Soft Computing (2021) 25:835–849
https://doi.org/10.1007/s00500-020-05441-2

FOUNDATIONS

A general approach to fuzzy regression models based on different loss functions

Amir Hamzeh Khammar[1] · Mohsen Arefi[1] · Mohammad Ghasem Akbari[1]

Published online: 12 January 2021

Abstract
In this paper, a new general approach is presented to fit a fuzzy regression model when the response variable and the parameters of model are as fuzzy numbers. In this approach, for estimating the parameters of fuzzy regression model, a new definition of objective function is introduced based on the different loss functions and under the averages of differences between the α-cuts of errors. The application of the proposed approach is studied using a simulated data set and some real data sets in the presence of different types of outliers.

Keywords Goodness of fit · Loss function · Outlier data · Robust fuzzy regression

1 Introduction

The fuzzy regression analysis is based on some regression models between a response variable and some explanatory variables when some quantities are as imprecise. Also, among the observed data, we may encounter with some outliers. In such situations, we need to introduce a regression model which supports these limitations. An approach for solving these limitations is using of the robust fuzzy regression models.

In the following, we review some works on robust fuzzy regression models. Arefi (2020) studied a robust fuzzy regression based on a generalized quantile loss function under the fuzzy outputs and fuzzy parameters. Chang and Lee (1994) presented fuzzy least absolute deviations regression based on the ranking of fuzzy numbers. An approach to fit a robust least squares fuzzy regression based on a kernel function is investigated by Khammar et al. (2020). Oussalah and De Schutter (2002) proposed a fuzzy regression model with the combination of least trimmed squares (LTS) and least median squares (LMS). They studied the performance of the proposed model when the data are contaminated by outliers. Sanli and Apaydin (2004) investigated a robust estimation procedure for fuzzy linear regression model with fuzzy input–output data based on the least median squares method. A robust approach to model a fuzzy linear regression based on M-estimators is studied by Shon (2005). Choi and Buckley (2008) utilized the least absolute deviations (LAD) method for estimating the parameters of a fuzzy regression model and investigated the performance of the proposed model under the fuzzy outlier data. D'Urso et al. (2011) and D'Urso and Massari (2013) proposed a robust fuzzy linear regression model with crisp inputs and fuzzy outputs based on the least median squares-weighted least squares (LMS-WLS) estimation procedure. Some approaches to fit the fuzzy regression models based the least absolutes method are investigated by Chachi and Taheri (2016), Taheri and Kelkinnama (2012), and Zeng et al. (2016). Based on the least trimmed squares estimation, Chachi and Roozbeh (2017) proposed a estimation procedure for determining the coefficients of a fuzzy regression model with crisp input–fuzzy output data. A weighted least-squares fuzzy regression model under crisp input–fuzzy output data and fuzzy coefficients is provided by Chachi (2019). For considering some other approaches of fuzzy regression models, see Chen and Hsueh (2007, 2009), Nasrabadi and Hashemi (2008), Hao and Chiang (2008), Kula et al. (2012), Mosleh et al. (2010), Arefi and Taheri (2015), Lopez et al. (2016), Hesamian and Akbari (2019), and Rapaic et al. (2019).

Communicated by A. Di Nola.

✉ Mohsen Arefi
arefi@birjand.ac.ir

Amir Hamzeh Khammar
amir.khmmar@birjand.ac.ir

Mohammad Ghasem Akbari
g_z_akbari@birjand.ac.ir

[1] Department of Statistics, Faculty of Mathematical Sciences and Statistics, University of Birjand, Birjand, Iran

This paper is organized as follows: In Sect. 2, some preliminary concepts about fuzzy sets and fuzzy numbers are provided. In Sect. 3, a new general approach based on the different loss functions is investigated to fit the fuzzy regression models when the response variable and the parameters of models are as fuzzy numbers. In this section, we present some indices of goodness of fit to evaluate the proposed fuzzy regression models. Also, the cross-validation method is provided to examine the predictive ability of the proposed fuzzy regression models. Some numerical examples to assess the effectiveness of the proposed method are presented in Sect. 4. Finally, in Sect. 5, some concluding remarks are provided.

2 Preliminary concepts

In this section, we recall some notations and preliminary concepts on fuzzy sets (see Zimmermann 2001).

Let Ω be an universal set. A fuzzy set $\tilde{N}$ of Ω is defined by the membership function $\tilde{N}:\Omega\to[0,1]$. The α-cut of $\tilde{N}$ is as $\tilde{N}[\alpha]=\{x\in R:\tilde{N}(x)\geq\alpha\}$ for $0<\alpha\leq 1$.

Definition 1 A fuzzy set $\tilde{N}$ on Ω is called a fuzzy number, if

(i) $\widetilde{N}(x)=1$ for some $x\in\Omega$,

(ii) $\widetilde{N}[\alpha]$ is a closed bounded interval for $0<\alpha\leq 1$.

Definition 2 Let $\tilde{N}$ be a fuzzy number, then a LR fuzzy number is defined by the following membership function

$$\tilde{N}(x)=\begin{cases}L\left(\frac{m-x}{l}\right) & x\leq m,\\ R\left(\frac{x-m}{r}\right) & x>m,\end{cases}$$

where $l,r\geq 0$ and $L(.)$ and $R(.)$ are the strictly decreasing functions as $L,R:R^+\to[0,1]$. It is denoted by $\tilde{N}=(m,l,r)_{LR}$.

Remark 1 In a LR fuzzy number $\tilde{N}$, if $L(x)=R(x)$, then $\tilde{N}$ is called the LL fuzzy number and is denoted as $\tilde{N}=(m,l,r)_{LL}$. For $L(x)=R(x)=1-x$ for all $x\in[0,1]$, $\tilde{N}$ is called a triangular fuzzy number and is denoted by $\tilde{N}=(m,l,r)_T$. Also, for $l=r$, $\tilde{N}$ is a symmetric triangular fuzzy number as $\tilde{N}=(m,l)_T$.

Proposition 1 *Assume that* $\tilde{A}=(m_a,l_a,r_a)_{LR}$ *and* $\tilde{B}=(m_b,l_b,r_b)_{LR}$ *are two LR fuzzy numbers and* $\lambda\in R-\{0\}$. *Some of the arithmetic operations based on extension principle* (Zimmermann 2001) *are given as follows*

$$\lambda\otimes\tilde{A}=\begin{cases}(\lambda m_a,\lambda l_a,\lambda r_a)_{LR} & \lambda>0,\\ (\lambda m_a,-\lambda r_a,-\lambda l_a)_{RL} & \lambda<0,\end{cases}$$

$$\tilde{A}\oplus\tilde{B}=(m_a+m_b,l_a+l_b,r_a+r_b)_{LR}.$$

3 Methodology

In this section, we introduce a new approach in fuzzy regression theory based on the concept of α-cuts for the crisp inputs and the fuzzy output when the parameters of model are as fuzzy quantities.

3.1 Fuzzy regression model

In the following, we introduce fuzzy regression model under crisp input–fuzzy output variables. Our aim is to fit an optimal fuzzy linear regression model to this data set. Suppose that the fuzzy linear regression model based on crisp input–fuzzy output variables and the fuzzy parameters is given as follows

$$\tilde{Y}_i=\tilde{\beta}_0\oplus(\tilde{\beta}_1\otimes x_{i1})\oplus(\tilde{\beta}_2\otimes x_{i2})\oplus\cdots\oplus(\tilde{\beta}_p\otimes x_{ip}),$$

where, $\tilde{\beta}_j=(\beta_j,\gamma_j)_{LL}$, $j=0,\ldots,p$, and $\widetilde{Y}_i=(y_i,s_i)_{LL}$, $i=1,\ldots,n$. The estimated fuzzy response variables are obtained as follows:

$$\begin{aligned}\hat{\tilde{Y}}_i&=\tilde{\beta}_0\oplus(\tilde{\beta}_1\otimes x_{i1})\oplus(\tilde{\beta}_1\otimes x_{i2})\oplus\cdots\oplus(\tilde{\beta}_p\otimes x_{ip})\\ &=\left(\sum_{j=0}^{p}\beta_j x_{ij},\sum_{j=0}^{p}\gamma_j x_{ij}\right)_{LL},\quad i=1,\ldots,n,\end{aligned}\tag{1}$$

where $x_{i0}=1$, $i=1,\ldots,n$.

3.2 Objective function

In the following, we first introduce some differences between the α-cuts of fuzzy numbers, and then the objective functions of fuzzy regression models are provided based on loss function on these differences as follows.

Definition 1 Let $\tilde{A}$ and $\tilde{B}$ be two fuzzy numbers with α-cuts $\tilde{A}[\alpha]=[\tilde{A}^L[\alpha],\tilde{A}^R[\alpha]]$ and $\tilde{B}[\alpha]=[\tilde{B}^L[\alpha],\tilde{B}^R[\alpha]]$. The averages of differences between the lower and upper α-cuts of $\tilde{A}$ and $\tilde{B}$ are defined as

$$DL(\tilde{A},\tilde{B})=\int_0^1(\tilde{A}^L[\alpha]-\tilde{B}^L[\alpha])d\alpha,$$

$$DR(\tilde{A},\tilde{B})=\int_0^1(\tilde{A}^R[\alpha]-\tilde{B}^R[\alpha])d\alpha.$$

Remark 2 Let $\tilde{A}$, $\tilde{B}$ and $\tilde{C}$ be three fuzzy numbers. Then, $DL(.,.)$ and $DR(.,.)$ in Definition 1 satisfy the following properties:

(i) $DL(\tilde{A},\tilde{A})=0$ and $DR(\tilde{A},\tilde{A})=0$,
(ii) $DL(\tilde{A},\tilde{B})=-DL(\tilde{B},\tilde{A})$ and $DR(\tilde{A},\tilde{B})=-DR(\tilde{B},\tilde{A})$,
(iii) $DL(\tilde{A},\tilde{B})+DL(\tilde{B},\tilde{C})=DL(\tilde{A},\tilde{C})$ and $DR(\tilde{A},\tilde{B})+DR(\tilde{B},\tilde{C})=DR(\tilde{A},\tilde{C})$.

