Computers and Mathematics with Applications 60 (2010) 1696–1705

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

A fuzzy varying coefficient model and its estimation

Si-Lian Shen, Chang-Lin Mei *, Jian-Ling Cui

School of Science, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, People's Republic of China

ARTICLE INFO

Article history:
Received 20 July 2009
Received in revised form 30 June 2010
Accepted 30 June 2010

Keywords:
Fuzzy regression
Fuzzy varying coefficient model
Restricted weighted least-squares
Fuzzy number
Smoothing parameter

ABSTRACT

The fuzzy linear regression model has been a useful tool for analyzing relationships between a set of variables in a fuzzy environment and has been extensively studied in the literature. However, this model may fail to reflect the more complicated regression relationships that are usually found in practice because of its simple and predefined linear structure. In order to enhance the feasibility and adaptability of the fuzzy linear models, we propose in this paper a fuzzy varying coefficient model in which the fuzzy coefficients in the fuzzy linear models are allowed to vary with a covariate. A restricted weighted least-squares estimation is suggested for locally fitting the model. Furthermore, some real-world datasets are analyzed in order to evaluate the performance of the proposed method, and the results show that the proposed model with its estimation approach performs satisfactorily in predicting the fuzzy response even in the case where the regression relationship is complicated.

1. Introduction

Regression analysis is a methodology commonly used for analyzing the relationship between a response variable and a set of explanatory variables. In the conventional regression analysis, the associated uncertainty is randomness and observations of both the response variable and the explanatory variables are assumed to be crisp values. This kind of regression analysis has been well developed in the framework of probability theory and statistics. In practice, however, we are frequently faced with another type of uncertainty, of vagueness or fuzziness. Specifically, observations of some variables cannot be collected precisely, especially when the data are influenced by subjective judgment or expressed linguistically. Therefore, it is necessary to develop some methods for performing regression analysis in a fuzzy environment.

On the basis of Zadeh's fuzzy theory [1], Tanaka et al. [2] were the first to develop the method of fuzzy regression analysis. Since then, many approaches have been proposed in the literature of fuzzy regression. Roughly speaking, these approaches can be classified into two categories. One is the mathematical programming based methods in which the total spread of the fuzzy parameters is minimized under some constraints to calibrate the fuzzy regression models (see, for example, [2–6]). The other is the fuzzy least-squares based methods that minimize the total square of the deviations between the observed and estimated values of the response to estimate the fuzzy parameters (see, for example, [7–14]). The fuzzy least-squares based methods are appealing because they can be implemented in the framework of traditional least-squares regression.

Most of the existing research on fuzzy regression analysis has focused on the fuzzy parametric models, especially on the fuzzy linear regression model. In many real-world situations, however, the functional relationship between the response variable and explanatory variables is difficult to predefine. If a preassumed parametric model does not follow the data generation mechanism well, a large estimation bias will be obtained. To overcome this problem, some model-free fuzzy regression methods have been developed, of which the most popular ones are perhaps the fuzzified neural networks [15–19], the support vector fuzzy regression machines [20–26] and some fuzzy nonparametric regression approaches [27,28]. These methods make fewer assumptions on the regression function and thus are more flexible in exploring hidden structures of

* Corresponding author. Tel.: +86 29 82665483; fax: +86 29 82663938.
E-mail addresses: slshen@stu.xjtu.edu.cn (S.-L. Shen), clmei@mail.xjtu.edu.cn (C.-L. Mei).

0898-1221/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2010.06.049

the regression relationships. However, they may fail to incorporate some prior information if available and suffer from the so-called curse of dimensionality which makes the analysis practically invalid when the number of the explanatory variables is large.

Ideally, a regression model with the simplicity of linear models and the flexibility of nonparametric models may be preferred. Along this line of thinking, we propose in this paper a fuzzy varying coefficient regression model with the explanatory variables being crisp and the response being assumed to be a symmetric fuzzy number. In the model, the basic structure of a linear fuzzy regression model is retained, but the regression coefficients are allowed to vary as some unknown fuzzy functions of a covariate. These varying coefficients can greatly increase model flexibility and adaptability while the local linear structure connecting the explanatory variables and the response can effectively avoid the curse of dimensionality encountered in fuzzy nonparametric models. In order to overcome the shortcoming that some least-squares based methods cannot guarantee automatically the non-negativity of the estimated spreads of the fuzzy coefficients (see, for example, [11,28–30]), we suggest a restricted weighted least-squares procedure for locally fitting the proposed fuzzy varying coefficient model. Some real-world datasets are further analyzed with the proposed model fitting method in order to evaluate its performance.

