Available online at www.sciencedirect.com
ScienceDirect
Fuzzy Sets and Systems 160 (2009) 2505 – 2523
FUZZY sets and systems
www.elsevier.com/locate/fss

An enhanced fuzzy linear regression model with more flexible spreads

Jingli Lu, Ruili Wang*

School of Engineering and Advance Technology, Massey University, Private Bag 11-222, Palmerston North, New Zealand

Received 15 May 2008; received in revised form 13 February 2009; accepted 18 February 2009
Available online 4 March 2009

Abstract

One of the deficiencies of previous fuzzy linear regression models is that with the increase of the magnitudes of independent variables, the spreads of estimated fuzzy dependent variables are increasing, even though the spreads of observed dependent variables actually decrease or remain unchanged. Some solutions have been proposed to solve this spreads increasing problem. However, those solutions still cannot model a decreasing trend in the spreads of the observed dependent variables as the magnitudes of the independent variables increase. In this paper we propose an enhanced fuzzy linear regression model (model FLR_{FS}), in which the spreads of the estimated dependent variables are able to fit the spreads of the observed dependent variables, no matter the spreads of the observed dependent variables are increased, decreased or unchanged as the magnitudes and spreads of the independent variables change. Four numerical examples are used to demonstrate the effectiveness of model FLR_{FS}.

Keywords: Fuzzy linear regression; Fuzzy number; Linear programming; Least-square method

1. Introduction

Fuzzy linear regression (FLR) was first proposed by Tanaka et al. [25] as an extension of the classical regression analysis, which is becoming a powerful tool to explore the vague relationship between dependent and independent variables [3]. In fuzzy regression, some elements of the regression models are represented by imprecise data.

General FLR models for crisp input–fuzzy output data [25] and fuzzy input–fuzzy output data [21] can be represented as follows, respectively:

$$\hat{\tilde{Y}}_i = \tilde{A}_0 + \tilde{A}_1 x_{i1} + \cdots + \tilde{A}_j x_{ij} + \cdots + \tilde{A}_m x_{im} \qquad (\mathrm{FLR}_{CF})$$

$$\hat{\tilde{Y}}_i = \tilde{A}_0 + \tilde{A}_1 \tilde{X}_{i1} + \cdots + \tilde{A}_j \tilde{X}_{ij} + \cdots + \tilde{A}_m \tilde{X}_{im} \qquad (\mathrm{FLR}_{FF})$$

where $\tilde{A}_j$ is the jth fuzzy regression coefficients, x_{ij} or $\tilde{X}_{ij}$ is the jth independent variable of the ith instance, $x_{i0}(\tilde{X}_{i0})$ is 1, $\hat{\tilde{Y}}_i$ is the ith estimated dependent variable, $i = 1, 2, \ldots, n$, $j = 0, 1, \ldots, m$. A tilde character ($\sim$) is placed above

* Corresponding author. Tel.: +64 21 077 2118.
E-mail address: r.wang@massey.ac.nz (R. Wang).

0165-0114/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.fss.2009.02.023

the name of a fuzzy variable to distinguish a fuzzy variable from a crisp variable. As crisp numbers are special fuzzy numbers, model FLR_{CF} can be treated as a special case of model FLR_{FF}.

The methods to estimate the fuzzy regression coefficients can be roughly categorized into two groups. One is the linear programming (LP) methods [18,23,25]; the other is the least-squares (LS) methods [1,4–6,16,27]. The LP methods minimize the total spread of the estimated dependent variables or that of the fuzzy regression coefficients, subject to the constraint that the estimated dependent variables include the observed dependent variables within a certain h-level. The advantage of the LP methods is low computational complexity. However, the LP methods have been criticized by Redden and Woodall [20] as (i) they are extremely sensitive to outliers [10]; (ii) they do not allow all observations to contribute to the estimation; and (iii) the estimated intervals become wider as more data are collected. Multi-objective fuzzy regression techniques are developed to overcome these deficiencies of the LP methods [18,19,22,26]. The LS methods minimize the total difference between the estimated dependent variables and their observed counterparts. Thus, compared with the estimations of the LP methods, the estimations of the LS methods have relatively small differences between the estimated dependent variables and the observed ones. However, the LS methods have relatively higher computational complexity. A comprehensive literature review of fuzzy regression can be found in [14].

As indicated in [2,11–13,17], a problem of model FLR_{FF} is that with the increase of the magnitudes of independent variables, the spreads of estimated dependent variables are increasing (refer to Section 2.3), even though the spreads of observed dependent variables are roughly constant or decreasing. We call it *spreads increasing problem* (refer to Section 2.3) in this paper. Some models [2,5,8,11–13,17], which address this problem, and their deficiencies are briefly discussed below. More details are given in Section 3.

FLR models presented in [2,5,8] can avoid the *spreads increasing problem* by modelling centres and spreads of dependent variables separately. However, the number of parameters to be estimated in model FLR_{CD08} [2] proportionally increases with the increase of the number of instances. Although more parameters involved in a regression model increase the model fitness, these also decrease the model generality [13]. Therefore, model FLR_{CD08} is unsuitable for large dataset regression (refer to Section 3.4). In models $FLR_{D'Urso03}$ [8] and $FLR_{Coppi06}$ [5], the spreads of estimated dependent variables are only determined by the centres of the estimated dependent variables. This limits the ability of $FLR_{D'Urso03}$ and $FLR_{Coppi06}$ to model the spreads of the dependent variables by independent variable (refer to Section 3.3).

Although solutions proposed in [11,12,17] also alleviate *the spreads increasing problem*, these solutions still cannot model a *decreasing* trend in the spreads of the observed dependent variables, as the magnitudes of the independent variables increase. For example, in these models [11,12,17], if the independent variables are crisp, the spreads of the estimated dependent variables can only be a constant (refer to Section 3), even though the spreads of the observed dependent variables are decreasing with the increase of the magnitudes of the independent variables, as shown in Example 1.

Example 1. In Table 1, the independent variable is the height of the male candidates; and the fuzzy dependent variable measures how a candidate's height belongs to the concept *high*. L-type fuzzy numbers in the form of (m_y, α_y) are used to describe *high* (for a detailed description of L-type fuzzy number, refer to Section 2). m_y is the centre of a fuzzy number, which measures the possibility of a given candidate's height belonging to *high*. In this example, m_y is not greater than 1. α_y is the spread of a fuzzy number, which describes the vagueness of m_y. The taller a candidate's height is, closer the possibility of the candidate's height is to 1, and lessens the vagueness of the candidate's height belonging to *high*. However, it is difficult to model this relationship between the candidates' heights and *high* by model FLR_{FF}, because of the *spreads increasing problem* in model FLR_{FF}, which is that the estimated dependent variables can only increase with the magnitudes of the independent variables. Moreover, neither the models proposed in [11,12] nor the model proposed in [17] can capture the relationship between height and *high*, because in these models, when the independent variables are crisp, the spread of the estimated dependent variable can only be a constant (refer to Section 3), which is not true for dataset1 in Table 1.

Note that another problem of modelling the relationship between the candidates' heights and *high* by FLR_{FF} is that the estimated spreads of *high* maybe negative, since the relationship between the candidates' heights and *high* is not strictly linear. When the heights are greater than 2.1, the spreads of observed *high* stop decreasing and the spreads of estimated *high* are negative. Following the arguments in D'Urso [8] and Coppi et al. [5], negative predicted spreads can be interpreted as a lack of uncertainty and set to 0.

Table 1
Dataset1.

i	Height	$High(m_y, \alpha_y)_L$
1	1.7	$(0.60, 0.30)_L$
2	1.8	$(0.70, 0.25)_L$
3	1.9	$(0.80, 0.10)_L$
4	2.0	$(0.90, 0.05)_L$
5	2.1	$(1.00, 0.00)_L$

From the above, we can see that in the previous FLR models the increasing trend in the spreads of the estimated dependent variable limits the ability of FLR to model the relationship between the dependent and independent variables. To alleviate this problem, in this paper, we propose a flexible spreads FLR model (FLR_{FS}). In our model FLR_{FS}, the spreads of estimated dependent variables are able to fit the spreads of observed dependent variables, no matter if the spreads of the observed dependent variables are increased, decreased or unchanged, as the magnitudes and the spreads of the independent variables change.

This paper is organized as follows. In Section 2, we provide a brief introduction to fuzzy numbers and FLR, then describe the *spreads increasing problem* in more detail. Related literatures to solve the *spreads increasing problem* are reviewed in Section 3. In Section 4, a new FLR model, FLR_{FS}, is proposed, which is able to model the linear relationship between the dependent and independent variables better than the previous models. Four numerical experiments are used to demonstrate the effectiveness of model FLR_{FS} in Section 5. Section 6 gives our conclusions and future work.

2. Fuzzy number and FLR

In this section, we briefly introduce fuzzy numbers and the arithmetic rules of fuzzy numbers; then describe the *spreads increasing problem.*

2.1. Fuzzy number

The definition of fuzzy numbers given by Dubois and Prade [7] is as follows.

Definition 2.1. A fuzzy number $\tilde{A}$ is a convex normalized fuzzy set of the real line **R**; its membership function $\mu_{\tilde{A}}(x)$ satisfies the following criteria:

(i) α-cut set of $\tilde{A}$, $\mu_\alpha = \{x|\mu_{\tilde{A}}(x)\geqslant\alpha\}$, is a closed interval;
(ii) $\mu_1 = \{x|\mu_{\tilde{A}}(x) = 1\}$ is non-empty;
(iii) convexity: for $\lambda \in [0, 1]$, $\mu_{\tilde{A}}(\lambda x_1 + (1-\lambda)x_2) \geqslant \min(\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_2))$.

Definition 2.2. As defined in [5,8,29], an *LR*-type fuzzy number $\tilde{A}$ is

$$\mu_{\tilde{A}}(x) = \begin{cases} L\left(\dfrac{m_a - x}{\alpha_a}\right) & \text{for } x \leqslant m_a \\ R\left(\dfrac{x - m_a}{\beta_a}\right) & \text{for } x > m_a \end{cases}$$

where m_a is called centre or mean value, and α_a and β_a are called left and right spreads, respectively, $\alpha_a, \beta_a > 0$. $L(z)$ and $R(z)$ are reference functions that map $\Re^+ \to [0, 1]$, and strictly decreasing for $z \geqslant 0$. Also, L (or R) satisfies the following conditions: if $L(0) = 1$, $L(x) < 1$ for $\forall x > 0$; $L(x) > 0$ for $\forall x < 1$; $L(1) = 0$, or $[L(x) > 0, \forall x$ and $L(+\infty) = 0]$. $\tilde{A}$ can be denoted as $\tilde{A} = (m_a, \alpha_a, \beta_a)_{LR}$. If $\alpha_a = \beta_a$, then $\tilde{A}$ is symmetric, $\tilde{A} = (m_a, \alpha_a)_L$, which is called L-type fuzzy number.

Definition 2.3. If $L(x) = R(x) = 1 - x$, $\tilde{A}$ is a triangular fuzzy number.[1] Furthermore, if $\alpha_a = \beta_a$, then $\tilde{A}$ is a symmetric triangular fuzzy number.

