boldsymbol{\delta}(\boldsymbol{Y_i})] = \boldsymbol{0}\ , \quad (4.\ 30)$$

$$\frac{\partial d}{\partial b} = a\Sigma[3x_i + \delta(X_i)] + b[x_i{}^2 + \left(x_i - \underline{\xi_i}\right)^2]$$

$$= \Sigma\left[\left(x_i - \underline{\xi_i}\right)\left(y_i - \underline{\eta_i}\right) + \left(x_i + \overline{\xi_i}\right)\left(y_i + \underline{\eta_i}\right) + x_i y_i\right] = 0$$

Here x_i (i=1,2, … , n) and,

$$\delta\ (X) = \overline{\xi} - \underline{\xi}$$

$$\delta\ (Y) = \overline{\eta} - \underline{\eta}$$

The minimum optimization of the model offered by Diamond is shown in Equation 4.39 is $\rho(\ E, b) = \Sigma(E + bX_i, Y_i)^2$. The term "E" in the equation can be defined as $E = (c, \underline{\gamma}, \overline{\gamma})_T$. There are two conditions in the model depicted by Equation 4.38. The conditions are $b \geq 0$ and $b<0$.

If $b \geq 0$, the equation 4.41 will be transformed the Equation 4.44 below:

$$\boldsymbol{d}(\boldsymbol{E} + \boldsymbol{bXi}, \boldsymbol{Yi}) = \left[\boldsymbol{c} + \boldsymbol{bx_i} - \boldsymbol{y_i} - \underline{\boldsymbol{\gamma}} - \left(\boldsymbol{b}\underline{\boldsymbol{\xi_i}} - \underline{\boldsymbol{\eta_i}}\right)\right]^2 + [\boldsymbol{c} + \boldsymbol{bx_i} - \boldsymbol{y_i} + \overline{\boldsymbol{\gamma}} + \left(\boldsymbol{b}\overline{\boldsymbol{\xi_i}} - \overline{\boldsymbol{\eta_i}}\right)]^2 + \ (\boldsymbol{c} + \boldsymbol{bx_i} - \boldsymbol{y_i})^2 \quad (4.\ 31)$$

In the condition of b < 0, the equation 4.42 will be transformed the Equation 4.45 below:

$$\boldsymbol{d}(\boldsymbol{E} + \boldsymbol{bXi}, \boldsymbol{Yi}) = \left[\boldsymbol{c} + \boldsymbol{bx_i} - \boldsymbol{y_i} - \underline{\boldsymbol{\gamma}} - \left(\boldsymbol{b}\underline{\boldsymbol{\xi_i}} + \underline{\boldsymbol{\eta_i}}\right)\right]^2 + [\boldsymbol{c} + \boldsymbol{bx_i} - \boldsymbol{y_i} + \overline{\boldsymbol{\gamma}} + \left(\boldsymbol{b}\overline{\boldsymbol{\xi_i}} + \overline{\boldsymbol{\eta_i}}\right)]^2 + \ (\boldsymbol{c} + \boldsymbol{bx_i} - \boldsymbol{y_i})^2 \quad (4.\ 32)$$

The inputs are crisp numbers that are x_i, (i=1,2…., N.) and the outputs are fuzzy numbers that are $Y_i = (y_i, \overline{\eta}, \underline{\eta})$ in the model proposed by Diamond 83. The model:

$$Y = A + xB,\ x \in R,\ A \in T(R),\ B \in \wp(\mathrm{R}),\ \mathrm{Y_i} \in F(R) \tag{4.33}$$

The parameters in the model are defined with $A = \left(a, \underline{a}, \overline{a}\right)$ and $B = \left(b, \underline{\beta}, \overline{\beta}\right)$, also A and B are symmetric fuzzy numbers. After applying the crisp inputs and fuzzy outputs to the model, it will be improved as below:

Minimum r(A,B) = $\Sigma d(A + x_i B, Y_i)^2$

When the parenthesis of the equation above is opened:

$$\boldsymbol{d(A + x_i, Y_i) = (a + bx_i - y_i)^2 + \left(a + bx_i - \underline{a} - \underline{\beta} x_i + y_i + \underline{\eta_i}\right)^2 + \left(a + bx_i + \overline{a} + \overline{\beta} x_i + y_i + \overline{\eta_i}\right)^2} \tag{4.34}$$

The parameters of the model (a and b) will be calculated when the equation derivation is taken by a $\left(\frac{\partial d}{\partial a}\right)$ and b $\left(\frac{\partial d}{\partial b}\right)$, after the results of the derivation equaled to zero. The calculated parameters' centers and fuzzy spread are given below:

A$= \hat{y} - b\hat{x}$

$a = \hat{\eta} - \beta\hat{x}$

$b = K/_{T^2}$

$\beta = k/_{T^2}$

The K, T2, $\hat{x}$,$\hat{y}$ are defined below with the known x, y, and other values [79]:

$K = \Sigma(\mathrm{x_i} - \hat{\mathrm{x}})(y_\mathrm{i} - \hat{\mathrm{y}})$

$k = \Sigma(\mathrm{x_i} - \hat{\mathrm{x}})(\eta_\mathrm{i} - \hat{\eta})$

$\hat{x} = \frac{\Sigma x_i}{N} \qquad \hat{y} = \frac{\Sigma y_i}{N}$

$T^2 = \Sigma(x_i - \hat{x})^2 \qquad a, \beta \geq 0$

The method proposed by Diamond [79] was developed by the studies of Wang and Tsaur [94]. Wang and Tsaur studied crisp input and fuzzy output regression which was proposed by Tanaka et al. [75], they applied the Least Square approach to Tanaka's method and reached a simpler and more predictable way according to Diamond's common method.

4.5 Comparison of The Fuzzy Linear Regression Models

The power, efficiency, or goodness of crisp regression models are probable to compare with each other, according to some statistical indicator. Thus, the most appropriate crisp regression model can be distinguished from the other models straightforwardly. On the contrary, fuzzy regression models work without distribution and they do not have detailed statistics like crisp regression models. The comparison of the fuzzy regression models is a necessity to prefer a more applicable model. In the study, two indicators would be used as distinguisher property.

Total Error Fit (TEF) is one of the statistics that calculates the difference between the membership functions of the observed and estimated output in Triangular Fuzzy Number form [109]. The total sum of errors can be defined as below:

$$\boldsymbol{E} = \sum_{i=1}^{n} |\mu_{Y}(x) - \mu_{\tilde{Y}}(x)| \qquad (4.35)$$

Another statistic is Goodness of Fit (GOF) which is the mean squared distance between response and prediction based on the concept of Diamond's distance [79] [110]. The error term can be calculated as in Equation 4.36 below:

$$GOF = \frac{1}{n}\sum_{i=1}^{n} d^2(Y_i, \ \hat{F}(x_i))$$

$$\boldsymbol{GOF} = \frac{1}{n}\sum_{i=1}^{n} ((l_{y_i} - \hat{l}(x_i))^2 + (c_{y_i} - \hat{c}(x_i))^2 + (u_{y_i} - \hat{u}(x_i))^2) \qquad (4.36)$$

5 APPLICATION

The requirement of humanity to energy increases parallel to the world population [111]. The huge consumption of fossil fuel to generate energy that causes global warming currently reveals the importance of clean and renewable energy sources [112]. Among renewable energy sources, wind energy is one of the most prominent ones.

The tools that alter wind energy to dynamic energy and consequently generate electrical energy are called “Wind Turbines” and the places where more than one wind turbine is constructed to generate cumulative electrical energy are named “Wind Plants” or “Wind Farms”. The wind energy plants have no harmful effect on the environments, human beings, and other living beings around the place they constructed.

Besides the significant advantages of wind energy, there are some disadvantages, such as its intermittent form and high construction costs of wind farms. Therefore, the selection of the construction zone is one of the most principal factors in the calculation of the cost and the performance of the wind turbine.

The values of the wind speed and the wind direction, temperature, air density, and other related parameters of the construction zone have to be observed before at least one year from the construction time of a wind turbine. Similarly, the distance between the construction zone and main transportation roads or residential area, the joint convenience to the national electric line, the slope of the terrain, the flora of the area are the other critical issues to construct a wind turbine on a selected zone [113].

5.1 The Calculation of The Power Curve

The power curve is the theoretical electric power that is expected to be generated by a wind turbine at any place on earth. It is generally used by wind turbine manufacturers as an empirical parameter to predict the power that would be generated. The theoretical electric power is explained in Equation 5.1.

$$P_w = 1/2\varepsilon \times \rho \times A \times V^3 \tag{5. 1}$$

Here, P_w is power (watt), ρ is the density of the air (kg/m^3), A is the area swept by the wing of the turbine (m^2), v is the speed of wind (m/s), ε is the efficiency of the turbine that is generally around %30. The area is:

$$\boldsymbol{A = \pi r^2} \qquad (5.\ 2)$$

where r is the radius of the turbine rotor (m). According to Equation 5.2, the theoretical power would increase to the square of the radius of the wind turbine rotor, and cubic of the speed of wind relatively. However, a wind turbine cannot generate electrical power when the wind speed is approximately below 3,5 m/s, and it cannot generate more electrical energy or it reaches maximum generated electricity capacity when wind speed is approximately above 13,5 m/s, due to its physical feature. Consequently, Figure 5.1 can be obtained by using the constraints and features explained. The obtained line looks like a curve, and it is called “Wind Turbine Power Curve (WTPC)”.

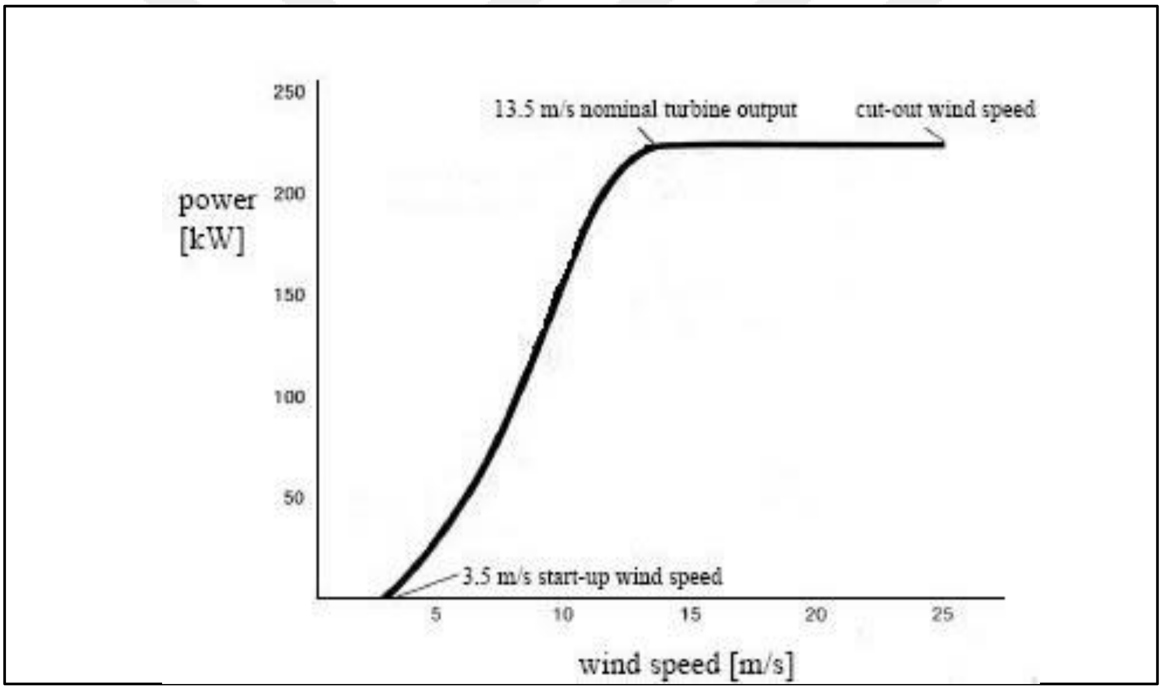

Figure 5. 1 Wind Turbine Power Curve (WTPC) [11].

A WTPC is determined by the manufacturer of a wind turbine according to the characteristics of the wind turbine. Nevertheless, the observed generated power curve is not usually matching to manufacturers’ one. Thus, short-term power prediction that gives instant information about the energy at a specific time is becoming an important subject to analyze the power curve. As a result of the propagation of wind energy and the necessity to minimize the costs, lots of models are created by scientists. However, it can be said that the best single way for all situations is not found, owing to numerous variables with different combinations like wind speed, temperature, and geographical structures. There are two types of wind power prediction methods. When the first group of methods uses statistical methods that are more efficient for short-term observations (1- 8 hours), the other group of methods is numeric models and they reach wind power by using wind speed, wind direction, temperature, air density, and similar quantities that are more effective with 8–12 hours. In other words, the second group methods easily study manufacturers' WTPC [114].

The differences between theoretical and observed power led researchers to apply many statistical methods to the cases. After examination of WTPC models, it is divulged that proposed models are frequently complex. Because estimated parameters depend on several factors, the situation requires the application of parametric or non-parametric techniques. The main parametric methods are the Linearized Model, Polynomial Model, Probabilistic Model, and Logistic function model. The main non-parametric methods are Neural Networks, Data Mining Algorithms, and Fuzzy Clustering Methods [11].

Consequently, many methods are applied to forecast the power generated by wind turbines, and important studies are revealed. For instance, Croonenbroeck and Ambach applied the time series model to wind power forecasting [115]. Villanueva and Feijoo used logistic functions for wind turbine power generation [116]. The non-linear regression methods are studied by Marciukaitis et al. [117]. Moreover, machine learning methods are practiced. Different techniques are tried by Marvuglia and Messineo [118]. Park and Hur concerned Support Vector Machines to Short-Term Power forecasting [119] [120].

6 IMPLEMENTATION OF THE PROPOSED MODELS

In the section, crisp models are implemented to the dataset of 2292 observations to understand whether the crisp models work or not. It is emphasized that other statistical models like non-linear regression and machine learning methods are applied to a similar dataset that belongs to a wind turbine. It is realized that fuzzy linear or fuzzy nonlinear regression methods are not utilized in a wind turbine dataset formerly. Here, fuzzy linear regression methods that are stated in Section 4 are implemented to the dataset in diverse conditions. Thereby, the convenience and functionality of the fuzzy methods are discussed. Moreover, the results of the different fuzzy linear regression methods are compared in the section.

6.1 Description of The Data Set and The Analyzed Case

A wind turbine from a wind farm constructed in Canakkale District of Turkey is selected for analysis for the study because there are plenty of wind speed and wind direction observations that have sufficient variability and availability. The real data is collected by the Supervisory Control and Data Acquisition (SCADA) system of the wind turbine. The components of the dataset are wind speed (m/s), theoretical power output (kW), generated power output (kW), and wind direction (°). The measurements include 1-hour intervals covering the period from 01-09-2018 to 31-12-2018. The theoretical power output (kW) is Power Curve given by the manufacturer. The data consist of 4 months, 2292 observations. The first three and last three observations of the data set are depicted in Table 6.1 below:

Table 6. 1 Sample of the Dataset.

Date/Time	Theoretical Power (kW)	Power Generated (kW)	Wind Speed (m/s)	Wind Direction (°)	Cosines of Wind Direction
01-09-2018 00:00	3588.3	3404.1	12,6	72.3	0.303
01-09-2018 01:00	3478.1	3102.2	11.7	70.6	0.332
01-09-2018 02:00	3572,1	3222.7	12.4	74.1	0.273
⋮	⋮	⋮	⋮	⋮	⋮
31-12-2018 21:00	2601.1	2309.9	9.7	80.3	0.167
31 12 2018 22:00	3025.2	2681.3	10.4	80.5	0.166
31 12 2018 23:00	3583.3	3514.3	12.5	80.5	0.165

The figure of the initial data set that contains power output, wind speed, and cosines of wind direction is shown in Figure 6.2. Normally, the wind directions 359° and 001° are neighbors to one other, cosine that is a method to linearize the direction. The wind direction is limited between -1 and 1.

