METRON (2013) 71:279–306
DOI 10.1007/s40300-013-0025-9

Weighted Least Squares and Least Median Squares estimation for the fuzzy linear regression analysis

Pierpaolo D'Urso · Riccardo Massari

Received: 6 September 2013 / Accepted: 11 October 2013 / Published online: 23 October 2013

Abstract In this paper, we discuss the problem of regression analysis in a fuzzy domain. By considering an iterative Weighted Least Squares estimation approach, we propose a general linear regression model for studying the dependence of a general class of fuzzy response variable, i.e., LR_2 fuzzy variable or trapezoidal fuzzy variable,on a set of crisp or LR_2 fuzzy explanatory variables. We also show some theoretical properties and a suitable generalization of the determination coefficient in order to investigate the goodness of fit of the regression model. Furthermore, we discuss some theoretical issues and an assessment of imprecision of the regression function. Finally, we suggest a robust version of the fuzzy regression model based on the Least Median Squares estimation approach which is able to neutralize and/or smooth the disruptive effects of possible crisp or fuzzy outliers in the estimation process. A simulation study and two empirical applications are presented.

Keywords Fuzzy input/output data · Fuzzy linear regression analysis · Robust fuzzy linear regression · Weighted Least Squares (WLS) · Least Median Squares (LMS)

1 Introduction

Linear regression model is a widely used statistics tool to evaluate the linear relationship between a quantitative dependent variable (output, or response variable), and one or more explanatory variables (inputs).

In linear regression modeling, two main issues have to be dealt with in practical problems:

1. imprecision or vagueness in the definition and/or observation of output and/or of inputs (see [10] for a more detailed discussion about the overall sources of uncertainty which may affect regression analysis);

P. D'Urso (✉) · R. Massari
Dipartimento di Scienze Sociali ed Economiche, Sapienza-Università di Roma,
P.za Aldo Moro, 5, 00185 Rome, Italy
e-mail: pierpaolo.durso@uniroma1.it

2. presence of outliers, which could cause the estimates of the regression coefficients to be bias.

As for the first issue, data imprecision may be due to several causes: (i) imprecision in measuring the empirical phenomena observed; (ii) vagueness of the variables of interest (inputs and/or outputs) when they are expressed in linguistic terms; (iii) partial or total ignorance about the variables' values on specific instances; (iv) granularity (categorization) of the variables of interest. When dealing with one or more of these situations, a fuzzification of the inputs and/or the output could suitably exploit the available information. Converting imprecise data into fuzzy data could be more effective than replacing them with a single value.

In this paper, imprecise data are then represented by fuzzy statistical variables.

The second issue regards the robustness of the estimates in a noisy environment. The Least Squares (LS) approach is one of the most popular methods for estimating linear regression coefficients, due to its theoretical and applicative advantages. However, the LS approach is not robust to the presence of outliers. This shortcoming of the LS approach undermines its application, even in presence of a small percentage of anomalous observations. In this paper, we cope with this issue by considering a robust estimation method, which is effective in reducing the distorting effect of outliers.

Following an iterative Weighted Least Squares (WLS) estimation approach, we propose a linear regression model for studying the dependence of a general class of fuzzy linear variables on a set of crisp or fuzzy explanatory variables (see Sect. 2). The proposed model represents a generalization of the fuzzy regression model suggested by Coppi et al. [10].

In order to investigate the goodness of fit of the regression model, some theoretical properties and a suitable generalization of the determination coefficient are described (Sect. 2).

Furthermore, we illustrate an assessment of the imprecision associated with the estimates of the regression coefficients of the proposed regression model (Sect. 3).

We then suggest a robust version of the fuzzy regression model based on the Least Median Squares (LMS) estimation approach, that is able to neutralize and/or smooth disruptive effects of possible crisp or fuzzy outliers in the estimation process (Sect. 4). The proposed robust model is a generalization of the robust model proposed by D'Urso et al. [20].

In order to illustrate the good performance of our model a simulation study and two empirical applications are presented (Sects. 5 and 6).

Some final remarks conclude the paper.

2 The linear regression model for LR_2 fuzzy inputs and output

Consider a fuzzy regression model with fuzzy/crisp output and fuzzy/crisp input.

Based on the traditional inferential approach, the expected value of the fuzzy/crisp output should be reparameterized in terms of a linear model involving the "regression effects" of p fuzzy/crisp explanatory variables. In literature, different theoretical contributions have been proposed for the fuzzy regression analysis based on a conjoint approach (inferential and fuzzy) (see, e.g., [2,3,22,23,25,26,29,32,36]).

The methodological approach considered in our paper is to select, among a class of possible linear models (expressing the relationship between the fuzzy/crisp output and the fuzzy/crisp inputs), the "best" linear model according to some specific criteria.

Following this approach, two main lines of research could be pursued in literature, the Possibilistic approach firstly introduced by Tanaka et al. [34], and the Least Squares (LS)

approach, based on suitable extensions of the well-known least squares criterion to the fuzzy setting (see, among others, [4,6,9–11,14–16,18,20,28,35]).

In the possibilistic framework, the fuzzy regression coefficients of a regression model are estimated by minimizing the fuzziness of the estimated response variable, conditionally on obtaining fuzzy response values which contain (to a certain possibility degree $0 \leq h \leq 1$) the observed fuzzy responses (see, for instance, [33]; and, in a comparative perspective, [5,12,24]).

In the LS approach the objective is to find the linear model which "best approximates" the observed data in a given metric space. The LS criterion is then conditional on the chosen metric.

Two main features characterize the adopted line of research, i.e., (a) the definition of the linear regression model, and (b) the specific metric space introduced for applying the LS criterion.

As for the first aspect, we will extend the linear regression models proposed by Coppi et al. [10] and D'Urso et al. [20] to the case when both the output and the inputs are fuzzy, in particular, fuzzy LR_2 variables (see the following Sect. 2.1), by setting up a procedure for estimating the two centers and the spreads of the regression coefficients. As for the second aspect, we will extend the distance function introduced by Coppi et al. [10] in our framework. Notice that our fuzzy regression models, which are explained in the following sections, are based on an exploratory approach.

2.1 Fuzzy data

We formalize imprecise data as LR_2 fuzzy data. In particular, LR_2 fuzzy data can be represented as $\tilde{y} \equiv (m_1, m_2, l, r)_{LR_2}$, where $m_1,\ m_2 \in \mathbb{R}\ (m_1 \leq m_2)$ denotes the centers, or the "modes", of the fuzzy data, while $l,\ r \in \mathbb{R}^+$ are the left and right spread, respectively, with the following membership functions [13,37]:

$$\mu_{\tilde{y}}(\omega) = \begin{cases} L\left(\frac{m_1-\omega}{l}\right) & \omega \leq m_1 \quad (l > 0) \\ 1 & m_1 \leq \omega \leq m_2 \\ R\left(\frac{\omega-m_2}{r}\right) & \omega \geq m_2 \quad (r > 0) \end{cases} \tag{1}$$

where L (and R) is a decreasing "shape" function from $\mathbb{R}^+$ to $[0, 1]$ with $L(0) = 1$; $L(\omega) < 1$ for all $\omega > 0$; $L(\omega) > 0$ for all ω; $L(1) = 0$(or $L(\omega) > 0$ for all ω and $L(+\infty) = 0$).

If $l = r$, we obtain symmetrical LR_2 fuzzy data.

When $m_1 = m_2 = m$ we obtain the LR_1 fuzzy data, which has the following membership function:

$$\mu_{\tilde{y}}(\omega) = \begin{cases} L\left(\frac{m-\omega}{l}\right) & \omega \leq m \quad (l > 0) \\ R\left(\frac{\omega-m}{r}\right) & \omega \geq m \quad (r > 0) \end{cases} \tag{2}$$

A particular case of LR_2 fuzzy data is the trapezoidal fuzzy data, whose membership function is:

$$\mu_{\tilde{y}}(\omega) = \begin{cases} 1 - \frac{m_1-\omega}{l} & m_1 - l \leq \omega \leq m_1 \quad (l > 0) \\ 1 & m_1 \leq \omega \leq m_2 \\ 1 - \frac{\omega-m_2}{r} & m_2 \leq \omega \leq m_2 + r \quad (r > 0) \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

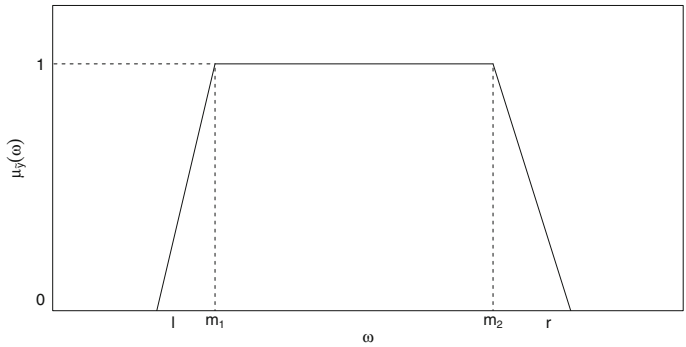

Fig. 1 Geometric representation of the trapezoidal membership function

Figure 1 shows the membership function of a trapezoidal fuzzy datum.

A particular case of LR_1 fuzzy data is the triangular fuzzy data, with the following membership function:

$$\mu_{\tilde{y}}(\omega) = \begin{cases} 1 - \frac{m-\omega}{l} & m - l \leq \omega \leq m \quad (l > 0) \\ 1 - \frac{\omega - m}{r} & m \leq \omega \leq m + r \quad (r > 0) \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

2.2 Model definition and estimation

Consider the linear dependence relationship between a LR_2 fuzzy output (or response variable) $\widetilde{Y} \equiv (m_1, m_2, l, r)$ and a set of p LR_2 fuzzy inputs $\{\widetilde{X}_j \equiv ({}_x m_{1j}, {}_x m_{2j}, {}_x l_j, {}_x r_j) : j = 1, \ldots, p\}$.

The proposed fuzzy linear regression model consists of modeling simultaneously the two centers of the LR_2 response variable by means of a multiple regression model on the LR_2 explanatory variables, and the left and right spreads of the response through two multiple linear regressions on the estimated centers.

Hence, the linear regression model with fuzzy response variable $\widetilde{Y}$ and fuzzy explanatory variables $\widetilde{X}_j,\ j = 1, \ldots, p$, can be formalized as follows, using a matrix notation:

$$\mathbf{m}_1 = \mathbf{m}_1^* + \boldsymbol{\varepsilon}_{\mathbf{m}_1} \quad \mathbf{m}_1^* = \mathbf{M}_1\boldsymbol{\alpha}_1 + \mathbf{M}_2\boldsymbol{\alpha}_2 + \mathbf{L}\boldsymbol{\alpha}_l + \mathbf{R}\boldsymbol{\alpha}_r \tag{5a}$$

$$\mathbf{m}_2 = \mathbf{m}_2^* + \boldsymbol{\varepsilon}_{\mathbf{m}_2} \quad \mathbf{m}_2^* = \mathbf{M}_1\boldsymbol{\beta}_1 + \mathbf{M}_2\boldsymbol{\beta}_2 + \mathbf{L}\boldsymbol{\beta}_l + \mathbf{R}\boldsymbol{\beta}_r \tag{5b}$$

$$\mathbf{l} = \mathbf{l}^* + \boldsymbol{\varepsilon}_{\mathbf{l}} \quad \mathbf{l}^* = \mathbf{1}\gamma_0 + \mathbf{m}_1^*\gamma_1 + \mathbf{m}_2^*\gamma_2 = \mathbf{M}^*\boldsymbol{\gamma} \tag{5c}$$

$$\mathbf{r} = \mathbf{r}^* + \boldsymbol{\varepsilon}_{\mathbf{r}} \quad \mathbf{r}^* = \mathbf{1}\delta_0 + \mathbf{m}_1^*\delta_1 + \mathbf{m}_2^*\delta_2 = \mathbf{M}^*\boldsymbol{\delta} \tag{5d}$$

where $\mathbf{m}_1$, $\mathbf{m}_2$ are the n-vectors of the left and right centers of the response fuzzy variables, m_1, m_2; $\mathbf{l}$, $\mathbf{r}$ are the n-vectors of the left and right spreads of the response fuzzy variables, l, r; $\mathbf{M}_1$, $\mathbf{M}_2$ are the $(n \times (p+1))$-matrices of the left and right centers of the input fuzzy variables (design matrices), ${}_x m_{1j}$, ${}_x m_{2j}$ $(j = 1, \ldots, p)$, respectively; $\mathbf{L}$, $\mathbf{R}$ are the $(n \times (p+1))$-matrices of the left and right spreads of the input fuzzy variables, ${}_x l_j$, ${}_x r_j$ $(j = 1, \ldots, p)$, respectively[1]; $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\boldsymbol{\alpha}_l$, $\boldsymbol{\alpha}_r$ are the $(p+1)$-vectors of coefficients of the models on the left centers $\mathbf{m}_1$; $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_l$, $\boldsymbol{\beta}_r$ are the $(p+1)$-vectors of coefficients of the models on the right centers $\mathbf{m}_2$; γ_0, γ_1 and γ_2 are the coefficients of the model on the left spreads, $\mathbf{l}$; δ_0, δ_1 and δ_2

[1] $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{L}$, and $\mathbf{R}$ also contain a vector of ones, related to the intercepts of the model.

are their counterparts for the model on the right spreads, $\mathbf{r}$; $\boldsymbol{\varepsilon}_{\mathbf{m}_1}$, $\boldsymbol{\varepsilon}_{\mathbf{m}_2}$ are the n-vectors of the error terms of the models on the left and right centers, respectively; $\boldsymbol{\varepsilon}_{\mathbf{l}}$, $\boldsymbol{\varepsilon}_{\mathbf{r}}$ are the n-vectors of the error terms of the models on the left and right spreads, respectively; $\mathbf{1}$ is the n-vector of ones. The theoretical values of the centers, and of the spreads are marked with an asterisk symbol (*). Finally, $\mathbf{M}^*$ is the $(n \times 3)$ matrix whose columns are the vector of ones and the vectors of the theoretical values of the left and right centers of the response variable; $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ are the (3×1) vectors of the coefficients of the model on the left and right spreads, respectively.