2.2. Arithmetic operations on fuzzy numbers

By applying Zadeh's extension principle [28], the arithmetic operations of fuzzy numbers can be expressed as follows:

$$(\tilde{A} + \tilde{B})(z) = \sup_{x+y=z} T(\tilde{A}(x) + \tilde{B}(y))$$

$$(\tilde{A} * \tilde{B})(z) = \sup_{x*y=z} T(\tilde{A}(x) + \tilde{B}(y))$$

where $T(\cdot)$ is a triangular norm. The T-norm based *LR*-type fuzzy number addition preserves the shape. However, multiplication is not shape preserving, namely, the product of two *LR*-type fuzzy numbers may not be *LR*-type.

Dubois and Prade [7] provided an approximation form for *LR*-type fuzzy number multiplication. According to their approximation formulas, the multiplication of two *LR*-type fuzzy numbers can be presented as follows:

(i) if $\tilde{A} > 0$ and $\tilde{B} > 0$,

$$(m_a, \alpha_a, \beta_a)_{LR} \cdot (m_b, \alpha_b, \beta_b)_{LR} \approx (m_a m_b, m_a\alpha_b + m_b\alpha_a, m_a\beta_b + m_b\beta_a)_{LR}$$

(ii) if $\tilde{A} < 0$ and $\tilde{B} > 0$,

$$(m_a, \alpha_a, \beta_a)_{LR} \cdot (m_b, \alpha_b, \beta_b)_{LR} \approx (m_a m_b, -m_a\beta_b + m_b\alpha_a, -m_a\alpha_b + m_b\beta_a)_{LR}$$

(iii) if $\tilde{A} < 0$ and $\tilde{B} < 0$,

$$(m_a, \alpha_a, \beta_a)_{LR} \cdot (m_b, \alpha_b, \beta_b)_{LR} \approx (m_a m_b, -m_a\beta_b - m_b\beta_a, -m_a\alpha_b - m_b\alpha_a)_{LR}$$

2.3. Spreads increasing problem

For simplicity, most research considers $\tilde{A}_j$, $\tilde{X}_{ij}$ and $\hat{\tilde{Y}}_i$ in model FLR_{FF} as *LR*-type fuzzy numbers or triangular fuzzy numbers. By using the approximation formulas of Dubois and Prade [7], Yang and Lin [27] described model FLR_{FF} as

$$\hat{\tilde{Y}}_i = \tilde{A}_0 + \tilde{A}_1\tilde{X}_{i1} + \cdots + \tilde{A}_m\tilde{X}_{im} \approx (m_{\hat{y}_i}, \alpha_{\hat{y}_i}, \beta_{\hat{y}_i})_{LR}$$

$$m_{\hat{y}_i} = m_{a_0} + \sum_{j=1}^{m} m_{a_j} m_{x_{ij}}$$

$$\alpha_{\hat{y}_i} = \alpha_{a_0} + \sum_{\tilde{A}_j>0, j=1}^{m} [s_{ij}(m_{a_j}\alpha_{x_{ij}} + m_{x_{ij}}\alpha_{a_j}) + (1 - s_{ij})(m_{a_j}\alpha_{x_{ij}} - m_{x_{ij}}\beta_{a_j})]$$
$$+ \sum_{\tilde{A}_j<0, j=1}^{m} [s_{ij}(-m_{a_j}\beta_{x_{ij}} + m_{x_{ij}}\alpha_{a_j}) + (1 - s_{ij})(-m_{a_j}\beta_{x_{ij}} - m_{x_{ij}}\beta_{a_j})] \quad (1)$$

$$\beta_{\hat{y}_i} = \beta_{a_0} + \sum_{\tilde{A}_j>0, j=1}^{m} [s_{ij}(m_{a_j}\beta_{x_{ij}} + m_{x_{ij}}\beta_{a_j}) + (1 - s_{ij})(m_{a_j}\beta_{x_{ij}} - m_{x_{ij}}\alpha_{a_j})]$$
$$+ \sum_{\tilde{A}_j<0, j=1}^{m} [s_{ij}(-m_{a_j}\alpha_{x_{ij}} + m_{x_{ij}}\beta_{a_j}) + (1 - s_{ij})(-m_{a_j}\alpha_{x_{ij}} - m_{x_{ij}}\alpha_{a_j})] \quad (2)$$

$$s_{ij} = 1 \text{ if } \tilde{X}_{ij} \geqslant 0; \quad s_{ij} = 0 \text{ if } \tilde{X}_{ij} < 0$$

[1] For easy explanation, we assume that all *LR*-type fuzzy numbers in this paper are triangular fuzzy numbers.

From Eqs. (1) and (2), we can see that as the magnitude of the independent variable (i.e., $|m_{x_{ij}}|$) increase, the spreads of the estimated dependent variable (i.e., $\alpha_{\hat{y}_i}$ and $\beta_{\hat{y}_i}$) increases. For example, when $\tilde{A}_j > 0$ and $\tilde{X}_{ij} > 0$, the left and right spreads of $\hat{\tilde{Y}}_i$ are

$$\alpha_{\hat{y}_i} = \alpha_{a_0} + \sum_{\tilde{A}_j>0,\, j=1}^{m} (m_{a_j}\alpha_{x_{ij}} + m_{x_{ij}}\alpha_{a_j}), \quad \beta_{\hat{y}_i} = \beta_{a_0} + \sum_{\tilde{A}_j>0,\, j=1}^{m} [s_{ij}(m_{a_j}\beta_{x_{ij}} + m_{x_{ij}}\beta_{a_j})]$$

which increase with the increase of $|m_{x_{ij}}|$. Similarly, we can deduce that the spreads of $\hat{\tilde{Y}}_i$ increase as $|m_{x_{ij}}|$ increase, when $\tilde{A}_j > 0$ and $\tilde{X}_{ij} < 0$ ($\tilde{A}_j < 0$ and $\tilde{X}_{ij} < 0$; or $\tilde{A}_j < 0$ and $\tilde{X}_{ij} > 0$).

It is the inherent property of model FLR_{FF} that determines the spreads of $\hat{\tilde{Y}}_i$ increasing with the increase of $|m_{x_{ij}}|$. This property will affect the regression performance of model FLR_{FF}, when the spreads of the observed dependent variable are not increasing as the magnitude of $\tilde{X}_{ij}$ increase. We name this property as *spreads increasing problem* of model FLR_{FF} in this paper.

3. Review on related literatures

The *spreads increasing problem* has been addressed in several papers [2,5,8,11,12,17], and some solutions have been proposed. However, the previous solutions still have some deficiencies.

3.1. Models FLR_{KC02} and FLR_{KC03}

Kao and Chyu [11] proposed a crisp coefficients FLR model (FLR_{KC02}) to tackle the *spreads increasing problem*, which can be expressed as follows:

$$\hat{\tilde{Y}}_i = a_0 + a_1\tilde{X}_{i1} + \cdots + a_j\tilde{X}_{ij} + \cdots + a_m\tilde{X}_{im} + \tilde{\varepsilon} \qquad (\mathrm{FLR}_{KC02})$$
$$\tilde{\varepsilon} = (0, l, r)_{LR}$$

where each coefficient a_j is a crisp number, $j = 0, 1, \ldots, m$; $\tilde{\varepsilon}$ is a triangular fuzzy error term; $\tilde{X}_{ij}=(m_{x_{ij}}, \alpha_{x_{ij}}, \beta_{x_{ij}})_{LR}$. A two-stage methodology is proposed to obtain the crisp coefficients and the fuzzy error term. The first stage is to estimate crisp coefficients a_j by applying the classical LS method to the defuzzified (such as centroids) independent and dependent variables. In the second stage, fuzzy error term $\tilde{\varepsilon}$ is determined by minimizing the total difference between the membership values of the estimated dependent variables and those of the observed dependent variables. Totally, there are $m + 3$ parameters to be estimated in model FLR_{KC02}, which are l, r, and a_j's.

For crisp independent variables, a deficiency of model FLR_{KC02} is that the spreads of each estimated response variable are the spreads of $\tilde{\varepsilon}$, which are always constants. An example of model FLR_{KC02} for a single crisp independent variable is as follows:

$$\hat{\tilde{Y}}_i = 4.95 + 1.71x_i + (0, 3.01, 1.80)_{LR} \qquad (3)$$

In the above model, the left and right spreads of all the estimated response variables are the spreads of $\tilde{\varepsilon}$, which is 3.01 and 1.80, respectively, even though the spreads of the observed response variables change as the independent variables change.

For fuzzy independent variables, a deficiency of model FLR_{KC02} is that the spread of each estimated response variable cannot be less than a constant. For instance, the left spread of the ith instance cannot be less than $a_1\alpha_{x_{i1}} + \cdots + a_m\alpha_{x_m}$. However, $a_1\alpha_{x_{i1}} + \cdots + a_m\alpha_{x_m}$ has no relationship with the left spread of the ith observed response variable. A numerical example of model FLR_{KC02} for a single fuzzy independent variable is as follows:

$$\hat{\tilde{Y}}_i = 3.5724 + 0.5193\tilde{X}_i + (0, 0.24, 0.24)_{LR} \qquad (4)$$

The left spreads of the estimated responses are $0.5193\alpha_{x_i} + 0.24$. However, crisp coefficient 3.5724 and 0.5193 are determined by applying the classical LS method to the centroids of the independent and dependent variables, which have no relationship with the spreads of the observed dependent variables.

The model proposed in [12] (FLR_{KC03}) has a similar form with model FLR_{KC02}. Except in FLR_{KC03}, as shown in the following, centre c of error term $\tilde{\varepsilon}$ can be any crisp number, not only the origin:

$$\hat{\tilde{Y}}_i = a_0 + a_1\tilde{X}_{i1} + \cdots + a_j\tilde{X}_{ij} + \cdots + a_m\tilde{X}_{im} + \tilde{\varepsilon} \qquad (FLR_{KC03})$$
$$\tilde{\varepsilon} = (c, l, r)_{LR}$$

Model FLR_{KC03} is not able to avoid the deficiencies of model FLR_{KC02} either, which are described above.

3.2. *Model* FLR_{NN04}

The *spreads increasing problem* in model FLR_{FF} is caused by the fuzzy arithmetic rules. To avoid the *spreads increasing problem*, Nasrabadi and Nasrabadi [17] defined new arithmetic operations for symmetric fuzzy numbers and used these operations in fuzzy regression analysis. The arithmetic operations defined in [17] are as follows:

For any L-type fuzzy numbers $\tilde{A} = (m_a, \alpha_a)_L$, $\tilde{B} = (m_b, \alpha_b)_L$, and an algebraic operation $\times$ on $\Re$, $\otimes$ is the corresponding algebraic operation of $\times$ on L-type fuzzy numbers. $\otimes$ is defined as: $\tilde{A} \otimes \tilde{B} = (m_a \times m_b, \alpha_a \times \alpha_b)_L$.

