Shahid Bahonar University of Kerman
Journal of Mahani Mathematical Research
Print ISSN: 2251-7952 Online ISSN: 2645-4505

FUZZY NONPARAMETRIC REGRESSION BASED ON K-NEAREST NEIGHBORS AND THE R-NEIGHBORHOOD RADIUS

E. Ebrahimi, M. Arefi ✉, and M.G. Akbari

Article type: Research Article
(Received: 21 August 2025, Received in revised form 03 December 2025)
(Accepted: 13 February 2026, Published Online: 13 February 2026)

Abstract. In this paper, we present four nonparametric methods to fit some fuzzy regression models, when both the explanatory and response variables are fuzzy quantities. In this approach, we first introduced a distance between triangular fuzzy numbers. Then, two fuzzy nonparametric regression models are presented based on the extended version of K-nearest neighbors (KNN) method on fuzzy data (with the same/modified weights). In addition, a new method is investigated to fit two fuzzy nonparametric regression models based on the R-neighborhood radius (RNR) method on fuzzy data (with the same/modified weights). Among these methods, the two methods of KNN and RNR with the modified weights have the better performances than the methods with the same weights. To evaluate the proposed fuzzy nonparametric regression models, two measures of goodness of fit are presented. The application of the proposed methods are studied in modelling some data sets.

Keywords: Fuzzy nonparametric regression, K-nearest neighbors (KNN), R-neighborhood radius (RNR), Goodness of fit.
2020 MSC: 62A86, 62G08, 62J86.

1. Introduction

Fuzzy regression analysis is a suitable and widely used tool for analyzing complex systems when the available variables are reported as imprecise quantities. After the introduction of fuzzy sets by Zadeh [33], several authors have presented different approaches on regression models in the fuzzy environment.

A fuzzy regression model based on the least absolute deviations is investigated by Kelkinnama and Taheri [26], and Zeng et al. [34]. Chachi [11], D'Urso et al. [21], and Hesamian and Akbari [23] studied some approaches to model the fuzzy linear regression based on least squares methods. Some approaches for modelling the clustering fuzzy regression are presented by Arefi [6], Bas and Egrioglu [10], and Dotto et al. [19]. The regression analysis in intuitionistic/type-2 fuzzy environment studied by Akbari and Hesamian [1], Arefi [4], Arefi and Taheri [8], and Hosseinzadeh et al. [25]. The problem of

✉ arefi@birjand.ac.ir, ORCID: 0000-0003-3187-9647
https://doi.org/10.22103/jmmr.2026.25799.1855
Publisher: Shahid Bahonar University of Kerman

How to cite: E. Ebrahimi, M. Arefi, M.G. Akbari, *Fuzzy nonparametric regression based on K-nearest neighbors and the R-neighborhood radius*, J. Mahani Math. Res. 2026; 15(2): 251-267.

robust estimations for parameters of the linear regression models in imprecise environments are studied by Arefi [5], Arefi and Khammar [7], Asadolahi et al. [9], Chachi and Roozbeh [14], Chachi et al. [16], Hesamian and Akbari [22], and Khammar et al. [27]. Some approaches on fuzzy nonparametric regression models are studied by Cheng and Lee [17], Danesh et al. [18], Hesamian et al. [24] and Kong et al. [30]. Also, the other approaches for analyzing the fuzzy linear regression are presented by Akbari and Hesamian [2], Arabpour and Amini [3], Chachi and Jalalvand [12], Chachi et al. [13], Chachi and Taheri [15], D'Urso [20], Khammar et al. [28, 29], and Lu and Wang [31].

In this paper, we present four fuzzy nonparametric regression models, when both the explanatory and response variables are fuzzy quantities. We extend the KNN method with the same and modified weights based on fuzzy data. Also, a new approach based on the R-neighborhood radius (RNR) with the same and modified weights is investigated under fuzzy data. These methods are based on the introduced distance between triangular fuzzy numbers. Among these methods, the KNN and RNR methods with the modified weights have the better performances than the methods with the same weights, because they assign more weights to closer data and less weights to farther data. Note that the KNN methods are based on the K elements in nearest neighbors and the RNR methods are based on the data enclosed in radius R. Therefore, the RNR methods can be performed with more speed and accuracy, when the outliers are in data set.

This paper is organized as follows. In Section 2, some preliminary concepts about fuzzy sets is recalled. Some fuzzy nonparametric regression models based on KNN and RNR methods with the same/modified weights are presented in Section 3. The applications of proposed methods are studied on some numerical data sets in Section 4. Finally, in Section 5, some concluding remarks are provided.

2. **Preliminary concepts**

In this section, we recall some preliminary concepts about fuzzy numbers (for more details, see Zimmermann [35]).
Let X be an universal set. A fuzzy set $\tilde{A}$ of X is defined by its membership function $\tilde{A} : X \to [0, 1]$. The α-cut of $\tilde{A}$ is defined by $\tilde{A}[\alpha] = \left\{ x \in R : \tilde{A}[x] \geq \alpha \right\}$ for $0 < \alpha \leq 1$.

Definition 2.1. A fuzzy set $\tilde{A}$ of the R (real line) is called a fuzzy number, if

i) $\tilde{A}[x] = 1$ for some $x \in R$.
ii) $\tilde{A}[\alpha]$ is a closed bounded interval for $0 < \alpha \leq 1$.

Definition 2.2. A fuzzy number $\tilde{A}$ is called a triangular fuzzy number, and denoted by $\tilde{A} = \left(a^l, a, a^u\right)_T$, if its membership function is defined as follows:

$$\tilde{A}(x) = \begin{cases} \frac{x-a^l}{a-a^l} & a^l < x \leq a, \\ \frac{a^u-x}{a^u-a} & a < x \leq a^u, \\ 0 & o.w. \end{cases}$$

Lemma 2.3. *Suppose that $\tilde{A} = (a^l, a, a^u)_T$ and $\tilde{B} = (b^l, b, b^u)_T$ are two triangular fuzzy numbers and $\lambda \in R - \{0\}$. Then, based on the extension principle, some arithmetic operators is as*

$$\tilde{A} \oplus \tilde{B} = (a^l + b^l, a + b, a^u + b^u)_T,$$

$$\lambda\tilde{A} = \begin{cases} (\lambda a^l, \lambda a, \lambda a^u)_T & \lambda > 0, \\ (\lambda a^u, \lambda a, \lambda a^l)_T & \lambda < 0. \end{cases}$$

Definition 2.4. Let $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_p)$ and $\tilde{\mathbf{y}} = (\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_p)$ be two vectors of triangular fuzzy numbers with $\tilde{x}_i = (x_i^l, x_i, x_i^u)_T$ and $\tilde{y}_i = (y_i^l, y_i, y_i^u)_T$, $i = 1, ..., p$. Then, the distance between $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ is defined as follows:

$$d(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = \sum_{i=1}^{p} \sqrt{\frac{|x_i^l - y_i^l|^2 + |x_i - y_i|^2 + |x_i^u - y_i^u|^2}{3}}. \tag{1}$$

Theorem 2.5. *Let $\tilde{\mathbf{x}}$, $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{z}}$ be three vectors of fuzzy numbers in R^p. Then, (R, d) is a metric space. It satisfy the following properties*

(i) $d(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \geq 0$,

(ii) $d(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = d(\tilde{\mathbf{y}}, \tilde{\mathbf{x}})$,

(iii) $d(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \leq d(\tilde{\mathbf{x}}, \tilde{\mathbf{z}}) + d(\tilde{\mathbf{z}}, \tilde{\mathbf{y}})$.