The remainder of this paper is organized as follows. In Section 2, the fuzzy varying coefficient model is proposed and the restricted weighted least-squares procedure is suggested for fitting the model. Three real-world datasets are analyzed in Section 3 with the proposed method and the results are compared with some other existing fuzzy regression methods. The paper is then ended with some final remarks.

2. The fuzzy varying coefficient model and its restricted weighted least-squares estimation

2.1. The fuzzy varying coefficient model

We introduce in this subsection the fuzzy varying coefficient model used to describe a regression problem where the explanatory variables are crisp and the response is a symmetric fuzzy number.

A fuzzy number $\tilde{A}$ is a fuzzy subset of the real line $\mathscr{R}$ with its membership function $\mu_{\tilde{A}}(x)$ satisfying the following conditions:

(i) for any $\alpha \in (0, 1)$, the α-level set of $\mu_{\tilde{A}}(x)$ is a closed interval;
(ii) there exists an $x \in \mathscr{R}$ such that $\mu_{\tilde{A}}(x) = 1$;
(iii) for any $\lambda \in [0, 1]$ and $x_1, x_2 \in \mathscr{R}$, $\mu_{\tilde{A}}(\lambda x_1 + (1-\lambda)x_2) \geq \min(\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_2))$. In particular, a fuzzy number $\tilde{A}$ with the following membership function:

$$\mu_{\tilde{A}}(x) = \begin{cases} L\left(\dfrac{|A-x|}{S}\right), & A - S \leq x \leq A + S; \\ 0, & \text{elsewhere} \end{cases} \tag{1}$$

is called a symmetric fuzzy number and is usually denoted by $(A, S)_L$, where A and S are the center and spread of $\tilde{A}$ respectively and $L(\cdot)$ is a strictly decreasing function on $[0, 1]$ with $L(0) = 1$ and $L(1) = 0$. The choice of $L(\cdot)$ is generally dependent upon subjective judgment.

Let $X_1, X_2, \ldots, X_p$ and U be crisp explanatory variables and $\tilde{Y} = (Y, S)_L$ be the symmetric fuzzy response with center Y and spread S. We propose a fuzzy varying coefficient model:

$$\tilde{Y} = \tilde{\beta}_1(U)X_1 + \tilde{\beta}_2(U)X_2 + \cdots + \tilde{\beta}_p(U)X_p. \tag{2}$$

In the model, $\tilde{\beta}_j(U) = \left(\beta_j(U), \sigma_j(U)\right)_L$ is assumed to be a symmetric fuzzy number with its center $\beta_j(U)$ and spread $\sigma_j(U)$ being unknown functions of U for each $j = 1, 2, \ldots, p$. The operations in the fuzzy regression function $\tilde{\beta}_1(U)X_1 + \tilde{\beta}_2(U)X_2 + \cdots + \tilde{\beta}_p(U)X_p$ are according to Zedeh's extension principle [31], in which it is assumed, without loss of generality, that all of the explanatory variables take non-negative values (otherwise, an appropriate transformation can be performed to make this assumption valid). Generally, we take $X_1 \equiv 1$ to make the model include a fuzzy varying intercept.

Suppose that $\{\tilde{y}_i = (y_i, s_i)_L;\ u_i, x_{i1}, x_{i2}, \ldots, x_{ip};\ i = 1, 2, \ldots, n\}$ are n observations of the symmetric fuzzy response $\tilde{Y}$ and the crisp explanatory variables $X_1, X_2, \ldots, X_p$ and U. The sample form of the model (2) is

$$\tilde{y}_i = \tilde{\beta}_1(u_i)x_{i1} + \tilde{\beta}_2(u_i)x_{i2} + \cdots + \tilde{\beta}_p(u_i)x_{ip}, \quad i = 1, 2, \ldots, n, \tag{3}$$

or

$$\begin{cases} y_i = \beta_1(u_i)x_{i1} + \beta_2(u_i)x_{i2} + \cdots + \beta_p(u_i)x_{ip}; \\ s_i = \sigma_1(u_i)x_{i1} + \sigma_2(u_i)x_{i2} + \cdots + \sigma_p(u_i)x_{ip}, \end{cases} \quad i = 1, 2, \ldots, n. \tag{4}$$

The main task of fitting the model is to estimate the fuzzy regression coefficients $\tilde{\beta}_j(U)$ $(j = 1, 2, \ldots, p)$ at any $U = u_0$ on the basis of the observations.

