Article

Boscovich Fuzzy Regression Line

Pavel Škrabánek [1,*] , **Jaroslav Marek** [2] and **Alena Pozdílková** [2]

1. Institute of Automation and Computer Science, Brno University of Technology, Technická 2896/2, 616 69 Brno, Czech Republic
2. Department of Mathematics and Physics, University of Pardubice, Studentská 95, 532 10 Pardubice, Czech Republic; jaroslav.marek@upce.cz (J.M.); alena.pozdilkova@upce.cz (A.P.)
* Correspondence: pavel.skrabanek@vut.cz

Abstract: We introduce a new fuzzy linear regression method. The method is capable of approximating fuzzy relationships between an independent and a dependent variable. The independent and dependent variables are expected to be a real value and triangular fuzzy numbers, respectively. We demonstrate on twenty datasets that the method is reliable, and it is less sensitive to outliers, compare with possibilistic-based fuzzy regression methods. Unlike other commonly used fuzzy regression methods, the presented method is simple for implementation and it has linear time-complexity. The method guarantees non-negativity of model parameter spreads.

Keywords: fuzzy linear regression; non-symmetric triangular fuzzy number; least absolute value; Boscovich regression line; outlier

1. Introduction

A regression model approximates the functional relationship between a dependent y and independent variables x. Parameters of a regression model are estimated using a set of observations of x and y. The model with the estimated parameters can be used to predict a dependent variable value for a specific combination of the independent variables. In practice, statistical regression models are most often used for this purpose, but their usage is limited by an assumption that any deviation of a prediction from a corresponding observation is due to a random error.

In many practical applications, the deviations are a result of imprecise observations, an indefiniteness of the system structure and parameters [1,2], or the vagueness of human perception of the model (in contrast with the statistical regression where the errors are associated with observations) [3]. There are also cases where the observations are inherently fuzzy, e.g., if the observations are described by linguistic terms [4–6]. In such cases, the deviations are not due to randomness, but they are due to fuzziness and fuzzy regression should be used. The fuzzy regression can be also used when statistical distributional assumptions cannot be justified, or if the representation of the regression model is poor [3]. The fuzzy regression is an efficient alternative to statistical regression whenever a dataset is insufficient to support statistical regression analysis [7].

The fuzzy regression approximates the relationship between the dependent y and the independent variables x using a fuzzy regression model. Once the underlying regression relationship is known, an appropriate parametric fuzzy regression model (e.g., linear [8], polynomial [9], and logistic [10]) can be used for the approximation of the relationship. If the relationship is unknown, it is possible to utilize a nonparametric fuzzy regression model. Such as a model based on k-nearest neighbour smoothing [11], kernel smoothing [11], local linear smoothing [12], and adaptive neuro-fuzzy inference systems [13].

In fuzzy regression analysis, the relationship is typically approximated as a linear dependence by a model with all fuzzy parameters [2,14–24], but other fuzzy linear models are used as well [8]. For example, the model with all fuzzy parameters can be extended

with a fuzzy error term [22], or some of the parameters can be real value numbers [8]. Models with all fuzzy parameters are known for a dependence of model prediction spreads on absolute values of x [8,23]. In the case that all parameters, except a y-intercept or the fuzzy error term, are real value numbers, the model prediction spreads are constant in the whole range of x values [8].

For estimation of unknown parameters of a fuzzy model, possibilistic and statistics-based approaches are frequently used. Possibilistic fuzzy estimators minimize a total spread of the fuzzy model predictions subject to constraints that arise from observations [1,3,15,16,19,21,25,26]. The statistics-based solutions adopt concepts that are used in statistics such as the ordinary least squares method [17,18,20,27–33], the least absolute deviations method [22–24,34–37], and adaptive smoothing methods [11,12]. Some fuzzy estimators combine the possibilistic and the statistics-based approaches [38].

Several limitations were observed for some fuzzy estimators. As pointed out by many authors [22,24,39–41], numerous fuzzy estimators are sensitive to outliers. This criticism resulted in a number of fuzzy estimators that are more or less resilient to outliers [3,11,12,24,26,41–44]. A serious issue of some estimators is the fact that they do not guarantee non-negativity of spreads [35,41,45,46]. A certain disadvantage is also the computational complexity of the estimators. Depending on a used fuzzy estimator, matrix operations [18], linear programming [1,3,15,16,19,21,24,44], quadratic programming [26,27,32,38], or a general constrained optimization problem [25,29–31,33] must be solved to obtain estimates of the unknown parameters. Some methods employ customized iterative optimization algorithms [20,41].

Herein, we propose a new fuzzy estimator intended for a simple fuzzy regression model with real value independent and fuzzy dependent variables. We based the estimator on the Boscovich regression line [47,48], hence we named it Boscovich fuzzy regression line. The original Boscovich regression line was a pioneering regression method based on minimization criterion, i.e., it is a predecessor of today's statistical regression methods. The method was characterized by extremely low computing demands. The presented Boscovich fuzzy regression line inherited this property. Moreover, the used fuzzy linear model guaranties non-negativity of the model parameter spreads by its nature, and spreads of model predictions are not influenced by independent variable values. Even with its simplicity, the presented fuzzy estimator provides regression models in a quality comparable to other fuzzy regression methods.

2. Materials and Methods

The Boscovich regression method [48] was designed for the simple theoretical linear regression model

$$y = \beta_0 + \beta_1 x, \tag{1}$$

where β_0 and β_1 are unknown parameters of the model. The estimation of β_0, β_1 is based on n imprecise observations (x_i, y_i), where $i \in I$, and $I = \{1, 2, \ldots, n\}$. The estimator introduced by Boscovich was based on two constraints [49]:

(1) the sum of the positive and negative residuals (in the sense of y axis) shall be equal,
(2) the sum of the best absolute values of all the residuals shall be as small as possible.

The constraint (1) implies that the regression line should pass through the centroid $(\overline{x}, \overline{y})$ formed by the observations, i.e.,

$$\overline{y} = \hat{\beta}_0^* + \hat{\beta}_1^* \overline{x}, \tag{2}$$

where $\hat{\beta}_0^*, \hat{\beta}_1^*$ are the best estimates of the unknown regression coefficients β_0, β_1 with respect to the constraints (1) and (2). As follows from the constraint (1), the coordinates of the centroid can be expressed as

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i, \qquad \overline{y} = \frac{1}{n} \sum_{i=1}^{n} y_i. \tag{3}$$

Thus, the best estimate of the unknown regression coefficient β_0 can be expressed on the base of (2) as

$$\hat{\beta}_0^* = \bar{y} - \hat{\beta}_1^* \bar{x}. \tag{4}$$

Naturally, only one point, $(\bar{x}, \bar{y})$, is insufficient to form any line. Therefore, at least one other point is needed. Boscovich suggested using one of the observations as the second point of the regression line. Whereas n observations are available, n regression lines can be constructed in such a way. The k-th regression line is clearly determined by

$$L_k : y_k = \bar{y} + \hat{\beta}_1^{(k)} \Delta x_k, \tag{5}$$

where $\hat{\beta}_1^{(k)}$ is the k-th estimate of β_1 based on the k-th observation (x_k, y_k), and $\Delta x_k = x_k - \bar{x}$. On the basis of (5), the k-th estimate of β_1 can be expressed as

$$\hat{\beta}_1^{(k)} = \left(\Delta x_k^{-1} \right) (y_k - \bar{y}). \tag{6}$$

Altogether, n estimates of β_1 are obtained in such a way. The selection of the best one can be carried out using an evaluation function J based on the constraint (2), i.e., for the k-th regression line, it holds that

$$J^{(k)} = \sum_{i=1}^{n} \left| y_i - \hat{y}_i^{(k)} \right| = \sum_{i=1}^{n} \left| y_i - \bar{y} - \hat{\beta}_1^{(k)} \Delta x_i \right|, \tag{7}$$

where $\hat{y}_i^{(k)}$ is the i-th prediction of the dependent variable y using the k-th regression model (5).

