Iranian Journal of Operations Research
Vol. 15, No. 1, 2024, pp. 1-11
DOI:

Neutrosophic Fuzzy Regression: A Linear Programming Approach

Zahra Behdani[1], Majid Darehmiraki[1]

Regression is a statistical technique used in finance, investment, and several other domains to assess the magnitude and precision of the association between a dependent variable (often represented as Y) and a set of other factors (referred to as independent variables). This work introduces a linear programming approach for constructing regression models for Neutrosophic data. To achieve this objective, we use the least absolute deviation approach to transform the regression issue into a linear programming problem. In this method, the minimum absolute error value method is used to estimate the model parameters. The feature of this method is the low effectiveness of estimates compared to outlier data. Ultimately, the efficacy of the suggested approach in resolving such problems has been shown through the presentation of a concrete illustration.
Keywords: *Neutrosophic fuzzy, Regression, Least absolute, Linear programming.*

Manuscript was received on 12/25/2023, revised on 01/21/2024 and accepted for publication on 06/03/2024.

1. Introduction

Regression analysis is a statistical method that explains and predicts changes in the dependent variable through the independent variable or variables. Regression methods are theoretically divided into three types: linear, non-linear, and non-parametric regression. Linear regression, according to the number of dependent and independent variables, can be included in the following three types of divisions:

A) Simple linear regression: In this case, there is a dependent variable and an independent variable.

b) Multiple regression: in this case, one dependent variable and more than one independent variable are assumed.

c) Multivariate regression: In this case, several dependent variables and several independent variables are assumed for analysis.

Classical linear regression has many applications, but sometimes building the model faces problems such as a small or inappropriate number of observations, problems in defining the appropriate distribution function, ambiguity in the relationship between dependent and independent variables, and ambiguity in the occurrence or degree of occurrence of events. Carelessness and error. Using statistical regression analysis in these situations may lead to wrong conclusions. To solve this problem, other methods such as robust regression and fuzzy linear regression can be used, where fuzzy linear regression is discussed with Neutrosophic data. Fuzzy regression is generally divided into three types:

A) Fuzzy regression in the case that the relationships between the variables (regression model coefficients) are assumed to be fuzzy.

[*] Corresponding Author.
[1] Department of mathematics and statistics, Behbahan Khatam Alanbia University of Technology, Khouzestan, Iran, Email: behdani@bkatu.ac.ir.

b) Fuzzy regression in cases where the observations in the dependent variable and independent variables are imprecise and fuzzy.

c) Fuzzy regression, in which both relationships between variables and fuzzy observations are considered.

In linear regression with fuzzy coefficients, it is assumed that observations and variables are accurate and that there is ambiguity in the regression model and coefficients.

Due to the indeterminacy in real world data, fuzzy logic is one of the most efficient tools to deal with these data [17,19,22]. The topic of fuzzy regression was noticed in the early 80s after the introduction of fuzzy set theory by Professor Zadeh [30]. For the first time, Tanaka and his colleagues [29] studied fuzzy regression in 1982. Tanaka assumed that the data were triangular fuzzy numbers and estimated the regression coefficients by minimizing a fuzzy index. The basis of his work was mathematical programming methods. In the same year, he discussed the simplest form of fuzzy regression. He predicted the value of the dependent variable with fuzzy observations. In 1986, Jajoga calculated the linear regression coefficients using a generalized version of the least squares method, while until then, the mathematical programming method was mainly used to analyze regression models [12]. In 1987, Celmens proposed a method for fitting a multivariate fuzzy model by minimizing a least squares objective function and presented a least squares method for fuzzy regression models [5]. Fuzzy regression is a method used in fuzzy inference systems to handle uncertainty and solve problems without using accurate mathematical models. It is based on fuzzy set theory and regression analysis. The fuzzy regression function approach is a type of fuzzy inference system that uses regression analysis and fuzzy clustering. However, the data set's outliers have an impact on this strategy. To overcome this limitation, robust intuitionistic fuzzy regression function approaches have been proposed, which use robust regression methods instead of ordinary least squares. These approaches have been successfully applied to forecast Bitcoin and gold time series data, even in the presence of outliers. It has also been used to guess the rheological properties of bio-nanocomposites. A fuzzy-hybrid model is better at this than a crisp model at both being accurate and general [10]. A new model called fuzzy data envelopment regression analysis (FDERA) has also been suggested. It combines fuzzy regression and data envelopment analysis (DEA) to figure out how efficient different businesses are [1]. In their study, the authors of reference [21] introduced a technique called kernel fuzzy c-regression (KFCR) to address fuzzy regression modeling. The LAD-KFCR incorporates the modified Huber function to implement the least absolute deviation (LAD) approach. Li et al. introduced a novel fuzzy regression model that utilizes trapezoidal fuzzy numbers and the least absolute deviation approach [18]. Initially, a novel distance metric was established to quantify the dissimilarity between trapezoidal fuzzy numbers. This distance metric serves as the foundation for many applications. Additionally, the least absolute deviation method was combined with the aforementioned distance metric to examine a fuzzy regression model. In this model, the parameters are represented as trapezoidal fuzzy numbers. Concurrently, they thoroughly examined the model algorithms for three specific scenarios, including various forms of inputs, outputs, and regression coefficients. The researchers in [2] used ordered weighted averaging operators to provide a comprehensive fuzzy regression modeling approach for crisp, fuzzy-input, or fuzzy-output data. The OWA-Least Trimmed Absolute Deviations (OWA-LTAD) fuzzy regression model is a specific instance of the OWA-based fuzzy regression model. It has been extensively examined for data with crisp inputs and fuzzy outputs. Hesameian et al. used an expanded center and range technique in a three-step process to look into nonparametric regression problems with fuzzy answers and precise predictors [9].They used the Nadaraya-Watson estimator to estimate the unidentified fuzzy smooth function. The unknown bandwidth at each step was established using generalized cross validation criteria [8]. The single-valued Neutrosophic number (SVNN) is very appropriate for expressing information on truth, falsehood, and indeterminacy in circumstances that are inconsistent and ambiguous. In [16], researchers came up with SVNN-GPR, a method that combines Support Vector Neural Network (SVNN) and Gaussian process regression to check how stable open-pit mining slopes are in an uncertain environment. The score values obtained from the SVNN output of the suggested approach were used to forecast the stability of individual slopes.