Based on the above definition of the arithmetic operations, FLR_{FF} can be written as follows:

$$\hat{\tilde{Y}}_i = \tilde{A}_0 + \tilde{A}_1\tilde{X}_{i1} + \cdots + \tilde{A}_m\tilde{X}_{im}$$
$$= \left(m_{a_0} + \sum_{p=1}^{m} m_{a_p} m_{x_{ip}}, \alpha_{a_0} + \sum_{p=1}^{m} \alpha_{a_p}\alpha_{x_{ip}} \right)_L \qquad (FLR_{NN04})$$

To some extent, model FLR_{NN04} can avoid the *spreads increasing problem*, because in model FLR_{NN04}, the spreads of estimated dependent variables have no relationship with the magnitudes of independent variables. However, a deficiency of model FLR_{NN04} is that the spreads of the estimated dependent variables can only depend on the spreads of the independent variables, because the spreads of the observed dependent variables may also depend on the magnitudes of the independent variables, such as the example shown in Example 1.

3.3. *Models* $FLR_{D'Urso03}$ *and* $FLR_{Coppi06}$

The models proposed by D'Urso [8] and Coppi et al. [5] are able to circumvent the *spreads increasing problem* by modelling the centres of dependent variables by classical regression methods, and meanwhile modelling the spreads of the dependent variables on their estimated centres.

For multiple independent variables $X_i = (\tilde{X}_{i1}, \ldots, \tilde{X}_{ij}, \ldots, \tilde{X}_{im})$ and single dependent variable $\tilde{Y}_i = (m_{y_i}, \alpha_{y_i}, \beta_{y_i})_{LR}$, estimated response $\hat{\tilde{Y}}_i = (m_{\hat{y}_i}, \alpha_{\hat{y}_i}, \beta_{\hat{y}_i})_{LR}$ obtained by the regression model proposed in [8] is as follows, in which $\tilde{X}_{ij} = (m_{x_{ij}}, \alpha_{x_{ij}}, \beta_{x_{ij}})_{LR}$:

$$m_{y_i} = m_{\hat{y}_i} + \varepsilon_i, \quad m_{\hat{y}_i} = \mathbf{M}_{xi}\mathbf{a} + \mathbf{A}_{xi}\mathbf{r} + \mathbf{B}_{xi}\mathbf{s} \qquad (5) \qquad (FLR_{D'Urso03})$$
$$\alpha_{y_i} = \alpha_{\hat{y}_i} + \lambda_i, \quad \alpha_{\hat{y}_i} = m_{\hat{y}_i} b + d \qquad (6)$$
$$\beta_{y_i} = \beta_{\hat{y}_i} + \rho_i, \quad \beta_{\hat{y}_i} = m_{\hat{y}_i} g + h \qquad (7)$$

where ε_i, λ_i and ρ_i are residuals; $\mathbf{M}_{xi} = (1, m_{x_{i1}}, \ldots, m_{x_{ij}}, \ldots, m_{x_{im}})$, $\mathbf{A}_{xi} = (1, \alpha_{x_{i1}}, \ldots, \alpha_{x_{ij}}, \ldots, \alpha_{x_{im}})$, $\mathbf{B}_{xi} = (1, \beta_{x_{i1}}, \ldots, \beta_{x_{ij}}, \ldots, \beta_{x_{im}})$; $(m+1)$ dimension vectors $\mathbf{a}$, $\mathbf{r}$ and $\mathbf{s}$ are the regression parameters for centres m_{y_i}, $\mathbf{a} = (\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_m)^{\mathrm{T}}$, $\mathbf{r} = (\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_m)^{\mathrm{T}}$, $\mathbf{s} = (\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_m)^{\mathrm{T}}$; $i = 1, 2, \ldots, n$; and $j = 1, 2, \ldots, m$; b and d (g and h) are the regression parameters to estimate left (right) spreads α_{y_i} (β_{y_i}). Model $FLR_{D'Urso03}$ is based on three sub-models. The first one as shown in Eq. (5) estimates the centres of the dependent variables. The other two sub-models in Eqs. (6) and (7) model the left and right spreads of the dependent variables based on the estimated centres that are obtained in Eq. (5).

Model $FLR_{D'Urso03}$ is able to avoid the *spreads increasing problem*, because of Eqs. (6) and (7), in which $\alpha_{\hat{y}_i}$ and $\beta_{\hat{y}_i}$ only depend on $m_{\hat{y}_i}$ that can increase or decrease with the increase of $\alpha_{x_{ij}}$'s and $\beta_{x_{ij}}$'s. For instance, assume that regression parameter $\mathbf{r}_k$ is negative and b is positive. Then, $m_{\hat{y}_i}$ decreases with the increase of $\alpha_{x_{ik}}$ (the left spread of the kth independent variable), and $\alpha_{\hat{y}_i}$ increase with the increase of $m_{\hat{y}_i}$. This makes $\alpha_{\hat{y}_i}$ decrease with the increase of $\alpha_{x_{ik}}$.

However, in model $\text{FLR}_{D\text{'}Urso03}$, $\alpha_{\hat{y}_i}$ ($\beta_{\hat{y}_i}$) is determined by $m_{\hat{y}_i}$. That limits the ability of $\text{FLR}_{D\text{'}Urso03}$ to model α_{y_i} (β_{y_i}) by independent variable X_i. For single independent variable, the three sub-models of $\text{FLR}_{D\text{'}Urso03}$ can be rewritten as follows:

$$\begin{aligned} m_{\hat{y}_i} &= (1, m_{x_i}) * (\mathbf{a}_0, \mathbf{a}_1)^{\mathrm{T}} + (1, \alpha_{x_i}) * (\mathbf{r}_0, \mathbf{r}_1)^{\mathrm{T}} + (1, \beta_{x_i}) * (\mathbf{s}_0, \mathbf{s}_1)^{\mathrm{T}} \\ &= \mathbf{a}_1 m_{x_i} + \mathbf{r}_1 \alpha_{x_i} + \mathbf{s}_1 \beta_{x_i} + k \end{aligned} \tag{8}$$

$$\begin{aligned} \alpha_{\hat{y}_i} &= ((1, m_{x_i}) * (\mathbf{a}_0, \mathbf{a}_1)^{\mathrm{T}} + (1, \alpha_{x_i}) * (\mathbf{r}_0, \mathbf{r}_1)^{\mathrm{T}} + (1, \beta_{x_i}) * (\mathbf{s}_0, \mathbf{s}_1)^{\mathrm{T}})b + d \\ &= \mathbf{a}_1 b m_{x_i} + \mathbf{r}_1 b \alpha_{x_i} + \mathbf{s}_1 b \beta_{x_i} + kb + d \end{aligned} \tag{9}$$

$$\begin{aligned} \beta_{\hat{y}_i} &= ((1, m_{x_i}) * (\mathbf{a}_0, \mathbf{a}_1)^{\mathrm{T}} + (1, \alpha_{x_i}) * (\mathbf{r}_0, \mathbf{r}_1)^{\mathrm{T}} + (1, \beta_{x_i}) * (\mathbf{s}_0, \mathbf{s}_1)^{\mathrm{T}})g + h \\ &= \mathbf{a}_1 b m_{x_i} + \mathbf{r}_1 b \alpha_{x_i} + \mathbf{s}_1 b \beta_{x_i} + kg + h \end{aligned} \tag{10}$$

where $k = \mathbf{a}_0 + \mathbf{r}_0 + \mathbf{s}_0$. Assume that in a single independent variable dataset, m_{y_i}'s (the centres of the dependent variables) increase with the increase of α_{x_i}'s and β_{x_i}'s (the spreads of the independent variables); and α_{y_i}'s (the spreads of the dependent variables) increase with the increase of α_{x_i}'s, but decrease when β_{x_i}'s increase. According to Eq. (8), $m_{\hat{y}_i}$ is able to describe the relationship between $m_{\hat{y}_i}$ and $\tilde{X}_i$ (i.e., $m_{\hat{y}_i}$ increases when α_{x_i} and β_{x_i} increase. This requires $\mathbf{r}_1$ and $\mathbf{s}_1$ to be positive.). However, $\alpha_{\hat{y}_i}$ is not able to describe the relationship between α_{y_i} and $\tilde{X}_i$ properly because when α_{y_i} increases with the increase of α_{x_i}, b needs to be positive, and meanwhile α_{y_i} decreases with the increase of β_{x_i} that requires b to be negative. Thus, in this case, model $\text{FLR}_{D\text{'}Urso03}$ is not able to properly describe the relationship between $\tilde{X}_i$ and $\tilde{Y}_i$.

Similarly, the model proposed by Coppi et al. [5] is also composed of three sub-models. For crisp inputs $x_i = (x_{i1}, \ldots, x_{im})$ and *LR*-type fuzzy outputs $\tilde{Y}_i = (m_{y_i}, \alpha_{y_i}, \beta_{y_i})_{LR}$, estimated responses $\hat{\tilde{Y}}_i = (m_{\hat{y}_i}, \alpha_{\hat{y}_i}, \beta_{\hat{y}_i})_{LR}$ obtained by the regression model proposed in [5] is

$$\begin{aligned} m_{y_i} &= m_{\hat{y}_i} + \varepsilon_i, \quad m_{\hat{y}_i} = \mathbf{F}(x_i)\mathbf{a} && (11) \\ \alpha_{y_i} &= \alpha_{\hat{y}_i} + \lambda_i, \quad \alpha_{\hat{y}_i} = m_{\hat{y}_i} b + d && (12) \\ \beta_{y_i} &= \beta_{\hat{y}_i} + \rho_i, \quad \beta_{\hat{y}_i} = m_{\hat{y}_i} g + h && (13) \end{aligned} \qquad (\text{FLR}_{Coppi06})$$

where $\mathbf{F}(x_i) = [f_1(x_i), \ldots, f_k\, x_i), \ldots, f_p)(x_i)]$, f_k's are suitably chosen functions. For crisp input–fuzzy output, model $\text{FLR}_{D\text{'}Urso03}$ is a specification of model $\text{FLR}_{Coppi06}$, in which $\mathbf{F}(x_i) = [1, x_{i1}, \ldots, x_{im}]$. Similar to model $\text{FLR}_{D\text{'}Urso03}$, the sub-models of $\text{FLR}_{Coppi06}$ shown in Eqs. (12) and (13) also depend on the sub-model given in Eq. (11). Thus, $\text{FLR}_{Coppi06}$ has the same problem as $\text{FLR}_{D\text{'}Urso03}$, which is that α_{y_i} (β_{y_i}) cannot be linearly modelled by X_i freely.

3.4. *Model* FLR_{CD08}

To address the *spreads increasing problem*, a variable spread FLR model FLR_{CD08} is proposed by Chen and Dang in [2], which is a three-phase method.

In the first phase, regression coefficients are treated as fuzzy numbers and the membership functions of the LS estimates of the regression coefficients are constructed, since Chen and Dang argue that the membership functions of fuzzy sets are more capable of capturing the relationship between independent variables and dependent variables than crisp numbers [2]. To avoid the *spreads increasing problem*, in the second phase, the fuzzy regression coefficients are defuzzified by the centre of gravity method to crisp regression coefficients. In the third phase, for each instance, fuzzy error term $\tilde{E}_i$ is determined by a mathematical programming method. The objective function of the mathematical programming method is to minimize the total difference between the estimated and observed membership values of response variables, E_{KC} (refer Section 4.2), subject to the constraints that the spreads of each estimated response variable are equal to those of the observed response variable. For predicting the response of an unseen instance, a Mamdani fuzzy inference system [29] is applied to the derived regression model.