Remark 2.6. In this paper, we present some approaches to fit the fuzzy nonparametric regression models based on this distance. This distance is an extended version of Euclidean distance on triangular fuzzy data (If $x_i^l = x_i = x_i^u$ in $\tilde{x}_i$ and $y_i^l = y_i = y_i^u$ in $\tilde{y}_i$, then the proposed distance is reduced to Euclidean distance). It is simple and fast in calculations. It is notable that the proposed method does not change by replacing other distances between fuzzy numbers (specially, to choose a suitable distance between LR fuzzy number).

3. Fuzzy Nonparametric Regression

In this section, we want to fit a fuzzy nonparametric regression model, when the available data set is as fuzzy quantities. Assume that the fuzzy nonparametric regression model is as follows:

$$\tilde{y}_i = f(\tilde{\mathbf{x}}_i) + \epsilon,$$

where, $f(.)$ is an unknown, continuous, and differentiable function, and ϵ is an error variable with zero mean.

Now, based on a random sample of size N, we have the observed fuzzy data as $(\tilde{\mathbf{x}}_i, \tilde{y}_i)$, $i = 1, 2, \cdots, N$, where $\tilde{y}_i = (y_i^l, y_i, y_i^u)_T$ and $\tilde{\mathbf{x}}_i = (\tilde{x}_{i1}, \tilde{x}_{i2}, \ldots, \tilde{x}_{ip})$ with $\tilde{x}_{ij} = (x_{ij}^l, x_{ij}, x_{ij}^u)_T$, $j = 1, 2, \cdots, p$. The aim is to fit a fuzzy nonparametric regression model to this fuzzy data set as follows:

$$\widehat{\tilde{y}}_i = \widehat{f}(\tilde{\mathbf{x}}_i).$$

Now, we want to present some new methods based on the K-nearest neighbors (KNN) and R-neighborhood radius (RNR) on fuzzy data as follows.

3.1. **K-Nearest Neighbors (KNN).** The KNN refers to the K elements in a data set that are closest to a given value x, based on a specified distance. In other words, for a given element x in a data set with N elements, the distance between x and every other element is calculated. The K elements with the smallest distances to x are the K-nearest neighbors of x (for more details, see Cheng and Lee [17]).

Now, we introduce two KNN methods based on fuzzy data with the same and modified weights as follows.

3.1.1. *KNN with the same weights.* Let the fuzzy data be as $(\tilde{\mathbf{x}}_i, \tilde{y}_i)$, $i = 1, 2, \cdots, N$. Then, the estimation of the response variable based on KNN method with the same weights is defined as the following procedure:

1) We first select the value $\tilde{\mathbf{x}}_i$, $i = 1, ..., N$.
2) Based on Definition 2.4, we calculate $d(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_i), d(\tilde{\mathbf{x}}_2, \tilde{\mathbf{x}}_i), ..., d(\tilde{\mathbf{x}}_N, \tilde{\mathbf{x}}_i)$, denoted by $d_{1i}, d_{2i}, ..., d_{Ni}$. Also, we order $d_{1i}, d_{2i}, ..., d_{Ni}$, denoted by $d_{(1)i} \leq d_{(2)i} \leq ... \leq d_{(N)i}$.
3) The weighted values $\omega_j(\tilde{\mathbf{x}}_i)$ are obtained as

$$\omega_j(\tilde{\mathbf{x}}_i) = \begin{cases} \frac{1}{K} & if\ j \in J_{\tilde{\mathbf{x}}_i}, \\ 0 & o.w., \end{cases} \qquad j = 1, 2, ..., N,$$

where, the set $J_{\tilde{\mathbf{x}}_i}$ is as

$$\begin{aligned} J_{\tilde{\mathbf{x}}_i} &= \{j : \tilde{\mathbf{x}}_j\ is\ one\ of\ the\ K - nearest\ observations\ to\ \tilde{\mathbf{x}}_i\} \\ &= \left\{j : d_{ji} \in \{d_{(1)i}, d_{(2)i}, ..., d_{(K)i}\}\right\}. \end{aligned}$$

4) Finally, the response value $\widehat{\tilde{y}}_i$ is estimated as

$$\hat{\tilde{y}}_i = \sum_{j=1}^{N} \omega_j(\tilde{\mathbf{x}}_i)\ \tilde{y}_j.$$

For selecting the optimal value of K, see Remark 3.5 in Subsection 3.3.

3.1.2. *KNN with the modified weights.* The estimation of the response variable based on KNN method with the modified weights under the fuzzy data $(\tilde{\mathbf{x}}_i, \tilde{y}_i)$, $i = 1, 2, \cdots, N$, is defined based on the following procedure:

1) We first select the value $\tilde{\mathbf{x}}_i$, $i = 1, ..., N$.
2) Based on Definition 2.4, we calculate $d(\tilde{\mathbf{x}_1}, \tilde{\mathbf{x}}_i), d(\tilde{\mathbf{x}_2}, \tilde{\mathbf{x}}_i), ..., d(\tilde{\mathbf{x}}_N, \tilde{\mathbf{x}}_i)$, denoted by $d_{1i}, d_{2i}, ..., d_{Ni}$. We order $d_{1i}, d_{2i}, ..., d_{Ni}$, denoted by $d_{(1)i} \leq d_{(2)i} \leq ... \leq d_{(N)i}$.
3) The weights $\omega_j(\tilde{\mathbf{x}}_i)$ are obtained as

$$\omega_j(\tilde{\mathbf{x}}_i) = \frac{W_j(\tilde{\mathbf{x}}_i)}{\sum_{j \in J_{\tilde{\mathbf{x}}_i}}^{K} W_j(\tilde{\mathbf{x}}_i)} \quad j = 1, 2, ..., N,$$

where, $W_j(\tilde{\mathbf{x}}_i) = 1 - \frac{d_{ji}}{d_{Ki}}$ and the set $J_{\tilde{\mathbf{x}}_i}$ is as

$$\begin{aligned} J_{\tilde{\mathbf{x}}_i} &= \{j : \tilde{\mathbf{x}}_j \ is\ one\ of\ the\ K - nearest\ observations\ to\ \tilde{\mathbf{x}}_i\} \\ &= \left\{j : d_{ji} \in \{d_{(1)i}, d_{(2)i}, ..., d_{(K)i}\}\right\}. \end{aligned}$$

4) Finally, the value of response variable is estimated as

$$\widehat{\tilde{y}}_i = \sum_{j=1}^{N} \omega_j(\tilde{\mathbf{x}}_i)\ \tilde{y}_j.$$

For selecting the optimal value of K, see Remark 3.5 in Subsection 3.3.

3.2. **R-Neighborhood Radius (RNR).** The primary advantage of the KNN algorithm is its conceptual and computational simplicity. But, it has some limitations as follows:

1- Its computational cost grows substantially with the number of training samples. So, it isn't suitable for large data sets.
2- The KNN is sensitive to outlier data. In particular, if the distribution of data is non-uniform, since the value of K is fixed, the outlier data affect on weights and the values of response variables will not be predicted correctly.