Boscovich expressed a premise that one of the n regression lines is the best approximation of the model (1). Considering constraint (2), the best estimate of the unknown regression coefficient β_1 is given by

$$\hat{\beta}_1^* = \underset{\hat{\beta}_1^{(k)}, \forall k \in I}{\arg \min} \sum_{i=1}^{n} \left| y_i - \bar{y} - \hat{\beta}_1^{(k)} \Delta x_i \right|. \tag{8}$$

3. Fuzzy Set Preliminaries

Definition 1. *A fuzzy subset $\tilde{A}$ defined on $\mathbb{R}$ with membership function $\mu_{\tilde{A}} : \mathbb{R} \to [0, 1]$ is called a fuzzy number if*

(a) $\mu_{\tilde{A}}(x)$ *is normal, i.e., $\exists x_0 \in \mathbb{R}$ with $\mu_{\tilde{A}}(x_0) = 1$,*
(b) $\mu_{\tilde{A}}(x)$ *is fuzzy convex,*
(c) $\mu_{\tilde{A}}(x)$ *is upper semi-continuous on $\mathbb{R}$,*
(d) $\mu_{\tilde{A}}(x)$ *is compactly supported, i.e., $cl\{x \in \mathbb{R}; \mu_{\tilde{A}}(x) > 0\}$ is compact, where $cl(A)$ denotes the closure of the set A.*

Definition 2. *Let L and R be continuous decreasing functions $L, R : [0, +\infty) \to [0, 1]$ fulfilling $L(0) = R(0) = 1$ and $L(1) = R(1) = 0$, invertible on $[0, 1]$. Moreover, let $m_{\tilde{A}}^L, m_{\tilde{A}}^R \in \mathbb{R}$, where $m_{\tilde{A}}^L \leq m_{\tilde{A}}^R$, and $\alpha_{\tilde{A}}, \beta_{\tilde{A}} \in \mathbb{R}^+$; then a fuzzy set $\tilde{A}$ is said to be an L-R type fuzzy number if its membership function is*

$$\mu_{\tilde{A}}(x) = \begin{cases} L\left(\frac{m_{\tilde{A}}^L - x}{\alpha_{\tilde{A}}} \right), & \text{for } m_{\tilde{A}}^L - \alpha_{\tilde{A}} < x < m_{\tilde{A}}^L, \\ 1, & \text{for } m_{\tilde{A}}^L \leq x \leq m_{\tilde{A}}^R, \\ R\left(\frac{x - m_{\tilde{A}}^R}{\beta_{\tilde{A}}} \right), & \text{for } m_{\tilde{A}}^R < x < m_{\tilde{A}}^R + \beta_{\tilde{A}}, \\ 0, & \text{otherwise,} \end{cases} \tag{9}$$

where $m_{\tilde{A}}^{L}$ and $m_{\tilde{A}}^{R}$ are the left and the right points of the core and $\alpha_{\tilde{A}}$ and $\beta_{\tilde{A}}$ are the left and the right spread of the fuzzy number $\tilde{A}$. Symbolically,

$$\tilde{A} = \left(m_{\tilde{A}}^{L}, m_{\tilde{A}}^{R}, \alpha_{\tilde{A}}, \beta_{\tilde{A}} \right)_{LR}. \tag{10}$$

In the case that the functions L and R are linear, and $m_{\tilde{A}}^{L} = m_{\tilde{A}}^{R} = m_{\tilde{A}}$, then the L-R fuzzy number is called a triangular fuzzy number and its membership function is

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{m_{\tilde{A}} - x}{\alpha_{\tilde{A}}}, & m_{\tilde{A}} - \alpha_{\tilde{A}} < x \le m_{\tilde{A}}, \\ \frac{x - m_{\tilde{A}}}{\beta_{\tilde{A}}}, & m_{\tilde{A}} < x < m_{\tilde{A}} + \beta_{\tilde{A}}, \\ 0, & otherwise, \end{cases} \tag{11}$$

where $m_{\tilde{A}}$ is the mean of the triangular fuzzy number $\tilde{A}$. The triangular fuzzy number $\tilde{A}$ is sufficiently represented by the triplet

$$\tilde{A} = (m_{\tilde{A}}, \alpha_{\tilde{A}}, \beta_{\tilde{A}})_{T}. \tag{12}$$

Let the set of all triangular fuzzy numbers be denoted as $\mathbb{R}_{T}$, where T is a symbol for a triangular fuzzy number.

Definition 3. *Let* $\tilde{A} = (m_{\tilde{A}}, \alpha_{\tilde{A}}, \beta_{\tilde{A}})_{T}$ *be the triangular fuzzy number where* $\tilde{A} \in \mathbb{R}_{T}$ *and* λ *is a real number. Then their scalar multiplication can be, on the basis of the extension principle, defined as*

$$\lambda \cdot \tilde{A} = \begin{cases} (\lambda m_{\tilde{A}}, \lambda \alpha_{\tilde{A}}, \lambda \beta_{\tilde{A}})_{T} & \lambda \ge 0, \\ (\lambda m_{\tilde{A}}, -\lambda \beta_{\tilde{A}}, -\lambda \alpha_{\tilde{A}})_{T} & \lambda < 0. \end{cases} \tag{13}$$

For any $\lambda, \nu \in \mathbb{R}$ *and any* $\tilde{A} \in \mathbb{R}_{T}$, *it holds that*

$$(\lambda \nu) \cdot \tilde{A} = \lambda \cdot (\nu \cdot \tilde{A}). \tag{14}$$

Definition 4. *The opposite of a triangular fuzzy number* $\tilde{A} = (m_{\tilde{A}}, \alpha_{\tilde{A}}, \beta_{\tilde{A}})_{T}$, *where* $\tilde{A} \in \mathbb{R}_{T}$, *is the triangular fuzzy number*

$$-\tilde{A} = -1 \cdot \tilde{A} = (-m_{\tilde{A}}, \beta_{\tilde{A}}, \alpha_{\tilde{A}})_{T}. \tag{15}$$

The fuzzy numbers $\tilde{A}$ *and* $-\tilde{A}$ *are symmetrical with respect to the axis* $x = 0$.

Definition 5. *For the triangular fuzzy numbers* $\tilde{A} = (m_{\tilde{A}}, \alpha_{\tilde{A}}, \beta_{\tilde{A}})_{T}$, $\tilde{B} = (m_{\tilde{B}}, \alpha_{\tilde{B}}, \beta_{\tilde{B}})_{T}$, *and* $\tilde{C} = (m_{\tilde{C}}, \alpha_{\tilde{C}}, \beta_{\tilde{C}})_{T}$, *where* $\tilde{A}, \tilde{B}, \tilde{C} \in \mathbb{R}_{T}$, *the sum of* $\tilde{A}$ *and* $\tilde{B}$ *can be, on the basis of the extension principle, defined as*

$$\tilde{A} \oplus \tilde{B} = (m_{\tilde{A}} + m_{\tilde{B}}, \alpha_{\tilde{A}} + \alpha_{\tilde{B}}, \beta_{\tilde{A}} + \beta_{\tilde{B}})_{T}. \tag{16}$$

The operation for $\tilde{A}, \tilde{B}, \tilde{C} \in \mathbb{R}_{T}$ *is:*

(a) *commutative, i.e.,* $\tilde{A} \oplus \tilde{B} = \tilde{B} \oplus \tilde{A}$,
(b) *associative, i.e.,* $\tilde{A} \oplus (\tilde{B} \oplus \tilde{C}) = (\tilde{A} \oplus \tilde{B}) \oplus \tilde{C}$.

Furthermore, it holds:

(a) *for any* $\lambda, \nu \in \mathbb{R}$ *with* $\lambda \nu \ge 0$ *and any* $\tilde{A} \in \mathbb{R}_{T}$ *that*

$$(\lambda + \nu) \cdot \tilde{A} = \lambda \cdot \tilde{A} \oplus \nu \cdot \tilde{A}, \tag{17}$$

(b) *for any* $\lambda \in \mathbb{R}$ *and any* $\tilde{A}, \tilde{B} \in \mathbb{R}_{T}$ *that*

$$\lambda \cdot (\tilde{A} \oplus \tilde{B}) = \lambda \cdot \tilde{A} \oplus \lambda \cdot \tilde{B}, \tag{18}$$

(c)   for any $\tilde{A} \in \mathbb{R}_T$ and $\lambda \in \mathbb{R}$ that

$$\tilde{A} - \lambda = (m_{\tilde{A}} - \lambda, \alpha_{\tilde{A}}, \beta_{\tilde{A}})_T, \tag{19}$$

and

$$\tilde{A} + \lambda = (m_{\tilde{A}} + \lambda, \alpha_{\tilde{A}}, \beta_{\tilde{A}})_T. \tag{20}$$

Definition 6. *The average value of n triangular fuzzy numbers $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_n \in \mathbb{R}_T$ is*

$$\tilde{\bar{A}} = \frac{1}{n} \sum_{i=1}^{n} \tilde{A}_i = \frac{1}{n} \cdot \left(\tilde{A}_1 \oplus \tilde{A}_2 \oplus \cdots \oplus \tilde{A}_n \right), \tag{21}$$

where $\tilde{A}_i = \left(m_{\tilde{A}_i}, \alpha_{\tilde{A}_i}, \beta_{\tilde{A}_i} \right)_T$ for $\forall i \in [1, n]$, $\tilde{\bar{A}} = \left(m_{\tilde{\bar{A}}}, \alpha_{\tilde{\bar{A}}}, \beta_{\tilde{\bar{A}}} \right)_T \in \mathbb{R}_T$.