A generic model of FLR_{CD08} is

$$\begin{aligned} \hat{\tilde{Y}}_i &= (b_0)_c + (b_1)_c \tilde{X}_{i1} + \cdots + (b_j)_c \tilde{X}_{ij} + \cdots + (b_m)_c \tilde{X}_{im} + \tilde{E}_i && (14) \\ \tilde{E}_i &= (0, \alpha_i, \beta_i)_{LR} && (15) \end{aligned} \qquad (\text{FLR}_{CD08})$$

where $(b_j)_c$'s are the defuzzified crisp regression coefficients; $\tilde{E}_i$ is the estimated error term of the ith instance; $i = 1, 2, \ldots, n$; $j = 0, 1, \ldots, m$.

The parameters to be estimated in FLR_{CD08} are regression coefficients, $(b_j)_c$'s, and left and right spreads of error terms, α_i's and β_i's, respectively. Totally, there are $m + 2n + 1$ parameters in FLR_{CD08}, which are proportional to the number of instances, n, and the dimension of the dataset, m. For large datasets, n is usually significantly greater than m. More parameters involved in a regression model increase the model fitness, but these also decrease the model generality [13]. Thus, FLR_{CD08} is unsuitable to large dataset regression, whose number of instances is large.

4. Flexible spreads FLR model FLR_{FS}

From the above sections, we have seen the shortcomings of the previous FLR models. In this section, we describe our flexible spreads FLR model (FLR_{FS}) that is able to overcome the problems mentioned above of the previous FLR models.

4.1. Description of model FLR_{FS}

We first describe model FLR_{FS} for single regression, then extend it to multiple regression.

In fuzzy regression analysis, the spreads and magnitudes of independent variables are all the information that can be obtained. A general case is that the spreads of dependent variables may depend on both the spreads and magnitudes of the independent variables. Thus, a general FLR model should be able to allow: (i) the spreads of estimated dependent variables depend on both the spreads and magnitudes of the independent variables; (ii) the spreads of the estimated dependent variables can change freely (increase, decrease or fixed) as the spreads and magnitudes of the independent variables change. FLR_{FS} is a model that possesses these properties.

Also, considering the relationship between fuzzy numbers and crisp numbers, in FLR_{FS} the centres of estimated dependent variables are modelled by the centres of independent variables, and the spreads of the estimated dependent variables are modelled by both the centres and spreads of the independent variables. Since the centre of a fuzzy number is the element belonging to a fuzzy concept with 100%, it can be treated as a crisp number. Thus, in FLR_{FS} the estimation of the centres of dependent variables is based on classical linear regression. The spreads of fuzzy numbers can be treated as the vagueness of the fuzzy numbers. The vagueness of dependent variables depends not only on the vagueness of independent variables but also on the centres of the independent variables, such as the dataset1 given in Table 1. To capture this relationship between the spreads of the dependent variables and the independent variables, in FLR_{FS} the spreads of the dependent variables are estimated by both the centres and spreads of the independent variables.

Model FLR_{FS} for single independent variable $\tilde{X}_i = (m_{x_i}, \alpha_{x_i}, \beta_{x_i})_{LR}$, can be described as follows:

$$
\begin{aligned}
&\hat{\tilde{Y}}_i = k_0 + k_1 m_{x_i} + \tilde{S}_i \qquad\qquad (\mathrm{FLR}_{FS}\ \text{single})\\
&\tilde{S}_i = (0, \alpha_{s_i}, \beta_{s_i})_{LR}\\
&\alpha_{s_i} = k_{ll}\alpha_{x_i} + k_{lm}m_{x_i} + k_{lr}\beta_{x_i} + c_l\\
&\beta_{s_i} = k_{rl}\alpha_{x_i} + k_{rm}m_{x_i} + k_{rr}\beta_{x_i} + c_r\\
&\alpha_{s_i} \geqslant 0, \quad \beta_{s_i} \geqslant 0, \ i = 1, 2, \ldots, n
\end{aligned}
$$

where k_0 and k_1 are crisp regression coefficients; $\tilde{S}_i$ is a fuzzy spread term for ith instance; $k_{ll}, k_{lm}, k_{lr}, k_{rl}, k_{rm}, k_{rr}$, c_l and c_r are crisp spread coefficients. To achieve the fuzzy regression model, the parameters (i.e., $k_0, k_1, k_{ll}, k_{lm}, k_{lr}$, k_{rl}, k_{rm}, k_{rr}, c_l and c_r) need to be determined, subject to the constraints that the spread of $\tilde{S}_i$ should be non-negative. Parameters k_{ll} and k_{rl} reflect the influence of the left spreads of the independent variables on the left and right spreads of the dependent variables, respectively. Similarly, parameters k_{lr} and k_{rr} show how the right spreads of the independent variables affect the left and right spreads of the dependent variables. Parameters k_{lm} and k_{rm} give the information of how the spreads of the dependent variables depend on the centres of the independent variables.

All parameters, $k_0, k_1, k_{ll}, k_{lm}, k_{lr}, k_{rl}, k_{rm}, k_{rr}, c_l$ and c_r, can be positive or negative. Thus, the spreads of the estimated dependent variables can increase or decrease freely as the spreads and magnitudes of the independent variables change. Thus, the model FLR_{FS} is able to avoid the *spreads increasing problem*.

For L-type independent variable $\tilde{X}_i = (m_{x_i}, \alpha_{x_i})_L$ and dependent variable $\tilde{Y}_i = (m_{y_i}, \alpha_{y_i})_L$, a simplified FLR$_{FS}$ model can be expressed as

$$\hat{\tilde{Y}}_i = k_0 + k_1 m_{x_i} + \tilde{S}_i$$

$$\tilde{S}_i = (0, \alpha_{s_i})_L, \quad \alpha_{s_i} = k_{ll}\alpha_{x_i} + k_{lm} m_{x_i} + c$$

$$\alpha_{s_i} \geqslant 0, \quad i = 1, 2, \ldots, n$$

A generalized model FLR$_{FS}$ for multiple independent variables $X_i = (\tilde{X}_{i1}, \ldots, \tilde{X}_{ij}, \ldots, \tilde{X}_{im})$, can be described as follows:

$$\begin{aligned}
&\hat{\tilde{Y}}_i = k_0 + k_1 m_{x_{i1}} + \cdots + k_j m_{x_{ij}} + \cdots + k_m m_{x_{im}} + \tilde{S}_i \qquad (\text{FLR}_{FS}\ \text{multiple})\\
&\tilde{S}_i = (0, \alpha_{s_i}, \beta_{s_i})_{LR}\\
&\alpha_{s_i} = \sum_{t=1}^{m} k_{llt}\alpha_{x_{it}} + \sum_{t=1}^{m} k_{lmt} m_{x_{it}} + \sum_{t=1}^{m} k_{lrt}\beta_{x_{it}} + c_l\\
&\beta_{s_i} = \sum_{t=1}^{m} k_{rlt}\alpha_{x_{it}} + \sum_{t=1}^{m} k_{rmt} m_{x_{it}} + \sum_{t=1}^{m} k_{rrt}\beta_{x_{it}} + c_r\\
&\alpha_{s_i} \geqslant 0, \quad \beta_{s_i} \geqslant 0, \quad i = 1, 2, \ldots, n, \; j = 0, 1, \ldots, m
\end{aligned}$$

where $\tilde{X}_{ij} = (m_{x_{ij}}, \ \alpha_{x_{ij}}, \ \beta_{x_{ij}})_{LR}$, k_j's are crisp regression coefficients; $\tilde{S}_i$ is a fuzzy spread term for ith instance; k_{llt}'s, k_{lmt}'s, k_{lrt}'s, k_{rlt}'s, k_{rmt}'s, k_{rrt}'s, c_l and c_r are crisp spread coefficients. Model FLR$_{FS}$ for multiple regression is a generalization of model FLR$_{FS}$ for single regression.

The parameters to be estimated in model FLR$_{FS}$ are k_j's, k_{llt}'s, k_{lmt}'s, k_{lrt}'s, k_{rlt}'s, k_{rmt}'s, k_{rrt}'s, c_l and c_r. The number of the parameters is $7m$+3, which is only proportional to the dimension of the dataset. Considering that more parameters in a model decrease the model generality, FLR$_{FS}$ is thus more suitable to low-dimensional dataset regression.

As we can see from above, model FLR$_{FS}$ can be easily extended from single regression to multiple regression. However, the computational complexity of model FLR$_{FS}$ will increase significantly with the increase of data size and the dimension of the independent variables. A gradient-descent optimization strategy proposed in [1] deals with the high-dimensional data linear regression.

Model FLR$_{FS}$ can be used for descriptive purposes to study the fuzzy relationship between dependent and independent variables. Also, it can be used for prediction purposes. Based on the arguments of D'Urso [8] and Coppi et al. [5], if non-positive predicted spreads are interpreted as a lack of uncertainty and can be set to 0 for practical purposes, model FLR$_{FS}$ can then be used for prediction purposes.

Although models FLR$_{FS}$ and FLR$_{KC02}$ have a similar form, they have several significant differences. (i) In model FLR$_{FS}$, fuzzy spread variable $\tilde{S}_i$ is different for each instance, which is determined by the centres and spreads of independent variables $\tilde{X}_i$. In model FLR$_{KC02}$, $\tilde{\varepsilon}$ is an error term, which is fixed for all instances. (ii) In model FLR$_{FS}$, regression coefficients, k_1 and k_2, describe the relationship between m_{y_i} (the centres of the dependent variables) and m_{x_i} (the centres of the independent variables). In model FLR$_{KC02}$, the regression coefficients are obtained from modelling the relationship between the centroids of $\tilde{Y}_i$ and $\tilde{X}_i$; but they are used to describe the relationship between $\tilde{Y}_i$ and $\tilde{X}_i$.

The similarity between models FLR$_{D'Urso03}$, FLR$_{Coppi06}$ and FLR$_{FS}$ is that all of them model the centres and spreads of dependent variables separately. There are also two differences between models FLR$_{D'Urso03}$, FLR$_{Coppi06}$ and FLR$_{FS}$. (i) In FLR$_{D'Urso03}$ and FLR$_{Coppi06}$, the centres of dependent variables are modelled by both the centres and spreads of independent variables. As mentioned above, the centre of a fuzzy number can be treated as a crisp number. Therefore, in FLR$_{FS}$ the centres of dependent variables are determined by classical linear regression. (ii) In FLR$_{D'Urso03}$ and FLR$_{Coppi06}$, the spreads of dependent variables are modelled by their corresponding estimated centres. In FLR$_{FS}$, the spreads of estimated responses variable are able to depend on both the centres and spreads of independent variables linearly.