Note that in the model (5a)–(5d) we assume that the estimates of both spreads depend on the estimates of both centers (see Eqs. (5c) and (5d)). We can interpret the left and right center of the dependent variable as the lower and the upper bound, respectively, of interval-valued data, and the spreads as the degree of imprecision of these interval-valued data. Hence, the assumption of linear dependency between spreads and centers is reasonable since in many instances the magnitude of the error depends on the size of the interval estimates.

We use the Weighted Least Squares (WLS) procedure to estimate the coefficients of the model [20]. In what follows, we refer to the model (5a)–(5d) as the WLS-based fuzzy regression model. Depending on the nature of the weighing matrix $\mathbf{W}$, we have different linear regression models. In particular, when $\mathbf{W} = \mathbf{I}$, we obtain the Least Squares (LS) based fuzzy regression model.

The objective function to be minimized is the weighted squared Euclidean distance between the observed fuzzy variables and their estimates, $\widetilde{\Delta}^2_{\mathbf{W}}$ [9].

Let $\|\mathbf{x}\|_{\mathbf{W}} = (\mathbf{x}'\mathbf{W}\mathbf{x})^{\frac{1}{2}}$ be the weighted norm of the generic vector $\mathbf{x}$, where $\mathbf{W}$ is a diagonal matrix, whose elements are the weights attached to each observation. Then, the weighted squared Euclidean distance $\widetilde{\Delta}^2_{\mathbf{W}}$ can be written as:

$$\begin{aligned}\widetilde{\Delta}^2_{\mathbf{W}} &= \|\mathbf{m}_1 - \mathbf{m}_1^*\|^2_{\mathbf{W}} + \|\mathbf{m}_2 - \mathbf{m}_2^*\|^2_{\mathbf{W}} \\ &\quad + \|(\mathbf{m}_1 - \lambda\mathbf{l}) - (\mathbf{m}_1^* - \lambda\mathbf{l}^*)\|^2_{\mathbf{W}} + \|(\mathbf{m}_2 + \rho\mathbf{r}) - (\mathbf{m}_2^* + \rho\mathbf{r}^*)\|^2_{\mathbf{W}} \end{aligned} \tag{6}$$

where $\lambda = \int_0^1 L^{-1}(\omega)\, d\omega$ and $\rho = \int_0^1 R^{-1}(\omega)\, d\omega$ are parameters which account for the shape of the membership function. In particular, if the membership function is trapezoidal, then $\lambda = \rho = 1/2$ [9].

Equation (6) can be developed as follows:

$$\begin{aligned}\widetilde{\Delta}^2_{\mathbf{W}} &= (\mathbf{m}_1 - \mathbf{m}_1^*)'\mathbf{W}(\mathbf{m}_1 - \mathbf{m}_1^*) + (\mathbf{m}_2 - \mathbf{m}_2^*)'\mathbf{W}(\mathbf{m}_2 - \mathbf{m}_2^*) \\ &\quad + [(\mathbf{m}_1 - \lambda\mathbf{l}) - (\mathbf{m}_1^* - \lambda\mathbf{l}^*)]'\mathbf{W}[(\mathbf{m}_1 - \lambda\mathbf{l}) - (\mathbf{m}_1^* - \lambda\mathbf{l}^*)] \\ &\quad + [(\mathbf{m}_2 + \rho\mathbf{r}) - (\mathbf{m}_2^* + \rho\mathbf{r}^*)]'\mathbf{W}[(\mathbf{m}_2 + \rho\mathbf{r}) - (\mathbf{m}_2^* + \rho\mathbf{r}^*)] \\ &= (\mathbf{m}_1 - \mathbf{m}_1^*)'\mathbf{W}(\mathbf{m}_1 - \mathbf{m}_1^*) + (\mathbf{m}_2 - \mathbf{m}_2^*)'\mathbf{W}(\mathbf{m}_2 - \mathbf{m}_2^*) \\ &\quad + [(\mathbf{m}_1 - \mathbf{m}_1^*) - \lambda(\mathbf{l} - \mathbf{l}^*)]'\mathbf{W}[(\mathbf{m}_1 - \mathbf{m}_1^*) - \lambda(\mathbf{l} - \mathbf{l}^*)] \\ &\quad + [(\mathbf{m}_2 - \mathbf{m}_2^*) + \rho(\mathbf{r} - \mathbf{r}^*)]'\mathbf{W}[(\mathbf{m}_2 - \mathbf{m}_2^*) + \rho(\mathbf{r} - \mathbf{r}^*)] \\ &= 2\,(\mathbf{m}_1'\mathbf{W}\mathbf{m}_1 - 2\,\mathbf{m}_1'\mathbf{W}\mathbf{m}_1^* + \mathbf{m}_1^{*\prime}\mathbf{W}\mathbf{m}_1^* + \mathbf{m}_2'\mathbf{W}\mathbf{m}_2 - 2\,\mathbf{m}_2'\mathbf{W}\mathbf{m}_2^* + \mathbf{m}_2^{*\prime}\mathbf{W}\mathbf{m}_2^*) \\ &\quad - 2\,\lambda(\mathbf{m}_1'\mathbf{W}\mathbf{l} - \mathbf{m}_1'\mathbf{W}\mathbf{l}^* - \mathbf{m}_1^{*\prime}\mathbf{W}\mathbf{l} + \mathbf{m}_1^{*\prime}\mathbf{W}\mathbf{l}^*) + \lambda^2(\mathbf{l}'\mathbf{W}\mathbf{l} - 2\,\mathbf{l}'\mathbf{W}\mathbf{l}^* + \mathbf{l}^{*\prime}\mathbf{W}\mathbf{l}^*) \\ &\quad + 2\,\rho(\mathbf{m}_2'\mathbf{W}\mathbf{r} - \mathbf{m}_2'\mathbf{W}\mathbf{r}^* - \mathbf{m}_2^{*\prime}\mathbf{W}\mathbf{r} + \mathbf{m}_2^{*\prime}\mathbf{W}\mathbf{r}^*) + \rho^2(\mathbf{r}'\mathbf{W}\mathbf{r} - 2\,\mathbf{r}'\mathbf{W}\mathbf{r}^* + \mathbf{r}^{*\prime}\mathbf{W}\mathbf{r}^*). \end{aligned} \tag{7}$$

By minimizing (7), we obtain the iterative solutions of the model (5a)–(5d), which are reported in Appendix.

2.3 Properties of the model

In this section, we illustrate some properties of the WLS-based fuzzy regression model (5a)–(5d), which will be useful in the following.

Proposition 1 *The weighted sums of the residuals of the left and right centers and of the left and right spreads are equal to 0:*

$$\begin{aligned}\mathbf{1}'\mathbf{W}(\mathbf{m}_1 - \hat{\mathbf{m}}_1) &= 0\\ \mathbf{1}'\mathbf{W}(\mathbf{m}_2 - \hat{\mathbf{m}}_2) &= 0\\ \mathbf{1}'\mathbf{W}(\mathbf{l} - \hat{\mathbf{l}}) &= 0\\ \mathbf{1}'\mathbf{W}(\mathbf{r} - \hat{\mathbf{r}}) &= 0\end{aligned}$$

where $\hat{\mathbf{m}}_1, \hat{\mathbf{m}}_2, \hat{\mathbf{l}}$ *and* $\hat{\mathbf{r}}$ *are the estimates of the left and right centers, and of the left and right spreads of the respondent variable.*

From this proposition we also derive that the weighted mean of the residuals is equal to 0.

Proposition 2 *The residuals of the left and right centers are uncorrelated with the estimates of the left and right centers, respectively:*

$$\begin{aligned}(\mathbf{m}_1 - \hat{\mathbf{m}}_1)'\mathbf{W}\hat{\mathbf{m}}_1 &= 0\\ (\mathbf{m}_2 - \hat{\mathbf{m}}_2)'\mathbf{W}\hat{\mathbf{m}}_2 &= 0\end{aligned}$$

Similarly, the residuals of the left and right spread are uncorrelated with the estimates of the left and right spreads, respectively:

$$\begin{aligned}(\mathbf{l} - \hat{\mathbf{l}})'\mathbf{W}\hat{\mathbf{l}} &= 0\\ (\mathbf{r} - \hat{\mathbf{r}})'\mathbf{W}\hat{\mathbf{r}} &= 0\end{aligned}$$

Note that, given the relationship between the sub-models in (5a)–(5d), it follows that:

$$\begin{aligned}(\mathbf{m}_1 - \hat{\mathbf{m}}_1)'\mathbf{W}\hat{\mathbf{l}} &= 0\\ (\mathbf{m}_2 - \hat{\mathbf{m}}_2)'\mathbf{W}\hat{\mathbf{r}} &= 0\\ (\mathbf{l} - \hat{\mathbf{l}})'\mathbf{W}\hat{\mathbf{m}}_1 &= 0\\ (\mathbf{r} - \hat{\mathbf{r}})'\mathbf{W}\hat{\mathbf{m}}_2 &= 0\end{aligned}$$

Proofs for Propositions 1–2 can be easily derived from the LS properties proved in Coppi et al. [10].

2.4 Goodness of fit

To evaluate the goodness of fit of the model (5a)–(5d) to the data, we propose a generalization of the determination coefficient R^2 for fuzzy regression models suggested by Coppi et al. [10].

First, define the following quantities:

– the *total weighted sum of squares*:

$$\begin{aligned}SST_{\mathbf{W}} = &\|\mathbf{m}_1 - \mathbf{1}\bar{m}_1\|^2_{\mathbf{W}} + \|\mathbf{m}_2 - \mathbf{1}\bar{m}_2\|^2_{\mathbf{W}}\\ &+\|(\mathbf{m}_1 - \lambda\mathbf{l}) - (\mathbf{1}\bar{m}_1 - \lambda\mathbf{1}\bar{l})\|^2_{\mathbf{W}} + \|(\mathbf{m}_2 + \rho\mathbf{r}) - (\mathbf{1}\bar{m}_2 + \rho\mathbf{1}\bar{r})\|^2_{\mathbf{W}}, \quad (8)\end{aligned}$$

– the *weighted explained sum of squares*:

$$SSE_{\mathbf{W}} = \|\hat{\mathbf{m}}_1 - \mathbf{1}\bar{m}_1\|^2_{\mathbf{W}} + \|\hat{\mathbf{m}}_2 - \mathbf{1}\bar{m}_2\|^2_{\mathbf{W}} + \|(\hat{\mathbf{m}}_1 - \lambda\hat{\mathbf{l}}) - (\mathbf{1}\bar{m}_1 - \lambda\mathbf{1}\bar{l})\|^2_{\mathbf{W}} + \|(\hat{\mathbf{m}}_2 + \rho\hat{\mathbf{r}}) - (\mathbf{1}\bar{m}_2 + \rho\mathbf{1}\bar{r})\|^2_{\mathbf{W}}, \quad (9)$$

– the *weighted residual sum of squares*:

$$SSR_{\mathbf{W}} = \|\mathbf{m}_1 - \hat{\mathbf{m}}_1\|^2_{\mathbf{W}} + \|\mathbf{m}_2 - \hat{\mathbf{m}}_2\|^2_{\mathbf{W}} + \|(\mathbf{m}_1 - \lambda\mathbf{l}) - (\hat{\mathbf{m}}_1 - \lambda\hat{\mathbf{l}})\|^2_{\mathbf{W}} + \|(\mathbf{m}_2 + \rho\mathbf{r}) - (\hat{\mathbf{m}}_2 + \rho\hat{\mathbf{r}})\|^2_{\mathbf{W}}, \quad (10)$$

where $\bar{m}_1, \bar{m}_2, \bar{l}$ and $\bar{r}$ are the sample means of the left and right centers and of the left and right spreads, respectively.

Based on the properties illustrated in Sect. 2.3, it can be shown that:

$$SST_{\mathbf{W}} = SSE_{\mathbf{W}} + SSR_{\mathbf{W}}. \quad (11)$$

Then, the determination coefficient for the weighted fuzzy linear regression model is defined as:

$$R^2_{\mathbf{W}} = \frac{SSE_{\mathbf{W}}}{SST_{\mathbf{W}}} = 1 - \frac{SSR_{\mathbf{W}}}{SST_{\mathbf{W}}}, \quad 0 \le R^2_{\mathbf{W}} \le 1. \quad (12)$$

As in the standard linear regression framework, the closer $R^2_{\mathbf{W}}$ approaches 1, the better the fit of the model to the data.

The analysis of the goodness of fit of a model is useful when one wants to select the model which provides the best fit to the data, in a class of parametric models.

However, it can be shown that $R^2_{\mathbf{W}}$ is not decreasing as the number of inputs in the model increases. For this reason, if the objective is to select the "best" model in a class of models, then $R^2_{\mathbf{W}}$ could be ineffective.

A better solution is to adopt the adjusted determination coefficient $\bar{R}^2_{\mathbf{W}}$, which adds a penalization term that takes into account the number of inputs. We indicate the number of parameters of the fuzzy regression model with $\bar{p}$. In particular, when both inputs and output are LR_2 fuzzy variables $\bar{p} = [8 \cdot (p + 1) + 6]$. Then, the adjusted determination coefficient is:

$$\bar{R}^2_{\mathbf{W}} = 1 - (1 - R^2_{\mathbf{W}})\frac{n - 1}{n - \bar{p}}. \quad (13)$$

$\bar{R}^2_{\mathbf{W}}$ increases only if the inclusion of a new input improves $R^2_{\mathbf{W}}$ more than would be expected by chance. The adjusted determination coefficient can be used to select the optimal number of inputs to be included in the model.