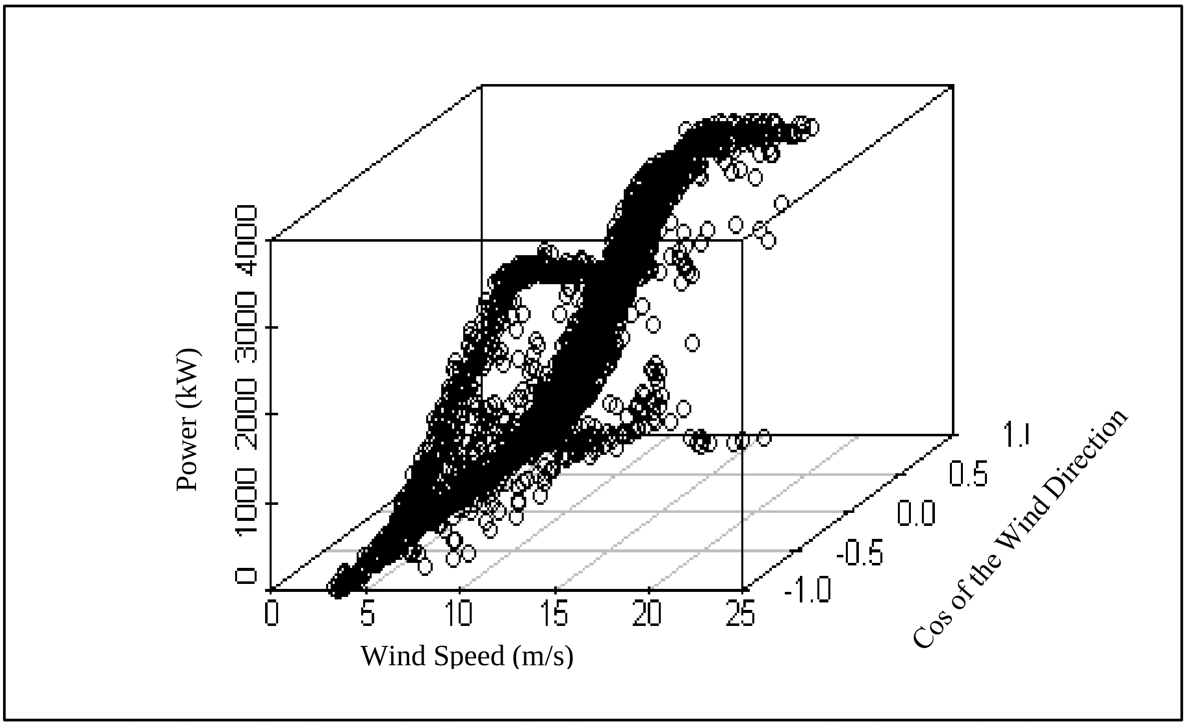

Figure 6. 1 The initial data set of power output, wind speed, and wind direction.

The three-dimension graphic of the data set is shown in Figure 6.1 above. Because cosines of wind directions are used in the study especially as an independent variable in regression models, the cosine values are preferred to show in the figure. Also, the fuzziness of the relations between variables is almost distinctive at first look.

The relation between theoretical power and wind speed, also between generated power and wind speed are illustrated in Figure 6.2 below [11]:

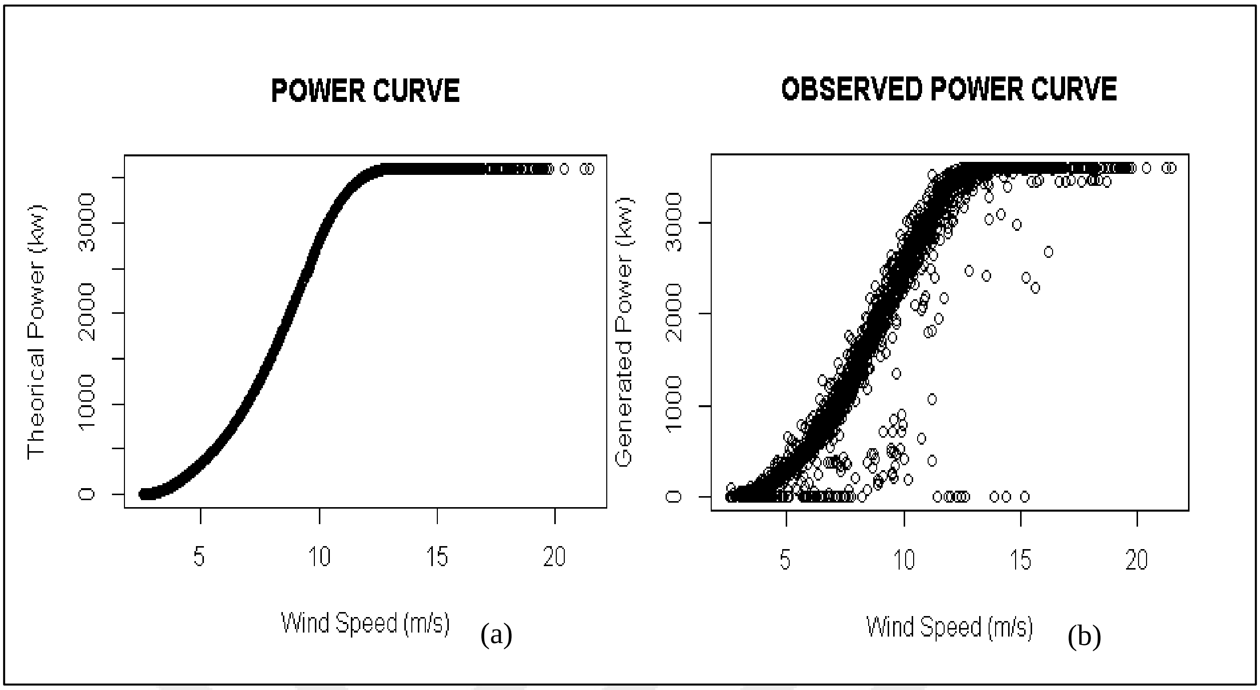

Figure 6. 2 (a) Wind Speed vs. Theoretical Power, (b) Wind Speed vs. Generated Power.

The curve shown in Figure 6.2 (a) is the power curve that belongs to the manufacturer of the wind turbine used in the study. As seen, it is similar to a generic WTPC example given in Figure 5.1. On the other hand, the observed power curve is shown in Figure 6.2 (b) is unlike common WTPC. The observed power curved behaves in that it has an upper and lower spread from the central tendency that is expected to be similar to the theoretical power curve. Also, there are outliers in Figure 6.2 (b) whose reasons cannot be explained. For instance, there are "0" kw generated power observations, although the wind speed increases. It can be assumed that the wind turbine is out of order when the observations are recorded. Because of such occasions, fuzzy regression models will be put on trial as explained before in the study.

6.2 Proposed Models

The modeling of proper location selection for wind turbines with fuzzy linear regression is the main purpose of the study. For the aim, fuzzy linear regression methods are applied to the wind turbine dataset not only to verify the reliability of the manufacturer's expected electric power generation curve but also to control the suitability of the place where the turbines were constructed in the study.

Even though it looks there is a relation between the wind speed and the generated power, there are also uncontrollable factors that affect the wind speed such as pressure, temperature, and terrain of the location, etc. The unexplained relationships at the background of the dataset steered to use of fuzzy regression in the study. Because it is assumed that there are strong relations between error terms and system fuzziness, applying the fuzzy regression models instead of crisp regression models is evaluated to be more efficient and powerful [104].

Although the data set has a vast number of observations with narrow periods like every hour in four months, the number of the independent variables could not ease to conclude about estimations of the generated energy values. The recent studies show that WTPC's have generally Weibull or Rayleigh distribution [11], therefore linear models are not effective to forecast wind turbine power. Although Piecewise linear regression appears more operative, the most suitable model cannot be detected. Since fuzzy regression methods do not consider distribution for modeling, the fuzzy regression methods look more applicable. Thus, the fuzzy regression models will be more effective to evaluate the outputs. When these conditions occur, the fuzzy regression models are especially recommended to use for estimation of generated power by a wind turbine:

- When a general frame of the wind turbine models is wanted to realize before complex calculations and modeling for a place,
- When there are a small number of observations,
- When it is hard to observe the parameters,
- When there is a more optional place to construct a wind farm.

CICO, CIFO, and FIFO datasets would also be used in models. So, these types of datasets would be produced with the method of fuzzification. The fuzzified observations would be in the Triangular Fuzzy Number (TFN) form. The symmetric form of the triangular fuzzy numbers would be used in both dependent and independent variables. When the application of non-symmetric triangular fuzzy numbers is just possible in the Fuzzy Least Square method as a dependent variable, it is not preferred due to the difficulty of comparability between other models. Compendiously, the dependent variable, and independent variables would have symmetric spread to lower and upper borders. The main output, generated power, in the dataset would be fuzzified with the help of the theoretical output data. It was mentioned that theoretical power is generally greater than generated power in the previous section. Hence, the difference between theoretical power and generated power would be used as a symmetric spread for the upper and lower side, but the generated power would be accepted as a central tendency. The main input or independent variable, wind speed, would be symmetrically fuzzified with 10 % below and above the observed value. As emphasized above, three groups of datasets (CICO, CIFO, FIFO) would be applied, and two models would be set up. The first model is:

$$\widetilde{Y} = \widetilde{A_o} + \widetilde{A_1}X_1$$

Here, $\widetilde{Y}$ is the predicted wind power (kW) in TFN form, X_1 is the wind speed (m/s) that can be crisp or symmetric TFN according to the applied method. Lastly, $\widetilde{A_o}, \widetilde{A_1}$ are the fuzzy coefficients. The second model is:

$$\widetilde{Y} = \widetilde{A_o} + \widetilde{A_1}X_1 + \widetilde{A_2}X_2$$

In the second model, there is an extra regressor X_2different from the first model. It is the cosine value of the observations' wind directions. The variable would be crisp form in applications.

The models are practiced with the "fuzzyreg" and "FuzzyNumbers" package in R Programming [121], [122]. The package includes the Lee and Tanaka method [107] that is an example of interval regression methods, Tanaka Method [96] that is an example of the Fuzzy Linear Programming Method, Fuzzy Least Square Method [77] and lastly Nasrabadi Method [91] that is an example of the Fuzzy Goal Programming Methods. The number of independent variables, type of variables, and type of predicted values is summarized below in Table 6.2. [121]**.**

Table 6. 2 The Available Method and their features in "fuzzyreg" package in R.

METHOD	***M***	***X***	***Y***	$\hat{Y}$ **predicted**
Lee & Tanaka	**∞**	crisp	crisp	nsTFN
Tanaka	**∞**	crisp	sTFN	sTFN
Fuzzy Least Square	1	crisp	nsTFN	nsTFN
Nasrabadi	∞	sTFN	sTFN	sTFN

In Table 6.2, *M* is the number of allowed independent variables, "crisp" means crisp numbers should be used as a variable, sTFN is symmetric Triangular Fuzzy Numbers (TFN), and nsTFN is non-symmetric TFN. The whole methods that are implemented to Model-1 and Model-2 with different type of input-output (CICO, CIFO, FIFO) in the study are detailed in Figure 6.3:

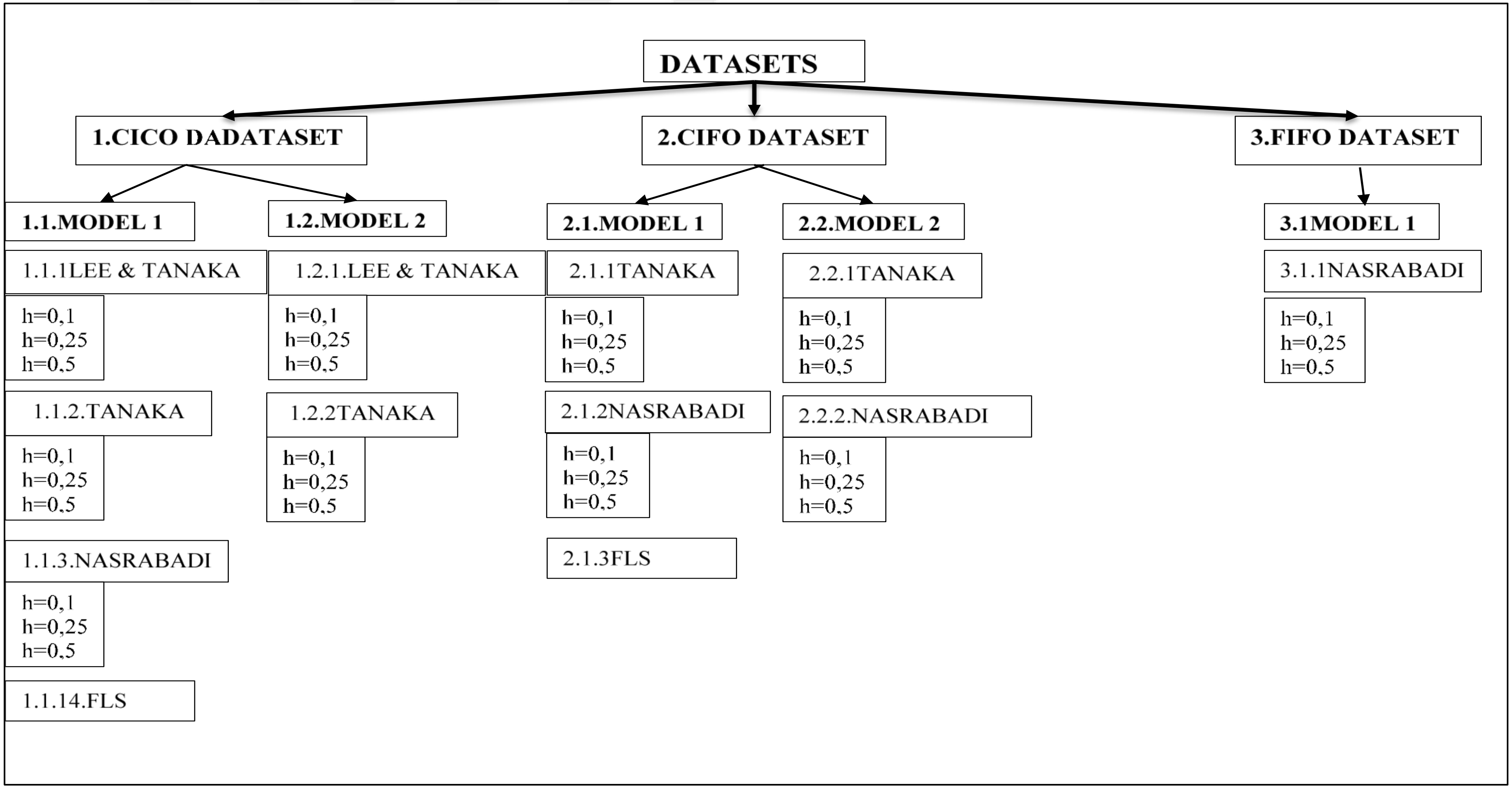

Figure 6. 3 Scheme of the Models that are applied to Dataset.

6.3 Crisp Models

Firstly, the classical (crisp) linear regression method is applied to understand the basics of the dataset and the effectiveness of the classical linear regression model to compare it with fuzzy models.

6.3.1 Implementation of Classical Linear Regression Method to Model 1.

Model 1 will be analyzed, next, Model 2 will be evaluated with the crisp method. Model 1, a crisp simple linear regression model, is stated again below:

$$\boldsymbol{Y = \beta_o + \beta_1 X_1}$$

Here, Y is the predicted wind power (kW), X_1 is the wind speed (m/s). The regression equation is:

$$Y = -1207 + 329X_1$$

The summary of the crisp linear regression model is as in Table 6.3 below:

Table 6. 3 Statistics of the Crisp Linear Regression Model.

Predictor	**Coefficients**	**Estimated Standard Error**	**T Value**	**P**
Intercepts	-1207.30	27.14	-44.49	0.00
Wind Speed (X_1)	328.96	2.88	114.24	0.00

The model is significant at 0.05 level and R-squared is 0.85. The ANOVA table of the model is shown below in Table 6.4:

Table 6. 4 ANOVA Table of the Crisp Linear Regression Model.

Source	**DF**	**Sum of Sq.**	**Mean Sq.**	**F**	**P**
Regression	1	3237788492	3237788492	13050	0.00
Residuals	2291	568411674	248106		
Total	2292	3806200166			

Under this circumstance, statistical transformations should be applied to variables. The crisp linear regression model significant and the R-squared statistics of the model is 0.85.