Remark 3 In a special case, if $\tilde{A} = (m_a, l_a, r_a)_{LR}$ and $\tilde{B} = (m_b, l_b, r_b)_{LR}$ are two LR fuzzy numbers, then

$$DL(\tilde{A}, \tilde{B}) = (m_a - m_b) - \lambda(l_a - l_b),$$
$$DR(\tilde{A}, \tilde{B}) = (m_a - m_b) + \rho(r_a - r_b),$$

where $\lambda = \int_0^1 L^{-1}(\alpha)d\alpha$ and $\rho = \int_0^1 R^{-1}(\alpha)d\alpha$. Furthermore, if $\tilde{A} = (m_a, l_a, r_a)_T$ and $\tilde{B} = (m_b, l_b, r_b)_T$ are two triangular fuzzy numbers, then

$$DL(\tilde{A}, \tilde{B}) = (m_a - m_b) - \frac{1}{2}(l_a - l_b),$$
$$DR(\tilde{A}, \tilde{B}) = (m_a - m_b) + \frac{1}{2}(r_a - r_b),$$

Remark 4 If the LR fuzzy numbers $\tilde{A} = (m_a, l_a, r_a)_{LR}$ and $\tilde{B} = (m_b, l_b, r_b)_{LR}$ are reduced to the crisp numbers (i.e. $l_a = r_a = l_b = r_b = 0$), then $DL(\tilde{A}, \tilde{B}) = DR(\tilde{A}, \tilde{B}) = m_a - m_b$.

Definition 3 Suppose that $\tilde{Y}_i$ and $\hat{\tilde{Y}}_i$ are the observed fuzzy response variable and the estimated fuzzy response variable, respectively. The objective function based on Definition 1 is defined as

$$O = \frac{1}{2}\sum_{i=1}^{n}\left[\psi(DL(\tilde{Y}_i, \hat{\tilde{Y}}_i)) + \psi((DR(\tilde{Y}_i, \hat{\tilde{Y}}_i))\right], \qquad (2)$$

where $\psi(.)$ is a loss function.

In the objective function O, the relation $\frac{1}{2}\big(\psi(DL(\tilde{Y}_i, \hat{\tilde{Y}}_i)) + \psi((DR(\tilde{Y}_i, \hat{\tilde{Y}}_i))\big)$ can be considered as the loss value of error between $\tilde{Y}_i$ and $\hat{\tilde{Y}}_i$. Note that based on the different loss functions $\psi(.)$, we can obtain the different fuzzy regression models. Some loss functions are listed as follows:

(1) Squared error loss function : $\psi^1(e) = e^2$,

(2) Absolute error loss function : $\psi^2(e) = |e|$,

(3) Quantile loss function (Koenker 2005) :
$\psi_\tau^3(e) = |e||\tau - I(e < 0)|, \quad 0 < \tau \leq 1$,

(4) Huber loss function (Huber 1981) :
$$\psi_c^4(e) = \begin{cases} \frac{1}{2}e^2, & |e| \leq c, \\ c|e| - \frac{1}{2}c^2, & |e| > c \end{cases}, c > 0.$$

Remark 5 In the loss function ψ_τ^3, τ is the quantile level and for $\tau = 0.5$, ψ_τ^3 is reduced to the loss function ψ^2. Also, in the loss function ψ_c^4, based on some of indices of goodness of fit, we choose the value of c in interval [0, 2] (Huber 1981 says that the good value of c is in interval [1, 2]. So, it is taken as $c = 1.5$).

From the relation (2), the objective functions based on the above loss functions are provided as follows:

(1) $$O_{\psi^1} = \frac{1}{2}\sum_{i=1}^{n}\Big((DL(\tilde{Y}_i, \hat{\tilde{Y}}_i))^2 + (DR(\tilde{Y}_i, \hat{\tilde{Y}}_i))^2\Big).$$

(2) $$O_{\psi^2} = \frac{1}{2}\sum_{i=1}^{n}\Big(|DL(\tilde{Y}_i, \hat{\tilde{Y}}_i)| + |DR(\tilde{Y}_i, \hat{\tilde{Y}}_i)|\Big).$$

(3) $$O_{\psi_\tau^3} = \frac{1}{2}\sum_{i=1}^{n}\Big[\Big(|DL(\tilde{Y}_i, \hat{\tilde{Y}}_i)| \times |\tau - I(DL(\tilde{Y}_i, \hat{\tilde{Y}}_i) < 0)|\Big) + \Big(|DR(\tilde{Y}_i, \hat{\tilde{Y}}_i)| \times |\tau - I(DR(\tilde{Y}_i, \hat{\tilde{Y}}_i) < 0)|\Big)\Big].$$

(4) $$O_{\psi_c^4} = \frac{1}{2}\sum_{i=1}^{n}\Big[\Big((DL(\tilde{Y}_i, \hat{\tilde{Y}}_i))^2/2 - \big((DL(\tilde{Y}_i, \hat{\tilde{Y}}_i))^2/2 - c|DL(\tilde{Y}_i, \hat{\tilde{Y}}_i)| + c^2/2\big)I(|DL(\tilde{Y}_i, \hat{\tilde{Y}}_i)| > c)\Big) + \Big((DR(\tilde{Y}_i, \hat{\tilde{Y}}_i))^2/2 - \big((DR(\tilde{Y}_i, \hat{\tilde{Y}}_i))^2/2 - c|DR(\tilde{Y}_i, \hat{\tilde{Y}}_i)| + c^2/2\big)I(|DR(\tilde{Y}_i, \hat{\tilde{Y}}_i)| > c)\Big)\Big].$$

By minimizing these objective functions under the parameters of model, we can obtain the optimal fuzzy regression models.

3.3 Estimation of model parameters

Based on the above objective functions, we can calculate the estimations of parameters of models as follows (assume that $x_{ij} \geq 0$).

(A) Squared error loss function: Based on relation (1), the objective function O_{ψ^1} is rewritten as

$$O_{\psi^1} = \frac{1}{2}\sum_{i=1}^{n}\left[\left[\left(y_i - \sum_{j=0}^{p}\beta_j x_{ij}\right) - \lambda\left(s_i - \sum_{j=0}^{p}\gamma_j x_{ij}\right)\right]^2 + \left[\left(y_i - \sum_{j=0}^{p}\beta_j x_{ij}\right) + \lambda\left(s_i - \sum_{j=0}^{p}\gamma_j x_{ij}\right)\right]^2\right]$$
$$= (\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta) + \lambda^2(\mathbf{S} - \mathbf{X}\gamma)'(\mathbf{S} - \mathbf{X}\gamma).$$

where

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \ldots & x_{1p} \\ 1 & x_{21} & \ldots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \ldots & x_{np} \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix},$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \quad \gamma = \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_p \end{bmatrix}.$$

By differentiating of O_{ψ^1} with respect to β_j and γ_j, the estimations of parameters are obtained as follows:

$$\hat{\beta} = (\mathbf{X}^{'}\mathbf{X})^{-1}\mathbf{X}^{'}\mathbf{Y},$$
$$\hat{\gamma} = (\mathbf{X}^{'}\mathbf{X})^{-1}\mathbf{X}^{'}\mathbf{S}.$$

Remark 6 If we encounter with some negative spreads, we can again run the fuzzy regression model with considering such parameters as crisp (i.e. the related spreads are zero). Thus, the centers are obtained as before, but the spreads are calculated as follows

$$\hat{\gamma}^* = (\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{X}^{*'}\mathbf{S},$$

where $\mathbf{X}^*$ is like $\mathbf{X}$ in which the columns corresponding to crisp parameters are removed.

(B) Absolute error loss function and quantile loss function: For minimizing the objective functions O_{ψ^2} and O_{ψ^3}, we can translate it into the linear programming problems as follows. Assume that $DL(\tilde{Y}_i, \hat{\tilde{Y}}_i) = d_i^{L+} - d_i^{L-}$ and $DR(\tilde{Y}_i, \hat{\tilde{Y}}_i) = d_i^{R+} - d_i^{R-}$ with $d_i^{L+} = \max(DL(\tilde{Y}_i, \hat{\tilde{Y}}_i), 0)$, $d_i^{L-} = -\min(DL(\tilde{Y}_i, \hat{\tilde{Y}}_i), 0)$, $d_i^{R+} = \max(DR(\tilde{Y}_i, \hat{\tilde{Y}}_i), 0)$, and $d_i^{R-} = -\min(DR(\tilde{Y}_i, \hat{\tilde{Y}}_i), 0)$, then the objective functions are presented as

$$O_{\psi^2} = \frac{1}{2}\sum_{i=1}^{n}\Big[|DL(\tilde{Y}_i, \hat{\tilde{Y}}_i)| + |DR(\tilde{Y}_i, \hat{\tilde{Y}}_i)|\Big]$$
$$= \frac{1}{2}\sum_{i=1}^{n}\Big[(d_i^{L+} + d_i^{L-}) + (d_i^{R+} + d_i^{R-})\Big],$$
$$O_{\psi_\tau^3} = \frac{1}{2}\sum_{i=1}^{n}\Big[\Big(|DL(\tilde{Y}_i, \hat{\tilde{Y}}_i)| \times |\tau - I(DL(\tilde{Y}_i, \hat{\tilde{Y}}_i) < 0)|\Big)$$
$$+ \Big(|DR(\tilde{Y}_i, \hat{\tilde{Y}}_i)| \times |\tau - I(DR(\tilde{Y}_i, \hat{\tilde{Y}}_i) < 0)|\Big)\Big]$$
$$= \frac{1}{2}\sum_{i=1}^{n}\Big[\tau(d_i^{L+} + d_i^{R+}) + (1-\tau)(d_i^{L-} + d_i^{R-})\Big].$$

Hence, the linear programming problems are given as follows:

$$\min \ O_{\psi^2} = \frac{1}{2}\sum_{i=1}^{n}\Big[d_i^{L+} + d_i^{L-} + d_i^{R+} + d_i^{R-}\Big]$$

s.t.

