Iranian Journal of Fuzzy Systems **Vol. 7, No. 2,** (2010) pp. 19-39

FUZZY LINEAR REGRESSION MODEL WITH CRISP COEFFICIENTS: A GOAL PROGRAMMING APPROACH

H. HASSANPOUR, H. R. MALEKI AND M. A. YAGHOOBI

ABSTRACT. The fuzzy linear regression model with fuzzy input-output data and crisp coefficients is studied in this paper. A linear programming model based on goal programming is proposed to calculate the regression coefficients. In contrast with most of the previous works, the proposed model takes into account the centers of fuzzy data as an important feature as well as their spreads in the procedure of constructing the regression model. Furthermore, the model can deal with both symmetric and non-symmetric triangular fuzzy data as well as trapezoidal fuzzy data which have rarely been considered in the previous works. To show the efficiency of the proposed model, some numerical examples are solved and a simulation study is performed. The computational results are compared with some earlier methods.

1. Introduction

Since Zadeh [52] introduced fuzzy set theory, it has been widely developed in theory and application (e.g. see [1, 2, 5, 34, 39, 43, 51]). Regression analysis is one of the areas in which fuzzy set theory has been used frequently. Since Tanaka et al. [42] initiated research on fuzzy linear regression (FLR) analysis, this area has been widely developed and a wide variety of methods have been proposed. Putting the exact phrase "fuzzy linear regression" on googlescholar search engine yields about 950 references. One approach to deal with FLR is linear programming (LP). This approach was first introduced by Tanaka et al. [42] and developed by others (e.g. see [16, 17, 31, 32, 36, 37, 38, 40, 41]). Another approach is least-squares method, which was first introduced by Celmins [6] and developed by others (see e.g. [9, 10, 13, 15, 20, 21, 27, 28, 29, 30, 47, 48, 50]). Some authors discussed features, advantages and shortcomings of different methods. To overcome the shortcomings, some new methods have been proposed (see e.g. [7, 9, 16, 17, 18, 21, 22, 33]). A relatively comprehensive literature review on FLR can be seen e.g. in [9, 16, 22, 44, 45, 46, 49].

FLR models can be classified into two general categories according to the type of dependent and independent variables:

(a) Input data are non-fuzzy and output data are fuzzy numbers.
(b) Both input and output data are fuzzy numbers.

Received: February 2008; Revised: August 2008 and April 2009; Accepted: September 2009
Key words and phrases: Fuzzy linear regression, Goal programming, Linear programming, Fuzzy number.

The FLR model in case (a) can be expressed as follows:

$$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 x_1 + \tilde{A}_2 x_2 + ... + \tilde{A}_p x_p \tag{1}$$

(The symbol " ˜ " over a letter indicates a fuzzy number). The FLR model in case (b) can be expressed as follows:

(b-i) FLR model with crisp coefficients:

$$\tilde{Y} = a_0 + a_1 \tilde{x}_1 + a_2 \tilde{x}_2 + \cdots + a_p \tilde{x}_p. \tag{2}$$

(b-ii) FLR model with fuzzy coefficients:

$$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 \tilde{x}_1 + \tilde{A}_2 \tilde{x}_2 + \cdots + \tilde{A}_p \tilde{x}_p. \tag{3}$$

The authors applied goal programming approach to construct the FLR model (1) in [14]. The present paper focuses on case (b-i) which has been presented in [11]. Further researches on case (b-i) can be found e.g. in [3, 7, 19, 20].

Diamond in [11] defined a metric on triangular fuzzy numbers and formulated a least-squares model to estimate the regression coefficients $a_0, a_1, \cdots, a_p$ of model (2).

Arabpour and Tata [3] considered the metric of Diamond [11] on triangular fuzzy numbers and extended it to trapezoidal fuzzy numbers. They calculated the coefficients of model (2) by using the normal equations corresponding to a least-squares model.

Kao and Chyu in [19] considered the FLR model (2) with an additional triangular fuzzy error term $E = (0, l, r)$ in the right-hand side. They proposed a two-stage approach to construct the FLR model. In the first stage, the fuzzy observations are defuzzified, using the centroid method, so that the traditional least-squares method can be applied to find a crisp regression line showing the general trend of the data. In the second stage, the error term of the fuzzy regression model, representing the fuzziness of the data in a general sense, is determined to give the regression model the best explanatory power for the data. To this end, the total non-overlapped area between the graphs of membership functions of the observed and estimated triangular fuzzy responses is minimized under some constraints. In [19], an algebraic formula has been presented to calculate the mentioned areas. However, it cannot be used in a general situation. Because, the relative status of the observed and estimated responses are not known. Furthermore, their method does not guarantee the non-negativity of the spreads l and r.

Also, Kao and Chyu in [20] proposed an idea, stemmed from the classical least squares, to handle fuzzy observations in regression analysis. They minimized fuzzy sum of squared errors via a nonlinear programming model.

Choi and Buckley [7] also considered the FLR model similar to Kao and Chyu's model [19] and proposed a fuzzy least absolute deviations method to calculate the regression coefficients. They first calculate the crisp coefficients $a_0, a_1, \cdots, a_p$ by defuzzifying the fuzzy data, using the centroid method, and minimizing a non-linear function. Then, to calculate the error term $(0, l, r)$, they solve two unconstrained non-linear programming problems. Their method does not guarantee the non-negativity of the spreads l and r, like Kao and Chyu's method.

In literature, various distances (metrics) between fuzzy numbers have been proposed which are appropriate for using the least-squares method (e.g. see [3, 11, 12, 13, 48]). However, as mentioned by Choi and Buckley [7], the least-squares method is so sensitive to the outliers that it could be greatly affected by a small number of outliers. Since outliers in the response variable represent model failure, there has been an increased interest in robust estimation procedures, which are insensitive to some outliers, applied to the regression analysis. Like to an ordinary regression, we need robust methods in order to estimate the fuzzy regression coefficients. The least absolute deviation estimators were proposed, based on the medians, as an alternative to least squares method. Under some conditions, the least absolute deviation estimators were more efficient than the least squares method in ordinary regression models [7, 8, 23]. Moreover, Choi and Buckley [7] pointed out that in the fuzzy regression model using the least squares method there is a tendency that the larger the values of independent variables, the wider the spreads of the estimated dependent variables. Thus, they suggested the least absolute deviation estimators using the L_1-norm instead of the least squares method in estimation of the regression coefficients. So they used the absolute deviations between the centers and end-points of h-level sets of two triangular fuzzy numbers in their models.

The authors formulated a simple goal programming (GP) model to calculate the fuzzy coefficients of model (1) in [14], when the input data are crisp and the output data are triangular fuzzy numbers. Their GP model minimizes the total absolute deviations between the middle points of observed and estimated responses, and absolute deviations between their spreads. Then they showed that the proposed model has better performance than some similar models based on both the Kim and Bishu's [21] criterion of goodness and closeness of the centers of observed and estimated responses (note that closeness of the centers of observed and estimated responses has a special importance because they have the highest membership function value). Also, they showed that the proposed model is less sensitive to outliers than some similar models in the literature. In this paper, the mentioned absolute deviations are used in both triangular and trapezoidal cases. Then, a simple GP model is formulated to calculate the coefficients of model (2), when the input data $\tilde{x}_{ij},\ i = 1, 2, \cdots, n,\ j = 1, 2, \cdots, p$ and the observed responses $\tilde{y}_i,\ i = 1, 2, \cdots, n$ are triangular (trapezoidal) fuzzy numbers (Note that n is the number of observations and p is the number of independent variables).

A triangular fuzzy number (TFN), say $\tilde{C}$, is denoted by $\tilde{C} = (c, \alpha, \beta)$ where c, α, and β are the center (middle point), left spread, and right spread of $\tilde{C}$, respectively. c is a real number and $\alpha, \beta > 0$. The membership function of $\tilde{C}$ is as follows:

$$\tilde{C}(x) = \begin{cases} \frac{x-(c-\alpha)}{\alpha} & c-\alpha \leq x \leq c, \\ \frac{c+\beta-x}{\beta} & c \leq x \leq c+\beta, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

The intervals $[c-(1-h)\alpha \ , \ c+(1-h)\beta]$ and $[c-\alpha \ , \ c+\beta]$ are called h-level set ($h \in (0,1]$) and support of $\tilde{C}$, respectively. If $\alpha = \beta$, then $\tilde{C}$ is a symmetric TFN and is denoted by $\tilde{C} = (c, \alpha)$. A real number can be considered as a degenerated TFN whose spreads are zero, and vice versa. The set of all TFNs is denoted by $T(\mathbb{R})$. The following formulas for addition of two TFNs and multiplication of a TFN by a scaler are drawn from the extension principle of Zadeh [53].