The residual versus fitted values and Normal Probability plots can be shown in Figure 6.4 (a) and Figure 6.4 (b):

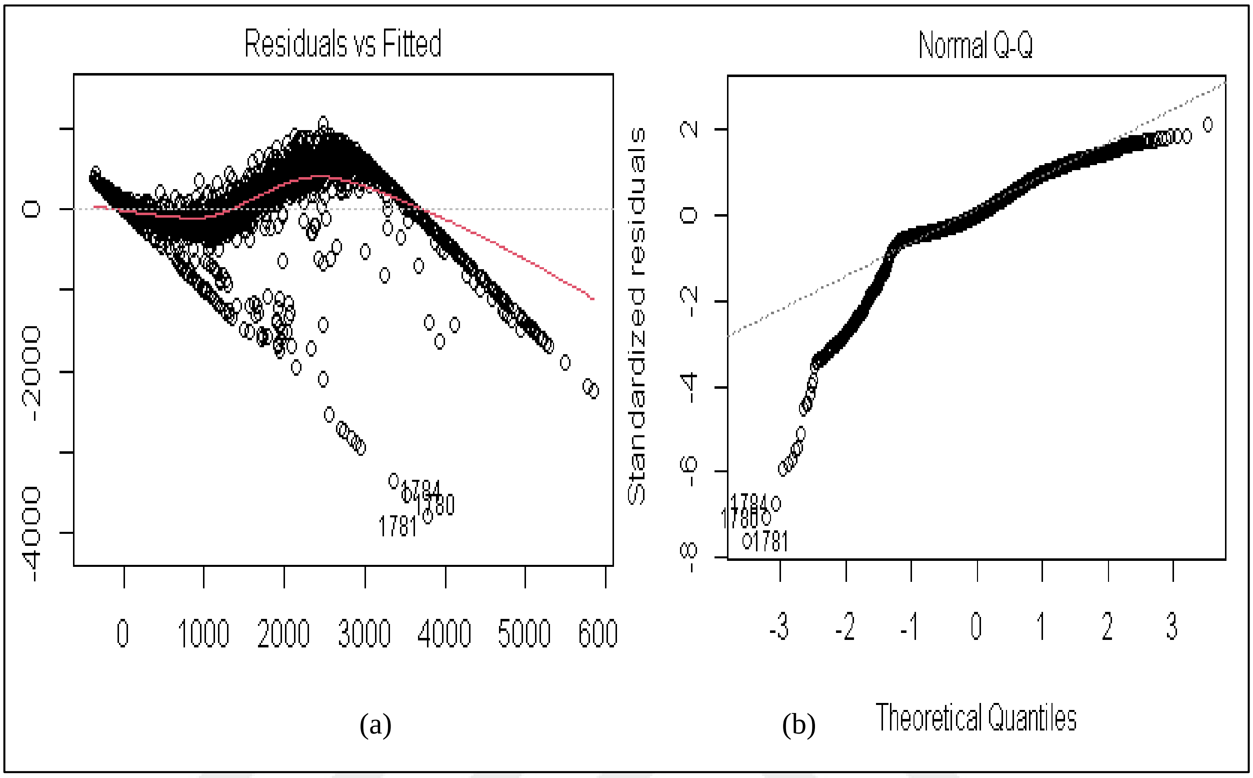

Figure 6. 4 (a) Residuals versus Fitted Values. (b) Normal Q-Q Plot.

When Figure 6.4 (a) Residuals versus Fitted Values plot is analyzed, it is seen that the residuals do not scatter randomly around the zero line, which means that the linearity assumption is not reasonable. When residuals are expected to be distributed equally around the zero line, the residuals create an unexpected pattern. It suggests that the variance of the error terms are not equal. Also, there are many outliers on the graph as an undesired situation. Therefore, the errors can be said errors are not distributed normally, moreover, the Normal Q-Q plot in Figure 6.4 (b) proves the non-normality.

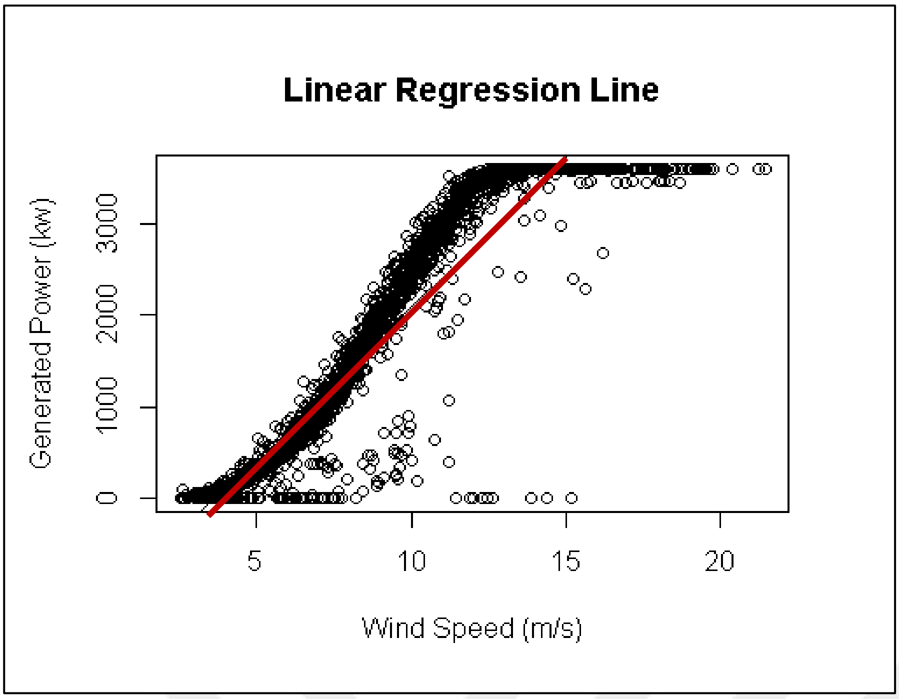

Figure 6. 5 The Linear Regression Line of the Crisp Model.

The linear regression line is depicted in Figure 6.5. The Box-Cox [123] transformation is also applied to the independent variable to enhance the linear regression line. However, the results are not at expected level. Adding one more independent variable to the model is an option to increase the effectiveness of the model.

6.3.2 Implementation of Classical Linear Regression Method to Model 2.

Model 2 is a crisp multiple linear regression model as stated below:

$$\boldsymbol{Y = \beta_o + \beta_1 X_1 + \beta_2 X_2}$$

Here, Y is the predicted wind power (kW), X_1 is the wind speed (m/s) and X_2 is the cosines of the wind direction. The regression equation is:

$$Y = -1218.68 + 329.65X_1 + 28.53X_2$$

This model is significant at 0.05 level, MSE is 497.8 and R-squared is calculated as 0.8509.

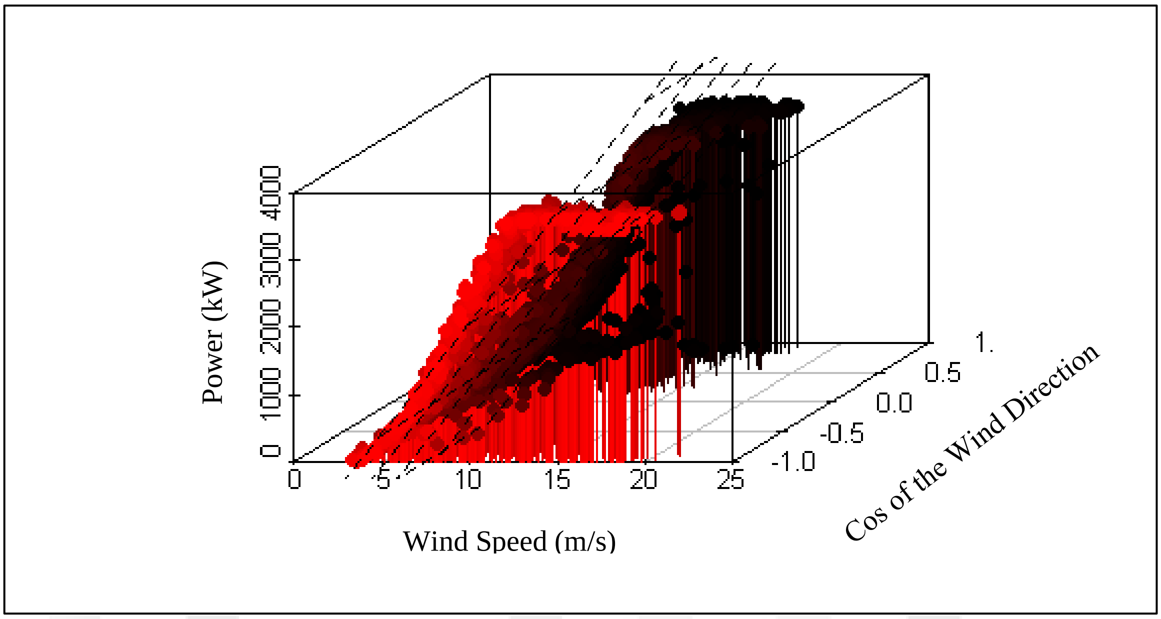

Figure 6. 6 The Multiple Linear Regression Plane of the Crisp Model.

The crisp multiple linear regression model is significant at 0.05 level and, the adjusted R-squared is 0.8509. Although an independent variable is added to the model, expected recovery at adjusted R-squared statics is not realized. Also, the non-normality problems of the errors continue. Consequently, it signifies that the crisp linear methods should be abandoned, and fuzzy linear methods should be adapted.

6.4 Implementation of Fuzzy Linear Regression Methods

In the subsection, the fuzzy linear regression models that are embedded in “fuzzyreg” package in R and detailed in Table 6.2 are implemented according to the plan in Figure 6.3.

6.4.1 Implementation of FLR Methods to Model 1 with CICO Dataset

As known crisp numbers are also fuzzy numbers with zero spread to lower and upper bound, so they can be evaluated as fuzzy numbers. Hence, all mentioned methods are used in the section. Here, the original dataset, whose inputs and outputs are crisp, is used in the section without fuzzification.

h-values defines the desired reliability level by obtaining the width or narrowness of the fuzzy spread of parameters in fuzzy regression model [93]. It means when *h*-value increases the width of the upper and lower bounders from center increase. Three different *h*-values at

0.01, 0.25, and 0.5 that signify different fuzziness levels are applied to models. When 0.01 *h*-value indicates the least fuzziness, 0.5 *h*-value shows the most fuzziness in the study.

6.4.1.1 Lee and Tanaka Method

At 0.01 *h* level, the results fuzzy linear model using Lee and Tanaka method are given below. The coefficients of the model in form **nsTFN** are as in Table 6.5.

Table 6.5 Values of Model 1 with Lee and Tanaka Method and CICO dataset at 0.01 *h*-value.

	center	Left Spread	Right Spread
Intercept	-1205.85	0	251.62
X_1	328.82	0	71.82

The central tendency of the fuzzy regression model:

$$\tilde{Y} = -1205.85 + 328.82\, X_1$$

The lower boundary of the model support interval:

$$\tilde{Y}^{L} = -1205.85 + 77.07 X_1$$

The upper boundary of the model support interval:

$$\tilde{Y}^{U} = -954.23 + 400.64 X_1$$

The total error of fit (TEF) is calculated as $3.517522{\times}10^{12}$ and the mean squared between response and prediction is 7208626.

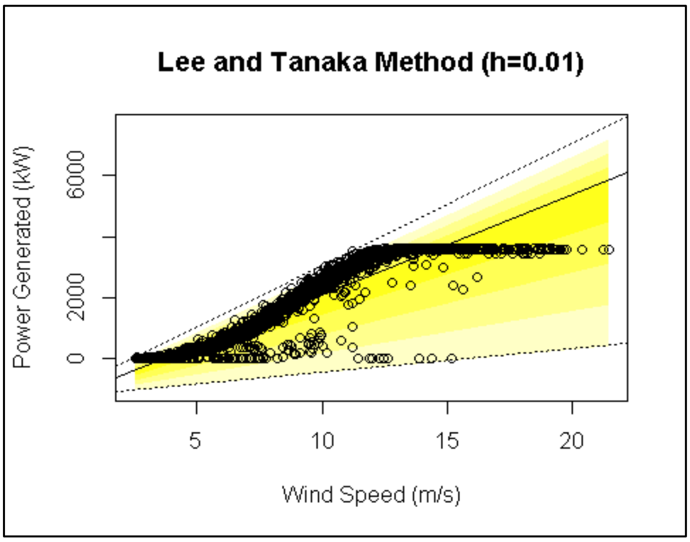

Figure 6. 7 Model 1 with Lee and Tanaka Method and CICO dataset (*h*=0.01).

The central tendency with bold line at the middle, lower and upper boundary with dashed lines are shown in Figure 6.7. As seen in the figure, Lee and Tanaka method inclines to include all observations between fuzzy upper and lower boundary. However, the outliers at the left or right spread do not the same distance from the center, the upper and lower bounds extend non-symmetrically. So, the parameters of the models $\widetilde{A_o}$ *and* $\widetilde{A_1}$ are non-symmetric triangular fuzzy numbers.

At 0.25 *h* level, the results fuzzy linear model using Lee and Tanaka method are given below. The coefficients of the model in form **nsTFN** are as in Table 6.6 :

Table 6.6 Values of Model 1 with Lee and Tanaka Method and CICO dataset at 0.25 *h*-value.

	center	Left Spread	Right Spread
Intercept	-1205.85	0	332.13
X1	328.82	332.31	94.8

The central tendency of the fuzzy regression model:

$$\tilde{Y} = -1205.85 + 328.82\, X_1$$

The lower boundary of the model support interval:

$$\tilde{Y}^L = -1205.85 - 3.5X_1$$

Upper boundary of the model support interval:

$$\tilde{Y}^U = -873.72 + 423.6X_1$$

The total error of fit (TEF) is calculated as 4.643129×10^{12} and the mean squared between response and prediction is 12008163.

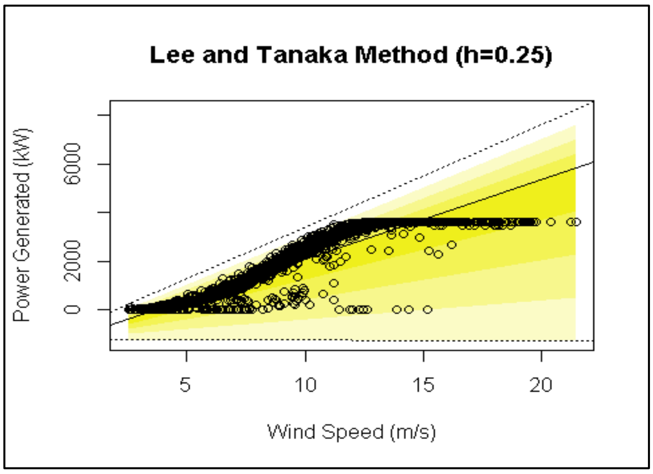

Figure 6. 8 Model 1 with Lee and Tanaka Method and CICO dataset (*h*=0.25)

The central tendency with bold line at the middle, lower and upper boundary with dashed lines are shown in Figure 6.8. The model with 0.25 *h*-value is fuzzier or has wider spreads than the model with 0.01 *h*-value depicted in Figure 6.7. Lee and Tanaka Method tends to include all observations into upper and lower boundary at minimum fuzziness level (h = 0.01). When the fuzziness increases, the whole fuzzy area that captures the observation increases unnecessarily, also the total error fit values and the mean squared distance values are increased depending on the increase of the *h*-values. Due to the tendency of the model to include all observations in the fuzzy area, usage of minimum *h*-value is evaluated to give accurate results.

6.4.1.2 Tanaka Method

At 0.01 *h* level, the results fuzzy linear model using Tanaka method are given below. The coefficients of the model in form **sTFN** are as in Table 6.7.

Table 6. 7 Values of Model 1 with Tanaka Method and CICO dataset at 0.01 *h* value.

	center	Left Spread	Right Spread
Intercept	-0.12	0	0
X_1	157.25	158.77	158.77

Central tendency of the fuzzy regression model:

$$\tilde{Y} = -0.12 + 157.25\, X_1$$

Lower boundary of the model support interval:

$$\tilde{Y}^{L} = -0.12 - 1.51X_1$$

Upper boundary of the model support interval:

$$\tilde{Y}^{U} = -0.12 + 316.02X_1$$

The total error of fit (TEF) is calculated as infinitive and the mean squared between response and prediction is 6623078.

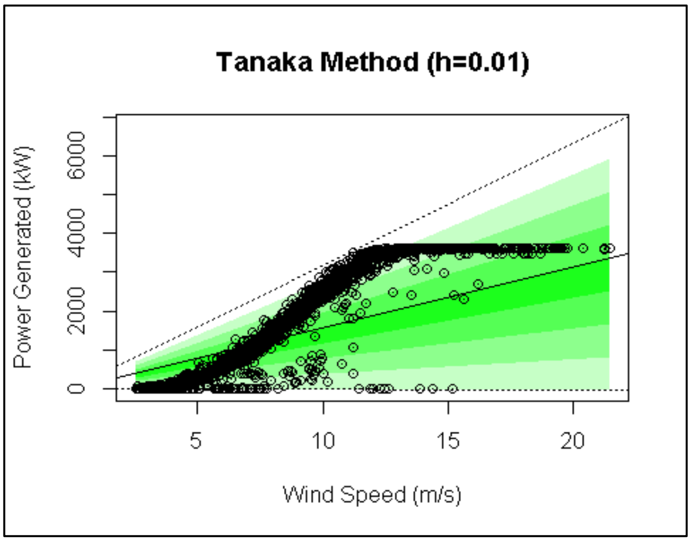

Figure 6. 9 Model 1 with Tanaka Method and CICO dataset (h=0.01)

The central tendency with bold line at the middle, lower and upper boundary with dashed lines are shown in Figure 6.9. As seen in the figure, Tanaka Method inclines to include all observation into spreads too. However, distinctly from Lee and Tanaka Method, the outliers at the left or right spread have the same distance from the center, the upper and lower bounds extend symmetrically. In other words, left and right spreads of intercept and independent variable are equal to each other. The parameters of the models $\widetilde{A_o}$ and $\widetilde{A_1}$ are symmetric triangular fuzzy numbers.

