Journal of Uncertain Systems
Vol.7, No.4, pp.265-276, 2013
Online at: www.jus.org.uk

A Least-Absolutes Regression Model for Imprecise Response Based on the Generalized Hausdorff-Metric

J. Chachi[a], S.M. Taheri[b,c,*]

[a] *Department of Mathematics, Statistics and Computer Sciences, Semnan University, Semnan 35195–363, Iran*

[b] *Department of Engineering Science, College of Engineering, University of Tehran, Tehran, Iran*

[c] *Department of Mathematical Sciences, Isfahan University of Technology, Isfahan 84156–83111, Iran*

Received 20 February 2012; Revised 16 July 2013

Abstract

This study attempts to develop a least-absolutes fuzzy regression model for crisp input-fuzzy output data. To estimate the parameters in the proposed model, the generalized Hausdorff-metric on the space of LR-fuzzy numbers is used. The problem of estimation of the interested parameters relies on a non-linear optimization problem, which is also translated to a linear optimization problem, making the computations of the proposed method very simple. For evaluating the performance of the model, three goodness-of-fit criteria are employed. Numerical comparative studies, based on four data sets including a real agricultural data set, indicate that the proposed model could be a rational substituted model of some common ones, especially for large sample data set.

Keywords: fuzzy regression, generalized Hausdorff-metric, goodness-of-fit, least-absolutes deviations, linear programming, non-linear programming, LR-fuzzy number

1 Introduction

Regression analysis has been the most frequently used technique in various areas of applied Statistics. There are situations in practice, however, in which the classical regression methods are confronted with imprecise data. Moreover, we encounter many situations where necessary assumptions for statistical regression analysis can not be met because they are not based on random uncertainty. Fuzzy regression methods were proposed to model the relationships between variables, when the available data are fuzzy, the relationships are imprecise, or the underlying statistical assumptions are not fulfilled.

Fuzzy regression methods have been previously treated in the literature considering different sources of uncertainty (i.e. possibilistic uncertainty and probabilistic uncertainty) as well as considering different kinds of input/output data (i.e. crisp data and fuzzy data). They also have wide applications in many areas including engineering, biology, business, and economics. In the following we briefly review some studies on this topic.

Tanaka et al. [40, 41] formulated the fuzzy linear regression problem for crisp explanatory variables and crisp/fuzzy response variables as linear programming problems. To estimate the parameters of the regression model, they minimize the fuzziness of model by minimizing the total spreads of its fuzzy coefficients, subject to including the data points of each sample within a specified feasible interval. Celminš [4] and Diamond [10], using certain distances between fuzzy numbers, proposed some fuzzy least-squares approaches to the problem of fuzzy regression modeling. Chang and Lee [6] proposed a modification of the fuzzy linear regression model based on the approach of Tanaka, by allowing the spreads of the parameters to be unrestricted in sign. Kim and Bishu [26] proposed an approach based on the criterion of minimizing the difference of membership values between the observed and estimated fuzzy dependent variable. Xu and Li [45] discussed the problem of multidimensional least-squares fitting, and proposed a fuzzy linear regression model as an analogue of the traditional linear least-squares method. Kao and Chyu [23] proposed a two-stage methodology to construct the fuzzy regression model. In the first stage, the crisp coefficients of the model are estimated, and in the second stage, the fuzzy error term is determined. They also investigated a least-squares method in fuzzy regression

*Corresponding author. Email: taheri@cc.iut.ac.ir (S.M. Taheri).

analysis for crisp/fuzzy input-fuzzy output data [24]. D'Urso [12] developed a certain adaptive least-squares estimation procedure for regression models with crisp/fuzzy inputs and crisp/fuzzy output. Nasrabadi and Nasrabadi [34] defined new arithmetic operations for symmetric fuzzy numbers and employed them in fuzzy regression analysis. Using the new arithmetic operations, their model could avoid the spreads increasing problem. Kim et al. [27] proposed a two-stage method to construct the fuzzy linear regression model. In the first stage, fuzzy observations were defuzzified to crisp values and the least-absolutes estimators were derived as the crisp regression coefficients. In the second stage, the fuzzy error term was calculated. Coppi et al. [9] investigated a linear regression model for studying the dependence of a LR-fuzzy response variable on a set of crisp explanatory variables, along with a suitable iterative least-squares estimation procedure. Arabpour and Tata [1], by using normal equations corresponding to some least-squares models, calculated the fuzzy regression coefficients for crisp/fuzzy input-fuzzy output data. Chen and Dang [7] investigated a variable spread fuzzy linear regression model, based on a three-phase method. In the first phase, regression coefficients were treated as fuzzy numbers and the membership functions of the least-squares estimates of the regression coefficients were constructed. In the second phase, the fuzzy regression coefficients were defuzzified to obtain the crisp regression coefficients. Finally, in the third phase, for each instance, the fuzzy error term was determined by a mathematical programming method. Guo et al. [18] proposed a scalar variable formation of fuzzy regression model based on the axiomatic credibility measure foundation. Lu and Wang [28] proposed an enhanced fuzzy linear regression model, in which the spreads of the estimated dependent variables are able to fit the spreads of the observed dependent variables, no matter the spreads of the observed dependent variables are increased, decreased, or unchanged as the magnitudes and spreads of the independent variables change. A revisited approach for possibilistic fuzzy regression methods is investigated by Bisserier et al. [3], in which a new modified fuzzy linear model form is introduced where the identified model output can envelop all the observed data and ensure a total inclusion property. Ferraro et al. [15] proposed a linear regression model for imprecise response, based on a least-squares method. They also analyzed limit distribution and asymptotic properties of the estimators and applied them to the determination of the confidence regions and hypothesis testing procedures (see also [14]). To handle the large variation issues in fuzzy input-output data, Yu and Lee [46] proposed a quadratic programming method to construct a piecewise regression model. Hassanpour et al. [19] used a least-absolutes approach to calculate the crisp regression coefficients of a fuzzy regression model for fuzzy input-output data. They also proposed a goal programming approach to determine the coefficients of fuzzy linear regression [20]. Using tabu search and harmony search methods, Mashinchi et al. [29] proposed a metaheuristic unconstrained global continuous optimization approach to the fuzzy regression problem. Under the considerations of fuzzy parameters and fuzzy arithmetic operations, Wu [44] proposed a fuzzy linear regression model which has the similar form as that of conventional one. He conducted the h-level (conventional) linear regression models of fuzzy linear regression model for the sake of invoking the statistical techniques in (conventional) linear regression analysis for real-valued data. Two fuzzy logistic regression models for the case when the explanatory variables are crisp and the value of the binary response variable is reported as a number between zero and one, or by some linguistic terms, are investigated by Pourahmad et al. [36, 37]. By introducing and applying new metrics on the space of fuzzy numbers, Taheri and Kelkinnama [25, 38, 39] developed least-absolutes deviations approaches to fuzzy regression analysis. D'Urso et al. [13] proposed a robust fuzzy linear regression model based on the so-called least median squares-weighted least squares estimation procedure to deal with data contaminated by outliers. Chachi et al. [5], by using the α-cuts of fuzzy input-fuzzy output observations, proposed a least-squares method to estimate the crisp parameters of a fuzzy regression model. Guo et al. [17] proposed uncertainty copula linked multivariate uncertainty distributional theory for developing an uncertainty distributional structure for uncertainty linear regression models. Tutmez [43] investigated a weighted fuzzy regression analysis based on spatial dependence measure of the memberships.