To reduce these limitations, we present a new approach based on the R-neighborhood radius (RNR). In this approach, we only select the nearest neighbors that fall within the fixed radius R. In RNR algorithm, the value of optimal radius R is first determined and then, the distances between a given element x and other elements x_j in data set are calculated. If the computed distance is less than R (i.e. $d(x_j, x) \leq R$), then this element is included in the neighborhood set of x. As a result, the number of elements within the neighborhood set of x may vary depending on the chosen radius R. This makes the calculations faster than the KNN methods.
Now, we introduce two methods based on RNR under fuzzy data with the same and modified weights as follows.

3.2.1. *RNR with the same weights.* Suppose that the available fuzzy data are as $(\tilde{\mathbf{x}}_i, \tilde{y}_i)$, $i = 1, 2, \cdots, N$. Then, the estimation of the response variable based on RNR method with the same weights is defined as the following procedure:

1) We first select the value of radius R, and also, consider the value of $\tilde{\mathbf{x}}_i$, $i = 1, ..., N$.
2) Based on Definition 2.4, we calculate $d(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_i), d(\tilde{\mathbf{x}}_2, \tilde{\mathbf{x}}_i), ..., d(\tilde{\mathbf{x}}_N, \tilde{\mathbf{x}}_i)$, denoted by $d_{1i}, d_{2i}, ..., d_{Ni}$. We order $d_{1i}, d_{2i}, ..., d_{Ni}$, denoted by $d_{(1)i} \leq d_{(2)i} \leq ... \leq d_{(N)i}$.
3) The weights $\omega_j(\tilde{\mathbf{x}}_i)$ are obtained as
$$\omega_j(\tilde{\mathbf{x}}_i) = \begin{cases} \frac{1}{t_i} & if\ j \in J_{\tilde{\mathbf{x}}_i}, \\ & \\ 0 & o.w., \end{cases} \qquad j = 1, 2, ..., N,$$
where, the set $J_{\tilde{\mathbf{x}}_i}$ is as
$$\begin{aligned} J_{\tilde{\mathbf{x}}_i} &= \{j : d_{ji} \leq R,\ j = 1, ..., N\} \\ &= \left\{j : d_{ji} \in \{d_{(1)i}, d_{(2)i}, ..., d_{(t_i)i}\}\right\}, \end{aligned}$$
where, t_i is the number of elements whose distances to the ith element is less than or equal to R.
4) Finally, the response value $\widehat{\tilde{y}}_i$ is estimated as
$$\hat{\tilde{y}}_i = \sum_{j=1}^{N} \omega_j(\tilde{\mathbf{x}}_i)\ \tilde{y}_j.$$

3.2.2. *RNR with the modified weights.* Let the fuzzy data be as $(\tilde{\mathbf{x}}_i, \tilde{y}_i)$, $i = 1, 2, \cdots, N$. Then, the estimation of the response variable based on RNR method with the same weights is defined as the following procedure:

1) We first select the value of radius R, and consider the value of $\tilde{\mathbf{x}}_i$, $i = 1, ..., N$.
2) Based on Definition 2.4, we calculate $d(\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_i), d(\tilde{\mathbf{x}}_2, \tilde{\mathbf{x}}_i), ..., d(\tilde{\mathbf{x}}_N, \tilde{\mathbf{x}}_i)$, denoted by $d_{1i}, d_{2i}, ..., d_{Ni}$. The ordered values of $d_{1i}, d_{2i}, ..., d_{Ni}$ are denoted by $d_{(1)i} \leq d_{(2)i} \leq ... \leq d_{(N)i}$.
3) The weights $\omega_j(\tilde{\mathbf{x}}_i)$ are obtained as
$$\omega_j(\tilde{\mathbf{x}}_i) = \frac{W_j(\tilde{\mathbf{x}}_i)}{\sum_{j \in J_{\tilde{\mathbf{x}}_i}}^{t_i} W_j(\tilde{\mathbf{x}}_i)},\ j = 1, ..., N,$$
where, $W_j(\tilde{\mathbf{x}}_i) = 1 - \frac{d_{ji}}{d_{t_i i}}$ and the set $J_{\tilde{\mathbf{x}}_i}$ is as
$$\begin{aligned} J_{\tilde{\mathbf{x}}_i} &= \{j : d_{ji} \leq R, j = 1, ..., N\} \\ &= \left\{j : d_{ji} \in \{d_{(1)i}, d_{(2)i}, ..., d_{(t_i)i}\}\right\}, \end{aligned}$$
where, t_i is the number of elements whose distances to the ith element is less than or equal to R.

4) Finally, the response value $\widehat{\tilde{y}}_i$ is estimated as

$$\hat{\tilde{y}}_i = \sum_{j=1}^{N} \omega_j(\tilde{\mathbf{x}}_i)\ \tilde{y}_j.$$

For selecting the optimal value of R, see Remark 3.6 in Subsection 3.3.

Remark 3.1. In the proposed methods (KNN or RNR) with the same weights, all neighbors about an element have the same effect. But, it is clear that the data closer to an element are more important than the data located at a further distance. To solve this problem, we present the improved KNN/RNR with the modified weights which the closer neighbors have a greater effect, while the more distant neighbors become less or ineffective.

3.3. **Goodness of fit of the model.** Assume that $\tilde{Y}_i$ and $\hat{\tilde{Y}}_i$, $i = 1, ..., N$ are the observed response values and the predicted response values, respectively. To evaluate the proposed fuzzy nonparametric regression models, we introduce two goodness of fit measures between $\tilde{Y}_i$ and $\hat{\tilde{Y}}_i$ as follows.

Definition 3.2. The mean of similarity measures (MSM) between $\tilde{Y}_i$ and $\hat{\tilde{Y}}_i$, $i = 1, ..., N$, is defined as (see Pappis and Karacapilidis [32])

$$MSM = \frac{1}{n}\sum_{j=1}^{n} S_{PK}(\tilde{Y}_i, \hat{\tilde{Y}}_i), \tag{2}$$

where, $S_{PK}(\tilde{Y}_i, \hat{\tilde{Y}}_i) = \frac{Card(\tilde{Y}_i \cap \hat{\tilde{Y}}_i)}{Card(\tilde{Y}_i \cup \hat{\tilde{Y}}_i)}$ and $Card(\tilde{A}) = \int \tilde{A}(x)dx$ denotes the cardinal number of fuzzy set $\tilde{A}$, $(\tilde{Y}_i \cap \hat{\tilde{Y}}_i)(x) = \min(\tilde{Y}_i(x), \hat{\tilde{Y}}_i(x))$, and $(\tilde{Y}_i \cup \hat{\tilde{Y}}_i)(x) = \max(\tilde{Y}_i(x), \hat{\tilde{Y}}_i(x))$. Also, $\tilde{Y}_i(x)$ and $\hat{\tilde{Y}}_i(x)$ are the membership functions of $\tilde{Y}_i$ and $\hat{\tilde{Y}}_i$, respectively.