If $\tilde{C}_1 = (c_1, \alpha_1, \beta_1)$ and $\tilde{C}_2 = (c_2, \alpha_2, \beta_2)$ be in $T(\mathbb{R})$ and $r \in \mathbb{R}$, then:

$$\tilde{C}_1 + \tilde{C}_2 = (c_1 + c_2, \alpha_1 + \alpha_2, \beta_1 + \beta_2), \tag{5}$$

$$r\tilde{C}_1 = \begin{cases} (rc_1, r\alpha_1, r\beta_1) & r \geq 0, \\ (rc_1, -r\beta_1, -r\alpha_1) & r < 0. \end{cases} \tag{6}$$

A trapezoidal fuzzy number (TPFN), say $\tilde{C}$, is denoted by $\tilde{C} = (c_L, c_U, \alpha, \beta)$ where c_L and c_U are lower and upper middle points and α, β are left and right spreads of $\tilde{C}$, respectively. c_L and c_U are real numbers ($c_L \leq c_U$) and $\alpha, \beta > 0$. The membership function of $\tilde{C}$ is as follows:

$$\tilde{C}(x) = \begin{cases} \frac{x-(c_L-\alpha)}{\alpha} & c_L - \alpha \leq x \leq c_L \quad \alpha > 0, \\ 1 & c_L \leq x \leq c_U, \\ \frac{c_U+\beta-x}{\beta} & c_U \leq x \leq c_U + \beta \quad \beta > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

The intervals $[c_L-(1-h)\alpha \ , \ c_U+(1-h)\beta]$ and $[c_L-\alpha \ , \ c_U+\beta]$ are called h-level set ($h \in (0,1]$) and support of $\tilde{C}$, respectively. If $\alpha = \beta$, then $\tilde{C}$ is a symmetric TPFN and is denoted by $\tilde{C} = (c_L, c_U, \alpha)$. Also, $\tilde{C}$ is a TFN if $c_L = c_U$. A real number can be considered as a degenerated TPFN whose spreads are zero, and its middle points are equal, and vice versa. The set of all TPFNs is denoted by $TP(\mathbb{R})$. The following formulas for addition of two TPFNs and multiplication of a TPFN by a scaler are drawn from the extension principle of Zadeh [53].

If $\tilde{C}_1 = (c_{1L}, c_{1U}, \alpha_1, \beta_1)$ and $\tilde{C}_2 = (c_{2L}, c_{2U}, \alpha_2, \beta_2)$ be in $TP(\mathbb{R})$ and $r \in \mathbb{R}$, then:

$$\tilde{C}_1 + \tilde{C}_2 = (c_{1L} + c_{2L}, c_{1U} + c_{2U}, \alpha_1 + \alpha_2, \beta_1 + \beta_2), \tag{8}$$

$$r\tilde{C}_1 = \begin{cases} (rc_{1L}, rc_{1U}, r\alpha_1, r\beta_1) & r \geq 0, \\ (rc_{1U}, rc_{1L}, -r\beta_1, -r\alpha_1) & r < 0. \end{cases} \tag{9}$$

The purpose of this paper is to estimate the FLR parameters $a_0, a_1, \cdots, a_p$ for both triangular and trapezoidal fuzzy input-output data so that the estimated responses $\tilde{Y}_i$,

$$\tilde{Y}_i = a_0 + a_1\tilde{x}_{i1} + a_2\tilde{x}_{i2} + \cdots + a_p\tilde{x}_{ip}, \quad i = 1, \cdots, n, \tag{10}$$

have the best fitness to the given data according to a criterion of goodness. Kim and Bishu [21] introduced a difference between two membership functions. We adopt it as a criterion of goodness. Also, a new parameter (named λ) which measures the total distance between the centers of observed and estimated responses has been introduced in [14]. We use it in this paper.

The paper is organized as follows: A GP approach is introduced in Section 2, the numerical examples and comparisons are presented in Section 3, and Section 4 is devoted to discussion and conclusion.

2. A Goal Programming Approach (GPA) for Estimating the Regression Coefficients

The purpose of classical regression is to fit a function to a set of real (crisp) data, so that the estimated responses from the regression model be close to the corresponding observed responses as much as possible. To this end, classical regression methods minimize a function of differences between the observed and estimated responses. Similarly, in fuzzy regression, we attempt to fit a fuzzy regression model to the given fuzzy data, so that the fuzzy estimated responses be close to the corresponding fuzzy observed responses, as much as possible. Therefore, we try to close the membership functions of observed and estimated responses as much as possible. Clearly, it is important that a point with the highest membership value in a fuzzy estimated response be close to the point with the highest membership value in the corresponding observed response, and the points with other membership values be, too. Since the observed data in our study are assumed to be triangular or trapezoidal fuzzy numbers, we try to close the membership functions of observed and estimated responses by closing their middle points as well as their spreads. To do this, we use the following definition of distance between two TFNs, and extend it to trapezoidal fuzzy numbers. Then we try to minimize the total distance between the observed and estimated responses.

Definition 2.1. Let $\tilde{C}_1 = (c_1, \alpha_1, \beta_1)$ and $\tilde{C}_2 = (c_2, \alpha_2, \beta_2)$ be in $T(\mathbb{R})$. Define $d : T(\mathbb{R}) \times T(\mathbb{R}) \to \mathbb{R}$ as follows:

$$d(\tilde{C}_1, \tilde{C}_2) = |c_1 - c_2| + |\alpha_1 - \alpha_2| + |\beta_1 - \beta_2|.$$

Proposition 2.2. *The function d defined in Definition 2.1 is a metric on $T(\mathbb{R}) \times T(\mathbb{R})$, i.e. for each $\tilde{A}, \tilde{B}, \tilde{C} \in T(\mathbb{R})$:*

1. $d(\tilde{A}, \tilde{B}) \geq 0$ *and* $d(\tilde{A}, \tilde{A}) = 0$.
2. $d(\tilde{A}, \tilde{B}) = d(\tilde{B}, \tilde{A})$.
3. $d(\tilde{A}, \tilde{C}) \leq d(\tilde{A}, \tilde{B}) + d(\tilde{B}, \tilde{C})$.

Proof. Straightforward. □

Remark 2.3. For two crisp numbers, the above distance is reduced to the absolute difference between them.

Remark 2.4. Some metrics on fuzzy numbers (such as Xu and Li's distance [48] and Yang and Ko's distance [49]) use the membership functions of fuzzy numbers explicitly. Although the membership functions did not appear in the above definition explicitly, it uses the membership functions implicitly. Because the TFNs are completely characterized by their centers and spreads. Therefore, closing the centers and the spreads of two TFNs is enough (in fact necessary and sufficient) to close their membership functions, which is the purpose of this paper.

Remark 2.5. Some other metrics on fuzzy numbers (such as Diamond and Kloeden's distance [12] and D'Urso's distance [13]) incorporate the middle points of triangular fuzzy numbers in the metric as well as our distance. Their metrics use the L_2-norm. So they use the least-squares method to calculate the FLR coefficients. As mentioned above, the solutions of least-squares method are so sensitive to the outliers that they could be greatly affected by a small number of outliers [7]. So using L_1-norm has been suggested in [7] instead of L_2-norm. Further, as seen in [14], using L_1-norm to a set of crisp input-fuzzy output data reduces the effect of outliers. Furthermore, as it will be seen later, by using L_1-norm we convert a non-linear programming problem to a linear one. The advantage of this conversion is that linear programming problems can be solved exactly by the Simplex method whereas most of algorithms for solving non-linear programming problems yield approximate solution.