4. Boscovich Fuzzy Regression Line

The presented fuzzy regression method is intended for datasets where observations of the independent variable x are real value numbers and observations of the dependent variable y are triangular fuzzy numbers $\tilde{Y}$, i.e., a set of n observations is given as

$$O = \left\{ (x_i, \tilde{Y}_i) \,\middle|\, \forall i \in I \right\}, \tag{22}$$

where x_i is the i-th real value observation of the independent variable x, $\tilde{Y}_i$ is the i-th fuzzy observation of the dependent variable y, $\tilde{Y}_i = (y_i, \underline{v}_i, \overline{v}_i)_T$, y_i is the mean, $\underline{v}_i$ is the left, and $\overline{v}_i$ is the right spread of $\tilde{Y}_i$.

For the approximation of a linear dependence between x and $\tilde{Y}$, we use a simple fuzzy linear model

$$\tilde{Y} = \tilde{A}_0 + a_1 x, \tag{23}$$

where $\tilde{A}_0$ and a_1 are unknown model parameters, $\tilde{A}_0 \in \mathbb{R}_T$, $a_1 \in \mathbb{R}$, $\tilde{A}_0 = (a_0, \underline{\alpha}_0, \overline{\alpha}_0)_T$, a_0 is the mean, $\underline{\alpha}_0$ is the left, and $\overline{\alpha}_0$ is the right spread of $\tilde{A}_0$. As the y-intercept $\tilde{A}_0$ is the only fuzzy parameter of the model (23), the fuzziness of model predictions is independent of model input x [8].

Following the Boscovich idea (see Section 2, constraint (1)), the best estimate of the fuzzy regression line (23) shall pass through a centroid formed by the observations O, i.e.,

$$\tilde{\bar{Y}} = \hat{\tilde{A}}_0^* + \hat{a}_1^* \overline{x}, \tag{24}$$

where $\hat{\tilde{A}}_0^*, \hat{a}_1^*$ are the best estimates of the unknown regression coefficients $\tilde{A}_0$ and a_1 according to the constraints (1) and (2); $\tilde{\bar{Y}}$ is the y coordinate of the centroid, and $\overline{x}$ is its x coordinate. The coordinates of the centroid formed by the observations (22) are given as

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i, \qquad \tilde{\bar{Y}} = \frac{1}{n} \sum_{i=1}^{n} \tilde{Y}_i, \tag{25}$$

therefore $\overline{x} \in \mathbb{R}, \tilde{\bar{Y}} \in \mathbb{R}_T$.

Let us express the estimate of the unknown fuzzy regression coefficient $\tilde{A}_0$, using the Formula (24), as

$$\hat{\tilde{A}}_0^* = \tilde{\bar{Y}} - \hat{a}_1^* \overline{x}. \tag{26}$$

As in the case of the Boscovich regression line, n fuzzy regression lines can be constructed using the observations (22). The k-th fuzzy regression line, based on the k-th observation $(x_k, \tilde{Y}_k)$ is given as

$$\tilde{L}_k : \tilde{Y}_k = \tilde{\bar{Y}} + \hat{a}_1^{(k)} \Delta x_k, \tag{27}$$

where $\hat{a}_1^{(k)}$ is the slope of the k-th fuzzy regression line (i.e., the k-th estimate of a_1), and $\Delta x_k = x_k - \overline{x}$.

The model (27) uses the trick used in the ordinary least squares method. Specifically, we relate the explanatory variable to the centre of gravity using the relation $\Delta x_k = x_k - \overline{x}$. The y coordinate of the centroid $\tilde{\overline{Y}}$ incorporates the fuzziness of the underling relationship (25), which is reflected in the intercept of the regression line $\hat{\tilde{A}}_0^*$ (26). The intercept is the only fuzzy coefficient of the linear function in our model. Such a constructed line, fulfilling the first Boscovich assumption (Section 2, constraint (1)), is going through the centroid, which is a necessary constraint for an unbiased estimate of the regression line. Considering this, we can construct an estimate of the slope of the k-th regression line on the basis of the mean values of the fuzzy numbers $\tilde{Y}_k$ and $\tilde{\overline{Y}}$

$$L_k' : m_{\tilde{Y}_k} = m_{\tilde{\overline{Y}}} + \hat{a}_1^{(k)} \Delta x_k, \tag{28}$$

where $m_{\tilde{Y}_k}$ is the mean of the k-th observation of $\tilde{Y}$, and $m_{\tilde{\overline{Y}}}$ is the mean of the y coordinate of the centroid $\tilde{\overline{Y}}$.

The k-th estimate of the slope a_1 is then given as

$$\hat{a}_1^{(k)} = \left(\Delta x_k^{-1} \right) \left(\tilde{Y}_k - m_{\tilde{\overline{Y}}} \right). \tag{29}$$

For the selection of the best estimate of the slope a_1, an appropriate evaluation function has to be formulated (see Section 2, constraint (2)). As follows from the relaxed Equation (28), the criterion is given as

$$J_F^{(k)} = \sum_{i=1}^n \left| m_{\tilde{Y}_i} - m_{\tilde{\overline{Y}}} - a_1^{(k)} \Delta x_i \right|, \tag{30}$$

and the best estimate of a_1 is given as

$$\hat{a}_1^* = \underset{\hat{a}_1^{(k)}, \forall k \in I}{\arg \min} \sum_{i=1}^n \left| m_{\tilde{Y}_i} - m_{\tilde{\overline{Y}}} - a_1^{(k)} \Delta x_i \right|. \tag{31}$$

The proposed fuzzy estimator can be written using a pseudocode as an Algorithm 1.

Algorithm 1 Boscovich fuzzy regression line

1: **function** BFRL(O)
Require: The set of n observations $O = \left\{ (x_i, \tilde{Y}_i) \big| \forall i \in I \right\}$
Ensure: The best estimates of $\tilde{A}_0$ and a_1
2:　　$\overline{x} \leftarrow \frac{1}{n} \sum_{i=1}^n x_i$
3:　　$\tilde{\overline{Y}} \leftarrow \frac{1}{n} \sum_{i=1}^n \tilde{Y}_i$
4:　　$\Delta x_i \leftarrow x_i - \overline{x}, \forall i \in I$
5:　　$\hat{a}_1^{(i)} \leftarrow \left(\Delta x_i^{-1} \right) \left(m_{\tilde{Y}_i} - m_{\tilde{\overline{Y}}} \right), \forall i \in I$
6:　　$J_F^{(k)} \leftarrow \sum_{i=1}^n \left| m_{\tilde{Y}_i} - m_{\tilde{\overline{Y}}} - \hat{a}_1^{(k)} \Delta x_i \right|, \forall k \in I$
7:　　$\hat{a}_1^* \leftarrow \underset{\hat{a}_1^{(k)}, \forall k \in I}{\arg \min} J_F^{(k)}$　　　　　　　　　　　　　　　　▷ Best estimate of a_1
8:　　$\hat{\tilde{A}}_0^* \leftarrow \tilde{\overline{Y}} - \hat{a}_1^* \overline{x}$　　　　　　　　　　　　　　　　　　　▷ Best estimate of $\tilde{A}_0$
9:　　**return** $\hat{\tilde{A}}_0^*, \hat{a}_1^*$
10: **end function**

---

5. Numerical Examples

We compared the presented Boscovich fuzzy regression line (BFRL) with several possibilistic and statistics-based fuzzy regression methods. As representatives of the possibilistic family, we selected a possibilistic linear regression (PLR) analysis [16], a PLR combined with an omission approach (OPRL) [44], and a multi-objective fuzzy linear regression (MOFLR) [26]. The OPRL and MOFLR were designed to be less sensitive to outliers. We implemented the OPRL for detection of one outlier. Each of these three methods requires setting of one parameter by a decision maker. The decision maker must set up a threshold value h in the cases of PRL and OPRL, where $h \in [0, 1)$. The threshold value indicates a degree of fitness of an estimated fuzzy regression model [16]. MOFLR requires presenting of a weighting coefficient ω, where $\omega \in (0, 1)$. The weighting coefficient determines a trade off between outlier penalization ($\omega \to 1$) and data fitting ($\omega \to 0$) [26]. From the statistics-based methods, we considered a fuzzy least squares (FLS) [18], a fuzzy least absolute linear regression (FLAR) [24], and a robust fuzzy regression (RFR) analysis [41]. Note that the RFR requires at least six observations to estimate parameters of a simple fuzzy linear model and it employs a customized iterative optimization method.