As observed by Coppi et al. [10], the denominator of the adjusting factor $(n - \bar{p})$ in (13) decreases more than proportionally as p increases, thus penalizing the model with $p + 1$ variables in a more severe way than the traditional (crisp) version of the adjusted determination coefficient. Then, it would be better to use an adjusting factor in which we consider only the number of parameters of one of the centers-model, p.

2.5 Some remarks

Remark 1 (*Generalization of the design matrices*) The design matrices $\mathbf{M}_1, \mathbf{M}_2, \mathbf{L}$ and $\mathbf{R}$ in the model (5a)–(5d) can be generalized by considering appropriate functions of the components of the fuzzy explanatory variables $\tilde{\mathbf{X}}$.

Let be $\mathbf{F}_1, \mathbf{F}_2, \mathbf{F}_l$ and $\mathbf{F}_r$ the "transformed" design matrices, where:

$$\mathbf{f}'_{1i} = [f_1({}_x\mathbf{m}_{1i}), \ldots, f_p({}_x\mathbf{m}_{1i})]$$
$$\mathbf{f}'_{2i} = [f_1({}_x\mathbf{m}_{2i}), \ldots, f_p({}_x\mathbf{m}_{2i})]$$
$$\mathbf{f}'_{li} = [f_1({}_x\mathbf{l}_i), \ldots, f_p({}_x\mathbf{l}_i)]$$
$$\mathbf{f}'_{ri} = [f_1({}_x\mathbf{r}_i), \ldots, f_p({}_x\mathbf{r}_i)]$$

are the generic rows of the transformed design matrices. Each row represents the regression "profile" of observation i in terms of suitably chosen functions of the observed vectors of the fuzzy explanatory variables. In this way, the model allows also for transformation of the original fuzzy variables, like the polynomial or the logarithmic transformations. By substituting $\mathbf{F}_1, \mathbf{F}_2, \mathbf{F}_l$ and $\mathbf{F}_r$ in (5a)–(5d) the properties of the model proposed can be easily extended to this more general case.

Remark 2 (*Local optima issues*) As for other iterative estimation algorithm, the solutions of the model (5a)–(5d) (see Appendix) do not guarantee the attainment of the global minimum. For this reason we initialize the iterative algorithm considering several different starting points in order to check the stability of the solution.

Remark 3 (*Negative spreads*) The iterative solutions of the model (5a)–(5d) (see Appendix) do not automatically guarantee the non negativity of the estimated spreads $\mathbf{l}^*$ and $\mathbf{r}^*$. To cope with this issue, one can adopt the approaches proposed by D'Urso [14]. In particular, among the different approaches for guaranteeing the non-negativity of the estimated spreads proposed by D'Urso [14] there is the so-called "unconstrained approach", in which a logarithmic transformation of the spreads is suggested (for more details, see [14]). In literature, this approach has been particularly successful and has been used afterwards by various authors in fuzzy-exploratory and fuzzy-inferential frameworks. For instance, Ferraro et al. [21]—following the idea of considering a modeling structure based on three sub-models proposed by D'Urso and Gastaldi [15] and D'Urso [14] and using the least-squares approach as in Coppi and D'Urso [9] and Coppi et al. [10]—formalized a linear regression model in a fuzzy-inferential framework using the logarithmic transformation of the spreads of the response as suggested in D'Urso [14] within an exploratory framework.

Remark 4 (*Particular cases of the model* (5a)–(5d)) The model (5a)–(5d) can be considered as the most general fuzzy regression model, with LR_2 fuzzy inputs and outputs. By combining

Table 1 Regression models with fuzzy/crisp output/inputs and mixed membership functions

Output	Input		
	Crisp	LR_1	LR_2
Crisp	$\mathbf{y}^* = \mathbf{X}\boldsymbol{\alpha}$	$\mathbf{y}^* = \mathbf{M}\boldsymbol{\alpha}_1 + \mathbf{L}\boldsymbol{\alpha}_l + \mathbf{R}\boldsymbol{\alpha}_r$	$\mathbf{y}^* = \mathbf{M}_1\boldsymbol{\alpha}_1 + \mathbf{M}_2\boldsymbol{\alpha}_2 + \mathbf{L}\boldsymbol{\alpha}_l + \mathbf{R}\boldsymbol{\alpha}_r$
LR_1	$\mathbf{m}^* = \mathbf{X}\boldsymbol{\alpha}$	$\mathbf{m}^* = \mathbf{M}\boldsymbol{\alpha} + \mathbf{L}\boldsymbol{\alpha}_l + \mathbf{R}\boldsymbol{\alpha}_r$	$\mathbf{m}^* = \mathbf{M}_1\boldsymbol{\alpha}_1 + \mathbf{M}_2\boldsymbol{\alpha}_2 + \mathbf{L}\boldsymbol{\alpha}_l + \mathbf{R}\boldsymbol{\alpha}_r$
	$\mathbf{l}^* = \mathbf{M}^*\boldsymbol{\gamma}$	$\mathbf{l}^* = \mathbf{M}^*\boldsymbol{\gamma}$	$\mathbf{l}^* = \mathbf{M}^*\boldsymbol{\gamma}$
	$\mathbf{r}^* = \mathbf{M}^*\boldsymbol{\delta}$	$\mathbf{r}^* = \mathbf{M}^*\boldsymbol{\delta}$	$\mathbf{r}^* = \mathbf{M}^*\boldsymbol{\delta}$
LR_2	$\mathbf{m}_1^* = \mathbf{X}\boldsymbol{\alpha}$	$\mathbf{m}_1^* = \mathbf{M}\boldsymbol{\alpha} + \mathbf{L}\boldsymbol{\alpha}_l + \mathbf{R}\boldsymbol{\alpha}_r$	$\mathbf{m}_1^* = \mathbf{M}_1\boldsymbol{\alpha}_1 + \mathbf{M}_2\boldsymbol{\alpha}_2 + \mathbf{L}\boldsymbol{\alpha}_l + \mathbf{R}\boldsymbol{\alpha}_r$
	$\mathbf{m}_2^* = \mathbf{X}\boldsymbol{\beta}$	$\mathbf{m}_2^* = \mathbf{M}\boldsymbol{\beta} + \mathbf{L}\boldsymbol{\beta}_l + \mathbf{R}\boldsymbol{\beta}_r$	$\mathbf{m}_2^* = \mathbf{M}_1\boldsymbol{\beta}_1 + \mathbf{M}_2\boldsymbol{\beta}_2 + \mathbf{L}\boldsymbol{\beta}_l + \mathbf{R}\boldsymbol{\beta}_r$
	$\mathbf{l}^* = \mathbf{M}^*\boldsymbol{\gamma}$	$\mathbf{l}^* = \mathbf{M}^*\boldsymbol{\gamma}$	$\mathbf{l}^* = \mathbf{M}^*\boldsymbol{\gamma}$
	$\mathbf{r}^* = \mathbf{M}^*\boldsymbol{\delta}$	$\mathbf{r}^* = \mathbf{M}^*\boldsymbol{\delta}$	$\mathbf{r}^* = \mathbf{M}^*\boldsymbol{\delta}$

different typologies of membership functions for the fuzzy/crisp inputs/outputs, we obtain different (fuzzy) regression models, outlined in Table 1.

For instance, by putting $\mathbf{M}_1 = \mathbf{M}_2 = \mathbf{X}, \mathbf{L} = \mathbf{R} = \mathbf{0}, \boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2 = \boldsymbol{\alpha}$ and $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \boldsymbol{\beta}$ one can easily obtain from (5a)–(5d) the fuzzy regression model with crisp inputs and LR_2 fuzzy output, and from the iterative solutions, reported in Appendix, the corresponding coefficients' estimates.

The models in Table 1 can be also generalized to the case in which the fuzzy output and/or the fuzzy inputs are symmetrical.

3 Assessment of imprecision of the regression function

As observed by Coppi et al. [10] in the case of crisp inputs and LR_1 fuzzy output, the estimation procedure of the fuzzy linear regression model provides a crisp evaluation of the regression coefficients. Since the response variable is fuzzy, the fuzzy linear regression model implicitly involves a fuzzy regression model expressed in terms of fuzzy regression coefficients. Thus, the crisp estimates of the fuzzy regression model involve a certain degree of imprecision. This observation can be extended also to all the models reported in Table 1 with fuzzy output, and in particular to our proposed model (5a)–(5d).

To evaluate the imprecision due to the crisp estimates of the fuzzy regression model, we exploit the "implicit" fuzzy model with fuzzy regression coefficients.

Following a similar line of reasoning as in Coppi et al. [10], we refer to the case with LR_2 response variable and crisp inputs, but our conclusions can be extended to more complex models.

The fuzzy regression model with LR_2 fuzzy response variable and crisp explanatory variables is:

$$\begin{aligned} \mathbf{m}_1^* &= \mathbf{X}\boldsymbol{\alpha} \\ \mathbf{m}_2^* &= \mathbf{X}\boldsymbol{\beta} \\ \mathbf{l}^* &= \mathbf{M}^*\boldsymbol{\gamma} \\ \mathbf{r}^* &= \mathbf{M}^*\boldsymbol{\delta} \end{aligned} \tag{14}$$

where, $\mathbf{M}^* = (\mathbf{1}, \mathbf{m}_1^*, \mathbf{m}_2^*)$, $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \gamma_2)'$ and $\boldsymbol{\delta} = (\delta_0, \delta_1, \delta_2)'$.

The implicit fuzzy model can be expressed as:

$$\tilde{y}_i^* = \tilde{\beta}_0 + \tilde{\beta}_1 x_{i1} \oplus \ldots \oplus \tilde{\beta}_p x_{ip}, \quad i = 1, \ldots, n \tag{15}$$

where: $\tilde{y}_i^* = (m_{1i}^*, m_{2i}^*, l_i^*, r_i^*)$ is the theoretical value of the LR_2 fuzzy response variable for the i-th unit; the coefficient $\tilde{\beta}_k = (\beta_{1k}, \beta_{2k}, \beta_{lk}, \beta_{rk})$, $k = 1, \ldots, p$ is a LR_2 fuzzy number, with the four components being the left and right centers, and the left right spread of the k-th coefficient, respectively; $\oplus$ denotes the addition of fuzzy numbers.

We express the model (15) in the following way:

$$\begin{aligned} m_{1i}^* &= \beta_{10} + \beta_{11} x_{i1} + \cdots + \beta_{1p} x_{ip} \quad \mathbf{m}_1^* = \mathbf{X}\boldsymbol{\beta}_1 \\ m_{2i}^* &= \beta_{20} + \beta_{21} x_{i1} + \cdots + \beta_{2p} x_{ip} \quad \mathbf{m}_2^* = \mathbf{X}\boldsymbol{\beta}_2 \\ l_i^* &= \beta_{l0} + \beta_{l1} |x_{i1}| + \cdots + \beta_{lp} |x_{ip}| \quad \mathbf{l}^* = |\mathbf{X}|\boldsymbol{\beta}_l \\ r_i^* &= \beta_{r0} + \beta_{r1} |x_{i1}| + \cdots + \beta_{rp} |x_{ip}| \quad \mathbf{r}^* = |\mathbf{X}|\boldsymbol{\beta}_r \end{aligned} \tag{16}$$

where $|\mathbf{X}|$ is the matrix of the absolute values of the inputs, $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_l$ and $\boldsymbol{\beta}_r$ are the $(p+1)$ vectors of the components of the vector of the fuzzy coefficients $\tilde{\beta}_k$.

Coppi et al. [10] observed that from (16) we obtain estimates of $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_l$ and $\boldsymbol{\beta}_r$ which are compatible with $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$, the coefficient of the model (14).

Let assume the fuzzy arithmetic relationships represented by the equations in (16) can be approximated as follows:

$$\begin{aligned} \mathbf{m}_1^{*p} &= \mathbf{X}\boldsymbol{\beta}_1 + \mathbf{u}_1 \\ \mathbf{m}_2^{*p} &= \mathbf{X}\boldsymbol{\beta}_2 + \mathbf{u}_2 \\ \mathbf{l}^{*p} &= |\mathbf{X}|\boldsymbol{\beta}_l + \mathbf{u}_l \\ \mathbf{r}^{*p} &= |\mathbf{X}|\boldsymbol{\beta}_r + \mathbf{u}_l \end{aligned} \tag{17}$$

where p indicates that these relationships are proxies of the real relationships, and where $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_l$ and $\mathbf{u}_r$ are vectors of residuals.

By means of Ordinary Least Squares (OLS), we obtain compatible estimate of the model (16). For instance, the OLS estimate of $\boldsymbol{\beta}_1$ is:

$$\boldsymbol{\beta}_1 = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{m}_1^{*p} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\alpha}}_1 = \hat{\boldsymbol{\alpha}}_1$$

where $\hat{\boldsymbol{\alpha}}_1$ is the LS estimate of $\boldsymbol{\alpha}_1$ from the model (14). In a similar way, we obtain a compatible estimate of $\boldsymbol{\beta}_2$.

As for the estimates of the spreads $\boldsymbol{\beta}_l$ and $\boldsymbol{\beta}_r$, one can adopt the non-negative Least Squares (NNLS) algorithm [27], to avoid negative estimates.

In conclusion, the model (14) provides both a good approximation of the centers of the fuzzy regression coefficients $\tilde{\beta}$ and of the fuzzy values of the fuzzy response variable, $\tilde{y}$. Moreover, by means of the LS approximation of the spreads in (17) we obtain reasonable estimates of the spreads of $\tilde{\beta}$.

Finally, note that another source of uncertainty in our framework is related to the data generation process [10]. One could take into account this type of uncertainty by means of the bootstrap procedure to evaluate the standard error of the regression coefficients estimates.

Results illustrated in this section could also be extended to the other fuzzy regression models shown in Table 1.

4 Robust fuzzy regression

It can be shown that the WLS-based fuzzy regression model (5a)–(5d) is generally not robust to the presence of outlier data.