Suppose the given inputs are non-symmetric TFNs $\tilde{x}_{ij} = (x_{ij}, l_{ij}, r_{ij})$ for $i = 1, \cdots, n$ and $j = 1, \cdots, p$, and the observed responses are non-symmetric TFNs $\tilde{y}_i = (y_i, l_i, r_i)$ for $i = 1, \cdots, n$. To calculate the coefficients of model (2), we need to multiply the TFN $\tilde{x}_j$ by the scaler a_j. This multiplication depends on the sign of the scaler, as seen in (9). Therefore, for different states of the signs of regression coefficients, different models (different LP models or different least-squares models,...) must be formulated and solved to estimate the coefficients. To avoid this, we propose a goal programming model to estimate the regression coefficients which is independent of the sign of them. Therefore, only one model is enough to be formulated and solved irrespective of the sign of coefficients.

Definition 2.6. [35] Let $\tilde{A}$ and $\tilde{B}$ be two fuzzy sets. $\tilde{A}$ is said to be a subset of $\tilde{B}$, denoted by $\tilde{A} \subseteq \tilde{B}$, if and only if its membership function is less than or equal to that of $\tilde{B}$ everywhere on X:

$$\tilde{A} \subseteq \tilde{B} \Leftrightarrow \tilde{A}(x) \leq \tilde{B}(x) \quad \forall\ x \in X. \tag{11}$$

Proposition 2.7. *Suppose r and s are real numbers and $\tilde{A}$ is a fuzzy number. Then*

$$(r+s)\tilde{A} \subseteq r\tilde{A} + s\tilde{A}. \tag{12}$$

Proof. We have to show that:

$$((r+s)\tilde{A})(x) \leq (r\tilde{A} + s\tilde{A})(x) \quad \forall x \in \mathbb{R}. \tag{13}$$

If o is a binary operation on real numbers and $\tilde{A}$ and $\tilde{B}$ are two fuzzy numbers with membership functions $\tilde{A}(x)$ and $\tilde{B}(x)$, respectively, then by the extension principle of Zadeh [53] we have:

$$(\tilde{A}o\tilde{B})(x) = \sup_{yoz=x} \{inf\{\tilde{A}(y), \tilde{B}(z)\}\}. \tag{14}$$

Let us denote the membership function of the real numbers r and s by their characteristic functions χ_r and χ_s, respectively. By using (14) we have:

$$\begin{aligned}
((r+s)\tilde{A})(x) &= \sup_{uv=x} \{\inf\{(\chi_r + \chi_s)(u), \tilde{A}(v)\}\} \\
&= \sup_{uv=x} \{\inf\{ \sup_{y+z=u} \{\inf\{\chi_r(y), \chi_s(z)\}\}, \tilde{A}(v)\}\} \\
&= \sup_{uv=x,\ r+s=u} \{\inf\{\chi_r(r), \chi_s(s), \tilde{A}(v)\}\} \\
&= \sup_{(r+s)v=x} \{\tilde{A}(v)\} \\
&= \sup_{(r+s)v=x} \{\inf\{\tilde{A}(v), \tilde{A}(v)\}\}.
\end{aligned} \tag{15}$$

On the other hand:

$$\begin{aligned}
(r\tilde{A} + s\tilde{A})(x) &= \sup_{y+z=x} \{\inf\{(\chi_r\tilde{A})(y), (\chi_s\tilde{A})(z)\}\} \\
&= \sup_{y+z=x} \{\inf\{\sup_{tu=y} \{\inf\{\chi_r(t), \tilde{A}(u)\}\}, \\
&\qquad \sup_{wv=z} \{\inf\{\chi_s(w), \tilde{A}(v)\}\}\}\} \\
&= \sup_{y+z=x,\ ru=y,\ sv=z} \{\inf\{\chi_r(r), \tilde{A}(u), \chi_s(s), \tilde{A}(v)\}\} \\
&= \sup_{ru+sv=x} \{\inf\{\tilde{A}(u), \tilde{A}(v)\}\}.
\end{aligned} \tag{16}$$

The Inequality (13) is obtained from (15) and (16) because

$$\{(v,v)|v \in \mathbb{R}\} \subseteq \{(u,v)|u,v \in \mathbb{R}\}.$$

□

A real number a has infinite representations in the form of $a = a' - a''$, where a' and a'' are nonnegative real numbers. This matter leads to the following lemma.

Lemma 2.8. *Suppose $a_j \in \mathbb{R}$, $j = 0, \cdots, p$ and $\tilde{x}_{ij} = (x_{ij}, l_{ij}, r_{ij}) \in T(\mathbb{R})$, $i = 1, \cdots, n$, $j = 1, \cdots, p$. Set $a_j = a'_j - a''_j$ for $j = 1, \cdots, p$ where $a'_j, a''_j \geq 0$. Then for each choice of a'_j and a''_j we have:*

$$\tilde{Y}_i \subseteq (a_0 + \sum_{j=1}^{p}(a'_j - a''_j)x_{ij}, \sum_{j=1}^{p}(a'_j l_{ij} + a''_j r_{ij}), \sum_{j=1}^{p}(a'_j r_{ij} + a''_j l_{ij})), \tag{17}$$

where $\tilde{Y}_i$ is given by (10).

Proof. By using Proposition 2.7 and Equations (8) and (9) we have:

$$\begin{aligned}
\tilde{Y}_i &= a_0 + \sum_{j=1}^{p} a_j \tilde{x}_{ij} \\
&= a_0 + \sum_{j=1}^{p} (a'_j - a''_j)\tilde{x}_{ij} \\
&\subseteq a_0 + \sum_{j=1}^{p} (a'_j \tilde{x}_{ij} - a''_j \tilde{x}_{ij}) \\
&= a_0 + \sum_{j=1}^{p} (a'_j (x_{ij}, l_{ij}, r_{ij}) - a''_j (x_{ij}, l_{ij}, r_{ij})) \\
&= a_0 + \sum_{j=1}^{p} (a'_j x_{ij}, a'_j l_{ij}, a'_j r_{ij}) + (-a''_j x_{ij}, a''_j r_{ij}, a''_j l_{ij})) \\
&= (a_0 + \sum_{j=1}^{p} (a'_j - a''_j) x_{ij}, \sum_{j=1}^{p} (a'_j l_{ij} + a''_j r_{ij}), \sum_{j=1}^{p} (a'_j r_{ij} + a''_j l_{ij})),
\end{aligned}$$

and the proof is completed. □

Let us denote the right hand side of relation (17) by $\tilde{Y}_i(\mathbf{a}', \mathbf{a}'')$ where $\mathbf{a}'$ and $\mathbf{a}''$ are nonnegative p-dimensional vectors with jth element a'_j and a''_j, respectively. Clearly, for each choice of $\mathbf{a}'$ and $\mathbf{a}''$, $\tilde{Y}_i(\mathbf{a}', \mathbf{a}'')$ is a TFN, and we consider it as an approximation for $\tilde{Y}_i$. Indeed, there are many approximations for $\tilde{Y}_i$, among which, we try to choose the best approximation. Figure 1 depicts $\tilde{y}_i, \tilde{Y}_i$, and $\tilde{Y}_i(\mathbf{a}', \mathbf{a}'')$ for a special choice of $\mathbf{a}'$ and $\mathbf{a}''$. We have to find $\tilde{Y}_i$ as close as possible to $\tilde{y}_i$. Instead, first we approximate $\tilde{Y}_i$ to its supersets $\tilde{Y}_i(\mathbf{a}', \mathbf{a}'')$ for different choices of $\mathbf{a}'$ and $\mathbf{a}''$, as seen in Lemma 2.8. Then, we try to find appropriate values of $\mathbf{a}'$ and $\mathbf{a}''$ so that $\tilde{Y}_i(\mathbf{a}', \mathbf{a}'')$ be close to $\tilde{y}_i$ as much as possible. To this end, we attempt to close the membership function of each approximated response $\tilde{Y}_i(\mathbf{a}', \mathbf{a}'')$ to that of corresponding observed response $\tilde{y}_i$ as much as possible. Therefore, we introduce the following mathematical programming model, which finds the best choices of $\mathbf{a}'$ and $\mathbf{a}''$ for the coefficients of model (2):

$$\begin{aligned}
\min \quad & \sum_{i=1}^{n} d(\tilde{Y}_i(\mathbf{a}', \mathbf{a}''), \tilde{y}_i) \\
s.t. \quad & a_0 \in \mathbb{R}, \quad a'_j, a''_j \geq 0 \quad j = 1, 2, \cdots, p,
\end{aligned} \tag{18}$$

where d was introduced in Definition 2.1.