As the reader could find from the above literature review, there have been a few works on the fuzzy regression modeling by robust methods (e.g. least absolutes methods). In this research, we introduce and develop a new robust approach to regression analysis in fuzzy environment. Several numerical investigations show the efficiency and performances of the robust proposed method. The results of comparative studies show that the proposed method is able to determine the regression coefficients with a good explanatory power, especially for the case with several input variables and large size sample data set.

The rest of this paper is organized as follows: In the next section, the concepts of Hausdorff-metric and its generalization, which are essential for the proposed method in this paper, are reviewed. Section 3 provides the formalization of the proposed multiple fuzzy least-absolutes regression model for studying the functional dependence of a LR-fuzzy response variable on a set of crisp explanatory variables. Afterwards, in Section 4, some indices are provided to evaluate the goodness-of-fit of the proposed model. In Section 5, by using the results of four numerical examples, we provide some comparative studies. Also the proposed method is illustrated in details, with reference to an application to a real data set in the framework of an environmental study. Finally, in Section 6, we make some concluding remarks.

Through this paper, it is assumed that the reader is familiar with elementary fuzzy sets and fuzzy arithmetics. A brief review of the necessary concepts and results are given in Appendix.

2 The Hausdorff-metric and Its Generalization

One of the most important aspects of the analysis of imprecise data, especially those related with formalization the parameter estimation methods of a regression model, is the usage of a convenient distance on the family of fuzzy numbers, which is easy to handle and interpret and which reflects the intuitive meaning of difference between fuzzy numbers.

Several metrics on the family of fuzzy numbers have been defined by authors (see, for example, [2, 11, 42]). But the generalized Hausdorff-metric, which is used in this paper, not only fulfills many good properties but is also easy to calculate and handle for statistical purposes, and is therefore useful from the practical point of view. More detailed discussions on this topic can be found in [16, 22, 33, 35].

Definition 2.1. *Let $\mathcal{K}_c(\mathbb{R}^d)$ be the family of all d-dimensional non-empty compact convex sets on $\mathbb{R}^d$. The Hausdorff-metric between sets $\mathbf{A}, \mathbf{B} \in \mathcal{K}_c(\mathbb{R}^d)$ is defined by*

$$d_H(\mathbf{A}, \mathbf{B}) = \max\{\sup_{\mathbf{b}\in\mathbf{B}} \inf_{\mathbf{a}\in\mathbf{A}} \|\mathbf{a}-\mathbf{b}\|, \sup_{\mathbf{a}\in\mathbf{A}} \inf_{\mathbf{b}\in\mathbf{B}} \|\mathbf{a}-\mathbf{b}\|\},$$

where $\|\cdot\|$ denotes the Euclidean norm.

Proposition 2.1. *Let $I_1 = [a_1, a_2]$ and $I_2 = [b_1, b_2]$ be two intervals on $\mathbb{R}$. Then, the Hausdorff-metric between I_1 and I_2 is given by*

$$\begin{aligned} d_H(I_1, I_2) &= \max\{|a_1 - b_1|, |a_2 - b_2|\} \\ &= |midI_1 - midI_2| + |sprI_1 - sprI_2|, \end{aligned}$$

where $midI_1 = (a_1 + a_2)/2$ and $sprI_1 = (a_2 - a_1)/2$.

Definition 2.2. *The generalized Hausdorff-metric between fuzzy numbers $\widetilde{A}$ and $\widetilde{B}$ is defined by*

$$\mathcal{D}_p(\widetilde{A}, \widetilde{B}) = \begin{cases} \left(\int_0^1 [d_H(A_\alpha, B_\alpha)]^p d\alpha\right)^{\frac{1}{p}} & \text{if} \quad p \in [1, \infty) \\ \sup_{\alpha\in[0,1]} d_H(A_\alpha, B_\alpha) & \text{if} \quad p = \infty, \end{cases}$$

where A_α and B_α are the α-cuts of the fuzzy numbers $\widetilde{A}$ and $\widetilde{B}$, respectively.

Specially, for symmetric LR-fuzzy numbers $\widetilde{M} = (m, \lambda_m)_L$ and $\widetilde{N} = (n, \lambda_n)_L$, by Proposition 2.1, we have

$$d_H(M_\alpha, N_\alpha) = |m - n| + L^{-1}(\alpha)|\lambda_m - \lambda_n|,$$

and, therefore,

$$\begin{aligned} \mathcal{D}_1(\widetilde{M}, \widetilde{N}) &= |m - n| + \mathcal{L}_1|\lambda_m - \lambda_n|; \qquad \mathcal{L}_1 = \int_0^1 L^{-1}(\alpha)d\alpha, \\ \mathcal{D}_\infty(\widetilde{M}, \widetilde{N}) &= |m - n| + \mathcal{L}_\infty|\lambda_m - \lambda_n|; \qquad \mathcal{L}_\infty = \sup_{\alpha\in[0,1]} L^{-1}(\alpha). \end{aligned}$$

Moreover, for symmetric triangular fuzzy numbers $\widetilde{M} = (m, \lambda_m)_T$ and $\widetilde{N} = (n, \lambda_n)_T$, we obtain

$$\begin{aligned} \mathcal{D}_1(\widetilde{M}, \widetilde{N}) &= |m - n| + 0.5|\lambda_m - \lambda_n|, \\ \mathcal{D}_\infty(\widetilde{M}, \widetilde{N}) &= |m - n| + |\lambda_m - \lambda_n|. \end{aligned}$$

3 The Proposed Model

In this section, using metric $\mathcal{D}_1$ on the set of all symmetric LR-fuzzy numbers, we introduce a least-absolutes method to multiple fuzzy regression analysis. A Similar method can be developed by using $\mathcal{D}_\infty$ instead of $\mathcal{D}_1$.