BFRL was designed as a parameter estimator of the fuzzy linear model (23). PLR, OPLR, MOFLR, FLS, FLAR and FLS expect a fuzzy linear model with all fuzzy parameters. For one real value independent variable x, the model is given as

$$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 \cdot x, \tag{32}$$

where $\tilde{A}_0$ and $\tilde{A}_1$ are unknown model parameters, $\tilde{A}_0, \tilde{A}_1 \in \mathbb{R}_T$, $\tilde{A}_1 = (a_1, \underline{\alpha}_1, \overline{\alpha}_1)_T$, a_1 is the mean, $\underline{\alpha}_1$ is the left, and $\overline{\alpha}_1$ is the right spread of $\tilde{A}_1$.

RFR uses a different approach to model prediction calculation. The i-th prediction of y, $\tilde{Y}_i = (y_i, \underline{v}_i, \overline{v}_i)_T$, is given as

$$\tilde{Y}_i = \begin{cases} y_i = \mathbf{fa}, \\ \underline{v}_i = by_i + d, \\ \overline{v}_i = gy_i + h, \end{cases} \tag{33}$$

where $\mathbf{a}, b, d, g$ and h are unknown model parameters, and $\mathbf{f} = \begin{bmatrix} 1 & x_i \end{bmatrix}$.

We evaluated performance of the aforementioned methods on various datasets. The possibilistic approaches (PRL, OPRL, MOFLR) were designed for symmetric fuzzy numbers. To allow comparison of all the methods, we involved datasets with symmetric fuzzy observations into the evaluation process. We used data from example 1 published in [16] and from example 2 published in [18]. We labelled the datasets SFN-1 and SFN-2, respectively. The presented BFRL, as well as the statistics-based methods (FLS, FLAR, RFR), are also capable of processing non-symmetric fuzzy numbers. To fully examine performance of these methods, we included three datasets with non-symmetric fuzzy observations of the dependent variable into the evaluation process. Specifically, we used data from example 1 published in [45], from example 2 published in [50], and from example 3 published in [51]. We labelled the datasets NFN-1, NFN-2, and NFN-3, respectively. The datasets SFN-1, SFN-2, NFN-1, NFN-2, and NFN-3 consist of 5, 8, 16, 8 and 8 observations, respectively.

We evaluated performance of the methods using a total error E which we defined as a sum of absolute errors, i.e.,

$$E = \sum_{i=1}^{n} D_i, \tag{34}$$

where D is a difference between membership functions of observed and estimated fuzzy responses. For the i-th observation, the difference is given as

$$D_i = \int_{S_{\tilde{Y}_i} \cup S_{\hat{\tilde{Y}}_i}} \left| \mu_{\tilde{Y}_i}(y) - \mu_{\hat{\tilde{Y}}_i}(y) \right| dy, \tag{35}$$

where $\mu_{\tilde{Y}_i}(y)$ and $\mu_{\hat{\tilde{Y}}_i}(y)$ are the membership functions of the *i*-th observed $\tilde{Y}_i$ and the *i*-th estimated response $\hat{\tilde{Y}}_i$, respectively; and $S_{\tilde{Y}_i}$ and $S_{\hat{\tilde{Y}}_i}$ represent supports of $\mu_{\tilde{Y}_i}(y)$ and $\mu_{\hat{\tilde{Y}}_i}(y)$, respectively [52]. Total errors (34) of the examined fuzzy regression methods are summarized in Table 1. The errors are organized with respect to fuzzy regression methods (columns) and datasets (rows). For PRL and OPRL, we used threshold values $h = 0$ and $h = 0.5$. For MOFLR, we considered $\omega \in \{0.1, 0.5, 0.99\}$. Note that RFR cannot be applied on SFN-1 since the dataset consists of 5 observations only. Since PRL, OPRL and MOFLR return biased results on datasets with non-symmetric fuzzy observations, the results were not included into Table 1.

Table 1. Total errors of the possibilistic linear regression (PLR), the PLR combined with the omission approach (OPRL), the multi-objective fuzzy linear regression (MOFLR), the fuzzy least squares (FLS), the fuzzy least absolute linear regression (FLAR), the robust fuzzy regression (RFR) analysis, and the Boscovich fuzzy regression line (BFRL). Testing done on datasets with symmetric (SFN-1, $n = 5$ and SFN-2, $n = 8$) and non-symmetric fuzzy numbers (NFN-1, $n = 16$; NFN-2, $n = 8$ and NFN-3, $n = 8$). Settings of adjustable parameters h and ω are stated under method abbreviations.

	PLR		OPRL		MOFLR			FLS	FLAR	RFR	BFRL
	$h = 0$	$h = 0.5$	$h = 0$	$h = 0.5$	$\omega = 0.1$	$\omega = 0.5$	$\omega = 0.99$				
SFN-1	12.40	17.89	12.40	17.89	220.42	35.35	14.53	10.14	9.50	-	9.17
SFN-2	3.86	6.45	3.51	4.78	73.73	11.89	5.09	3.23	3.24	3.23	3.61
NFN-1	-	-	-	-	-	-	-	144.13	133.81	151.98	161.09
NFN-2	-	-	-	-	-	-	-	15.21	14.09	14.57	15.30
NFN-3	-	-	-	-	-	-	-	2.98	2.29	2.90	3.03

We also investigated the sensitivity of the methods on outliers. We considered three types of outliers: (a) outliers with respect to centres of the fuzzy dependent variable $\tilde{Y}$ (o1), (b) outliers with respect to spreads of $\tilde{Y}$ (o2), and (c) outliers with respect to both the centres and the spreads (o3). To examine the sensitivity, we created from each above stated dataset, three new sets where each new set contained one outlier. The outliers in the datasets are specified in Table 2 by their serial numbers i, means y_i, left spreads $\underline{v}_i$ and right spreads $\overline{v}_i$. Their original values are written in normal text while the changed ones are in bold.

Table 2. Outliers of type o1 (columns 3–5), o2 (columns 6 - 8) and o3 (columns 9–11) in datasets (column 1) are identified by their serial numbers i (column 2). Original values of means $m_{\tilde{Y}}$, right spreads $\alpha_{\tilde{Y}}$ and left spreads $\beta_{\tilde{Y}}$ of the observations $\tilde{Y}$ are in normal text and the changed values are in bold.

		o1			o2			o3		
Dataset	i	y_i	$\underline{v}_i$	$\overline{v}_i$	y_i	$\underline{v}_i$	$\overline{v}_i$	y_i	$\underline{v}_i$	$\overline{v}_i$
SFN-1	3	**1**	1.8	1.8	8	**35**	**35**	**1**	**35**	**35**
SFN-2	4	**4**	0.4	0.4	2	**1.6**	**1.6**	**4**	**1.6**	**1.6**
NFN-1	7	**170**	12	12	70.9	**65**	**77**	**170**	**65**	**77**
NFN-2	6	**3**	1.5	1.7	22	**10**	**0.01**	**3**	**10**	**0.01**
NFN-3	1	**9.5**	0.17	0.4	2.5	**3**	**4**	**9.5**	**3**	**4**

Total errors of acquired fuzzy regression models on the modified datasets with symmetric and non-symmetric fuzzy observations are summarized in Tables 3 and 4, respectively. The results are organized with respect to fuzzy regression methods (columns) and datasets (rows). Settings of adjustable parameters are given under method abbreviations. Asterisk marked results point to the situation when outliers were correctly recognized by OPRL (Table 3).