For instance, when $\mathbf{W} = \mathbf{I}$, the WLS-based fuzzy regression model reduces itself to the LS-based fuzzy regression model, which is extremely sensitive to the presence of outliers, yielding a distortion in the parameter estimates [20].

In fuzzy regression, we could have different types of outliers in the dataset with respect to: one or more crisp explanatory variables; the centers of one or more fuzzy explanatory variables; the spreads of one or more fuzzy explanatory variables; the centers of the fuzzy dependent variable; the spreads of the fuzzy dependent variable; the fuzzy regression lines; more aspects.

We denote the data observed on a generic unit, where both output and p inputs are fuzzy LR_2 variables, with $(\tilde{y}_i, \tilde{\mathbf{x}}_i)$. Let us also suppose that the theoretical relationship between the fuzzy output $\tilde{y}_i$ and the fuzzy inputs $\tilde{\mathbf{x}}_i$ can be described by the model (5a)–(5d). A sample of such observations is depicted in Fig. 2a, in the case when there is a single fuzzy input ($p = 1$). Solid lines represent the interval-valued data given by the left and right centers of $\tilde{y}$ and $\tilde{x}$, dashed lines represents the spreads, i.e. the uncertainty, around the centers.

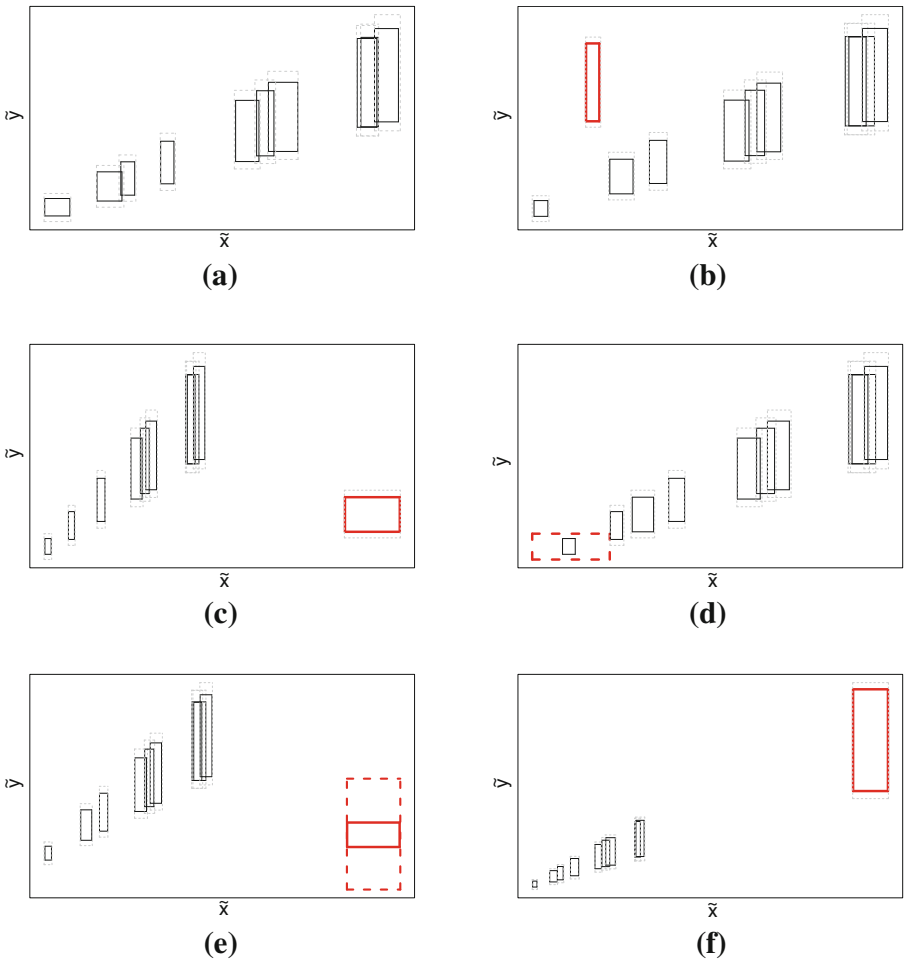

Fig. 2 Example of outliers for the (5a)–(5c). **a** No outlier. **b** Outlier with respect to the relationship between $\tilde{y}$ and $\tilde{x}$, but not with respect to the two variables. **c** Outlier with respect to the relationship between $\tilde{y}$ and $\tilde{x}$, and with respect to the centers of $\tilde{x}$. **d** Outlier with respect to the spreads of $\tilde{x}$. **e** Outlier with respect to the centers of $\tilde{x}$ to the spreads of $\tilde{y}$ and to the relationship between $\tilde{y}$ and $\tilde{x}$. **f** Outlier with respect to the centers of $\tilde{y}$ and $\tilde{x}$, but not with respect to the relationship between $\tilde{y}$ and $\tilde{x}$

As observed above, different types of outliers could occur in the dataset. Consider, for instance, Fig. 2b where there is an outlier (depicted with a bolder line) with respect to the relationship between $\tilde{y}$ and $\tilde{x}$. A closer inspection reveals that the unit is not an outlier with respect to the two fuzzy variables. Indeed, the value of the left and right centers (and of the left and right spreads) are in the range of the values observed for the other units.

In Fig. 2c we have a different case, since the anomalous unit is an outlier with respect to both the relationship between $\tilde{y}$ and $\tilde{x}$, and to the fuzzy input. As can be seen, the values of the left and right centers of $\tilde{x}$ lie outside the range of the left and right centers of the fuzzy input for the remaining units.

In Fig. 2d the outlier is with respect to the spreads of $\tilde{x}$. This outlier partially undermines also the relationship between $\tilde{y}$ and $\tilde{x}$, at least for the part of the model (5a)–(5d) devoted to the spreads.

Figure 2e shows a more general case in which the unit is an outlier with respect to the centers of $\tilde{x}$, the spreads of $\tilde{y}$ and the relationship between the fuzzy variables.

Finally, in Fig. 2f we show a situation in which the observation is an outlier with respect to both $\tilde{y}$ and $\tilde{x}$, but not with respect to their relationship.

In this section we propose a robust version of the fuzzy regression model (5a)–(5d). The proposed model is based on the Least Median Squares (LMS) estimation method [30], which relies on the minimization of the median of squared residuals:

$$\begin{aligned}\widetilde{\Delta}^2_{med} = \underset{i}{median}\,\big\{&(m_{1i} - m^*_{1i})^2 + (m_{2i} - m^*_{2i})^2 \\ &+[(m_{1i} - \lambda l_i) - (m^*_{1i} - \lambda l^*_i)]^2 + [(m_{2i} + \rho r_i) - (m^*_{2i} + \rho r^*_i)]^2\big\} \qquad (18)\end{aligned}$$

The two-steps estimation procedure can be illustrated as follows [20].

In the first step we apply a random re-sampling procedure [31], in which, we consider several subsets of $\bar{p} = [8 \cdot (p + 1) + 6]$ observations, where $\bar{p}$ is the number of unknown parameters of the model (5a)–(5d). As the number of subsets increases, the probability of extracting at least one subset without outliers, increases.

Let $\mathbf{M}_{1s}$, $\mathbf{M}_{2s}$, $\mathbf{L}_s$ and $\mathbf{R}_s$ be the $[\bar{p} \times (p + 1)]$ matrices extracted from the matrices $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{L}$ and $\mathbf{R}$ defined in Sect. 2.2, whose rows match up to the randomly selected observations. Let also $\mathbf{m}_{1s}$, $\mathbf{m}_{2s}$, $\mathbf{r}_s$ and $\mathbf{l}_s$ be the corresponding sub-vectors $(\bar{p} \times 1)$ of $\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{l}$ and $\mathbf{r}$, respectively.

For each subset, the regression coefficients are estimated using the iterative solutions illustrated in Appendix, by putting $\mathbf{W} = \mathbf{I}_{\bar{p}}$, thus obtaining the estimated values $\hat{\mathbf{m}}_{1s}$, $\hat{\mathbf{m}}_{2s}$, $\hat{\mathbf{l}}_s$ and $\hat{\mathbf{r}}_s$. These estimates are employed to compute the median of squared residuals (18).

Since the optimal solution of LMS employs only a subset of observations, it is likely that a great deal of the remaining observations are not outliers. Hence, in the second step of the procedure we improve the estimates by considering all observations, assigning low weights to data identified as outliers. The identification of these observations is based on the robust residuals from LMS (see [20]).

In particular, we adopt the following weights for our analysis:

$$w_i = \begin{cases} 1, & |r_i/\hat{\sigma}| \leq c_1 \\ 0.5, & c_1 \leq |r_i^2/\hat{\sigma}| \leq c_2 \\ 0, & |r_i/\hat{\sigma}| \leq c_2, \end{cases} \qquad (19)$$

where r_i is the square root of the i-th squared residual from LMS:

$$\begin{aligned} r_i^2 = \underset{i}{median}\,\big\{&(m_{1i} - m^*_{1i})^2 + (m_{2i} - m^*_{2i})^2 \\ &+[(m^*_{1i} - \lambda l_i) - (m^*_{1i} - \lambda l^*_i)]^2 + [(m_{2i} + \rho r_i) - (m^*_{2i} + \rho r^*_i)]^2\big\}, \end{aligned}$$

$\hat{\sigma}$ is the robust estimate of the scale of residuals, $\hat{\sigma} = \sqrt{median(r_i^2)}$, $r_i^2/\hat{\sigma}$, $i = 1, \ldots, n$, are the standardized residuals, and c_1 and $c_2$are constants. For our applications we set $c_1 = 2.5$ and $c_2 = 3.5$.

The final estimates of the coefficients of the WLS-based fuzzy regression model (5a)–(5d) are derived by applying formulae (20a)–(20n) (see Appendix) to the whole sample, by setting $\mathbf{W} = diag(w_i)$.

In what follows, we refer to this robust model as the LMS–WLS-based fuzzy regression model.

The weights (19) suitably tune the effect of outliers, removing the units with weight equal to 0 from the optimization process, and reducing the impact of those units with weights 0.5, which deviate less than the former from the bulk of data.

Finally, note that the presence of outlier entails, *ceteris paribus*, an increase in $R^2_{\mathbf{W}}$, since outliers have weights equal to 0, or at most 0.5, thus yielding a decrease of $SSR_{\mathbf{W}}$. Hence, we expect that the robust LMS–WLS-based fuzzy regression model performs better, in terms of goodness of fit to data, with respect to the LS-based fuzzy regression model.

5 A simulation study

5.1 Fuzzy simple linear regression model: crisp input, LR_2 fuzzy output

To illustrate the main features of the LS ($\mathbf{W} = \mathbf{I}$) and of the LMS–WLS-based fuzzy regression models proposed, we first consider a simulated dataset with a LR_2 fuzzy response variable and one crisp explanatory variable ($p = 1$). As observed above, the LS and the LMS–WLS-based fuzzy regression models could be easily derived from the model (5a)–(5d) (see Remark 4).

We generated 40 observations on the crisp variable from $U[1, 2]$. Then we generated the LR_2 fuzzy output as follows:

$$\begin{cases} m_{1i} = 1.61 + 3.50x_i + N(0,1) = m^*_{1i} + N(0,1) \\ m_{2i} = 1.99 + 4.18x_i + N(0,1) = m^*_{2i} + N(0,1) \\ l_i = 0.24 + 0.01m^*_{1i} + 0.04m^*_{2i} + N(0,1) = l^*_i + N(0,1) \\ r_i = 0.16 + 0.04m^*_{1i} + 0.06m^*_{2i} + N(0,1) = r^*_i + N(0,1) \end{cases}$$

Figures 3a–b show the results of the two models fitted to the generated dataset. The value of the determination coefficient is reported in the top left of each figure. Each LR_2 fuzzy output value is represented by a solid line (centers) and two dashed lines (spreads). The fitted model is represented by two solid lines for the models on the left and right centers, and two dotted lines for the models on the left and right spreads.

The two models provide similar results, as can be seen also by the value of $R^2_{\mathbf{W}}$.

To evaluate the different behaviour of the LS and of the LMS–WLS-based fuzzy regression model in presence of anomalous data, we have also contaminated the simulated dataset with three different kind of outliers: one outlier in the input; one outlier in both centers; one outlier in both spreads.

Figures 3c–h refer to the cases in which each kind of outlier is considered, one at a time. The outlier is highlighted with a thicker line. In Figs. 4a–h different combinations of the three type of outliers are considered.

Some consideration follows:

1. the LS-based fuzzy regression model is heavily affected by the presence of a single outlier in the input or in the centers of the output, especially when the two contaminations are combined;
2. the effect of the presence of anomalous values in the spread is mitigated by the presence of the weights λ and ρ of the spreads in the objective function (6);
3. the LMS–WLS-based fuzzy regression model performance is not affected by the presence of a single outlier, irrespective of which element (centers and/or spreads of the response variable variable, and/or explanatory variable) of the generated dataset is contaminated.