Assume that the observed data on n statistical units are denoted as $(\widetilde{y}_1, \mathbf{x}_1), \ldots, (\widetilde{y}_n, \mathbf{x}_n)$, where $\widetilde{\mathbf{y}}_{n\times 1} = [\widetilde{y}_1, \ldots, \widetilde{y}_n]^t$ is the vector of symmetric LR-fuzzy numbers, i.e. $\widetilde{y}_i = (y_i, s_i)_L$ $(i = 1, \ldots, n)$, which determines the fuzzy observed of the dependent variable, and $\mathbf{x}_i = [x_{0i}, x_{1i}, \ldots, x_{ki}] \in \mathbb{R}^{k+1}$ $(i = 1, \ldots, n; k < n; x_{0i} = 1)$, forms the $(k+1)$-dimensional vector of crisp observed independent variables. Without loss of generality, we can assume that $x_{ji} > 0$, by a simple translation of all data if necessary. Based on the aforementioned data set, we will consider the following functional dependence between $\widetilde{\mathbf{y}}_{n\times 1}$ and $\mathbf{X}_{n\times(k+1)}$

$$\begin{aligned}
\widetilde{\mathbf{y}}_{n\times 1} &= \mathbf{X}_{n\times(k+1)} \otimes \widetilde{\boldsymbol{\beta}}_{(k+1)\times 1}, \\
\begin{bmatrix} \widetilde{y}_1 \\ \vdots \\ \widetilde{y}_n \end{bmatrix} &= \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} \otimes \begin{bmatrix} (\beta_0, \sigma_0)_L \\ \vdots \\ (\beta_k, \sigma_k)_L \end{bmatrix} = \begin{bmatrix} (\sum_{j=0}^k x_{j1}\beta_j, \sum_{j=0}^k x_{j1}\sigma_j)_L \\ \vdots \\ (\sum_{j=0}^k x_{jn}\beta_j, \sum_{j=0}^k x_{jn}\sigma_j)_L \end{bmatrix}.
\end{aligned}$$

The procedure for estimating the parameter $\widetilde{\boldsymbol{\beta}}_{(k+1)\times 1}$ is based on choosing the best candidate $\widehat{\widetilde{\boldsymbol{\beta}}}_{(k+1)\times 1}$ instead of $\widetilde{\boldsymbol{\beta}}_{(k+1)\times 1}$, consisting of minimizing the total difference between the observed values of the response variable, $\widetilde{\mathbf{y}}_{n\times 1}$, and its theoretical counterpart, $\widehat{\widetilde{\mathbf{y}}}_{n\times 1}$, defined by

$$\widehat{\widetilde{\mathbf{y}}}_{n\times 1} = \mathbf{X}_{n\times(k+1)} \otimes \widehat{\widetilde{\boldsymbol{\beta}}}_{(k+1)\times 1},$$

with respect to the distance $\mathcal{D}_1$. Thus, we consider the least-absolutes optimization problem as follows

$$\begin{aligned}
&\min_{\widetilde{\boldsymbol{\beta}}} \mathcal{D}_1(\widetilde{\mathbf{y}}, \mathbf{X} \otimes \widetilde{\boldsymbol{\beta}}) \\
&s.t. \qquad \boldsymbol{\sigma} \in {\mathbb{R}^+}^{k+1}, \boldsymbol{\beta} \in \mathbb{R}^{k+1},
\end{aligned}$$

or, equivalently

$$\begin{aligned}
&\min_{\widetilde{\boldsymbol{\beta}}} \sum_{i=1}^n \left| y_i - \sum_{j=0}^k x_{ji}\beta_j \right| + \mathcal{L}_1 \sum_{i=1}^n \left| s_i - \sum_{j=0}^k x_{ji}\sigma_j \right| \\
&s.t. \qquad \sigma_j \in \mathbb{R}^+, \beta_j \in \mathbb{R}, \quad j = 0, 1, \ldots, k,
\end{aligned}$$

which is a constrained non-linear programming problem. The minimization of $\mathcal{D}_1$ over $\mathbb{R}^{k+1} \times {\mathbb{R}^+}^{k+1}$ can separately be solved: once for all possible candidates for $\boldsymbol{\beta} = [\beta_0, \beta_1, \ldots, \beta_k] \in \mathbb{R}^{k+1}$, and then for all possible candidates for $\boldsymbol{\sigma} = [\sigma_0, \sigma_1, \ldots, \sigma_k] \in {\mathbb{R}^+}^{k+1}$, which are the center values and the spread values of the fuzzy coefficients $\widetilde{\boldsymbol{\beta}}$, respectively. Thus, the above optimization problem can be rewritten as the following two sub-optimization non-linear programming problems

$$\min_{\widetilde{\boldsymbol{\beta}}} \mathcal{D}_1(\widetilde{\mathbf{y}}, \mathbf{X} \otimes \widetilde{\boldsymbol{\beta}}) \equiv \begin{cases} (A) & \min_{\boldsymbol{\beta}} \sum_{i=1}^n \left| y_i - \sum_{j=0}^k x_{ji}\beta_j \right|, \quad s.t.\ \beta_j \in \mathbb{R} \\ (B) & \mathcal{L}_1 \min_{\boldsymbol{\sigma}} \sum_{i=1}^n \left| s_i - \sum_{j=0}^k x_{ji}\sigma_j \right|, \quad s.t.\ \sigma_j \in \mathbb{R}^+. \end{cases}$$

In order to simplify the above optimization problems, we show how by introducing additional variables, two linear programming problems can handle the optimization problems (A) and (B).

First, we consider the sub-optimization problem (A). Let ε_i^+ and ε_i^-, $i = 1, \ldots, n$, represent two nonnegative variables such that

$$\left| y_i - \sum_{j=0}^k x_{ji}\beta_j \right| = \varepsilon_i^+ + \varepsilon_i^-, \qquad y_i - \sum_{j=0}^k x_{ji}\beta_j = \varepsilon_i^+ - \varepsilon_i^-.$$

