Journal of Statistical Modelling: Theory and Applications
Vol. 4, No. 2, 2023, pp. 17-34
Yazd University Press 2023

JSMTA

Research Paper

Least square regression in intuitionistic fuzzy environment with crisp coefficients with the ability to determine the decision level

Zahra Behdani*, Majid Darehmiraki
Department of Mathematics and Statistics, Behbahan Khatam Alanbia University of Technology, Khouzestan, Iran

Received: February 04, 2024/ Revised: April 20, 2024/ Accepted: May 22, 2024

Abstract: In practical problems where imprecise components are an inseparable part of them, using fuzzy sets to check their features and measure their efficiency solves some of the current limitations. In this article, we examine the problem of linear regression in an intuitionistic fuzzy environment. Here, we consider the input and output data as intuitionistic fuzzy numbers and assume that the coefficients of the model are crisp numbers. Since to correctly determine the coefficients, we need to calculate the distance between the output of the model and the output data, we present a new parametric distance to calculate the distance between intuitionistic fuzzy numbers. The salient point in this article is that the researcher can estimate the parameter values of the model based on different decision-making levels. To show the efficiency and test the proposed method, several examples are presented at the end of the article. The results are calculated and compared based on the values of different levels of decision-making.

Keywords: Distance; Fuzzy regression; Intuitionistic fuzzy.
Mathematics Subject Classification (2010): 03E72, 62A86.

1 Introduction

Finding the relationship between independent and dependent variables is one of the concerns of data science experts. Regression models are among the statistical methods that have the ability to find this relationship. Regression models are used in various sciences such as medicine, engineering, social sciences, etc Khasanzoda et al. (2022). In many cases, the data we are facing is inaccurate or the relationship between them is not

*Corresponding author: behdani@bkatu.ac.ir

accurate. This inaccuracy may be caused by the way of measurement or related to the essence of the problem. In order to deal with this indeterminacy correctly, it is necessary to go to fuzzy sets or fuzzy logic in such cases and use fuzzy numbers and relationships. After fuzzy logic was first introduced by Zadeh in 1968, researchers and scientists active in this field presented various extensions for it, including: intuitive fuzzy, bipolar fuzzy, hesitant fuzzy, second type fuzzy and... . In this paper, we investigate regression models in an intuitive fuzzy environment with a proposed distance measure. In the following, we will review some of the works done in this field.

Hazarika et al. (2021) used kernel ridge regression in intuitionistic fuzzy enviroment. Egrioglu and Bas (2023) presented a robust regression model in an intuitive fuzzy environment, which is more robust to outlier data. Arefi and Taheri (2014) used a least squares estimation to solve the regression model with intuitionistic fuzzy data. Hassamian and Akbari proposed a semi-parametric method to investigate logistic regression in intuitive fuzzy space with exact inputs and intuitionistic fuzzy output Hesamian et al. (2020). The authors of paper Hesamian and akbari (2023) used the ridge estimator for partial logistic regression with intuitionistic fuzzy outputs and coefficients. Chakravarty et al. (2022) proposed a new regression function by using a noise cluster to deal with outlier data. Pang et al. (2022) presented a polynomial regression model based on fuzzy c-means clustering. They applied polynomial regression as center of each cluster. Hesamian and akbari (2023) proposed a new fuzzy regression model for exact inputs and fuzzy output using Gaussian process and non-parametric Bayesian method. Choi et al. (2019) estimated the parameters of the ridge regression model using the α-level estimation method. Nguyen et al. (2023) used Generative Adversarial Networks for the first time in order to improve the fuzzy regression model.

Both the distance between IFSs and the degree of similarity between them are relative concepts that can be calculated. In most cases, we can take use of IFSs' potential utility across disciplines by employing distance and similarity measurements. As a result, a wide range of approaches are being used to construct these tools. The hunt for novel distance and similarity measures continues, despite the fact that a fair number of them already exist. Researchers then started looking into how comparable and different IFSs could be compared to one another. Szmidt and Kacprzyk (2000) created an IFS membership parameter distance in 2000. In other circumstances, Wang and Xin (2005) said their metric is insufficient to measure distances. Thus, Grzegorzewski (2004) created Hausdorff-based distance metrics. Chen (2007) later asserted that Grzegorzewski's distance formula's distance property inequalities are invalid. Hung and Yang (2007) introduced the Lp metric for IFS distance measures. Hung and Yang (2008) created fuzzy entropy-based distance measurements and successfully induced specific similarity measures on IFSs. Yang and Chiclana Yang and Chiclana (2012) created a 3D distance using the Hausdorff metric and proved its compatibility with 2D. Hatzimichailidis et al. (2012) created a distance metric using membership and non membership degree matrices' eigenvalues. Chen and Deng (2020) introduced three distance measures: Hamming-hesitation, Hamming-cosine, and Euclidean-hesitation. In addition, Mahanta and Panda Mahanta and Panda (2021) created a novel distance metric to address IFSs' high information reluctance. In a recent study on distance and similarity metrics, Garg and Rani (2021) updated Jiang et al. (2019) by employing right-angled triangles instead of isosceles triangles. Arefi and Taheri (2014) applied the least square regression for intuitionistic fuzzy data. Leski (2023) proposed a double-ordered fuzzy c-regression models based on fuzzy S-estimator. Very recently, Hesamian et al. (2024) presented a fuzzy nonparametric fuzzy regression via the Nadaraya–Watson estimator.

Considering the increasing use of data science tools for data analysis, in this article the multivariate regression model has been investigated in the fuzzy environment. Although, as mentioned earlier, this problem has been investigated by other researchers, but in this article, by defining a parametric distance measure, we tried to examine this problem from a new perspective. Since our distance measure is a parametric distance measure and the distance between two fuzzy numbers is measured at different levels, therefore, when estimating the parameters in our regression model, we can estimate the parameters at different decision levels. For example, if the alpha cut of fuzzy numbers is done with values close to 1 for the membership function and close to zero for the non-member functions, then the fitted model can be called a model with a high decision level. In summary, here by changing the values of alpha cut, we can estimate different models for a set of fuzzy data.

In the continuation of the article, in the next section, some of the prerequisites required in relation to the calculation of fuzzy numbers are stated. In the Section 3, a new fuzzy distance is introduced and with its help, an intuitive fuzzy regression model with fuzzy inputs and outputs is proposed. Two numerical examples, one of which is practical, are provided to show the efficiency of the model in the Section 4. Finally, in the Section 5 of the article, the conclusion is drawn.

2 Intuitionistic fuzzy arithmetic

In this section, we will state the fuzzy definitions needed in the continuation of the work.

Definition 2.1. *Suppose X be a universe. An intuitionistic fuzzy set (IFS) is defined as*

$$\hat{A} = \{\langle x, \mu_{\hat{A}}(x), \nu_{\hat{A}}(x) | x \in X\rangle\},$$

where $\mu_{\hat{A}}(x)$ and $\nu_{\hat{A}}(x)$ are the degrees of membership and nonmembership functions, respectively.

It should also be noted that

$$0 \leq \mu_{\hat{A}}(x), \nu_{\hat{A}}(x) \leq 1, \quad 0 \leq \mu_{\hat{A}}(x) + \nu_{\hat{A}}(x) \leq 1.$$

In the case of intuitive fuzzy sets, sometimes another component called the hesitation function is introduced as follows

$$\pi_{\hat{A}}(x) = 1 - \mu_{\hat{A}}(x) - \nu_{\hat{A}}(x).$$

Definition 2.2. *If $\hat{A} = (\mu_{\hat{A}}, \nu_{\hat{A}})$ be an IFN in set of real number $\boldsymbol{R}$ then we have*

$$\mu_{\hat{A}}(x) = \begin{cases} f(x), & a \leq x \leq b, \\ 1, & b \leq x \leq c, \\ g(x), & c \leq x \leq d, \\ 0, & \text{otherwise}, \end{cases}$$

$$\nu_{\hat{A}}(x) = \begin{cases} h(x), & a' \le x \le b', \\ o, & b' \le x \le c', \\ r(x), & c' \le x \le d', \\ 1, & \text{otherwise}, \end{cases}$$

where $f(x)$ and $r(x)$ are nondecreasing continuous functions and $g(x)$ and $h(x)$ are nonincreasing continuous functions. Also $a' \le a, b' \le b \le c \le c', d \le d'$.

The authors of Ma et al. (1999) proposed another form for displaying fuzzy numbers in which the dependent and independent variables have been changed. By extending this representational form, we can define IFNs as follows.