The least-squares method has many benefits, such as being able to deal with non-Markovian data, path-dependent pay-offs, and a flexible way to estimate conditional expectations [20].It allows for specific implementations with increased freedom, and even modest extensions of standard linear regression can produce satisfactory results in practical applications [2]. The method is robust against noise and can accurately reconstruct multiple unknown coefficients from noisy data [3]. It also avoids the need for assigning experimental values and is robust against intensity fluctuations [4]. Additionally, the least-squares method can be generalized to multidimensional space, expanding its capabilities for geometric modeling of multifactor processes and phenomena [5].

So far, various fuzzy regression models have been examined by researchers, considering the accuracy of Neutrosophic fuzzy numbers in modeling uncertain values, in this article we present the fuzzy regression model in the Neutrosophic environment, as far as we know, this regression model with the method The least absolute deviation has not been investigated so far. The subsequent sections of this work are structured as follows. The next part introduces some fundamental principles of fuzzy Neutrosophic. Section 3 presents the introduction of the Neutrosophic fuzzy regression model, along with the implementation of the least absolute deviation approach. The efficacy and effectiveness of the suggested approach is shown in Section 4. The overview of this study is finally offered in Section 5.

2. Preliminiries

Fuzzy Neutrosophic sets are used to represent uncertainty in decision making by substituting binary components with Neutrosophic triplets in the tabular format of soft sets. This enables a more effective management of situations in which the decision maker harbors uncertainty about the accuracy of the fuzzy/qualitative descriptions supplied to the items of the universal set [24]. Neutrosophic decision making approaches, using these sets, have been formulated and used across several fields, including the selection of a new player by a soccer team [17]. Fuzzy Neutrosophic sets have the benefit of effectively handling both uncertain and inconsistent information concurrently, rendering them well-suited for decision-making issues that include fuzzy or qualitative parameters [28]. Some people have also suggested using aggregation operators that work with single-valued Neutrosophic fuzzy sets to help with the problems that come up when you have to make choices for groups with more than one attribute [20].

The Neutrosophic sets were presented in order to address the weaknesses of the fuzzy sets. Smarandache [25,26] defined Neutrosophic set as follows:

Definition 2.1. Let X be a universe of discourse, with a class of elements in X denoted by x. A Neutrosophic set B in X is summarized by a truth-membership function $\mu_B(x)$, an indeterminacy-membership function $\nu_B(x)$, and a falsity membership function $\eta_B(x)$. The functions $\mu_B(x), \nu_B(x)$ and $\eta_B(x)$ are real standard or nonstandard subset of $]0^-, 1^+[$. That is $\mu_B(x): X \to]0^-, 1^+[$, $\nu_B(x): X \to]0^-, 1^+[$, and $\eta_B(x): X \to]0^-, 1^+[$.

Definition 2.2. Let X be a universe of discourse. A single valued Neutrosophic set A over X is an object having the form $A = \{(x, \mu_A(x), \nu_A(x), \eta_A(x)): x \in X\}$ where $\mu_A(x): X \to [0,1], \nu_A(x): X \to [0,1]$ and $\eta_A(x): X \to [0,1]$ with $0 \leq \mu_A(x) + \nu_A(x) + \eta_A(x) \leq 3$ for all $x \in X$. The functions $\mu_A(x), \nu_A(x)$ and $\eta_A(x)$ denote the truth-membership degree, the indeterminacy-membership degree and the falsity-membership degree of x to A, respectively. For convenience, a SVN number is denoted by $A = (a, b, c)$, where $a, b, c \in [0,1]$ and $a + b + c \leq 3$.