Let us consider the following matrix notations

$$\begin{aligned}
\boldsymbol{\varepsilon}^+_{n\times 1} &= [\varepsilon^+_1, \ldots, \varepsilon^+_n]^t, \\
\boldsymbol{\varepsilon}^-_{n\times 1} &= [\varepsilon^-_1, \ldots, \varepsilon^-_n]^t, \\
\mathbf{e}_{(k+1+2n)\times 1} &= [\boldsymbol{\beta}_{1\times(k+1)}\ (\boldsymbol{\varepsilon}^+_{n\times 1})^t\ (\boldsymbol{\varepsilon}^-_{n\times 1})^t]^t, \\
\mathbf{H}_{n\times(k+1+2n)} &= [\mathbf{X}_{n\times(k+1)}\ \mathbf{I}_{n\times n}\ \ -\mathbf{I}_{n\times n}], \\
\mathbf{h}_{(k+1+2n)\times 1} &= [\mathbf{0}_{1\times(k+1)}\ \mathbf{J}_{1\times 2n}]^t,
\end{aligned}$$

where $\mathbf{I}_{n\times n}$ is an identity matrix of order n, $\mathbf{0}_{1\times(k+1)}$ denotes the $(1\times(k+1))$-vector of 0's and $\mathbf{J}_{1\times 2n}$ denotes the $(1\times 2n)$-vector of 1's. Now, the non-linear optimization problem (A) becomes equivalent to the following linear optimization problem

$$\begin{aligned}
\min_{\mathbf{e}_{(k+1+2n)\times 1}} \quad & \mathbf{h}^t_{(k+1+2n)\times 1}\mathbf{e}_{(k+1+2n)\times 1} \\
s.t. \quad & \mathbf{H}_{n\times(k+1+2n)}\mathbf{e}_{(k+1+2n)\times 1} = \mathbf{y}_{n\times 1}, \\
& \boldsymbol{\varepsilon}^+_{n\times 1} \in \mathbb{R}^{+^n},\ \boldsymbol{\varepsilon}^-_{n\times 1} \in \mathbb{R}^{+^n},\ \boldsymbol{\beta}_{1\times(k+1)} \in \mathbb{R}^{k+1}.
\end{aligned}$$

This problem can be solved by the common softwares. In the present article, we used the software MATLAB [30] for numerical studies given in this paper.

The same method may be easily used to solve the optimization problem (B). In this case, we replace $\boldsymbol{\beta}_{1\times(k+1)} \in \mathbb{R}^{k+1}$ with $\boldsymbol{\sigma}_{1\times(k+1)} \in \mathbb{R}^{+^{k+1}}$ in the above optimization problem, which means that all the variables are assumed to be nonnegative.

4 Goodness-of-Fit Criteria

Several goodness-of-fit criteria have been introduced by authors to evaluate the performance of a fuzzy regression model. Here, we employ three criteria which have been proposed for evaluation a fuzzy regression model with fuzzy response data.

I) The relative error of estimation: This criteria, which is introduced by Kim and Bishu [26], is defined as follows

$$E_1(i) = \int \frac{|\widetilde{y}_i(x) - \widehat{\widetilde{y}}_i(x)|}{\int \widetilde{y}_i(x)\,dx}\,dx.$$

The $E_1(i)$ is the ratio of the total difference between the estimated and observed membership values of response variable to the total observed membership values of the response variable. This criteria was also used in [7, 28].

II) The error of estimation: A variation of $E_1(i)$, which is used in [7, 23, 24, 28], is proposed as follows

$$E_2(i) = \int |\widetilde{y}_i(x) - \widehat{\widetilde{y}}_i(x)|\,dx.$$

This index measures the difference between the estimated and observed membership values of response variable.

III) The similarity measure: This index is defined based on the similarity of fuzzy numbers as follows

$$S(i) = \frac{\mathrm{Card}(\widetilde{y}_i \cap \widehat{\widetilde{y}}_i)}{\mathrm{Card}(\widetilde{y}_i \cup \widehat{\widetilde{y}}_i)} = \frac{\int \min\{\widetilde{y}_i(x), \widehat{\widetilde{y}}_i(x)\}\,dx}{\int \max\{\widetilde{y}_i(x), \widehat{\widetilde{y}}_i(x)\}\,dx}.$$

Such a similarity measure have been used for evaluating the performance of a fuzzy regression model in [21, 28, 38, 39].

The ranges of E_1 and E_2 are $[0,\infty)$, while that of S is $[0,1]$. Thus, in order to compare the indices, we modify the ranges of E_1 and E_2 by taking into account $G_1(i) = 1/(1+E_1(i))$ and $G_2(i) = 1/(1+E_2(i))$. In practice, we use $G_1 = \sum_{i=1}^n G_1(i)/n$, $G_2 = \sum_{i=1}^n G_2(i)/n$, and $S = \sum_{i=1}^n S(i)/n$ to evaluate the goodness-of-fit of the models. The model with higher values of G_1, G_2, and S provides a better goodness-of-fit to the data.

Table 1: Comparison between various models in Example 5.1

The model(s) proposed by	S	G_1	G_2
Ferraro et al. [15] $\widehat{\widetilde{y}} = (4.95 + 1.71x, \exp(0.6098 + 0.0742x))_T$	$0.4079_{(9)}$	$0.3620_{(6)}$	$0.5534_{(6)}$
Choi and Buckley [8] $\widehat{\widetilde{y}} = (5.5211 + 1.4958x, 2.0060 + 0.0788x, 2.2369 + 0.0326x)_T$	$0.4889_{(2)}$	$0.4702_{(1)}$	$0.6194_{(2)}$
Modarres et al. [31] $\widehat{\widetilde{y}} = (4.82 + 1.66x, 1.84 + 0.16x)_T$	$0.4670_{(3)}$	$0.4387_{(3)}$	$0.6034_{(3)}$
Nasrabadi and Nasrabadi [34] $\widehat{\widetilde{y}} = (4.6812 + 1.7306x, 2.3221)_T$	$0.4408_{(5)}$	$0.3958_{(5)}$	$0.5765_{(5)}$
Kao and Chyu [24] $\widehat{\widetilde{y}} = (4.926 + 1.718x, 2.32)_T$	$0.4095_{(7)}$	$0.3596_{(8)}$	$0.5511_{(8)}$
Kao and Chyu [23] $\widehat{\widetilde{y}} = (4.95 + 1.71x, 3.01, 1.8)_T$	$0.4663_{(4)}$	$0.4273_{(4)}$	$0.5947_{(4)}$
Tanaka et al. [40] $\widehat{\widetilde{y}} = (3.85 + 2.1x, 3.85)_T$	$0.4313_{(6)}$	$0.2971_{(9)}$	$0.4898_{(9)}$
Xu and Li [45], Kim and Bishu [26], and Diamond [10] $\widehat{\widetilde{y}} = (4.95 + 1.71x, 1.84 + 0.16x)_T$	$0.4087_{(8)}$	$0.3619_{(7)}$	$0.5533_{(7)}$
The new least-absolutes model $\widehat{\widetilde{y}} = (6.444 + 1.3112x, 1.8 + 0.2x)_T$	$0.5008_{(1)}$	$0.4480_{(2)}$	$0.6209_{(1)}$

5 Illustrative Examples and Competitive Studies

In this section four numerical examples, based on some well-known data sets, are investigated to illustrate the efficiency of the proposed fuzzy linear regression model with respect to some other methods.