Definition 2.3. *Let $\hat{A} = (\mu', \nu')$ be an IFN, where μ' is a pair $(\underline{p}, \overline{p})$ of functions $0 \le r \le 1$ are satisfied in following conditions*
1. $\underline{p}$ is a bounded monotonic increasing left continuous function,
2. $\overline{p}$ is a bounded monotonic decreasing left continuous function,
3. $\underline{p} \le \overline{p}$, $0 \le r \le 1$,
and ν' is a pair $(\underline{q}, \overline{q})$ of functions $0 \le r \le 1$ are satisfied in following conditions
1. $\underline{q}$ is a bounded monotonic decreasing left continuous function,
2. $\overline{q}$ is a bounded monotonic increasing left continuous function,
3. $\underline{q} \le \overline{q}$, $0 \le r \le 1$.

According to Definition 2.3, trapezoidal intuitionistic fuzzy numbers (TraIFN) $\hat{A} = (\mu', \nu')$ can be defined as follows: $\mu'(r) = (\underline{p}(r), \overline{p}(r)) = (a + r(b-a), d + r(c-d))$ and $\nu'(r) = (\underline{q}, \overline{q}) = (a' + (1-r)(b'-a'), d' + (1-r)(c'-d'))$, $0 \le r \le 1$.

Definition 2.4. *Let $\tilde{A} = (\underline{u}_1, \overline{u}_1, \underline{v}_1, \overline{v}_1)$ and $\tilde{B} = (\underline{u}_2, \overline{u}_2, \underline{v}_2, \overline{v}_2)$ be two IFNs. The parametric distance between two numbers is defined as follows*

$$d_{\alpha,\beta}(\tilde{A}, \tilde{B}) = \left[\int_\alpha^1 (\overline{u}_1 - \overline{u}_2)^2 + (\underline{u}_1 - \underline{u}_2)^2 dr + \int_0^\beta (\underline{v}_1 - \underline{v}_2)^2 + (\overline{v}_1 - \overline{v}_2)^2 dr \right]^{\frac{1}{2}}. \quad (1)$$

Theorem 2.5. *The defined function (1) is a meter, which means it applies to the following properties:*
1. $d_{\alpha,\beta}(\tilde{A}, \tilde{B}) \ge 0$, $d(\tilde{A}, \tilde{A}) = 0$.
2. $d_{\alpha,\beta}(\tilde{A}, \tilde{B}) = d(\tilde{B}, \tilde{A})$.
3. $d_{\alpha,\beta}(\tilde{A}, \tilde{C}) \le d(\tilde{A}, \tilde{B}) + d(\tilde{B}, \tilde{C})$.

Proof. Items 1 and 2 are clear. For item 3, we have:

$$\begin{aligned} d_{\alpha,\beta}(\tilde{A}, \tilde{C}) &= \left[\int_\alpha^1 (\overline{u}_{\tilde{A}} - \overline{u}_{\tilde{C}})^2 + (\underline{u}_{\tilde{A}} - \underline{u}_{\tilde{C}})^2 dr + \int_0^\beta (\underline{v}_{\tilde{A}} - \underline{v}_{\tilde{C}})^2 + (\overline{v}_{\tilde{A}} - \overline{v}_{\tilde{C}})^2 dr \right]^{\frac{1}{2}} \\ &= \left[\int_\alpha^1 (\overline{u}_{\tilde{A}} - \overline{u}_{\tilde{B}} + \overline{u}_{\tilde{B}} - \overline{u}_{\tilde{C}})^2 + (\underline{u}_{\tilde{A}} - \underline{u}_{\tilde{B}} + \underline{u}_{\tilde{B}} - \underline{u}_{\tilde{C}})^2 dr \right. \\ &\quad \left. + \int_0^\beta (\underline{v}_{\tilde{A}} - \underline{v}_{\tilde{B}} + \underline{v}_{\tilde{B}} - \underline{v}_{\tilde{C}})^2 + (\overline{v}_{\tilde{A}} - \overline{v}_{\tilde{B}} + \overline{v}_{\tilde{B}} - \overline{v}_{\tilde{C}})^2 dr \right]^{\frac{1}{2}} \end{aligned}$$

$$\begin{aligned}
&\leq \Big[\int_\alpha^1 (\overline{u}_{\tilde{A}} - \overline{u}_{\tilde{B}})^2 + (\overline{u}_{\tilde{B}} - \overline{u}_{\tilde{C}})^2 + (\underline{u}_{\tilde{A}} - \underline{u}_{\tilde{B}})^2 + (\underline{u}_{\tilde{B}} - \underline{u}_{\tilde{C}})^2 dr \\
&\quad + \int_0^\beta (\underline{v}_{\tilde{A}} - \underline{v}_{\tilde{B}})^2 + (\underline{v}_{\tilde{B}} - \underline{v}_{\tilde{C}})^2 + (\overline{v}_{\tilde{A}} - \overline{v}_{\tilde{B}})^2 + (\overline{v}_{\tilde{B}} - \overline{v}_{\tilde{C}})^2 dr\Big]^{\frac{1}{2}} \\
&\leq d(\tilde{A}, \tilde{B}) + d(\tilde{B}, \tilde{C}).
\end{aligned}$$

□

Assume $\tilde{A} = (a, p, q, p', q')$ and $\tilde{B} = (b, t, s, t', s')$ be two LR IFNs. Then

$$\begin{aligned}
d^2_{\alpha,\beta}(\tilde{A}, \tilde{B}) &= 2(a-b)^2 + (t-p)^2 \int_\alpha^1 (R^{-1}(r))^2 dr + (t'-p')^2 \int_0^\beta (R^{-1}(r))^2 dr \\
&\quad + (q-s)^2 \int_\alpha^1 (L^{-1}(r))^2 dr + (q'-s')^2 \int_0^\beta (L^{-1}(r))^2 dr \\
&\quad + (a-b)(t-p) \int_\alpha^1 L^{-1}(r) dr + (a-b)(t'-p') \int_0^\beta L^{-1}(r) dr \\
&\quad - (a-b)(s-q) \int_\alpha^1 L^{-1}(r) dr - (a-b)(s'-q') \int_0^\beta L^{-1}(r) dr.
\end{aligned}$$

In a special case where two numbers are triangular, we have

$$\begin{aligned}
d^2_{\alpha,\beta}(\tilde{A}, \tilde{B}) &= 2(a-b)^2[1-\alpha] + (t-p)^2(\frac{(1-\alpha)^3}{3}) + (t'-p')^2(\frac{\beta^3}{3}) \\
&\quad + (q-s)^2(\frac{(1-\alpha)^3}{3}) + 2(a-b)^2\beta + (q'-s')^2(\frac{\beta^3}{3}) \\
&\quad + 2(a-b)(t-p)(\frac{1}{2} - \alpha + \frac{\alpha^2}{2}) + 2(a-b)(t'-p')(\frac{\beta^2}{2}) \\
&\quad - 2(a-b)(s-q)(\frac{1}{2} - \alpha + \frac{\alpha^2}{2}) - 2(a-b)(s'-q')(\frac{\beta^2}{2}).
\end{aligned}$$

With simplify the above equation we have

$$\begin{aligned}
d^2_{\alpha,\beta}(\tilde{A}, \tilde{B}) &= 2(1-\alpha+\beta)[(a-b)^2] + \frac{(1-\alpha)^3}{3}[(p-t)^2 + (q-s)^2] \\
&\quad + \frac{\beta^3}{3}[(p'-t')^2 + (q'-s')^2] + (1-\alpha)^2(a-b)[(t-p) - (s-q)] \\
&\quad + \beta^2(a-b)[(t'-p') - (s'-q')].
\end{aligned}$$

Remark 2.6. *To use this meter, we first scale the data.*

Example 2.7. *Let $\tilde{A} = (3, 2, 1, 1, 2)$ and $\tilde{B} = (5, 1, 1, 1, 1)$ be two TriIFNs. In this case, we have*

$$d^2_{\alpha,\beta}(\tilde{A}, \tilde{B}) = \frac{(1-\alpha)^3}{3} \times 0.2355 + \frac{\beta^3}{3} \times 0.2355.$$

Table 1 shows the value of distance $\tilde{A}$ and $\tilde{B}$ for different α and β. Also, in Figure 1, you can see the membership function of two IFNs $\tilde{A}$ and $\tilde{B}$. that alpha and beta *represent different levels of decision making. As alpha increases, a higher decision level is selected, and conversely, if alpha is a lower value, it means that this problem is solved at a lower level of decision making. It is the opposite for beta, i.e. a low beta value indicates problem solving at a higher level of decision making. The data in the Table 1 shows that the distance between these two fuzzy numbers for members of the set that have a membership function of* 0.8 *and a non-membership function of* 0.2 *is equal to* 0.0012 *(second row of the table). The members of the set whose membership is less than* 0.4 *and their non-membership is* 0.8 *is estimated at* 0.0572*.*