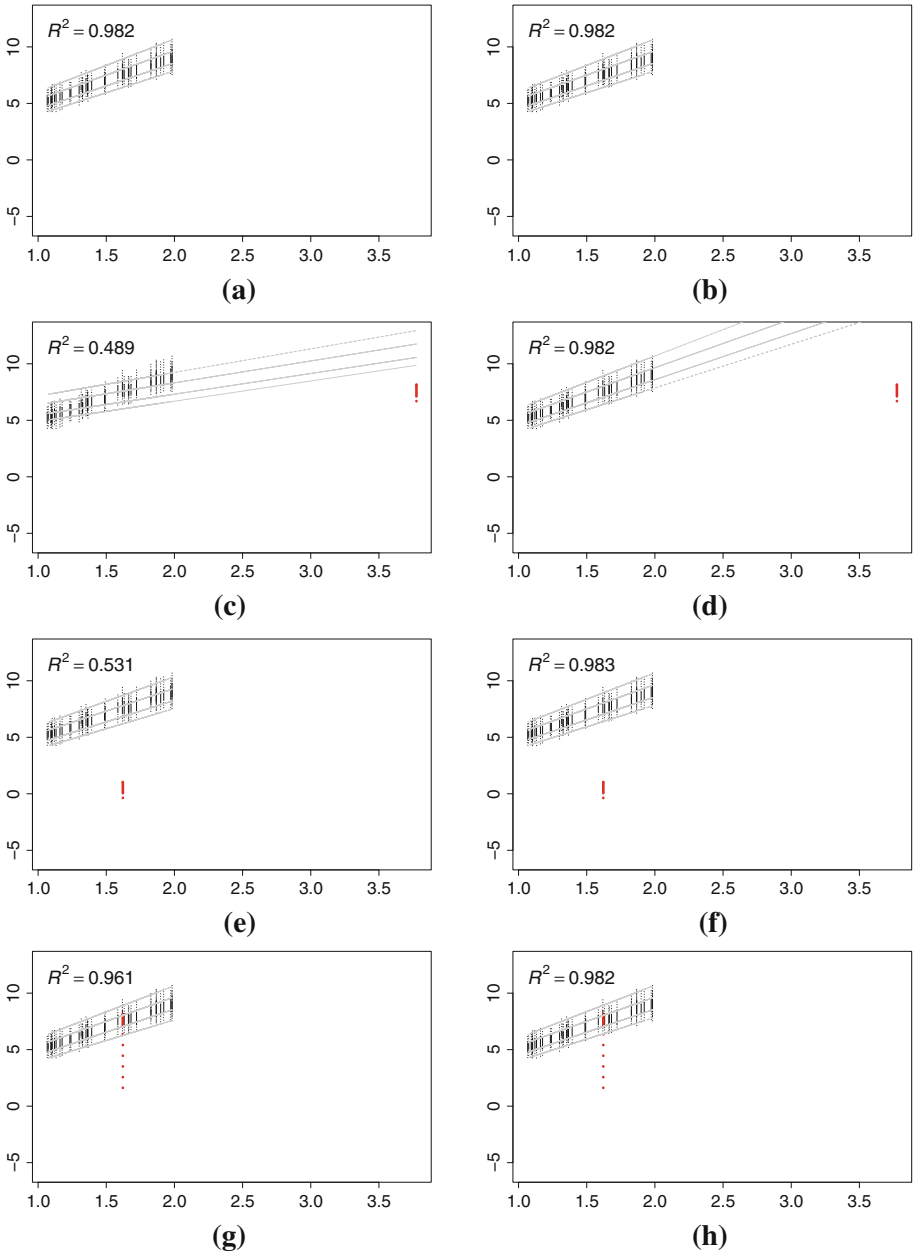

Fig. 3 Fitting of the LS and LMS–WLS-based fuzzy regression models to the simulated dataset: no contamination or contamination in one element. **a** LS: no outlier. **b** LMS–WLS: no outlier. **c** LS: one outlier in the input. **d** LMS–WLS: one outlier in the input. **e** LS: one outlier in both centers. **f** LMS–WLS: one outlier in both centers. **g** LS: one outlier in both spreads. **h** LMS–WLS: one outlier in both spreads

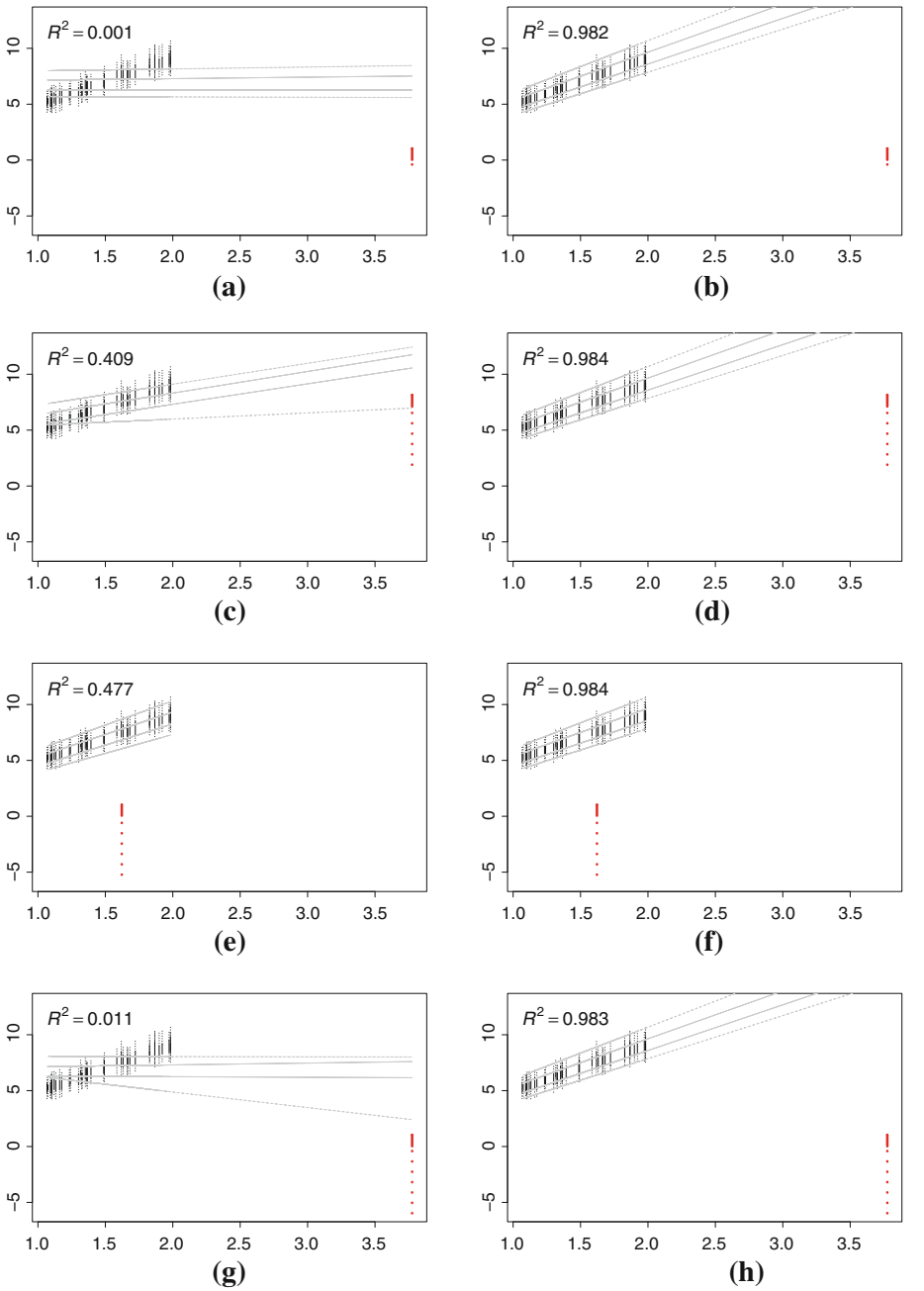

Fig. 4 Fitting of the LS and LMS–WLS-based fuzzy regression models to the simulated dataset: contamination in two or more elements. **a** LS: one outlier in the input and in the centers. **b** LMS–WLS: one outlier in the input and in the centers. **c** LS: one outlier in the input and in the spreads. **d** LMS–WLS: one outlier in the input and in the spreads. **e** LS: one outlier in the centers and in the spreads. **f** LMS–WLS: one outlier in the centers and in the spreads. **g** LS: one outlier in all elements. **h** LMS–WLS: one outlier in all elements

5.2 Fuzzy simple linear regression model: LR_2 fuzzy input/output

We now consider a simulation carried out on a fuzzy linear regression model with a fuzzy LR_2 response variable and one fuzzy LR_2 explanatory variable ($p = 1$). Fuzzy data for the simulation study are generated with the following scheme:

$$\begin{aligned}
\mathbf{m}_1 &= (\mathbf{1},\ U[1,2])(1.42,\ 1.53)' + (\mathbf{1},\ U[2.5,3.5])(0.16,\ 0.88)' \\
&\quad + (\mathbf{1},\ U[0.1,0.2])(0.20,\ 0.58)' + (\mathbf{1},\ U[0.15,0.25])(0.80,\ 0.26)' + N(0,1) \\
&= \mathbf{m}_1^* + N(0,1) \\
\mathbf{m}_2 &= (\mathbf{1},\ U[1,2])(1.00,\ 2.25)' + (\mathbf{1},\ U[2.5,3.5])(2.61,\ 1.38)' \\
&\quad + (\mathbf{1},\ U[0.1,0.2])(0.24,\ 0.48)' + (\mathbf{1},\ U[0.15,0.25])(0.01,\ 0.48)' + N(0,1) \\
&= \mathbf{m}_2^* + N(0,1) \\
\mathbf{l} &= 1.1751 + \mathbf{m}_1^* \cdot 0.044 + \mathbf{m}_2^* \cdot 0.014 + N(0,1) = \mathbf{l}^* + N(0,1) \\
\mathbf{r} &= 1.093 + \mathbf{m}_1^* \cdot 0.022 + \mathbf{m}_2^* \cdot 0.026 + N(0,1) = \mathbf{r}^* + N(0,1)
\end{aligned}$$

We generated 100 datasets of 200 observations. In each dataset the regression coefficients were held constant, while the values of the fuzzy inputs were randomly generated. We fitted to each generated dataset both the LS and the LMS–WLS-based fuzzy regression model. Finally, we computed the mean and the median of $R^2_{\mathbf{W}}$ to evaluate the average and the median fitting performance of the two models over the simulation cycle. Results are reported in the fourth column of Table 3. As expected, both models provide a good fitting performance.

Then we contaminated each dataset by adding an increasing percentage of outliers (from 5 to 30 %, by steps of 5 %) in the input or in the output centers following different contamination schemes, summarized in Table 2.

In Table 3 (columns fifth to tenth) we also report the mean and the median of $R^2_{\mathbf{W}}$ computed over the 100 datasets generated for each outlier generation scheme.

Table 2 Outlier generation schemes

Scheme	Outliers in the input centers	Outliers in the output centers
	${}_x m_1 \sim U[4,5]$	$m_1^* \sim U[0,1]$
	${}_x m_1 \sim U[6.5,8.5]$	$m_2^* \sim U[14.5,15]$

Table 3 Simulation results: mean and median of $R^2_{\mathbf{W}}$ computed over 100 simulated datasets

Model	Outlier scheme		Percentage of outliers						
			0 %	5 %	10 %	15 %	20 %	25 %	30 %
LS	Input centers	Mean	0.978	0.330	0.264	0.240	0.234	0.199	0.168
	Output centers			0.238	0.142	0.107	0.077	0.071	0.073
	Input centers	Median	0.978	0.322	0.269	0.241	0.249	0.200	0.198
	Output centers			0.234	0.147	0.115	0.081	0.073	0.069
LMS–WLS	Input centers	Mean	0.978	0.976	0.978	0.978	0.977	0.977	0.978
	Output centers			0.979	0.978	0.978	0.977	0.976	0.979
	Input centers	Median	0.979	0.977	0.980	0.978	0.979	0.980	0.980
	Output centers			0.980	0.979	0.978	0.977	0.977	0.978

Some consideration follow:

1. the LS-based fuzzy regression model is heavily affected by the presence of outliers in the centers of the input variable, even when there are only 5 % of outliers;
2. the LMS–WLS-based fuzzy regression model is not affected by the presence of outliers in the centers of the input variable, irrespective of the percentage of outliers,
3. similar conclusions can be deduced when there are outliers in the centers of the output variable.

6 Applications

6.1 Daily variation of pollutant concentration

In this application we examine the dependence relationship between the atmospheric concentration of carbon monoxide (CO) and other pollutants, namely mono-nitrogen oxides (NO_x), which, in atmospheric chemistry, correspond to the total concentration of nitric oxide (NO) and nitrogen dioxide (NO_2), and ozone (O_3).

The original data was collected in Rome during the year 1999. The original dataset provides hourly values of all the variables considered. Prior to the analysis, we have standardized all variables.

We are interested in detecting the effect that the daily variation of the inputs (NO_x and O_3) exerts on the daily variation of the concentration of CO. Missing data prevent us from contrasting daily variation of CO with daily variation of NO_x and O_3. Therefore, we compute the weekly averages of the daily minimum and maximum of each variable.

To cope with the loss of information due to summarizing the data, we consider the variables as LR_2 fuzzy variables. Then, the two centers of each LR_2 fuzzy variable are given by the mean values of the minimum and maximum value recorded each day of the week; the left (right) spreads are the mean deviations from the average minimum (maximum) values, of those values which are lower (higher) than the average minimum (maximum) values. We further assume that the shape of the LR_2 membership function is trapezoidal, which implies $\lambda = \rho = 1/2$. See Coppi et al. [10] for a similar fuzzy formalization of the data.

Having considered weekly data, we end up with 53 observations. The obtained fuzzy data matrix is reported in Table 4

The determination coefficient computed for the LS-based fuzzy regression model is equal to 0.879, while that of the LMS–WLS-based fuzzy regression model is 0.880. Both models provide a good fit to data, even if the LMS–WLS-based fuzzy regression model slightly outperforms the non-robust model due to the presence of three outliers.

The estimates of the coefficients are reported in Tables 5 (models on the centers) and 6 (models on the spreads). As can be seen, the estimates are similar between the two models.

With the proposed model, it is possible to highlight which component of each fuzzy explanatory variables mainly affects each component of the response variable.

Consider, for instance, the influence of the left centers of the explanatory variables on the left and right centers of the response, given respectively by α_1 and β_1. In both cases, the main effect is exerted by the weekly average minimum values of the concentration of NO_x. The greater is this value, the greater are both the weekly average minimum values of CO. Moreover, since the effect on the right center is greater, as the minimum value of the concentration mono-nitrogen oxides raises, the daily variation of carbon monoxideincreases.