At 0.25 h level, the results fuzzy linear model using Tanaka method are given below. The coefficients of the model in form sTFN are as in Table 6.8 :

Table 6. 8 Values of Model 1 with Tanaka Method and CICO dataset at 0.25 h-value.

	center	Left Spread	Right Spread
Intercept	-0.12	0	0
X_1	157.25	209.57	209.57

Central tendency of the fuzzy regression model:

$\widetilde{Y} = -0.12 + 157.25\, X_1$

Lower boundary of the model support interval:

$\widetilde{Y}^{\text{L}} = -0.12 - 52.32 X_1$

Upper boundary of the model support interval:

$\widetilde{Y}^{\text{U}} = -0.12 + 366.83 X_1$

The total error of fit (TEF) is calculated as infinitive and the mean squared between response and prediction is 9947141.

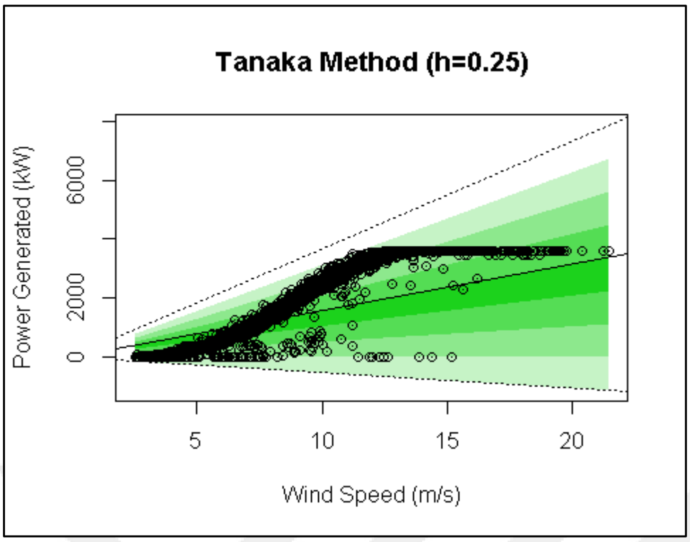

Figure 6. 10 Model 1 with Tanaka Method and CICO dataset (h=0.25)

The central tendency with bold line at the middle, lower and upper boundary with dashed lines are shown in Figure 6.10. The model with 0.25 h-value is fuzzier than the model with 0.01 h-value depicted in Figure 6.9. The model has also the tendency to include all observations into fuzzy spreads symmetrically. Thus, the model can be said fuzzier than the previous model with 0.01 h value. Due to the inclusion tendency of the model, it is also an efficient way to use minimum h-value like in Lee and Tanaka Method.

6.4.1.3 Nasrabadi Method

At 0.01 h level, the results fuzzy linear model using Nasrabadi method are given below. The coefficients of the model in form **sTFN** are as in Table 6.9 :

Table 6. 9 Values of Model 1 with Nasrabadi Method and CICO dataset at 0.25 h-value.

	center	Left Spread	Right Spread
Intercept	-1207.3	0	0
X_1	328.96	0	0

Central tendency of the fuzzy regression model:

$$\tilde{Y} = -1207.3 + 328.96\, X_1$$

Lower boundary of the model support interval:

$\widetilde{Y}^{\text{L}} = -1207.3 + 328.96X_1$

Upper boundary of the model support interval:

$\widetilde{Y}^{\text{U}} = -1207.3 + 328.96X_1$

The total error of fit (TEF) is calculated as infinitive and the mean squared between response and prediction is 743670.

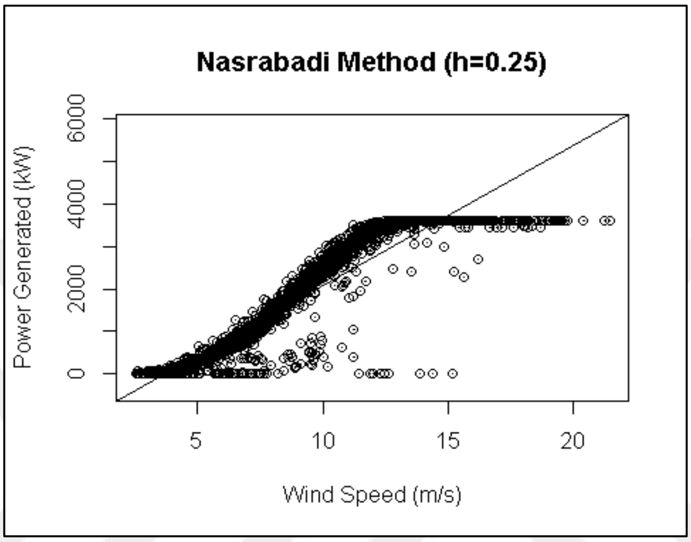

Figure 6. 11 Model 1 with Nasrabadi Method and CICO dataset (h=0.25)

The central tendency and lower and upper boundary collided with each other, and the borders can be seen with the bold line at the middle shown in Figure 6.11. As seen in the figure, Nasrabadi Method does not incline to include all observation contrary to Lee and Tanaka, Tanaka Method. However, the outliers at the left or right boundary have not spread from the center. So, the parameters of the models $\widetilde{A_o}$ *and* $\widetilde{A_1}$ are crisp numbers or symmetric triangular fuzzy numbers with a “0” spread.

Because the results of models with h=0.01 and h=0.5 gave us the same results with h=0.25, the results of the other h values were not put here. Due to the nature of the Nasrabadi Method and usage of crisp numbers for independent and dependent variables, the method behaved such it is a crisp least square method and gave the same results as it. Normally, the Nasrabadi method provided left and right spread with small values like 10^{-3}, but the spreads were dismissed due to being small values that do not affect the results. Three results with different h values have the same results.

6.4.1.4 Fuzzy Least Squares Method

As stated before, in the method *h* value is not used. The results of fuzzy linear model using Fuzzy Least Squares Method are given below. The Coefficients of the model in form of **sTFN** are as in Table 6.10 :

Table 6. 10 Values of Model 1 with Lee and Tanaka Method and CICO dataset.

	center	Left Spread	Right Spread
Intercept	-1207.3	0	0
X_1	328.96	0	0

Central tendency of the fuzzy regression model:

$\tilde{Y} = -1207.3 + 328.96 X_1$

Lower boundary of the model support interval:

$\tilde{Y}^{L} = -1207.3 + 328.96 X_1$

Upper boundary of the model support interval:

$\tilde{Y}^{U} = -1207.3 + 328.96 X_1$

The total error of fit (TEF) is calculated as infinitive and the mean squared between response and prediction is 743670.

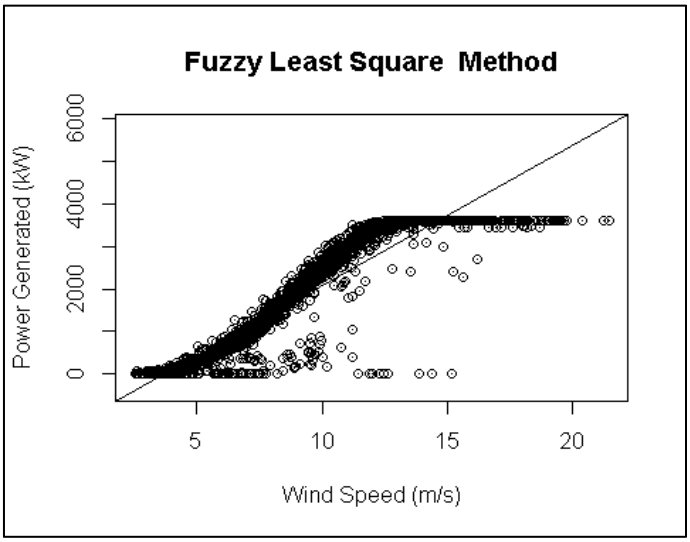

Figure 6. 12 Model 1 with FLS Method and CICO dataset.

The central tendency and lower and upper boundary collided with each other, and the borders can be seen with the bold line at the middle shown in Figure 6.12. As seen in the Figure 6.11

of the Nasrabadi Method, Fuzzy Least Square Method is not inclined to contain all observations. However, the outliers at the left or right boundary have not spread from the center. So, the parameters of the models $\widetilde{A_o}$ and $\widetilde{A_1}$ are crisp numbers.

The fuzzy Least Square Method panned out the same result with the Nasrabadi method and crisp least square method. The crisp values for independent and dependent variables are estimated to cause the situation.

6.4.1.5 Comparison of FLR Methods to Model 1 with CICO Dataset

The results of fuzzy linear regression models that are applied to crisp input and crisp output data with "wind speed" as an independent variable and "generated power" as a dependent variable (Model 1) are compared in Table 6.11 below. The total error fit (TEF) and the mean squared distance statistics are given for each method with *h*-values.

Table 6. 11 Comparison of FLR Methods with Model 1 and CICO dataset.

Method	*h*-val.	Total Error Fit	Mean Square Distance	Central		Left Sp.		Right Sp.	
				Intercept	X_1	Intercept	X_1	Intercept	X_1
Lee and Tanaka	**0.01**	3.5×10^{12}	7.2×10^{6}	-1205.9	328.8	0	0	251.6	72
	0.25	4.6×10^{12}	1.2×10^{7}	-1205.9	328.8	0	0	332	95
	0.5	7×10^{12}	2.6×10^{7}	-1205.9	328.8	0	499	498	142
Tanaka	**0.01**	∞	6.6×10^{6}	-0.1	157	0	158	0	158
	0.25	∞	1×10^{7}	-0.1	157	0	210	0	210
	0.5	∞	2×10^{7}	-0.1	157	0	314	0	314
Nasrabadi	**0.01**	∞	7.4×10^{5}	-1207	328.9	0	0	0	0
	0.25	∞	7.4×10^{5}	-1207	328.9	0	0	0	0
	0.5	∞	7.4×10^{5}	-1207	328.9	0	0	0	0
FLS	**-**	∞	7.4×10^{5}	-1207	328.9	0	0	0	0

When models are compared, the model with the least TEF or the mean squared distance (GOF) value should be preferred. After analyzing the results of models, due to assessment with the vast number of observations some methods gave infinitive total error fit value. Nasrabadi and FLS Methods gave the least Mean Square Distance value. However, these methods behaved like crisp linear regression models. If high fuzziness is not preferred, the models can be selected. But Tanaka or Lee and Tanaka Method with minimum *h*-value can opt when optimal fuzziness is wanted. Because GOF values of Lee and Tanaka Method and Tanaka Method are 7.2×10^{6} and 6.6×10^{6} with 0.01 *h*-value, respectively. These results are the minimum among the other options with fuzzy spreads.

6.4.2 Implementation of FLR Methods to Model 2 with CICO Dataset

In the section, the original dataset has crisp input and crisp output variables in the dataset without fuzzification. Lee and Tanaka method and Tanaka method are used in the section. Also, three different *h* values that signify different fuzziness levels are applied to models. These are h=0.01, h=0.25, and h=0.5 . However, the details and graphs of the *h* value with 0.25 and 0.5 are not given like in the previous section. The results of the h=0.25 and h=0.5 are given to compare with other *h*-value results. Because they gave similar but fuzzier results as h=0.01 value. Model 2 would be as given below:

$$\widetilde{Y} = \widetilde{A_o} + \widetilde{A_1}X_1 + \widetilde{A_2}X_2$$

6.4.2.1 Lee and Tanaka Method

At 0.01 *h* level, the results fuzzy linear model using Lee and Tanaka method are given below. The coefficients of the model in form n**sTFN** are as in Table 6.12.

Table 6. 12 Values of Model 2 with Lee and Tanaka Method and CICO dataset at 0.01 *h*-value.

	center	Left Spread	Right Spread
Intercept	-1216.94	0	262
X_1	329.5	253.4	59.87
X_2	28	0	105.4

Central tendency of the fuzzy regression model:

$$\widetilde{Y} = -1216.94 + 329.5X_1 + 28X_2$$

Lower boundary of the model support interval:

$$\widetilde{Y}^L = -1216.94 + 76.1X_1 + 28X_2$$

Upper boundary of the model support interval:

$$\widetilde{Y}^U = -954.94 + 389.37X_1 + 133.4X_2$$

The total error of fit (TEF) is calculated as infinitive and the mean squared between response and prediction is infinitive.

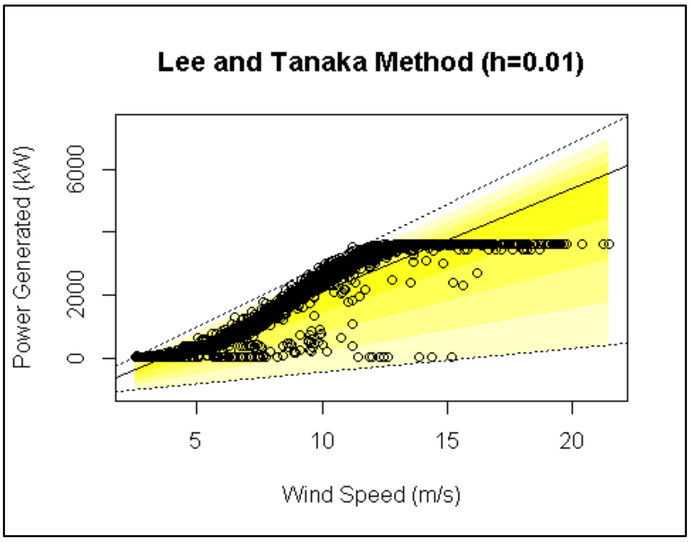

Figure 6. 13 Wind Speed vs. Power Generated in Lee and Tanaka Method with CICO dataset (*h*=0.01).

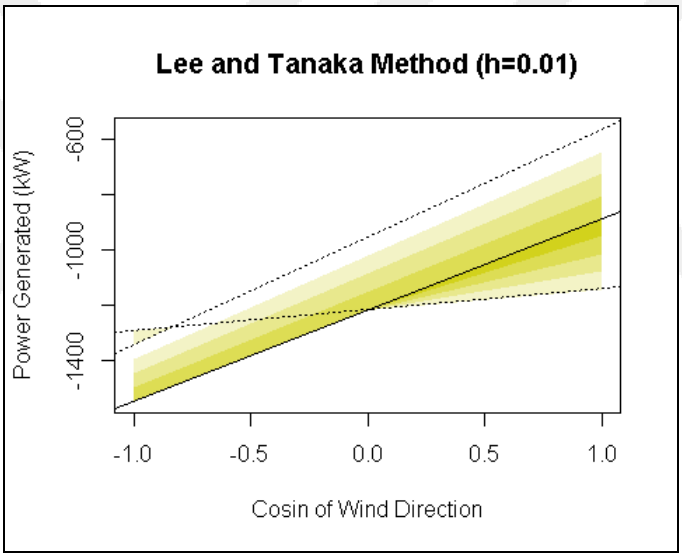

Figure 6. 14 Model 2 with Lee and Tanaka Method and CICO dataset (*h*=0.01).

The central tendency with bold line at the middle, lower and upper boundary with dashed lines are shown in Figure 6.13 and Figure 6.14. The outliers at the left or right spread have not the same distance from the center, the upper and lower bounds extend non-symmetrically for the first input "wind speed" and the second input. The parameters of the models $\widetilde{A_o}, \widetilde{A_1}$ *and* $\widetilde{A_2}$ are non-symmetric triangular fuzzy numbers. When Figure 6.13 and Figure 6.7 that belong to Model 1 resemble, on the other hand, Figure 6.14 cannot incline the observations. In every *h*-values, the second input cannot reflect the observation, as seen in Figure 6.14. Consequently, Lee and Tanaka's method does not give appropriate results for Model 2.

6.4.2.2 Tanaka Method

At 0.01 h level, the results fuzzy linear model using Tanaka method are given below. The coefficients of the model in form **sTFN** are as in Table 6.13 :

Table 6. 13 Values of Model 2 with Tanaka Method and CICO dataset at 0.01 h-value.

	center	Left Spread	Right Spread
Intercept	-0.7	0	0
X_1	151.4	152.9	152.9
X_2	69.6	71	71

Central tendency of the fuzzy regression model:

$$\tilde{Y} = -0.7 + 151.4\, X_1 + 69.6\, X_2$$

Lower boundary of the model support interval:

$$\tilde{Y}^{L} = -0.7 - 1.5X_1 - 1.4X_2$$

Upper boundary of the model support interval:

$$\tilde{Y}^{U} = -0.7 + 304.3X_1 + 140.6X_2$$

The total error of fit (TEF) is calculated as infinitive and the mean squared between response and prediction is infinitive.

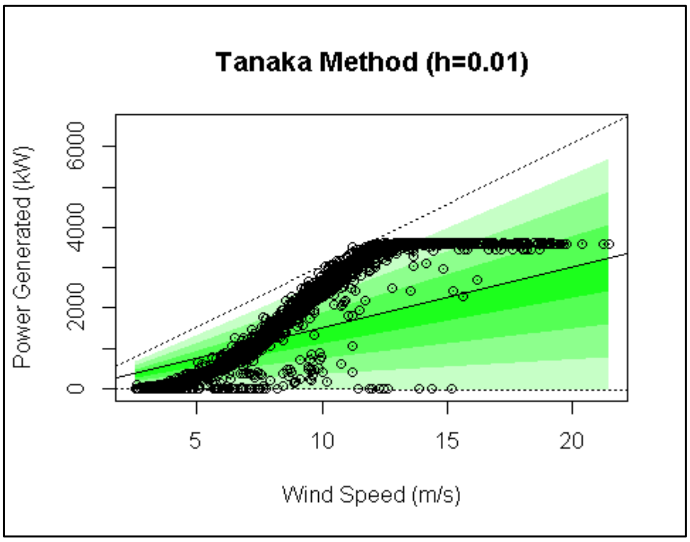

Figure 6. 15 Wind Speed vs. Power Generated in Tanaka Method with CICO dataset (h=0.01).