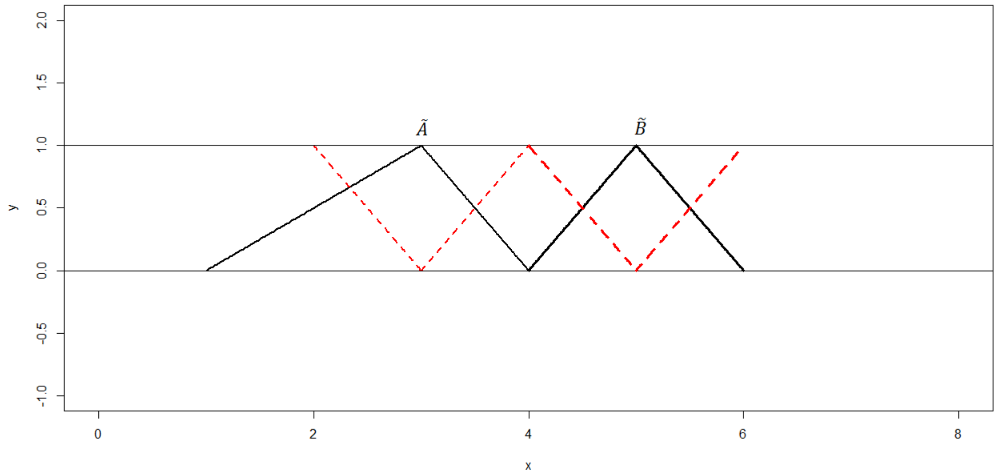

Figure 1: Membership functions of $\tilde{A}$ and $\tilde{B}$ in Example 2.7.

Table 1: $\tilde{A}$ and $\tilde{B}$ distance values for different α and β in Example 2.7.

	α	β	$d^2_{\alpha,\beta}(\tilde{A},\tilde{B})$
1	0.2	0.8	0.0804
2	0.8	0.2	0.0012
3	0.5	0.5	0.0196
4	0.4	0.8	0.0572
5	0.6	0.6	0.0220
6	0.3	0.4	0.0320
7	0.0	1.0	0.1570
8	1	0	0

3 Regression model

Regression models provide models based on which the relationship between independent variables (explanatory) and dependent variable (response) can be investigated. These functional models are designed to predict and control the response variable. If we encounter imprecise variables or imprecise relationships, or if (for example, due to a

small sample size) it is not possible to ensure the correctness of some assumptions, common models lose their validity and efficiency. Then fuzzy regression can provide us with a more suitable model. Based on the fact that each of the received quantities are fuzzy or accurate, the four different categories of the regression model are obtained
1. Fuzzy regression in the case that the relationships between the variables are assumed to be fuzzy.
2. Fuzzy regression in the case that the observations in the response variable and independent variables are imprecise and fuzzy.
3. Fuzzy regression in the case where independent observations are exact numbers, model coefficients and observed responses are fuzzy numbers.
4. Fuzzy regression in the case that both the relationships between the variables and the observations are fuzzy.

In classical regression, it is assumed that the variables and their related observations are accurate. Also the difference between the observed value for the dependent variable and the value obtained through the model, and the overall error of the model, is attributed to the random error related to observations and measurements, the absence of some variables, etc. About these random error statements and its possible distribution, assumptions such as normality, non-correlation, stability of variance... are considered. Based on these assumptions, statistical analysis such as estimation of parameters, prediction of the value of the dependent variable and hypothesis tests related to the model can be performed. But in many cases, one or more of the above assumptions may not be true, or, for example, due to the small sample size, it is not possible to ensure the correctness of some assumptions. For example, in a study, observations related to variables may be inaccurate or inaccurately reported; Or the variables under study may have an imprecise relationship. Also, assumptions such as normality and non-correlation of random error sentences may not hold. In such a situation, classical tools cannot provide suitable criteria for data modelling. One of the possible ways is to use the concept of fuzzy sets to model data in such conditions.

In linear regression with fuzzy input and output, it is assumed that coefficients are positive crisp numbers and ambiguity there exist in the input and output.

$$\tilde{Y}_i = c_0 + c_1\tilde{X}_{i1} + c_2\tilde{X}_{i2} + c_3\tilde{X}_{i3} + \ldots + c_m\tilde{X}_{im} + \epsilon_i, \quad i = 1, 2, \ldots, n,$$

or

$$\tilde{Y}_i = c_0 + \sum_{j=1}^{m} c_j\tilde{X}_{ij} + \epsilon_i, \quad i = 1, 2, \ldots, n.$$

Based on observations

$$(\tilde{Y}_i, \tilde{X}_{i_1}, \tilde{X}_{i_2}, \ldots, \tilde{X}_{i_m}), \quad i = 1, 2, \ldots, n,$$

where $c_0, c_1, \ldots, c_m$ are model coefficients. To simplify the relationships, we assume that the regression and independent variables of both intuitive triangular symmetric fuzzy numbers are as follows

$$\begin{aligned} \tilde{Y}_i &= (y_i, l_i, r_i, l'_i, r'_i), \quad i = 1, 2, \ldots, n, \\ \tilde{X}_{ij} &= (x_{ij}, s_{ij}, w_{ij}, s'_{ij}, w'_{ij}), \quad i = 1, 2, \ldots, n,, \quad i = j, 2, \ldots, m. \end{aligned}$$

Our goal here is to find the coefficients c_i, $i = 1, 2, \ldots, m$ so that the distance between the $\hat{\tilde{Y}}_i = (\mu_{\hat{\tilde{Y}}_i}, \nu_{\hat{\tilde{Y}}_i})$ and $\tilde{Y}_i = (\mu_{\tilde{Y}_i}, \nu_{\tilde{Y}_i})$ is the smallest. For this purpose, the following minimization problem based on the distance defined in Definition 2.3 should be solved

$$\text{minimize} \sum_{i=1}^{n} d^2(\hat{\tilde{Y}}_i, \tilde{Y}_i).$$

The occurrence of multiple collinearity between the predictor variables in the linear regression analysis may cause severe instability in the estimates of the least squares of the regression parameters, which means that the magnitude and sign of the parameters in different samples are significantly unstable. It will be that as a result, the estimation of the lowest second power obtained will not be reliable. The least squares function is written as follows

$$\begin{aligned} S(c_0, c_1, \ldots, c_m) \quad = \quad & \sum_{i=1}^{n}(d^2_{\alpha,\beta}(\hat{\tilde{Y}}_i, \tilde{Y}_i)) = \sum_{i=1}^{n} 2(1-\alpha+\beta)(y_i - c_0 - \sum_{j=1}^{m} c_j x_{ij})^2 \\ & +(\frac{(1-\alpha)^3}{3})[(l_i - c_0 - \sum_{j=1}^{m} c_j s_{ij})^2 + (r_i - c_0 - \sum_{j=1}^{m} c_j w_{ij})^2] \\ & +(\frac{\beta^3}{3})[(l'_i - c_0 - \sum_{j=1}^{m} c_j s'_{ij})^2 + (r'_i - c_0 - \sum_{j=1}^{m} c_j w'_{ij})^2] \\ & +(1-\alpha)^2(y_i - c_0 - \sum_{j=1}^{m} c_j x_{ij})[(l_i - c_0 - \sum_{j=1}^{m} c_j s_{ij}) \\ & -(r_i - c_0 - \sum_{j=1}^{m} c_j w_{ij})] + (\beta^2)(y_i - c_0 - \sum_{j=1}^{m} c_j x_{ij}) \\ & \times[(l'_i - c_0 - \sum_{j=1}^{m} c_j s'_{ij}) - (r'_i - c_0 - \sum_{j=1}^{m} c_j w'_{ij})]. \end{aligned}$$