In a special case a single valued triangular Neutrosophic number is defined as:

Definition 2.3. Let $v_{\tilde{a}}, w_{\tilde{a}}$ and $t_{\tilde{a}}$ and $a_1'', a_1, a_1', a_2, a_3, a_4', a_4, a_4'' \in \mathbb{R}$ such that $a_1'' \leq a_1 \leq a_1' \leq a_2 \leq a_3 \leq a_4' \leq a_4 \leq a_4''$. . Then a single valued trapezoidal Neutrosophic number (SVTNN), $\tilde{a} = \langle (a_1'', a_1, a_1', a_2, a_3, a_4', a_4, a_4''), v_{\tilde{a}}, w_{\tilde{a}}, t_{\tilde{a}} \rangle$ is a special Neutrosophic set on the real line set $\mathbb{R}$, whose truth-membership, indeterminacy-membership and falsity membership functions are given as follows:

$$\mu_{\tilde{a}}(x) = \begin{cases} \dfrac{(x-a_1)}{(a_2-a_1)} v_{\tilde{a}} & a_1 \leq x \leq a_2 \\ v_{\tilde{a}} & a_2 \leq x \leq a_3 \\ \dfrac{(a_4-x)}{(a_4-a_3)} v_{\tilde{a}} & a_3 \leq x \leq a_4 \\ 0 & otherwise \end{cases}$$

$$\nu_{\tilde{a}}(x) = \begin{cases} \dfrac{(a_2 - x + w_{\tilde{a}}(x-a_1))}{(a_2-a_1)} & a_1' \leq x \leq a_2 \\ w_{\tilde{a}} & a_2 \leq x \leq a_3 \\ \dfrac{(x - a_3 + w_{\tilde{a}}(a_4'-x))}{(a_4'-a_3)} & a_3 \leq x \leq a_4' \\ 1 & otherwise \end{cases}$$

$$\eta_{\tilde{a}}(x) = \begin{cases} \dfrac{(a_2 - x + t_{\tilde{a}}(x-a_1))}{(a_2-a_1)} & a_1'' \leq x \leq a_2 \\ t_{\tilde{a}} & a_2 \leq x \leq a_3 \\ \dfrac{(x - a_3 + t_{\tilde{a}}(a_4''-x))}{(a_4''-a_3)} & a_3 \leq x \leq a_4'' \\ 1 & otherwise \end{cases}$$

where $v_{\tilde{a}}, w_{\tilde{a}}$ and $t_{\tilde{a}}$ denote the maximum truth-membership degree, minimum indeterminacy-membership degree and minimum falsity-membership degree, respectively.

Definition 2.4. Let $\tilde{a} = \langle (a_1, a_2, a_3, a_4), v_{\tilde{a}}, w_{\tilde{a}}, t_{\tilde{a}} \rangle$ and $\tilde{b} = \langle (b_1, b_2, b_3, b_4), v_{\tilde{b}}, w_{\tilde{b}}, t_{\tilde{b}} \rangle$ be two SVTNNs and $\lambda \geq 0$ be any real number. Then,

1) Addition of two SVTNNs

$$\tilde{a} + \tilde{b} = \langle (a_1 + b_1, a_2 + b_2, a_3 + b_3, , a_4 + b_4), v_{\tilde{a}} \wedge v_{\tilde{b}}, w_{\tilde{a}} \vee w_{\tilde{b}}, t_{\tilde{a}} \vee t_{\tilde{b}} \rangle.$$

2) Subtraction of two SVTNNs

$$\tilde{a} - \tilde{b} = \langle (a_1 - b_4, a_2 - b_3, a_3 - b_2, , a_4 - b_1), v_{\tilde{a}} \wedge v_{\tilde{b}}, w_{\tilde{a}} \vee w_{\tilde{b}}, t_{\tilde{a}} \vee t_{\tilde{b}} \rangle.$$

3) Inverse of a SVTNN

$$\tilde{a}^{-1} = \langle \left(\frac{1}{a_4}, \frac{1}{a_3}, \frac{1}{a_2}, \frac{1}{a_1} \right), v_{\tilde{a}}, w_{\tilde{a}}, t_{\tilde{a}} \rangle, \; where \; \tilde{a} \neq 0.$$

4) Multiplication of a SVTNN by constant value

$$\lambda \tilde{a} = \begin{cases} \langle (\lambda a_1, \lambda a_2, \lambda a_3, \lambda a_4), v_{\tilde{a}}, w_{\tilde{a}}, t_{\tilde{a}} \rangle & \lambda > 0, \\ \langle (\lambda a_4, \lambda a_3, \lambda a_2, \lambda a_1), v_{\tilde{a}}, w_{\tilde{a}}, t_{\tilde{a}} \rangle & \lambda < 0. \end{cases}$$

Let $\tilde{a} = \langle (a_1, a_2, a_3, a_4), v_{\tilde{a}}, w_{\tilde{a}}, t_{\tilde{a}} \rangle$ and $\tilde{b} = \langle (b_1, b_2, b_3, b_4), v_{\tilde{b}}, w_{\tilde{b}}, t_{\tilde{b}} \rangle$ be two SVTNNs .To calculate the distance between two Neutrosophic fuzzy numbers, we develop the distance measure suggested by Diamond [6,7] for Neutrosophic numbers as follows:

$$d(\tilde{a}, \tilde{b}) = |a_1 - b_1| + |a_2 - b_2| + |a_3 - b_3| + |a_4 - b_4|.$$