Similar evidence can be drawn from the effect of the right centers of the inputs, in particular of NO_x. Given the almost nil effect on the left center of the output and the strong influence on

Table 4 Standardized pollution data: left (right) centers are the mean values of the minimum (maximum) value recorded each day of the week; left (right) spreads are the mean deviations from the average minimum (maximum) values, considering only values lower (higher) than the average minimum (maximum)

Week	CO_2				NO_X				O_3			
	Centers		Spreads		Centers		Spreads		Centers		Spreads	
	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right
1	−1.161	1.241	1.268	0.948	−1.151	1.822	1.241	0.971	−0.835	0.037	0.838	0.830
2	−0.957	3.560	1.123	0.890	−0.678	4.120	0.809	0.503	−0.841	−0.434	0.846	0.835
3	−1.139	3.419	1.336	0.992	−0.941	3.606	1.212	0.832	−0.833	0.320	0.838	0.827
4	−0.932	4.690	1.036	0.793	−0.606	4.126	0.713	0.525	−0.765	−0.150	0.819	0.439
5	−1.280	2.481	1.378	1.036	−1.074	2.230	1.285	0.794	−0.584	1.059	0.831	0.035
6	−1.189	2.971	1.375	1.050	−0.882	3.599	1.160	0.674	−0.782	0.510	0.818	0.567
7	−1.316	2.190	1.428	1.094	−1.247	1.649	1.341	1.060	−0.645	1.343	0.785	0.505
8	−1.056	2.290	1.224	0.832	−0.870	2.496	1.062	0.613	−0.773	0.368	0.796	0.742
9	−0.923	2.730	1.181	0.280	−0.866	2.407	1.147	0.655	−0.691	1.034	0.759	0.280
10	−1.214	2.556	1.355	1.026	−1.070	2.147	1.153	0.863	−0.736	1.478	0.776	0.684
11	−1.006	2.938	1.123	0.861	−0.805	3.092	0.898	0.681	−0.756	1.648	0.776	0.729
12	−1.297	1.235	1.390	1.142	−1.113	1.083	1.341	0.810	−0.368	2.043	0.635	0.299
13	−1.463	2.248	1.481	1.355	−1.087	1.617	1.208	0.996	−0.734	1.573	0.758	0.710
14	−1.113	1.231	1.268	0.803	−1.132	1.585	1.422	0.842	−0.714	1.702	0.776	0.590
15	−1.289	1.111	1.370	1.181	−1.103	1.225	1.212	1.021	−0.699	2.222	0.747	0.579
16	−1.287	1.658	1.394	1.181	−1.097	1.834	1.290	0.905	−0.731	1.404	0.790	0.702
17	−1.278	0.757	1.384	1.065	−1.041	0.803	1.176	0.773	−0.843	1.597	0.846	0.838
18	−1.272	1.576	1.312	1.220	−1.092	1.722	1.289	0.828	−0.846	1.611	0.846	0.846
19	−1.272	1.277	1.394	1.181	−1.092	0.467	1.294	0.941	−0.825	1.281	0.843	0.718
20	−1.280	1.036	1.355	1.224	−1.101	0.739	1.218	0.945	−0.740	1.913	0.820	0.256
21	−1.289	1.377	1.355	1.123	−1.120	0.885	1.251	1.022	−0.827	1.543	0.841	0.794
22	−1.206	0.953	1.316	1.123	−0.985	0.578	1.116	0.887	−0.833	3.251	0.843	0.826
23	−1.222	1.269	1.283	1.142	−1.004	0.964	1.093	0.884	−0.831	2.961	0.838	0.826
24	−1.247	1.318	1.316	1.195	−1.017	0.761	1.070	0.947	−0.825	2.934	0.833	0.806
25	−1.210	0.970	1.297	1.123	−1.026	1.067	1.157	0.894	−0.834	2.531	0.838	0.830
26	−1.094	0.505	1.239	0.948	−1.027	0.672	1.075	0.979	−0.846	0.594	0.846	0.846
27	−1.045	0.728	1.123	0.658	−0.954	0.550	1.065	0.842	−0.766	3.218	0.834	0.630
28	−1.289	0.779	1.309	1.239	−1.087	0.747	1.171	1.024	−0.835	2.260	0.842	0.827
29	−1.256	0.986	1.297	1.200	−1.108	0.370	1.235	1.013	−0.810	3.374	0.836	0.746
30	−1.189	0.546	1.413	1.099	−1.043	0.135	1.271	0.952	−0.457	2.162	0.825	0.462
31	−1.239	0.729	1.274	1.210	−1.152	0.597	1.290	1.097	−0.797	2.602	0.818	0.769
32	−1.231	0.372	1.262	1.152	−1.104	0.453	1.201	1.031	−0.835	3.028	0.840	0.830
33	−1.231	0.015	1.367	0.890	−1.092	0.307	1.229	0.750	−0.780	2.101	0.832	0.710
34	−1.297	−0.280	1.297	1.297	−1.248	0.027	1.279	1.186	−0.774	2.079	0.822	0.678
35	−1.161	0.011	1.227	0.832	−1.109	0.039	1.181	1.037	−0.608	3.019	0.779	0.247
36	−1.307	0.698	1.394	1.220	−1.091	0.750	1.205	0.977	−0.761	1.670	0.790	0.731
37	−1.169	1.423	1.278	1.006	−1.060	0.985	1.156	0.996	−0.806	2.111	0.822	0.790
38	−1.189	1.460	1.326	1.134	−1.043	1.534	1.247	0.889	−0.842	2.281	0.846	0.834

Table 4 continued

Week	CO_2				NO_x				O_3			
	Centers		Spreads		Centers		Spreads		Centers		Spreads	
	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right
39	−1.173	1.501	1.297	1.079	−0.976	1.617	1.150	0.846	−0.840	1.670	0.846	0.832
40	−1.214	2.091	1.283	1.123	−1.016	1.470	1.126	0.868	−0.843	0.905	0.846	0.838
41	−1.272	2.938	1.413	1.084	−1.013	2.261	1.213	0.863	−0.835	1.208	0.842	0.827
42	−1.164	2.199	1.227	1.006	−0.834	2.389	0.902	0.742	−0.841	1.078	0.846	0.838
43	−1.297	1.767	1.378	1.161	−1.144	1.422	1.227	1.034	−0.835	0.283	0.839	0.826
44	−1.156	2.805	1.251	0.919	−0.877	2.439	1.020	0.686	−0.840	0.206	0.846	0.832
45	−1.256	2.091	1.413	1.137	−1.010	1.876	1.260	0.822	−0.835	0.805	0.844	0.824
46	−1.264	2.107	1.399	1.084	−0.960	1.806	1.304	0.823	−0.755	0.275	0.834	0.555
47	−1.264	2.398	1.384	1.103	−1.043	2.104	1.190	0.848	−0.827	0.185	0.836	0.806
48	−1.197	2.506	1.355	1.079	−0.883	2.700	1.073	0.741	−0.830	−0.118	0.838	0.822
49	−0.905	2.873	1.142	0.193	−0.576	2.965	0.866	0.293	−0.832	−0.583	0.838	0.830
50	−1.079	2.786	1.326	0.832	−0.732	2.840	1.081	0.383	−0.826	−0.250	0.830	0.822
51	−1.206	3.120	1.344	0.861	−0.968	2.605	1.275	0.738	−0.821	−0.235	0.831	0.794
52	−1.314	2.116	1.471	1.195	−1.112	1.735	1.281	0.888	−0.641	0.561	0.801	0.240
53	−1.332	1.214	1.413	1.210	−1.098	1.711	1.250	0.871	−0.798	0.535	0.824	0.758

Table 5 Coefficients' estimates for the LS and the LMS–WLS-based fuzzy regression models: Models on the centers (pollution data)

Model	α_1	α_2	α_l	α_r	α_1	α_2	α_l	α_r
LMS								
Int.	−4.172	2.082	0.773	1.027	−2.366	3.905	−0.784	−0.540
O_3	−0.166	0.007	−0.231	−0.102	0.052	−0.013	0.603	−0.358
NO_x	0.540	−0.013	−0.148	−0.061	2.218	0.931	0.656	1.517
LMS–WLS								
Int.	−3.553	1.222	1.173	1.099	−1.654	2.825	−0.937	−0.679
O_3	−0.078	0.016	−0.462	−0.073	−0.221	−0.045	1.279	−0.444
NO_x	0.577	−0.010	−0.093	−0.105	2.078	0.919	0.458	1.650

Table 6 Coefficients' estimates for the LS and the LMS–WLS-based fuzzy regression models: Models on the spreads (pollution data)

Model	δ_0	δ_1	δ_2	γ_0	γ_1	γ_2
LS	−0.774	−1.508	−0.002	0.397	−0.743	0.014
LMS–WLS	−0.898	−1.602	0.005	0.397	−0.742	0.014

the right center, we can derive a positive influence of the maximum value of the concentration mono-nitrogen oxides on the daily variation of CO.

Overall, we observe a direct relationship between the daily variation of NO_x and that of carbon monoxide.

6.2 Attitude towards traditional vs. "creative" advertising

The aim of this application is to illustrate how to cope with various source of the uncertainty that may affect the regression analysis: fuzziness of the response and of the explanatory variables; uncertainty about the values of regression coefficients; uncertainty about the choice of a specific model in a class of parametric model.

Data for our analysis are drawn from a survey on a sample of 103 students from Sapienza University and LUISS University, in Rome, interviewed about their opinions about traditional and new media. A section of the survey was devoted to the respondents' opinions towards traditional vs. innovative advertising campaigns. Respondents are asked to report their degree of agreement towards these seven statements (in brackets are reported the names of each variable):

- I am sensitive to traditional advertising campaigns, i.e. campaigns broadcast by TV and/or radio, published on newspaper or magazines, etc. (**sens-tr**; response variable).
- I am tired of traditional advertising campaigns (**tired-tr**).
- I do not pay attention to traditional advertising campaigns (**not-tr**).
- I try to avoid traditional advertising campaigns (**avoid-tr**).
- I am impressed by "creative" advertising campaigns, e.g., via blog and/or social networks, sponsorship of public events, etc. (**impr-cre**).
- Creative advertising campaigns are more effective in capturing my attention (**eff-cre**).
- I better remember a creative advertising campaigns with respect to a more traditional one (**rem-cre**).

The degree of agreement were reported on a 4-item scale, from "I totally disagree" (1) to "I totally agree" (4).

The complete dataset is reported in Table 7.

Coppi and D'Urso [8] observed that the subjective evaluation of a qualitative scale could be better represented in a fuzzy framework, which takes into account the uncertainty and the heterogeneity in individual evaluation.

Hence, we adopted a fuzzy coding for describing the subjective judgements reported in the survey. In particular, we recoded the qualitative variables as LR_1 fuzzy variables. The LR_1 fuzzy recoding of these linguistic variables is reported in Table 8 [1], and represented in Fig. 5, in which it is also shown the membership function of each fuzzy value.

Then, we analysed the relationship between the variable **sens-tr** and the remaining variables by means of a fuzzy linear regression model with LR_1 fuzzy output and LR_1 fuzzy outputs. As observed in the Remark 4, this model is a particular case of the more general model (5a)–(5d).

To select the optimal model we employed a procedure based on the maximization of the value of the adjusted determination coefficient, $\bar{R}^2_{\mathbf{W}}$. Notice that in this case each added variable involves the estimation of 7 additional coefficients ($\bar{p} = [3 * (p + 1) + 4]$). Then the penalization factor increases more than proportionally for each added variable, as observed in Sect. 2.4. For this reason we considered the following expression for the adjusted determination coefficient:

$$\bar{R}^2_{\mathbf{W}} = 1 - (1 - R^2_{\mathbf{W}}) \frac{n-1}{n-p}$$

The selection procedure adopted is backward-type and can be illustrated as follows.

For a model with k fuzzy inputs we compute $\bar{R}^2_{\mathbf{W}}(k)$. Then, we compute $\bar{R}^2_{\mathbf{W},j}(k-1)$ for all the k models derived from the first model by dropping one variable at time ($j = 1, \ldots, k$).