This function should be minimized with respect to c_0, $c_1, \ldots, c_m$. If we write these equations in the form of a matrix, the data and results will be displayed more compactly. For this purpose, we write the least squares equation as follows

$$\begin{aligned} S(c_0, c_1, \ldots, c_m) = & \, 2(1-\alpha+\beta)((\boldsymbol{Y} - \boldsymbol{XC})'(\boldsymbol{Y} - \boldsymbol{XC})) + A((\boldsymbol{L} - \boldsymbol{SC})'(\boldsymbol{L} - \boldsymbol{SC}) \\ & +(\boldsymbol{R} - \boldsymbol{WC})'(\boldsymbol{R} - \boldsymbol{WC})) + B((\boldsymbol{L}^\star - \boldsymbol{S}^\star\boldsymbol{C})'(\boldsymbol{L}^\star - \boldsymbol{S}^\star\boldsymbol{C}) \\ & +(\boldsymbol{R}^\star - \boldsymbol{W}^\star\boldsymbol{C})'(\boldsymbol{R}^\star - \boldsymbol{W}^\star\boldsymbol{C})) + D[(\boldsymbol{Y} - \boldsymbol{XC})'((\boldsymbol{L} - \boldsymbol{SC}) \\ & -(\boldsymbol{R} - \boldsymbol{WC}))] + E[(\boldsymbol{Y} - \boldsymbol{XC})'((\boldsymbol{L}^\star - \boldsymbol{S}^\star\boldsymbol{C}) - (\boldsymbol{R}^\star - \boldsymbol{W}^\star\boldsymbol{C}))], \end{aligned}$$

where in $A = \dfrac{(1-\alpha)^3}{3}$, $B = \dfrac{\beta^3}{3}$, $D = (1-\alpha)^2$ and $E = \beta^2$

$$\boldsymbol{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \boldsymbol{L} = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{bmatrix}, \quad \boldsymbol{R} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}, \quad \boldsymbol{L}^\star = \begin{bmatrix} l'_1 \\ l'_2 \\ \vdots \\ l'_n \end{bmatrix}, \quad \boldsymbol{R}^\star = \begin{bmatrix} r'_1 \\ r'_2 \\ \vdots \\ r'_n \end{bmatrix},$$

$$\boldsymbol{X}=\begin{bmatrix}1 & x_{11} & x_{12} & \dots & x_{1m}\\ 1 & x_{21} & x_{22} & \dots & x_{2m}\\ . & . & . & \dots & .\\ . & . & . & \dots & .\\ . & . & . & \dots & .\\ 1 & x_{n1} & x_{n2} & \dots & x_{nm}\end{bmatrix},\quad \boldsymbol{S}=\begin{bmatrix}1 & s_{11} & s_{12} & \dots & s_{1m}\\ 1 & s_{21} & s_{22} & \dots & s_{2m}\\ . & . & . & \dots & .\\ . & . & . & \dots & .\\ . & . & . & \dots & .\\ 1 & s_{n1} & s_{n2} & \dots & s_{nm}\end{bmatrix},$$

$$\boldsymbol{W}=\begin{bmatrix}1 & w_{11} & w_{12} & \dots & w_{1m}\\ 1 & w_{21} & w_{22} & \dots & w_{2m}\\ . & . & . & \dots & .\\ . & . & . & \dots & .\\ . & . & . & \dots & .\\ 1 & w_{n1} & w_{n2} & \dots & w_{nm}\end{bmatrix},\quad \boldsymbol{S}^{\star}=\begin{bmatrix}1 & s'_{11} & s'_{12} & \dots & s'_{1m}\\ 1 & s'_{21} & s'_{22} & \dots & s'_{2m}\\ . & . & . & \dots & .\\ . & . & . & \dots & .\\ . & . & . & \dots & .\\ 1 & s'_{n1} & s'_{n2} & \dots & s'_{nm}\end{bmatrix},$$

$$\boldsymbol{W}^{\star}=\begin{bmatrix}1 & w'_{11} & w'_{12} & \dots & w'_{1m}\\ 1 & w'_{21} & w'_{22} & \dots & w'_{2m}\\ . & . & . & \dots & .\\ . & . & . & \dots & .\\ . & . & . & \dots & .\\ 1 & w'_{n1} & w'_{n2} & \dots & w'_{nm}\end{bmatrix},\quad \boldsymbol{C}=\begin{bmatrix}c_0\\ c_1\\ .\\ .\\ .\\ c_m\end{bmatrix}.$$

The least squares estimators should apply in the following relationship

$$\left.\frac{\partial S(c_0,c_1,\dots,c_m)}{\partial \boldsymbol{C}}\right|_{\hat{\boldsymbol{C}}}=0.$$

Therefore, the least squares estimator of C, under the condition of having the inverse of D, becomes as follows.

$$\begin{aligned}\hat{\boldsymbol{C}}=&\Big(4(1-\alpha+\beta)\boldsymbol{X}'\boldsymbol{X}+2A(\boldsymbol{S}'\boldsymbol{S}+\boldsymbol{W}'\boldsymbol{W})+2B(\boldsymbol{S}^{\star\prime}\boldsymbol{S}^{\star}+\boldsymbol{W}^{\star\prime}\boldsymbol{W}^{\star})\\ &+D(\boldsymbol{X}'\boldsymbol{S}-\boldsymbol{X}'\boldsymbol{W})+E(\boldsymbol{X}'\boldsymbol{S}^{\star}-\boldsymbol{X}'\boldsymbol{W}^{\star})\Big)^{-1}\\ &\times\Big(4(1-\alpha+\beta)\boldsymbol{X}'\boldsymbol{Y}+2A(\boldsymbol{S}'\boldsymbol{L}+\boldsymbol{W}'\boldsymbol{R})+2B(\boldsymbol{S}^{\star\prime}\boldsymbol{L}^{\star}+\boldsymbol{W}^{\star\prime}\boldsymbol{R}^{\star})\\ &+D((\boldsymbol{S}-\boldsymbol{W})'\boldsymbol{Y}+\boldsymbol{X}'(\boldsymbol{L}-\boldsymbol{R}))+E((\boldsymbol{S}^{\star}-\boldsymbol{W}^{\star})'\boldsymbol{Y}+\boldsymbol{X}'(\boldsymbol{L}^{\star}-\boldsymbol{R}^{\star}))\Big).\end{aligned}$$

If the triangular fuzzy numbers are intuitively symmetric, then $\boldsymbol{L}=\boldsymbol{L}^{\star}$, $\boldsymbol{R}=\boldsymbol{R}^{\star}$, $\boldsymbol{S}=\boldsymbol{S}^{\star}$ and $\boldsymbol{W}=\boldsymbol{W}^{\star}$. In this case, the least squares estimator is equal to:

$$\begin{aligned}\hat{\boldsymbol{C}}=&(4(1-\alpha+\beta)\boldsymbol{X}'\boldsymbol{X}+2(\boldsymbol{S}'\boldsymbol{S}+\boldsymbol{W}'\boldsymbol{W})(A+B)\\ &+(\boldsymbol{X}'\boldsymbol{S}-\boldsymbol{X}'\boldsymbol{W})(D+E))^{-1}(4(1-\alpha+\beta)\boldsymbol{X}'\boldsymbol{Y}+2(\boldsymbol{S}'\boldsymbol{L}+\boldsymbol{W}'\boldsymbol{R})(A+B)\\ &+((\boldsymbol{S}-\boldsymbol{W})'\boldsymbol{Y}+\boldsymbol{X}'(\boldsymbol{L}-\boldsymbol{R}))(D+E)).\end{aligned}\tag{2}$$

Considering that to use the meter introduced in definition 2.4, it is necessary to divide the data of each variable by the maximum value, the middle points of each variable, so the parameters estimated by this relationship are slightly different from the actual parameters. The actual parameters can be found using the following transformation:

$$\begin{aligned}\hat{C_{org_j}}&=\hat{C_{new j}}*\max y_i/\max x_{ij},\quad j=1,2,...,m,\quad i=1,2,...,n\\ \hat{C_{org_0}}&=\text{mean}(y)-\sum_{j=1}^{m}\hat{C_{org_j}}*\text{mean}(x_j).\end{aligned}$$

In this model, the sum of squared error and mean squared error are calculated respectively using the following relations.

$$SSE = \sum_{i=1}^{n} d^2_{\alpha,\beta}(\tilde{Y}, \hat{\tilde{Y}}), \qquad MSE = \frac{1}{n} SSE.$$

4 Numerical examples

This section illustrates how to use the proposed intuitionistic fuzzy linear least squares regression model by presenting two numerical examples, one with simulated fuzzy numbers and the other with actual fuzzy numbers. The capabilities of the suggested regression model are evaluated alongside those of the another model already in existence.

Example 4.1. *The inputs and outputs are all TriFN and the sample data are from Table 3 in Sakawa and Yano (1992) where they were ordinary triangular numbers but we have converted them to intuitive fuzzy triangular numbers here. In this example, all the data are intuitive symmetric triangular fuzzy numbers whose non-membership function is the same as the membership function (Table 2). In Table 3, the estimation of the parameters of the data is recorded. Also, the predicted values for two points* $x_1 = (6, 0.75)$ *and* $x_2 = (9, 0.43)$ *have been calculated and the obtained values are recorded in the 5th and 6th columns of the Table 3. Values are calculated for different decision levels of* α *and* β*. As the value of alpha increases, the estimate of the slope of the line decreases while the width from the origin increases. while with the increase of beta, the slope increases and the width from the intercept decreases. Figures 2 to 5 show the observed and fitted values at different alpha and beta levels. Figure 1 shows the membership functions for the estimated values and the observed values. As can be seen, the membership functions of the two values are relatively close to each other.*

Table 2: Sample data in Sakawa and Yano (1992).