$$d_i^{L+} - d_i^{L-} = DL(\tilde{Y}_i, \hat{\tilde{Y}}_i) = (y_i - \textstyle\sum_{j=0}^{p}\beta_j x_{ij})$$
$$- \lambda(s_i - \textstyle\sum_{j=0}^{p}\gamma_j x_{ij}),$$
$$d_i^{R+} - d_i^{R-} = DR(\tilde{Y}_i, \hat{\tilde{Y}}_i) = (y_i - \textstyle\sum_{j=0}^{p}\beta_j x_{ij})$$
$$+ \lambda(s_i - \textstyle\sum_{j=0}^{p}\gamma_j x_{ij}),$$
$$d_i^{L+} \geq 0, \quad d_i^{L-} \geq 0, \quad d_i^{R+} \geq 0, \quad d_i^{R-} \geq 0, \quad i = 1, 2, \ldots, n,$$
$$\beta_j \in R, \quad \gamma_j \geq 0, \quad j = 0, \ldots, p.$$

and

$$\min \ O_{\psi_\tau^3} = \frac{1}{2}\sum_{i=1}^{n}\Big[\tau(d_i^{L+} + d_i^{R+}) + (1-\tau)(d_i^{L-} + d_i^{R-})\Big]$$

s.t.

$$d_i^{L+} - d_i^{L-} = DL(\tilde{Y}_i, \hat{\tilde{Y}}_i) = (y_i - \textstyle\sum_{j=0}^{p}\beta_j x_{ij})$$
$$- \lambda(s_i - \textstyle\sum_{j=0}^{p}\gamma_j x_{ij}),$$
$$d_i^{R+} - d_i^{R-} = DR(\tilde{Y}_i, \hat{\tilde{Y}}_i) = (y_i - \textstyle\sum_{j=0}^{p}\beta_j x_{ij})$$
$$+ \lambda(s_i - \textstyle\sum_{j=0}^{p}\gamma_j x_{ij}),$$
$$d_i^{L+} \geq 0, \quad d_i^{L-} \geq 0, \quad d_i^{R+} \geq 0, \quad d_i^{R-} \geq 0, \quad i = 1, 2, \ldots, n,$$
$$\beta_j \in R, \quad \gamma_j \geq 0, \quad j = 0, \ldots, p.$$

Remark 7 Note that for solving the above linear programming problems, we can use "LINGO" software (Schrage 2006) or "Mathematica" software (Wolfram 2015).

(C) Huber loss function: It is simple to verify that the Huber loss function can be equivalently rewritten as

$$\psi_c^4(e) = \frac{1}{2}\big[e^2 - (e-c)_+^2 - (-e-c)_+^2\big].$$

Hence, the objective functions $O_{\psi_c^4}$ based on matrix forms are obtained as follows

$$O_{\psi_c^4} = \frac{1}{4}\Big[\big(\mathbf{Y} - \mathbf{X}\beta - \lambda(S - \mathbf{X}\gamma)\big)^{'}\big(\mathbf{Y} - \mathbf{X}\beta - \lambda(\mathbf{S} - \mathbf{X}\gamma)\big)$$
$$- \big(\mathbf{Y} - \mathbf{X}\beta - \lambda(\mathbf{S} - \mathbf{X}\gamma) - c\mathbf{L}\big)_+^{'}$$
$$\big(\mathbf{Y} - \mathbf{X}\beta - \lambda(\mathbf{S} - \mathbf{X}\gamma) - c\mathbf{L}\big)_+$$
$$- \big(\mathbf{X}\beta - \mathbf{Y} - \lambda(\mathbf{X}\gamma - \mathbf{S}) - c\mathbf{L}\big)_+^{'}$$
$$\big(\mathbf{X}\beta - \mathbf{Y} - \lambda(\mathbf{X}\gamma - \mathbf{S}) - c\mathbf{L}\big)_+$$
$$+ \big(\mathbf{Y} - \mathbf{X}\beta + \lambda(S - \mathbf{X}\gamma)\big)^{'}\big(\mathbf{Y} - \mathbf{X}\beta + \lambda(\mathbf{S} - \mathbf{X}\gamma)\big)$$
$$- \big(\mathbf{Y} - \mathbf{X}\beta + \lambda(\mathbf{S} - \mathbf{X}\gamma) - c\mathbf{L}\big)_+^{'}$$
$$\big(\mathbf{Y} - \mathbf{X}\beta + \lambda(\mathbf{S} - \mathbf{X}\gamma) - c\mathbf{L}\big)_+$$

Table 1 Data set in Example 1

No.	CEC (Cmol(+)/kg) y_i	SAND (%) x_{i1}	OM (%) x_{i2}	No.	CEC (Cmol(+)/kg) y_i	SAND (%) x_{i1}	OM (%) x_{i2}
1	16.50	35.00	0.88	13	24.40	31.00	3.52
2	18.60	37.00	1.13	14	21.80	31.00	2.33
3	19.30	27.00	1.31	15	23.80	17.00	1.71
4	20.30	29.00	1.98	16	20.80	14.00	1.14
5	17.30	38.00	1.02	17	17.50	19.00	0.99
6	20.40	32.00	1.29	18	17.80	28.00	1.14
7	19.30	29.00	1.52	19	20.20	26.00	1.46
8	21.90	18.00	1.33	20	20.00	32.00	1.81
9	15.90	40.00	1.71	21	22.80	10.00	1.38
10	18.30	28.00	2.00	22	19.10	38.00	0.84
11	22.60	13.00	1.68	23	12.10	49.00	1.48
12	23.70	19.00	2.15	24	12.80	42.00	1.08

$$- \big(\mathbf{X}\beta - \mathbf{Y} + \lambda(\mathbf{X}\gamma - \mathbf{S}) - c\mathbf{L}\big)_+^{'} \big(\mathbf{X}\beta - \mathbf{Y} + \lambda(\mathbf{X}\gamma - \mathbf{S}) - c\mathbf{L}\big)_+ \Big], \quad (3)$$

where $\mathbf{L} = (1, 1, \ldots, 1)'$. By differentiating (3) with respect to the vectors β and γ, we have

$$\frac{\partial O_{\psi_c^4}}{\partial \beta} = -\mathbf{X}^{'}\mathbf{Y} + \mathbf{X}^{'}\mathbf{X}\beta + \frac{1}{2}\mathbf{X}^{'}[H_1(\beta,\gamma) + H_2(\beta,\gamma)] = 0,$$

$$\frac{\partial O_{\psi_c^4}}{\partial \gamma} = -\lambda^2\mathbf{X}^{'}\mathbf{S} + \lambda^2\mathbf{X}^{'}\mathbf{X}\gamma - \frac{1}{2}\lambda\mathbf{X}^{'}[H_1(\beta,\gamma) - H_2(\beta,\gamma)] = 0,$$

where

$$H_1(\beta,\gamma) = (\mathbf{Y} - \mathbf{X}\beta - \lambda(\mathbf{S} - \mathbf{X}\gamma) - c\mathbf{L})_+ - (\mathbf{X}\beta - \mathbf{Y} - \lambda(\mathbf{X}\gamma - \mathbf{S}) - c\mathbf{L})_+,$$

$$H_2(\beta,\gamma) = (\mathbf{Y} - \mathbf{X}\beta + \lambda(\mathbf{S} - \mathbf{X}\gamma) - c\mathbf{L})_+ - (\mathbf{X}\beta - \mathbf{Y} + \lambda(\mathbf{X}\gamma - \mathbf{S}) - c\mathbf{L})_+.$$

Now, the estimations of parameters are obtained by the following simple iterative algorithm:

$$\beta^{m+1} = (\mathbf{X}^{'}\mathbf{X})^{-1}\mathbf{X}^{'}\Big[\mathbf{Y} - \frac{1}{2}\big(H_1(\beta^m,\gamma^m) + H_2(\beta^m,\gamma^m)\big)\Big],$$

$$\gamma^{m+1} = (\mathbf{X}^{'}\mathbf{X})^{-1}\mathbf{X}^{'}\Big[\mathbf{S} + \frac{1}{2\lambda}\big(H_1(\beta^m,\gamma^m) - H_2(\beta^m,\gamma^m)\big)\Big],$$

$$m = 0, 1, \ldots.$$

3.4 Goodness of fit of the model

In order to evaluate the proposed fuzzy regression models, we introduce some indices of goodness of fit. Also, using cross-validation method, the performance and predictive ability of proposed regression models are evaluated (see Geisser 1993; Wasserman 2006)

3.4.1 Goodness of fit

Definition 4 (Pappis and Karacapilidis 1993) Suppose that $\tilde{Y}_i$ and $\hat{\tilde{Y}}_i$ are the observed fuzzy response variable and the estimated fuzzy response variable, respectively. The mean of similarity measures (MSM) is defined as

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

where $S_{PK}(\hat{\tilde{Y}}_i, \tilde{Y}_i) = \frac{\text{Card}(\hat{\tilde{Y}}_i \cap \tilde{Y}_i)}{\text{Card}(\hat{\tilde{Y}}_i \cup \tilde{Y}_i)}$ and $\text{Card}(\tilde{A})$ denotes the cardinal number of $\tilde{A}$ as $\text{Card}(\tilde{A}) = \int_R \tilde{A}(x)dx$.