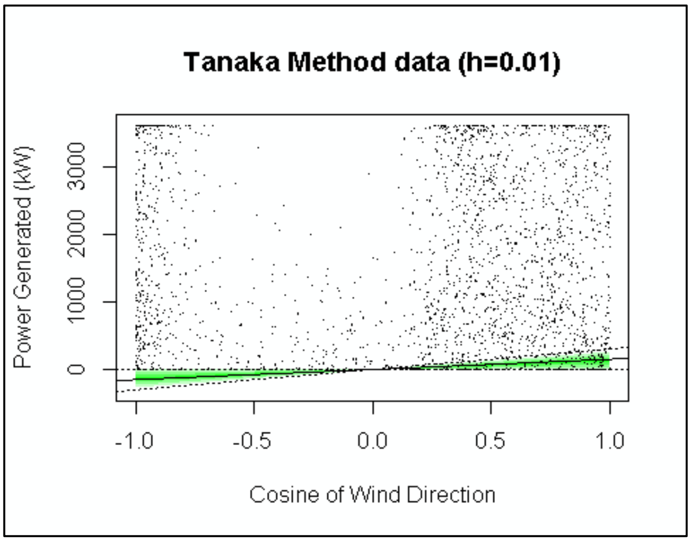

Figure 6. 16 Model 2 with Tanaka Method and CICO dataset (h=0.01).

The fuzzy regression lines are shown in Figure 6.15 and Figure 6.16. The outliers at the left or right spread have the same distance from the center, the parameters of the models $\widetilde{A_o}, \widetilde{A_1}\ and\ \widetilde{A_2}$ are symmetric triangular fuzzy numbers. Figure 6.16 can incline the observations contrary to Lee and Tanaka method. So, model 2 is more preferable to Lee and Tanaka method.

6.4.2.3 Comparison of FLR Methods to Model 2 with CICO Dataset

The results of fuzzy linear regression models that are applied to crisp input and crisp output data with "wind speed" and "cosine of wind direction "as independent variables and "generated power" as dependent variables (Model 2) are compared in Table 6.14 below.

Table 6. 14 Comparison of FLR Methods with Model 2 and CICO dataset.

Method	**h-val.**	**Total Error Fit**	**MSE**	**Central**			**Left Sp.**			**Right Sp.**		
				Intrcpt.	**X_1**	**X_2**	**Intrcpt.**	**X_1**	**X_2**	**Intrcpt.**	**X_1**	**X_2**
Lee and Tanaka	**0.01**	∞	∞	-1217	330	28	0	253	0	262	60	105
	0.25	∞	∞	-1217	330	28	0	335	0	346	79	139
	0.5	∞	∞	-1217	330	28	0	502	0	519	119	209
Tanaka	**0.01**	∞	∞	-0.7	151	70	0	153	71	0	153	71
	0.25	∞	∞	-0.5	151	70	0	202	94	0	202	94
	0.5	∞	∞	-0.2	151	70	0	303	141	0	303	141

After analyzing the results of models, due to assessment with the vast number of observations and two regressors, all methods gave infinitive total error fit and Mean Squared

Distance values. Nonetheless, Lee and Tanaka's method cannot include a second independent variable (cosine of wind direction) into the model. Consequently, the methods are not suggested when a vast number of observations are used with many regressors, but Lee and Tanaka Method is not strictly recommended.

6.4.3 Implementation of FLR Methods to Model 1 with CIFO Dataset

In the section, the original dataset, however, the dependent variable is fuzzified with the help of the theoretical output data. The symmetrical spreads of the dependent variables are the numeric difference between theoretical power and generated power and the generated power is accepted as a central tendency. The sample of CIFO data set is given in Table 6.15. Tanaka, Nasrabadi and, FLS methods are used in the section. Also, three different *h* values that signify different fuzziness levels are applied to models. However, the details and graphs of the *h* value with 0.5 are not given. The results of the *h*=0.5 are given to compare with other *h*-value results.

Table 6. 15 Sample CIFO Data Set.

Date/Time	**Theoretical Power (kW)**	**Power Generated (kW)**	**Symmetric Spread of Power Generated**	**Wind Speed (m/s)**	**Cosines of Wind Direction**
01-09-2018 00:00	3588.3	3404.1	184.2	12.6	0.303
01-09-2018 01:00	3478.1	3102.2	375.9	11.7	0.332
01-09-2018 02:00	3572.1	3222.7	349.4	12.4	0.273
⋮	⋮	⋮	⋮	⋮	⋮
31-12-2018 21:00	2601.1	2309.9	291.2	9.7	0.167
31 12 2018 22:00	3025.2	2681.3	343.9	10.4	0.166
31 12 2018 23:00	3583.3	3514.3	69.0	12.5	0.165

6.4.3.1 Tanaka Method

At 0.01 *h* level, the results fuzzy linear model using Tanaka method are given below. The coefficients of the model in form **sTFN** are as in Table 6.16.

Table 6. 16 Values of Model 1 with Tanaka Method and CIFO dataset at 0.01 *h*-value.

	center	Left Spread	Right Spread
Intercept	-1.1	0	0
X_1	157.3	158.8	158.8

Central tendency of the fuzzy regression model:

$\widetilde{Y} = -1.1 + 157.3\, X_1$

Lower boundary of the model support interval:

$\widetilde{Y}^{\mathrm{L}} = -1.1 - 1.5X_1$

Upper boundary of the model support interval:

$\widetilde{Y}^{\mathrm{U}} = -1.1 + 316.1X_1$

The total error of fit (TEF) is calculated as infinitive and the mean squared between response and prediction is 5970442.

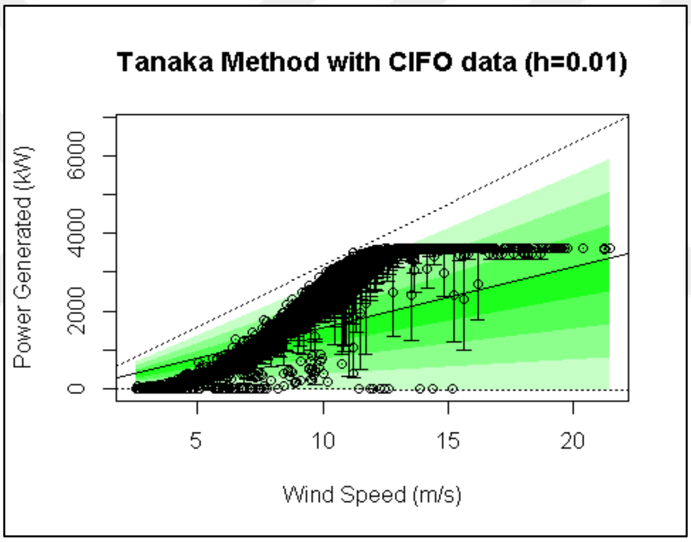

Figure 6. 17 Model 1 with Tanaka Method and CIFO dataset (h=0.01)

The fuzzy regression lines are shown in Figure 6.17. The outliers at the left or right spread has the same distance from the center, the upper and lower bounds extend symmetrically. So, the parameters of the models $\widetilde{A_o}$ *and* $\widetilde{A_1}$ are symmetric triangular fuzzy numbers. Also, the symbols dots at the middle with “T” shaped lines signify the fuzzy outputs that are also triangular fuzzy number.

At 0.25 *h* level, the results fuzzy linear model using Tanaka method are given below. The coefficients of the model in form **sTFN** are as in Table 6.17.

Table 6. 17 Values of Model 1 with Tanaka Method and CIFO dataset at 0.25 *h*-value.

	center	Left Spread	Right Spread
Intercept	-0.87	0	0
X_1	157.29	209.62	209.62

Central tendency of the fuzzy regression model:

$$\tilde{Y} = -0.87 + 157.29\, X_1$$

Lower boundary of the model support interval:

$$\tilde{Y}^{L} = -0.87 - 52.33 X_1$$

Upper boundary of the model support interval:

$$\tilde{Y}^{U} = -0.87 + 366.91 X_1$$

The total error of fit (TEF) is calculated as infinitive and the mean squared between response and prediction is 9058213.

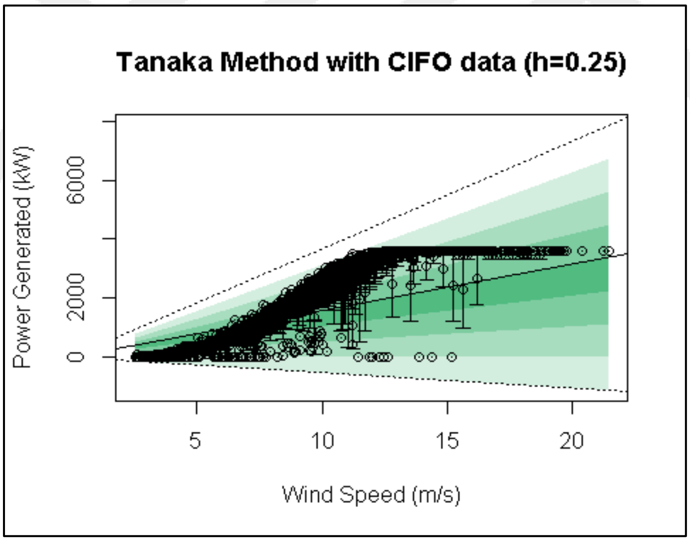

Figure 6. 18 Model 1 with Tanaka Method and CIFO dataset (*h*=0.25).

The central tendency with bold line at the middle, lower and upper boundary with dashed lines are shown in Figure 6.18. The model is fuzzier than previous model.

6.4.3.2 Nasrabadi Method

At 0.01 *h* level, the results fuzzy linear model using Nasrabadi method are given below. The coefficients of the model in form **sTFN** are as in Table 6. 18.

Table 6. 18 Values of Model 1 with Nasrabadi Method and CIFO dataset at 0.01 h-value.

	center	Left Spread	Right Spread
Intercept	-1206.86	99.2	99.2
X_1	328.92	16.8	16.8

Central tendency of the fuzzy regression model:

$\tilde{Y} = -1206.86 + 328.92\, X_1$

Lower boundary of the model support interval:

$\tilde{Y}^{L} = -1306.1 + 312.12 X_1$

Upper boundary of the model support interval:

$\tilde{Y}^{U} = -1107.65 + 345.73 X_1$

The total error of fit (TEF) is calculated as infinitive and the mean squared between response and prediction is 830139.

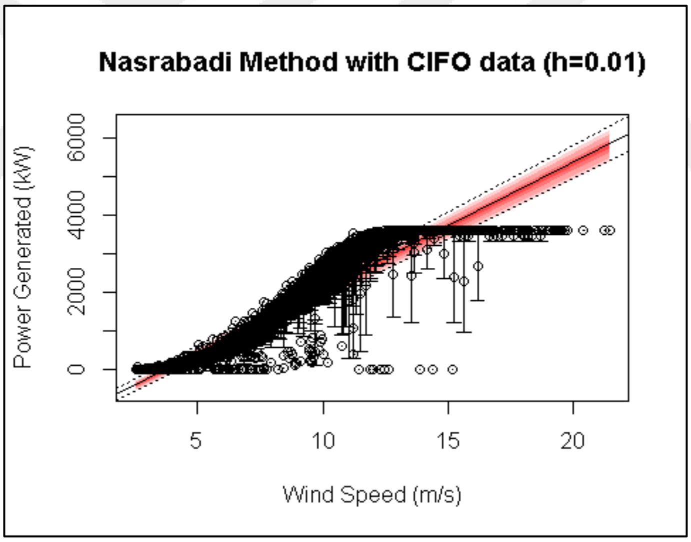

Figure 6. 19 Model 1 with Nasrabadi Method and CIFO dataset (h=0.01).

The central tendency with bold line at the middle, lower and upper boundary with dashed lines are shown in Figure 6.19. As seen in the figure, Nasrabadi Method does not incline to include all observation into spreads. Nasrabadi Method specifies optimal spread and does not prefer to put all observation into the fuzzy area. The outliers at the left or right spread have the same distance from the center, the upper and lower bounds extend symmetrically like in Tanaka Method. In other words, left and right spreads of intercept and independent variable are equal to each other. The parameters of the models $\widetilde{A_o}$ and $\widetilde{A_1}$ are symmetric triangular fuzzy numbers.

At 0.25 h level, the results fuzzy linear model using Nasrabadi method are given below. The coefficients of the model in form **sTFN** are as in Table 6.19 :

Table 6. 19 Values of Model 1 with Tanaka Method and CIFO dataset at 0.25 h-value.

	center	Left Spread	Right Spread
Intercept	-1207.3	130.96	130.96
X_1	328.96	22.18	22.18

Central tendency of the fuzzy regression model:

$$\tilde{Y} = -1207.3 + 328.96\, X_1$$

Lower boundary of the model support interval:

$$\tilde{Y}^{L} = -1338.26 + 306.78 X_1$$

Upper boundary of the model support interval:

$$\tilde{Y}^{U} = -1076.34 + 351.14 X_1$$

The total error of fit (TEF) is calculated as infinitive and the mean squared between response and prediction is 884561.

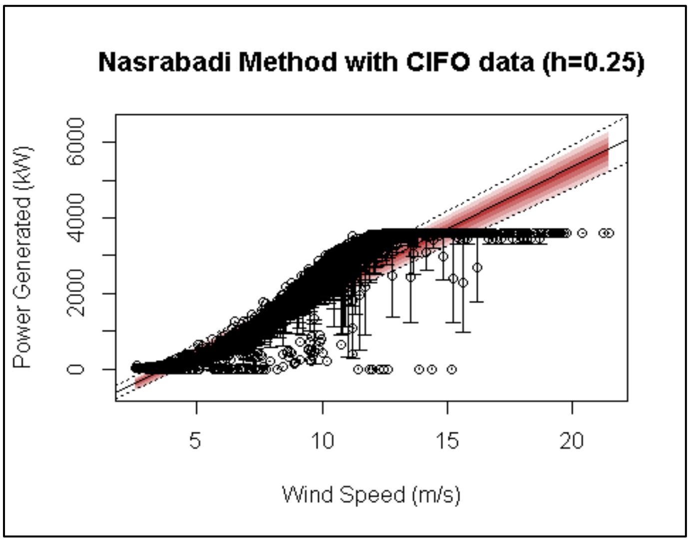

Figure 6. 20 Model 1 with Nasrabadi Method and CIFO dataset (h=0.25).

The fuzzy regression lines are shown in Figure 6.20. The model is fuzzier than the previous model. When h-value is increased, the spreads from the center are increased.

6.4.3.3 Fuzzy Least Square Method

The results of the Fuzzy Least Squares Method are given below. The Coefficients of the model are in form of **sTFN** due to observations that have symmetric spreads. The spreads of the Values are as in Table 6.20 :

Table 6. 20 Values of Model 1 with FLS Method and CIFO dataset.

	center	Left Spread	Right Spread
Intercept	-1207.3	49.1	49.1
X_1	328.96	8.32	8.32

Central tendency of the fuzzy regression model:

$$\tilde{Y} = -1207.3 + 328.96\, X_1$$

Lower boundary of the model support interval:

$$\tilde{Y}^{L} = -1256.4 + 320.65 X_1$$

Upper boundary of the model support interval:

$$\tilde{Y}^{U} = -1158.2 + 337.28 X_1$$

The total error of fit (TEF) is calculated as infinitive and the mean squared between response and prediction is 29642706.

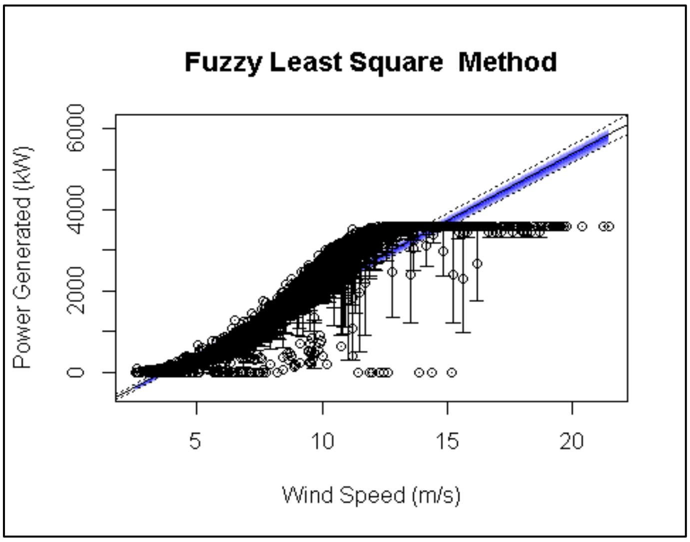

Figure 6. 21 Model 1 with FLS Method and CIFO dataset.