Table 3. Total errors of the possibilistic linear regression (PLR), the PLR combined with the omission approach (OPRL), the multi-objective fuzzy linear regression (MOFLR), the fuzzy least squares (FLS), the fuzzy least absolute linear regression (FLAR), the robust fuzzy regression (RFR) analysis, and the Boscovich fuzzy regression line (BFRL). Method tested on datasets with symmetric fuzzy numbers SFN-1 and SFN-2 affected by outliers of type o1, o2 and o3 (first column). Settings of adjustable parameters h and ω are stated under method abbreviations. Asterisk indicates correctly recognized outliers by OPRL.

	PLR		OPRL		MOFLR			FLS	FLAR	RFR	BFRL
	$h = 0$	$h = 0.5$	$h = 0$	$h = 0.5$	$\omega = 0.1$	$\omega = 0.5$	$\omega = 0.99$				
SFN-1-o1	35.99	56.88	35.99	23.03 *	220.58	39.57	22.50	15.74	14.17	-	14.59
SFN-2-o1	7.58	12.70	4.33 *	6.29 *	73.74	12.29	5.82	3.76	3.01	3.76	3.49
SFN-1-o2	131.05	131.05	42.03 *	44.30 *	836.00	132.99	79.48	52.97	41.08	-	53.02
SFN-2-o2	8.21	10.17	-	5.62 *	96.52	15.44	6.51	3.90	3.82	4.60	3.97
SFN-1-o3	132.57	132.38	42.56 *	45.06 *	836.13	145.87	81.79	54.64	42.06	-	54.61
SFN-2-o3	10.70	15.72	5.23 *	6.96 *	96.54	16.10	7.69	4.80	3.93	5.64	4.60

Table 4. Total errors of the fuzzy least squares (FLS), the fuzzy least absolute linear regression (FLAR), the robust fuzzy regression (RFR) analysis, and the Boscovich fuzzy regression line (BFRL) (first row) tested on datasets with non-symmetric fuzzy numbers, NFN-1, NFN-2, and NFN-3, affected by outliers of type o1, o2 and o3 (first column).

	FLS	FLAR	RFR	BFRL
NFN-1-o1	180.45	146.88	166.43	189.74
NFN-2-o1	19.23	15.34	20.14	19.14
NFN-3-o1	5.30	2.29	5.32	5.03
NFN-1-o2	207.69	183.84	192.53	214.43
NFN-2-o2	15.74	14.89	17.04	16.54
NFN-3-o2	5.86	4.94	6.57	6.15
NFN-1-o3	259.93	205.94	393.57	259.24
NFN-2-o3	25.98	19.47	28.62	25.64
NFN-3-o3	11.63	5.58	11.69	10.60

The obtained parameter estimates by PRL and OPRL on datasets with symmetric fuzzy numbers are summarized in Table 5. Estimates provided by MOFLR on the same datasets are given in Table 6. Settings of adjustable parameters h and ω are stated under method abbreviations. The estimates of model parameters generated by FLS, FLAR, RFR and BFRL on datasets with symmetric and non-symmetric fuzzy observations are summarized in Tables 7 and 8, respectively.

We implemented all the fuzzy regression methods in MATLAB R2016a. We used default setting of optimization functions which were used within the calculations. It means that interior point methods were utilized to solve both linear and quadratic programming problems.

Table 5. Parameters of fuzzy regression models $\tilde{A}_0$ and $\tilde{A}_1$ estimated by the possibilistic linear regression (PLR) and the PLR combined with the omission approach (OPRL) on datasets (first column) with symmetric fuzzy observations. Settings of the adjustable parameter h are stated in the second line.

	PLR				OPRL			
	$h=0$		$h=0.5$		$h=0$		$h=0.5$	
	$\tilde{A}_0$	$\tilde{A}_1$	$\tilde{A}_0$	$\tilde{A}_1$	$\tilde{A}_0$	$\tilde{A}_1$	$\tilde{A}_0$	$\tilde{A}_1$
SFN-1	$(3.85,3.85,3.85)_T$	$(2.10,0.00,0.00)_T$	$(4.15,5.57,5.57)_T$	$(1.97,0.00,0.00)_T$	$(3.85,3.85,3.85)_T$	$(2.10,0.00,0.00)_T$	$(4.15,5.57,5.57)_T$	$(1.97,0.00,0.00)_T$
SFN-2	$(1.28,0.83,0.83)_T$	$(0.13,0.00,0.00)_T$	$(1.39,1.23,1.23)_T$	$(0.11,0.00,0.00)_T$	$(1.03,0.63,0.63)_T$	$(0.14,0.01,0.01)_T$	$(0.94,0.98,0.98)_T$	$(0.15,0.00,0.00)_T$
SFN-1-o1	$(1.68,6.02,6.02)_T$	$(1.51,0.59,0.59)_T$	$(2.27,9.32,9.32)_T$	$(1.33,1.27,1.27)_T$	$(1.68,6.02,6.02)_T$	$(1.51,0.59,0.59)_T$	$(4.15,5.57,5.57)_T$	$(1.20,0.00,0.00)_T$
SFN-2-o1	$(2.85,1.25,1.25)_T$	$(0.03,0.00,0.00)_T$	$(3.05,2.00,2.00)_T$	$(0.02,0.00,0.00)_T$	$(1.29,0.82,0.82)_T$	$(0.13,0.00,0.00)_T$	$(1.44,1.12,1.12)_T$	$(0.11,0.00,0.00)_T$
SFN-1-o2	$(5.53,26.47,26.47)_T$	$(1.32,2.84,2.84)_T$	$(5.52,25.25,25.25)_T$	$(1.33,3.25,3.25)_T$	$(3.85,3.85,3.85)_T$	$(2.10,0.00,0.00)_T$	$(4.15,5.57,5.57)_T$	$(1.98,0.00,0.00)_T$
SFN-2-o2	$(0.53,1.60,1.60)_T$	$(0.16,0.00,0.00)_T$	$(1.09,1.83,1.83)_T$	$(0.11,0.00,0.00)_T$	-	-	$(1.44,1.12,1.12)_T$	$(0.11,0.00,0.00)_T$
SFN-1-o3	$(6.68,26.19,26.19)_T$	$(-1.89,2.94,2.94)_T$	$(4.13,24.07,24.07)_T$	$(-1.04,3.64,3.64)_T$	$(3.85,3.85,3.85)_T$	$(2.10,0.00,0.00)_T$	$(4.15,5.57,5.57)_T$	$(1.97,0.00,0.00)_T$
SFN-2-o3	$(4.01,1.77,1.77)_T$	$(-0.02,0.00,0.00)_T$	$(3.63,2.52,2.52)_T$	$(-0.01,0.00,0.00)_T$	$(1.29,0.82,0.82)_T$	$(0.13,0.00,0.00)_T$	$(1.44,1.12,1.12)_T$	$(0.11,0.00,0.00)_T$

Table 6. Parameters of fuzzy regression models $\tilde{A}_0$ and $\tilde{A}_1$ estimated by the multi-objective fuzzy linear regression (MOFLR) on datasets (first column) with symmetric fuzzy observations. Settings of the adjustable parameter ω are stated in the second line. Negative spreads are in bold.