3. Fuzzy least absolute regression

Fuzzy least absolute regression utilizes a probabilistic linguistic term set to construct a fuzzy regression model that can handle inputs of mixed kinds. This study presents the notion of a double-cut set for probabilistic linguistic term sets and explores novel operations for probabilistic linguistic phrase sets using double-cut sets. The suggested model is a combination of fuzzy and least absolute regression techniques. To calculate the fuzzy regression parameters, a linear programming approach is used. The suggested model is suitable for cases when the data exhibits fat-tailed or non-linear characteristics [13]. A different research used intuitionistic fuzzy numbers (IFNs) to represent the explanatory and response variables in intuitionistic fuzzy regression models. Mathematical programming problems are constructed using the criteria of least absolute deviations to create intuitionistic fuzzy regression models with intuitionistic fuzzy parameters [3,8].

We assume that we have a set of n observations in the form of $(X_{i0}, \ldots, X_{ik}, Y_i)$, $(i = 1, \ldots, n)$ so that $X_{i0} = 1$. In this model, the relationship between the independent variable and the dependent variable is presented as follows:

$$Y_i = a_0 X_{i0} + a_1 X_{i1} + \cdots + a_k X_{ik}.$$

Independent and dependent variables are Neutrosophic fuzzy numbers while coefficients $a_j, j = 0, 1, \ldots, k$ are crisp numbers.

In order to get the following objective function, we first compute each estimated value a_j of the regression coefficient a_j based on the criteria of the least absolute deviation. This is accomplished by minimizing the total absolute error in accordance with the suggested absolute distance.

$$\min \sum_{i=1}^{n} d(Y_i, a_0 X_{i0} + a_1 X_{i1} + \cdots + a_k X_{ik}).$$

Let $X_{ij} = \langle (X_{ij_1}, X_{ij_2}, X_{ij_3}, X_{ij_4}), v_{X_{ij}}, w_{X_{ij}}, t_{X_{ij}} \rangle$ and $X_{ij} = \langle (Y_{i_1}, Y_{i_2}, Y_{i_3}, Y_{i_4}), v_{Y_i}, w_{Y_i}, t_{Y_i} \rangle$. Here we consider two cases:

(1) $a_j > 0$.

$$\min \sum_{i=1}^{n} \left[\left| Y_{i_1} - \sum_{j=0}^{k} a_j X_{ij_1} \right| + \left| Y_{i_2} - \sum_{j=0}^{k} a_j X_{ij_2} \right| + \left| Y_{i_3} - \sum_{j=0}^{k} a_j X_{ij_3} \right| + \left| Y_{i_4} - \sum_{j=0}^{k} a_j X_{ij_4} \right| \right]$$

Now we turn the above minimization problem into a linear programming problem.

$$\min \sum_{i=1}^{n} (\eta_{i1} + \eta_{i2} + \theta_{i1} + \theta_{i2} + \rho_{i1} + \rho_{i2} + \kappa_{i1} + \kappa_{i2})$$

$$S.t.$$

$$Y_{i_1} - \sum_{j=0}^{k} a_j X_{ij_1} = \eta_{i1} - \eta_{i2},$$

$$Y_{i_2} - \sum_{j=0}^{k} a_j X_{ij_2} = \theta_{i1} - \theta_{i2},$$

$$Y_{i_3} - \sum_{j=0}^{k} a_j X_{ij_3} = \rho_{i1} - \rho_{i2},$$

$$Y_{i_4} - \sum_{j=0}^{k} a_j X_{ij_4} = \kappa_{i1} - \kappa_{i2},$$

$$\eta_{i1}, \eta_{i2}, \theta_{i1}, \theta_{i2}, \rho_{i1}, \rho_{i2}, \kappa_{i1}, \kappa_{i2}, a_j \geq 0.$$

(2) $a_j < 0$.

$$\min\sum_{i=1}^{n}\left[\left|Y_{i_1}+\sum_{j=0}^{k}a_jX_{ij_4}\right|+\left|Y_{i_2}+\sum_{j=0}^{k}a_jX_{ij_3}\right|+\left|Y_{i_3}+\sum_{j=0}^{k}a_jX_{ij_2}\right|+\left|Y_{i_4}+\sum_{j=0}^{k}a_jX_{ij_1}\right|\right]$$

Now we turn the above minimization problem into a linear programming problem.

$$\min\sum_{i=1}^{n}(\eta_{i1}+\eta_{i2}+\theta_{i1}+\theta_{i2}+\rho_{i1}+\rho_{i2}+\kappa_{i1}+\kappa_{i2})$$

$$S.t.$$

$$Y_{i_1}+\sum_{j=0}^{k}a_jX_{ij_4}=\eta_{i1}-\eta_{i2},$$

$$Y_{i_2}+\sum_{j=0}^{k}a_jX_{i3}=\theta_{i1}-\theta_{i2},$$

$$Y_{i_3}+\sum_{j=0}^{k}a_jX_{ij_2}=\rho_{i1}-\rho_{i2},$$

$$Y_{i_4}+\sum_{j=0}^{k}a_jX_{ij_1}=\kappa_{i1}-\kappa_{i2},$$

$$\eta_{i1},\eta_{i2},\theta_{i1},\theta_{i2},\rho_{i1},\rho_{i2},\kappa_{i1},\kappa_{i2},a_j\geq 0.$$