4.2. Property of model FLR_{FS}

Property of FLR_{FS}**.** The feasible region of model FLR_{FS} contains the feasible regions of models FLR_{KC02}, FLR_{KC03} and FLR_{NN04}. Thus, models FLR_{KC02}, FLR_{KC03} and FLR_{NN04} can be seen as special cases of model FLR_{FS}.

Proof. For simplicity, we only proof the above property of FLR_{FS} for single regression. For multiple regression, the property can be proved in a similar way.

Model FLR_{FS} for single regression can be rewritten as

$$\hat{\tilde{Y}}_i = (k_0 + k_1 m_{x_i}, k_{ll}\alpha_{x_i} + k_{lm} m_{x_i} + k_{lr}\beta_{x_i} + c_l, k_{rl}\alpha_{x_i} + k_{rm} m_{x_i} + k_{rr}\beta_{x_i} + c_r)_{LR} \tag{16}$$

(a) For single regression, both model FLR_{KC02} and model FLR_{KC03} can be written as

$$\hat{\tilde{Y}}_i = a_0 + a_1\tilde{X}_i + \tilde{\varepsilon} = a_0 + a_1(m_{x_i}, \alpha_{x_i}, \beta_{x_i})_{LR} + (m_\varepsilon, \alpha_\varepsilon, \beta_\varepsilon)_{LR}$$

$$= \begin{cases} (a_0 + a_1 m_{x_i} + m_\varepsilon, a_1\alpha_{x_i} + \alpha_\varepsilon, a_1\beta_{x_i} + \beta_\varepsilon)_{LR} & \text{if } a_1 \geqslant 0 \quad (17) \\ (a_0 + a_1 m_{x_i} + m_\varepsilon, a_1\beta_{x_i} + \alpha_\varepsilon, a_1\alpha_{x_i} + \beta_\varepsilon)_{LR} & \text{if } a_1 < 0 \quad (18) \end{cases}$$

The feasible region of model FLR_{FS} contains the feasible regions of models FLR_{KC02}, FLR_{KC03}, since for any solution of model FLR_{KC02} or FLR_{KC03} $(a_0, a_1, m_\varepsilon, \alpha_\varepsilon, \beta_\varepsilon)$, we can always find an equivalent solution in model FLR_{FS} by comparing the coefficients of Eq. (16) with those of Eq. (17), and the coefficients of Eq. (16) and those of Eq. (18):

$$\begin{aligned} &\text{for } a_1 \geqslant 0: \\ &\quad k_0 = a_0 + m_\varepsilon, \quad k_1 = k_{ll} = k_{rr} = a_1 \\ &\quad k_{lm} = k_{lr} = k_{rm} = k_{rl} = 0, \quad c_l = \alpha_\varepsilon, \quad c_r = \beta_\varepsilon \end{aligned} \tag{19}$$

$$\begin{aligned} &\text{for } a_1 < 0: \\ &\quad k_0 = a_0 + m_\varepsilon, \quad k_1 = k_{lr} = k_{rl} = a_1 \\ &\quad k_{lm} = k_{ll} = k_{rm} = k_{rr} = 0, \quad c_l = \alpha_\varepsilon, \quad c_r = \beta_\varepsilon \end{aligned} \tag{20}$$

(b) In [17], model FLR_{NN04} is only considered with L-type fuzzy numbers. For a single independent variable, $\tilde{X}_i = (m_{x_i}, \alpha_{x_i})_L$, model FLR_{NN04} can be expressed as

$$\begin{aligned} \hat{\tilde{Y}}_i &= \tilde{A}_0 + \tilde{A}_1\tilde{X}_i = (m_{a_0}, \alpha_{a_0})_L + (m_{a_1}, \alpha_{a_1})_L(m_{x_i}, \alpha_{x_i})_L \\ &= (m_{a_0} + m_{a_1} m_{x_i}, \alpha_{a_0} + \alpha_{a_1}\alpha_{x_i})_L \end{aligned} \tag{21}$$

When all fuzzy numbers are L-type, model FLR_{FS} can be simplified as

$$\hat{\tilde{Y}}_i = (k_0 + k_1 m_{x_i}, k_{ll}\alpha_{x_i} + k_{lm} m_{x_i} + c)_L \tag{22}$$

The feasible region of model FLR_{FS} contains the feasible region of model FLR_{NN04}, since for any solution of model FLR_{NN04} $(m_{a_0}, \alpha_{a_0}, m_{a_1}, \alpha_{a_1})$, we can always find an equivalent solution in model FLR_{FS} by comparing the coefficients of Eq. (21) and those of Eq. (22):

$$k_0 = m_{a_0}, \quad k_1 = m_{a_1}, \quad k_{ll} = \alpha_{a_1}, \quad k_{lm} = 0, \quad c = \alpha_{a_0} \tag{23}$$

However, not all solutions of model FLR_{FS} can have an equivalent solution in models FLR_{KC02}, FLR_{KC03} or FLR_{NN04}. For instance, no solutions of models FLR_{KC02}, FLR_{KC03} or FLR_{NN04} are equivalent to the solutions of model FLR_{FS}, when the spreads of the dependent variables are determined or partially determined by the centres of the independent variables.

Thus, for single regression, the feasible region of model FLR_{FS} contains the feasible region of models FLR_{KC02}, FLR_{KC03} and FLR_{NN04}. Also, the conversion formulas from the parameters of solutions of models FLR_{KC02}, FLR_{KC03} and FLR_{NN04} to the parameters of solutions of model FLR_{FS} are given in Eqs. (19), (20) and (23). For single regression, models FLR_{KC02}, FLR_{KC03} and FLR_{NN04} can be seen as the special cases of model FLR_{FS}. □

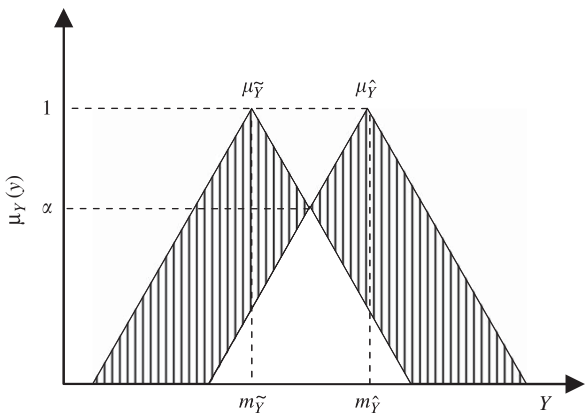

Fig. 1.

4.3. Parameters estimation

When model FLR_{FS} is adopted, the following task is to estimate the parameters. Minimizing the total difference between estimated and observed response variables is a common criterion of parameters estimation. Various distance measurements have been proposed to measure the total difference between the estimated and observed response variables in fuzzy regression.

In [15], the error of estimation, E_{KB}, is defined as the ratio of the total difference between the estimated and observed membership values of response variables to the total observed membership values of the response variables, which is the shaded areas over the left triangle area, in Fig. 1. A formularized definition of E_{KB} is given by

$$E_{KB} = \frac{\int_{S_{\tilde{Y}} \cup S_{\hat{\tilde{Y}}}} |\mu_{\tilde{Y}}(x) - \mu_{\hat{\tilde{Y}}}(x)|\,dx}{\int_{S_{\tilde{Y}}} \mu_{\tilde{Y}}(x)\,dx} \tag{24}$$

where $\mu_{\tilde{Y}}(x)$ and $\mu_{\hat{\tilde{Y}}}(x)$ are the estimated and observed membership functions of the response variables, $S_{\tilde{Y}}$ and $S_{\hat{\tilde{Y}}}$ are the supports of $\mu_{\tilde{Y}}(x)$ and $\mu_{\hat{\tilde{Y}}}(x)$.

E_{KC}, a variation of E_{KB}, was used in [11,12]. E_{KC} measures the total difference between the estimated and observed membership values of response variables, which include all the shaded areas in Fig. 1:

$$E_{KC} = \int_{S_{\tilde{Y}} \cup S_{\hat{\tilde{Y}}}} |\mu_{\tilde{Y}}(x) - \mu_{\hat{\tilde{Y}}}(x)|\,dx \tag{25}$$

In [9], the similarity of fuzzy numbers is used as a measurement to evaluate the effectiveness of regression, which is defined as follows:

$$S_H = \frac{\int \min(\mu_{\tilde{Y}}(x), \mu_{\hat{\tilde{Y}}}(x))\,dx}{\int \max(\mu_{\tilde{Y}}(x), \mu_{\hat{\tilde{Y}}}(x))\,dx} \tag{26}$$

In [8], the squared Euclidean distance between two fuzzy numbers $\tilde{A}_1 = (m_1, \alpha_1, \beta_1)_{LR}$ and $\tilde{A}_2 = (m_2, \alpha_2, \beta_2)_{LR}$ is defined as

$$d^2(\tilde{A}_1, \tilde{A}_2) = \|m_1 - m_2\|^2 \pi_c + \|(m_1 - \alpha_1) - (m_2 - \alpha_2)\|^2 \pi_\alpha + \|(m_1 + \beta_1) - (m_2 + \beta_2)\|^2 \pi_\beta$$

where π_c, π_α and π_β are arbitrary positive weights.

In [5], a generalized squared Euclidean distance is used, which can be described as

$$\Delta^2(\tilde{A}_1, \tilde{A}_2) = \|m_1 - m_2\|^2 + \|(m_1 - \lambda\alpha_1) - (m_2 - \lambda\alpha_2)\|^2 + \|(m_1 + \rho\beta_1) - (m_2 + \rho\beta_2)\|^2$$

where $\lambda = \int_0^1 L^{-1}(\omega)\,d\omega$, $\rho = \int_0^1 R^{-1}(\omega)\,d\omega$. From the definition of Δ^2, we can see that Δ^2 weights the centres and spreads differently by means of λ and ρ.