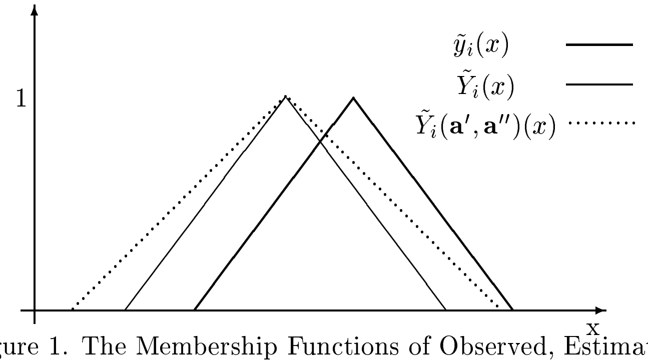

Figure 1. The Membership Functions of Observed, Estimated, and Approximated Responses

The model (18) can be converted to a GP model by choosing appropriate deviation variables. To this end, set $\tilde{Y}_i(\mathbf{a}', \mathbf{a}'') = (Y_i(\mathbf{a}', \mathbf{a}''), L_i(\mathbf{a}', \mathbf{a}''), R_i(\mathbf{a}', \mathbf{a}''))$ for $i = 1, \cdots, n$ and define

$$n_{ic} = \{|Y_i(\mathbf{a}', \mathbf{a}'') - y_i| - (Y_i(\mathbf{a}', \mathbf{a}'') - y_i)\}/2, \quad (19)$$

$$p_{ic} = \{|Y_i(\mathbf{a}', \mathbf{a}'') - y_i| + (Y_i(\mathbf{a}', \mathbf{a}'') - y_i)\}/2, \quad (20)$$

$$n_{il} = \{|L_i(\mathbf{a}', \mathbf{a}'') - l_i| - (L_i(\mathbf{a}', \mathbf{a}'') - l_i)\}/2, \quad (21)$$

$$p_{il} = \{|L_i(\mathbf{a}', \mathbf{a}'') - l_i| + (L_i(\mathbf{a}', \mathbf{a}'') - l_i)\}/2, \quad (22)$$

$$n_{ir} = \{|R_i(\mathbf{a}', \mathbf{a}'') - r_i| - (R_i(\mathbf{a}', \mathbf{a}'') - r_i)\}/2, \quad (23)$$

$$p_{ir} = \{|R_i(\mathbf{a}', \mathbf{a}'') - r_i| + (R_i(\mathbf{a}', \mathbf{a}'') - r_i)\}/2. \quad (24)$$

In fact, n_{ic} and p_{ic} are the negative and positive deviations between the centers of ith estimated and observed response, respectively. Also, n_{il} and p_{il} (n_{ir} and p_{ir}) are the negative and positive deviations between their left spreads (right spreads), respectively. It can be easily seen that

$$n_{ic} = \begin{cases} y_i - Y_i(\mathbf{a}', \mathbf{a}'') & \text{if} \quad Y_i(\mathbf{a}', \mathbf{a}'') \leq y_i, \\ 0 & \text{if} \quad Y_i(\mathbf{a}', \mathbf{a}'') > y_i, \end{cases} \quad (25)$$

$$p_{ic} = \begin{cases} Y_i(\mathbf{a}', \mathbf{a}'') - y_i & \text{if} \quad y_i \leq Y_i(\mathbf{a}', \mathbf{a}''), \\ 0 & \text{if} \quad y_i > Y_i(\mathbf{a}', \mathbf{a}''). \end{cases} \quad (26)$$

Similar relations are hold true for the other deviation variables. By using the above deviation variables, the model (18) converts to the following GP model:

(GP1):

$$\min \quad \sum_{i=1}^{n}(n_{ic}+p_{ic}+n_{il}+p_{il}+n_{ir}+p_{ir}) \tag{27}$$

$$s.t. \quad a_0+\sum_{j=1}^{p}(a'_j-a''_j)x_{ij}+n_{ic}-p_{ic}=y_i \qquad i=1,2,\cdots,n, \tag{28}$$

$$\sum_{j=1}^{p}(a'_j l_{ij}+a''_j r_{ij})+n_{il}-p_{il}=l_i \qquad i=1,2,\cdots,n, \tag{29}$$

$$\sum_{j=1}^{p}(a'_j r_{ij}+a''_j l_{ij})+n_{ir}-p_{ir}=r_i \qquad i=1,2,\cdots,n, \tag{30}$$

$$n_{ik}p_{ik}=0 \qquad i=1,2,\cdots,n, \quad k=l,c,r, \tag{31}$$

$$a_0\in\mathbb{R}, \quad a'_j,a''_j\geq 0 \qquad j=1,2,\cdots,p, \tag{32}$$

$$n_{ik},p_{ik}\geq 0 \qquad i=1,2,\cdots,n, \quad k=l,c,r. \tag{33}$$

Now we point out to some properties of GP1. The constraints (31) can be removed and the obtained LP model can be solved by the Simplex method [4]. Moreover, since for symmetric data $l_i = r_i$ and $l_{ij} = r_{ij}$ for each i and j, the constraints (29) and (30) are the same. So, one of them (in fact n constraints) can be removed. Note that if we use the Choi and Buckley's deviations (i.e., the absolute deviations between both the centers and the end-points of the h-level sets or supports of two TFNs), then the constraints (29) and (30) will not be the same for symmetric data, so the reduction of the constraints will not be possible. Another feature of GP1 is that in spite of some other methods such as [7] and [19], it does not produce negative spreads for TFNs. This fact has been shown in the following proposition.

Proposition 2.9. *The estimated responses from GP1 have non-negative spreads.*

Proof. The estimated responses from GP1 are:

$$\tilde{Y}_i(\mathbf{a}',\mathbf{a}'')=(Y_i(\mathbf{a}',\mathbf{a}''),L_i(\mathbf{a}',\mathbf{a}''),R_i(\mathbf{a}',\mathbf{a}'')),\ i=1,\cdots,n.$$

We have to show that $L_i(\mathbf{a}',\mathbf{a}''), R_i(\mathbf{a}',\mathbf{a}'')\geq 0$ for each i. According to the proof of Lemma 2.8, $L_i(\mathbf{a}',\mathbf{a}'')=\sum_{j=1}^{p}(a'_j l_{ij}+a''_j r_{ij})$ and $R_i(\mathbf{a}',\mathbf{a}'')=\sum_{j=1}^{p}(a'_j r_{ij}+a''_j l_{ij})$, which are both non-negative. Because a'_j, a''_j, l_{ij} and r_{ij} are all non-negative. □

Although the estimated responses $\tilde{Y}_i$, $i=1,\cdots,n$ have been approximated to $\tilde{Y}_i(\mathbf{a}',\mathbf{a}'')$, $i=1,\cdots,n$ (see Figure 1 and the relation (17)), the following theorem shows that GP1 yields the exact regression coefficients and estimated responses if the given fuzzy input-output data satisfy in a fuzzy linear model.

Theorem 2.10. *Suppose that the observed fuzzy input-output data satisfy in the fuzzy linear model:*

$$\tilde{y} = \bar{a}_0 + \bar{a}_1\tilde{x}_1 + \bar{a}_2\tilde{x}_2 + \cdots + \bar{a}_p\tilde{x}_p.$$

Then:
(i) The optimal solution of GP1 contains $\bar{a}_0, \cdots, \bar{a}_p$.
(ii) The estimated responses from GP1 and the observed responses are exactly the same (i.e., $\tilde{Y}_i(\boldsymbol{a}', \boldsymbol{a}'') = \tilde{y}_i$*).*