The fuzzy varying coefficient model is an extension of the fuzzy linear regression model, by allowing the fuzzy regression coefficients to be dependent upon another explanatory variable U to increase the flexibility and adaptability of the fuzzy linear model. Therefore, it is possible to expect a less biased estimate of the response variable. Additionally, as the regression coefficients in model (2) are one-dimensional, the curse of dimensionality problem can be avoided.

From the practical point of view, one of the most important applications of the fuzzy varying coefficient model is perhaps in exploring a dynamic relationship between a fuzzy variable and a set of crisp covariates by taking the values of U in the regression coefficients to be the time points where the observations of $\tilde{Y}$ and $X_1, X_2, \ldots, X_p$ are made. In many real-world problems from econometrics, finance, epidemiology, environmental science and so on, fitting this kind of dynamic relationship is especially helpful for understanding how the association between the response and explanatory variables varies over time.

2.2. The restricted weighted least-squares estimation of the fuzzy varying coefficient model

Before embarking on calibrating the fuzzy varying coefficient model, we briefly introduce the distance from [32] that will be used to formulate the estimation equations of the fuzzy varying coefficient model.

Let $\tilde{A}$ and $\tilde{B}$ be two fuzzy numbers. Xu [32] proposed a distance between $\tilde{A}$ and $\tilde{B}$ as

$$d^2(\tilde{A}, \tilde{B}) = \int_0^1 f(\lambda)\bar{d}^2(A_\lambda, B_\lambda)\mathrm{d}\lambda, \tag{5}$$

where $f(\lambda)$ is an increasing function on $[0, 1]$ satisfying $f(0) = 0$ and $\int_0^1 f(\lambda)\mathrm{d}\lambda = 1/2$, $A_\lambda = [a_1(\lambda), a_2(\lambda)]$ and $B_\lambda = [b_1(\lambda), b_2(\lambda)]$ are λ-level sets of $\tilde{A}$ and $\tilde{B}$ respectively and

$$\bar{d}^2(A_\lambda, B_\lambda) = [a_1(\lambda) - b_1(\lambda)]^2 + [a_2(\lambda) - b_2(\lambda)]^2.$$

As pointed out in [32], this distance incorporates, through the λ-level sets A_λ and B_λ, the information of the membership functions of fuzzy numbers $\tilde{A}$ and $\tilde{B}$. Furthermore, the degree of nearness of and overlap between the λ-level sets A_λ and B_λ, measured by $\bar{d}^2(A_\lambda, B_\lambda)$, is weighted by the monotonically increasing function $f(\lambda)$ to emphasize the contribution of the higher values of λ to the distance between $\tilde{A}$ and $\tilde{B}$.

In particular, when $\tilde{A} = (A, S_A)_L$ and $\tilde{B} = (B, S_B)_L$ are two symmetric fuzzy numbers, the λ-level sets of $\tilde{A}$ and $\tilde{B}$ are respectively

$$A_\lambda = [A - L^{-1}(\lambda)S_A, A + L^{-1}(\lambda)S_A] \quad \text{and} \quad B_\lambda = [B - L^{-1}(\lambda)S_B, B + L^{-1}(\lambda)S_B],$$

where $L^{-1}(\lambda)$ is the inverse function of $L(\lambda)$. Therefore

$$\bar{d}^2(A_\lambda, B_\lambda) = 2(A - B)^2 + 2(L^{-1}(\lambda))^2(S_A - S_B)^2$$

and the distance (5) becomes

$$d^2(\tilde{A}, \tilde{B}) = (A - B)^2 + \triangle(L, f)(S_A - S_B)^2, \tag{6}$$

where $\triangle(L, f) = 2\int_0^1 f(\lambda)(L^{-1}(\lambda))^2\mathrm{d}\lambda$. Furthermore, when $\tilde{A}$ and $\tilde{B}$ are symmetric triangular fuzzy numbers (that is, $L(t) = 1 - t \quad 0 \le t \le 1$) and $f(\lambda) = \lambda$, we have $\triangle(L, f) = 1/6$.