Definition 3.3. The mean squared error (MSE) measure is defined as follows

$$MSE = \frac{1}{n}\sum_{j=1}^{n} d^2(\tilde{Y}_i, \hat{\tilde{Y}}_i), \tag{3}$$

where, $d(\tilde{Y}_i, \hat{\tilde{Y}}_i) = \sqrt{(\hat{Y}_i^l - Y_i^l)^2 + (\hat{Y}_i - Y_i)^2 + (\hat{Y}_i^u - Y_i^u)^2}$.

Remark 3.4. The indices MSM and MSE are on the intervals $[0, 1]$ and $[0, \infty)$, respectively. The optimal fuzzy regression model is the model with maximum value of MSM and the minimum value of MSE.

Remark 3.5. For selecting the optimal value of K, we obtain the estimated response variables $\hat{\tilde{y}}_i$, $i = 1, ..., N$, based on the KNN method for $K = 2, ..., N$. Then, based on Definition 3.3, we calculate MSE between $\hat{\tilde{y}}_i$ and $\tilde{y}_i$ for $K = 2, ..., N$. The optimal value is K with minimum value of MSE.

Remark 3.6. For selecting the optimal value of R, we perform the following procedure:

1) First, we calculate the distances $d_{ij} = d(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_j)$, $i, j = 1, 2, ..., N$, and $j > i$. We order d_{ij}, denoted by $d^*_{(1)} \leq d^*_{(2)} \leq ... \leq d^*_{(N(N-1)/2)}$.
2) We obtain the estimated response variables $\hat{\tilde{y}}_i$, $i = 1, ..., N$, based on the RNR method for $R \in \{d^*_{(1)}, d^*_{(2)}, ..., d^*_{(N(N-1)/2)}\}$.
3) Based on Definition 3.3, we calculate MSE between $\hat{\tilde{y}}_i$ and $\tilde{y}_i$ for $R \in \{d^*_{(1)}, d^*_{(2)}, ..., d^*_{(N(N-1)/2)}\}$.
4) The optimal value is R with minimum value of MSE.

4. Numerical examples

In this section, we present some numerical examples to illustrate our proposed approach, using the goodness of fit measures.

Example 4.1. *Consider a study for evaluating the employees engagement by assessing 25 officers working at an inland revenue board of a city in Iran (see Hesamian and Akbari [22]). The main purpose of this study was to provide a regression model which capable of enhancing the understanding of uncertainty about the main individual factors of the employees engagement in the final work outcomes. Four independent variables were used for subjective evaluation of the employees work quality ($\tilde{y}_i$) as follows:*
$\tilde{\mathbf{x}}_1$*: Inability to endure job stress,*
$\tilde{\mathbf{x}}_2$*: Frequency of delay,*
$\tilde{\mathbf{x}}_3$*: Communication and coordination ability,*
$\tilde{\mathbf{x}}_4$*: Performance.*
The data set is listed in Table 1. We want to fit a fuzzy nonparametric regression model as follows

$$\hat{\tilde{y}}_i = \sum_{j=1}^{N} \omega_j(\tilde{\mathbf{x}}_i)\ \tilde{y}_j.$$

First, based on Remarks 3.5 and 3.6, we obtain the optimal values $K = 4$ *and* $R = 21.49$*. Then, we fit some fuzzy nonparametric regression models based on KNN and RNR methods. We compare the proposed methods with some other methods. The results are summarized in Table 2.*

To compare the proposed methods with other methods based on the goodness of fit measures, it can be concluded that the proposed method has a suitable performance. The performances of the proposed fuzzy nonparametric regression models are shown in Figures 1-4. Figures 1 and 3 show the KNN and RNR methods with the same weights and Figures 2 and 4 show the KNN and RNR methods with the modified weights (the + signs are the bounds and centers of observed response variables, and the drawn lines are the fuzzy nonparametric regression models).

Table 1. Data set in Example 4.1.

i	$\tilde{x}_{i1} = (x_{i1}^l, x_{i1}, x_{i1}^u)_T$	$\tilde{x}_{i2} = (x_{i2}^l, x_{i2}, x_{i2}^u)_T$	$\tilde{x}_{i3} = (x_{i3}^l, x_{i3}, x_{i3}^u)_T$	$\tilde{x}_{i4} = (x_{i4}^l, x_{i4}, x_{i4}^u)_T$	$\tilde{y}_i = (y_i^l, y_i, y_i^u)_T$
1	$(79, 87, 95)_T$	$(65, 71, 73)_T$	$(82, 86, 91)_T$	$(86, 93, 96)_T$	$(90, 95, 98)_T$
2	$(71, 79, 87)_T$	$(74, 80, 83)_T$	$(71, 75, 79)_T$	$(65, 67, 70)_T$	$(70, 77, 83)_T$
3	$(14, 24, 32)_T$	$(59, 65, 67)_T$	$(69, 71, 75)_T$	$(29, 35, 40)_T$	$(23, 26, 32)_T$
4	$(51, 60, 67)_T$	$(56, 65, 69)_T$	$(66, 68, 71)_T$	$(61, 66, 70)_T$	$(58, 62, 66)_T$
5	$(87, 92, 93)_T$	$(75, 78, 81)_T$	$(66, 70, 76)_T$	$(81, 86, 89)_T$	$(85, 90, 92)_T$
6	$(55, 59, 61)_T$	$(44, 53, 63)_T$	$(59, 61, 64)_T$	$(64, 66, 68)_T$	$(17, 20, 24)_T$
7	$(78, 82, 86)_T$	$(75, 77, 79)_T$	$(75, 80, 83)_T$	$(70, 72, 76)_T$	$(72, 77, 80)_T$
8	$(48, 52, 56)_T$	$(55, 57, 59)_T$	$(55, 60, 63)_T$	$(67, 69, 73)_T$	$(18, 21, 25)_T$
9	$(82, 85, 87)_T$	$(71, 77, 83)_T$	$(66, 71, 76)_T$	$(61, 68, 71)_T$	$(66, 72, 75)_T$
10	$(63, 66, 70)_T$	$(58, 60, 63)_T$	$(53, 56, 60)_T$	$(69, 73, 75)_T$	$(58, 62, 64)_T$
11	$(65, 69, 72)_T$	$(56, 59, 62)_T$	$(70, 72, 76)_T$	$(66, 70, 73)_T$	$(58, 63, 68)_T$
12	$(57, 59, 61)_T$	$(51, 54, 56)_T$	$(62, 68, 71)_T$	$(62, 67, 70)_T$	$(55, 58, 61)_T$
13	$(89, 90, 96)_T$	$(84, 88, 96)_T$	$(77, 82, 87)_T$	$(80, 85, 87)_T$	$(83, 88, 91)_T$
14	$(66, 70, 76)_T$	$(69, 71, 76)_T$	$(66, 68, 71)_T$	$(60, 64, 66)_T$	$(61, 65, 70)_T$
15	$(51, 55, 59)_T$	$(61, 63, 65)_T$	$(58, 61, 64)_T$	$(66, 68, 70)_T$	$(24, 26, 28)_T$
16	$(86, 90, 93)_T$	$(73, 75, 77)_T$	$(60, 64, 68)_T$	$(73, 78, 83)_T$	$(77, 82, 85)_T$
17	$(56, 60, 64)_T$	$(58, 62, 64)_T$	$(63, 66, 70)_T$	$(63, 65, 69)_T$	$(18, 22, 24)_T$
18	$(66, 68, 71)_T$	$(60, 63, 68)_T$	$(70, 75, 79)_T$	$(75, 80, 86)_T$	$(67, 71, 74)_T$
19	$(60, 67, 72)_T$	$(50, 56, 62)_T$	$(63, 65, 68)_T$	$(39, 42, 47)_T$	$(21, 24, 28)_T$
20	$(82, 88, 93)_T$	$(63, 69, 73)_T$	$(59, 63, 67)_T$	$(73, 76, 80)_T$	$(68, 72, 77)_T$
21	$(88, 90, 93)_T$	$(82, 85, 89)_T$	$(58, 62, 66)_T$	$(74, 78, 81)_T$	$(76, 82, 87)_T$
22	$(55, 59, 62)_T$	$(68, 71, 74)_T$	$(62, 64, 68)_T$	$(54, 57, 60)_T$	$(61, 64, 69)_T$
23	$(83, 87, 91)_T$	$(49, 51, 53)_T$	$(62, 67, 70)_T$	$(58, 60, 62)_T$	$(20, 23, 25)_T$
24	$(70, 74, 80)_T$	$(59, 61, 66)_T$	$(59, 61, 66)_T$	$(77, 82, 85)_T$	$(69, 74, 78)_T$
25	$(60, 63, 65)_T$	$(49, 52, 54)_T$	$(56, 58, 61)_T$	$(60, 62, 64)_T$	$(19, 22, 25)_T$