3.1. Error Analysis

Within the framework of conventional regression analysis, the evaluation of a model is based on the degree of dissimilarity between the data that is seen and the data that is estimated. Instead of using points on a single axis, fuzzy regression models make use of membership functions that are distributed over two axes. As a result, the evaluation of fuzzy regression models needs to be based on the discrepancy that exists between the membership functions. In the event where the anticipated membership function and the actual membership function are in perfect agreement with one another, there will be no difference in the membership numbers. Four distinct locations may be used to create a trapezoid-shaped membership function: the left and right extremities, as well as the beginning and finish of the peak value. This means that the estimated fuzzy number's membership function will be very similar to the observed fuzzy data's membership function if the two endpoints and two peak values are very close to the corresponding points of the observed fuzzy number.

4. Numerical example

This section presents numerical illustrations using Neutrosophic fuzzy numbers to showcase the use of the proposed fuzzy multiple linear least absolute regression model.

Due to the fact that most previous research on fuzzy regression analysis has focused on triangular fuzzy numbers, we present an example of how our new fuzzy multiple linear least absolute regression model works with trapezoidal Neutrosophic fuzzy numbers. These data include a response variable and three independent variables that are recorded in Table 1.

Table 1. The observed fuzzy data

The observed fuzzy outpu Y	X_1	X_2	X_3

⟨(27.80,29.3,30.8,32); 0.7,0.4,0.3⟩	⟨(2,3,4,5); 0.7,0.4,0.3⟩	⟨(2.1,2.6,2.8,3); 0.7,0.4,0.3⟩	⟨(2.9,3,3.2,3.4); 0.7,0.4,0.3⟩
⟨(23.80,25.3,26.8,28.1); 0.7,0.4,0.3⟩	⟨(1,1.3,2,2.5); 0.7,0.4,0.3⟩	⟨(2.1,2.3,2.7,3.1); 0.7,0.4,0.3⟩	⟨(3,3.1,3.8,3.9); 0.7,0.4,0.3⟩
⟨(26.80,30.3,32.8,33); 0.7,0.4,0.3⟩	⟨(1.1,1.7,2,2.4); 0.7,0.4,0.3⟩	⟨(1.6.1.7,1.9,2.2); 0.7,0.4,0.3⟩	⟨(2.1,2.3,2.6,2.8); 0.7,0.4,0.3⟩
⟨(27,29,30,32); 0.7,0.4,0.3⟩	⟨(3.2,3.9,4.5,5); 0.7,0.4,0.3⟩	⟨(**2**, 2.6,2.9,3.1); 0.7,0.4,0.3⟩	⟨(1.9,2.3,2.5,2.8); 0.7,0.4,0.3⟩
⟨(32,34,36,38); 0.7,0.4,0.3⟩	⟨(3,3.2,3.8,4.3); 0.7,0.4,0.3⟩	⟨(3.1,3.3,3.4,3.7); 0.7,0.4,0.3⟩	⟨(2.6.3.3.1.3.8); 0.7,0.4,0.3⟩
⟨(23.2,24.3,26.2,28); 0.7,0.4,0.3⟩	⟨(1.5,2,2.4,3); 0.7,0.4,0.3⟩	⟨(3.9,4,4.1,4.3); 0.7,0.4,0.3⟩	⟨(2.8,3,3.6,3.7); 0.7,0.4,0.3⟩
⟨(20.80,23.3,25,27); 0.7,0.4,0.3⟩	⟨(1.8,2.1,2.4,2.8); 0.7,0.4,0.3⟩	⟨(1.8,1.9,2,2.3); 0.7,0.4,0.3⟩	⟨(1.7,1.9,2.4,2.8); 0.7,0.4,0.3⟩
⟨(17.7,20.3,22,24); 0.7,0.4,0.3⟩	⟨(**2**. 8,3.1,3.4,4); 0.7,0.4,0.3⟩	⟨(**2**. 9,3,3.2,3.3); 0.7,0.4,0.3⟩	⟨2.5,2.8,2.9,3.1; 0.7,0.4,0.3⟩
⟨(24,25.3,26.8.,29); 0.7,0.4,0.3⟩	⟨(3.8,4.2,4.9,5); 0.7,0.4,0.3⟩	⟨(1.2.1.3,1.4,1.9); 0.7,0.4,0.3⟩	⟨(1,1.7,2.7,3); 0.7,0.4,0.3⟩
⟨(23,25.3,27.6,28.9); 0.7,0.4,0.3⟩	⟨(3,3.1,3.4,4); 0.7,0.4,0.3⟩	⟨(2.5,2.8,3.4,3.6); 0.7,0.4,0.3⟩	⟨(3.1,3.3,3.9,4); 0.7,0.4,0.3⟩

For Neutrosophic fuzzy data in Table 1, we fitted the following regression model:

$$\tilde{y} = \beta_{0+}\beta_1\tilde{x}_1 + \beta_2\tilde{x}_2 + \beta_3\tilde{x}_3$$

In this model, the independent variables and the dependent variable are Neutrosophic fuzzy numbers and the coefficients of the model are crisp numbers. As mentioned before, the method of the least absolute error introduced in section 3 is used to estimate the coefficients. In order to find estimates, the problem has been converted into a linear programming problem and the estimation of parameters has been calculated using R software.