In consideration of the advantages of Xu's distance and its simple form for the symmetric fuzzy numbers, the distance in (6) is used henceforth to formulate the objective function with the restrictions that the spreads of the fuzzy coefficients are non-negative for calibrating the fuzzy varying coefficient model.

Suppose that the domain of variable U is $\mathscr{U}$ and u_0 is a given point in $\mathscr{U}$. On the basis of the distance (6) and the principle of the kernel smoothing in statistics, we formulate the following restricted weighted least-squares problem. That is, the objective function is

$$\begin{cases} \Psi\left(\beta_1(u_0), \ldots, \beta_p(u_0); \sigma_1(u_0), \ldots, \sigma_p(u_0)\right) \\ \quad = \displaystyle\sum_{i=1}^n d^2\left((y_i, s_i)_L, \left(\sum_{j=1}^p \beta_j(u_0)x_{ij}, \sum_{j=1}^p \sigma_j(u_0)x_{ij}\right)_L\right) K_h(u_i - u_0) \\ \quad = \displaystyle\sum_{i=1}^n \left(\left(y_i - \sum_{j=1}^p \beta_j(u_0)x_{ij}\right)^2 + \triangle(L, f)\left(s_i - \sum_{j=1}^p \sigma_j(u_0)x_{ij}\right)^2\right) K_h(u_i - u_0) \\ \quad = \displaystyle\sum_{i=1}^n \left(y_i - \sum_{j=1}^p \beta_j(u_0)x_{ij}\right)^2 K_h(u_i - u_0) + \triangle(L, f)\sum_{i=1}^n \left(s_i - \sum_{j=1}^p \sigma_j(u_0)x_{ij}\right)^2 K_h(u_i - u_0) \\ \quad \text{subject to} \quad \sigma_j(u_0) \ge 0, \quad j = 1, 2, \ldots, p. \end{cases} \tag{7}$$

And it is minimized with respect to $\{\beta_j(u_0), \sigma_j(u_0), j = 1, 2, \ldots, p\}$, where $K_h(\cdot) = K(\cdot/h)/h$, with $K(\cdot)$ being a given kernel function and h being the smoothing parameter.

It can be observed that the objective function in Eq. (7) is a summation of two separate parts in which each part includes a different group of unknown parameters, that is, $\{\beta_j(u_0),\ j = 1, 2, \dots, p\}$ and $\{\sigma_j(u_0),\ j = 1, 2, \dots, p\}$, and the factor $\triangle(L, f)$ is independent of the unknown parameters. Therefore, the restricted weighted least-squares problem is equivalent to minimizing the following two equations:

$$\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p}\beta_j(u_0)x_{ij}\right)^2 K_h(u_i - u_0) \tag{8}$$

and

$$\begin{cases}\displaystyle\sum_{i=1}^{n}\left(s_i - \sum_{j=1}^{p}\sigma_j(u_0)x_{ij}\right)^2 K_h(u_i - u_0)\\ \text{subject to} \quad \sigma_j(u_0) \geq 0, \quad j = 1, 2, \dots, p.\end{cases} \tag{9}$$

Let

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p}\\ x_{21} & x_{22} & \cdots & x_{2p}\\ \vdots & \vdots & \vdots & \vdots\\ x_{n1} & x_{n2} & \cdots & x_{np}\end{pmatrix}, \qquad \mathbf{Y} = \begin{pmatrix} y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}, \qquad \mathbf{S} = \begin{pmatrix} s_1\\ s_2\\ \vdots\\ s_n\end{pmatrix}, \tag{10}$$

$$\mathbf{W}(u_0) = \mathrm{Diag}\left(K_h(u_1 - u_0), K_h(u_2 - u_0), \dots, K_h(u_n - u_0)\right), \tag{11}$$

$$\boldsymbol{\beta}(u_0) = \left(\beta_1(u_0), \beta_2(u_0), \dots, \beta_p(u_0)\right)^{\mathrm{T}},$$

$$\boldsymbol{\sigma}(u_0) = \left(\sigma_1(u_0), \sigma_2(u_0), \dots, \sigma_p(u_0)\right)^{\mathrm{T}}.$$