	MOFLR					
	$\omega=0.1$		$\omega=0.5$		$\omega=0.99$	
	$\tilde{A}_0$	$\tilde{A}_1$	$\tilde{A}_0$	$\tilde{A}_1$	$\tilde{A}_0$	$\tilde{A}_1$
SFN-1	$(4.95,36.80,36.80)_T$	$(1.71,3.20,3.20)_T$	$(4.95,7.36,7.36)_T$	$(1.71,0.64,0.64)_T$	$(4.95,1.84,1.84)_T$	$(1.71,0.16,0.16)_T$
SFN-2	$(1.38,2.95,2.95)_T$	$(0.12,0.50,0.50)_T$	$(1.38,0.59,0.59)_T$	$(0.12,0.10,0.10)_T$	$(1.37,0.30,0.30)_T$	$(0.12,0.05,0.05)_T$
SFN-1-o1	$(3.25,36.80,36.80)_T$	$(1.71,3.20,3.20)_T$	$(3.25,7.36,7.36)_T$	$(1.71,0.64,0.64)_T$	$(3.25,3.72,3.72)_T$	$(1.71,0.32,0.32)_T$
SFN-2-o1	$(2.38,2.95,2.95)_T$	$(0.06,0.50,0.50)_T$	$(2.38,0.59,0.59)_T$	$(0.06,0.10,0.10)_T$	$(2.38,0.30,0.30)_T$	$(0.07,0.05,0.05)_T$
SFN-1-o2	$(4.95,166.40,166.40)_T$	$(1.71,3.20,3.20)_T$	$(4.95,33.28,33.28)_T$	$(1.71,0.64,0.64)_T$	$(4.94,16.81,16.81)_T$	$(1.71,0.32,0.32)_T$
SFN-2-o2	$(1.38,14.95,14.95)_T$	$(0.12,\mathbf{-0.17},\mathbf{-0.17})_T$	$(1.38,2.99,2.99)_T$	$(0.12,\mathbf{-0.03},\mathbf{-0.03})_T$	$(1.35,1.51,1.51)_T$	$(0.12,\mathbf{-0.02},\mathbf{-0.02})_T$
SFN-1-o3	$(3.25,166.40,166.40)_T$	$(1.71,3.20,3.20)_T$	$(3.25,33.28,33.28)_T$	$(1.71,0.64,0.64)_T$	$(3.26,16.81,16.81)_T$	$(1.71,0.32,0.32)_T$
SFN-2-o3	$(2.38,14.95,14.95)_T$	$(0.06,\mathbf{-0.17},\mathbf{-0.17})_T$	$(2.38,2.99,2.99)_T$	$(0.06,\mathbf{-0.03},\mathbf{-0.03})_T$	$(2.38,1.51,1.51)_T$	$(0.07,\mathbf{-0.02},\mathbf{-0.02})_T$

Table 7. Parameters of fuzzy regression models $\tilde{A}_0$, $\tilde{A}_1$, $\mathbf{a}$, b, g, d, h and a_1 estimated by the fuzzy least squares (FLS), the fuzzy least absolute linear regression (FLAR), the robust fuzzy regression (RFR) analysis, and the Boscovich fuzzy regression line (BFRL) on datasets (first column) with symmetric fuzzy observations. Negative spreads are in bold.

	FLS		FLAR		RFR					BFRL	
	$\tilde{A}_0$	$\tilde{A}_1$	$\tilde{A}_0$	$\tilde{A}_1$	$\mathbf{a}^T$	b	g	d	h	$\tilde{A}_0$	a_1
SFN-1	$(4.95,1.84,1.84)_T$	$(1.71,0.16,0.16)_T$	$(5.61,1.80,1.80)_T$	$(1.48,0.20,0.20)_T$	-	-	-	-	-	$(5.70,2.32,2.32)_T$	1.46
SFN-2	$(1.38,1.48,1.48)_T$	$(0.12,0.03,0.03)_T$	$(1.41,0.10,0.10)_T$	$(0.12,0.03,0.03)_T$	$[0.81 \quad 0.16]$	0.18	0.18	-0.03	-0.02	$(1.20,0.49,0.49)_T$	0.13
SFN-1-o1	$(3.25,1.84,1.84)_T$	$(1.71,0.16,0.16)_T$	$(4.14,1.80,1.80)_T$	$(1.77,0.20,0.20)_T$	-	-	-	-	-	$(2.44,2.32,2.32)_T$	1.98
SFN-2-o1	$(2.38,0.15,0.15)_T$	$(0.06,0.03,0.03)_T$	$(2.00,0.10,0.10)_T$	$(0.08,0.03,0.03)_T$	$[2.37 \quad 0.06]$	0.39	0.39	-0.77	-0.77	$(1.90,0.49,0.49)_T$	0.10
SFN-1-o2	$(4.95,8.32,8.32)_T$	$(1.71,0.16,0.16)_T$	$(5.60,1.80,1.80)_T$	$(1.48,0.20,0.20)_T$	-	-	-	-	-	$(5.70,8.80,8.80)_T$	1.46
SFN-2-o2	$(1.38,0.75,0.75)_T$	$(0.12,\mathbf{-0.01},\mathbf{-0.01})_T$	$(1.42,0.21,0.21)_T$	$(0.12,0.02,0.02)_T$	$[1.57 \quad 0.10]$	0.18	0.19	0.14	0.12	$(1.20,0.64,0.64)_T$	0.13
SFN-1-o3	$(3.25,8.32,8.32)_T$	$(1.71,0.16,0.16)_T$	$(4.05,1.80,1.80)_T$	$(1.79,0.20,0.20)_T$	-	-	-	-	-	$(2.44,8.80,8.80)_T$	1.98
SFN-2-o3	$(2.38,0.75,0.75)_T$	$(0.06,\mathbf{-0.01},\mathbf{-0.01})_T$	$(2.00,0.21,0.21)_T$	$(0.08,0.02,0.02)_T$	$[2.78 \quad 0.03]$	0.18	0.18	0.03	0.03	$(1.90,0.64,0.64)_T$	0.10

Table 8. Parameters of fuzzy regression models $\tilde{A}_0, \tilde{A}_1, \mathbf{a}, b, g, d, h$ and a_1 estimated by the fuzzy least squares (FLS), the fuzzy least absolute linear regression (FLAR), the robust fuzzy regression (RFR) analysis, and the Boscovich fuzzy regression line (BFRL) on datasets (first column) with non-symmetric fuzzy observations. Negative spreads are in bold.

	FLS		FLAR		RFR						BFRL	
	$\tilde{A}_0$	$\tilde{A}_1$	$\tilde{A}_0$	$\tilde{A}_1$	$\mathbf{a}^{\mathbf{T}}$		b	g	d	h	$\tilde{A}_0$	a_1
NFN-1	$(24.47, 4.85, 4.46)_T$	$(34.05, 4.95, 5.80)_T$	$(25.46, 4.68, 4.82)_T$	$(32.90, 4.90, 5.45)_T$	$[21.66$	$35.16]$	0.16	0.17	0.67	0.39	$(27.07, 12.28, 13.16)_T$	32.32
NFN-2	$(12.93, 1.29, 1.70)_T$	$(0.54, 0.04, 0.01)_T$	$(12.86, 1.30, 1.58)_T$	$(0.57, 0.02, 0.02)_T$	$[12.65$	$0.55]$	0.09	0.09	0.13	0.13	$(12.93, 1.63, 1.79)_T$	0.54
NFN-3	$(1.31, 0.16, 0.29)_T$	$(0.13, 0.01, 0.02)_T$	$(0.51, 0.15, 0.29)_T$	$(0.17, 0.01, 0.02)_T$	$[1.31$	$0.13]$	0.06	0.15	0.08	0.10	$(1.09, 0.26, 0.54)_T$	0.14
NFN-1-o1	$(33.94, 4.85, 4.46)_T$	$(31.87, 4.95, 5.80)_T$	$(25.46, 4.70, 4.82)_T$	$(32.90, 4.94, 5.45)_T$	$[35.02$	$27.99]$	0.16	0.18	0.17	0.02	$(37.05, 12.28, 13.16)_T$	29.79
NFN-2-o1	$(13.61, 1.29, 1.70)_T$	$(0.20, 0.04, 0.01)_T$	$(13.00, 1.30, 1.58)_T$	$(0.50, 0.02, 0.02)_T$	$[13.25$	$0.19]$	0.10	0.12	0.06	0.02	$(13.61, 1.63, 1.79)_T$	0.20
NFN-3-o1	$(6.80, 0.16, 0.29)_T$	$(-0.22, 0.01, 0.02)_T$	$(0.51, 0.15, 0.29)_T$	$(0.17, 0.01, 0.02)_T$	$[8.62$	$-0.36]$	0.01	0.05	0.18	0.26	$(5.10, 0.26, 0.54)_T$	-0.09
NFN-1-o2	$(24.47, 9.92, 10.67)_T$	$(34.05, 3.78, 4.36)_T$	$(25.46, 4.55, 4.82)_T$	$(32.90, 5.03, 5.45)_T$	$[34.14$	$29.79]$	0.15	0.17	0.73	0.49	$(27.07, 15.59, 17.22)_T$	32.32
NFN-2-o2	$(12.93, 0.99, 1.76)_T$	$(0.54, 0.19, \mathbf{-0.02})_T$	$(12.86, 0.95, 1.58)_T$	$(0.57, 0.13, 0.02)_T$	$[12.86$	$0.47]$	0.10	0.10	0.07	0.11	$(12.93, 2.69, 1.58)_T$	0.54
NFN-3-o2	$(1.31, 2.38, 3.11)_T$	$(0.13, \mathbf{-0.13}, \mathbf{-0.16})_T$	$(0.51, 0.20, 0.41)_T$	$(0.17, 0.01, 0.01)_T$	$[1.50$	$0.11]$	-0.01	0.11	0.68	0.71	$(1.09, 0.61, 0.99)_T$	0.14
NFN-1-o3	$(33.94, 9.92, 10.67)_T$	$(31.87, 3.78, 4.36)_T$	$(25.46, 4.71, 4.82)_T$	$(32.90, 4.97, 5.45)_T$	$[53.26$	$25.81]$	0.19	0.23	3.44	3.25	$(37.05, 15.59, 17.22)_T$	29.79
NFN-2-o3	$(13.61, 0.99, 1.76)_T$	$(0.20, 0.19, \mathbf{-0.02})_T$	$(13.00, 0.95, 1.58)_T$	$(0.50, 0.13, 0.02)_T$	$[13.69$	$0.10]$	0.21	0.11	-0.12	-0.03	$(13.61, 2.69, 1.58)_T$	0.20
NFN-3-o3	$(6.80, 2.38, 3.11)_T$	$(-0.22, \mathbf{-0.13}, \mathbf{-0.16})_T$	$(0.51, 0.20, 0.41)_T$	$(0.17, 0.01, 0.01)_T$	$[8.44$	$-0.37]$	0.34	0.42	-0.55	-0.44	$(5.10, 0.61, 0.99)_T$	-0.09