E_{KB}, E_{KC}, d^2 and Δ^2 range from zero to positive infinity, while S_H ranges from 0 to 1. Thus, compared with E_{KB}, E_{KC}, d^2 and Δ^2, S_H can better describe the total difference between the estimated and observed response variables. Therefore, to estimate the parameters of model FLR$_{FS}$, our objective function is set to maximize the average similarity between the estimated and observed response variables. This is referred to as *MaxSim* solution for FLR$_{FS}$, which can be described as

$$\begin{aligned}
\text{Max} \quad & \frac{1}{n}\sum_{i=1}^{n} S_{H_i} \\
\text{s.t.} \quad & S_{H_i} = \frac{\int \min(\mu_{\tilde{Y}_i}(x), \mu_{\hat{\tilde{Y}}_i}(x))\,dx}{\int \max(\mu_{\tilde{Y}_i}(x), \mu_{\hat{\tilde{Y}}_i}(x))\,dx} \\
& \hat{\tilde{Y}}_i = k_0 + k_1 m_{x_{i1}} + \cdots + k_j m_{x_{ij}} + \cdots + k_m m_{x_{im}} + \tilde{S}_i \\
& \tilde{S}_i = (0, \alpha_{s_i}, \beta_{s_i})_{LR} \\
& \alpha_{s_i} = \sum_{t=1}^{m} k_{llt}\alpha_{x_{it}} + \sum_{t=1}^{m} k_{lmt} m_{x_{it}} + \sum_{t=1}^{m} k_{lrt}\beta_{x_{it}} + c_l \\
& \beta_{s_i} = \sum_{t=1}^{m} k_{rlt}\alpha_{x_{it}} + \sum_{t=1}^{m} k_{rmt} m_{x_{it}} + \sum_{t=1}^{m} k_{rrt}\beta_{x_{it}} + c_r \\
& \alpha_{s_i} \geqslant 0, \quad \beta_{s_i} \geqslant 0, \quad i = 1, 2, \ldots, n, \quad j = 0, 1, \ldots, m
\end{aligned} \tag{27}$$

5. Numerical examples

Note that the solution of the optimization problem *MaxSim* for model FLR$_{FS}$ depends on the initial values, since the feasible region of solutions may not be continuous. In this section, we first provide a strategy of setting initial values, which is adopted in our experiments. Then, the effectiveness of model FLR$_{FS}$ will be demonstrated on four datasets: dataset1 that has been shown in Table 1 in Section 1, and other three commonly used datasets (one is a single crisp input–fuzzy output dataset (dataset2) from [24]; another is a single fuzzy input–fuzzy output dataset (dataset3) from [22], and the other is a multiple fuzzy inputs–fuzzy output real world dataset (dataset4) from [8]).

5.1. Initial value setting

For practical reasons, in this section we provide an initial value setting strategy. The experimental results in Section 5.2 are based on this strategy. The purpose to provide this initial value setting strategy here is neither to demonstrate it is the best strategy to set initial values nor to claim that it guarantees to achieve the global optimization.

For simplicity, in this section we only introduce the initial value setting strategy for single regression FLR$_{FS}$. For model FLR$_{FS}$ dealing with multiple regression, the initial values can be set in a similar way.

Given observations $(\tilde{X}_i, \tilde{Y}_i)$, where $\tilde{X}_i = (m_{x_i}, \alpha_{x_i}, \beta_{x_i})_{LR}$, $\tilde{Y}_i = (m_{y_i}, \alpha_{y_i}, \beta_{y_i})_{LR}$, $i = 1, 2, \ldots, n$, the task of fuzzy regression is to find the parameters of model FLR$_{FS}$, which maximizes Eq. (27), subject to its constraints.

In model FLR$_{FS}$, line $k_0 + k_1 m_{x_i}$ describes the relationship between m_{x_i} (the centre of independent variable) and m_{y_i} (the centre of dependent variable). Thus, we apply the conventional LS estimation to get the linear relationship between m_{x_i} and m_{y_i}, i.e., $m_{y_i} = b_0 + b_1 m_{x_i}$. b_0 and b_1 can be set as the initial values of k_0 and k_1.

In model FLR$_{FS}$, there are three factors that can affect the spreads of response variables: α_{x_i}, β_{x_i} and m_{x_i}. k_{ll} describes how α_{y_i} (the left spread of dependent variable) depends on α_{x_i} (the left spread of independent variable). Then, the conventional LS estimation is applied to α_{y_i} and α_{x_i} to get their linear relationship: $\alpha_{y_i} = p_{ll} + b_{ll}\alpha_{x_i}$. If equal weights is set to the three factors, α_{x_i}, β_{x_i} and m_{x_i}, $b_{ll}/3$ can be set as the initial value of k_{ll}. Similarly, k_{lr} describes how α_{y_i} depends on β_{x_i}. The LS estimation is applied to α_{y_i} and β_{x_i} to get their relationship: $\alpha_{y_i} = p_{lr} + b_{lr}\beta_{x_i}$. $b_{lr}/3$ can be set as the initial value of k_{lr}. Parameter k_{lm} describes how α_{y_i} depends on m_{x_i}. By applying the LS

Table 2
Dataset2.

i	x_i	$\tilde{Y}_i = (m_{y_i}, \alpha_{y_i})_L$	$\hat{\tilde{Y}}_i = (m_{\hat{y}_i}, \alpha_{\hat{y}_i})_L$	S_H	E_{KC}
1	1	$(8.0, 1.8)_L$	$(6.00, 2.80)_L$	0.19	3.13
2	2	$(6.4, 2.2)_L$	$(7.75, 2.70)_L$	0.36	2.33
3	3	$(9.5, 2.6)_L$	$(9.50, 2.60)_L$	1.00	0.00
4	4	$(13.5, 2.6)_L$	$(11.25, 2.50)_L$	0.19	3.51
5	5	$(13.0, 2.4)_L$	$(13.00, 2.40)_L$	1.00	0.00
Average				0.5462	1.7932

estimation to α_{y_i} and m_{x_i}, their relationship $\alpha_{y_i} = p_{lm} + b_{lm} m_{x_i}$ is obtained. The initial value of k_{lm} can be set as $b_{lm}/3$. $(p_{ll}/3 + p_{lr}/3 + p_{lm}/3)$ can be set as the initial value of c_l. Unequal weights can also be assigned to the three factors α_{x_i}, β_{x_i} and m_{x_i}. For example, if the effect of m_{x_i} is significantly greater than the effects of both α_{x_i} and β_{x_i} (i.e., $b_{lm} \gg b_{ll}$ and b_{lr}), the initial spreads of the estimated response variables can be set only based on m_{x_i}.

Similarly, we can set the initial values of k_{rl}, k_{rm}, k_{rr} and c_r in the same way.

5.2. Examples

The following experimental results are based on the initial value setting strategy described in Section 5.1.

Example 1 (*Continue*). As it is shown in Section 1, the dataset1 cannot be modelled by FLR_{FF}, FLR_{KC02}, FLR_{KC03} and FLR_{NN04} properly, because the spreads of the observed response variables tend to decrease as the magnitudes of the independent variable increase. This decreasing trend in the spreads of the observed response variables can be fitted by model FLR_{FS}. Applying L-type fuzzy numbers based FLR_{FS} model and the initial value setting strategy given in Section 5.1, the following regression model is obtained:

$$\hat{\tilde{Y}}_i = -1.10 + m_{x_i} + (0, -0.5m_{x_i} + 1.05)_L \tag{28}$$

It can be seen from Eq. (28) that the spreads of the estimated response variables are decreasing with the increase of m_{xi}'s. The estimated response variables for the five instances are $(0.60, 0.20)_L$, $(0.70, 0.15)_L$, $(0.80, 0.10)_L$, $(0.90, 0.05)_L$ and $(1.00, 0.00)_L$, respectively.

Note that when m_x is greater than 2.10, the centre of the estimated dependent variable is greater than 1.00 and the spreads of the estimated dependent variable are less than 0, which can be interpreted as the height belongs to the concept *high* with full confidence and a lack of uncertainty. For practical reasons, the estimated centres that are greater than 1.00 can be set to 1.00, and the estimated spreads that are less than 0 can be set to 0.

Example 2. In this example, we consider the crisp input–fuzzy output dataset given by Tanaka et al. [24], which is shown in the left half of Table 2. x_i is the observed independent variable, $\tilde{Y}_i$ is the observed dependent variable, $i = 1, 2, \ldots, 5$.

As the observed dependent variables in dataset2 are symmetric, L-type fuzzy numbers based FLR_{FS} model is adopted. According to the initial value setting strategy given in Section 5.1, the initial values are set as

$$\text{Initial values (a)}: \; k_0 = 4.95; \quad k_1 = 1.71; \quad k_{ll} = 0; \quad k_{lm} = 0.08; \quad c = 2.08$$

Then, the following regression model is obtained:

$$\hat{\tilde{Y}}_i = (4.25 + 1.75m_{x_i}, -0.10m_{x_i} + 2.90)_L \tag{29}$$

Estimated response variable $\hat{\tilde{Y}}_i$, S_H and E_{KC} for each instance of dataset2 are listed in the right half of Table 2.

Since model FLR_{FS} is proposed to solve the *spreads increasing problem*, in this example model FLR_{FS} is compared with models FLR_{KC02} [11], FLR_{KC03} [12], FLR_{NN04} [17] and FLR_{CD08} [2], which are able to avoid the *spreads*

Table 3
Fuzzy regression models of dataset2.

$FLR_{FS}(MaxSim)$ (given initial values (a))	$\hat{\tilde{Y}}_i = (4.25 + 1.75m_{x_i}, -0.10m_{x_i} + 2.90)_L$
FLR_{KC02} [11]	$\hat{\tilde{Y}}_i = (4.95 + 1.71m_{x_i}, 3.01, 1.80)_{LR}$
FLR_{KC03} [12]	$\hat{\tilde{Y}}_i = (4.926 + 1.718m_{x_i}, 2.320)_L$
FLR_{NN04} [17]	$\hat{\tilde{Y}}_i = (4.6812 + 1.73306m_{x_i}, 2.3221)_L$
FLR_{CD08} [2]	$\hat{\tilde{Y}}_i = 4.95 + 1.71m_{x_i} + \tilde{E}_i$

Table 4
Comparison of the performance of difference methods on Dataset2.

Models	$AveS_H$	$TotE_{KC}$
$FLR_{FS}(MaxSim)$ (given the initial values (a))	0.5462	8.9659
FLR_{KC02} [11]	0.4663	9.679
FLR_{KC03} [12]	0.4095	10.089
FLR_{NN04} [17]	0.4388	9.771
FLR_{CD08} [2]	0.5198	7.857

increasing problem and also are provided experimental results on dataset2 by their authors. These five models are listed in Table 3,[2] where $\tilde{E}_i$'s in model FLR_{CD08}, are $\tilde{E}_1 = (-1.8, 0, 1.8)$, $\tilde{E}_2 = (-2.6, 0, 1.8)$, $\tilde{E}_3 = (-3.4, 0, 1.8)$, $\tilde{E}_4 = (-1.8, 0, 3.4)$ and $\tilde{E}_5 = (-3, 0, 1.8)$.

It is worth to note that the solution of a model also depends on the objective function to be optimized. Thus, it is not easy to compare the solutions of different models that optimize different objective functions.

However, if the evaluation results of model M_A are better than these of model M_B when both objective functions of M_A and M_B are used as the evaluation measurements, we can then say M_A outperforms M_B in terms of these two evaluation measurements.

In this experiment, in order to compare model FLR_{FS} with models FLR_{KC02}, FLR_{KC03} and FLR_{NN04}, both average similarity $AveS_H$ and total error $TotE_{KC}$ are used as the evaluation measurements, which are defined as follows:

$$AveS_H = \frac{1}{n}\sum_{i=1}^{n} S_{Hi}, \quad TotE_{KC} = \sum_{i=1}^{n} E_{KC_i}$$

because: (i) the objective function of model FLR_{FS} is to maximize $AveS_H$ between observed response variables $\tilde{Y}$ and their estimated counterparts $\hat{\tilde{Y}}$ and (ii) the objective function of models FLR_{KC02}, FLR_{KC03}, FLR_{NN04} and FLR_{CD08} is to minimize $TotE_{KC}$ between $\tilde{Y}$ and $\hat{\tilde{Y}}$.