Table 2. Goodness of fit measures in different regression models in Example 4.1.

Model	*MSE*	*MSM*
KNN with the same weights	312.19	0.36
KNN with the modified weights	69.08	0.49
RNR with the same weights	170.85	0.52
RNR with the modified weights	81.47	0.71
Chachi et al. [16]	1130.84	0.33
Akbari and Hesamian [2]	681.21	0.42
D'Urso et al. [21]	718.24	0.40
Hesamian and Akbari [22]	674.96	0.47

Example 4.2. *Consider the data set given in Table 3 (see D'Urso [20]). In this study, we have two the fuzzy independent variables $\tilde{\mathbf{x}}_1$ as "decision on cooking" and $\tilde{\mathbf{x}}_2$ as "decision on environment", and a fuzzy response variable $\tilde{y}_i$ as "decision on cellar". This data set is the performances of the 30 good-quality Roman restaurants. We want to fit a fuzzy nonparametric regression*

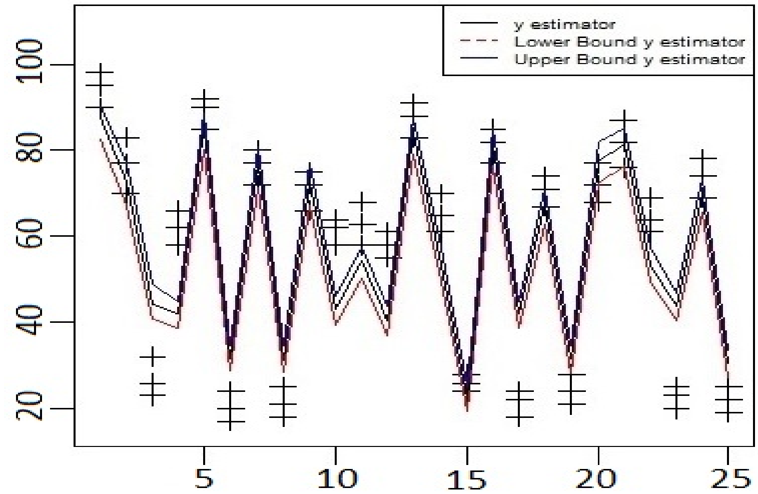

Figure 1. KNN method with the same weights in Example 4.1.

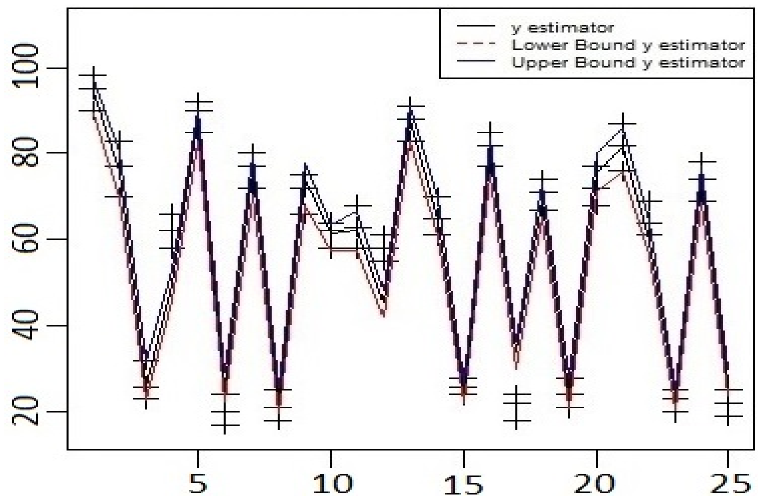

Figure 2. KNN method with the modified weights in Example 4.1.

model as follows

$$\hat{\tilde{y}}_i = \sum_{j=1}^{N} \omega_j(\tilde{\mathbf{x}}_i)\tilde{y}_j.$$

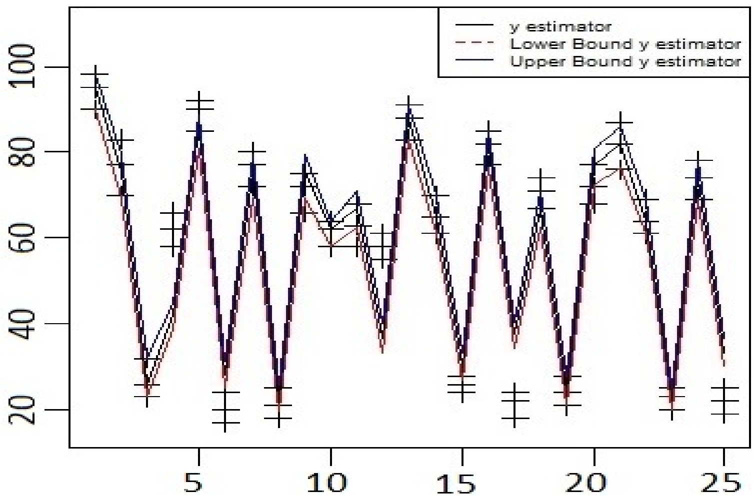

FIGURE 3. RNR method with the same weights in Example 4.1.