Definition 5 (Chen and Dang 2008) Suppose that $\tilde{Y}_i$ and $\hat{\tilde{Y}}_i$ are the observed fuzzy response variable and the estimated fuzzy response variable, respectively. The index for goodness of fit of regression model is defined as

$$\bar{G}_1 = \frac{1}{n}\sum_{i=1}^{n} S_{CD}(\hat{\tilde{Y}}_i, \tilde{Y}_i),$$

where $S_{CD}(\hat{\tilde{Y}}_i, \tilde{Y}_i) = \frac{1}{1+E(\hat{\tilde{Y}}_i, \tilde{Y}_i)}$ and $E(\hat{\tilde{Y}}_i, \tilde{Y}_i) = \frac{\int_{S_{\tilde{Y}_i} \cup S_{\hat{\tilde{Y}}_i}} |\hat{\tilde{Y}}_i(y) - \tilde{Y}_i(y)|dy}{\int_{S_{\tilde{Y}_i}} \tilde{Y}_i(y)dy}$. Also, $\hat{\tilde{Y}}_i(y)$ and $\tilde{Y}_i(y)$ are the mem-

Table 2 Comparison study between different models in Example 1

Method	w	Constant	Fuzzy linear regression model	$\bar{G}_1$	$\bar{G}_2$	MSM
SLF	0.05		$\hat{\hat{Y}}_{SLF} = (21.977, 0.760)_T \oplus ((-0.222x_1) \oplus ((2.473, 0.139)_T \otimes x_2)$	0.340	0.496	0.284
	0.1		$\hat{\hat{Y}}_{SLF} = (21.977, 1.520)_T \oplus ((-0.222x_1) \oplus ((2.473, 0.278)_T \otimes x_2)$	0.464	0.546	0.435
	0.2		$\hat{\hat{Y}}_{SLF} = (21.977, 3.039)_T \oplus ((-0.222x_1) \oplus ((2.473, 0.556)_T \otimes x_2)$	0.599	0.545	0.606
ALF	0.05		$\hat{\hat{Y}}_{ALF} = (21.208, 0.786)_T \oplus ((-0.201x_1) \oplus ((2.675, 0.123)_T \otimes x_2)$	**0.403**	**0.576**	**0.334**
	0.1		$\hat{\hat{Y}}_{ALF} = (21.318, 1.628)_T \oplus ((-0.205x_1) \oplus ((2.678, 0.231)_T \otimes x_2)$	**0.490**	**0.571**	**0.466**
	0.2		$\hat{\hat{Y}}_{ALF} = (21.560, 3.260)_T \oplus ((-0.210x_1) \oplus ((2.700, 0.460)_T \otimes x_2)$	**0.609**	**0.557**	**0.618**
HLF	0.05	$c = 0.5$	$\hat{\hat{Y}}_{HFL} = (21.144, 0.821)_T \oplus ((-0.202x_1) \oplus ((2.727, 0.119)_T \otimes x_2)$	**0.390**	**0.573**	**0.331**
		$c = 1.5$	$\hat{\hat{Y}}_{HFL} = (21.758, 0.825)_T \oplus ((-0.221x_1) \oplus ((2.622, 0.126)_T \otimes x_2)$	0.358	0.551	0.294
		$c = 2$	$\hat{\hat{Y}}_{HFL} = (21.869, 0.819)_T \oplus ((-0.223x_1) \oplus ((2.573, 0.125)_T \otimes x_2)$	0.355	0.547	0.278
	0.1	$c = 0.5$	$\hat{\hat{Y}}_{HFL} = (21.156, 1.641)_T \oplus ((-0.202x_1) \oplus ((2.725, 0.241)_T \otimes x_2)$	**0.486**	**0.569**	**0.467**
		$c = 1.5$	$\hat{\hat{Y}}_{HFL} = (21.743, 1.650)_T \oplus ((-0.220x_1) \oplus ((2.622, 0.251)_T \otimes x_2)$	0.465	0.552	0.447
		$c = 2$	$\hat{\hat{Y}}_{HFL} = (21.869, 1.638)_T \oplus ((-0.223x_1) \oplus ((2.573, 0.249)_T \otimes x_2)$	0.459	0.546	0.439
	0.2	$c = 0.5$	$\hat{\hat{Y}}_{HFL} = (21.185, 3.285)_T \oplus ((-0.203x_1) \oplus ((2.719, 0.482)_T \otimes x_2)$	**0.613**	**0.562**	**0.623**
		$c = 1.5$	$\hat{\hat{Y}}_{HFL} = (21.670, 3.292)_T \oplus ((-0.218x_1) \oplus ((2.621, 0.503)_T \otimes x_2)$	0.601	0.548	0.615
		$c = 2$	$\hat{\hat{Y}}_{HFL} = (21.846, 3.240)_T \oplus ((-0.222x_1) \oplus ((2.565, 0.512)_T \otimes x_2)$	0.544	0.597	0.610
QLF	0.05	$\tau = 0.25$	$\hat{\hat{Y}}_{QFL} = (20.664, 0.954)_T \oplus ((-0.271x_1) \oplus ((3.450, 0.076)_T \otimes x_2)$	0.347	0.517	0.257
		$\tau = 0.5$	$\hat{\hat{Y}}_{QFL} = (21.208, 0.786)_T \oplus ((-0.201x_1) \oplus ((2.675, 0.123)_T \otimes x_2)$	**0.403**	**0.576**	**0.334**
		$\tau = 0.75$	$\hat{\hat{Y}}_{QFL} = (22.164, 1.095)_T \oplus ((-0.180x_1) + 2.239x_2$	0.360	0.532	0.284
	0.1	$\tau = 0.25$	$\hat{\hat{Y}}_{QFL} = (20.446, 1.415)_T \oplus ((-0.266x_1) \oplus ((3.464, 0.291)_T \otimes x_2)$	0.420	0.504	0.367
		$\tau = 0.5$	$\hat{\hat{Y}}_{QFL} = (21.318, 1.628)_T \oplus ((-0.205x_1) \oplus ((2.678, 0.231)_T \otimes x_2)$	**0.490**	**0.571**	**0.469**
		$\tau = 0.75$	$\hat{\hat{Y}}_{QFL} = (22.195, 2.141)_T \oplus ((-0.185x_1) \oplus ((2.286, 0.017)_T \otimes x_2)$	0.452	0.534	0.432
	0.2	$\tau = 0.25$	$\hat{\hat{Y}}_{QFL} = (20.443, 3.488)_T \oplus ((-0.252x_1) \oplus ((3.341, 0.396)_T \otimes x_2)$	0.545	0.497	0.552
		$\tau = 0.5$	$\hat{\hat{Y}}_{QFL} = (21.560, 3.260)_T \oplus ((-0.210x_1) \oplus ((2.700, 0.460)_T \otimes x_2)$	**0.609**	**0.557**	**0.618**
		$\tau = 0.75$	$\hat{\hat{Y}}_{QFL} = (22.100, 4.173)_T \oplus ((-0.175x_1) \oplus ((2.225, 0.136)_T \otimes x_2)$	0.574	0.523	0.588
AR	0.05	$\tau = 0.25$	$\hat{\hat{Y}}_{AR} = (22.633, 0.150)_T \oplus ((-0.185x_1) \oplus ((2.134, 0.304)_T \otimes x_2)$	0.351	0.503	0.191
		$\tau = 0.5$	$\hat{\hat{Y}}_{AR} = (21.078, 0.150)_T \oplus ((-0.195x_1) \oplus ((2.661, 0.304)_T \otimes x_2)$	0.380	0.560	0.255
		$\tau = 0.75$	$\hat{\hat{Y}}_{AR} = (22.178, 0.150)_T \oplus ((-0.176x_1) \oplus ((2.185, 0.304)_T \otimes x_2)$	0.510	0.451	0.426
	0.1	$\tau = 0.25$	$\hat{\hat{Y}}_{AR} = (20.673, 0.318)_T \oplus ((-0.273x_1) \oplus ((3.461, 0.603)_T \otimes x_2)$	0.418	0.487	0.306
		$\tau = 0.5$	$\hat{\hat{Y}}_{AR} = (21.156, 0.318)_T \oplus ((-0.200x_1) \oplus ((2.688, 0.603)_T \otimes x_2)$	0.446	0.535	0.355
		$\tau = 0.75$	$\hat{\hat{Y}}_{AR} = (22.158, 0.318)_T \oplus ((-0.176x_1) \oplus ((2.185, 0.603)_T \otimes x_2)$	0.428	0.505	0.323
	0.2	$\tau = 0.25$	$\hat{\hat{Y}}_{AR} = (20.655, 0.635)_T \oplus ((-0.270x_1) \oplus ((3.439, 1.206)_T \otimes x_2)$	0.524	0.462	0.446
		$\tau = 0.5$	$\hat{\hat{Y}}_{AR} = (22.944, 0.635)_T \oplus ((-0.230x_1) \oplus ((2.724, 1.206)_T \otimes x_2)$	0.486	0.424	0.400
		$\tau = 0.75$	$\hat{\hat{Y}}_{AR} = (22.633, 0.635)_T \oplus ((-0.185x_1) \oplus ((2.134, 1.206)_T \otimes x_2)$	0.510	0.451	0.426

Table 3 Values of response variable, corresponding predictions by cross-validation method, and indices of goodness of fit of Mohammadi and Taheri's data set in Example 1

No.	$\tilde{Y}_i$	$\hat{\tilde{Y}}_i^{(-i)}$	No.	$\tilde{Y}_i$	$\hat{\tilde{Y}}_i^{(-i)}$	Indices
1	$(16.50, 3.30)_T$	$(15.98, 3.53)_T$	13	$(24.40, 4.88)_T$	$(24.33, 4.96)_T$	
2	$(18.60, 3.72)_T$	$(18.04, 4.11)_T$	14	$(21.80, 4.36)_T$	$(21.61, 4.35)_T$	
3	$(19.30, 3.86)_T$	$(19.17, 3.70)_T$	15	$(23.80, 4.76)_T$	$(23.16, 3.81)_T$	MSM=0.776
4	$(20.30, 4.06)_T$	$(20.21, 4.28)_T$	16	$(20.80, 4.16)_T$	$(20.44, 3.79)_T$	
5	$(17.30, 3.46)_T$	$(16.37, 3.72)_T$	17	$(17.50, 3.50)_T$	$(18.99, 3.51)_T$	
6	$(20.40, 4.08)_T$	$(19.70, 4.19)_T$	18	$(17.80, 3.56)_T$	$(17.05, 3.30)_T$	$\bar{G}_1$=0.734
7	$(19.30, 3.86)_T$	$(19.44, 4.22)_T$	19	$(20.20, 4.04)_T$	$(20.21, 3.79)_T$	
8	$(21.90, 4.38)_T$	$(22.00, 4.44)_T$	20	$(20.00, 4.00)_T$	$(20.06, 4.70)_T$	
9	$(15.90, 3.18)_T$	$(15.90, 3.11)_T$	21	$(22.80, 4.56)_T$	$(22.96, 3.66)_T$	$\bar{G}_2$=0.700
10	$(18.30, 3.66)_T$	$(19.51, 3.81)_T$	22	$(19.10, 3.82)_T$	$(17.67, 3.26)_T$	
11	$(22.60, 4.52)_T$	$(22.93, 4.40)_T$	23	$(12.10, 2.42)_T$	$(12.54, 2.98)_T$	
12	$(23.70, 4.74)_T$	$(23.06, 4.45)_T$	24	$(12.80, 2.56)_T$	$(13.60, 3.15)_T$	

Table 4 Data set in Example 2

$No.$	$\tilde{Y}_i$	x_i
1	$(22.50, 7.50)_T$	1
2	$(28.75, 8.75)_T$	2
3	$(25.00, 10.00)_T$	3
4	$(42.50, 17.50)_T$	4
5	$(40.00, 15.00)_T$	5
6	$(52.50, 12.50)_T$	6
7	$(75.00, 20.00)_T$	7
8	$(85.00, 15.00)_T$	8

Table 5 Different outliers in Example 2

Outlier's types	Data No.	The value of outliers
Centers of $\tilde{Y}$	2	$\tilde{Y} = (140.00, 8.75)_T$
	6	$\tilde{Y} = (120.00, 12.50)_T$
Spreads of $\tilde{Y}$	4	$\tilde{Y} = (42.50, 35.00)_T$
	7	$\tilde{Y} = (75.00, 1.00)_T$
Centers and spreads of $\tilde{Y}$	5	$\tilde{Y} = (140.00, 45.00)_T$
	7	$\tilde{Y} = (120.00, 40.00)_T$

bership functions of $\hat{\tilde{Y}}_i$ and $\tilde{Y}_i$ with the supports $S_{\hat{\tilde{Y}}_i}$ and $S_{\tilde{Y}_i}$, respectively.