According to the definitions of $AveS_H$ and $TotE_{KC}$, a better method in terms of these two measurements should have a higher $AveS_H$ value and a lower $TotE_{KC}$ value.

Table 4 shows the performances of model FLR_{FS}, and models FLR_{KC02}, FLR_{KC03}, FLR_{NN04} and FLR_{CD08} on $AveS_H$ and $TotE_{KC}$. From Table 4, we can see that the *MaxSim* solution of model FLR_{FS} outperforms all other four models in terms of $AveS_H$, and outperforms the other three models in terms of $TotE_{KC}$ except model FLR_{CD08}.

Although model FLR_{FS} achieves better performances than other three models (FLR_{KC02}, FLR_{KC03} and FLR_{NN04}) in terms of $AveS_H$ and $TotE_{KC}$, and a better performance than FLR_{CD08} in terms of $AveS_H$, it cannot guarantee that the solution of model FLR_{FS} given in Eq. (29) is the global optimization. The FLR model given by Hojati et al. [9] is shown as follows:

$$\hat{\tilde{Y}}_i = (6.75 + 1.25m_{x_i}, 1.65 + 0.15m_{x_i})_L$$

[2] In some papers, a fuzzy number is expressed by its lower bound, centre and upper bound. To keep the notation consistency in this paper, we express a fuzzy number by its centre, left and right spreads.

Table 5
Dataset3.

i	$\tilde{X}_i = (m_{x_i}, \alpha_{x_i})_L$	$\tilde{Y}_i = (m_{y_i}, \alpha_{y_i})_L$	$\hat{\tilde{Y}}_i = (m_{\hat{y}_i}, \alpha_{\hat{y}_i})_L$	S_H	E_{KC}
1	$(2.0, 0.5)_L$	$(4.0, 0.5)_L$	$(4.0, 0.5)_L$	1.0000	0.0000
2	$(3.5, 0.5)_L$	$(5.5, 0.5)_L$	$(4.7857, 0.5)_L$	0.0426	0.9184
3	$(5.5, 1.0)_L$	$(7.5, 1.0)_L$	$(5.833, 1.6758)_L$	0.0766	2.2952
4	$(7.0, 0.5)_L$	$(6.5, 0.5)_L$	$(6.619, 0.5)_L$	0.6341	0.2239
5	$(8.5, 0.5)_L$	$(8.5, 1.0)_L$	$(7.4048, 0.5)_L$	0.0000	1.0000
6	$(10.5, 1.0)_L$	$(8.0, 1.0)_L$	$(8.4524, 1.6758)_L$	0.4957	0.9021
7	$(11.0.0.5)_L$	$(10.5, 0.5)_L$	$(8.7143, 0.5)_L$	0.0000	1.0000
8	$(12.5, 0.5)_L$	$(9.5, 0.5)_L$	$(9.5, 0.5)_L$	1.0000	0.0000
Average				0.4061	0.7925

Their model obtains a higher $AveS_H$ value (0.5515) and a lower $TotE_{KC}$ value (8.3986) than model FLR$_{FS}$ given in Eq. (29). This is because the *MaxSim* solution of model FLR$_{FS}$ obtained in this experiment is a local maximum, not a global one. Given appropriate initial values, model FLR$_{FS}$ can also find the solution obtained in [9]. This is because the solution given in [9] can be expressed by model FLR$_{FS}$ as follows:

$$K_0 = 6.75; \quad k_1 = 1.25; \quad k_{ll} = 0; \quad k_{lm} = 0.15; \quad c = 1.65$$

Although the solution given in [9] outperforms the solution of model FLR$_{FS}$ given in Eq. (29) in terms of $AveS_H$ and $TotE_{KC}$, the model proposed in [9] has the *spreads increasing problem*.

From this example, we can see that initial values are important for finding the *MaxSim* solution of FLR$_{FS}$. In Section 6, we will give a potential solution on how to set initial values in order to find the global optimization.

Example 3. In this example, we compare model FLR$_{FS}$ with other models using the fuzzy input–fuzzy output dataset given by Sakawa and Yano in [22], which is shown in the left half of Table 5. $\tilde{X}_i$'s are the observed independent variables. $\tilde{Y}_i$'s are the observed dependent variables.

The independent variables and the observed dependent variables in dataset3 are L-type. Thus, L-type fuzzy number based FLR$_{FS}$ model is adopted. According to the initial value setting strategy given in Section 5.1, the initial values can be set as[3]:

Initial values (b): $k_0 = 3.5724; \quad k_1 = 0.5193; \quad k_{ll} = 1; \quad k_{lm} = 0; \quad c = 0$

Then, the following regression model is obtained:

$$\hat{\tilde{Y}}_i = (2.9524 + 0.5238 m_{x_i}, 2.3516\alpha_{x_i} - 0.6758)_L \tag{30}$$

Estimated response $\hat{\tilde{Y}}_i$, S_H and E_{KC} for each instance of dataset3 are listed in the right half of Table 5.

The *MaxSim* solution of model FLR$_{FS}$ is compared with the other four models: FLR$_{KC02}$, FLR$_{KC03}$, FLR$_{NN04}$ and FLR$_{CD08}$, which are listed in Table 6. $\tilde{E}_i$'s in model FLR$_{CD08}$, are $\tilde{E}_1 = (-0.234, 0, 0.234)$, $\tilde{E}_2 = (-0.234, 0, 0.234)$, $\tilde{E}_3 = (0, 0, 0.935)$, $\tilde{E}_4 = (-0.234, 0, 0.234)$, $\tilde{E}_5 = (-0.234, 0, 0.234)$, $\tilde{E}_6 = (-0.935, 0, 0)$, $\tilde{E}_7 = (-0.234, 0, 0.234)$ and $\tilde{E}_8 = (-0.234, 0, 0.234)$.

To compare the performances of model FLR$_{FS}$, and models FLR$_{KC02}$ [11], FLR$_{KC03}$ [12], FLR$_{NN04}$ [17] and FLR$_{CD08}$ [2], $AveS_H$ and $TotE_{KC}$ are also used as evaluation measurements in this example.

The comparison of the performances of different methods on $AveS_H$ and $TotE_{KC}$ is given in Table 7. From Table 7, we can see that the average similarity of FLR$_{FS}$ is significantly higher than that of the other methods. Also, the total

[3] By applying the LS estimation to α_{y_i} and α_{x_i}, and α_{y_i} and m_{x_i}, we get $\alpha_{y_i} = 1.0 * \alpha_{x_i}$ and $\alpha_{y_i} = 0.5911 + 0.0045 m_{x_i}$. Since $0.0045 \ll 1$, we set the initial spreads of estimated response only depend on α_{x_i}.

Table 6
Fuzzy regression models of dataset3.

$\text{FLR}_{FS}(MaxSim)$ (given initial values (b))	$\hat{\tilde{Y}}_i = (2.9524 + 0.5238m_{x_i}, 2.3516\alpha_{x_i} - 0.6758)_L$
FLR_{KC02} [11]	$\hat{\tilde{Y}}_i = (3.5724 + 0.5193m_{x_i}, 0.5193\alpha_{x_i} + 0.24)_L$
FLR_{KC03} [12]	$\hat{\tilde{Y}}_i = (3.554 + 0.522m_{x_i}, 0.522\alpha_{x_i} + 0.951, 0.522\alpha_{x_i} + 0.949)_{LR}$
FLR_{NN04} [17]	$\hat{\tilde{Y}}_i = (3.5767 + 0.5467m_{x_i}, \alpha_{x_i})_L$
FLR_{CD08} [2]	$\hat{\tilde{Y}}_i = 3.5284 + 0.5298m_{x_i} + \tilde{E}_i$

Table 7
Comparison of the performance of difference methods on Dataset3.

Models	$AveS_H$	$TotE_{KC}$
$\text{FLR}_{FS}(MaxSim)$ (given the initial values (b))	0.4061	6.3396
FLR_{KC02} [11]	0.1499	7.470
FLR_{KC03} [12]	0.2351	9.363
FLR_{NN04} [17]	0.2026	7.541
FLR_{CD08} [2]	0.1854	7.000

error of FLR_{FS} is significantly lower than that of the other methods. Thus, model FLR_{FS} outperforms all four models (FLR_{KC02}, FLR_{KC03}, FLR_{NN04} and FLR_{CD08}) in terms of both $AveS_H$ and $TotE_{KC}$ on dataset3.

Example 4. In this example, we investigate the effectiveness of model FLR_{FS} for multiple regression on the restaurants data given by D'Urso [8]. The restaurants data listed in Table 8 are drawn from an Italian specialized book, which concerns the performances of the 30 good-quality Roman restaurants, where fuzzy inputs $\tilde{X}_{i1}$'s and $\tilde{X}_{i2}$'s are *decision on cooking* and *decision on cellar*, respectively, and $\tilde{Y}$ is *decision on service*.

The independent variables are two dimensional. Thus, model FLR_{FS} for multiple regression is adopted. According to the initial value setting strategy given in Section 5.1, the following parameters of model FLR_{FS} are obtained:

$$k_0 = -1.66; \quad k_1 = 1.33; \quad k_2 = 0.00$$

$$k_{ll1} = -0.91; \quad k_{lr1} = 0.00; \quad k_{lm1} = 0.63; \quad k_{ll2} = -0.65; \quad k_{lr2} = -1.48; \quad k_{lm2} = -0.28; \quad c_l = 0.80$$

$$k_{rl1} = 0.37; \quad k_{rr1} = 1.00; \quad k_{rm1} = -0.12; \quad k_{rl2} = 0.09; \quad k_{rr2} = 0.20; \quad k_{rm2} = 0.03; \quad c_r = 0.08 \tag{31}$$

The regression parameters of model $\text{FLR}_{D'Urso03}$ on dataset4 are estimated by minimizing the total squared Euclidean distance d^2 between the estimated and observed response variables, which are listed as follows [8]:

$$\mathbf{a} = (0.6498399, 0.4542534, 0.4924441)^{\mathrm{T}}$$

$$\mathbf{r} = (-1.868527, 2.3604004, 0.7392849)^{\mathrm{T}}$$

$$\mathbf{s} = (-0233325, -0.13392, 0.1271022)^{\mathrm{T}}$$

$$b = 0.1173197, \quad d = -0.401173, \quad g = 0.2306911, \quad h = -0.650102 \tag{32}$$

In order to compare model FLR_{FS} with model $\text{FLR}_{D'Urso0}$, both $AveS_H$ and *sum of* d^2 are used as the evaluation measurements, because: (i) the objective function of model FLR_{FS} is to maximize $AveS_H$ between $\tilde{Y}$ and $\hat{\tilde{Y}}$ and (ii) the objective function of $\text{FLR}_{D'Urso03}$ is to minimize total squared Euclidean distance d^2 between $\tilde{Y}$and $\hat{\tilde{Y}}$.

The comparison of the performances of FLR_{FS} and $\text{FLR}_{D'Urso03}$ on both $AveS_H$ and total d^2 is given in Table 9. From Table 9, we can see that model FLR_{FS} outperforms model $\text{FLR}_{D'Urso03}$ in terms of both $AveS_H$ and *sum of* d^2.