We here assume that the inverse matrix of $\mathbf{X}^{\mathrm{T}}\mathbf{W}(u_0)\mathbf{X}$ exists for each $u_0 \in \mathscr{U}$. Then the solution of the weighted least-squares problem (8), that is, the estimate of the center vector $\boldsymbol{\beta}(u_0)$ of the fuzzy coefficients, can be obtained using matrix notation as

$$\begin{aligned}\hat{\boldsymbol{\beta}}(u_0) &= \left(\hat{\beta}_1(u_0), \hat{\beta}_2(u_0), \dots, \hat{\beta}_p(u_0)\right)^{\mathrm{T}}\\ &= \left(\mathbf{X}^{\mathrm{T}}\mathbf{W}(u_0)\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}(u_0)\mathbf{Y}\\ &= \mathbf{H}(u_0)\mathbf{Y},\end{aligned} \tag{12}$$

where

$$\mathbf{H}(u_0) = \left(\mathbf{X}^{\mathrm{T}}\mathbf{W}(u_0)\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}(u_0) \tag{13}$$

is a $p \times n$ matrix. Let $\mathbf{e}_j$ denote a p-dimensional column vector with its jth element being 1 and the others being 0. Then we obtain from (12) that

$$\hat{\beta}_j(u_0) = \mathbf{e}_j^{\mathrm{T}}\hat{\boldsymbol{\beta}}(u_0) = \mathbf{e}_j^{\mathrm{T}}\mathbf{H}(u_0)\mathbf{Y}, \quad j = 1, 2, \dots, p. \tag{14}$$

The spread estimates of the fuzzy coefficients at $u_0 \in \mathscr{U}$ can be obtained by solving the restricted weighted least-squares problem (9). In solving the least-squares problem with the same restrictions as in Eq. (9), Waterman [33] proposed an algorithm without the use of linear programming in which the solution can be obtained by performing the regression on all possible subsets of the explanatory variables and solving a series of unrestricted least-squares problems. This algorithm is easy to carry out with some modern computer software and can directly be extended to the case of the restricted weighted least-squares problem (9). The extended procedure can be described in detail as follows.

(i) Solving the weighted least-squares problem (9) without considering the restrictions, we obtain, with the same procedure as was used for solving (8),

$$\hat{\sigma}_j(u_0) = \mathbf{e}_j^{\mathrm{T}}\left(\mathbf{X}^{\mathrm{T}}\mathbf{W}(u_0)\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}(u_0)\mathbf{S}, \quad j = 1, 2, \dots, p. \tag{15}$$

If $\hat{\sigma}_j(u_0) \geq 0$ for all $j = 1, 2, \dots, p$, then no further work needs to be done.

(ii) If at least one of $\hat{\sigma}_j(u_0)$ ($j = 1, 2, \dots, p$) is negative, then we need to solve a number of unrestricted weighted least-squares problems. Specifically, let

$$\mathscr{C} = \{C : C \subset \{1, 2, \dots, p\} \text{ and } C \neq \{1, 2, \dots, p\}\}$$

denote the index set with its elements being all possible subsets of $\{1, 2, \dots, p\}$ except for $\{1, 2, \dots, p\}$ itself. For each $C = \{l_1, l_2, \dots, l_k\} \in \mathscr{C}$, where $1 \leq k < p$, let

$$Q_C\left(\sigma_{l_1}(u_0), \sigma_{l_2}(u_0), \dots, \sigma_{l_k}(u_0)\right) = \sum_{i=1}^{n}\left(s_i - \sum_{j=1}^{k}\sigma_{l_j}(u_0)x_{il_j}\right)^2 K_h(u_i - u_0). \tag{16}$$