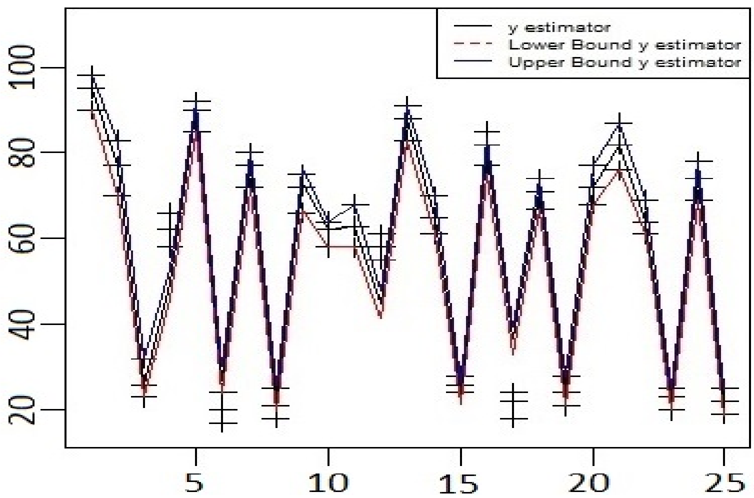

FIGURE 4. RNR method with the modified weights in Example 4.1.

We obtain the optimal value $K = 4$ *in KNN method and* $R = 0.82$ *in RNR method. The other authors have presented some fuzzy regression model on this fuzzy data set. We compare the proposed methods with these methods in Table 4. Based on the goodness of fit measures, the proposed methods (specially in*

KNN method with the modified weights) have the suitable performances. The performances of the proposed fuzzy nonparametric regression models are shown in Figures 5-8. Figures 5 and 7 show the KNN and RNR methods with the same weights and Figures 6 and 8 show the KNN and RNR methods with the modified weights (the + signs are the bounds and centers of observed response variables, and the drawn lines are the fuzzy nonparametric regression models).

TABLE 3. Data set in Example 4.2.

No.	$\tilde{x}_{i1}$	$\tilde{x}_{i2}$	$\tilde{y}_i$	No.	$\tilde{x}_{i1}$	$\tilde{x}_{i2}$	$\tilde{y}_i$
1	$(6.5, 7, 8.25)_T$	$(7.25, 8, 9)_T$	$(7.25, 8, 9)_T$	16	$(6.5, 7, 8.25)_T$	$(6.5, 7, 8.25)_T$	$(6.5, 7, 8.25)_T$
2	$(6.5, 7, 8.25)_T$	$(6.5, 7, 8.25)_T$	$(5.75, 6, 6.5)_T$	17	$(5.75, 6, 6.5)_T$	$(6.5, 7, 8.25)_T$	$(5.75, 6, 6.5)_T$
3	$(5.75, 6, 6.5)_T$	$(6.5, 7, 8.25)_T$	$(5.75, 6, 6.5)_T$	18	$(6.5, 7, 8.25)_T$	$(7.25, 8, 9)_T$	$(6.5, 7, 8.25)_T$
4	$(7.25, 8, 9)_T$	$(9, 9, 10)_T$	$(9, 9, 10)_T$	19	$(6.5, 7, 8.25)_T$	$(6.5, 7, 8.25)_T$	$(7.25, 8, 9)_T$
5	$(7.25, 8, 9)_T$	$(7.25, 8, 9)_T$	$(7.25, 8, 9)_T$	20	$(6.5, 7, 8.25)_T$	$(9, 9, 10)_T$	$(6.5, 7, 8.25)_T$
6	$(5.75, 6, 6.5)_T$	$(6.5, 7, 8.25)_T$	$(5, 5, 6)_T$	21	$(6.5, 7, 8.25)_T$	$(7.25, 8, 9)_T$	$(6.5, 7, 8.25)_T$
7	$(6.5, 7, 8.25)_T$	$(7.25, 8, 9)_T$	$(6.5, 7, 8.25)_T$	22	$(6.5, 7, 8.25)_T$	$(7.25, 8, 9)_T$	$(5.75, 6, 6.5)_T$
8	$(6.5, 7, 8.25)_T$	$(6.5, 7, 8.25)_T$	$(5, 5, 6)_T$	23	$(6.5, 7, 8.25)_T$	$(9, 9, 10)_T$	$(6.5, 7, 8.25)_T$
9	$(6.5, 7, 8.25)_T$	$(7.25, 8, 9)_T$	$(6.5, 7, 8.25)_T$	24	$(6.5, 7, 8.25)_T$	$(6.5, 7, 8.25)_T$	$(7.25, 8, 9)_T$
10	$(5.75, 6, 6.5)_T$	$(6.5, 7, 8.25)_T$	$(5.75, 6, 6.5)_T$	25	$(6.5, 7, 8.25)_T$	$(6.5, 7, 8.25)_T$	$(5.75, 6, 6.5)_T$
11	$(6.5, 7, 8.25)_T$	$(7.25, 8, 9)_T$	$(7.25, 8, 9)_T$	26	$(6.5, 7, 8.25)_T$	$(6.5, 7, 8.25)_T$	$(5.75, 6, 6.5)_T$
12	$(6.5, 7, 8.25)_T$	$(5.75, 6, 6.5)_T$	$(5.75, 6, 6.5)_T$	27	$(6.5, 7, 8.25)_T$	$(6.5, 7, 8.25)_T$	$(6.5, 7, 8.25)_T$
13	$(6.5, 7, 8.25)_T$	$(7.25, 8, 9)_T$	$(9, 9, 10)_T$	28	$(6.5, 7, 8.25)_T$	$(7.25, 8, 9)_T$	$(6.5, 7, 8.25)_T$
14	$(6.5, 7, 8.25)_T$	$(7.25, 8, 9)_T$	$(7.25, 8, 9)_T$	29	$(6.5, 7, 8.25)_T$	$(6.5, 7, 8.25)_T$	$(6.5, 7, 8.25)_T$
15	$(6.5, 7, 8.25)_T$	$(6.5, 7, 8.25)_T$	$(6.5, 7, 8.25)_T$	30	$(5.75, 6, 6.5)_T$	$(6.5, 7, 8.25)_T$	$(5.75, 6, 6.5)_T$

TABLE 4. Goodness of fit measures in different fuzzy regression models in Example 4.2.

Model	MSE	MSM
KNN with the same weights	1.32	0.31
KNN with the modified weights	0.37	0.55
RNR with the same weights	1.55	0.40
RNR with the modified weights	1.17	0.43
Lu and Wang [31]	2.15	0.56
D'Urso [20]	2.45	0.22
Chachi and Taheri [15]	2.01	0.38

5. Conclusion

In this paper, we present four methods to fit some nonparametric regression models based on the extended versions of KNN and RNR on fuzzy data. These methods have certain merits as follows

1) In the proposed nonparametic regression models, we are assume that both the explanatory and response variables are fuzzy quantities.

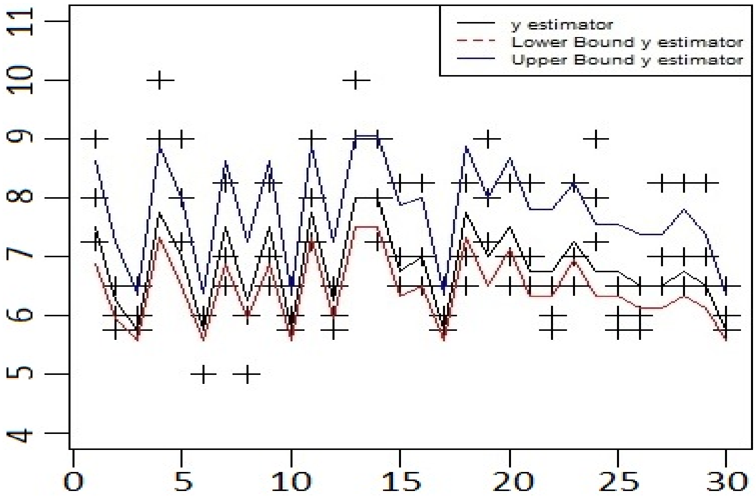

FIGURE 5. KNN method with the same weights in Example 4.2.