No.	X	Y
1	$(2; 0.5)_T$	$(4.5, 0.5)_T$
2	$(3.5; 0.5)_T$	$(5.5, 0.5)_T$
3	$(5.5, 1)_T$	$(7.5, 1)_T$
4	$(7, 0.5)_T$	$(6.5, 0.5)_T$
5	$(8.5, 0.5)_T$	$(8.5, 1)_T$
6	$(10.5, 1)_T$	$(8, 1)_T$
7	$(11, 0.5)_T$	$(10.5, 0.5)_T$
8	$(12.5, 0.5)_T$	$(9.5, 0.5)_T$

For example, at the decision level greater than 0.99*, the values of* C_0 *and* C_1 *are equal to* 3.8497 *and* 0.4909*, respectively. Also, the prediction of two* $x_1 = (6, 0.75)$ *and* $x_2 = (9, 0.43)$ *points is* $(6.7954, 0.3582)$ *and* $(8.2682, 0.2111)$ *based on this model. As you can see, as the decision level and alpha value increase, the median value increases while the width of the triangle decreases. Figures 3, 4, and 5 show the estimated values and the observed values for different values of* α *and* β*, which are actually different levels of decision making.*

Table 4 contains information about sum of squared error and MSE for Example 4.1.

Table 3: Estimated model parameters and prediction values for two $x_1 = (6, 0.75)$ and $x_2 = (9, 0.43)$ points at different levels of α and β decision making for the data in Table 2.

	α	β	$\hat{C}_{New} = (C_0, C_1)_{New}$	$\hat{C} = (C_0, C_1)$	$\hat{\tilde{Y}}_1 = (\hat{y}, \hat{l})$	$\hat{\tilde{Y}}_2 = (\hat{y}, \hat{l})$
1	0.2	0.8	(0.1849,0.8281)	(2.3018,0.6956)	(6.4756,0.5217)	(8.5624,0.2991)
2	0.8	0.2	(0.3448,0.6137)	(3.6636,0.5155)	(6.7569,0.3867)	(8.3036,0.2217)
3	0.5	0.5	(0.2634,0.7229)	(2.9701,0.6073)	(6.6136,0.4554)	(8.4354,0.2611)
4	0.4	0.8	(0.2033,0.8035)	(2.4582,0.6749)	(6.5079,0.5062)	(8.5327,0.2902)
5	0.6	0.6	(0.2549,0.7343)	(2.8978,0.6168)	(6,5987,0.4626)	(8.4492,0.2652
6	0.3	0.4	(0.2326,0.7642)	(2.7076,0.6420)	(6.5594,0.4815)	(8.4853,0.2760)
7	0	1	(0.1463,0.8799)	(1.9729,0.7391)	(6.4076,0.5543)	(8.6250,0.3178)
8	0.99	0.01	(0.3666,0.5845)	(3.8497,0.4909)	(6.7954,0.3582)	(8.2682,0.2111)
9	0.4	0.3	(0.2574,0.7306)	(2.9213,0.6137)	(6.6035,0.4603)	(8.4447,0.2639)

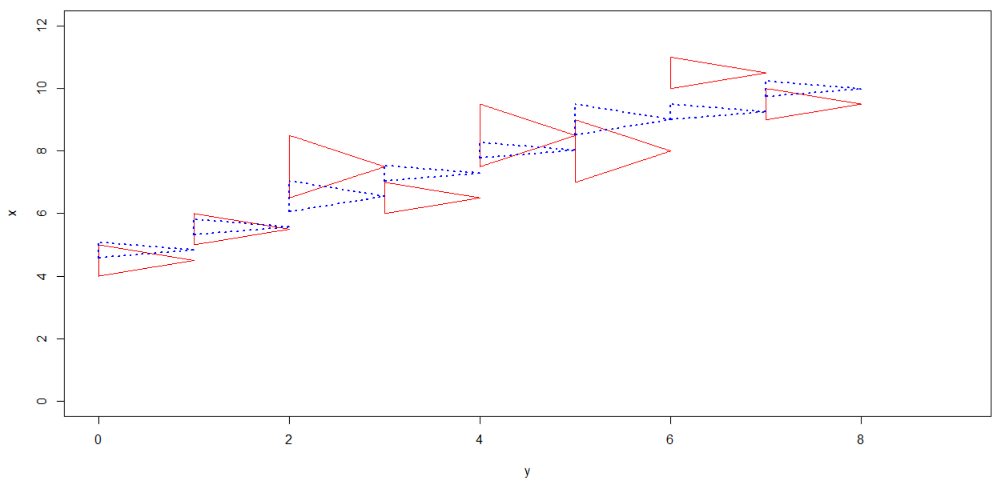

Figure 2: Membership functions of $\tilde{y}$ and $\hat{\tilde{y}}$, $\alpha = 0.99$, $\beta = 0.01$ in Example 4.1.

Table 4: $\tilde{Y}$ and $\hat{\tilde{Y}}$ distance values for different α and β in Example 4.1.

	α	β	$\sum_{i=1}^{8} d^2_{\alpha,\beta}(\tilde{Y}, \hat{\tilde{Y}})$	MSE
1	0.2	0.8	2.8349	0.3543
2	0.8	0.2	0.0816	0.0102
3	0.5	0.5	0.7460	0.0932
4	0.4	0.8	2.0408	0.2551
5	0.6	0.6	0.8251	0.1031
6	0.3	0.4	1.1690	0.1461
7	0.0	1.0	5.4078	0.6760
8	0.99	0.01	0.0022	0.0003
9	0.4	0.3	0.7189	0.0898

Remark 4.2. *To check the stability of the coefficients with respect to data changes*

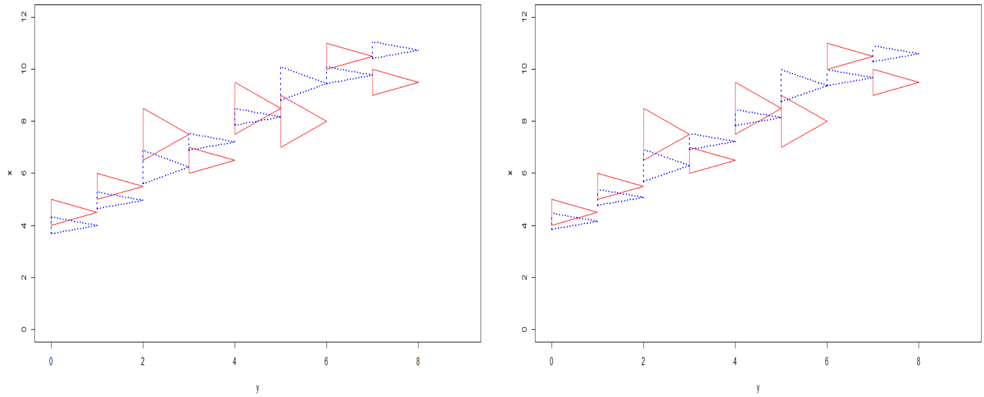

Figure 3: Membership functions of $\tilde{y}$ and $\tilde{\hat{y}}$, $\alpha = 0.3$, $\beta = 0.4$ (left) and $\alpha = 0.4$, $\beta = 0.3$ (right) in Example 4.1.

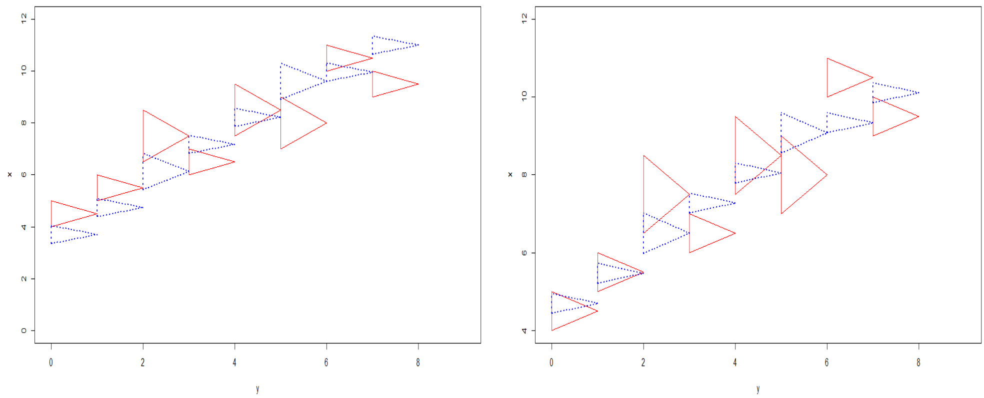

Figure 4: Membership functions of $\tilde{y}$ and $\tilde{\hat{y}}$, $\alpha = 0.2$, $\beta = 0.8$ (left) and $\alpha = 0.8$, $\beta = 0.2$ (right) in Example 4.1.