Proof. (i). Set $a_0 = \bar{a}_0$,

$$a'_j = \begin{cases} \bar{a}_j & \text{if} \quad \bar{a}_j \geq 0 \\ 0 & \text{if} \quad \bar{a}_j < 0 \end{cases} \quad \text{and} \quad a''_j = \begin{cases} 0 & \text{if} \quad \bar{a}_j \geq 0 \\ -\bar{a}_j & \text{if} \quad \bar{a}_j < 0 \end{cases}$$

for $j = 1, \cdots, p$, and $n_{ik} = p_{ik} = 0$, for $i = 1, 2, \cdots, n$ and $k = l, c, r$. First, we show that these values yield a feasible solution for GP1. Clearly, they satisfy in the constraints (31)-(33). Without loss of generality, for simplicity, suppose that $\bar{a}_j \geq 0,\ j = 1, \cdots, p$. By the assumption we have:

$$a_0 + \sum_{j=1}^{p}(a'_j - a''_j)x_{ij} + n_{ic} - p_{ic} = \bar{a}_0 + \sum_{j=1}^{p}\bar{a}_j x_{ij} = y_i, \tag{34}$$

$$\sum_{j=1}^{p}(a'_j l_{ij} + a''_j r_{ij}) + n_{il} - p_{il} = \sum_{j=1}^{p}\bar{a}_j l_{ij} = l_i, \tag{35}$$

$$\sum_{j=1}^{p}(a'_j r_{ij} + a''_j l_{ij}) + n_{ir} - p_{ir} = \sum_{j=1}^{p}\bar{a}_j r_{ij} = r_i. \tag{36}$$

The Equations (34), (35), and (36) show that the suggested values satisfy in the constraints (28), (29), and (30) (i.e., we have a feasible solution for GP1). To show the optimality, note that the objective function value of GP1 for this feasible solution is 0, which is the least possible value for a non-negative function. Therefore, the suggested values yield the optimal solution of GP1. Indeed, in the optimal solution we have $a'_j - a''_j = \bar{a}_j$.

(ii). The second part of the theorem immediately follows from (34)-(36). In fact, from (34)-(36) we have:

$$Y_i(\mathbf{a}', \mathbf{a}'') = y_i, \quad L_i(\mathbf{a}', \mathbf{a}'') = l_i, \quad R_i(\mathbf{a}', \mathbf{a}'') = r_i.$$

and the proof is completed. □

The explained method can be applied to trapezoidal fuzzy numbers easily. It is enough to modify the distance introduced in Definition 2.1 for two TPFNs $\tilde{C}_i = (c_{iL}, c_{iU}, \alpha_i, \beta_i),\ i = 1, 2$ as follows:

$$d(\tilde{C}_1, \tilde{C}_2) = |c_{1L} - c_{2L}| + |c_{1U} - c_{2U}| + |\alpha_1 - \alpha_2| + |\beta_1 - \beta_2|. \tag{37}$$

In the case of trapezoidal fuzzy input-output data, suppose $\tilde{x}_{ij} = (x_{ijL}, x_{ijU}, l_{ij}, r_{ij})$ and $\tilde{y}_i = (y_{iL}, y_{iU}, l_i, r_i),\ i = 1, \cdots, n,\ j = 1, \cdots, p$. Following a process similar to the triangular case leads to the following GP model to estimate the coefficients of model (2):

$$
\begin{aligned}
(GP2):\quad \min\quad & z=\sum_{i=1}^{n}(n_{iL}+p_{iL}+n_{iU}+p_{iU}+n_{il}+p_{il}+n_{ir}+p_{ir})\\
s.t.\quad & a_0+\sum_{j=1}^{p}(a'_j x_{ijL}-a''_j x_{ijU})+n_{iL}-p_{iL}=y_{iL} && i=1,2,\cdots,n,\\
& a_0+\sum_{j=1}^{p}(a'_j x_{ijU}-a''_j x_{ijL})+n_{iU}-p_{iU}=y_{iU} && i=1,2,\cdots,n,\\
& \sum_{j=1}^{p}(a'_j l_{ij}+a''_j r_{ij})+n_{il}-p_{il}=l_i && i=1,2,\cdots,n,\\
& \sum_{j=1}^{p}(a'_j r_{ij}+a''_j l_{ij})+n_{ir}-p_{ir}=r_i && i=1,2,\cdots,n,\\
& n_{ik}p_{ik}=0 && i=1,2,\cdots,n,\quad k=L,U,l,r,\\
& a_0\in\mathbb{R},\quad a'_j,a''_j\geq 0 && j=1,2,\cdots,p,\\
& n_{ik},p_{ik}\geq 0 && i=1,2,\cdots,n,\quad k=L,U,l,r.
\end{aligned}
$$

The deviation variables in GP2 have a similar definition as in GP1. In a similar way, Proposition 2.9, Theorem 2.10, and the other properties of GP1 can be proved for GP2.

3. **Numerical Results and Comparisons**

In this section two criteria of goodness are presented. Then, GPA is compared with some similar methods by solving some numerical examples and performing a simulation study.

3.1. **Performance Evaluation.** To evaluate the performance of fuzzy regression model, Kim and Bishu [21] introduced the following scaler as the difference between the triangular membership functions $\tilde{Y}_i(x)$ and $\tilde{y}_i(x)$:

$$E_i=\int_{S_{\tilde{Y}_i}\bigcup S_{\tilde{y}_i}}|\tilde{Y}_i(x)-\tilde{y}_i(x)|dx,\quad i=1,2,\cdots,n, \tag{38}$$

where $S_{\tilde{Y}_i}$ and $S_{\tilde{y}_i}$ are the supports of the TFNs $\tilde{Y}_i$ and $\tilde{y}_i$, respectively. Note that E_i is the non-overlapped area between the graphs of $\tilde{Y}_i(x)$ and $\tilde{y}_i(x)$ (Figure 2). It is clear that a smaller value of E_i denotes a better estimation, or a better fitness of the ith estimated response to the ith observed response. We use the total errors (TE) in our comparisons:

$$TE=\sum_{i=1}^{n}E_i. \tag{39}$$

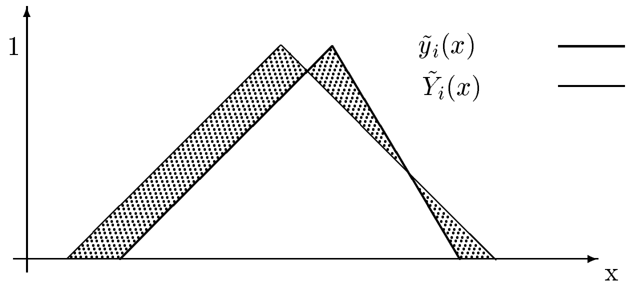

Figure 2. The Difference between Two Membership Functions

It is notable that the centers of TFNs have a special importance since they have the highest membership value. Therefore, to evaluate the performance of fuzzy regression model a new parameter named λ has been introduced in [14]. λ measures the total difference between the centers of observed and estimated responses as follows:

$$\lambda = \sum_{i=1}^{n} \lambda_i \quad \text{where} \quad \lambda_i = |a_0 + \sum_{j=1}^{p} a_j x_{ij} - y_i|. \tag{40}$$

Similarly, in the trapezoidal case we use the non-overlapped area between the membership functions of the ith observed and estimated response, as the error of the ith estimated response. Furthermore, similar to the triangular case, to measure the fitness of the middle points, we use two parameters named λ_L and λ_U. λ_L (λ_U) is the total difference between the lower (upper) middle points of observed and estimated responses:

$$\lambda_L = \sum_{i=1}^{n} \lambda_{iL} \quad \text{where} \quad \lambda_{iL} = |a_0 + \sum_{j=1}^{p} (a'_j x_{ijL} - a''_j x_{ijU}) - y_{iL}|, \tag{41}$$

$$\lambda_U = \sum_{i=1}^{n} \lambda_{iU} \quad \text{where} \quad \lambda_{iU} = |a_0 + \sum_{j=1}^{p} (a'_j x_{ijU} - a''_j x_{ijL}) - y_{iU}|. \tag{42}$$

3.2. **Numerical Examples and Simulation Study.** In this subsection, some numerical examples are solved. One of them (Example 3.1) has been presented in [19] which contains symmetric triangular fuzzy input-output data. The crisp coefficients have been considered in the model of Diamond [11] (D), the models of Kao and Chyu in [19] and [20] (KC1 and KC2, respectively), the model (7) of Arabpour and Tata [3] (AT), and the model of Choi and Buckley [7] (CB). Therefore, based on Example 3.1, GPA is compared with D, KC1, KC2, AT and CB. Moreover, a simulation study is made to compare GPA with D and AT. Also, an example having trapezoidal fuzzy input-output data is solved and compared with [3]. Finally, an application is presented.

Example 3.1. The data of Table 1 are those of [19]. They have also been used in [3, 20]. The errors of estimations are presented in Table 1.