Example 5.1. *[40] Consider the following crisp input-fuzzy output data given in [40]*

$$(\widetilde{y}; x) = ((8.0, 1.8)_T; 1), ((6.4, 2.2)_T; 2), ((9.5, 2.6)_T; 3), ((13.5, 2.6)_T; 4), ((13.0, 2.4)_T; 5).$$

By applying the new proposed approach described in Section 3, the fuzzy regression model is derived as

$$\widehat{\widetilde{y}} = (6.4440, 1.8)_T \oplus (1.3112, 0.2)_T x.$$

A summary of the results of various models as well as their performances are given in Table 1. From Table 1, one can see that with regard to the criteria S and G_2, the least-absolutes model proposed in this paper performs the best model among all investigated models. Moreover, it is better than the other ones in terms of G_1, except the model proposed by Choi and Buckley [8].

Example 5.2. *[15] In this example we are interested in analyzing the dependence relationship of the Retail Trade Sales of the U.S. in 2002 by kind of business on the number of employees. The Retail Trade Sales has been in the period January 2002 through December 2002 (see the data set in Table 2, given in [15]). The results of fitting the new proposed model and for some common models are presented in Table 3. The results show that, while the index S for some models are close to that of the proposed model, the proposed model has better performance concerning G_1 and G_2. It is noticeable that, the amount of G_1 for the proposed model is three times better than that of the model with the second order (i.e. Modarres et al.'s model [31]).*

Table 2: The retail trade sales and the number of employees of 22 kinds of business in the U.S. in 2002.

Kind of business	Retail trade sales	Number of employees
Automotive parts, acc., and tire stores	4638-5795	453,468
Furniture stores	4054-4685	249,807
Home furnishings stores	2983-5032	285,222
Household appliance stores	1035-1387	69,168
Computer and software stores	1301-1860	73,935
Building mat. and supplies dealers	14508-20727	988,707
Hardware stores	1097-1691	142,881
Beer, wine, and liquor stores	2121-3507	133,035
Pharmacies and drug stores	11964-14741	783,392
Gasoline stations	16763-23122	926,792
Mens clothing stores	532-1120	62,223
Family clothing stores	3596-9391	522,164
Shoe stores	1464-2485	205,067
Jewelry stores	1304-5810	148,752
Sporting goods stores	1748-3404	188,091
Book stores	968-1973	133,484
Discount dept. stores	9226-17001	762,309
Department stores	5310-14057	668,459
Warehouse clubs and superstores	13162-22089	830,845
All other gen. merchandize stores	2376-4435	263,116
Miscellaneous store retailers	7862-10975	792,361
Fuel dealers	1306-3145	98,574

Table 3: Performance of various models in Example 5.2

The model(s) proposed by	S	G_1	G_2
Ferraro et al. [15] $\widehat{\widetilde{y}} = (-672.731 + 0.0181x, \exp\{5.9244 + 0.000002482x\})_T$	$0.2135_{(7)}$	$0.0010_{(3)}$	$0.4386_{(2)}$
Choi and Buckley [8] $\widehat{\widetilde{y}} = (-254.7958 + 0.01737x, 361.8300 + 0.00446x, 0.00385x)_T$	$0.2578_{(1)}$	$0.0010_{(3)}$	$0.4236_{(5)}$
Modarres et al. [31] $\widehat{\widetilde{y}} = (-188.8609 + 0.01605x, 194.6657 + 0.00348x)_T$	$0.2435_{(2)}$	$0.0012_{(2)}$	$0.4305_{(4)}$
Nasrabadi and Nasrabadi [34] $\widehat{\widetilde{y}} = (-1714.8222 + 0.01813x, 1675.4064)_T$	$0.1834_{(8)}$	$0.0005_{(7)}$	$0.3546_{(7)}$
Kao and Chyu [24] $\widehat{\widetilde{y}} = (-1137.961 + 0.019234x, 1570.7727)_T$	$0.2253_{(6)}$	$0.0006_{(6)}$	$0.3807_{(6)}$
Tanaka et al. [40] $\widehat{\widetilde{y}} = (1162.6583 + 0.015314x, 1097.1898 + 0.008552x)_T$	$0.2428_{(4)}$	$0.0005_{(7)}$	$0.2865_{(8)}$
Xu and Li [45], Kim and Bishu [26], and Diamond [10] $\widehat{\widetilde{y}} = (-672.731 + 0.01807x, 185.7144 + 0.00347x)_T$	$0.2431_{(3)}$	$0.0010_{(3)}$	$0.4357_{(3)}$
The new least-absolutes model $\widehat{\widetilde{y}} = (-254.7958 + 0.01737x, 66.4543 + 0.00336x)$	$0.2352_{(5)}$	$0.0032_{(1)}$	$0.4518_{(1)}$

Table 4: Data set in Example 5.3

No.	$\widetilde{y}$	x_1	x_2	x_3
1	$(5.83, 3.56)_T$	2.00	0.00	15.25
2	$(0.85, 0.52)_T$	0.00	5.00	14.13
3	$(13.93, 8.50)_T$	1.13	1.50	14.13
4	$(4.00, 2.44)_T$	2.00	1.25	13.63
5	$(1.65, 1.01)_T$	2.19	3.75	14.75
6	$(1.58, 0.96)_T$	0.25	3.50	13.75
7	$(8.18, 4.99)_T$	0.75	5.25	15.25
8	$(1.85, 1.13)_T$	4.25	2.00	13.50

Example 5.3. *The data set given in Table 4 is used in [26], where the observations of independent variables are crisp and the observations of the dependent variable are presented as symmetric triangular fuzzy numbers. Choi and Buckley [8] claimed that the third observation in the data set is fuzzy outlier and for modeling this data set they proposed a least-absolutes fuzzy regression method.*

The fuzzy regression models based on the various methods are given in Table 5. As can be seen, our proposed model has more mean of similarity measure than the other models and ranks the second when compared with G_1 *and* G_2 *indices.*

Example 5.4. *(An application in soil science [32]) One of the classical problems in soil sciences is the measurement of physical, chemical, and biological soil properties. The problem results from the difficulty, time and cost of direct measurements. Pedomodels, which have become a popular topic in soil science and environmental research, are predictive functions of certain soil properties based on other easily or cheaply measured properties. The common method for fitting pedomodels is to use classical regression analysis, based on the assumption of data crispness. In modeling natural systems such as soil system, however, we may come across the imprecise observations. Here, we consider such a case in which the observations of the response variable are imprecise.*

Based on a study in a part of Silakhor plain (situated in the province of Lorestan, west of Iran), different soil physical and chemical properties were measured using standard procedures. But, due to some impreciseness in experimental environment, the observed data were reported as fuzzy numbers given in Table 6. The data set show soils saturated by water `(SP)` $(\widetilde{y})$*, as symmetric triangular fuzzy observations of the dependent variable, organic matter content* `(OM)` (x_1)*, sand content percentage* `(SAND)` (x_2)*, and silt* `(SILT)` (x_3) *as the crisp observations of the independent variables. Based on such data set, we wish to model the relationship between the response variable* `SP` *and explanatory variables* `OM`*,* `SAND`*, and* `SILT` *by the fuzzy regression model.*

The various models as well as their performances are summarized in Table 7. As can be seen, among the various models, our proposed model has more mean of similarity measures than the other models. The proposed model has also better performances than the other ones, considering the criteria G_1 *and* G_2*.*

6 Concluding Remarks

A least-absolutes fuzzy multiple linear regression model is proposed by using the generalized Hausdorff-metric on the space of fuzzy numbers. Using this model we can deal with multiple linear regression problem with crisp input-fuzzy output observations. The solution of the proposed methodology relies on a non-linear optimization problem, which is also translated to a linear optimization problem, making the computations of the proposed method very simple. The computations of the proposed method are rather simple and can be easily used in practical studies.