6. Discussion

We demonstrated in the numerical examples that the proposed Boscovich fuzzy regression line (BFRL) is capable of approximating a fuzzy linear relationship between the dependent y and one independent variable x, where the independent variable is a real value number and the dependent variable is a triangular fuzzy number $\tilde{Y}$. We compared BFRL with several other fuzzy linear regression methods. Most of the reference methods (PRL, OPRL, MOFLR, FLS and FLAR) approximate the relationship using the fuzzy linear model (32), while RFR uses the model (33). Prediction spreads of both models are dependent on x. BFRL estimates parameters of the fuzzy linear model (23). Prediction spreads of this model are independent of x. This fact predetermines BFRL for applications where the fuzziness of model predictions is independent of model inputs. An example of such an application is approximation of the dependence of water level, in an uncovered channel, on the opening of a floodgate. The level is measured using a perpendicularly positioned scale.

We studied performance of the fuzzy regression methods on twenty datasets where fifteen of them were affected by outliers. We measured prediction errors of acquired fuzzy regression models using the total error (34). We found that, in most cases, BFRL based models show lower total errors compared to PLR, OPRL, MOFLR and RFR models. Total errors of BFRL and FLS models are comparable but, in comparison, FLAR models always show lower errors (Tables 1, 3 and 4). Sensitivity of BFRL to all types of outliers is comparable with other statistics-based methods (FLS, FLAR, RFR) but considerably lower compared to PLR and MOFLR (Tables 3 and 4). If an outlier is correctly recognized by OPRL, the total error of a model produced by OPRL on the outlier affected dataset, is comparable with the total error of a BFRL based model (Table 3).

Unlike some reference methods, BFRL proved to be generally applicable and reliable. BFRL provides sensible parameter estimates for datasets with symmetric as well as with non-symmetric triangular fuzzy numbers (as compared to PLR, OPRL and MOFLR). BFRL is capable of operating on datasets with only two observations. For example, RFR requires at least six observations. In contrast to OPRL (SFN-2-o2 for $h = 0$, Table 5), BFRL always provided estimates of parameters (Tables 5 and 6). In contrast to RFR, BFRL always returned the same estimates on the same dataset. The iterative manner of optimization in RFR leads generally to various parameter estimates and various total errors. BFRL also guaranteed non-negativity of parameter spreads (Tables 7 and 8). This basic requirement is not guaranteed by MOFLR and FLS.

Fuzzy regression methods have mostly large computational complexity, which makes them difficult to use especially with regard to their not easy implementation. For example, PLR, OPRL, and FLAR result in a linear programming problem, MOFLR results in a quadratic one. An analytical solution must be expressed in the case of FLS, and a cus-

tomized iterative optimization method must be implemented in the case of RFR. Moreover, time-complexity of these methods is exponential. A pleasant feature of the method BFRL (Algorithm 1) is straightforward implementation and it has linear time-complexity.

7. Conclusions

Our approach is based on a fuzzy modification of the first statistical method for data fitting based on criterion minimization. The generalization of the Boschovich line to fuzzy regression does not allow an approximation of data by a model other than the regression line. Other predecessors of the ordinary least squares method (Mayer's method of averages, Lambert's line, and Laplace's method, see [53]) could enhance the family of fuzzy linear regression methods in future research. For example, a similarity between Lambert's ideas and robust non-parametric methods such as repeated median regression and the Theil–Stein estimate [54] might provide future directions for further research.

The proposed Boscovich fuzzy regression line is a simple fuzzy regression method. We demonstrated on the numerical examples that BFRL is a reliable method which provides good parameter estimates including those estimated from datasets affected by outliers. Prediction errors of BFRL models are smaller when compared to models proposed by the possibilistic-based methods, and favourably comparable with prediction errors of models produced by statistics-based methods. In comparison with the other methods, BFRL has two major advantages: small time-complexity and straightforward implementation. Moreover, the method guarantees non-negativity of model parameter spreads. The robustness and user-friendliness of the method makes it relevant in non-mathematical research fields.

Author Contributions: Conceptualization, J.M.; methodology, J.M. and P.Š.; software, P.Š.; validation, J.M., P.Š. and A.P.; formal analysis, J.M., P.Š. and A.P.; investigation, P.Š. and A.P.; data curation, P.Š.; writing—original draft preparation, P.Š.; writing—review and editing, J.M., P.Š. and A.P.; supervision, J.M. and P.Š. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This study is based on previously published data to which the authors continuously refer in the text.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bisserier, A.; Boukezzoula, R.; Galichet, S. A revisited approach to linear fuzzy regression using trapezoidal fuzzy intervals. *Inf. Sci.* **2010**, *180*, 3653–3673. [CrossRef]
2. Škrabánek, P.; Martínková, N. Algorithm xxxx: fuzzyreg: An R package for fitting fuzzy regression models. *ACM Trans. Math. Soft.* **2021**, in press.
3. Özelkan, E.C.; Duckstein, L. Multi-objective fuzzy regression: A general framework. *Comput. Oper. Res.* **2000**, *27*, 635–652. [CrossRef]
4. Yager, R.R. Fuzzy prediction based on regression models. *Inf. Sci.* **1982**, *26*, 45–63. [CrossRef]
5. Toyoura, Y.; Watada, J.; Khalid, M.; Yusof, R. Formulation of linguistic regression model based on natural words. *Soft Comput.* **2004**, *8*, 681–688. [CrossRef]
6. Pan, N.F.; Lin, T.C.; Pan, N.H. Estimating bridge performance based on a matrix-driven fuzzy linear regression model. *Autom. Constr.* **2009**, *18*, 578–586. [CrossRef]
7. Kim, K.J.; Moskowitz, H.; Koksalan, M. Fuzzy versus statistical linear regression. *Eur. J. Oper. Res.* **1996**, *92*, 417–434. [CrossRef]
8. Škrabánek, P.; Marek, J. Models used in fuzzy linear regression. In Proceedings of the 17th Conference on Applied Mathematics—APLIMAT 2018, Bratislava, Slovak Republic, 6–8 February 2018; pp. 955–964.
9. Mosleh, M.; Otadi, M.; Abbasbandy, S. Fuzzy polynomial regression with fuzzy neural networks. *Appl. Math. Model.* **2011**, *35*, 5400–5412. [CrossRef]
10. Pourahmad, S.; Ayatollahi, S.M.T.; Taheri, S.M.; Agahi, Z.H. Fuzzy logistic regression based on the least squares approach with application in clinical studies. *Comput. Math. Appl.* **2011**, *62*, 3353–3365. [CrossRef]

11. Cheng, C.B.; Lee, E. Nonparametric fuzzy regression—$k − NN$ and kernel smoothing techniques. *Comput. Math. Appl.* **1999**, *38*, 239–251. [CrossRef]

12. Wang, N.; Zhang, W.X.; Mei, C.L. Fuzzy nonparametric regression based on local linear smoothing technique. *Inf. Sci.* **2007**, *177*, 3882–3900. [CrossRef]

13. Danesh, S.; Farnoosh, R.; Razzaghnia, T. Fuzzy nonparametric regression based on an adaptive neuro-fuzzy inference system. *Neurocomputing* **2016**, *173*, 1450–1460. [CrossRef]