By using linear programming and using R software, the model parameters were estimated and these parameters are equal to: $\beta_0 = 16.96$, $\beta_1 = 1.60$, $\beta_2 = 0$, $\beta_3 = 1.74$

In other words, the appropriate model for the table data is as follows:

$$\tilde{y} = 16.96 + 1.6\tilde{x}_1 + 1.74\tilde{x}_3$$

Figure 1 shows the membership function plot of the observed value (continuous line) versus the fitted value (dashed line) for the fourth data. For simplicity, this diagram is drawn only for the membership function.

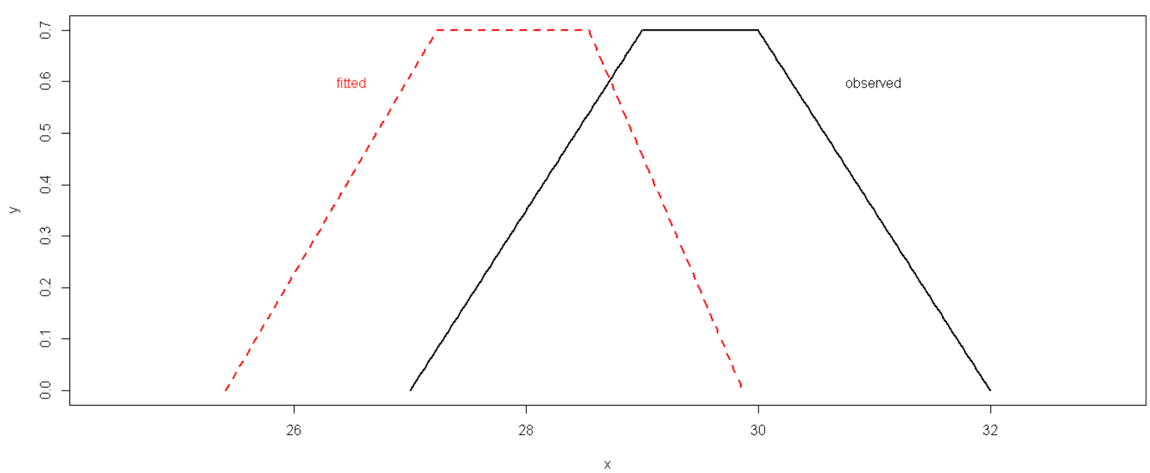

Figure 1. the membership function of the observed value (black) and the fitted value (red)

Using this model, it is possible to make predictions and estimates. For example, using this model, the fitted value for the 4th table data is equal to:

$$\tilde{y}_4 = ((25.41163, 27.23256, 28.54419, 29.86977); 0.7,0.4,0.3)$$

Also, for a new observation with values $\tilde{x}_1 = \left((4,2.1,3.7,2); 0.7,0.4,0.3\right), \tilde{x}_2 = \left((2.1,3.4,1.9,2.8); 0.7,0.4,0.3\right)$ and $\tilde{x}_3 = (2.8,2.1,3.4,4.5); 0.7,0.4,0.3)$ the predicted value using the model is equal to:

$$\tilde{y}_{new} = ((28.26512, 23.99535, 28.83023, 28.02093); 0.7,0.4,0.3)$$

Table 2 shows the fitted values for all ten data samples in Table 1. This Table shows all the fitted values of the data in Table 1 using the obtained model.

Table 2. The observed fuzzy data and the estimated fuzzy outputs for the data in Table1

The observed fuzzy outpu *Y*	**The fitted fuzzy output $\tilde{\tilde{y}}$**
$\langle(27.80,29.3,30.8,32); 0.7,0.4,0.3\rangle$	$\langle(25.23,30.50,28.96,30.91); 0.7,0.4,0.3\rangle$
$\langle(23.80,25.3,26.8,28.1); 0.7,0.4,0.3\rangle$	$\langle(23.8,23.4,26.8,27.78); 0.7,0.4,0.3\rangle$
$\langle(26.80,30.3,32.8,33); 0.7,0.4,0.3\rangle$	$\langle(22.39,23.87,24.7,25.69); 0.7,0.4,0.3\rangle$
$\langle(27,29,30,32); 0.7,0.4,0.3\rangle$	$\langle(25.41,31.94,28.54,29.86); 0.7,0.4,0.3\rangle$
$\langle(32,34,36,38); 0.7,0.4,0.3\rangle$	$\langle(26.31,29.60,28.47,30.49); 0.7,0.4,0.3\rangle$
$\langle(23.2,24.3,26.2,28); 0.7,0.4,0.3\rangle$	$\langle(24.25,25.40,27.09,28.23); 0.7,0.4,0.3\rangle$
$\langle(20.80,23.3,25,27); 0.7,0.4,0.3\rangle$	$\langle(22.82,25.22,25,26.34); 0.7,0.4,0.3\rangle$
$\langle(17.7,20.3,22,24); 0.7,0.4,0.3\rangle$	$\langle(25.82,28.91,27.48,28.79); 0.7,0.4,0.3\rangle$
$\langle(24,25.3,26.8.,29); 0.7,0.4,0.3\rangle$	$\langle(28.29,32.42,33.55,35.45); 0.7,0.4,0.3\rangle$
$\langle(23,25.3,27.6,28.9); 0.7,0.4,0.3\rangle$	$\langle(27.18,28.91,29.22,30.36); 0.7,0.4,0.3\rangle$