Three indices, based on the error of estimation and the similarity measure, are proposed to evaluate the goodness-of-fit of the models. By using such criteria, the efficiency of the proposed model is investigated by four well-known data sets. The results of comparative studies and numerical examples show that the proposed least-absolutes method is able to determine the regression coefficients with a good explanatory power, especially for the case with several input variables and large size sample data set.

Table 5: Performance of various models in Example 5.3

The model(s) proposed by	S	G_1	G_2
Ferraro et al. [15] $\widehat{y} = -16.7957 - 1.0989x_1 - 1.1798x_2 + 1.8559x_3$ $\widehat{l} = \exp(-6.4974 - 0.1362x_1 - 0.2904x_2 + 0.5726x_3)$	$0.1674_{(5)}$	$0.2786_{(4)}$	$0.4076_{(4)}$
Choi and Buckley [8] $\widehat{y} = -2.8273 - 0.3878x_1 - 1.0125x_2 + 0.6185x_3$ $\widehat{l} = 0.2133x_1 + 0.0368x_3$ $\widehat{r} = 0.2111x_3$	$0.3544_{(2)}$	$0.3768_{(1)}$	$0.5038_{(1)}$
Modarres et al. [31] $\widehat{y} = -23.1596 - 1.0758x_1 - 1.6443x_2 + 2.3866x_3$ $\widehat{l} = -10.2245 - 0.6825x_1 - 0.7282x_2 + 1.1353x_3$	$0.1679_{(4)}$	$0.2489_{(5)}$	$0.3731_{(5)}$
Nasrabadi-Nasrabadi [34] $\widehat{y} = -7.9334 - 0.4020x_1 - 1.3553x_2 + 1.2199x_3$ $\widehat{l} = 2.9182$	$0.1014_{(7)}$	$0.1946_{(6)}$	$0.3201_{(6)}$
Kao and Chyu [23] $\widehat{y} = -16.7957 - 1.0989x_1 - 1.1798x_2 + 1.8559x_3$ $\widehat{l} = 0.52$	$0.0615_{(8)}$	$0.3229_{(3)}$	$0.4571_{(3)}$
Tanaka et al. [40] $\widehat{y} = 37.9577 - 2.5418x_1 - 2.3226x_2 - 1.3538x_3$ $\widehat{l} = 9.9581$	$0.1896_{(3)}$	$0.1162_{(7)}$	$0.2412_{(7)}$
Mohammadi and Taheri [32], Xu and Li [45], Kim and Bishu [26], and Diamond [10] $\widehat{y} = -16.7957 - 1.0989x_1 - 1.1798x_2 + 1.8559x_3$ $\widehat{l} = 1.1588x_3$	$0.1440_{(6)}$	$0.0653_{(8)}$	$0.1563_{(8)}$
The new least-absolutes model $\widehat{y} = -2.8273 - 0.3878x_1 - 1.0125x_2 + 0.6185x_3$ $\widehat{l} = 0.1790x_3$	$0.3632_{(1)}$	$0.3534_{(2)}$	$0.4776_{(2)}$

Table 6: Soil data in Example 5.4

No.	OM	SAND	SILT	SP	No.	OM	SAND	SILT	SP
1	0.88	35	45	$(38, 3.8)_T$	14	2.33	31	42	$(52, 5.2)_T$
2	1.13	37	42	$(41, 4.1)_T$	15	1.71	17	50	$(52, 5.2)_T$
3	1.31	27	43	$(47.5, 4.75)_T$	16	1.14	14	53	$(49, 4.9)_T$
4	1.98	29	41	$(51, 5.1)_T$	17	0.99	19	44	$(49, 4.9)_T$
5	1.02	38	39	$(35, 3.5)_T$	18	1.14	28	43	$(44, 4.4)_T$
6	1.29	32	39	$(43, 4.3)_T$	19	1.46	26	44	$(49, 4.9)_T$
7	1.52	29	37	$(54, 5.4)_T$	20	1.18	32	42	$(50.3, 5.03)_T$
8	1.33	18	45	$(52, 5.2)_T$	21	1.38	10	49	$(52, 5.2)_T$
9	1.17	40	38	$(45, 4.5)_T$	22	0.84	38	43	$(42, 4.2)_T$
10	2.00	28	46	$(50, 5.0)_T$	23	1.48	49	35	$(40, 4.0)_T$
11	1.68	13	40	$(58.6, 5.86)_T$	24	1.08	42	44	$(37, 3.7)_T$
12	2.15	19	41	$(62, 6.2)_T$	25	0.36	79	14	$(21.2, 2.12)_T$
13	3.52	31	41	$(60, 6.0)_T$					