The fuzzy regression lines are shown in Figure 6.21. As seen in the figure Fuzzy Least Square Method does not tend to include all observation. The parameters of the models $\widetilde{A_o}$ *and* $\widetilde{A_1}$ are symmetric triangular fuzzy numbers.

6.4.3.4 Comparison of FLR Methods to Model 1 with CIFO Dataset

Table 6. 21 Comparison of FLR Methods to Model 1 with CIFO Dataset.

Method	h-val.	Total Error Fit	Mean Square Distance	Central		Left Sp.		Right Sp.	
				Intercept	X_1	Intercept	X_1	Intercept	X_1
Tanaka	**0.01**	∞	6×10^6	-1	157	0	159	0	159
	0.25	∞	9×10^6	-0.9	157	0	210	0	210
	0.5	∞	1.8×10^7	-0.6	157	0	314	0	314
Nasrabadi	**0.01**	∞	8.3×10^5	-1207	328.9	100	17	100	17
	0.25	∞	8.8×10^5	-1207	328.9	131	22	131	22
	0.5	∞	1.1×10^6	-1207	328.9	196	33	196	33
Diamond	-	∞	3×10^7	-1207	328.9	49	8	49	8

After analyzing the results of models, due to assessment with the vast number of observations all methods gave infinitive total error fit value. Because Mean Square Distance values of Nasrabadi Method are lower than the other methods. The Nasrabadi method can be recommended when outputs are fuzzy.

6.4.4 Implementation of FLR Methods to Model 2 with CIFO Dataset

Tanaka and Nasrabadi Methods are used in the section. Also, three different *h* values that signify different fuzziness levels are applied to models. However, the details and graphs of the *h* value with 0.25 and 0.5 are not given. Because they gave similar but fuzzier results as h=0.01 value.

6.4.4.1 Tanaka Method

At 0.01 *h* level, the results fuzzy linear model using Tanaka method are given below. The coefficients of the model in form **sTFN** are as in Table 6.22.

Table 6. 22 Values of Model 2 with Tanaka Method and CIFO dataset at 0.01 *h*-value.

	center	Left Spread	Right Spread
Intercept	-0.7	0	0
X_1	151	152	152
X_2	69	71	71

Central tendency of the fuzzy regression model:

$$\tilde{Y} = -0.7 + 151X_1 + 69X_2$$

Lower boundary of the model support interval:

$$\tilde{Y}^{L} = -0.7 - 1X_1 - 2X_2$$

Upper boundary of the model support interval:

$\tilde{Y}^{U} = -0.7 + 303X_1 + 140X_2$

The total error of fit (TEF) is calculated as infinitive and the mean squared between response and prediction is infinitive.

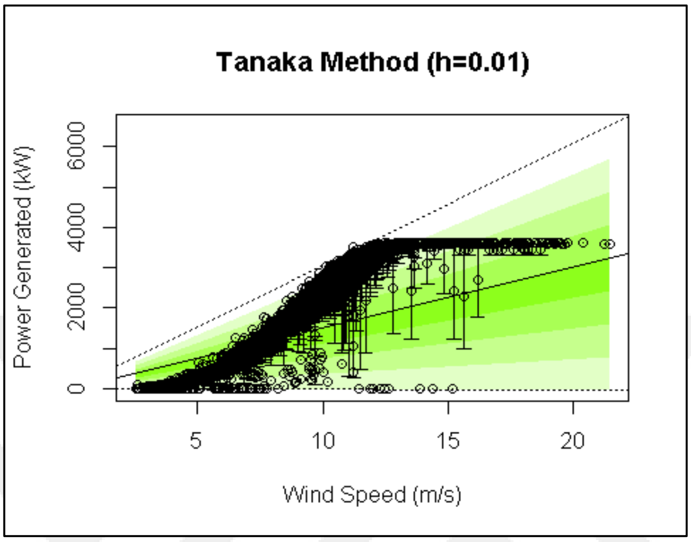

Figure 6. 22 Wind Speed vs. Power Generated in Tanaka Method with CIFO dataset (h=0.01)

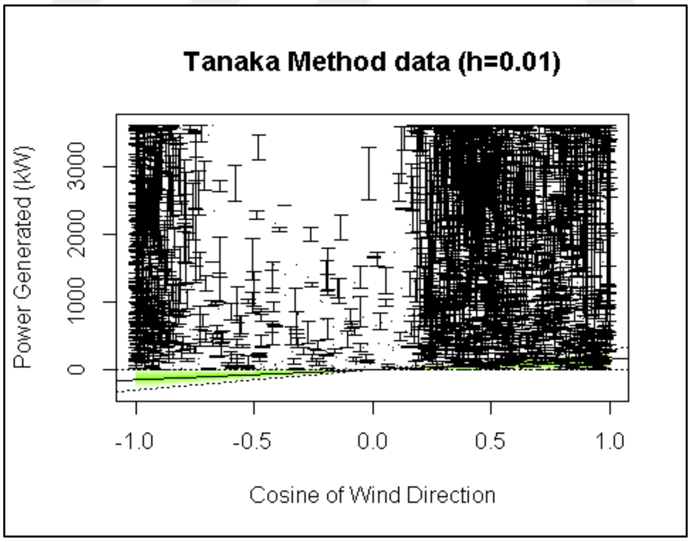

Figure 6. 23 Model 2 with Tanaka Method and CIFO dataset (h=0.01)

The fuzzy regression lines are shown in Figure 6.22 and Figure 6.23. The parameters of the models $\widetilde{A_o}, \widetilde{A_1}$ and $\widetilde{A_2}$ are symmetric triangular fuzzy numbers. Figure 6.23 can incline the observations partially.

6.4.4.2 Nasrabadi Method

At 0.01 *h* level, the results fuzzy linear model using Nasrabadi method are given below. The coefficients of the model in form **sTFN** are as in Table 6.23.

Table 6. 23 Values of Model 2 with Nasrabadi Method and CIFO dataset at 0.01 *h* value.

	center	Left Spread	Right Spread
Intercept	-1218	176	176
X_1	329.6	20	20
X_2	28	-156	-156

Central tendency of the fuzzy regression model:

$$\tilde{Y} = -1218 + 329.6X_1 + 28X_2$$

Lower boundary of the model support interval:

$$\tilde{Y}^{\mathrm{L}} = -1394 + 309.6X_1 + 184X_2$$

Upper boundary of the model support interval:

$$\tilde{Y}^{\mathrm{U}} = -1042 + 349.6X_1 - 128X_2$$

The total error of fit (TEF) is calculated as infinitive and the mean squared between response and prediction is infinitive.

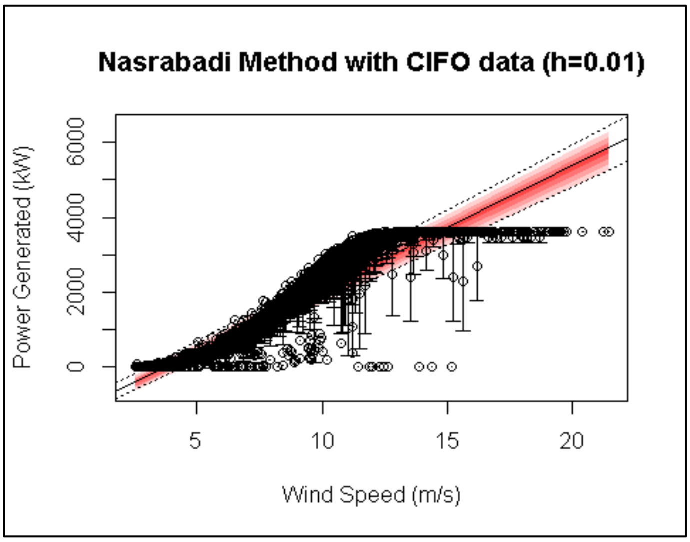

Figure 6. 24 Wind Speed vs. Power Generated in Nasrabadi Method with CIFO dataset (h=0.01).

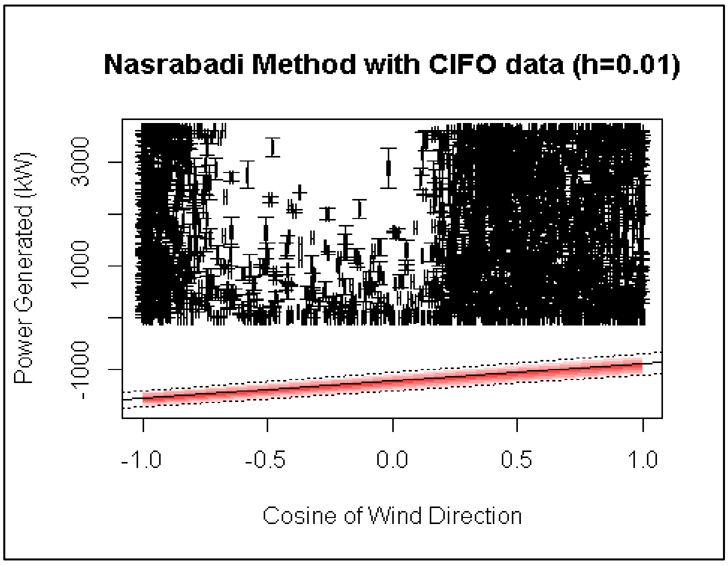

Figure 6. 25 Model 2 with Nasrabadi Method and CIFO dataset (h=0.01).

The central tendency with bold line at the middle, lower and upper boundary with dashed lines are shown in Figure 6.24 and Figure 6.25. As seen in the figures, Nasrabadi Method does not incline to include all observation into spreads. The parameters of the models $\widetilde{A_o}, \widetilde{A_1}\ and\ \widetilde{A_2}$ are symmetric triangular fuzzy numbers. However, Figure 6.25 cannot incline the observations. It means second independent variable is not effective to fit the model. So, Nasrabadi Method is not suggested for Model 2.

6.4.4.3 Comparison of FLR Methods to Model 2 with CIFO Dataset

The results of FLR models that are applied to crisp input and fuzzy output data with “wind speed” and “cosine of wind direction ”as independent variables and “generated power” as dependent variable (Model 2) are compared in Table 6.24.

Table 6. 24 Comparison of FLR Methods to Model 2 with CIFO Dataset.

Method	**h-val.**	**Total Error Fit**	**MSE**	**Central**			**Left Sp.**			**Right Sp.**		
				Intrcpt.	**X_1**	**X_2**	**Intrcpt.**	**X_1**	**X_2**	**Intrcpt.**	**X_1**	**X_2**
Tanaka	**0.01**	∞	∞	0	151	69	0	152	71	0	152	71
	0.25	∞	∞	0	151	69	0	201	94	0	201	94
	0.5	∞	∞	0	151	69	0	302	141	0	302	141
Nasrabadi	**0.01**	∞	∞	-1218	330	28	176	20	-156	176	20	-156
	0.25	∞	∞	-1218	330	28	232	26	-206	232	26	-206
	0.5	∞	∞	-1218	330	28	348	39	-309	348	39	-309

After analyzing the results of models, due to assessment with the vast number of observations and two regressors, all methods gave infinitive total error fit and Mean Squared Distance values. Nonetheless, the Nasrabadi method cannot include a second independent variable (cosine of wind direction) into the model. Consequently, the methods are not suggested when a vast number of observations are used with many regressors, but Tanaka Method looks more preferable to another.

6.4.5 Implementation of FLR Methods to Model 1 with FIFO Dataset

The dependent variable in the original dataset was fuzzified with the help of the theoretical output data. The symmetrical spreads of the dependent variables are the numeric difference between theoretical power and generated power and the generated power is accepted as a central tendency. The main input, wind speed, is symmetrically fuzzified with 10 % below and above the observed value. The sample of FIFO data set is given in Table 6.25. Only Nasrabadi Method is used in the section. Also, three different *h* values that signify different fuzziness levels are applied to the model. The results of 0.01 and 0.25 *h*-values are given.

Table 6. 25 The sample of FIFO data set.

Date/Time	**Theoretical Power (kW)**	**Power Generated (kW)**	**Symmetric Spread of Power Generated**	**Wind Speed (m/s)**	**Symmetric Spread of Wind Speed (%10)**
01-09-2018 00:00	3588.3	3404.1	184.2	12.6	1.26
01-09-2018 01:00	3478.1	3102.2	375.9	11.7	11.7
01-09-2018 02:00	3572.1	3222.7	349.4	12.4	1.24
⋮	⋮	⋮	⋮	⋮	⋮
31-12-2018 21:00	2601.1	2309.9	291.2	9.7	0.97
31 12 2018 22:00	3025.2	2681.3	343.9	10.4	1.04
31 12 2018 23:00	3583.3	3514.3	69.0	12.5	1.25

6.4.5.1 Nasrabadi Method

At 0.01 *h* level, the results fuzzy linear model using Nasrabadi method are given below. The coefficients of the model in form **sTFN** are as in Table 6.26.

Table 6. 26 Values of Model 1 with Nasrabadi Method and FIFO dataset at 0.01 h-value.

	center	Left Spread	Right Spread
Intercept	-84	-32	-32
X_1	200	10	10

Central tendency of the fuzzy regression model:

$\widetilde{Y} = -84 + 200 X_1$

Lower boundary of the model support interval:

$\widetilde{Y}^{L} = -52 + 190X_1$

Upper boundary of the model support interval:

$\widetilde{Y}^{U} = -116 + 210X_1$

The total error of fit (TEF) is calculated as 216964 and the mean squared between response and prediction is 502468.

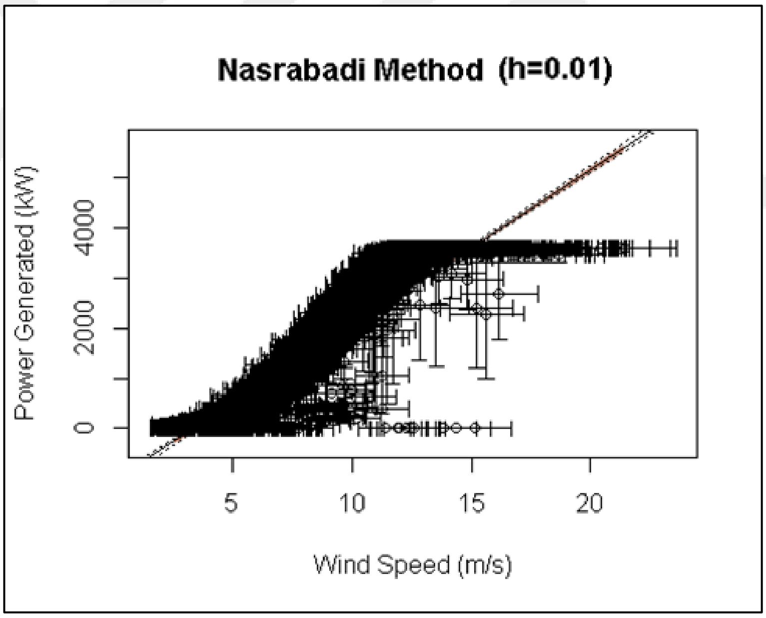

Figure 6. 26 Model 1 with Nasrabadi Method and FIFO dataset (h=0.01).

As seen in Figure 6.26, Nasrabadi Method does not incline to include all observation into spreads. Nasrabadi Method specifies optimal spread and does not prefer to put all observation into the fuzzy area. The parameters of the models $\widetilde{A_o}$ and $\widetilde{A_1}$ are symmetric triangular fuzzy numbers. Also, the symbols dots at the middle with “+” shaped lines signify FIFO data, the fuzzy inputs, and outputs that are also triangular fuzzy numbers.

At 0.01 h level, the results fuzzy linear model using Nasrabadi method are given below. The coefficients of the model in form **sTFN** are as in Table 6. 27.

Table 6. 27 Values of Model 1 with Nasrabadi Method and FIFO dataset at 0.25 *h*-value.

	center	Left Spread	Right Spread
Intercept	-1053	-88	-88
X_1	311	8	8

Central tendency of the fuzzy regression model:

$\tilde{Y} = -1053 + 311\, X_1$

Lower boundary of the model support interval:

$\tilde{Y}^{L} = -965 + 303X_1$

Upper boundary of the model support interval:

$\tilde{Y}^{U} = -1141 + 319X_1$

The total error of fit (TEF) is calculated as 262460 and the mean squared between response and prediction is 910200.