Select $\sigma_{l_1}(u_0), \sigma_{l_2}(u_0), \ldots, \sigma_{l_k}(u_0)$ to minimize $Q_C\left(\sigma_{l_1}(u_0), \sigma_{l_2}(u_0), \ldots, \sigma_{l_k}(u_0)\right)$; then the solution of this unrestricted weighted least-squares problem can be obtained by using

$$\left(\hat{\sigma}_{l_1}(u_0), \hat{\sigma}_{l_2}(u_0), \ldots, \hat{\sigma}_{l_k}(u_0)\right)^{\mathrm{T}} = \left(\mathbf{X}_C^{\mathrm{T}}\mathbf{W}(u_0)\mathbf{X}_C\right)^{-1}\mathbf{X}_C^{\mathrm{T}}\mathbf{W}(u_0)\mathbf{S}, \tag{17}$$

where $\mathbf{X}_C$ is an $n \times k$ matrix formulated from the l_1th, l_2th, $\ldots$, l_kth columns of $\mathbf{X}$, and $\mathbf{X}$, $\mathbf{S}$ and $\mathbf{W}(u_0)$ are shown respectively in Eqs. (10) and (11). If $\hat{\sigma}_{l_j}(u_0) \geq 0$ for all $j = 1, 2, \ldots, k$, then compute the value of $Q_C\left(\hat{\sigma}_{l_1}(u_0), \hat{\sigma}_{l_2}(u_0), \ldots, \hat{\sigma}_{l_k}(u_0)\right)$ according to Eq. (16); otherwise, nothing is done. For $C = \emptyset \in \mathscr{C}$, let $Q_C(\emptyset) = \sum_{i=1}^{n} s_i^2 K_h(u_i - u_0)$.

Let

$$\Gamma = \left\{Q_C\left(\hat{\sigma}_{l_1}(u_0), \ldots, \hat{\sigma}_{l_k}(u_0)\right) : C = \{l_1, \ldots, l_k\} \in \mathscr{C} \text{ and } \hat{\sigma}_{l_j}(u_0) \geq 0 \text{ for all } j = 1, \ldots, k\right\} \cup \{Q_C(\emptyset)\}.$$

Suppose that the minimal element of Γ is $Q_D\left(\hat{\sigma}_{q_1}(u_0), \ldots, \hat{\sigma}_{q_{k_0}}(u_0)\right)$ with $D = \{q_1, \ldots, q_{k_0}\} \in \mathscr{C}$ and define

$$\hat{\sigma}_j(u_0) = \begin{cases} \hat{\sigma}_j(u_0), & j \in \{q_1, q_2, \ldots, q_{k_0}\}, \\ 0, & j \notin \{q_1, q_2, \ldots, q_{k_0}\}, \end{cases} \tag{18}$$

for $j = 1, 2, \ldots, p$. Then $\hat{\sigma}_j(u_0)$ $(j = 1, 2, \ldots, p)$ are the solutions of the restricted weighted least-squares problem (9) and therefore they are the spread estimates for the coefficients at $u_0 \in \mathscr{U}$.

Performing the above estimation procedure at $u_0 = u_1, u_2, \ldots, u_n$ respectively, we can obtain the fitted values of the center of the fuzzy response $\tilde{Y}$ as

$$\begin{aligned} \hat{y}_i &= \hat{\beta}_1(u_i)x_{i1} + \hat{\beta}_2(u_i)x_{i2} + \cdots + \hat{\beta}_p(u_i)x_{ip} \\ &= (x_{i1}, x_{i2}, \ldots, x_{ip})\left(\hat{\beta}_1(u_i), \hat{\beta}_2(u_i), \ldots, \hat{\beta}_p(u_i)\right)^{\mathrm{T}} \\ &= \mathbf{x}_i^{\mathrm{T}}\mathbf{H}(u_i)\mathbf{Y}, \quad i = 1, 2, \ldots, n, \end{aligned} \tag{19}$$

where $\mathbf{x}_i^{\mathrm{T}} = (x_{i1}, x_{i2}, \ldots, x_{ip})$ is the ith row of $\mathbf{X}$. Similarly, the fitted values of the spread of $\tilde{Y}$ are

$$\hat{s}_i = \hat{\sigma}_1(u_i)x_{i1} + \hat{\sigma}_2(u_i)x_{i2} + \cdots + \hat{\sigma}_p(u_i)x_{ip}, \quad i = 1, 2, \ldots, n. \tag{20}$$

Therefore, the estimates of the fuzzy response $\tilde{Y}$ at u_i $(i = 1, 2, \ldots, n)$ are

$$\hat{\tilde{y}}_i = \left(\hat{y}_i, \hat{s}_i\right)_L, \quad i = 1, 2, \ldots, n. \tag{21}$$