Data			Errors (E_i)					
i	$\tilde{x}_i$	$\tilde{y}_i$	D	KC1	KC2	AT	CB	GPA
1	(2 , 0.5)	(4 , 0.5)	0.704	0.849	1.004	0.704	0.722	0.722
2	(3.5 , 0.5)	(5.5 , 0.5)	0.306	0.208	0.723	0.306	0.722	0.278
3	(5.5 , 1)	(7.5 , 1)	1.406	1.490	1.683	1.406	1.444	1.368
4	(7 , 0.5)	(6.5 , 0.5)	0.760	0.914	1.122	0.760	0.722	0.684
5	(8.5 , 0.5)	(8.5 , 0.5)	0.675	0.763	0.924	0.675	0.722	0.722
6	(10.5 , 1)	(8 , 1)	1.385	1.453	1.636	1.385	1.444	0.964
7	(11 , 0.5)	(10.5 , 0.5)	0.765	1.000	1.561	0.765	0.722	0.722
8	(12.5 , 0.5)	(9.5 , 0.5)	0.738	0.809	0.986	0.738	0.722	0.278
TE			6.739	7.486	9.640	6.739	7.220	5.737
λ			5.857	5.818	5.827	5.857	5.556	5.556

Table 1. The Data and Errors of Example 3.1

The last two rows of Table 1 show better estimation of GPA in comparison with the other methods according to both TE and λ in this example. The different regression models are as follows: (As mentioned above, KC1 cannot be used practically. The values of E_i and TE for KC1, given in Table 1, and the FLR model $\tilde{Y}_{KC1}$, given in the following, have been taken from [19].)

$$\begin{aligned}
\tilde{Y}_D &= 3.4877 + 0.5306\tilde{x}\\
\tilde{Y}_{KC1} &= 3.5724 + 0.5193\tilde{x} + (0, 0.24, 0.24)\\
\tilde{Y}_{KC2} &= 3.565 + 0.522\tilde{x} + (-0.011, 0.951, 0.949)\\
\tilde{Y}_{AT} &= 3.4877 + 0.5306\tilde{x}\\
\tilde{Y}_{CB} &= 3.9443 + 0.4445\tilde{x} + (0, 0.2777, 0.2777)\\
\tilde{Y}_{GPA} &= 3.9444 + 0.4444\tilde{x}.
\end{aligned}$$

Note that the results of AT and D are the same. Because, AT uses the metric and method of D, when the data are TFNs. Among the methods which have considered crisp coefficients AT and CB are more up to date. However, as mentioned previously, CB does not guarantee the nonnegativity of the spreads of TFNs (see also the result of CB in the application (restaurant data) in the following). Therefore, a simulation study is used in the following example to compare GPA with AT (and D which have the same results as AT).

Example 3.2. To compare GPA with AT and D a simulation study was performed as follows:

N=100 samples were generated, each containing n=100 observations. The centers x_{ij} of fuzzy inputs $\tilde{x}_{ij} = (x_{ij}, l_{ij}, r_{ij})$ were drawn from the uniform distribution on the interval (0,100), and the spreads (l_{ij}, r_{ij}) were random points in the unit square. Also, for each i, the center of ith observed response $\tilde{y}_i = (y_i, l_i, r_i)$ was chosen as:

$$y_i = A_0 + A_1 x_i + \varepsilon_i, \tag{43}$$

where $\varepsilon_i \overset{iid}{\sim} N(0,1)$, $A_0 = 1$, and A_1 varies over 0.6, 0.8, 1, 1.2, 1.4. Further, (l_i, r_i) corresponding to $\tilde{y}_i$ was a random point in the unit square. The results of simulation study are as follows:

Figure 3 depicts TE and λ in GPA versus AT and D in 10 runs of a program written by MATLAB software [25] (i.e. in solving 1000 samples). More details of the results are given in Table 2. In each run, k1 is the number of samples in which TE in GPA is better than TE in AT and D, and k2 is the number of samples in which λ in GPA is better than λ in AT and D. In 602 out of 1000 samples (i.e. in 60.2% of instances), TE in GPA is better than TE in AT and D. Also, in 970 out of 1000 samples (i.e. in 97% of instances) λ in GPA is better than λ in AT and D (Number of points over the bisectors).

Run no.	1	2	3	4	5	6	7	8	9	10	Average
k1	64	64	52	63	61	65	54	62	60	57	60.2
k2	98	97	98	95	97	95	99	96	95	100	97

Table 2. The Results of Simulation Study in 10 Runs to Compare GPA with AT and D

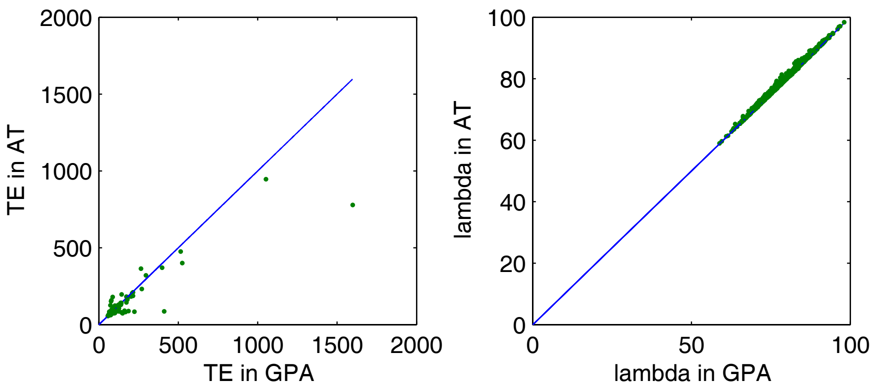

Figure 3. The Results of Simulation Study in Example 3.2 to Compare GPA with AT and D

Example 3.3. Consider the 8 trapezoidal fuzzy data which have been presented in [3], and given in the left half of Table 3. The errors of estimations in AT and GPA (the FLR model obtained from GP2) are presented in the right half of Table 3.

Data			Errors	
i	$\tilde{x}_i$	$\tilde{y}_i$	AT	GPA
1	(1.75, 2.25, 0.25, 0.25)	(3.75, 4.25, 0.25, 0.25)	1.604	1.083
2	(3.25, 3.75, 0.25, 0.25)	(5.25, 5.75, 0.25, 0.25)	0.363	0.417
3	(5.00, 6.00, 0.50, 0.50)	(7.00, 8.00, 0.50, 0.50)	2.051	2.013
4	(6.75, 7.25, 0.25, 0.25)	(6.25, 6.75, 0.25, 0.25)	1.308	1.006
5	(8.25, 8.75, 0.25, 0.25)	(8.25, 8.75, 0.25, 0.25)	0.966	1.083
6	(10.00, 11.00,0.50, 0.50)	(7.50, 8.50, 0.50, 0.50)	3.236	1.221
7	(10.75, 11.25,0.25, 0.25)	(10.25, 10.75, 0.25, 0.25)	1.149	1.083
8	(12.25, 12.75,0.25, 0.25)	(9.25 , 9.75, 0.25, 0.25)	1.463	0.417
TE			12.140	8.323
λ_L			5.862	5.833
λ_U			5.862	5.834

Table 3. The Data and Errors of Example 3.3

The last three rows of Table 3 show that the values of TE, λ_L and λ_U in GPA are better than their values in AT. The FLR models derived from AT and GPA are as follows:

$$\tilde{Y}_{AT} = 3.4773 + 0.5319\tilde{x}$$
$$\tilde{Y}_{GPA} = 3.9444 + 0.4444\tilde{x}.$$

3.2.1. **An Application: Restaurants Data.** Consider the data of Table 4 which have been presented in [13]. The data contains n=30 observations of a fuzzy output (decision on cellar) and two fuzzy input variables (decision on cooking and decision on environment) in a set of restaurants in Rome.

Comparing the obtained results (the last two rows of Table 4) shows that both TE and λ in GPA are better than those of D and AT (which are the same) for these data. This results confirm again better performance of GPA than D and AT which have used similar FLR model. The regression models of the compared methods are as follows:

$$\tilde{Y}_D = 0.0164 + 0.6021\tilde{x}_1 + 0.3623\tilde{x}_2$$
$$\tilde{Y}_{AT} = 0.0164 + 0.6021\tilde{x}_1 + 0.3623\tilde{x}_2$$
$$\tilde{Y}_{GPA} = -0.0000 + 1.0000\tilde{x}_1 + 0.0000\tilde{x}_2.$$

Also, the regression model of CB is $\tilde{Y}_{CB} = -0.0006 + 0.9999\tilde{x}_1 + 0.0002\tilde{x}_2 + (0, -0.0070, -0.0083)$ (The non-linear programming models of Choi and Buckley [7] have been solved by LINGO software [24]). Note that the error term of their model has negative spreads in this example.