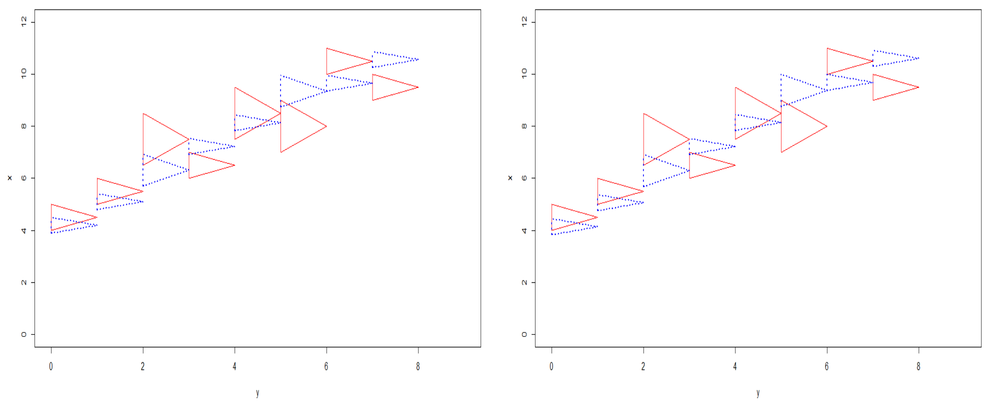

Figure 5: Membership functions of $\tilde{y}$ and $\tilde{\hat{y}}$, $\alpha = 0.5$, $\beta = 0.5$ (left) and $\alpha = 0.6$, $\beta = 0.6$ (right) in Example 4.1.

or the impact of outliers on the coefficients, we changed the values of x and y and re-estimated the regression coefficients for different values of α and β. The results and its comparison with the initial coefficients can be seen in Table 5. $x_1 = (15, 1)_T$ and $y_1 = (8, 0.5)_T$. We also changed the values $x_7 = (2.5, 0.5)_T$ and $y_7 = (3, 0.5)_T$.

Table 5: Estimate the parameters of the model by changing the first and seventh samples separately with fuzzy numbers $x_1 = (15, 1)_T$ and $y_1 = (8, 0.5)_T$ and $x_7 = (2.5, 0.5)_T$ and $y_7 = (3, 0.5)_T$ in Example 4.1.

	α	β	$\hat{C}_{New(1)} = (C_0, C_1)_{New(1)}$	$\hat{C} = (C_0, C_1)$	$\hat{C}_{New(7)} = (C_0, C_1)_{New(7)}$
1	0.2	0.8	$(2.7432, 0.5722)$	$(2.3018, 0.6956)$	$(2.2263, 0.6767)$
2	0.8	0.2	$(5.0346, 0.3228)$	$(3.6636, 0.5155)$	$(3.1541, 0.5340)$
3	0.5	0.5	$(3.7831, 0.4590)$	$(2.9701, 0.6073)$	$(2.7114, 0.6021)$
4	0.4	0.8	$(2.9737, 0.5471)$	$(2.4582, 0.6749)$	$(2.3456, 0.6584)$
5	0.6	0.6	$(3.6634, 0.4720)$	$(2.8978, 0.6168)$	$(2.6619, 0.6097)$
6	0.3	0.4	$(3.3572, 0.5053)$	$(2.7076, 0.6420)$	$(2.5283, 0.6302)$
7	0	1	$(2.2814, 0.6224)$	$(1.9729, 0.7391)$	$(1.9628, 0.7172)$
8	0.99	0.01	$(5.4057, 0.2824)$	$(3.8497, 0.4909)$	$(3.2637, 0.5171)$
9	0.4	0.3	$(3.7022, 0.4678)$	$(2.9213, 0.6137)$	$(2.6781, 0.6072)$

As it is known, the stability of the model decreases with the increase of alpha value. But instead, for small alpha values, the model is almost robust. Also, when the beta values are large, the parameter estimates have changed less.

Example 4.3. *Despite the fact that it is essential, measuring the cation exchange capacity of soils, also known as CEC, which is one of the significant chemical properties that affects other physical, hydraulic, fertility, and biological characteristics of the soil, is highly expensive and time-consuming. It is also one of the important chemical properties that affects other physical properties of the soil. When trying to accurately anticipate cation exchange capacity, it might be helpful to make use of certain physical and chemical features that are simple to test and do not need a significant financial investment. The goal of this study is to develop a regression model that is appropriate for estimating the cation exchange capacity of tillage in the western areas of Iran by making use of the quantity of sand and organic matter.*
The dataset presented in Table 6 displays the triangular IFNs for cation exchange capacity (CEC), sand content (SAND), and organic matter content (OM). Based on the available data, our study aims to construct a regression model that can effectively forecast the Cation Exchange Capacity (CEC) by using the explanatory variables of Sand (SAND) and Organic Matter (OM).

Table 7 shows the estimated values and the given values for Example 4.3 at different decision levels. We displayed the error values between the given values and the estimated values by two methods of sum of squares and mean of squares for Example 4.3 in Table 8.

Figures 6 to 8 also show the difference between the observed values and the fitted values at different levels of decision making.

For example, if in a new case, we observe $X_1 = (28; 2.8, 4.2)_T$ and $X_2 = (2; 0.2, 0.3)_T$, then for different values for α and β we predict CEC as the IFN in Table 6. For example, at the decision level of 0.4, or in other words, with members who have a membership function greater than0.4 and a non-membership function less than 0.8 (fourth

Table 6: Some of measured soil characteristics (CEC,SAND and OM) as triangular A-IFNs.

No.	$CEC(C\,mol\,(+)/kg)$	$SAND(\%)$	$OM(\%)$
	$\tilde{Y}_i = (y_i; l_i, r_i)_T$	$\tilde{X}_{i1} = (x_{i1}; s_{i1}, w_{i1})_T$	$\tilde{X}_{i2} = (x_{i2}; s_{i2}, w_{i2})_T$
1	$(16.5; 0.82, 1.65)_T$	$(35; 3.5, 5.25)_T$	$(0.88; 0.09, 0.13)_T$
2	$(18.6; 0.93, 1.86)_T$	$(37; 3.7, 5.55)_T$	$(1.13; 0.11, 0.17)_T$
3	$(19.3; 0.97, 1.93)_T$	$(27; 2.7, 4.05)_T$	$(1.31; 0.13, 0.2)_T$
4	$(20.3; 1.02, 2.03)_T$	$(29; 2.9, 4.35)_T$	$(1.98; 0.2, 0.3)_T$
5	$(17.3; 0.87, 1.73)_T$	$(38; 3.8, 5.70)_T$	$(1.02; 0.1, 0.15)_T$
6	$(19.3; 0.97, 1.93)_T$	$(29; 2.9, 4.35)_T$	$(1.52; 0.15, 0.23)_T$
7	$(20.4; 1.02, 2.04)_T$	$(32; 3.2, 4.8)_T$	$(1.29; 0.13, 0.19)_T$
8	$(21.9; 1.09, 2.19)_T$	$(18; 1.8, 2.7)_T$	$(1.33; 0.13, 0.2)_T$
9	$(15.9; 0.8, 1.59)_T$	$(40; 4.0, 6.0)_T$	$(1.71; 0.17, 0.26)_T$
10	$(18.3; 0.92, 1.83)_T$	$(28; 2.8, 4.2)_T$	$(2; 0.2, 0.3)_T$
11	$(22.6; 1.13, 2.26)_T$	$(13; 1.3, 1.95)_T$	$(1.68; 0.17, 0.25)_T$
12	$(23.7; 1.19, 2.37)_T$	$(19; 1.9, 2.85)_T$	$(3.15; 0.22, 0.32)_T$
13	$(24.4; 1.22, 2.44)_T$	$(31; 3.1, 4.65)_T$	$(3.52; 0.35, 0.53)_T$
14	$(21.8; 1.09, 2.18)_T$	$(31; 3.1, 4.65)_T$	$(2.33; 0.23, 0.35)_T$
15	$(23.8; 1.19, 2.38)_T$	$(17; 1.7, 2.55)_T$	$(1.71; 0.17, 0.26)_T$
16	$(20.8; 1.04, 2.08)_T$	$(14; 1.4, 2.1)_T$	$(1.14; 0.11, 0.17)_T$
17	$(17.5; 0.88, 1.75)_T$	$(19; 1.9, 2.85)_T$	$(0.99; 0.1, 0.15)_T$
18	$(17.8; 0.89, 1.78)_T$	$(28; 2.8, 4.2)_T$	$(1.14; 0.11, 0.17)_T$
19	$(20.2; 1.01, 2.02)_T$	$(26; 2.6, 3.9)_T$	$(1.46; 0.15, 0.22)_T$
20	$(20.0; 1.00, 2.00)_T$	$(32; 3.2, 4.8)_T$	$(1.81; 0.18, 0.27)_T$
21	$(22.8; 1.14, 2.28)_T$	$(10; 1.0, 1.5)_T$	$(1.38; 0.14, 0.21)_T$
22	$(19.1; 0.96, 1.91)_T$	$(38; 3.8, 5.7)_T$	$(0.84; 0.08, 0.13)_T$
23	$(12.1; 0.6, 1.21)_T$	$(49; 4.9, 7.35)_T$	$(1.48; 0.15, 0.22)_T$
24	$(12.8; 0.64, 1.28)_T$	$(42; 4.2, 6.3)_T$	$(1.08; 0.11, 0.16)_T$