14. Maturo, F.; Hošková-Mayerová, Š. Fuzzy Regression Models and Alternative Operations for Economic and Social Sciences. In *Recent Trends in Social Systems: Quantitative Theories and Quantitative Models*, 1st ed.; Springer International Publishing: Cham, Switzerland, 2017; pp. 235–247. [CrossRef]

15. Tanaka, H.; Uejima, S.; Asai, K. Linear Regression Analysis with Fuzzy Model. *IEEE Trans. Syst. Man Cybern.* **1982**, *12*, 903–907. [CrossRef]

16. Tanaka, H.; Hayashi, I.; Watada, J. Possibilistic linear regression analysis for fuzzy data. *Eur. J. Oper. Res.* **1989**, *40*, 389–396. [CrossRef]

17. Chang, P.T.; Lee, E.S. A generalized fuzzy weighted least-squares regression. *Fuzzy Sets Syst.* **1996**, *82*, 289–298. [CrossRef]

18. Diamond, P. Fuzzy Least Squares. *Inf. Sci.* **1988**, *46*, 141–157. [CrossRef]

19. Yen, K.; Ghoshray, S.; Roig, G. A linear regression model using triangular fuzzy number coefficients. *Fuzzy Sets Syst.* **1999**, *106*, 167–177. [CrossRef]

20. D'Urso, P.; Gastaldi, T. A least-squares approach to fuzzy linear regression analysis. *Comput. Stat. Data Anal.* **2000**, *34*, 427–440. [CrossRef]

21. Hojati, M.; Bector, C.; Smimou, K. A simple method for computation of fuzzy linear regression. *Eur. J. Oper. Res.* **2005**, *166*, 172–184. [CrossRef]

22. Choi, S.H.; Buckley, J.J. Fuzzy regression using least absolute deviation estimators. *Soft Comput.* **2008**, *12*, 257–263. [CrossRef]

23. Kelkinnama, M.; Taheri, S. Fuzzy least-absolutes regression using shape preserving operations. *Inf. Sci.* **2012**, *214*, 105–120. [CrossRef]

24. Zeng, W.; Feng, Q.; Li, J. Fuzzy least absolute linear regression. *Appl. Soft Comput.* **2017**, *52*, 1009–1019. [CrossRef]

25. Nasrabadi, M.M.; Nasrabadi, E. A mathematical-programming approach to fuzzy linear regression analysis. *Appl. Math. Comput.* **2004**, *155*, 873–881. [CrossRef]

26. Nasrabadi, M.M.; Nasrabadi, E.; Nasrabady, A.R. Fuzzy linear regression analysis: A multi-objective programming approach. *Appl. Math. Comput.* **2005**, *163*, 245–251. [CrossRef]

27. Celmiņš, A. Least squares model fitting to fuzzy vector data. *Fuzzy Sets Syst.* **1987**, *22*, 245–269. [CrossRef]

28. Diamond, P.; Körner, R. Extended fuzzy linear models and least squares estimates. *Comput. Math. Appl.* **1997**, *33*, 15–32. [CrossRef]

29. Kao, C.; Chyu, C.L. A fuzzy linear regression model with better explanatory power. *Fuzzy Sets Syst.* **2002**, *126*, 401–409. [CrossRef]

30. Kao, C.; Chyu, C.L. Least-squares estimates in fuzzy regression analysis. *Eur. J. Oper. Res.* **2003**, *148*, 426–435. [CrossRef]

31. Wu, H.C. Fuzzy estimates of regression parameters in linear regression models for imprecise input and output data. *Comput. Stat. Data Anal.* **2003**, *42*, 203–217. [CrossRef]

32. Hassanpour, H.; Maleki, H.; Yaghoobi, M. A note on evaluation of fuzzy linear regression models by comparing membership functions. *Iran. J. Fuzzy Syst.* **2009**, *6*, 1–6. [CrossRef]

33. Wu, H.C. The construction of fuzzy least squares estimators in fuzzy linear regression models. *Expert Syst. Appl.* **2011**, *38*, 13632–13640. [CrossRef]

34. Chen, L.H.; Hsueh, C.C. A mathematical programming method for formulating a fuzzy regression model based on distance criterion. *IEEE Trans. Syst. Man, Cybern. Part B (Cybern.)* **2007**, *37*, 705–712. [CrossRef]

35. Hassanpour, H.; Maleki, H.; Yaghoobi, M. Fuzzy linear regression model with crisp coefficients: A goal programming approach. *Iran. J. Fuzzy Syst.* **2010**, *7*, 1–153. [CrossRef]

36. Li, J.; Zeng, W.; Xie, J.; Yin, Q. A new fuzzy regression model based on least absolute deviation. *Eng. Appl. Artif. Intell.* **2016**, *52*, 54–64. [CrossRef]

37. Al-Qudaimi, A.; Kumar, A. A note on "A new fuzzy regression model based on absolute deviation". *Eng. Appl. Artif. Intell.* **2017**, *66*, 30–32. [CrossRef]

38. Lee, H.; Tanaka, H. Fuzzy approximations with non-symmetric fuzzy parameters in fuzzy regression analysis. *J. Oper. Res. Soc. Jpn.* **1999**, *42*, 98–112. [CrossRef]

39. Redden, D.T.; Woodall, W.H. Properties of certain fuzzy linear regression methods. *Fuzzy Sets Syst.* **1994**, *64*, 361–375. [CrossRef]

40. Redden, D.T.; Woodall, W.H. Further examination of fuzzy linear regression. *Fuzzy Sets Syst.* **1996**, *79*, 203–211. [CrossRef]

41. D'Urso, P.; Massari, R.; Santoro, A. Robust fuzzy regression analysis. *Inf. Sci.* **2011**, *181*, 4154–4174. [CrossRef]

42. Peters, G. Fuzzy linear regression with fuzzy intervals. *Fuzzy Sets Syst.* **1994**, *63*, 45–55. [CrossRef]

43. Kim, K.J.; Chen, H.R. A comparison of fuzzy and nonparametric linear regression. *Comput. Oper. Res.* **1997**, *24*, 505–519. [CrossRef]

44. Hung, W.L.; Yang, M.S. An omission approach for detecting outliers in fuzzy regression models. *Fuzzy Sets Syst.* **2006**, *157*, 3109–3122. [CrossRef]

45. Chang, P.T.; Lee, E. Fuzzy least absolute deviations regression and the conflicting trends in fuzzy parameters. *Comput. Math. Appl.* **1994**, *28*, 89–101. [CrossRef]

46. Lu, J.; Wang, R. An enhanced fuzzy linear regression model with more flexible spreads. *Fuzzy Sets Syst.* **2009**, *160*, 2505–2523. [CrossRef]

47. Boscovich, R.J. De litteraria expeditione per pontificiam ditioned, et synopsis amplioris operis, ac habentur plura ejus ex exemplaria etiam sensorum impressa. *Bononiensi Sci. Artium Inst. Atque Acad. Comment.* **1757**, *4*, 353–396.

48. Howarth, R.J. A History of Regression and Related Model-Fitting in the Earth Sciences (1636–2000). *Nat. Resour. Res.* **2001**, *10*, 241–286. [CrossRef]

49. Ku, H. *Precision Measurement and Calibration: Selected NBS Papers on Statistical Concepts and Procedures*; NBS Special Publication 300; U.S. Government Printing Office: Washington, DC, USA, 1969; Volume 1.

50. Chang, Y.H.O. Hybrid fuzzy least-squares regression analysis and its reliability measures. *Fuzzy Sets Syst.* **2001**, *119*, 225–246. [CrossRef]

51. Körner, R.; Näther, W. Linear regression with random fuzzy variables: Extended classical estimates, best linear estimates, least squares estimates. *Inf. Sci.* **1998**, *109*, 95–118. [CrossRef]

52. Kim, B.; Bishu, R.R. Evaluation of fuzzy linear regression models by comparing membership functions. *Fuzzy Sets Syst.* **1998**, *100*, 343–352. [CrossRef]

53. Hald, A. *A History of Probability and Statistics and Their Applications before 1750*, 1st ed.; Wiley series in probability and statistics; John Wiley & Sons: Hoboken, NJ, USA, 2003.

54. Wilcox, R. *Introduction to Robust Estimation and Hypothesis Testing*, 4th ed.; Elsevier/Academic Press: London, UK, 2016.