2.3. Selection of the kernel function and the smoothing parameter

In order to implement the above calibration procedure of the fuzzy varying coefficient model, we have to determine the kernel function $K(\cdot)$ and the smoothing parameter h in advance. The role of $K(\cdot)$ is to place more emphasis on the observations close to u_0 than those farther away in generating the estimates of the center and spread of the fuzzy coefficients at u_0. As in statistical nonparametric regression, two commonly used kernel functions are the Gaussian kernel

$$K(t) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}t^2\right) \tag{22}$$

and the Epanecknikov kernel

$$K(t) = \begin{cases} \frac{3}{4}(1 - t^2), & \text{if } |t| \leq 1; \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

The role of the smoothing parameter h in the weight $K_h(\cdot) = K(\cdot/h)/h$ is to adjust the degree of smoothness of the estimates of the center and spread of the fuzzy coefficients. Here, we use the distance (6) used to fuzzify the cross-validation procedure (see, for example, [34]) in statistics for selecting the optimal value of the smoothing parameter h. Specifically, for each of $i = 1, 2, \ldots, n$, delete the ith observation $\tilde{y}_i = (y_i, s_i)_L$ and compute the estimates of the center and spread of the fuzzy coefficients $\tilde{\beta}_j(u)$ $(j = 1, 2, \ldots, p)$ at $u = u_i$ according to the procedure described above. Let $\hat{\beta}_j^{(-i)}(u_i, h)$ and $\hat{\sigma}_j^{(-i)}(u_i, h)$ $(j = 1, 2, \ldots, p)$ be the resulting estimates of the centers and spreads of the fuzzy coefficients under h and

$$\hat{y}_{(-i)}(h) = \sum_{j=1}^{p}\hat{\beta}_j^{(-i)}(u_i, h)x_{ij} \quad \text{and} \quad \hat{s}_{(-i)}(h) = \sum_{j=1}^{p}\hat{\sigma}_j^{(-i)}(u_i, h)x_{ij}$$

Table 1
The GDP data and the fitted values of $\widetilde{GDP}$ obtained by fitting the fuzzy varying coefficient model with the proposed method and by fitting its linear counterpart with the Xu and Li and the Coppi et al. methods.

Data				Fitted values of $\widetilde{GDP}$		
U	I	WP	$\widetilde{GDP}$	Our method	The Xu–Li method	The Coppi et al. method
1	137.0	102.5	$(124.5,\ 6.23)_T$	$(124.461,\ 6.651)_T$	$(123.415,\ 6.174)_T$	$(123.415,\ 6.174)_T$
2	138.1	103.5	$(129.4,\ 6.47)_T$	$(129.006,\ 6.787)_T$	$(127.585,\ 6.382)_T$	$(127.585,\ 6.383)_T$
3	141.5	104.8	$(135.1,\ 6.76)_T$	$(135.836,\ 7.055)_T$	$(135.922,\ 6.800)_T$	$(135.921,\ 6.799)_T$
4	144.8	106.1	$(142.3,\ 7.12)_T$	$(142.815,\ 7.338)_T$	$(144.110,\ 7.209)_T$	$(144.110,\ 7.209)_T$
5	148.1	107.5	$(150.1,\ 7.51)_T$	$(150.529,\ 7.642)_T$	$(152.553,\ 7.632)_T$	$(152.553,\ 7.631)_T$
6	146.0	108.7	$(154.3,\ 7.72)_T$	$(152.942,\ 7.684)_T$	$(152.496,\ 7.627)_T$	$(152.496,\ 7.628)_T$
7	148.8	109.5	$(159.2,\ 7.96)_T$	$(158.325,\ 7.996)_T$	$(158.673,\ 7.937)_T$	$(158.673,\ 7.936)_T$
8	151.0	110.7	$(164.0,\ 8.20)_T$	$(164.660,\ 8.291)_T$	$(164.979,\ 8.252)_T$	$(164.979,\ 8.252)_T$
9	151.4	112.5	$(167.9,\ 8.40)_T$	$(169.839,\ 8.498)_T$	$(170.147,\ 8.509)_T$	$(170.147,\ 8.510)_T$
10	153.7	113.2	$(174.4,\ 8.72)_T$	$(175.039,\ 8.821)_T$	$(175.331,\ 8.769)_T$	$(175.330,\ 8.769)_T$
11	155.6	114.0	$(182.1,\ 9.11)_T$	$(179.978,\ 9.127)_T$	$(180.176,\ 9.011)_T$	$(180.176,\ 9.012)_T$
12	158.8	114.9	$(187.4,\ 9.37)_T$	$(187.534,\ 9.514)_T$	$(187.200,\ 9.363)_T$	$(187.199,\ 9.362)_T$
13	161.9	116.0	$(195.2,\ 9.76)_T$	$(195.396,\ 9.894)_T$	$(194.584,\ 9.732)_T$	$(194.583,\ 9.732)_T$
14	166.1	118.0	$(207.3,\ 10.37)_T$	$(206.920,\ 10.336)_T$	$(205.883,\ 10.297)_T$	$(205.883,\ 10.296)_T$
15	169.1	120.3	$(217.3,\ 10.87)_T$	$(216.753,\ 10.693)_T$	$(216.170,\ 10.811)_T$	$(216.170,\ 10.811)_T$
16	173.0	122.6	$(228.3,\ 11.42)_T$	$(228.145,\ 11.091)_T$	$(227.789,\ 11.392)_T$	$(227.788,\ 11.391)_T$
17	176.0	125.0	$(237.0,\ 11.85)_T$	$(238.141,\ 11.414)_T$	$(238.330,\ 11.918)_T$	$(238.330,\ 11.918)_T$
18	174.8	126.3	$(239.4,\ 11.97)_T$	$(238.997,\ 11.444)_T$	$(239.859,\ 11.994)_T$	$(239.859,\ 11.995)_T$