Definition 6 Suppose that $\tilde{Y}_i$ and $\hat{\tilde{Y}}_i$ are the observed fuzzy response variable and the estimated fuzzy response variable, respectively. Based on the objective function O_{ψ^2}, the index of goodness of fit of the model is defined as

$$\bar{G}_2 = \frac{1}{n}\sum_{i=1}^{n} S_O(\hat{\tilde{Y}}_i, \tilde{Y}_i),$$

where $S_O(\hat{\tilde{Y}}_i, \tilde{Y}_i) = \frac{1}{1+O_{\psi^2}(\hat{\tilde{Y}}_i, \tilde{Y}_i)}$.

Remark 8 The indices MSM, $\bar{G}_1$ and $\bar{G}_2$ are on the interval [0, 1]. The optimal model is the model with maximum values of MSM, $\bar{G}_1$ or $\bar{G}_2$.

3.4.2 Cross-validation

The cross-validation method is a well-known method for evaluating the performance and predictive ability of regression models. For doing this purpose, we divide the data set with size n into two sets. The first set contains the training data of size $n-1$, which we can use to develop the regression model, and the second set is the testing data of size $k=1$, which is used to evaluate the predictive ability of the presented regression model.

In this method, we consider n steps. In each step, the ith observation is first deleted from the data set for $i = 1, 2, \ldots, n$, and then, the fuzzy regression model is obtained based on the remaining observations (training data set). In final, the value of the ith response variable is predicted based on the proposed fuzzy regression model and is denoted as $\hat{\tilde{Y}}_i^{(-i)}$. To evaluate the performance of fuzzy regression model, we calculate the indices of goodness of fit MSM, $\bar{G}_1$ and $\bar{G}_2$ between $\tilde{Y}_i$ and $\hat{\tilde{Y}}_i^{(-i)}$.

4 Numerical examples

In this section, we present some numerical and simulation examples to illustrate our proposed approach.

Example 1 One of the classical problems in soil science is the measurement of physical, chemical, and/or biological soil properties. The problem results from the difficulty, time, and

Table 6 Performance of fuzzy regression models in Example 2

Outlier's type	Loss function	Optimal fuzzy regression model	$\bar{G}_1$	$\bar{G}_2$	MSM
Non-outliers	SLF	$\hat{\tilde{Y}}_{SLF} = (22.500, 7.188)_T \oplus ((0.360, 1.354)_T \otimes x) \oplus ((0.533, 0.000)_T \otimes x^2) \oplus ((0.054, 0.000)_T \otimes x^3)$	0.598	0.248	0.596
	ALF	$\hat{\tilde{Y}}_{ALF} = (13.407, 6.660)_T \oplus ((11.093, 0.898)_T \otimes x) \oplus ((-2.248, 0.000)_T \otimes x^2) \oplus ((0.247, 0.003)_T \otimes x^3)$	**0.755**	**0.556**	**0.699**
	HLF	$\hat{\tilde{Y}}_{HLF} = (13.714, 6.600)_T \oplus ((10.747, 0.898)_T \otimes x) \oplus ((-2.166, 0.000)_T \otimes x^2) \oplus ((0.242, 0.002)_T \otimes x^3)$	0.750	0.536	0.696
	QLF	$\hat{\tilde{Y}}_{QLF} = (25.714, 7.000)_T \oplus ((-5.113, 1.000)_T \otimes x) \oplus ((1.661, 0.000)_T \otimes x^2) \oplus ((-0.010, 0.000)_T \otimes x^3)$	0.743	0.500	0.694
Centers of $\tilde{Y}$	SLF	$\hat{\tilde{Y}}_{SLF} = (45.179, 7.188)_T \oplus ((11.921, 1.354)_T \otimes x) \oplus ((-2.865, 0.000)_T \otimes x^2) \oplus ((0.265, 0.000)_T \otimes x^3)$	0.327	0.059	0.137
	ALF	$\hat{\tilde{Y}}_{ALF} = (55.657, 6.429)_T \oplus ((-47.961, 1.071)_T \otimes x) \oplus ((16.339, 0.000)_T \otimes x^2) \oplus ((-1.235, 0.000)_T \otimes x^3)$	0.580	0.403	0.448
	HLF	$\hat{\tilde{Y}}_{HLF} = (32.706, 6.474)_T \oplus ((-15.643, 1.085)_T \otimes x) \oplus ((5.852, 0.000)_T \otimes x^2) \oplus (-0.385, 0.000)_T \otimes x^3)$	0.585	0.341	0.489
	QLF	$\hat{\tilde{Y}}_{QLF} = (25.914, 6.600)_T \oplus ((-5.062, 0.898)_T \otimes x) \oplus ((1.661, 0.000)_T \otimes x^2) \oplus ((-0.013, 0.000)_T \otimes x^3)$	**0.656**	**0.500**	**0.548**
Spreads of $\tilde{Y}$	SLF	$\hat{\tilde{Y}}_{SLF} = (22.500, 0.000)_T \oplus ((0.360, 2.314)_T \otimes x) \oplus ((0.533, 0.000)_T \otimes x^2) \oplus ((0.054, 0.000)_T \otimes x^3)$	0.527	0.205	0.412
	ALF	$\hat{\tilde{Y}}_{ALF} = (26.821, 6.357)_T \oplus ((-6.232, 1.214)_T \otimes x) \oplus ((1.982, 0.000)_T \otimes x^2) \oplus ((-0.036, 0.000)_T \otimes x^3)$	0.714	0.503	0.647
	HLF	$\hat{\tilde{Y}}_{HLF} = (26.861, 6.563)_T \oplus ((-6.206, 1.171)_T \otimes x) \oplus ((1.975, 0.000)_T \otimes x^2) \oplus ((-0.035, 0.000)_T \otimes x^3)$	0.711	0.491	0.645
	QLF	$\hat{\tilde{Y}}_{QLF} = (26.089, 6.250)_T \oplus ((-5.238, 1.250)_T \otimes x) \oplus ((1.660, 0.000)_T \otimes x^2) \oplus ((-0.012, 0.000)_T \otimes x^3)$	**0.715**	**0.499**	**0.648**
Centers and Spreads of $\tilde{Y}$	SLF	$\hat{\tilde{Y}}_{SLF} = (48.214, 12.936)_T \oplus ((-39.672, 0.000)_T \otimes x) \oplus ((15.722, 0.259)_T \otimes x^2) \oplus ((-1.272, 0.000)_T \otimes x^3)$	0.371	0.092	0.266
	ALF	$\hat{\tilde{Y}}_{ALF} = (55.357, 6.429)_T \oplus ((-47.961, 1.071)_T \otimes x) \oplus ((16.339, 0.000)_T \otimes x^2) \oplus ((-1.235, 0.000)_T \otimes x^3)$	0.618	0.403	0.466
	HLF	$\hat{\tilde{Y}}_{HLF} = (55.868, 6.549)_T \oplus ((-48.609, 1.062)_T \otimes x) \oplus ((16.532, 0.000)_T \otimes x^2) \oplus ((-1.250, 0.000)_T \otimes x^3)$	0.615	0.396	0.465
	QLF	$\hat{\tilde{Y}}_{QLF} = (25.914, 6.600)_T \oplus ((-5.062, 0.898)_T \otimes x) \oplus ((1.661, 0.000)_T \otimes x^2) \oplus ((-0.013, 0.000)_T \otimes x^3)$	**0.685**	**0.500**	**0.549**