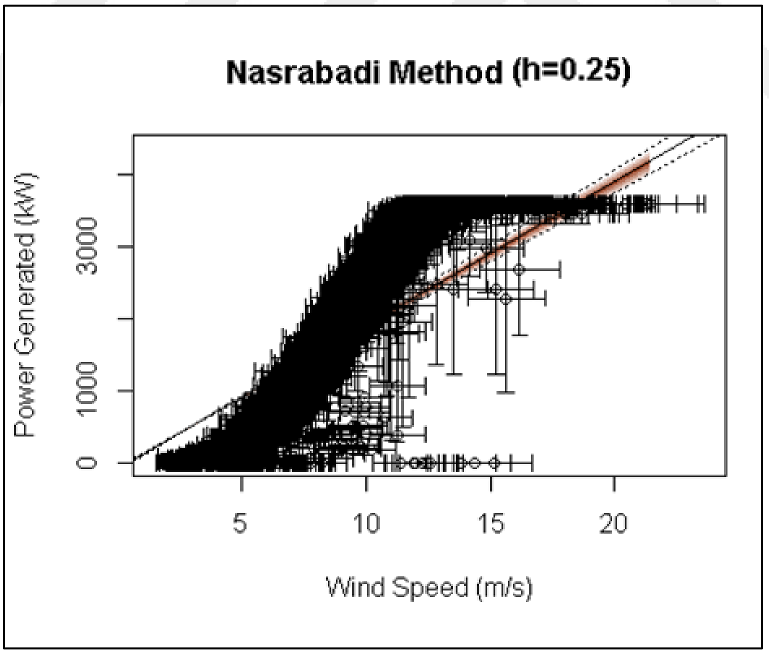

Figure 6. 27 Model 1 with Nasrabadi Method and CIFO dataset (*h*=0.25).

The fuzzy regression lines are shown in Figure 6.27. The model is fuzzier than the previous model. When *h*-value is increased, the spreads from the center are increased. In the section, when the dataset includes fuzzy input for fuzzy outputs, it is studied the applicability of the situation.

6.5 Results of Applied Models

In the section, the comparative two statistical indicators of all applied methods to all types of datasets are compared to see the whole picture of the study.

Table 6. 28 Comparison of all Applied Models.

Method	h-val.	DATA TYPES									
		CICO				CIFO				FIFO	
		Model 1		Model 2		Model 1		Model 2		Model 1	
Error Terms		TEF	GOF	TEF	GOF	TEF	GOF	TEF	GOF	TEF	GOF
Lee and Tanaka	**0.01**	3.5×10^{12}	7.2×10^{6}	∞	∞	-	-	-	-	-	-
	0.25	4.6×10^{12}	1.2×10^{7}	∞	∞	-	-	-	-	-	-
	0.5	7×10^{12}	2.6×10^{7}	∞	∞	-	-	-	-	-	-
Tanaka	**0.01**	∞	6.6×10^{6}	∞	∞	∞	6×10^{6}	∞	∞	-	-
	0.25	∞	1×10^{7}	∞	∞	∞	9×10^{6}	∞	∞	-	-
	0.5	∞	2×10^{7}	∞	∞	∞	1.8×10^{7}	∞	∞	-	-
Nasrabadi	**0.01**	∞	7.4×10^{5}	∞	∞	∞	8.3×10^{5}	∞	∞	2.2×10^{5}	9×10^{5}
	0.25	∞	7.4×10^{5}	∞	∞	∞	8.8×10^{5}	∞	∞	2.6×10^{5}	1×10^{6}
	0.5	∞	7.4×10^{5}	∞	∞	∞	1.1×10^{6}	∞	∞	4×10^{5}	1.5×10^{6}
FLS	**-**	∞	7.4×10^{5}	-	-	∞	3×10^{7}	-	-	-	-

CICO, CIFO, and FIFO datasets are used for modeling with different methods. It is an advantage for researchers to have the ability to use different datasets. Because, according to the situation, dependent or independent variables can be observed partially or observation can be mismeasured, and it can be needed to fuzzify the defective variables. Hence, different methods are applied to different datasets. All datasets are considered, Nasrabadi and FLS methods are generally more successful than the other methods. However, these methods are more complex for calculations. Nasrabadi and FLS methods do not include all observations into fuzzy upper and lower boundary, contrary to Lee and Tanaka, and Tanaka Methods.

In CICO dataset analysis, Nasrabadi and FLS behaved like crisp models, but it does not mean that they will behave like a crisp method in every CICO dataset. So, these models also can be applied to CICO dataset, for different cases. On the other, Lee and Tanaka, Tanaka methods can be applied to CICO dataset when fuzziness is desired. Because Lee and Tanaka's method has already been designed for CICO dataset. In brief, Lee, and Tanaka, Tanaka methods look more efficient with minimum *h*-value in CICO dataset case, due to other models' crispness. In other words, in a case whose observations are distributed proximately linear, Lee and Tanaka, Tanaka methods would be more powerful and suitable.

When the results of Model 2 are compared, the descriptive statistics are not distinguishing. However, the Tanaka method is more preferable to Lee and Tanaka method. Because Tanaka Method included a second independent variable when Lee and Tanaka's method did not.

In CIFO dataset, the applied methods in minimum *h*-value gave better results. But Nasrabadi Method is the winner of the case with a minimum GOF value. When the results of Model 2 are compared, the descriptive statistics are not distinctive too. However, the Tanaka method is more preferable to Nasrabadi Method. Because Tanaka Method included second independent variable again. Consequently, when there is more than one independent variable, Tanaka Method can be more accurate than the other methods. The Nasrabadi Method applied to FIFO dataset in the study. The results showed that Nasrabadi Method is appropriate for FIFO dataset cases.

7 CONCLUSION

Gathering data or observing parameters for all types of statistical analysis is always demanding and effortful. The processes are generally expensive and collecting accurate data is a probable problem. Besides the significant advantages of wind energy, it is emphasized that the high construction costs of wind farms are one of the most important disadvantages. So, wind energy investors are obliged to have detailed, precise, accurate, inclusive research done to professionals. Furthermore, all possible models should apply to observations. Otherwise, faults in observations or in models will cause a waste of economic resources.

It is aforementioned that numerous studies have been generated and many different statistical methods have been applied to predict accurate generated power by wind turbines, due to the importance of wind energy investments. The proposed methods to forecast power generation are frequently complicated. As various conditions influence approximate observations, parametric or non-parametric methods are used for the aim. The main parametric methods are the Linearized Model, Polynomial Model, Probabilistic Model, and Logistic function model. The main non-parametric methods are Neural Networks, Data Mining Algorithms, and Fuzzy Clustering Methods [11].

While there is a relationship between wind speed and produced electrical power, there are also unexplainable factors that influence wind speed and power production or their relationship. The study was motivated by the dataset's ambiguous context, which led to the use of fuzzy regression. Due to the fact that fuzzy regression methods do not take account of the distribution of observations and a fuzzy relation is assumed between wind speed, wind direction, and power generation. Also, the application of fuzzy set theory is naturally common in such subjects. The other motivation source to apply fuzzy linear regression is having a dataset that includes just four months of observation or seasonal (autumn) not annual. The annual dataset is obviously more available for crisp modeling, however fuzzy regression is more efficient for seasonal, limited, or partial datasets. Thus, the situations fulfilled the assumption to apply fuzzy regression modeling.

In the study, two models were proposed to estimate the amount of produced electrical energy by a wind turbine. The first model used only wind speed as a regressor, the second model used wind speed and wind direction as regressors. The original dataset had crisp input and crisp output. However, firstly the output of the dataset was fuzzified, so the CIFO dataset was generated. Secondly, both output and the first input (wind speed) were fuzzified, so the FIFO dataset was generated. Thereby, the applicability of the fuzzy regression models in the different datasets was evaluated.

Lee and Tanaka, Tanaka, Nasrabadi, and FLS method were applied to datasets and their results were compared. Hereby, the behavior and suitability of the fuzzy regression models were observed. When the results were examined, all suggested methods were quite successful for CICO dataset for Model 1. If the distribution of the observations is close to the normal distribution, the inclusive methods (Tanaka, Lee and Tanaka) would be more successful. In CIFO dataset for Model 1, the applied methods in minimum *h*-value gave proper results. But Nasrabadi Method gave the best results, so the method proved the accessibility in the situation.

The second regressor (wind direction) , which was not an explanatory variable in the crisp linear regression model is applied in the study. Due to not having another regressor in dataset, it was compulsory to use wind direction. So, it was a limitation for the study. The all proposed fuzzy regression method except Tanaka Method were ineffective for Model 2 in every dataset. The Tanaka method included the second regressor in the fuzzy area at least.

Accordingly, it is evaluated in the study that the fuzzy regression methods to predict the power generation are especially influential and advantageous when a general frame of the wind turbine models is wanted to realize before complex calculations and modeling for a place, or there are a small number of observations, or it is hard to observe the parameters or the observations are not trustworthy, or there are more optional places to construct a wind farm and decision-maker do not have sufficient time to take action.

Consequently, the suggested models provided alternative solutions for different situations, as well as a more flexible decision area and a theoretical basis for wind energy investors and researchers. This study will expand the horizon for studies such as the fuzzy piecewise regression method and fuzzy nonlinear regression methods that are expected to be used in the future in estimating the energy produced by wind turbines.

REFERENCES

[1] H. Ritchie and M. Roser, “Energy,” *Our World in Data*, Nov. 2020, Accessed: May 16, 2021. [Online]. Available: https://ourworldindata.org/energy.

[2] A. N. Penna, *A History of energy flows: From human labor to renewable power*. Taylor and Francis, 2019.

[3] X. Wu *et al.*, “Foundations of offshore wind turbines: A review,” *Renewable and Sustainable Energy Reviews*, vol. 104, pp. 379–393, Apr. 2019, doi: 10.1016/j.rser.2019.01.012.

[4] GWEC, *Global Wind Report 2021*. 2021, p. 75.

[5] WindEurope, “Wind energy in Europe in 2020 Trends and statistics,” Brussels, 2020.

[6] B. A. Schuelke-Leech, B. Barry, M. Muratori, and B. J. Yurkovich, “Big Data issues and opportunities for electric utilities,” *Renewable and Sustainable Energy Reviews*, vol. 52, no. 52. Elsevier Ltd, pp. 937–947, Aug. 22, 2015, doi: 10.1016/j.rser.2015.07.128.

[7] C. Croonenbroeck and G. Stadtmann, “Minimizing asymmetric loss in medium-term wind power forecasting,” *Renewable Energy*, vol. 81, pp. 197–208, Sep. 2015, doi: 10.1016/j.renene.2015.03.049.

[8] Y. Feng *et al.*, “Overview of wind power generation in China: Status and development,” *Renewable and Sustainable Energy Reviews*, vol. 50. Elsevier Ltd, pp. 847–858, Jun. 04, 2015, doi: 10.1016/j.rser.2015.05.005.

[9] C. Stathopoulos, A. Kaperoni, G. Galanis, and G. Kallos, “Wind power prediction based on numerical and statistical models,” *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 112, pp. 25–38, Jan. 2013, doi: 10.1016/j.jweia.2012.09.004.

[10] Q. Dong, Y. Sun, and P. Li, “A novel forecasting model based on a hybrid processing strategy and an optimized local linear fuzzy neural network to make wind power forecasting: A case study of wind farms in China,” *Renewable Energy*, vol. 102, pp. 241–257, 2017, doi: 10.1016/j.renene.2016.10.030.

[11] M. Lydia, S. S. Kumar, A. I. Selvakumar, and G. E. Prem Kumar, “A comprehensive review on wind turbine power curve modeling techniques,” *Renewable and Sustainable Energy Reviews*, vol. 30. Pergamon, pp. 452–460, Feb. 01, 2014, doi: 10.1016/j.rser.2013.10.030.

[12] C. Montgomery, Douglas and C. Runger, Runger, *Applied Statistics and Probability for Engineers*, Sixth. Wiley, 2014.

[13] J. M. Stanton, “Galton, Pearson, and the Peas: A Brief History of Linear Regression for Statistics Instructors,” *Journal of Statistics Education*, 2001, doi: 10.1080/10691898.2001.11910537.

[14] C. Montgomery, Douglas, E. A. Peck, and G. Vining, G., *Introduction to Linear Regression*. Wiley, 2012.

[15] A. Sen and M. Srivastava, *Regression Analysis Theory, Methods, and Applications*. New York: Springer, 1990.

[16] R. Alpar, *Uygulamalı Çok Değişkenli İstatistiksel Yöntemlere Giriş I*. Ankara: Nobel Yayın Dağıtım, 2003.

[17] W. W. J Neter, MH Kutner, CJ Nachtsheim, "Applied Linear Statistical Models. Fourth Edition," *Journal of Education*, vol. 36, no. 3, pp. 59–60, 1996, doi: 10.1177/002205749203600311.

[18] B. David and Dodge Yadolah, *Alternative Methods of Regression*. Wiley, 1993.

[19] S. Chatterjee and J. S. Simonoff, *Handbook of Regression Analysis*. Haboken, New Jersey: John Wiley and Sons, Inc., 2013.

[20] S. R. Jammaladaka and A. SenGupta, *Topics in Circular Statistics*, Series on. Singapore: World Scientific Publishing Co. Ptc.Ltd., 2001.

[21] E. J. Gumbel, J. A. Greenwood, and D. Durand, "The Circular Normal Distribution: Theory and Tables," *Journal of the American Statistical Association*, vol. 48, no. 261, pp. 131–152, 1953, doi: 10.1080/01621459.1953.10483462.

[22] A. L. Gould, "A Regression Technique for Angular Variates," *Biometrics*, vol. 25, no. 4, p. 683, Dec. 1969, doi: 10.2307/2528567.

[23] R. A. Johnson and T. E. Wehrly, "Some angular-linear distributions and related regression models," *Journal of the American Statistical Association*, 1978, doi: 10.1080/01621459.1978.10480062.

[24] K. v. Mardia, "9 Tests of unvariate and multivariate normality," *Handbook of Statistics*, vol. 1. Elsevier, pp. 279–320, Jan. 01, 1980, doi: 10.1016/S0169-7161(80)01011-5.

[25] U. Lund, "Least circular distance regression for directional data," *Journal of Applied Statistics*, vol. 26, no. 6, pp. 723–733, 1999, doi: 10.1080/02664769922160.

[26] S. Jammalamadaka and YR. Sarma, "Circular regression." VSP, pp. 109–128, 1993, Accessed: May 11, 2021. [Online].

[27] K. v. Mardia and P. E. Jupp, *Directional Statistics*. 2000.

[28] B. T. D. Downs and K.V Mardia, "Circular regression," 2002. [Online]. Available: https://academic.oup.com/biomet/article/89/3/683/252192.

[29] Mardia K. V., "Statistics of Directional Data," *Probability and mathematical statistics*, vol. 13, 1972, Accessed: May 11, 2021. [Online]. Available: https://www.scirp.org/(S(i43dyn45teexjx455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=386171.

[30] A. Jammalamadaka, S. Rao and SenGupta, *Topics in Circular Statistics*. 2001.

[31] C. Ley and T. Verdebout, *Modern directional statistics*. CRC Press, 2017.

[32] A. Jammalamadaka, S. Rao, and Y. R. Sarma, "Statistical Theory and Data Analysis II," 1988.

[33] M. I. Nurhab, B. Nurhab, T. Purwaningsih, and M. F. Teng, "Circular(2)-linear regression analysis with iteration order manipulation," *International Journal of Advances in Intelligent Informatics*, 2017, doi: 10.26555/ijain.v3i2.90.

[34] S. Kim and A. SenGupta, "Inverse Circular-Linear/Linear-Circular Regression," *Commun. Stat. Methods*, vol. 44, no. 22, pp. 4772–4782, 2015.

[35] L. A. Zadeh, "Fuzzy sets," *Information and Control*, 1965, doi: 10.1016/S0019-9958(65)90241-X.

[36] Z. Şen, "Fuzzy algorithm for estimation of solar irradiation from sunshine duration," *Solar Energy*, 1998, doi: 10.1016/S0038-092X(98)00043-7.

[37] V. B. Rao, *C++ Neural Networks and Fuzzy Logic: Introduction to Neural Networks*. 1995.

[38] L. A. Zadeh, "Fuzzy logic = computing with words," *IEEE Transactions on Fuzzy Systems*, 1996, doi: 10.1109/91.493904.

[39] Z. Şen, *Bulanık Mantık ve Modelleme İlkeleri*. İstanbul: Bilge Kültür Sanat, 2001.

[40] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *International Journal of Man-Machine Studies*, 1975, doi: 10.1016/S0020-7373(75)80002-2.

[41] L. A. Zadeh, "Is there a need for fuzzy logic?," *Information Sciences*, vol. 178, no. 13, pp. 2751–2779, Jul. 2008, doi: 10.1016/j.ins.2008.02.012.

[42] J. C. Bezdek, "Numerical taxonomy with fuzzy sets," *Journal of Mathematical Biology*, vol. 1, no. 1, pp. 57–71, 1974, doi: 10.1007/BF02339490.

[43] D. Dubois and H. Prade, "Operations in a fuzzy-valued logic," *Information and Control*, vol. 43, no. 2, pp. 224–240, Nov. 1979, doi: 10.1016/S0019-9958(79)90730-7.

[44] R. Lowen, "Convex fuzzy sets," *Fuzzy Sets and Systems*, vol. 3, no. 3, pp. 291–310, May 1980, doi: 10.1016/0165-0114(80)90025-1.

[45] W. Pedrycz, "Statistically grounded logic operators in fuzzy sets," *European Journal of Operational Research*, vol. 193, no. 2, pp. 520–529, Mar. 2009, doi: 10.1016/j.ejor.2007.12.009.