Data				Errors		
i	$\tilde{x}_{i1}$	$\tilde{x}_{i2}$	$\tilde{y}_i$	D	AT	GPA
1	(7,0.5 ,1.25)	(8,0.75,1)	(8,0.75,1)	1.1888	1.1888	1.250
2	(7,0.5 ,1.25)	(7,0.5 ,1.25)	(6,0.25,0.5)	1.1717	1.1717	1.250
3	(6,0.25,0.5)	(7,0.5 ,1.25)	(6,0.25,0.5)	0.3832	0.3832	0.000
4	(8,0.75,1)	(9,0 ,1)	(9,0 ,1)	1.2045	1.2045	1.375
5	(8,0.75,1)	(8,0.75,1)	(8,0.75,1)	0.4996	0.4996	0.000
6	(6,0.25,0.5)	(7,0.5 ,1.25)	(5,0 ,1)	1.0219	1.0219	0.825
7	(7,0.5 ,1.25)	(8,0.75,1)	(7,0.5 ,1.25)	0.1457	0.1457	0.000
8	(7,0.5 ,1.25)	(7,0.5 ,1.25)	(5,0 ,1)	1.3438	1.3438	1.375
9	(7,0.5 ,1.25)	(8,0.75,1)	(7,0.5 ,1.25)	0.1457	0.1457	0.000
10	(6,0.25,0.5)	(7,0.5 ,1.25)	(6,0.25,0.5)	0.3832	0.3832	0.000
11	(7,0.5 ,1.25)	(8,0.75,1)	(8,0.75,1)	1.1888	1.1888	1.250
12	(7,0.5 ,1.25)	(6,0.25,0.5)	(6,0.25,0.5)	0.7718	0.7718	1.250
13	(7,0.5 ,1.25)	(8,0.75,1)	(9,0 ,1)	1.3438	1.3438	1.375
14	(7,0.5 ,1.25)	(8,0.75,1)	(8,0.75,1)	1.1888	1.1888	1.250
15	(7,0.5 ,1.25)	(7,0.5 ,1.25)	(7,0.5 ,1.25)	0.4475	0.4475	0.000
16	(7,0.5 ,1.25)	(7,0.5 ,1.25)	(7,0.5 ,1.25)	0.4475	0.4475	0.000
17	(6,0.25,0.5)	(7,0.5 ,1.25)	(6,0.25,0.5)	0.3832	0.3832	0.000
18	(7,0.5 ,1.25)	(8,0.75,1)	(7,0.5 ,1.25)	0.1457	0.1457	0.000
19	(7,0.5 ,1.25)	(7,0.5 ,1.25)	(8,0.75,1)	1.4519	1.4519	1.250
20	(7,0.5 ,1.25)	(9,0 ,1)	(7,0.5 ,1.25)	0.8592	0.8592	0.000
21	(7,0.5 ,1.25)	(8,0.75,1)	(7,0.5 ,1.25)	0.1457	0.1457	0.000
22	(7,0.5 ,1.25)	(8,0.75,1)	(6,0.25,0.5)	1.2188	1.2188	1.250
23	(7,0.5 ,1.25)	(9,0 ,1)	(7,0.5 ,1.25)	0.8592	0.8592	0.000
24	(7,0.5 ,1.25)	(7,0.5 ,1.25)	(8,0.75,1)	1.4519	1.4519	1.250
25	(7,0.5 ,1.25)	(7,0.5 ,1.25)	(6,0.25,0.5)	1.1717	1.1717	1.250
26	(7,0.5 ,1.25)	(7,0.5 ,1.25)	(6,0.25,0.5)	1.1717	1.1717	1.250
27	(7,0.5 ,1.25)	(7,0.5 ,1.25)	(7,0.5 ,1.25)	0.4475	0.4475	0.000
28	(7,0.5 ,1.25)	(8,0.75,1)	(7,0.5 ,1.25)	0.1457	0.1457	0.000
29	(7,0.5 ,1.25)	(7,0.5 ,1.25)	(7,0.5 ,1.25)	0.4475	0.4475	0.000
30	(6,0.25,0.5)	(7,0.5 ,1.25)	(6,0.25,0.5)	0.3832	0.3832	0.000
TE				23.1594	23.1594	17.450
λ				18.1118	18.1118	16.000

Table 4. The Restaurants Data and Errors of Estimations

4. **Concluding Remarks**

A fuzzy linear regression model with fuzzy input-output data and crisp coefficients has been considered in this paper. Some available methods which have considered similar model have been introduced and some shortcomings of them have been discussed. Then, a new distance between triangular and trapezoidal fuzzy numbers has been introduced. The proposed distance takes into account both the middle points and the spreads of fuzzy data. Based on the suggested distance, a non-linear programming model has been formulated to calculate the regression coefficients. It minimizes the total absolute deviations between the middle points (and spreads) of observed and estimated responses. Some features of the proposed approach are as follows:

1. The nonlinear programming model which has been presented to calculate the regression coefficients is converted to a goal programming model and then to a linear programming (LP) model easily. The advantage of this conversion is that simplex method yields the exact solution od LP models. Whereas, the available algorithms for solving nonlinear programming models often give approximate solutions.

2. The estimated responses from the proposed model are very close to the observed responses in comparison with some other methods in view of membership function.

3. Closeness of the centers of estimated responses to the centers of corresponding observed responses has a special importance. Because, they have the highest membership value. A new parameter (λ) has been used to measure the total distance between the centers of observed and estimated responses. This parameter in the proposed approach is better than that of some other methods in numerical examples and simulation study.

4. Handling both the problems with triangular fuzzy numbers (symmetric and asymmetric) and the problems with trapezoidal fuzzy numbers (symmetric and asymmetric) is another feature of the proposed approach.

5. Obtaining the negative spreads for fuzzy numbers is one of the shortcomings of some previous methods, which is not encountered in the proposed method.

6. The multiplication of trapezoidal and triangular fuzzy numbers by a scaler depend on the sign of the scaler. Therefore, different models (least-squares or linear programming, $\cdots$) are required to be formulated and solved based on the sign of regression coefficients. The advantage of the proposed model in this paper is that it is independent of the sign of regression coefficients. So, only one linear programming model is solved for any status of the sign of regression coefficients.

7. One shortcoming of LP-based approaches in general is that the LP models which are used to estimate the parameters of FLR models have a large number of constraints, specially for a large number of observations. Although it seems that our approach also has this shortcoming, with available LP solvers (e.g. LINGO [24] and MPL [26]), the size of LP is not important.

Acknowledgement. The authors would like to express their gratitude to the anonymous referees for their comments and suggestions on the first version of this paper. They are also grateful to Mr. Ehsan Mehrabi Kermani for editing the English text. Also, partial support by the Fuzzy Systems and Applications Center of Excellence at Shahid Bahonar University of Kerman is gratefully acknowledged.

References

[1] S. Abbasbandy and M. Alavi, *A method for solving fuzzy linear systems*, Iranian Journal of Fuzzy Systems, **2(2)** (2005), 37-44.

[2] S. M. Abu Nayeem and M. Pal, *The p-center problem on fuzzy networks and reduction of cost*, Iranian Journal of Fuzzy Systems, **5(1)** (2008), 1-26.

[3] A. R. Arabpour and M. Tata, *Estimating the parameters of a fuzzy linear regression model*, Iranian Journal of Fuzzy Systems, **5(2)** (2008), 1-19.

[4] M. S. Bazaraa, J. J. Jarvis and H. D. Sherali, *Linear programming and network flows*, Jhon Wiley and Sons, New York, 1990.

[5] J. M. Cadenas and J. L. Verdegay, *A primer on fuzzy optimization models and methods*, Iranian Journal of Fuzzy Systems, **3(1)** (2006), 1-22.

[6] A. Celmins, *Least squares model fitting to fuzzy vector data*, Fuzzy Sets and Systems, **22** (1987), 245-269.