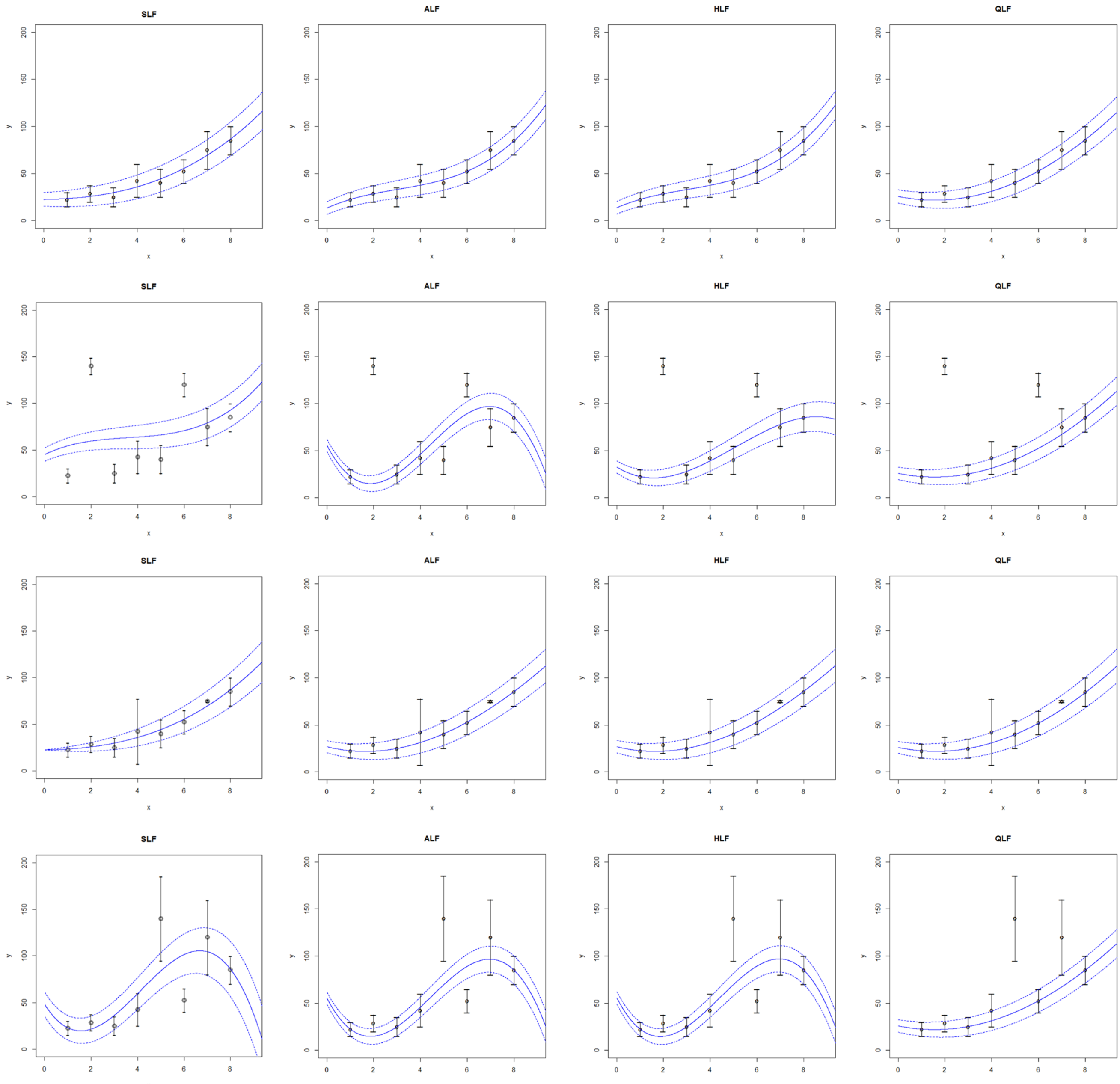

Fig. 1 Comparison between fuzzy regression models based on the different loss functions and outliers in Example 2

cost of direct measurements. Mohammadi and Taheri (2004) provided a data set that it includes some soil properties such as the cation exchange capacity (CEC), the sand content percentage (SAND), and the organic matter content (OM) (see Table 1). Now, we wish to model a relationship between CEC (as the response variable) and SAND and OM (as the explanatory variables) by the regression model:

$$\tilde{Y}_i = \tilde{\beta}_0 \oplus (\tilde{\beta}_1 \otimes x_{i1}) \oplus (\tilde{\beta}_2 \otimes x_{i2}), \quad i = 1, \ldots, 24,$$

where $\tilde{\beta}_j = (\beta_j, \gamma_j)_T$, $j = 0, 1, 2$. But due to some impreciseness in related experimental environment, the response variable $\tilde{Y}_i$ is reported as a symmetric triangular fuzzy number $(y_i, s_i)_T$, in which the spreads are proportional to centers. Here, we consider the sensitivity analysis of model by fuzzifying for different values of spreads as $s_i = wy_i$ with $w = 0.05, 0.10, 0.20$. Note that Arefi (2020) recently presented a quantile fuzzy linear regression model on this data set. Since Arefi's approach (AR) provided a optimal fuzzy regression model on this data set, we compare our work with this approach (the results are listed in Table 2). Results show that our proposed models based on the quantile loss function with $\tau = 0.5$ and Huber loss function with $c = 0.5$ have better performance than the fuzzy model given by Arefi (2020) (for each $w = 0.05, 0.1, 0.2$).

Now, to evaluate the performance and predictive ability of optimal regression model based on the quantile loss function with $\tau = 0.5$, we apply the cross-validation method on data set in Table 1 (with $s_i = 0.2y_i$). The results are given in Table 3. Based on the values of goodness-of-fit indices (see Table 3), we can suggest the proposed fuzzy regression model has a suitable performance for predicting the response variables.

Example 2 Consider the data set in Table 4 (Tanaka and Lee 1998). To study the effect of outliers on the our proposed models, we have listed some of different types of outliers in Table 5. Assume that the original fuzzy regression model is as

$$\tilde{Y}_i = \tilde{\beta}_0 \oplus (\tilde{\beta}_1 \otimes x_{i1}) \oplus (\tilde{\beta}_2 \otimes x_{i1}^2) \oplus (\tilde{\beta}_3 \otimes x_{i1}^3), \quad i = 1, \ldots, 8.$$

Table 6 lists the optimal fitted fuzzy regression models in the presence of outlier data. In original model, the models proposed based on the absolute error loss function and quantile loss functions (at $\tau = 0.75$) have better performance than other models. In the presence of outlier data, the fitted fuzzy regression model based on the quantile loss function at $\tau = 0.75$ is robust because the obtained fuzzy regression models based on the quantile loss function at $\tau = 0.75$ are approximately similar in all cases (see Fig. 1).

Now, we apply the cross-validation method on data set in Table 4 based on the optimal fuzzy regression model with the quantile loss function at $\tau = 0.75$. The results are given in Table 7. The values of goodness-of-fit indices show that the predictive ability of the proposed regression model is suitable.

Table 7 Values of response variable, corresponding predictions by cross-validation method, and indices of goodness of fit on Tanaka and Lee's data set in Example 2

$No.$	$\tilde{Y}_i$	$\hat{\tilde{Y}}_i^{(-i)}$	Indices
1	$(22.50, 7.50)_T$	$(24.25, 6.50)_T$	MSM= 0.540
2	$(28.75, 8.75)_T$	$(26.19, 9.58)_T$	
3	$(25.00, 10.00)_T$	$(32.67, 9.64)_T$	$\bar{G}_1$= 0.580
4	$(42.50, 17.50)_T$	$(35.52, 10.92)_T$	
5	$(40.00, 15.00)_T$	$(40.49, 11.39)_T$	$\bar{G}_2$= 0.239
6	$(52.50, 12.50)_T$	$(54.00, 12.86)_T$	
7	$(75.00, 20.00)_T$	$(66.71, 15.00)_T$	
8	$(85.00, 15.00)_T$	$(89.88, 15.13)_T$	

Table 8 Data set in Example 3

No.	$\tilde{Y}_i$	x_i
1	$(8.00, 1.8)_T$	1
2	$(6.48, 2.2)_T$	2
3	$(9.50, 2.6)_T$	3
4	$(13.50, 2.6)_T$	4
5	$(13.00, 2.4)_T$	5

Example 3 (Tanaka et al. 1982) Consider the data set in Table 8. In this data set, the observations of independent variable are crisp and the observations of dependent variable are presented as the symmetric triangular fuzzy numbers. This data set has been considered by many numbers of researchers. Here, we compare our approach with some other approaches introduced by Diamond (1988) (DM), Chen and Hsueh (2007) (CH_2), Chen and Hsueh (2009) (CH_1), Mosleh et al. (2010) (ME), Taheri and Kelkinnama (2012) (TK), Roldan et al. (2012) (RE), Zeng et al. (2016) (ZE), and Lopez et al. (2016) (HE). Results are given in Table 9. By using the indices of goodness of fit, we can suggest the following cases:

- In between the fuzzy linear regression models based on the least squared errors, the models (SLF), (DM), and (CH_1) have the approximately similar results.
- In between the fuzzy linear regression models based on the least absolute errors, the model (ALF) has better performance than the models (CH_2), (TK), and (ZE).
- The fuzzy linear regression model based on the quantile loss function (QLF) at $\tau = 0.6$ provides a model with better performance than other models considered in this example.

Example 4 (Simulation study) In this example, we want to design a data set that empirically assesses the robust performance of the proposed fuzzy regression model based on the quantile loss function. Based on a sample of size $n = 100$ simulated under a fuzzy regression model, we obtain a data set $(x_i, \tilde{Y}_i), i = 1, \ldots, 100$. Also, we add three different percentages of simulated outliers as $p.out = \{0.05, 0.10, 0.20\}$ in this the data set, and we want to evaluate the proposed models. Note that we assume that the data set is simulated based on the following model

$$\tilde{Y}_i = (13, 4)_T \oplus ((7, 1)_T \otimes \sin(\frac{3.14}{24} x_i)) \oplus ((0.33, 0.01)_T \otimes x_i) \oplus \varepsilon_i, i = 1, \ldots, n,$$

where x_i and ε_i are generated from $U(-40, 50)$ and $N(0, 1.5)$, respectively. The outlier observations are generated according to two different types: Type I is the outliers in the centers of $\tilde{Y}_i$, and Type II is the outliers in the spreads of $\tilde{Y}_i$. We generate these two types of outliers from the bivari-