Table 7: Performance of various models in Example 5.4

The model(s) proposed by	S	G_1	G_2
Ferraro et al. [15] $\widehat{y} = 70.5873 + 6.9253x_1 - 0.5628x_2 - 0.3958x_3$ $\widehat{l} = \exp(1.7897 + 0.1538x_1 - 0.0120x_2 - 0.0028x_3)$	$0.5830_{(3)}$	$0.4043_{(3)}$	$0.6826_{(3)}$
Choi and Buckley [8] $\widehat{y} = 72.0466 + 6.4557x_1 - 0.5523x_2 - 0.4278x_3$ $\widehat{l} = 3.8955 + 0.5599x_1$ $\widehat{r} = 0.3901 + 0.8664x_1 + 0.0665x_3$	$0.6002_{(2)}$	$0.4360_{(2)}$	$0.6990_{(2)}$
Modarres et al. [31] $\widehat{y} = 69.3354 + 9.5098x_1 - 0.6193x_2 - 0.3972x_3$ $\widehat{l} = 7.0587 + 0.6925x_1 - 0.0563x_2 - 0.0396x_3$	$0.4412_{(5)}$	$0.2814_{(5)}$	$0.5838_{(5)}$
Nasrabadi-Nasrabadi [34] $\widehat{y} = 74.0638 + 9.0493x_1 - 0.6855x_2 - 0.4406x_3$ $\widehat{l} = 4.7053$	$0.3663_{(7)}$	$0.2161_{(6)}$	$0.5283_{(6)}$
Tanaka et al. [40] $\widehat{y} = 61.3916 + 7.6673x_1 - 0.5015x_2 - 0.2188x_3$ $\widehat{l} = 0.0870x_2 + 0.1749x_3$	$0.4030_{(6)}$	$0.1440_{(7)}$	$0.4321_{(7)}$
Mohammadi and Taheri [32], Xu and Li [45], Kim and Bishu [26], and Diamond [10] $\widehat{y} = 70.5873 + 6.9253x_1 - 0.5628x_2 - 0.3958x_3$ $\widehat{l} = 3.19688 + 1.0407x_1$	$0.5727_{(4)}$	$0.3707_{(4)}$	$0.6696_{(4)}$
The new least-absolutes model $\widehat{y} = 72.0466 + 6.4557x_1 - 0.5523x_2 - 0.4278x_3$ $\widehat{l} = 1.0240 + 0.7163x_1 + 0.0599x_3$	$0.6014_{(1)}$	$0.4462_{(1)}$	$0.7031_{(1)}$

Further research would be centered on modeling fuzzy input-output data based on the proposed approach. Moreover, the sensitivity analysis with respect to outliers is a potential subject for further research.

Appendix: Fuzzy Sets and Fuzzy Arithmetic

A fuzzy set $\widetilde{A}$ on the universal set $\mathbb{X}$ is described by its membership function $\widetilde{A}(x) : \mathbb{X} \to [0,1]$. In this paper we assume that $\mathbb{X} = \mathbb{R}$, the set of real numbers. The crisp set $A_\alpha = \{x \in \mathbb{R} : \widetilde{A}(x) \geq \alpha\}$, $\alpha \in (0,1]$, is called the α-cut of $\widetilde{A}$, and for $\alpha = 0$ we assume $A_0 = \text{cl}\{x \in \mathbb{R} : \widetilde{A}(x) > 0\}$, where *cl* is the closure operator.

A specific class of fuzzy sets on $\mathbb{R}$, which is rich and flexible enough to cover most of the applications, is the so-called LR-fuzzy numbers $\widetilde{N} = (n,l,r)_{LR}$ with central value $n \in \mathbb{R}$, left and right spreads $l \in \mathbb{R}^+$, $r \in \mathbb{R}^+$, decreasing left and right shape functions $L : \mathbb{R}^+ \to [0,1]$, $R : \mathbb{R}^+ \to [0,1]$, with $L(0) = R(0) = 1$. Typically, the LR-fuzzy number $\widetilde{N}$ has the following membership function [47]

$$\widetilde{N}(x) = \begin{cases} L(\frac{n-x}{l}) & \text{if} \quad x \leq n, \\ R(\frac{x-n}{r}) & \text{if} \quad x \geq n. \end{cases}$$

We can easily obtain the α-cut of $\widetilde{N}$ as follows

$$N_\alpha = [n - L^{-1}(\alpha)l, n + R^{-1}(\alpha)r], \quad \alpha \in [0,1].$$

An LR-fuzzy number $\widetilde{N} = (n,l,r)_{LR}$ with $L = R$ and $l = r = \lambda$ is called symmetric and is abbreviated by $\widetilde{N} = (n,\lambda)_L$. In practice, it is usually preferred to use simple shapes for functions L and R such as triangular, i.e. $L(x) = R(x) = \max\{1-x, 0\}$, or normal, i.e. $L(x) = R(x) = \exp\{-x^2\}$.

For the algebraic operations of LR-fuzzy numbers, we have the following result on the basis of Zadeh's extension principle (for more details, see [47]).

Let $\widetilde{M} = (m, l_m, r_m)_{LR}$ and $\widetilde{N} = (n, l_n, r_n)_{LR}$ be two LR-fuzzy numbers and λ be a real number. Then

$$\lambda \otimes \widetilde{M} = \begin{cases} (\lambda m, \lambda l_m, \lambda r_m)_{LR} & \text{if} \quad \lambda > 0 \\ \mathcal{I}_{\{0\}} & \text{if} \quad \lambda = 0 \\ (\lambda m, |\lambda| r_m, |\lambda| l_m)_{RL} & \text{if} \quad \lambda < 0, \end{cases}$$

$$\widetilde{M} \oplus \widetilde{N} = (m+n, l_m + l_n, r_m + r_n)_{LR}.$$

where $\mathcal{I}_{\{0\}}$ stands for the indicator function of the crisp zero.

References

[1] Arabpour, A.R., and M. Tata, Estimating the parameters of a fuzzy linear regression model, *Iranian Journal of Fuzzy Systems*, vol.5, pp.1–19, 2008.

[2] Bertoluzza, C., N. Corral, and A. Salas, On a new class of distances between fuzzy numbers, *Mathware Soft Computing*, vol.2, pp.71–84, 1995.

[3] Bisserier, A., R. Boukezzoula, and S. Galichet, Linear fuzzy regression using trapezoidal fuzzy intervals, *Journal of Uncertain Systems*, vol.4, pp.59–72, 2010.

[4] Celminš, A., Multidimensional least-squares fitting of fuzzy models, *Mathematical Modelling*, vol.9, pp.669–690, 1987.

[5] Chachi, J., S.M. Taheri, and H.R. Pazhand, An interval-based approach to fuzzy regression for fuzzy input-output data, *Proceedings of the IEEE International Conference on Fuzzy Systems*, Taipei, Taiwan, pp.2859–2863, 2011.

[6] Chang, P.T., and S. Lee, Fuzzy linear regression with spreads unrestricted in sign, *Computers and Mathematics with Applications*, vol.28, pp.61–70, 1994.

[7] Chen, S.P., and J.F. Dang, A variable spread fuzzy linear regression model with higher explanatory power and forecasting accuracy, *Information Sciences*, vol.178, pp.3973–3988, 2008.

[8] Choi, S.H., and J.J. Buckley, Fuzzy regression using least absolute deviation estimators, *Soft Computing*, vol.12, pp.257–263, 2008.

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

[10] Diamond, P., Fuzzy least squares, *Information Sciences*, vol.46, pp.141–157, 1988.

[11] Diamond, P., and P. Kloeden, *Metric Spaces of Fuzzy Sets*, World Scientific, Singapore, 1994.

[12] D'Urso, P., Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data, *Computational Statistics and Data Analysis*, vol.42, pp.47–72, 2003.

[13] D'Urso, P., R. Massari, and A. Santoro, Robust fuzzy regression analysis, *Information Sciences*, vol.181, pp.4154–4174, 2011.