[7] S. H. Choi and J. J. Buckley, *Fuzzy regression using least absolute deviation estimators*, Soft Computing, **12(3)** (2008), 257-263.

[8] S. H. Choi, H. K. Kim and K. O. Park, *Nonlinear regression quantiles estimation*, Journal of the Korean Statistical Society, **29** (2000), 187-199.

[9] R. Coppi, P. D'Urso, P. Giordani and A. Santoro, *Least squares estimation of a linear regression model with LR fuzzy response*, Computational Statistics & Data Analysis, **51** (2006), 267-286.

[10] P. Diamond, *Fuzzy least squares*, Information Sciences, **46** (1988), 141-157.

[11] P. Diamond, *Least squares fitting of several variables*, Preprint of Second IFSA Congress, Tokyo, July (1987), 20-25.

[12] P. Diamond and P. Kloeden, *Metrics spaces of fuzzy sets: theory and applications*, World Scientific, 1994.

[13] P. D'Urso, *Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data*, Computational Statistics & Data Analysis, **42** (2003), 47-72.

[14] H. Hassanpour, H. R. Maleki and M. A. Yaghoobi, *A goal programming approach to fuzzy linear regression with non-fuzzy input and fuzzy output data*, APJOR, **26(5)** (2009), 587-604.

[15] H. Hassanpour, H. R. Maleki and M. A. Yaghoobi, *A note on evaluation of fuzzy linear regression models by comparing membership functions*, Iranian Journal of Fuzzy Systems, **6(2)** (2009), 1-6.

[16] M. Hojati, C. R. Bector and K. Smimou, *A simple method for computation of fuzzy linear regression*, EJOR, **166** (2005), 172-184.

[17] W. L. Hung and M. S. Yang, *An omission approach for detecting outliers in fuzzy regression models*, Fuzzy Sets and Systems, **157** (2006), 3109-3122.

[18] S. Jozsef, *On the effect of linear data transformations in possibilistic fuzzy linear regression*, Fuzzy Sets and Systems, **45** (1992), 185-188.

[19] C. Kao and C. L. Chyu, *A fuzzy linear regression model with better explanatory power*, Fuzzy Sets and Systems, **126** (2002), 401-409.

[20] C. Kao and C. L. Chyu, *Least-squares estimates in fuzzy regression analysis*, EJOR, **148** (2003), 426-435.

[21] B. Kim and R. R. Bishu, *Evaluation of fuzzy linear regression models by comparing membership functions*, Fuzzy Sets and Systems, **100** (1998), 343-352.

[22] H. K. Kim, J. H. Yoon and Y. Li, *Asymptotic properties of least squares estimation with fuzzy observations*, Information Sciences, **178** (2008), 439-451.

[23] R. Koenker and G. Bassett, *Regression quantiles*, Econometrica, **46** (1978), 33-50.

[24] L. Schage, *LINGO release 6.0*, LINDO System, Inc., 1999.

[25] *MATLAB 7.0*, The Mathworks, (www.mathworks.com).

[26] Maximal software inc., (www.maximalsoftware.com).

[27] M. Ming, M. Friedman and A. Kandel, *General fuzzy least-squares*, Fuzzy Sets and Systems, **88** (1997), 107-118.

[28] M. Modarres, E. Nasrabadi and M. M. Nasrabadi, *Fuzzy linear regression analysis from the point of view risk*, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, **12(5)** (2004), 635-649.

[29] M. Modarres, E. Nasrabadi and M. M. Nasrabadi, *Fuzzy linear regression models with least square errors*, Applied Mathematics and Computation, **163** (2005), 977-989.

[30] J. Mohammadi and S. M. Taheri, *Pedomodels fitting with fuzzy least squares regression*, Iranian Journal of Fuzzy Systems, **1(29)** (2004), 45-62.

[31] E. C. Özelkan and L. Duckstein, *Multi-objective fuzzy regression: a general framework*, Computers & Operations Research, **27** (2000), 635-652.

[32] G. Peters, *Fuzzy linear regression with fuzzy intervals*, Fuzzy Sets and Systems, **63** (1994), 45-55.

[33] D. T. Redden and W. H. Woodall, *Properties of certain fuzzy linear regression models*, Fuzzy Sets and Systems, **64** (1994), 361-375.

[34] M. R. Safi, H. R. Maleki and E. Zaeimazad, *A note on zimmermann method for solving fuzzy linear programming problems*, Iranian Journal of Fuzzy Systems, **4(2)** (2007), 31-46.

[35] M. Sakawa, *Fuzzy sets and interactive multiobjective optimization*, Plenum Press, New York, 1993.

[36] M. Sakawa and H. Yano, *Fuzzy linear regression and its applications*, In: Studies in Fuzziness, Fuzzy Regression Analysis, J. Kacprzyk and M. Fedrizzi, Omnitech Press, Warsaw, Poland, **1** (1992), 61-80.

[37] M. Sakawa and H. Yano, *Multiobjective fuzzy linear regression analysis for fuzzy input-output data*, Fuzzy Sets and Systems, **47** (1992), 173-181.

[38] D. A. Savic and W. Pedrycz, *Evaluation of fuzzy linear regression models*, Fuzzy Sets and Systems, **39** (1991), 51-63.

[39] E. Shivanian, E. Khorram and A. Ghodousian, *Optimization of linear objective function subject to fuzzy relation inequalities constraints with max-average composition*, Iranian Journal of Fuzzy Systems, **4(2)** (2007), 15-30.

[40] H. Tanaka, *Fuzzy data analysis by possibilistic linear models*, Fuzzy Sets and Systems, **24** (1987), 363-375.

[41] H. Tanaka, I. Hayashi and J. Watada, *Possibilistic linear regression analysis for fuzzy data*, EJOR, **40** (1989), 389-396.

[42] H. Tanaka, S. Uejima and K. Asai, *Linear regression analysis with fuzzy model*, IEEE Transactions on Systems, Man and Cybernetics, **12** (1982), 903-907.

[43] S. M. Vaezpour and F. Karimi, *T-best approximation in fuzzy normed spaces*, Iranian Journal of Fuzzy Systems, **5(2)** (2008), 93-99.

[44] N. Wang, W. X. Zhang and C. L. Mei, *Fuzzy nonparametric regression based on local linear smoothing technique*, Information Sciences, **177** (2007), 3882-3900.

[45] H. C. Wu, *Fuzzy estimates of regression parameters in linear regression models for imprecise input and output data*, Computational Statistics & Data Analysis, **42** (2003), 203 - 217.

[46] H. C. Wu, *Fuzzy linear regression model based on fuzzy scalar product*, Soft Computing, **12** (2008), 469-477.

[47] A. Wünsche and W. Näther, *Least-squares fuzzy regression with fuzzy random variables*, Fuzzy Sets and Systems, **130** (2002), 43-50.

[48] R. Xu and C. Li, *Multidimensional least-squares fitting with a fuzzy model*, Fuzzy Sets and Systems, **119** (2001), 215-223.

[49] M. S. Yang and C. H. Ko, *On a class of fuzzy c-numbers clustering procedures for fuzzy data*, Fuzzy sets and Systems, **84** (1996), 49-60.

[50] M. S. Yang and T. S. Lin, *Fuzzy least squares linear regression analysis for fuzzy input-output data*, Fuzzy Sets and Systems, **126** (2002), 389-399.

[51] M. A. Yaghoobi and M. Tamiz, *A short note on the relationship between goal programming and fuzzy programming for vectormaximum problems*, Iranian Journal of Fuzzy Systems, **2(2)** (2005), 31-36.

[52] L. A. Zadeh, *Fuzzy sets*, Information and Control, **8** (1965), 338-353.

[53] L. A. Zadeh, *Fuzzy sets as a basis for a theory of possibility*, Fuzzy Sets and Systems, **1** (1978), 3-28.

H. Hassanpour*, Department of Mathematics, University of Birjand, Birjand, Iran
E-mail address: hassanhassanpur@yahoo.com

H. R. Maleki, Faculty of Basic Sciences, Shiraz University of Technology, Shiraz, Iran (Corresponding author)
E-mail address: maleki@sutech.ac.ir

M. A. Yaghoobi, Department of Statistics, Shahid Bahonar University of Kerman, Kerman, Iran
E-mail address: Yaghoobi@mail.uk.ac.ir

*Corresponding author