Table 9 Comparison study between different models in Example 3

Method	Constant	Fuzzy linear regression model	$\bar{G}_1$	$\bar{G}_2$	MSM
DM		$\hat{\tilde{Y}}_{DM} = (4.950, 1.840)_T \oplus ((1.710, 0.160)_T \otimes x)$	0.459	0.488	0.412
CH_1		$\hat{\tilde{Y}}_{CH_1} = 1.710x \oplus (4.950, 2.320)_T$	0.459	0.488	0.414
CH_2		$\hat{\tilde{Y}}_{CH_2} = 4.750 + 1.650x \oplus (0, 2.400)_T$	0.569	0.585	0.512
ME		$\hat{\tilde{Y}}_{ME} = (4.9499, 1.8399, 1.8398)_T \oplus ((1.710, 0.160, 0.1601)_T \otimes x)$	0.459	0.488	0.412
TK		$\hat{\tilde{Y}}_{TK} = 6.375 + 1.325x \oplus (0, 2.400)_T$	0.562	0.589	0.519
RE		$\hat{\tilde{Y}}_{RE} = (4.950 + 1.710 * x, \frac{1}{0.344+0.208/x})_T$	0.459	0.488	0.410
ZE		$\hat{\tilde{Y}}_{ZE} = (4.250, 1.800)_T \oplus ((1.750, 0.200)_T \otimes x)$	0.543	0.565	0.487
HE		$\hat{\tilde{Y}}_{HE} = (\sqrt{45.470 + 5.800 * x^2}, \hat{s}(x))_T$; $\hat{s}(x)$ is Bernstein polynomials of degree $N = 10$ (see Example 4 of Lopez et al. 2016)	0.474	0.508	0.438
SLF		$\hat{\tilde{Y}}_{SLF} = (4.990, 1.840)_T \oplus ((1.702, 0.160)_T \otimes x)$	0.459	0.488	0.411
ALF		$\hat{\tilde{Y}}_{QLF} = (4.500, 2.400)_T \oplus ((1.700, 0.000)_T \otimes x)$	0.580	0.598	0.520
HLF	$c = 0.2$	$\hat{\tilde{Y}}_{HFL} = (5.480, 1.878)_T \oplus ((1.504, 0.104)_T \otimes x)$	0.584	0.566	0.483
	$c = 0.5$	$\hat{\tilde{Y}}_{HFL} = (5.944, 2.011)_T \oplus ((1.411, 0.078)_T \otimes x)$	0.551	0.570	0.496
	$c = 1.3$	$\hat{\tilde{Y}}_{HFL} = (5.442, 1.817)_T \oplus ((1.575, 0.150)_T \otimes x)$	0.480	0.510	0.433
	$c = 2$	$\hat{\tilde{Y}}_{HFL} = (4.990, 1.840)_T \oplus ((1.702, 0.160)_T \otimes x)$	0.459	0.488	0.411
QLF	$\tau = 0.25$	$\hat{\tilde{Y}}_{QLF} = (2.133, 2.067)_T \oplus ((2.173, 0.067)_T \otimes x)$	0.610	0.605	0.527
	$\tau = 0.4$	$\hat{\tilde{Y}}_{QLF} = (4.400, 2.600)_T \oplus ((1.700, 0.000)_T \otimes x)$	0.525	0.550	0.496
	$\tau = 0.6$	$\hat{\tilde{Y}}_{QLF} = (6.750, 1.650)_T \oplus ((1.250, 0.150)_T \otimes x)$	**0.634**	**0.626**	**0.553**
	$\tau = 0.75$	$\hat{\tilde{Y}}_{QLF} = (6.167, 1.533)_T \oplus ((1.833, 0.267)_T \otimes x)$	0.584	0.495	0.478

Table 10 Performance of fuzzy regression model in Example 4

Outlier's types	p.out	Quantile level	Optimal model	MSM	MSM without outliers
No-outliers	–	0.25	$\hat{\tilde{Y}}_{QLF} = (12.019, 3.999)_T \oplus ((6.400, 0.998)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.323, 0.010)_T \otimes x_i)$	0.584	–
		0.5	$\hat{\tilde{Y}}_{QLF} = (12.837, 3.910)_T \oplus ((6.171, 1.027)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.318, 0.012)_T \otimes x_i)$	**0.630**	–
		0.75	$\hat{\tilde{Y}}_{QLF} = (13.648, 4.010)_T \oplus ((6.312, 0.935)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.320, 0.012)_T \otimes x_i)$	0.592	–
Centers of $\tilde{Y}$	0.05	0.25	$\hat{\tilde{Y}}_{QLF} = (11.995, 4.000)_T \oplus ((6.546, 1.000)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.321, 0.010)_T \otimes x_i)$	0.551	0.579
		0.5	$\hat{\tilde{Y}}_{QLF} = (12.849, 4.001)_T \oplus ((6.366, 0.999)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.318, 0.010)_T \otimes x_i)$	**0.603**	**0.634**
		0.75	$\hat{\tilde{Y}}_{QLF} = (13.696, 4.096)_T \oplus ((6.233, 0.995)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.321, 0.005)_T \otimes x_i)$	0.563	0.591
	0.10	0.25	$\hat{\tilde{Y}}_{QLF} = (12.103, 4.001)_T \oplus ((6.392, 0.996)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.315, 0.010)_T \otimes x_i)$	0.536	0.590
		0.5	$\hat{\tilde{Y}}_{QLF} = (13.009, 3.995)_T \oplus ((5.885, 0.997)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.309, 0.010)_T \otimes x_i)$	**0.582**	**0.641**
		0.75	$\hat{\tilde{Y}}_{QLF} = (13.725, 3.999)_T \oplus ((6.049, 0.994)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.314, 0.010)_T \otimes x_i)$	0.536	0.590
	0.20	0.25	$\hat{\tilde{Y}}_{QLF} = (11.993, 4.000)_T \oplus ((6.551, 1.002)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.321, 0.010)_T \otimes x_i)$	0.550	0.578
		0.5	$\hat{\tilde{Y}}_{QLF} = (13.058, 3.974)_T \oplus ((6.067, 0.996)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.314, 0.011)_T \otimes x_i)$	**0.530**	**0.635**
		0.75	$\hat{\tilde{Y}}_{0.75} = (14.087, 4.000)_T \oplus ((6.111, 1.000)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.336, 0.010)_T \otimes x_i)$	0.442	0.530
Spreads of $\tilde{Y}$	0.05	0.25	$\hat{\tilde{Y}}_{QLF} = (12.037, 4.171)_T \oplus ((6.382, 0.806)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.323, 0.011)_T \otimes x_i)$	0.577	0.590
		0.5	$\hat{\tilde{Y}}_{QLF} = (12.870, 3.977)_T \oplus ((6.093, 1.500)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.315, 0.014)_T \otimes x_i)$	**0.611**	**0.627**
		0.75	$\hat{\tilde{Y}}_{QLF} = (13.682, 4.258)_T \oplus ((6.092, 0.934)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.323, 0.012)_T \otimes x_i)$	0.584	0.597
	0.10	0.25	$\hat{\tilde{Y}}_{QLF} = (11.909, 4.432)_T \oplus ((6.545, 0.509)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.324, 0.010)_T \otimes x_i)$	0.555	0.578
		0.5	$\hat{\tilde{Y}}_{QLF} = (12.923, 4.495)_T \oplus ((6.183, 0.888)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.316, 0.009)_T \otimes x_i)$	**0.611**	**0.640**
		0.75	$\hat{\tilde{Y}}_{QLF} = (13.806, 4.542)_T \oplus ((6.160, 0.892)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.326, 0.031)_T \otimes x_i)$	0.558	0.580
	0.20	0.25	$\hat{\tilde{Y}}_{QLF} = (11.691, 5.026)_T \oplus ((6.699, 0.152)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.327, 0.000)_T \otimes x_i)$	0.520	0.553
		0.5	$\hat{\tilde{Y}}_{QLF} = (12.929, 4.702)_T \oplus ((6.160, 0.811)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.317, 0.008)_T \otimes x_i)$	**0.585**	**0.636**
		0.75	$\hat{\tilde{Y}}_{QLF} = (14.107, 5.189)_T \oplus ((6.009, 0.565)_T \otimes \sin(\frac{3.14}{24}x_i)) \oplus ((0.322, 0.022)_T \otimes x_i)$	0.513	0.542

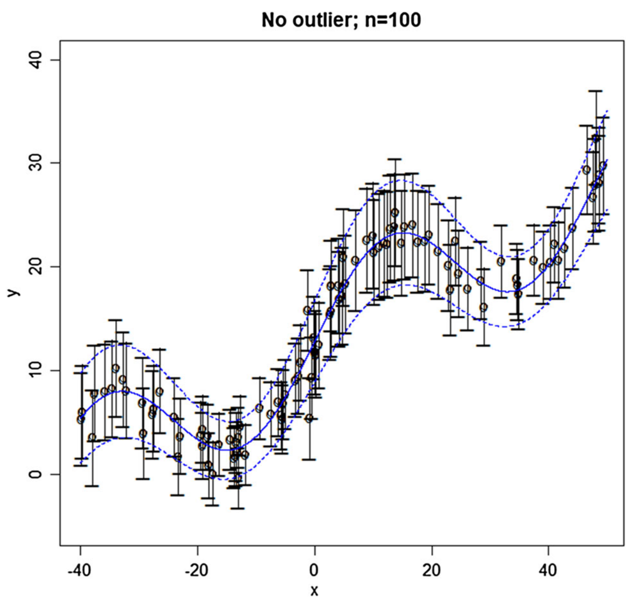

Fig. 2 Fuzzy regression model under original data and based on quantile loss function with $\tau = 0.5$ in Example 4

ate Gaussian distribution $N_2\left(\begin{bmatrix}\mu_x\\ \mu_y\end{bmatrix}, \begin{bmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y\\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{bmatrix}\right)$ and uniform distribution, respectively. In Type I, we generate observations (x_i, y_i) from two bivariate Gaussian distributions $N_2\left(\begin{bmatrix}\bar{x}+\varepsilon_i\\ max(y_i)+1.5s_y\end{bmatrix}, \begin{bmatrix}20 & 0\\ 0 & 5\end{bmatrix}\right)$ and $N_2\left(\begin{bmatrix}\bar{x}+\varepsilon_i\\ min(y_i)-1.5s_y\end{bmatrix}, \begin{bmatrix}20 & 0\\ 0 & 5\end{bmatrix}\right)$, where ε_i is generated from $U(-20, 40)$ and $s_y = \sqrt{\frac{1}{n}\sum_{i=1}^{100}(y_i - \overline{y})^2}$. In Type II, we generate the spreads of triangular fuzzy response observations $(y_i, s_i)_T$ based on $U(9, 13)$. The results of the fitted fuzzy regression models based on the quantile loss function with $\tau = 0.25, 0.5, 0.75$ are given in Table 10. Based on the index of MSM, the model based on the quantile loss function with $\tau = 0.5$ has the best results. See Figs. 2 and 3 for the performance of the best fuzzy regression models (using quantile loss function at $\tau = 0.5$). They show the robustness performances from our proposed model in original data set and in the models with the presence of outliers. Also, we can consider the similar results based on the results given in Tables 10 and 11. Based on these tables, we have

- The MSM in the original data and the MSM without outliers (in the models with outliers) are approximately similar (see Table 10).