[14] Ferraro, M.B., A. Colubi, G. González-Rodríguez, and R. Coppi, A determination coefficient for a linear regression model with imprecise response, *Environmetrics*, vol.22, pp.516–529, 2011.

[15] Ferraro, M.B., R. Coppi, G. González-Rodríguez, and A. Colubi, A linear regression model for imprecise response, *International Journal of Approximate Reasoning*, vol.51, pp.759–770, 2010.

[16] Grzegorzewski, P., Distances between intuitionistic fuzzy sets and/or interval-valued fuzzy sets based on the Hausdorff metric, *Fuzzy Sets and Systems*, vol.148, pp.319–328, 2004.

[17] Guo, R., Y. Cui, and D. Guo, Uncertainty linear regression models, *Journal of Uncertain Systems*, vol.5, pp.286–304, 2011.

[18] Guo, R. , D. Guo, and C. Thiart, Scalar fuzzy regression models, *Journal of Uncertain Systems*, vol.2, pp.177–186, 2008.

[19] Hassanpour, H., H.R. Maleki, and M.A. Yaghoobi, Fuzzy linear regression model with crisp coefficients: a programming approach, *Iranian Journal of Fuzzy Systems*, vol.7, pp.19–39, 2010.

[20] Hassanpour, H., H.R. Maleki, and M.A. Yaghoobi, A goal programming approach to fuzzy linear regression with fuzzy input-output data, *Soft Computing*, vol.15, pp.1569–1580, 2011.

[21] Hojati, M., C.R. Bector, and K. Smimou, A simple method for computation of fuzzy linear regression, *European Journal of Operational Research*, vol.166, pp.172–184, 2005.

[22] Hung, W.L., and M.S. Yang, Similarity measures of intuitionistic fuzzy sets based on Hausdorff distance, *Pattern Recognition Letters*, vol.25, pp.1603–1611, 2004.

[23] Kao, C., and C.L. Chyu, A fuzzy linear regression model with better explanatory power, *Fuzzy Sets and Systems*, vol.126, pp.401–409, 2002.

[24] Kao, C., and C.L. Chyu, Least-squares estimates in fuzzy regression analysis, *European Journal of Operational Research*, vol.148, pp.426–435, 2003.

[25] Kelkinnama, M., and S.M. Taheri, Fuzzy least-absolutes regression using shape preserving operations, *Information Sciences*, vol.214, pp.105–120, 2012.

[26] Kim, B., and R.R. Bishu, Evaluation of fuzzy linear regression models by comparison membership function, *Fuzzy Sets and Systems*, vol.100, pp.343–352, 1998.

[27] Kim, K.J., D.H. Kim, and S.H. Choi, Least absolute deviation estimator in fuzzy regression, *Journal of Applied Mathematics and Computing*, vol.18, pp.649–656, 2005.

[28] Lu, J., and R. Wang, An enhanced fuzzy linear regression model with more flexible spreads, *Fuzzy Sets and Systems*, vol.160, pp.2505–2523, 2009.

[29] Mashinchi, M.H., M.A. Orgun, M. Mashinchi, and W. Pedrycz, A tabu-harmony search-based approach to fuzzy linear regression, *IEEE Transactions on Fuzzy Systems*, vol.19, pp.432-448, 2011.

[30] MATLAB, *The Language of Technical Computing*, The MathWorks Inc., MA, 2009.

[31] Modarres, M., E. Nasrabadi, and M.M. Nasrabadi, Fuzzy linear regression models with least square errors, *Applied Mathematics and Computation*, vol.163, pp.977–989, 2005.

[32] Mohammadi, J., and S.M. Taheri, Pedomodels fitting with fuzzy least squares regression, *Iranian Journal of Fuzzy Systems*, vol.1, pp.45–62, 2004.

[33] Nadler, S.B., *Hyperspaces of Sets*, Marcel Dekker, New York, 1978.

[34] Nasrabadi, M.M., and E. Nasrabadi, A mathematical-programming approach to fuzzy linear regression analysis, *Applied Mathematics and Computation*, vol.155, pp.873–881, 2004.

[35] Näther, W., Regression with fuzzy random data, *Computational Statistics and Data Analysis*, vol.51, pp.235–252, 2006.

[36] Pourahmad, S., S.M.T. Ayatollahi, and S.M. Taheri, Fuzzy logistic regression: a new possibilistic model and its application in clinical vague status, *Iranian Journal of Fuzzy Systems*, vol.8, pp.1–17, 2011.

[37] Pourahmad, S., S.M.T. Ayatollahi, S.M. Taheri, and Z. Habib-Agahi, Fuzzy logistic regression based on the least squares approach with application in clinical studies, *Computers and Mathematics with Applications*, vol.62, pp.3353–3365, 2011.

[38] Taheri, S.M., and M. Kelkinnama, Fuzzy least absolutes regression, *Proceedings of the 4th International IEEE Conference on Intelligent Systems*, Varna (Bulgaria), vol.11, pp.55–58, 2008.

[39] Taheri, S.M., and M. Kelkinnama, Fuzzy linear regression based on least absolute deviations, *Irannian Journal of Fuzzy Systems*, vol.9, pp.121–140, 2012.

[40] Tanaka, H., I. Hayashi, and J. Watada, Possibilistic linear regression analysis for fuzzy data, *European Journal of Operational Research*, vol.40, pp.389–396, 1989.

[41] Tanaka, H., S. Uegima, and K. Asai, Linear regression analysis with fuzzy model, *IEEE Transactions on Systems, Man, and Cybernetics*, vol.12, pp.903–907, 1982.

[42] Trutschnig, W., G. González-Rodríguez, A. Colubi, and M.A. Gil, A new family of metrics for compact, convex (fuzzy) sets based on a generalized concept of mid and spread, *Information Sciences*, vol.179, pp.3964–3972, 2009.

[43] Tutmez, B., Spatial dependence-based fuzzy regression clustering, *Applied Soft Computing*, vol.12, pp.1–13, 2012.

[44] Wu, H.C., The construction of fuzzy least squares estimators in fuzzy linear regression models, *Expert Systems with Applications*, vol.38, pp.13632–13640, 2011.

[45] Xu, R., and C. Li, Multidimensional least-squares fitting with a fuzzy model, *Fuzzy Sets and Systems*, vol.119, pp.215–223, 2001.

[46] Yu, J.-R., and C.-W. Lee, Piecewise regression for fuzzy input-output data with automatic change-point detection by quadratic programming, *Applied Soft Computing*, vol.10, pp.111–118, 2010.

[47] Zimmermann, H.J., *Fuzzy Set Theory and Its Applications*, 4th Edition, Kluwer Nihoff, Boston, 2001.