The sum absolute error and the mean absolute error are two model evaluation criteria. These two criteria have been calculated for the data in table one and the model obtained for them.

$$\mathrm{SAE} = \sum_{i=1}^{n}\left[\left|Y_{i_1} - \sum_{j=0}^{k} a_j X_{ij_1}\right| + \left|Y_{i_2} - \sum_{j=0}^{k} a_j X_{ij_2}\right| + \left|Y_{i_3} - \sum_{j=0}^{k} a_j X_{ij_3}\right| + \left|Y_{i_4} - \sum_{j=0}^{k} a_j X_{ij_4}\right|\right]$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left[\left|Y_{i_1} - \sum_{j=0}^{k} a_j X_{ij_1}\right| + \left|Y_{i_2} - \sum_{j=0}^{k} a_j X_{ij_2}\right| + \left|Y_{i_3} - \sum_{j=0}^{k} a_j X_{ij_3}\right| + \left|Y_{i_4} - \sum_{j=0}^{k} a_j X_{ij_4}\right|\right]$$

$$SAE = 136.84, \quad MAE = 13.68$$

5. Conclusion

Fuzzy regression is a method used to estimate the relationship between variables when accurate mathematical models are not available. It is a fuzzy inference system based on regression analysis, using techniques such as ordinary least squares and fuzzy clustering. The goal of fuzzy regression is to minimize the vagueness of the model by estimating fuzzy coefficients. However, traditional fuzzy regression methods may not guarantee the desired level of generalization. To address this, a novel methodology called EFLR (Enhanced Fuzzy Linear Regression) has been proposed. EFLR aims to maximize the reliability of the constructed models instead of minimizing vagueness. It achieves this by minimizing the weighted variance of different ambiguities in each validation data situation. The proposed EFLR method has shown better generalizability and more accurate results compared to classic versions of fuzzy linear regression. Fuzzy regression offers several advantages over linear regression. Firstly, fuzzy regression models can handle uncertain relationships between variables and estimate the uncertain relationship more effectively. This is particularly useful in situations where the underlying relationship is not precise or crisp. Additionally, fuzzy regression models can handle multicollinearity, a common issue in predictive modeling, more efficiently. Fuzzy regression also provides higher reliability and robustness, as it is less sensitive to outliers and can identify poorly fitted data points. Furthermore, fuzzy regression models can provide more accurate results and narrower prediction intervals, leading to better generalization ability. Overall, fuzzy regression offers improved performance and flexibility compared to traditional linear regression models.

The Least Absolute Deviations approach is a crucial estimate technique in regression analysis due to its resilience and widespread use in many scenarios. A trapezoidal Neutrosophic fuzzy number is a frequently seen sort of Neutrosophic fuzzy number that may serve as a substitute for or symbolize other fuzzy numbers.

Therefore, using the Least Absolute Deviations method and trapezoidal Neutrosophic fuzzy number to construct a fuzzy regression model and perform further analysis has significant practical importance and is a worthwhile pursuit. This article demonstrates the computation of the precise magnitude between two trapezoidal Neutrosophic fuzzy numbers. Furthermore, a unique fuzzy regression model is introduced, which relies on the least absolute deviation technique using trapezoidal Neutrosophic fuzzy numbers. The study also investigates the algorithms used in the model and does a comprehensive analysis of the model. The experimental results indicate that our proposed model is applicable not only to trapezoidal Neutrosophic fuzzy numbers but also to triangular Neutrosophic fuzzy numbers, and it consistently produces better fitting results.

References

[1] Azadeh, Ali, et al. (2012), An integrated fuzzy regression-data envelopment analysis algorithm for optimum oil consumption estimation with ambiguous data, *Applied Soft Computing* ,12.8, 2614-2630.

[2] D'Urso, P., and Chachi, J., (2022), OWA fuzzy regression. *International journal of approximate reasoning*, 142, 430-450.

[3] Chen, L. H., and Nien, S. H. (2020), Mathematical programming approach to formulate intuitionistic fuzzy regression model based on least absolute deviations. *Fuzzy Optimization and Decision Making*, 19, 191-210.