Table 7: Estimate the parameters of the linear regression model for the data in Table 6 at different decision levels and predict the response variable with $X_1 = (28; 2.8, 4.2)_T$ and $X_2 = (2; 0.2, 0.3)_T$.

	α	β	$\hat{C}_{New} = (C_0, C_1, C_2)_{New}$	$\hat{C}_{org} = (C_0, C_1, C_2)_{org}$	$\hat{\tilde{Y}} = (\hat{y}, \hat{l}, \hat{r})$
1	0.2	0.8	$(0.2371, 0.1849, 0.8910)$	$(5.7843, 0.0921, 6.1764)$	$(20.7149, 1.4931, 2.2396)$
2	0.8	0.2	$(0.7874, -0.3223, 0.4140)$	$(19.2124, -0.1606, 2.8701)$	$(20.4584, 0.1246, 0.1869)$
3	0.5	0.5	$(0.4404, -0.0025, 0.7156)$	$(10.7457, -0.0012, 4.9608)$	$(20.6325, 0.9887, 1.4830)$
4	0.4	0.8	$(0.2770, 0.1482, 0.8570)$	$(6.7585, 0.07382, 5.9406)$	$(20.7067, 1.3948, 2.0922)$
5	0.6	0.6	$(0.4132, 0.02271, 0.7394)$	$(10.0824, 0.01131, 5.1256)$	$(20.6503, 1.05679, 1.5852)$
6	0.3	0.4	$(0.3489, 0.0821, 0.7952)$	$(8.5145, 0.0409, 5.5123)$	$(20.6836, 1.2169, 1.8254)$
7	0	1	$(0.1630, 0.2527, 0.9535)$	$(3.9775, 0.1258, 6.6098)$	$(20.7200, 1.6742, 2.5114)$
8	0.99	0.01	$(0.9089, -0.432, 0.3109)$	$(22.1776, -0.215, 2.1550)$	$(20.4616, -0.172, -0.257)$
9	0.4	0.3	$(0.4210, 0.01568, 0.7329)$	$(10.2720, 0.0078, 5.0804)$	$(20.6514, 1.0379, 1.5569)$

row), the estimation of linear regression parameters is equal to $c_0 = 6.7585$, $c_1 = 0.07382$, $c_2 = 5.9406$. *Also, the response variable in these levels is estimated as* $(20.7067, 1.3948, 2.0922)$ *intuitive fuzzy number for* $X_1 = (28; 2.8, 4.2)_T$ *and* $X_2 = (2; 0.2, 0.3)_T$. *You can see the difference between the observed values and the fitted values in this level in Figure 9 on the left.*

Table 8: $\tilde{Y}$ and $\hat{\tilde{Y}}$ distance values for different α and β in Example 4.3.

	α	β	$\sum_{i=1}^{24} d^2_{\alpha,\beta}(\tilde{Y},\hat{\tilde{Y}})$	MSE
1	0.2	0.8	48.2016	2.0084
2	0.8	0.2	0.8228	0.0343
3	0.5	0.5	12.6027	0.5251
4	0.4	0.8	34.9224	1.4551
5	0.6	0.6	14.0269	0.5844
6	0.3	0.4	20.0345	0.8348
7	0	1	89.9194	3.7466
8	0.99	0.01	0.0044	0.00018
9	0.4	0.3	12.164	0.5068

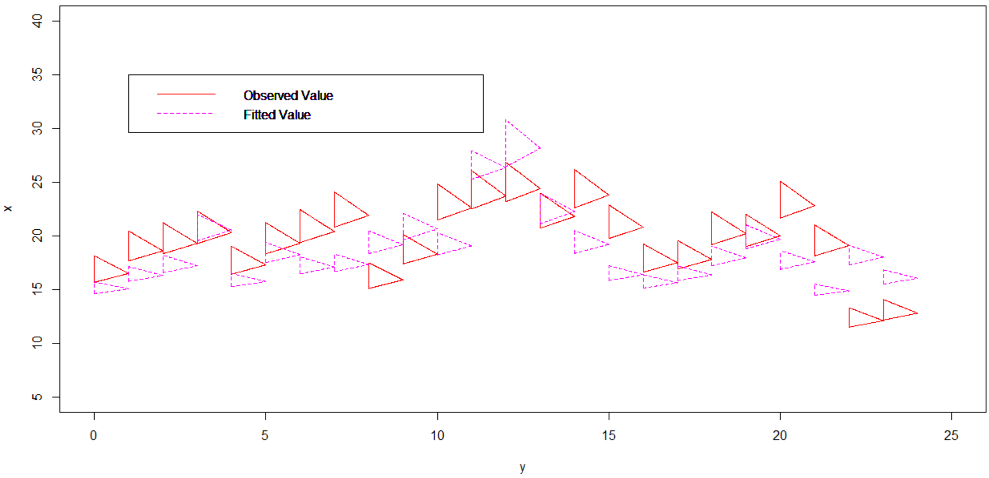

Figure 6: Membership functions of $\tilde{y}$ and $\hat{\tilde{y}}$, $\alpha = 0.5$, $\beta = 0.5$ in Example 4.3

5 Conclusion

Usually, in many statistical analyses where we are faced with real data sets, accurate measurement is not possible and the data is ambiguous. One of the most useful and widely used statistical analyses in finding a relationship between two or more variables based on a sample of observations from the community is linear or non-linear regression analysis. Now, if the observed samples of the variables are imprecise or there are ambiguities in the relationships between them (model coefficients), then fuzzy regression can be used. In this article, the problem of linear regression in the intuitive fuzzy space and providing a new parametric distance measure was investigated. One of the outstanding features of this model was that the decision maker could determine the parameters of the model at different levels of decision-making by choosing appropriate cuts. In this way, different decision makers can estimate the parameters of the model according to their desired level of decision making. At the end, by solving a practical example, the ability and accuracy of the presented model are shown.

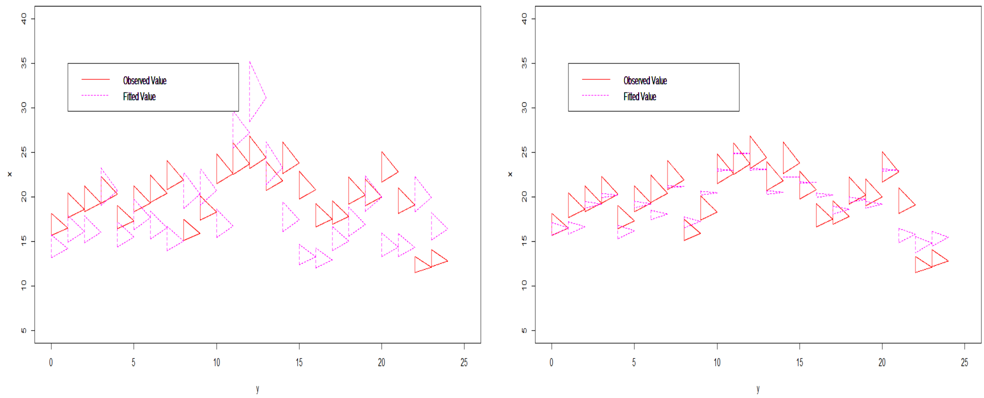

Figure 7: Membership functions of $\tilde{y}$ and $\tilde{\hat{y}}$, $\alpha = 0$, $\beta = 1$ (left) and $\alpha = 0.99$, $\beta = 0.01$ (right) in Example 4.3.