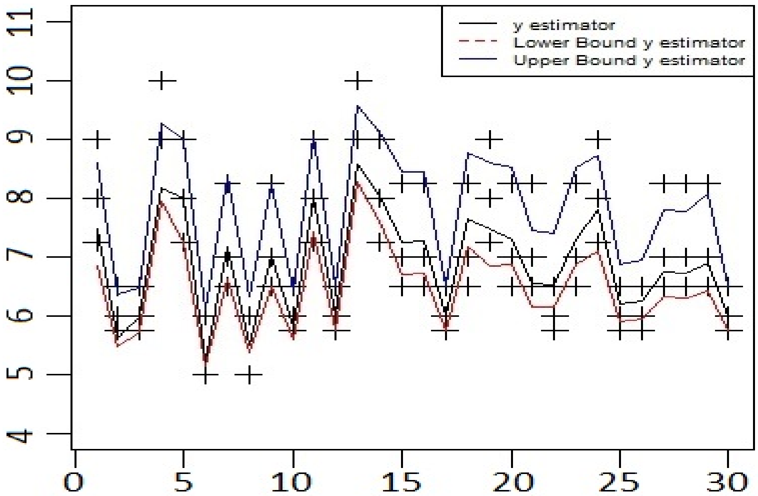

FIGURE 6. KNN method with the modified weights in Example 4.2.

2) The KKN method with the same weigths is an extended version of ordinal KNN method on fuzzy data.

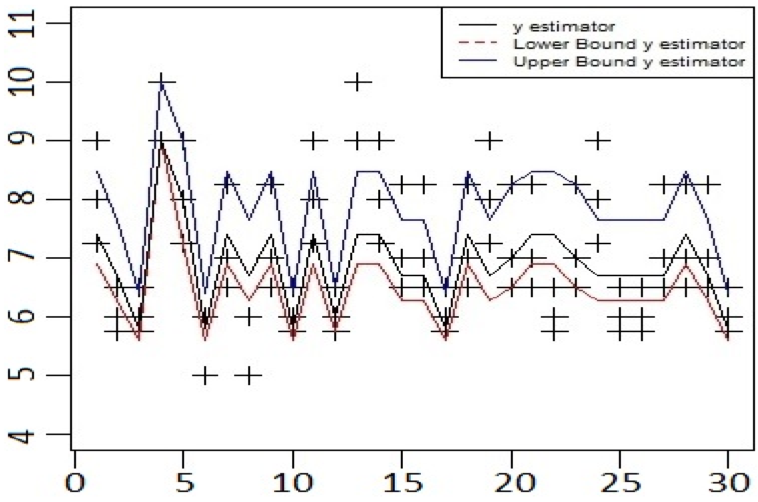

Figure 7. RNR method with the same weights in Example 4.2.

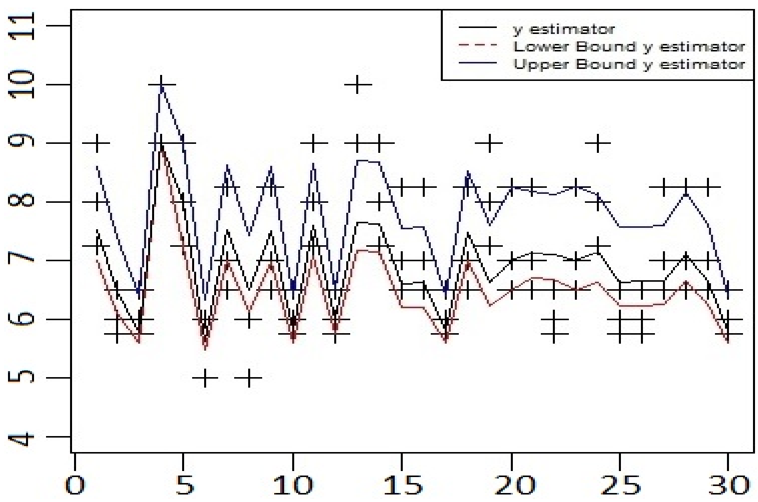

Figure 8. RNR method with the modified weights in Example 4.2.

3) The RNR method is a new method based on based on the data enclosed in radius R. Therefore, the RNR method can be performed with more speed and accuracy (especially with the presence of outliers).

4) The KNN and RNR with the modified weigths have the better performances than the methods with the same weights, because they assign more weights to closer data and less weights to farther data.

An important method for fitting a nonparametric regression model is to use the kerenl functions. We can extend the KNN and RNR methods based on some kernel functions. Based on these methods, we can obtain more suitable and flexible performances. This topic can be invistigated in future researches.

6. Aknowledgement

The authors would like to thank the reviewers for their valuable comments and suggestions in improving our manuscript.

References

[1] Akbari, M. G. & Hesamian, G. (2018). Linear model with exact inputs and interval-valued fuzzy outputs. IEEE Transactions on Fuzzy Systems, 26, 518-530. https://doi.org/10.1109/TFUZZ.2017.2686356

[2] Akbari, M. G. & Hesamian, G. (2019). A partial-robust-ridge-based regression model with fuzzy predictors-responses. Journal of Computational and Applied Mathematics, 351, 290-301. https://doi.org/10.1016/j.cam.2018.11.006

[3] Arabpour, A. & Amini, M. (2014). Weighted Linear Regression Model For Imprecise Response. Journal of Mahani Mathematical Research, 3, 1-17. DOI: 10.22103/jmmrc.2014.1399

[4] Arefi, M. (2016). Clustering regression based on interval-valued fuzzy outputs and interval-valued fuzzy parameters. Journal of Intelligent and Fuzzy Systems, 30, 1339-1351. https://doi.org/10.3233/IFS-152048

[5] Arefi, M. (2020). Quantile fuzzy regression based on fuzzy outputs and fuzzy parameters. Soft Computing, 24, 311-320. https://doi.org/10.1007/s00500-019-04424-2

[6] Arefi, M. (2025). Geometric clustering fuzzy regression based on c-means clustering. Iranian Journal of Fuzzy Systems, 22(3), 87-101. DOI: 10.22111/ijfs.2025.51386.9078

[7] Arefi, M. & Khammar, A. H. (2024). Nonlinear prediction of fuzzy regression model based on quantile loss function. Soft Computing, 28(6), 4861-4871. https://doi.org/10.1007/s00500-023-09190-w

[8] Arefi, M. & Taheri, S. M. (2015). Least squares regression based on Atanassov's intuitionistic fuzzy inputs-outputs and Atanassov's intuitionistic fuzzy parameters. IEEE Transactions on Fuzzy Systems, 23, 1142-1154. https://doi.org/10.1109/TFUZZ.2014.2346246

[9] Asadolahi, M., Akbari, M. Gh., Hesamian, G., & Arefi, M. (2021). A robust support vector regression with exact predictors and fuzzy responses. International Journal of Approximate Reasoning, 132, 206-225. https://doi.org/10.1016/j.ijar.2021.02.006