[4] Choi, S. H., and Buckley, J. J. (2008), Fuzzy regression using least absolute deviation estimators. *Soft computing*, 12, 257-263.

[5] Celmiņš, A. (1987), Least squares model fitting to fuzzy vector data. *Fuzzy sets and systems*, 22(3), 245-269.

[6] Diamond, P., and Tanaka, H. (1998), Fuzzy regression analysis. In Fuzzy sets in decision analysis, operations research and statistics. Boston, MA: Springer US. pp. 349-387

[7] Diamond, P. (1988), Fuzzy least squares, *Information sciences*, 46(3), 141-157.

[8] Egrioglu, E., and Bas, E. (2023), Robust intuitionistic fuzzy regression functions approaches. *Information Sciences*, 638, 118992.

[9] Hesamian, G., Torkian, F., Johannssen, A., and Chukhrova, N. (2024), A fuzzy nonparametric regression model based on an extended center and range method. *Journal of Computational and Applied Mathematics*, 436, 115377.

[10] Insel, M. A., Karakuş, S., Temelcan, G., Kocken, H. G., and Albayrak, I. (2023), Handling uncertainty in rheological properties of green eggshell nanocomposites by a fuzzy-hybrid modeling approach: a comparative study. *Physica Scripta*, 98(3), 035001.

[11] Ito, K., Liang, Y., and Zou, J. (2022), Least-Squares Method for Inverse Medium Problems. *arXiv preprint arXiv*:2201.00280.

[12] Jajuga, K. (1986), Linear fuzzy regression. *Fuzzy Sets and Systems*, 20(3), 343-353.

[13] Jiang, L., and Liao, H. (2020), Mixed fuzzy least absolute regression analysis with quantitative and probabilistic linguistic information. *Fuzzy Sets and Systems*, 387, 35-48.

[14] Kropat, E., Özmen, A., Weber, G. W., Meyer-Nieberg, S., and Defterli, O. (2016), Fuzzy prediction strategies for gene-environment networks–Fuzzy regression analysis for two-modal regulatory systems. *RAIRO-Operations Research-Recherche Opérationnelle*, 50(2), 413-435.

[15] Kao, C., and Chyu, C. L. (2003). Least-squares estimates in fuzzy regression analysis. *European Journal of Operational Research*, 148(2), 426-435.

[16] Klimek, M., and Pitera, M. (2018). The least squares method for option pricing revisited. *Applicationes Mathematicae*, 45, 5-29.

[17] Khabiri, B., & Iranmanesh, M. (2023). Fully fuzzy solid transportation problems with k-scale trapezoidal fuzzy numbers. *Iranian Journal of Operations Research*, *14*(1), 104-117.

[18] Li, J., Zeng, W., Xie, J., and Yin, Q. (2016), A new fuzzy regression model based on least absolute deviation. *Engineering Applications of Artificial Intelligence*, 52, 54-64.

[19] Mehrjerdi, Y. Z. (2023). A Fuzzy Multi Attribute-Simulation decision making approach to study healthcare performance in RFID based systems. *Iranian Journal of Operations Research*, *14*(2), 1-31.

[20] Naderi, M., Chakareski, J., and Ghanbari, M. (2023), Hierarchical Q-learning-enabled Neutrosophic AHP scheme in candidate relay set size adaption in vehicular networks. *Computer Networks*, 235, 109968.

[21] Oi, Y., and Endo, Y. (2019), Kernel fuzzy c-regression based on least absolute deviation with modified huber function. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 23(3), 571-576.

[22] Pourmahmoud, J., & Sharak, N. B. (2020). Evaluating cost efficiency using fuzzy data envelopment analysis method. *Iranian journal of operations research*, *11*(1), 25-42.

[23] Qin, J., Ye, J., Sun, X., Yong, R., and Du, S. (2023), A single-valued Neutrosophic Gaussian process regression approach for stability prediction of open-pit mine slopes. *Applied Intelligence*, 53(11), 13206-13223.

[24] Salama, A. A., Khaled, O. M., and Mahfouz, K. M. (2014), Neutrosophic correlation and simple linear regression. *Neutrosophic Sets and Systems*, 5, 3-8.

[25] Smarandache, F. (2006, May). Neutrosophic set-a generalization of the intuitionistic fuzzy set. In 2006 IEEE international conference on granular computing, IEEE. pp. 38-42

[26] Smarandache, F. (2016), of Neutrosophic Numbers. Critical Review, 13, 103.

[27] Voskoglou, M. G. (2023), An Application of Neutrosophic Sets to Decision Making. *arXiv preprint arXiv*:2306.01746.

[28]	Wu, M., Zhao, T., and Fan, J. (2023), Multiple Attribute Group Decision-Making Model Based on Single-Valued Neutrosophic Fuzzy Sets and VIKOR.
[29]	Tanaka, H., Uejima, S., and Asai, K. (1982), Linear regression analysis with fuzzy model. IEEE transactions systems man and cybernetics, 12(6), 903–907
[30]	Zadeh, L. A. (1965), Fuzzy sets. Information and control, 8(3), 338-353.