Table 8
Dataset4: restaurants data.

i	$\tilde{X}_{i1} = (m_{x_{i1}}, \alpha_{x_{i1}}, \beta_{x_{i1}})_{LR}$	$\tilde{X}_{i2} = (m_{x_{i2}}, \alpha_{x_{i2}}, \beta_{x_{i2}})_{LR}$	$\tilde{Y}_i = (m_{y_i}, \alpha_{y_i}, \beta_{y_i})_{LR}$
1	$(7, 0.5, 1.25)_{LR}$	$(8, 0.75, 1)_{LR}$	$(8, 0.75, 1)_{LR}$
2	$(7, 0.5, 1.25)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(6, 0.25, 0.5)_{LR}$
3	$(6, 0.25, 0.5)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(6, 0.25, 0.5)_{LR}$
4	$(8, 0.75, 1)_{LR}$	$(9, 0, 1)_{LR}$	$(9, 0, 1)_{LR}$
5	$(8, 0.75, 1)_{LR}$	$(8, 0.75, 1)_{LR}$	$(8, 0.75, 1)_{LR}$
6	$(6, 0.25, 0.5)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(5, 0, 1)_{LR}$
7	$(7, 0.5, 1.25)_{LR}$	$(8, 0.75, 1)_{LR}$	$(7, 0.5, 1.25)_{LR}$
8	$(7, 0.5, 1.25)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(5, 0, 1)_{LR}$
9	$(7, 0.5, 1.25)_{LR}$	$(8, 0.75, 1)_{LR}$	$(7, 0.5, 1.25)_{LR}$
10	$(6, 0.25, 0.5)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(6, 0.25, 0.5)_{LR}$
11	$(7, 0.5, 1.25)_{LR}$	$(8, 0.75, 1)_{LR}$	$(8, 0.75, 1)_{LR}$
12	$(7, 0.5, 1.25)_{LR}$	$(6, 0.25, 0.5)_{LR}$	$(6, 0.25, 0.5)_{LR}$
13	$(7, 0.5, 1.25)_{LR}$	$(8, 0.75, 1)_{LR}$	$(9, 0, 1)_{LR}$
14	$(7, 0.5, 1.25)_{LR}$	$(8, 0.75, 1)_{LR}$	$(8, 0.75, 1)_{LR}$
15	$(7, 0.5, 1.25)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(7, 0.5, 1.25)_{LR}$
16	$(7, 0.5, 1.25)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(7, 0.5, 1.25)_{LR}$
17	$(6, 0.25, 0.5)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(6, 0.25, 0.5)_{LR}$
18	$(7, 0.5, 1.25)_{LR}$	$(8, 0.75, 1)_{LR}$	$(7, 0.5, 1.25)_{LR}$
19	$(7, 0.5, 1.25)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(8, 0.75, 1)_{LR}$
20	$(7, 0.5, 1.25)_{LR}$	$(9, 0, 1)_{LR}$	$(7, 0.5, 1.25)_{LR}$
21	$(7, 0.5, 1.25)_{LR}$	$(8, 0.75, 1)_{LR}$	$(7, 0.5, 1.25)_{LR}$
22	$(7, 0.5, 1.25)_{LR}$	$(8, 0.75, 1)_{LR}$	$(6, 0.25, 0.5)_{LR}$
23	$(7, 0.5, 1.25)_{LR}$	$(9, 0, 1)_{LR}$	$(7, 0.5, 1.25)_{LR}$
24	$(7, 0.5, 1.25)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(8, 0.75, 1)_{LR}$
25	$(7, 0.5, 1.25)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(6, 0.25, 0.5)_{LR}$
26	$(7, 0.5, 1.25)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(6, 0.25, 0.5)_{LR}$
27	$(7, 0.5, 1.25)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(7, 0.5, 1.25)_{LR}$
28	$(7, 0.5, 1.25)_{LR}$	$(8, 0.75, 1)_{LR}$	$(7, 0.5, 1.25)_{LR}$
29	$(7, 0.5, 1.25)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(7, 0.5, 1.25)_{LR}$
30	$(6, 0.25, 0.5)_{LR}$	$(7, 0.5, 1.25)_{LR}$	$(6, 0.25, 0.5)_{LR}$

Table 9
Comparison of the performance of difference methods on Dataset4.

Models	$AveS_H$	Sum of d^2
FLR_{FS}(*MaxSim*)	0.5675	64.5683
$\text{FLR}_{D'Urso03}$	0.2263	73.6945

6. Conclusions and future work

In this paper, we proposed model FLR_{FS} that has more flexible spreads compared with previous FLR models. A property of model FLR_{FS} is that the spreads of estimated response variables are able to fit the spreads of observed response variables, no matter if the spreads of the observed response variables are increased, decreased or unchanged when the spreads and magnitudes of the independent variables change. This property makes model FLR_{FS} be able to avoid the *spreads increasing problem* that exists in model FLR_{FF} and overcome the deficiencies of models FLR_{KC02}, FLR_{KC03}, FLR_{NN04}, and $\text{FLR}_{D'Urso03}$, which are mentioned in Section 3. The number of parameters in model FLR_{FS} only proportionally increases with the increase of the dimension of independent variables, while the number of parameters in model FLR_{CD08} increases with the increase of both the dimension and the number of independent variables. This makes FLR_{FS} more suitable to dataset with large instance number than FLR_{CD08}.

The parameters of model FLR_{FS} are estimated by maximizing the average similarity between the estimated and observed response variables. The experimental results show that model FLR_{FS} has a better performance than previous models in terms of $AveS_H$, $TotE_{KC}$ and $sum\ d^2$. Also, the experimental results show that initial value setting is important for parameter estimation of FLR_{FS} when the objective function is set as maximizing the average similarity between the estimated and observed response variables. Although we have given a strategy for the initial value setting, it cannot guarantee that the generated regression model is the global optimization. Our future work is to find a more sophisticated initial value setting strategy to achieve better solutions of model FLR_{FS}. Genetic algorithms may be potential solutions for the initial value setting problem of model FLR_{FS}.

Acknowledgements

The authors gratefully acknowledge the comments and suggestions of the anonymous reviewers, which helped in improving the clarity and the quality of the paper.

References

[1] A. Bargiela, W. Pedrycz, T. Nakashima, Multiple regression with fuzzy data, Fuzzy Sets and Systems 158 (19) (2007) 2169–2188.
[2] S.P. Chen, J.F. Dang, A variable spread fuzzy linear regression model with higher explanatory power and forecasting accuracy, Information Sciences 178 (2008) 3973–3988.
[3] R. Coppi, Management of uncertainty in statistical reasoning: the case of regression analysis, International Journal of Approximate Reasoning 47 (3) (2008) 284–305.
[4] R. Coppi, P. D'Urso, Regression analysis with fuzzy informational paradigm: a least-squares approach using membership function information, International Journal of Pure and Applied Mathematics 8 (2003) 279–306.
[5] R. Coppi, P. D'Urso, P. Giordani, A. Santoro, Least squares estimation of a linear regression model with LR fuzzy response, Computational Statistics and Data Analysis 51 (2006) 267–286.
[6] P. Diamond, Fuzzy least squares, Information Science 46 (1988) 141–157.
[7] D. Dubois, H. Prade, Fuzzy Sets and Systems: Theory and Application, Academic Press, New York, 1980.
[8] P. D'Urso, Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data, Computational Statistics and Data Analysis 42 (1–2) (2003) 47–72.
[9] M. Hojati, C.R. Bector, K. Smimou, A simple method for computation of fuzzy linear regression, European Journal of Operational Research 166 (1) (2005) 172–184.
[10] W.L. Hung, M.S. Yang, An omission approach for detecting outliers in fuzzy regression models, Fuzzy Sets and Systems 157 (23) (2006) 3109–3122.
[11] C. Kao, C.-L. Chyu, A fuzzy linear regression model with better explanatory power, Fuzzy Sets and Systems 126 (2002) 401–409.
[12] C. Kao, C.-L. Chyu, Least-squares estimates in fuzzy regression analysis, European Journal of Operational Research 148 (2003) 426–435.
[13] C. Kao, P.H. Lin, Entropy for fuzzy regression analysis, International Journal of Systems Science 36 (2005) 869–876.
[14] C. Kahraman, A. Beskese, F.T. Bozbura, Fuzzy regression approaches and applications, in: Fuzzy Applications in Industrial Engineering, 2006, pp. 589–615.
[15] B. Kim, R.R. Bishu, Evaluation of fuzzy linear regression models by comparing membership functions, Fuzzy Sets and Systems 100 (1998) 343–352.
[16] M. Modarres, E. Nasrabadi, M.M. Nasrabadi, Fuzzy linear regression models with least square errors, Applied Mathematics and Computation 163 (2005) 977–989.
[17] M.M. Nasrabadi, E. Nasrabadi, A mathematical-programming approach to fuzzy linear regression analysis, Applied Mathematics and Computation 155 (3) (2004) 873–881.
[18] M.M. Nasrabadi, E. Nasrabadi, A.R. Nasrabady, Fuzzy linear regression analysis: a multi-objective programming approach, Applied Mathematics and Computation 163 (2005) 245–251.
[19] E.C. Özelkan, L. Duckstein, Multi-objective fuzzy regression: a general framework, Computers & Operations Research 27 (2000) 635–652.
[20] D.T. Redden, W.H. Woodall, Properties of certain fuzzy linear regression methods, Fuzzy Sets and Systems 64 (1994) 361–375.
[21] M. Sakawa, H. Yano, Fuzzy linear regression and its applications, in: J. Kacprzyk, M. Fedrizzi (Eds.), Fuzzy Regression Analysis, Physica-Verlag, Heidelberg, 1992, pp. 61–80.
[22] M. Sakawa, H. Yano, Multiobjective fuzzy linear regression analysis for fuzzy input–output data, Fuzzy Sets and Systems 47 (1992) 173–181.
[23] H. Shakouri, G.R. Nadimi, F. Ghaderi, Fuzzy linear regression models with absolute errors and optimum uncertainty, in: IEEE Internat. Conf. on Industrial Engineering and Engineering Management, 2007, pp. 917–921.
[24] H. Tanaka, I. Hayashi, J. Watada, Possibilistic linear regression analysis for fuzzy data, European Journal of Operational Research 40 (1989) 389–396.
[25] H. Tanaka, S. Uejima, K. AsaL, Linear regression analysis with fuzzy model, IEEE Transactions on Systems, Man and Cybernetics 12 (1982) 903–907.

[26] L. Tran, L. Duckstein, Multiobjective fuzzy regression with central tendency and possibilistic properties, Fuzzy Sets and Systems 130 (2002) 21–31.
[27] M.S. Yang, T.S. Lin, Fuzzy least-squares linear regression analysis for fuzzy input–output data, Fuzzy Sets and Systems 126 (2002) 389–399.
[28] L.A. Zadeh, Fuzzy sets, Information and Control 8 (1965) 338–353.
[29] H.J. Zimmermann, Fuzzy Set Theory and its Application, Kluwer Academic Press, Dordrecht, 1991.