[10] Bas, E. & Egrioglu, E. (2022). A fuzzy regression functions approach based on Gustafson-Kessel clustering algorithm. Information Sciences, 592, 206-214. https://doi.org/10.1016/j.ins.2022.01.057

[11] Chachi, J. (2019). A weighted least-squares fuzzy regression for crisp input-fuzzy output data. IEEE Transactions on Fuzzy Systems, 27, 739-748. https://doi.org/10.1109/TFUZZ.2018.2868554

[12] Chachi, J. & Jalalvand, M. (2025). Fuzzy Modeling Using The Similarity-based Approximate Reasoning System. Journal of Mahani Mathematical Research, 14, 165-187. DOI: 10.22103/jmmr.2024.22970.1587

[13] Chachi, J., Kazemifard, A., & Jalalvand, M. (2021). A multi-attribute assessment of fuzzy regression models. Iranian Journal of Fuzzy Systems, 18, 131-148. https://doi.org/10.22111/ijfs.2021.6181

[14] Chachi, J. & Roozbeh, M. (2017). A fuzzy robust regression approach applied to bedload transport data. Communications in Statistics - Simulation and Computation, 46, 1703-1714. https://doi.org/10.1080/03610918.2015.1010002

[15] Chachi, J. & Taheri, S. M. (2016). Multiple fuzzy regression model for fuzzy input-output data. Iranian Journal of Fuzzy Systems, 13, 63-78. https://doi.org/10.22111/ijfs.2016.2596

[16] Chachi, J., Taheri, S. M., Fattahi, S., & Ravandi, A. H. (2016). Two robust fuzzy regression models and their applications in predicting imperfections of cotton yarn. Journal of Textiles and Polymers, 4(2), 60-68. https://www.academia.edu/download/115503970/article_ 45940_ acece29211ec3d99f9913182697ba712.pdf

[17] Cheng, C. B. & Lee, E. S. (1999). Nonparametric fuzzy regression k-NN and kernel smoothing techniques. Computers and Mathematics with Applications, 38, 239-251. https://doi.org/10.1016/S0898-1221(99)00198-4

[18] Danesh, M., Danesh, S., Razzaghnia, T., & Maleki, A. (2021). Prediction of Fuzzy Nonparametric Regression Function: A Comparative Study of a New Hybrid Method and Smoothing Methods. Global Analysis and Discrete Mathematics, 6(1), 143–178. DOI: 10.22128/gadm.2021.387.1038

[19] Dotto, F., Farcomeni, A., Garcia-Escudero, L. A., & Mayo-Iscar, A. (2017). A fuzzy approach to robust regression clustering. Advances in Data Analysis and Classiffication, 11, 691-710. https://doi.org/10.1007/s11634-016-0271-9

[20] D'Urso, P. (2003). Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data. Computational Statistics and Data Analysis, 42, 47-72. https://doi.org/10.1016/S0167-9473(02)00117-2

[21] D'Urso, P. & Gastaldi, T. (2000). A least-squares approach to fuzzy linear regression analysis. Computational Statistics and Data Analysis, 34, 427-440. https://doi.org/10.1016/S0167-9473(99)00109-7

[22] Hesamian, G. & Akbari, M. Gh. (2021). A robust multiple regression model based on fuzzy random variables. Journal of Computational and Applied Mathematics, 338, Article No.: 113270. DOI:10.1016/j.cam.2020.113270

[23] Hesamian, G. & Akbari, M. Gh. (2020). A fuzzy additive regression model with exact predictors and fuzzy responses. Applied Soft Computing, 95, Article No.: 106507. https://doi.org/10.1016/j.asoc.2020.106507

[24] Hesamian, G., Torkian, F., Johannssen, A., & Chukhrova, N. (2024). A fuzzy nonparametric regression model based on an extended center and range method. Journal of Computational and Applied Mathematics, 436, Article No.: 115377. https://doi.org/10.1016/j.cam.2023.115377

[25] Hosseinzadeh, E., Hassanpour, H., & Arefi, M. (2015). A weighted goal programming approach to fuzzy linear regression with crisp inputs and type-2 fuzzy outputs. Soft Computing, 19, 1143-1151. https://doi.org/10.1007/s00500-014-1328-3

[26] Kelkinnama, M. & Taheri, S. M. (2012). Fuzzy least-absolutes regression using shape preserving operations. Information Sciences, 214, 105-120. https://doi.org/10.1016/j.ins.2012.04.017

[27] Khammar, A. H., Arefi, M., & Akbari, M.Gh. (2020). A robust least squares fuzzy regression model based on kernel function. Iranian Journal of Fuzzy Systems, 17(4), 105-119. https://doi.org/10.22111/ijfs.2020.5409

[28] Khammar, A. H., Arefi, M., & Akbari, M. Gh. (2021). A general approach to fuzzy regression models based on different loss functions. Soft Computing, 25, 835-849. https://doi.org/10.1007/s00500-020-05441-2

[29] Khammar, A. H., Arefi, M., & Akbari, M.Gh. (2021). Quantile fuzzy varying coefficient regression based on kernel function. Applied Soft Computing, 107, Article No.: 107313. https://doi.org/10.1016/j.asoc.2021.107313

[30] Kong, L., Song, X., & Wang, X. (2022). Nonparametric regression for interval-valued data based on local linear smoothing approach. Neurocomputing, 501, 834-843. https://doi.org/10.1016/j.neucom.2022.06.073

[31] Lu, J. & Wang, R. (2009). An enhanced fuzzy linear regression model with more flexible spreads. Fuzzy Sets and Systems, 160, 2505-2523. https://doi.org/10.1016/j.fss.2009.02.023

[32] Pappis, C. P. & Karacapilidis, N. I. (1993). A comparative assessment of measure of similarity of fuzzy values. Fuzzy Sets and Systems, 56, 171-174. https://doi.org/10.1016/0165-0114(93)90141-4

[33] Zadeh, L. A. (1965). Fuzzy sets. Information and Control, 8(3), 338-353. DOI: 10.2307/2272014

[34] Zeng, W., Feng, Q., & Lia, J. (2016). Fuzzy least absolute linear regression. Applied Soft Computing, 52, 1009-1019. https://doi.org/10.1016/j.asoc.2016.09.029

[35] Zimmermann, H. J. (2001). Fuzzy Set Theory and Its Applications, fourth ed., Kluwer Nihoff, Boston. DOI: 10.1007/978-94-010-0646-0

Ebrahim Ebrahimi
Orcid number: 0009-0007-3182-6090
Department of Statistics
Faculty of Mathematical Sciences and Statistics
University of Birjand
Birjand, Iran.
Email address: ebrahim-ebrahimi@birjand.ac.ir

Mohsen Arefi
Orcid number: 0000-0003-3187-9647
Department of Statistics
Faculty of Mathematical Sciences and Statistics
University of Birjand
Birjand, Iran.
Email address: arefi@birjand.ac.ir

Mohhamad Ghasem Akbari
Orcid number: 0000-0001-5064-8065
Department of Statistics
Faculty of Mathematical Sciences and Statistics
University of Birjand
Birjand, Iran.
Email address: g_z_akbari@birjand.ac.ir