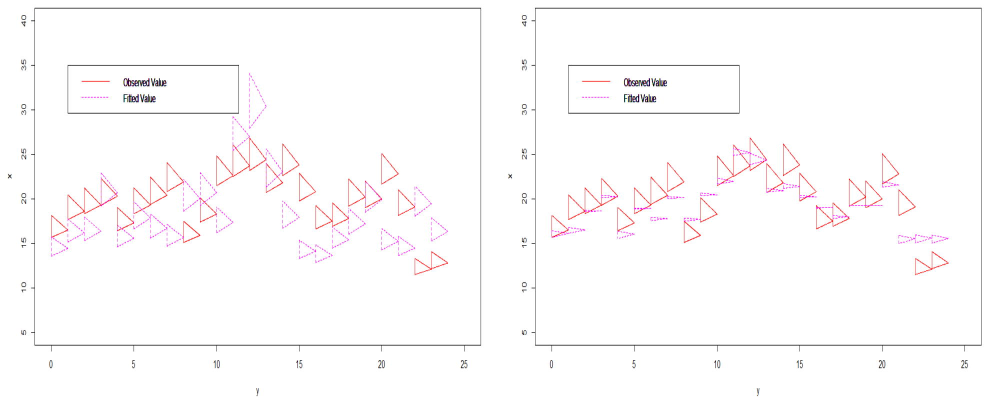

Figure 8: Membership functions of $\tilde{y}$ and $\tilde{\hat{y}}$, $\alpha = 0.2$, $\beta = 0.8$ (left) and $\alpha = 0.8$, $\beta = 0.2$ (right) in Example 4.3.

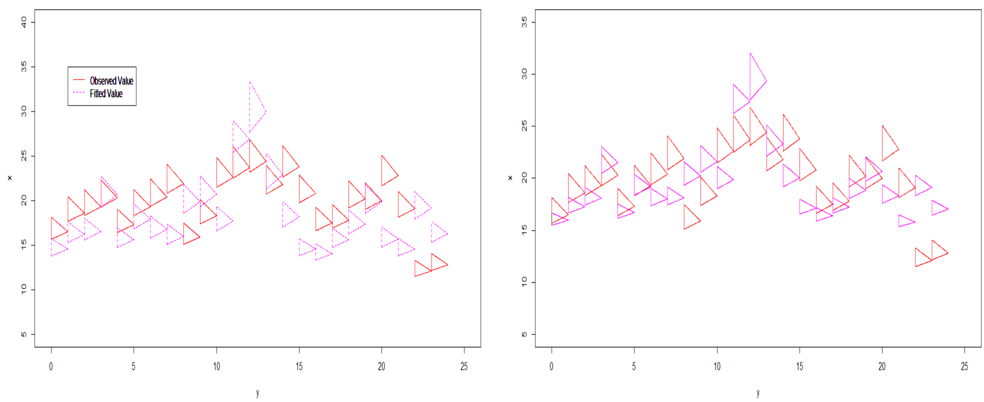

Figure 9: Membership functions of $\tilde{y}$ and $\tilde{\hat{y}}$, $\alpha = 0.4$, $\beta = 0.8$ (left) and $\alpha = 0.4$, $\beta = 0.3$ (right) in Example 4.3.

References

Arefi, M. and Taheri, S.M. (2014). Least-squares regression based on atanassov's intuitionistic fuzzy inputs-outputs and atanassov's intuitionistic fuzzy parameters. *IEEE Transactions on Fuzzy Systems*, **23**(4):1142–1154.

Choi, S.H., Jung, H.Y. and Kim, H. (2019). Ridge fuzzy regression model. *International Journal of Fuzzy Systems*, **21**(7):2077–2090.

Chakravarty, S., Demirhan, H. and Baser, F. (2022). Modified fuzzy regression functions with a noise cluster against outlier contamination. *Expert Systems with Applications*, 205:117717.

Chen, T.Y. (2007). A note on distances between intuitionistic fuzzy sets and/or interval-valued fuzzy sets based on the Hausdorff metric. *Fuzzy Sets and Systems*, **158**(22):2523–2525.

Chen, C. and Deng, X. (2020). Several new results based on the study of distance measures of intuitionistic fuzzy sets. *Iranian Journal of Fuzzy Systems*, **17**(2):147–163.

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

Garg, H. and Rani, D. (2021). Novel similarity measure based on the transformed right-angled triangles between intuitionistic fuzzy sets and its applications. *Cognitive Computation*, **13**(2):447–465.

Grzegorzewski, P. (2004). Distances between intuitionistic fuzzy sets and/or interval-valued fuzzy sets based on the Hausdorff metric. *Fuzzy Sets and Systems*, **148**(2):319–328.

Hazarika, B.B., Gupta, D. and Borah, P. (2021). An intuitionistic fuzzy kernel ridge regression classifier for binary classification. *Applied Soft Computing*, **112**:107816.

Hatzimichailidis, A.G., Papakostas, G.A. and Kaburlasos, V.G. (2012). A novel distance measure of intuitionistic fuzzy sets and its application to pattern recognition problems. *International Journal of Intelligent Systems*, **27**(4):396–409.

Hesamian, G., Akbari, M.G. and Roozbeh, M. (2020). Intuitionistic fuzzy partial logistic regression model using ridge methodology. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **28**(04):527–543.

Hesamian, G. and Akbari, M.G. (2017). Semi-parametric partially logistic regression model with exact inputs and intuitionistic fuzzy outputs. *Applied Soft Computing*, **58**:517–526.

Hesamian, G. and Akbari, M.G. (2023). A fuzzy Bayesian regression model with Gaussian process prior based on exact predictors and fuzzy responses. *Artificial Intelligence Review*, **56**(11):13765–13785.

Hesamian, G., Johannssen, A. and Chukhrova, N. (2024). An explainable fused lasso regression model for handling high-dimensional fuzzy data. *Journal of Computational and Applied Mathematics*, **441**:115721.

Hung, W.L. and Yang, M.S. (2007). Similarity measures of intuitionistic fuzzy sets based on Lp metric. *International Journal of Approximate Reasoning*, **46**(1):120–136.

Hung, W. L. and Yang, M. S. (2008). On the J-divergence of intuitionistic fuzzy sets with its application to pattern recognition. *Information Sciences*, **178**(6):1641–1650.

Jiang, Q., Jin, X., Lee, S.J. and Yao, S. (2019). A new similarity/distance measure between intuitionistic fuzzy sets based on the transformed isosceles triangles and its applications to pattern recognition. *Expert Systems with Applications*, **116**:439–453.

Khasanzoda, N., Zicmane, I., Beryozkina, S., Safaraliev, M., Sultonov, S. and Kirgizov, A. (2022). Regression model for predicting the speed of wind flows for energy needs based on fuzzy logic. *Renewable Energy*, **191**:723–731.

Leski, J.M. (2023). Fuzzy double-ordered c-regression models based on fuzzy S-estimators. *Fuzzy Sets and Systems*, **465**:108531.

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.

Ma, M., Friedman, M., and Kandel, A. (1999). A new fuzzy arithmetic. *Fuzzy Sets and Systems*, **108**(1):83–90.

Mahanta, J. and Panda, S. (2021). A novel distance measure for intuitionistic fuzzy sets with diverse applications. *International Journal of Intelligent Systems*, **36**(2):615–627.

Nguyen, R., Singh, S.K. and Rai, R. (2023). FuzzyGAN: Fuzzy generative adversarial networks for regression tasks. *Neurocomputing*, **525**:88–110.

Pang, Y., Shi, M., Zhang, L., Song, X. and Sun, W. (2022). PR-FCM: A polynomial regression-based fuzzy C-means algorithm for attribute-associated data. *Information Sciences*, **585**:209–231.

Sakawa, M. and Yano, H. (1992). Multiobjective fuzzy linear regression analysis for fuzzy input-output data. *Fuzzy Sets System*, **47**(2):173–181.

Szmidt, E. and Kacprzyk, J. (2000). Distances between intuitionistic fuzzy sets. *Fuzzy Sets and Systems*, **114**(3):505–518.

Yang, Y. and Chiclana, F. (2012). Consistency of 2D and 3D distances of intuitionistic fuzzy sets. *Expert Systems with Applications*, **39**(10):8665–8670.

Wang, W. and Xin, X. (2005). Distance measure between intuitionistic fuzzy sets. *Pattern Recognition Letters*, **26**(13):2063–2069.