[46] R. R. Yager, "Connectives and quantifiers in fuzzy sets," *Fuzzy Sets and Systems*, vol. 40, no. 1, pp. 39–75, Mar. 1991, doi: 10.1016/0165-0114(91)90046-S.

[47] P. J. M. van Laarhoven and W. Pedrycz, "A fuzzy extension of Saaty's priority theory," *Fuzzy Sets and Systems*, vol. 11, no. 1–3, pp. 229–241, 1983, doi: 10.1016/S0165-0114(83)80082-7.

[48] T. L. Saaty, "A scaling method for priorities in hierarchical structures," *Journal of Mathematical Psychology*, vol. 15, no. 3, pp. 234–281, Jun. 1977, doi: 10.1016/0022-2496(77)90033-5.

[49] J. J. Buckley, "Fuzzy hierarchical analysis," *Fuzzy Sets and Systems*, vol. 17, no. 3, pp. 233–247, Dec. 1985, doi: 10.1016/0165-0114(85)90090-9.

[50] N. Y. Pehlivan and A. Apaydın, "Bulanık k-En Yakın Komşuluk Tahmin Edicisi ve Bulanık Radyal Tabanlı Fonksiyon Ağları," *Selçuk Üniversitesi Fen Fakültesi Fen Dergisi*, vol. 1, no. 26, pp. 19–32, Jun. 2005, Accessed: May 08, 2021. [Online]. Available: https://dergipark.org.tr/tr/pub/sufefd/247085.

[51] W. Karwowski and G. W. Evans, "Fuzzy concepts in production management research: A review," *International Journal of Production Research*, vol. 24, no. 1, pp. 129–147, 1986, doi: 10.1080/00207548608919718.

[52] R. C. Bansal, “Bibliography on the Fuzzy Set Theory Applications in Power Systems (1994-2001),” *IEEE Transactions on Power Systems*, vol. 18, no. 4, pp. 1291–1299, 2003, doi: 10.1109/TPWRS.2003.818595.

[53] R. E. Bellman and L. A. Zadeh, “Decision-Making in a Fuzzy Environment,” *Management Science*, vol. 17, no. 4, p. B-141-B-164, Dec. 1970, doi: 10.1287/mnsc.17.4.b141.

[54] J. F. Baldwin and N. C. F. Guild, “Comparison of fuzzy sets on the same decision space,” *Fuzzy Sets and Systems*, vol. 2, no. 3, pp. 213–231, Jul. 1979, doi: 10.1016/0165-0114(79)90028-9.

[55] A. Mardani, A. Jusoh, and E. K. Zavadskas, “Fuzzy multiple criteria decision-making techniques and applications - Two decades review from 1994 to 2014,” *Expert Systems with Applications*, vol. 42, no. 8. Elsevier Ltd, pp. 4126–4148, May 15, 2015, doi: 10.1016/j.eswa.2015.01.003.

[56] C. Kahraman, S. C. Onar, and B. Oztaysi, “Fuzzy Multicriteria Decision-Making: A Literature Review *,” 2015. Accessed: May 09, 2021. [Online].

[57] T. Takagi and M. Sugeno, “Fuzzy Identification of Systems and Its Applications to Modeling and Control,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-15, no. 1, pp. 116–132, 1985, doi: 10.1109/TSMC.1985.6313399.

[58] D. K. Chaturvedi, “Introduction to soft computing,” *Studies in Computational Intelligence*, vol. 103, pp. 1–10, 2008, doi: 10.1007/978-3-540-77481-5_1.

[59] C. Kahraman, B. Öztayşi, and S. Çevik Onar, “A Comprehensive Literature Review of 50 Years of Fuzzy Set Theory,” *International Journal of Computational Intelligence Systems*, vol. 9, no. April, pp. 3–24, 2016, doi: 10.1080/18756891.2016.1180817.

[60] N. Baykal and T. Beyan, *Bulanık Mantık İlke ve Temelleri*. Ankara: Bıçaklar Kitabevi, 2004.

[61] M. Hanss, *Applied fuzzy arithmetic: An introduction with engineering applications*. 2005.

[62] M. Sugeno, “Fuzzy Measures and fuzzy integrals-a survey,” in *Fuzzy Automata and Decision Process*, M. M. Gupta, N. Saridis, G, and R. Gaines B, Eds. New York, 1977, pp. 89–102.

[63] R. R. Yager, “On a general class of fuzzy connectives,” *Fuzzy Sets and Systems*, 1980, doi: 10.1016/0165-0114(80)90013-5.

[64] D. Dubois and H. Prade, “Systems of linear fuzzy constraints,” *Fuzzy Sets and Systems*, 1980, doi: 10.1016/0165-0114(80)90004-4.

[65] D. Dubois and H. Prade, “New Results about Properties and Semantics of Fuzzy Set-Theoretic Operators,” in *Fuzzy Sets*, 1980.

[66] V. V. Cross and T. A. Sudkamp, *Similarity and Compatibility in Fuzzy Set Theory*, vol. 93. 2002.

[67] L. A. Zadeh, “The concept of a linguistic variable and its application to approximate reasoning-I,” *Information Sciences*, 1975, doi: 10.1016/0020-0255(75)90036-5.

[68] T. Chaira, *Fuzzy Set and Its Extension*. 2019.

[69] H.-J. Zimmermann, *Fuzzy Set Theory—and Its Applications*. 1996.

[70] J. S. R. Jang, C. T. Sun, and E. Mizutani, “Jyh-Shing Roger Jang, Chuen-Tsai Sun, Eiji Mizutani - Neuro-Fuzzy and Soft Computing_ A Computational Approach to Learning and Machine Intelligence -Prentice Hall (1997).pdf.” p. 614, 1997.

[71] H. Garg, “Some arithmetic operations on the generalized sigmoidal fuzzy numbers and its application,” *Granular Computing*, vol. 3, no. 1, pp. 9–25, 2018, doi: 10.1007/s41066-017-0052-7.

[72] Y.-J. Lai and C.-L. Hwang, “Fuzzy Mathematical Programming,” 1992, pp. 74–186.

[73] A. Kaufmann and M. M. Gupta, *Introduction to Fuzzy Arithmetic, Theory and Applications*. New York: Van Nostrand Reinold, 1985.

[74] D. Dubois and H. Prade, *Possibility Theory*. 1988.

[75] H. Tanaka, S. Uejima, and K. Asai, “LINEAR REGRESSION ANALYSIS WITH FUZZY MODEL.,” *IEEE Transactions on Systems, Man and Cybernetics*, 1982, doi: 10.1109/tsmc.1982.4308925.

[76] L. İ. Yücel, *Bulanık Teorinin Ekonometrik Uygulamaları*. İstanbul: DER Yayınları, 2018.

[77] N. Chukhrova and A. Johannssen, “Fuzzy regression analysis: Systematic review and bibliography,” *Applied Soft Computing Journal*. 2019, doi: 10.1016/j.asoc.2019.105708.

[78] H. Tanaka, “Fuzzy data analysis by possibilistic linear models,” *Fuzzy Sets and Systems*, 1987, doi: 10.1016/0165-0114(87)90033-9.

[79] P. Diamond, “Fuzzy least squares,” *Information Sciences*, vol. 46, no. 3, pp. 141–157, 1988, doi: 10.1016/0020-0255(88)90047-3.

[80] H. Moskowitz and K. Kim, “On assessing the *H* value in fuzzy linear regression,” *Fuzzy Sets and Systems*, 1993, doi: 10.1016/0165-0114(93)90505-C.

[81] G. Peters, “Fuzzy linear regression with fuzzy intervals,” *Fuzzy Sets and Systems*, vol. 63, no. 1, pp. 45–55, Apr. 1994, doi: 10.1016/0165-0114(94)90144-9.

[82] K. J. Kim, H. Moskowitz, and M. Koksalan, “Fuzzy versus statistical linear regression,” *European Journal of Operational Research*, vol. 92, no. 2, pp. 417–434, Jul. 1996, doi: 10.1016/0377-2217(94)00352-1.

[83] K. J. Kim and H. R. Chen, “A comparison of fuzzy and nonparametric linear regression,” *Computers and Operations Research*, vol. 24, no. 6, pp. 505–519, Jun. 1997, doi: 10.1016/S0305-0548(96)00075-5.

[84] H. Ishibuchi and M. Nii, “Fuzzy regression using asymmetric fuzzy coefficients and fuzzified neural networks,” *Fuzzy Sets and Systems*, vol. 119, no. 2, pp. 273–290, Apr. 2001, doi: 10.1016/S0165-0114(98)00370-4.

[85] Y. H. O. Chang, “Hybrid fuzzy least-squares regression analysis and its reliability measures,” *Fuzzy Sets and Systems*, vol. 119, no. 2, pp. 225–246, Apr. 2001, doi: 10.1016/S0165-0114(99)00092-5.

[86] P. D’Urso and T. Gastaldi, “A least-squares approach to fuzzy linear regression analysis,” *Computational Statistics and Data Analysis*, vol. 34, no. 4, pp. 427–440, Oct. 2000, doi: 10.1016/S0167-9473(99)00109-7.

[87] T. E. Dielman, “A comparison of forecasts from least absolute value and least squares regression,” *Journal of Forecasting*, vol. 5, no. 3, pp. 189–195, Jul. 1986, doi: 10.1002/for.3980050305.

[88] P. T. Chang and E. S. Lee, “Fuzzy least absolute deviations regression and the conflicting trends in fuzzy parameters,” *Computers and Mathematics with Applications*, vol. 28, no. 5, pp. 89–101, Sep. 1994, doi: 10.1016/0898-1221(94)00143-X.

[89] K. J. Kim, D. H. Kim, and S. H. Choi, “LEAST ABSOLUTE DEVIATION ESTIMATOR IN FUZZY REGRESSION,” Accessed: May 02, 2021. [Online].

[90] H. Tau Lee and S. Hua Chen, “Fuzzy regression model with fuzzy input and output data for manpower forecasting,” *Fuzzy Sets and Systems*, vol. 119, no. 2, pp. 205–213, Apr. 2001, doi: 10.1016/S0165-0114(98)00382-0.

[91] M. M. Nasrabadi, E. Nasrabadi, and A. R. Nasrabady, “Fuzzy linear regression analysis: A multi-objective programming approach,” *Applied Mathematics and Computation*, vol. 163, no. 1, pp. 245–251, 2005, doi: 10.1016/j.amc.2004.02.008.

[92] J. Watada, “Fuzzy Time-series Analysis and Its Forecasting of Sales Volume,” *Fuzzy Regression Analysis*, 1992, Accessed: May 02, 2021. [Online]. Available: https://ci.nii.ac.jp/naid/10021314274.

[93] T. J. Ross *et al.*, “Fuzzy Logic with Engineering Applications,” *IEEE Transactions on Information Theory*, 2004.

[94] H. F. Wang and R. C. Tsaur, “Resolution of fuzzy regression model,” *European Journal of Operational Research*, 2000, doi: 10.1016/S0377-2217(99)00317-3.

[95] H. Tanaka and J. Watada, “Possibilistic linear systems and their application to the linear regression model,” *Fuzzy Sets and Systems*, 1988, doi: 10.1016/0165-0114(88)90054-1.

[96] H. Tanaka, I. Hayashi, and J. Watada, “Possibilistic linear regression analysis for fuzzy data,” *European Journal of Operational Research*, vol. 40, no. 3, pp. 389–396, 1989, doi: 10.1016/0377-2217(89)90431-1.

[97] D. T. Redden and W. H. Woodall, “Properties of certain fuzzy linear regression methods,” *Fuzzy Sets and Systems*, 1994, doi: 10.1016/0165-0114(94)90159-7.

[98] Y. H. O. Chang and B. M. Ayyub, “Fuzzy regression methods - a comparative assessment,” *Fuzzy Sets and Systems*, 2001, doi: 10.1016/S0165-0114(99)00091-3.

[99] M. Sakawa and H. Yano, “Multiobjective fuzzy linear regression analysis and its application,” *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)*, vol. 73, no. 12, pp. 1–10, Jan. 1990, doi: 10.1002/ecjc.4430731201.

[100] M. Sakawa and H. Yano, “Fuzzy linear regression analysis for fuzzy input-output data,” *Information Sciences*, 1992, doi: 10.1016/0020-0255(92)90069-K.

[101] D. T. Redden and W. H. Woodall, “Further examination of fuzzy linear regression,” *Fuzzy Sets and Systems*, vol. 79, no. 2, pp. 203–211, Apr. 1996, doi: 10.1016/0165-0114(95)00176-X.

[102] E. Nasrabadi, S. M. Hashemi, and M. Ghatee, “AN LP-based approach to outliers detection in fuzzy regression analysis,” *International Journal of Uncertainty, Fuzziness and Knowlege-Based Systems*, vol. 15, no. 4, pp. 441–456, Aug. 2007, doi: 10.1142/S0218488507004789.

[103] M. M. Nasrabadi and E. Nasrabadi, “A Mathematical-Programming Approach to Fuzzy Linear Regression Analysis,” *Applied Mathematics and Computation*, vol. 155, pp. 873–881, 2004.

[104] E. C. Özelkan and L. Duckstein, “Multi-objective fuzzy regression: A general framework,” *Computers and Operations Research*, 2000, doi: 10.1016/S0305-0548(99)00110-0.

[105] H. Tanaka and H. Lee, “Interval regression analysis by quadratic programming approach,” *IEEE Transactions on Fuzzy Systems*, 1998, doi: 10.1109/91.728436.

[106] H. Lee and H. Tanaka, “Upper and lower approximation models in interval regression using regression quantile techniques,” *European Journal of Operational Research*, vol. 116, no. 3, pp. 653–666, Aug. 1999, doi: 10.1016/S0377-2217(98)00191-X.

[107] L. Haekwan and H. Tanaka, “Fuzzy approximations with non-symmetric fuzzy parameters in fuzzy regression analysis,” *Journal of the Operations Research Society of Japan*, vol. 42, no. 1, pp. 98–112, 1999, doi: 10.15807/jorsj.42.98.

[108] P. Diamond and H. Tanaka, “Fuzzy Regression Analysis,” 1998.

[109] B. Kim and R. R. Bishu, “Evaluation of fuzzy linear regression models by comparing membership functions,” *Fuzzy Sets and Systems*, vol. 100, no. 1–3, pp. 343–352, 1998, doi: 10.1016/S0165-0114(97)00100-0.

[110] N. Wang, W. X. Zhang, and C. L. Mei, “Fuzzy nonparametric regression based on local linear smoothing technique,” *Information Sciences*, vol. 100, no. 18, pp. 3882–3900, 2007, doi: 10.1016/j.ins.2007.03.002.

[111] A. Hepbasli and O. Ozgener, “A review on the development of wind energy in Turkey,” *Renewable and Sustainable Energy Reviews*, vol. 8, pp. 257–276, 2004.

[112] C. Ilkilic, “Wind energy and assessment of wind energy potential in Turkey,” *Renewable and Sustainable Energy Reviews*, vol. 16, pp. 1165–1173, 2012.

[113] S. Durak and S. Özer, *Rüzgar Enerjisi Teori ve Uygulama*. Ankara: Impress Yayinevi, 2008.

[114] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh, “Current methods and advances in forecasting of wind power generation,” *Renewable Energy*, vol. 37, no. 1, pp. 1–8, 2012, doi: http://dx.doi.org/10.1016/j.renene.2011.05.033.

[115] C. Croonenbroeck and D. Ambach, “A selection of time series models for short-to medium-term wind power forecasting,” *J. Wind Eng. Ind. Aerodyn*, vol. 136, pp. 201–210, 2015.

[116] D. Villanueva and A.E. Feijoo, “Reformulation of parameters of the logistic function applied to power curves of wind turbines,” *Electr. Pow. Syst. Res.*, vol. 137, pp. 51–58, 2016.

[117] M. Marciukaitis, I. Zutautaite, L. Martisauskas, B. Joksas, and A. S. Giedrius Gecevicius, “Non-linear regression model for wind turbine power curve,” *Renewable Energy*, vol. 113, pp. 732–741, 2017, doi: 10.1016/j.renene.2017.06.039.

[118] A. Marvuglia and A. Messineo, "Monitoring of wind farms' power curves using machine learning techniques," *Appl. Energy*, vol. 98, pp. 574–583, 2012.

[119] B. Park and J. Hur, "Accurate Short-Term Power Forecasting of Wind Turbines : The Case of Jeju Island ' s Wind Farm," *Energies*, vol. 10, no. 812, 2017, doi: 10.3390/en10060812.

[120] M. Ranaboldo, "Multiple linear regression MOS for short-term wind power forecast," 2013.

[121] P. Skrabanek and N. Martínková, "An R Package for Fuzzy Linear Regression," in *ENBIK*, 2018.

[122] M. Gagolewski and J. Caha, "A Guide to the FuzzyNumbers Package for R," pp. 1–38, 2019.

[123] R. I. Kabacoff, *Data analysis and graphics with R*.