Table 7 Student data

Student	sens-tr	tired-tr	not-tr	avoid-tr	impr-cre	eff-cre	rem-cre
1	2	3	2	3	1	3	1
2	1	3	4	3	3	4	4
3	2	4	1	2	4	4	4
4	2	2	2	2	3	4	4
5	2	3	2	3	3	3	3
6	3	2	2	3	2	3	2
7	2	3	3	3	4	4	4
8	3	2	2	2	2	2	2
9	2	3	3	4	3	3	3
10	3	2	2	3	2	4	3
11	2	4	2	3	4	3	4
12	2	2	3	3	4	3	2
13	3	4	1	2	3	4	4
14	2	3	2	3	4	3	3
15	2	3	2	4	3	4	4
16	2	4	3	3	3	4	4
17	1	2	2	4	4	4	4
18	3	2	3	4	2	3	3
19	4	3	1	1	4	4	4
20	2	3	3	4	3	3	3
21	3	3	1	2	2	2	3
22	3	2	3	4	4	4	4
23	4	1	1	1	4	3	3
24	3	3	2	3	3	4	3
25	3	1	1	2	4	4	4
26	3	4	3	4	4	4	4
27	3	4	4	4	4	4	4
28	2	4	2	3	3	3	3
29	1	4	4	4	2	4	3
30	2	2	4	3	1	1	1
31	2	1	3	3	2	3	2
32	2	3	3	3	3	4	3
33	2	3	3	1	4	4	4
34	3	2	2	2	3	3	2
35	2	4	4	4	3	3	3
36	2	3	1	1	4	4	4
37	3	2	1	2	1	2	2
38	3	2	2	3	3	3	4
39	3	2	2	4	1	2	3
40	3	1	2	2	2	4	3
41	3	3	2	2	4	4	4
42	3	2	3	4	3	3	4
43	2	3	2	2	3	2	2
44	3	1	1	2	3	4	4

Table 7 continued

Student	sens-tr	tired-tr	not-tr	avoid-tr	impr-cre	eff-cre	rem-cre
45	3	1	1	2	2	2	2
46	2	4	4	4	3	4	4
47	3	2	2	2	3	3	4
48	2	4	3	3	4	4	4
49	3	4	2	3	4	4	4
50	1	4	4	4	4	4	4
51	3	4	1	1	4	4	4
52	3	3	2	4	4	4	4
53	1	4	4	4	1	2	2
54	2	3	2	3	2	3	3
55	3	2	4	1	4	4	4
56	2	3	2	3	3	3	3
57	2	4	4	4	3	4	4
58	2	2	4	2	2	3	3
59	2	3	3	3	4	4	4
60	1	3	3	2	3	3	3
61	3	2	2	4	3	3	3
62	2	3	3	2	3	3	4
63	2	4	3	3	1	2	2
64	2	2	3	1	1	1	1
65	2	4	2	3	2	3	3
66	3	2	3	3	3	3	4
67	2	2	2	3	3	3	3
68	2	4	2	4	2	2	4
69	2	2	2	4	4	3	3
70	2	3	3	2	3	3	2
71	2	3	2	3	2	2	2
72	2	3	3	4	4	4	3
73	3	3	1	3	2	2	2
74	1	4	2	3	1	3	2
75	2	2	3	4	3	3	3
76	2	3	3	3	4	4	4
77	1	4	4	2	2	2	3
78	3	4	3	4	4	4	4
79	3	4	2	4	3	3	3
80	4	1	1	1	1	1	1
81	2	3	2	2	2	2	2
82	3	2	2	2	3	3	3
83	4	4	1	4	4	4	3
84	3	2	1	4	4	4	4
85	2	4	4	4	4	4	4
86	1	3	3	2	1	2	2
87	3	2	2	2	3	3	3
88	4	3	2	4	4	4	4

Table 7 continued

Student	sens-tr	tired-tr	not-tr	avoid-tr	impr-cre	eff-cre	rem-cre
89	2	4	3	3	4	4	4
90	2	1	4	2	3	3	1
91	3	3	2	1	4	4	4
92	2	3	2	3	2	3	3
93	2	4	3	4	2	2	2
94	2	3	3	2	3	3	3
95	3	1	1	1	3	4	4
96	2	3	3	3	2	3	4
97	3	4	4	4	3	4	4
98	2	2	2	1	3	4	4
99	2	3	3	4	3	3	4
100	3	2	3	4	3	3	4
101	4	1	1	1	3	2	3
102	2	2	2	2	2	2	2
103	2	4	3	3	4	4	4

Table 8 Linguistic variables and corresponding fuzzy values (center, left spread, right spread)

Linguistic variables	Code	Fuzzy values
I totally disagree	1	(3, 3, 1)
I partially disagree	2	(4, 1.5, 1.5)
I partially agree	3	(6, 1, 0.5)
I totally agree	4	(8, 1.75, 0.25)

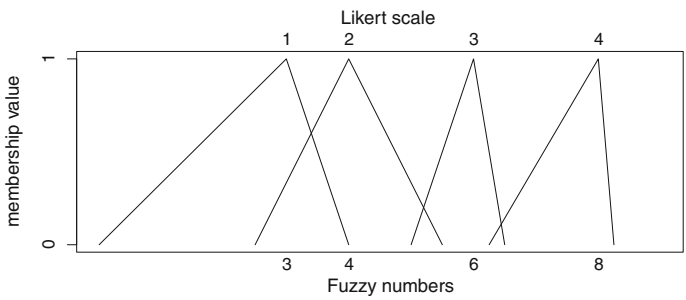

Fig. 5 Fuzzy recoding of the 4-item scale, with membership function

If $max_j \bar{R}^2_{\mathbf{W},j}(k-1) > \bar{R}^2_{\mathbf{W}}(k)$, we consider the model j' with $k-1$ inputs such that $j' = \text{argmax}_j \bar{R}^2_{\mathbf{W},j}(k-1)$ and we continue the procedure. Otherwise, we select the model with k fuzzy inputs.

The model selection procedure is summarized in Table 9.

As can be seen, the best model is the LMS–WLS-based fuzzy regression model with six fuzzy inputs. Note also that all the LS-based fuzzy regression models are severely affected by the presence of outliers, as can be seen by the low values of $\bar{R}^2_{\mathbf{W}}$.

As for the estimates of the coefficients of the fuzzy regression model, as observed in Sect. 3, one has to take into account the imprecision due to the ignorance about the data generation process. Hence we generated 100 bootstrap samples. We then fitted both models to these

Table 9 Model selection

Variable excluded	LS		LMS–WLS		Outliers	(%)
	$R^2_{\mathbf{W}}$	$\bar{R}^2_{\mathbf{W}}$	$R^2_{\mathbf{W}}$	$\bar{R}^2_{\mathbf{W}}$		
None	0.462	0.429	0.984	0.983	40	38.83
tired-tr	0.397	0.366	0.851	0.843	37	35.92
not-tr	0.287	0.250	0.540	0.517	24	23.30
avoid-tr	0.426	0.397	0.850	0.842	39	37.86
impr-cre	0.452	0.424	0.845	0.837	41	39.81
eff-cre	0.457	0.429	0.883	0.877	42	40.78
rem-cre	0.460	0.433	0.936	0.933	44	42.72

Table 10 Coefficients' estimates for the LS and the LMS–WLS-based fuzzy regression models: Models on the center (advertising data)

Model	$\boldsymbol{\alpha}$ (SE)	$\boldsymbol{\alpha}_l$ (SE)	$\boldsymbol{\alpha}_r$ (SE)
LS			
Int.	1.424 (1.228)	1.388 (0.734)	1.446 (0.527)
$\tilde{\mathbf{x}}_1$	−0.016 (0.098)	0.330 (0.119)	0.562 (0.331)
$\tilde{\mathbf{x}}_2$	−0.442 (0.075)	0.259 (0.104)	−0.432 (0.219)
$\tilde{\mathbf{x}}_3$	0.145 (0.089)	0.398 (0.160)	0.053 (0.283)
$\tilde{\mathbf{x}}_4$	0.144 (0.076)	−0.153 (0.116)	0.213 (0.226)
$\tilde{\mathbf{x}}_5$	−0.054 (0.128)	0.252 (0.178)	−0.104 (0.391)
$\tilde{\mathbf{x}}_6$	−0.019 (0.099)	0.030 (0.137)	−0.228 (0.330)
LMS–WLS			
Int.	1.645 (0.710)	1.019 (0.353)	0.505 (0.190)
$\tilde{\mathbf{x}}_1$	−0.056 (0.038)	0.589 (0.052)	1.503 (0.133)
$\tilde{\mathbf{x}}_2$	−0.376 (0.053)	0.643 (0.124)	−1.091 (0.144)
$\tilde{\mathbf{x}}_3$	−0.028 (0.054)	0.003 (0.039)	−0.096 (0.126)
$\tilde{\mathbf{x}}_4$	0.336 (0.059)	−0.546 (0.067)	0.977 (0.156)
$\tilde{\mathbf{x}}_5$	−0.050 (0.059)	0.009 (0.128)	−0.157 (0.188)
$\tilde{\mathbf{x}}_6$	0.119 (0.045)	−0.130 (0.077)	0.388 (0.180)

bootstrap samples. The standard deviation of the estimates of the regression coefficients provided us a measure of the accuracy of the estimates obtained with both models.

The estimates of the coefficients are reported in Tables 10 (models on the center) and 11 (models on the spreads). The bootstrap estimates of the standard errors are reported in brackets.

As expected, the presence of outliers produces some bias in the estimates of the LS-based fuzzy regression model. Consider, for instance, the effect that the center of the third explanatory variables exerts on the centers of the response. The two models return estimates with opposite signs. However, one would expect that the more the respondent tries to avoid traditional campaigns, the less sensitive is she or he to these types of campaigns, i.e., we expect a negative sign of the coefficient, as in the LMS–WLS-based fuzzy regression model, while the sign for the LS-based fuzzy regression model is positive.

Focusing only on the LMS–WLS-based fuzzy regression model and on the model which relates the centers of the explanatory variables to the center of the output, we notice that the second, the fourth and the sixth variable are significant. Thus, not paying attention to

Table 11 Coefficients' estimates for the LS and the LMS–WLS-based fuzzy regression models: Models on the spreads (advertising data)

	LS Est. (SE)	LMS–WLS Est (SE)
δ_0	1.763 (0.057)	2.121 (0.019)
δ_1	−0.174 (0.015)	−0.251 (0.006)
γ_0	2.051 (0.276)	1.937 (0.060)
γ_1	−0.095 (0.058)	−0.067 (0.013)

traditional campaigns, being impressed by creative campaign, and better recalling creative campaigns affect the most the sensitiveness to traditional campaigns.

7 Final remarks

In this paper, a generalization of the fuzzy regression model proposed by Coppi et al. [10] has been discussed. In particular, by considering an iterative Weighted Least Squares estimation approach, a general linear regression model for studying the dependence of a general class of fuzzy response variable ,i.e., LR_2 fuzzy variable or trapezoidal fuzzy variable, on a set of crisp or LR_2 fuzzy explanatory variables has been proposed. Furthermore, some theoretical properties and a suitable generalization of the determination coefficient to investigate the goodness of fit of the regression model, have been illustrated. To neutralize and/or smooth disruptive effects of possible crisp or fuzzy outliers in the estimation process, a robust version of the fuzzy regression model based on the Least Median Squares estimation approach has been suggested. Finally, some theoretical remarks and an assessment of imprecision of the regression function have been illustrated. The good performance of our models are shown by means of a simulation study and some applications to real cases.

In future, the proposed fuzzy regression model and its robust version might be improved in several directions. In particular, an interesting aspect is related to the modelization of the regression relationship between the spreads of the fuzzy response variable and the respective estimated centers. In model (5a)–(5d) these are assumed in a simple linear form. A more complex relationship could be considered, in order to cope with observational studies where the simple linear assumption is not suitable.

Another interesting issue is to utilize our models in a clusterwise context [17,19].

Furthermore,to improve the capability of managing the uncertainty due to randomness of the data, a further line of research that deserves careful attention in future research consists of making our fuzzy regression models probabilistic, by using the notion of fuzzy random variable (see, e.g., [7]). We will investigate the above lines of research in future works.

Appendix: Iterative solutions of the LR_2 output–LR_2 inputs regression model

By substituting in (7) the expressions (5a)–(5d), and by putting the first partial derivatives with respect to each coefficient equal to zero, the following iterative solutions are obtained.

$$\begin{aligned}\boldsymbol{\alpha}_1 = {} & (2-2\lambda\gamma_1+\lambda^2\gamma_1^2+\rho^2\delta_1^2)^{-1}(\mathbf{M}_1'\mathbf{W}\mathbf{M}_1)^{-1}\mathbf{M}_1'\mathbf{W} \\ & \cdot\{2\mathbf{m}_1-(\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)(2-2\lambda\gamma_1+\lambda^2\gamma_1^2+\rho^2\delta_1^2) \\ & -\lambda[\mathbf{m}_1\gamma_1+\mathbf{l}-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\gamma_2-\mathbf{1}\gamma_0] \\ & +\lambda^2\gamma_1[\mathbf{l}-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\gamma_2-\mathbf{1}\gamma_0]\end{aligned}$$

$$
\begin{aligned}
&+\rho\delta_1[\mathbf{m}_2-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)]\\
&+\rho^2\delta_1[\mathbf{r}-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\delta_2-\mathbf{1}\delta_0]\}
\end{aligned}
\tag{20a}
$$

$$
\begin{aligned}
\boldsymbol{\alpha}_2 = {}&(2-2\lambda\gamma_1+\lambda^2\gamma_1^2+\rho^2\delta_1^2)^{-1}(\mathbf{M}_2'\mathbf{W}\mathbf{M}_2)^{-1}\mathbf{M}_2'\mathbf{W}\\
&\cdot\{2\mathbf{m}_1-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)(2-2\lambda\gamma_1+\lambda^2\gamma_1^2+\rho^2\delta_1^2)\\
&-\lambda[\mathbf{m}_1\gamma_1+\mathbf{l}-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\gamma_2-\mathbf{1}\gamma_0]\\
&+\lambda^2\gamma_1[\mathbf{l}-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\gamma_2-\mathbf{1}\gamma_0]\\
&+\rho\delta_1[\mathbf{m}_2-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)]\\
&+\rho^2\delta_1[\mathbf{r}-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\delta_2-\mathbf{1}\delta_0]\}
\end{aligned}
\tag{20b}
$$

$$
\begin{aligned}
\boldsymbol{\alpha}_l = {}&(2-2\lambda\gamma_1+\lambda^2\gamma_1^2+\rho^2\delta_1^2)^{-1}(\mathbf{L}'\mathbf{W}\mathbf{L})^{-1}\mathbf{L}'\mathbf{W}\\
&\cdot\{2\mathbf{m}_1-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{R}\boldsymbol{\alpha}_r)(2-2\lambda\gamma_1+\lambda^2\gamma_1^2+\rho^2\delta_1^2)\\
&-\lambda[\mathbf{m}_1\gamma_1+\mathbf{l}-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\gamma_2-\mathbf{1}\gamma_0]\\
&+\lambda^2\gamma_1[\mathbf{l}-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\gamma_2-\mathbf{1}\gamma_0]\\
&+\rho\delta_1[\mathbf{m}_2-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)]\\
&+\rho^2\delta_1[\mathbf{r}-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\delta_2-\mathbf{1}\delta_0]\}
\end{aligned}
\tag{20c}
$$