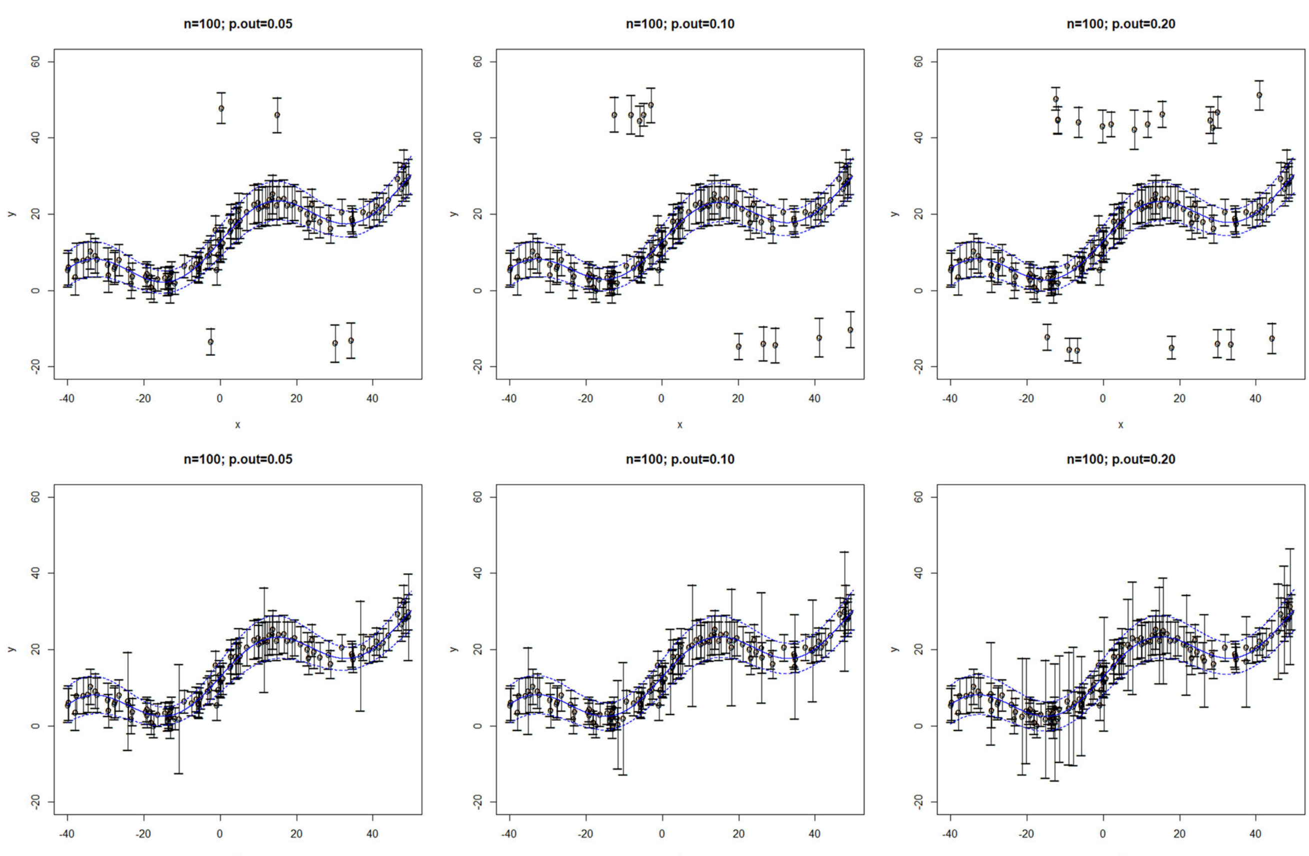

Fig. 3 Fuzzy regression models in the presence of outliers and based on quantile loss function with $\tau = 0.5$ in Example 4

Table 11 Mean values of the estimated parameters in Example 4

Outlier's types	Quantile level	Means of $\tilde{\beta}_0$		Means of $\tilde{\beta}_1$		Means of $\tilde{\beta}_2$	
		β_0	γ_0	β_1	γ_1	β_2	γ_2
Centers of $\tilde{Y}$	0.25	12.030	4.000	6.496	0.999	0.319	0.010
	0.50	12.972	3.990	6.106	0.997	0.314	0.010
	0.75	13.835	4.032	6.131	0.996	0.324	0.008
Spreads of $\tilde{Y}$	0.25	11.879	4.543	6.542	0.489	0.325	0.007
	0.50	12.907	4.391	6.155	1.066	0.316	0.010
	0.75	13.865	4.663	6.089	0.799	0.324	0.022

- The values of the estimated parameters with original data and the mean values of the estimated parameters in the presence of outliers are approximately similar (see Table 11).

5 Conclusion

In this paper, we present a new approach to fit the fuzzy regression models based on some loss functions when the response variable and the parameters of model are as fuzzy numbers. Some of certain merits in this approach are as follows:

(1) It is a general approach for fitting the fuzzy regression models based on the different types of loss functions.
(2) A new definition of the objective function is presented based on loss function and the average of differences between the α-cuts of errors.
(3) Among the proposed loss functions, the fuzzy regression models based on the quantile loss function and Huber loss function are robust under outlier data.
(4) To evaluate the goodness of fit of the proposed fuzzy regression models, we introduce three indices.

Compliance with ethical standards

Conflict of interest The authors declare that there is no conflict of interest regarding the publication of this paper.

Ethical approval This article does not contain any studies with human participants or animals.

References

Arefi M (2020) Quantile fuzzy regression based on fuzzy outputs and fuzzy parameters. Soft Comput 24:311–320
Arefi M, Taheri SM (2015) Least-squares regression based on Atanassov's intuitionistic fuzzy inputs–outputs and Atanassov's intuitionistic fuzzy parameters. IEEE Trans Fuzzy Syst 23:1142–1154
Chachi J (2019) A weighted least-squares fuzzy regression for crisp input-fuzzy output data. IEEE Trans Fuzzy Syst 27:739–748
Chachi J, Roozbeh M (2017) A fuzzy robust regression approach applied to bedload transport data. Commun Stat Simul Comput 47:1703–1714
Chachi J, Taheri SM (2016) Multiple fuzzy regression model for fuzzy input–output data. Iran J Fuzzy Syst 13:63–78
Chang PT, Lee CH (1994) Fuzzy least absolute deviations regression based on the ranking of fuzzy numbers. In: Proceedings of the third IEEE world congress on computational intelligence, Orlando, FL, vol 2, pp 1365–1369
Chen SP, Dang JF (2008) A variable spread fuzzy linear regression model with higher explanatory power and forecasting accuracy. Inf Sci 178:3973–3988
Chen LH, Hsueh CC (2007) A mathematical programming method for formulating a fuzzy regression model based on distance criterion. IEEE Trans Cybern 37:705–12
Chen LH, Hsueh CC (2009) Fuzzy regression models using the least-squares method based on the concept of distance. IEEE Trans Fuzzy Syst 17:1259–1272
Choi SH, Buckley JJ (2008) Fuzzy regression using least absolute deviation estimators. Soft Comput 12:257–263
Diamond P (1988) Fuzzy least squares. Inf Sci 46:141–157
D'Urso P, Massari R (2013) Weighted least squares and least median squares estimation for the fuzzy linear regression analysis. Metron 71:279–306
D'Urso P, Massari R, Santoro A (2011) Robust fuzzy regression analysis. Inf Sci 181:4154–4174
Geisser S (1993) Predictive inference, vol 55. CRC Press, Boca Raton
Hao PY, Chiang JH (2008) Fuzzy regression analysis by support vector learning approach. IEEE Trans Fuzzy Syst 16:428–441
Hesamian G, Akbari MG (2019) Fuzzy quantile linear regression model adopted with a semi-parametric technique based on fuzzy predictors and fuzzy responses. Expert Syst Appl 118:585–597
Huber PJ (1981) Robust statistics. Wiley, New York, pp 153–195
Khammar AH, Arefi M, Akbari MG (2020) A robust least squares fuzzy regression model based on kernel function. Iran J Fuzzy Syst 17:105–119
Koenker R (2005) Quantile regression. Cambridge University Press, Cambridge
Kula KS, Tank F, Dalkyly TE (2012) A study on fuzzy robust regression and its application to insurance. Math Comput Appl 17:223–234
Lopez R, de Hierro AF, Martinez-Morenob J, Aguilar-Pena C, Lopez R, de Hierro C (2016) A fuzzy regression approach using Bernstein polynomials for the spreads: computational aspects and applications to economic models. Math Comput Simul 128:13–25
Mohammadi J, Taheri SM (2004) Pedomodels fitting with fuzzy least squares regression. Iran J Fuzzy Syst 1:45–61

Mosleh M, Otadi M, Abbasbandy S (2010) Evaluation of fuzzy regression models by fuzzy neural network. J Comput Appl Math 234:825–834

Nasrabadi E, Hashemi SM (2008) Robust fuzzy regression analysis using neural networks. Int J Uncertainty Fuzziness Knowl Based Syst 16:579–598

Oussalah M, De Schutter J (2002) Robust fuzzy linear regression and application for contact identification. Intell Autom Soft Comput 8:31–39

Pappis CP, Karacapilidis NI (1993) A comparative assessment of measure of similarity of fuzzy values. Fuzzy Sets Syst 56:171–174

Rapaic D, Krstanovic L, Ralevic N, Obradovic R, Klipa C (2019) Sparse regularized fuzzy regression. Appl Anal Discrete Math 13:583–604

Roldan C, Roldan A, Martinez-Moreno J (2012) A fuzzy regression model based on distances and random variables with crisp input and fuzzy output data: a case study in biomass production. Soft Comput 16:785–795

Sanli K, Apaydin A (2004) The fuzzy robust regression analysis, the case of fuzzy data set has outlier. Gazi Univ J Sci 17:71–84

Schrage L (2006) Optimization Modeling with Lingo, 6th edn. Lindo Systems, Chicago

Shon BY (2005) Robust fuzzy linear regression based on M-estimators. Appl Math Comput 18:591–601

Taheri SM, Kelkinnama M (2012) Fuzzy linear regression based on least absolute deviations. Iran J Fuzzy Syst 9:121–140

Tanaka H, Lee H (1998) Interval regression analysis by quadratic programming approach. IEEE Trans Syst 6:473–481

Tanaka H, Uegima S, Asai K (1982) Linear regression analysis with fuzzy model. IEEE Trans Syst 12:903–907

Wasserman L (2006) All of nonparametric statistics. Springer, Berlin

Wolfram S (2003) The Mathematica Book, 5th edn. Wolfram Media Inc, USA

Zeng W, Feng Q, Lia J (2016) Fuzzy least absolute linear regression. Appl Soft Comput 52:1009–1019

Zimmermann HJ (2001) Fuzzy set theory and its applications, 4th edn. Kluwer Nihoff, Boston