$$
\begin{aligned}
\boldsymbol{\alpha}_r = {}&(2-2\lambda\gamma_1+\lambda^2\gamma_1^2+\rho^2\delta_1^2)^{-1}(\mathbf{R}'\mathbf{W}\mathbf{R})^{-1}\mathbf{R}'\mathbf{W}\\
&\cdot\{2\mathbf{m}_1-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l)(2-2\lambda\gamma_1+\lambda^2\gamma_1^2+\rho^2\delta_1^2)\\
&-\lambda[\mathbf{m}_1\gamma_1+\mathbf{l}-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\gamma_2-\mathbf{1}\gamma_0]\\
&+\lambda^2\gamma_1[\mathbf{l}-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\gamma_2-\mathbf{1}\gamma_0]\\
&+\rho\delta_1[\mathbf{m}_2-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)]\\
&+\rho^2\delta_1[\mathbf{r}-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\delta_2-\mathbf{1}\delta_0]\}
\end{aligned}
\tag{20d}
$$

$$
\begin{aligned}
\boldsymbol{\beta}_1 = {}&(2+\lambda^2\gamma_2+2\rho\delta_2+\rho^2\delta_2^2)^{-1}(\mathbf{M}_1'\mathbf{W}\mathbf{M}_1)^{-1}\mathbf{M}_1'\mathbf{W}\\
&\cdot\{2\mathbf{m}_2-(\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)(2+\lambda^2\gamma_2+2\rho\delta_2+\rho^2\delta_2^2)\\
&-\lambda\gamma_2[\mathbf{m}_1-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)]\\
&+\lambda^2\gamma_2[\mathbf{l}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\gamma_1-\mathbf{1}\gamma_0]\\
&+\rho[\mathbf{m}_2\delta_2+\mathbf{r}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\delta_1-\mathbf{1}\delta_0]\\
&+\rho^2\delta_2[\mathbf{r}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\delta_1-\mathbf{1}\delta_0]\}
\end{aligned}
\tag{20e}
$$

$$
\begin{aligned}
\boldsymbol{\beta}_2 = {}&(2+\lambda^2\gamma_2+2\rho\delta_2+\rho^2\delta_2^2)^{-1}(\mathbf{M}_2'\mathbf{W}\mathbf{M}_2)^{-1}\mathbf{M}_2'\mathbf{W}\\
&\cdot\{2\mathbf{m}_2-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)(2+\lambda^2\gamma_2+2\rho\delta_2+\rho^2\delta_2^2)\\
&-\lambda\gamma_2[\mathbf{m}_1-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)]\\
&+\lambda^2\gamma_2[\mathbf{l}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\gamma_1-\mathbf{1}\gamma_0]\\
&+\rho[\mathbf{m}_2\delta_2+\mathbf{r}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\delta_1-\mathbf{1}\delta_0]\\
&+\rho^2\delta_2[\mathbf{r}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\delta_1-\mathbf{1}\delta_0]\}
\end{aligned}
\tag{20f}
$$

$$
\begin{aligned}
\boldsymbol{\beta}_l = {}&(2+\lambda^2\gamma_2+2\rho\delta_2+\rho^2\delta_2^2)^{-1}(\mathbf{L}'\mathbf{W}\mathbf{L})^{-1}\mathbf{L}'\mathbf{W}\\
&\cdot\{2\mathbf{m}_2-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{R}\boldsymbol{\beta}_r)(2+\lambda^2\gamma_2+2\rho\delta_2+\rho^2\delta_2^2)\\
&-\lambda\gamma_2[\mathbf{m}_1-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)]\\
&+\lambda^2\gamma_2[\mathbf{l}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\gamma_1-\mathbf{1}\gamma_0]
\end{aligned}
$$

$$+\rho[\mathbf{m}_2\delta_2+\mathbf{r}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\delta_1-\mathbf{1}\delta_0]$$
$$+\rho^2\delta_2[\mathbf{r}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\delta_1-\mathbf{1}\delta_0]\} \tag{20g}$$

$$\begin{aligned}\boldsymbol{\beta}_r=&(2+\lambda^2\gamma_2+2\rho\delta_2+\rho^2\delta_2^2)^{-1}(\mathbf{R}'\mathbf{W}\mathbf{R})^{-1}\mathbf{R}'\mathbf{W}\\&\cdot\{2\mathbf{m}_2-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l)(2+\lambda^2\gamma_2+2\rho\delta_2+\rho^2\delta_2^2)\\&-\lambda\gamma_2[\mathbf{m}_1-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)]\\&+\lambda^2\gamma_2[\mathbf{l}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\gamma_1-\mathbf{1}\gamma_0]\\&+\rho[\mathbf{m}_2\delta_2+\mathbf{r}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\delta_1-\mathbf{1}\delta_0]\\&+\rho^2\delta_2[\mathbf{r}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\delta_1-\mathbf{1}\delta_0]\}\end{aligned} \tag{20h}$$

$$\begin{aligned}\gamma_1=&\lambda^{-1}[(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)'\mathbf{W}(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)]^{-1}\\&(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)'\mathbf{W}\{(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)-\mathbf{m}_1\\&+\lambda[\mathbf{l}-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\gamma_2-\mathbf{1}\gamma_0]\}\end{aligned} \tag{20i}$$

$$\begin{aligned}\gamma_2=&\lambda^{-1}[(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)'\mathbf{W}(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)]^{-1}\\&(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)'\mathbf{W}\{(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)-\mathbf{m}_1\\&+\lambda[\mathbf{l}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\gamma_1-\mathbf{1}\gamma_0]\}\end{aligned} \tag{20j}$$

$$\begin{aligned}\gamma_0=&\lambda^{-1}(\mathbf{1}'\mathbf{W}\mathbf{1})^{-1}\mathbf{1}'\mathbf{W}\{(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)-\mathbf{m}_1\\&+\lambda[\mathbf{l}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\gamma_1-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\gamma_2]\}\end{aligned} \tag{20k}$$

$$\begin{aligned}\delta_1=&\rho^{-1}[(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)'\mathbf{W}(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)]^{-1}\\&(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)'\mathbf{W}\{\mathbf{m}_2-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\\&+\rho[\mathbf{r}-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\delta_2-\mathbf{1}\delta_0]\}\end{aligned} \tag{20l}$$

$$\begin{aligned}\delta_2=&\rho^{-1}[(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)'\mathbf{W}(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)]^{-1}\\&(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)'\mathbf{W}\{\mathbf{m}_2-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\\&+\rho[\mathbf{r}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\delta_1-\mathbf{1}\delta_0]\}\end{aligned} \tag{20m}$$

$$\begin{aligned}\delta_0=&\rho^{-1}(\mathbf{1}'\mathbf{W}\mathbf{1})^{-1}\mathbf{1}'\mathbf{W}\{\mathbf{m}_2-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\\&+\rho[\mathbf{r}-(\mathbf{M}_1\boldsymbol{\alpha}_1+\mathbf{M}_2\boldsymbol{\alpha}_2+\mathbf{L}\boldsymbol{\alpha}_l+\mathbf{R}\boldsymbol{\alpha}_r)\delta_1-(\mathbf{M}_1\boldsymbol{\beta}_1+\mathbf{M}_2\boldsymbol{\beta}_2+\mathbf{L}\boldsymbol{\beta}_l+\mathbf{R}\boldsymbol{\beta}_r)\delta_2]\}.\end{aligned} \tag{20n}$$

References

1. Anand Raj, P., Nagesh Kumar, D.: Ranking alternatives with fuzzy weights using maximizing set and minimizing set. Fuzzy Sets Syst. **105**(3), 365–375 (1999)
2. Blanco-Fernández, A., Colubi, A., García-Bárzana, M.: A set arithmetic-based linear regression model for modelling interval-valued responses through real-valued variables. Inf. Sci. **247**(20), 109–122 (2013)
3. Blanco-Fernández, A., Casals, M.R., Colubi, A., Corral, N., García-Bárzana, M., Gil, M.A., González-Rodríguez, G., López, M.T., Lubiano, M.A., Montenegro, M., Ramos-Guajardo, A.B., de la Rosa de Sáa, S., Sinova, B.: A distance-based statistical analysis of fuzzy number-valued data. Int. J. Approx. Reason. doi:10.1016/j.ijar.2013.09.020 (2014)
4. Celmiņš, A.: Least squares model fitting to fuzzy vector data. Fuzzy Sets Syst. **22**(3), 245–269 (1987)
5. Chang, O.Y.H., Ayyub, M.B.: Fuzzy regression methods—a comparative assessment. Fuzzy Sets Syst. **119**(2), 187–203 (2001)

6. Chang, P.T., Stanley Lee, E.: A generalized fuzzy weighted least-squares regression. Fuzzy Sets Syst. **82**(3), 289–298 (1996)
7. Colubi, A., Coppi, R., D'Urso, P., Gil, M.A.: Statistics with fuzzy random variable. Metron **LXV**(3), 277–303 (2007)
8. Coppi, R., D'Urso, P.: Fuzzy k-mean clustering models for triangular fuzzy time trajectories. Stat. Methods Appl. **11**, 21–24 (2002)
9. Coppi, R., D'Urso, P.: Regression analysis with fuzzy informational paradigm: a least-squares approach using membership function information. Int. J. Pure Appl. Math. **8**, 279–306 (2003)
10. Coppi, R., D'Urso, P., Giordani, P., Santoro, A.: Least squares estimation of a linear regression model with LR fuzzy response. Comput. Stat. Data Anal. **51**(1), 267–286 (2006)
11. Diamond, P.: Fuzzy least squares. Inf. Sci. **46**(3), 141–157 (1988)
12. Diamond, P., Tanaka, H.: Fuzzy regression analysis. In: Slowinski, R. (ed.) Fuzzy Sets in Decision Analysis, Operations Research and Statistics, pp. 349–387. Kluwer Academic Publishers, Massachusetts (1998)
13. Dubois, D., Prade, H.: Possibility Theory. Plenum Press, New York (1988)
14. D'Urso, P.: Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data. Comput. Stat. Data Anal. **42**(1), 47–72 (2003)
15. D'Urso, P., Gastaldi, T.: A least-squares approach to fuzzy linear regression analysis. Comput. Stat. Data Anal. **34**(4), 427–440 (2000)
16. D'Urso, P., Giordani, P.: Fitting of fuzzy linear regression models with multivariate response. Int. Math. J. **3**(6), 655–664 (2003)
17. D'Urso, P., Santoro, A.: Fuzzy clusterwise linear regression analysis with symmetrical fuzzy output variable. Comput. Stat. Data Anal. **51**(1), 287–313 (2006)
18. D'Urso, P., Santoro, A.: Goodness of fit and variable selection in the fuzzy multiple linear regression. Fuzzy Sets Syst. **157**, 2627–2647 (2006)
19. D'Urso, P., Massari, R., Santoro, A.: A class of fuzzy clusterwise regression models. Inf. Sci. **180**(24), 4737–4762 (2010)
20. D'Urso, P., Massari, R., Santoro, A.: Robust fuzzy regression analysis. Inf. Sci. **181**(19), 4154–4174 (2011)
21. Ferraro, M.B., Coppi, R., González Rodríguez, G., Colubi, A.: A linear regression model for imprecise response. Int. J. Approx. Reason. **51**(7), 759–770 (2010)
22. González-Rodríguez, G., Blanco, Á., Corral, N., Colubi, A.: Least squares estimation of linear regression models for convex compact random sets. Adv. Data Anal. Classif. **1**(1), 67–81 (2007)
23. González-Rodríguez, G., Blanco, Á., Colubi, A., Lubiano, M.A.: Estimation of a simple linear regression model for fuzzy random variables. Fuzzy Sets Syst. **160**(3), 357–370 (2009)
24. Kim, K.J., Moskowitz, H., Koksalan, M.: Fuzzy versus statistical linear regression. Eur. J. Oper. Res. **92**(2), 417–434 (1996)
25. Körner, R., Näther, W.: Linear regression with random fuzzy variables: extended classical estimates, best linear estimates, least squares estimates. Inf. Sci. **109**(1–4), 95–118 (1998)
26. Krätschmer, V.: Least squares estimation in linear regression models with vague concepts. In: Lopez-Diaz, M., Gil, M., Grzegorzewski, P., Hryniewicz, P., Lawry, J. (eds.) Soft Methodology and Random Information Systems, pp. 407–414. Springer, Heidelberg (2004)
27. Lawson, C.L., Hanson, R.J.: Solving least squares problems. SIAM, Philadelphia (1995)
28. Ma, M., Friedman, M., Kandel, A.: General fuzzy least squares. Fuzzy Sets Syst. **88**(1), 107–118 (1997)
29. Näther, W.: On random fuzzy variables of second order and their application to linear statistical inference with fuzzy data. Metrika **51**(3), 201–221 (2000)
30. Rousseeuw, P.J.: Least median of squares regression. J. Am. Stat. Assoc. **79**(388), 871–880 (1984)
31. Rousseeuw, P.J., Leroy, A.M.: Robust regression and outlier detection, vol. 589. Wiley, New York (2005)
32. Sinova, B., Colubi, A., Gil, M., et al.: Interval arithmetic-based simple linear regression between interval data: discussion and sensitivity analysis on the choice of the metric. Inf. Sci. **199**, 109–124 (2012)
33. Tanaka, H., Watada, J.: Possibilistic linear systems and their application to the linear regression model. Fuzzy Sets Syst. **27**(3), 275–289 (1988)
34. Tanaka, H., Uejima, S., Asai, K.: Linear regression analysis with fuzzy model. IEEE Trans. Syst. Man Cybern. **12**(6), 903–907 (1982)
35. Wu, H.C.: Fuzzy estimates of regression parameters in linear regression models for imprecise input and output data. Comput. Stat. Data Anal. **42**(1), 203–217 (2003)
36. Wünsche, A., Naether, W.: Least-squares fuzzy regression with fuzzy random variables. Fuzzy Sets Syst. **130**(1), 43–50 (2002)
37. Zimmermann, H.J.: Fuzzy Set Theory and Its Applications. Kluver Academic Press, Norwell (2011)