be the predicted values of the center and spread of the fuzzy response $\tilde{Y}$ at u_i. The CV score is formulated by using the distance (6) as

$$\begin{aligned} \mathrm{CV}(h) &= \sum_{i=1}^{n} d^2\left((y_i, s_i)_L, (\hat{y}_{(-i)}(h), \hat{s}_{(-i)}(h))_L\right) \\ &= \sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p}\hat{\beta}_j^{(-i)}(u_i, h)x_{ij}\right)^2 + \triangle(L,f)\sum_{i=1}^{n}\left(s_i - \sum_{j=1}^{p}\hat{\sigma}_j^{(-i)}(u_i, h)x_{ij}\right)^2. \end{aligned} \tag{24}$$

Then, select h_0 as the optimal value of the smoothing parameter such that

$$\mathrm{CV}(h_0) = \min_{h>0}\mathrm{CV}(h). \tag{25}$$

3. Examples on application of the proposed method

Three real-world datasets are analyzed in this section in order to illustrate some applications of the proposed fuzzy varying coefficient model with its estimation method. Furthermore, we compare the results with those obtained by some other methods for fitting the corresponding fuzzy linear model in order to evaluate the performance of the proposed method.

In order to make a comprehensive comparison, we first introduce a metric for evaluating the accuracy of a method in predicting the fuzzy response as

$$\begin{aligned} \mathrm{GOF} &= \left(\frac{1}{n}\sum_{i=1}^{n} d^2((y_i, s_i)_L, (\hat{y}_i, \hat{s}_i)_L)\right)^{\frac{1}{2}} \\ &= \left(\frac{1}{n}\sum_{i=1}^{n}\left((y_i - \hat{y}_i)^2 + \triangle(L,f)(s_i - \hat{s}_i)^2\right)\right)^{\frac{1}{2}}, \end{aligned} \tag{26}$$

where $\hat{y}_i$ and $\hat{s}_i$ are the fitted values of the center and spread of the fuzzy response $\tilde{Y}$ at the ith designed point with a given method. This quantity gives an overall measurement of the difference between the observed values of the fuzzy response $\tilde{Y}$ and its corresponding estimates.

Example 1. The dataset is taken from [35] and consists of 18 observations of the income (I), working population (WP) and gross domestic product ($\widetilde{GDP}$) of Japan from 1975 to 1992, in which the $\widetilde{GDP}$ is assumed to be a symmetric triangular fuzzy number. The data are shown in the second to fourth columns of Table 1 where we use the lower subscript "T" to denote a symmetric triangular fuzzy number and we refer the reader to [35] for the details for formulating the dataset.

Now, we take the variable U to be the time order of each year and consider the following fuzzy varying coefficient model:

$$\widetilde{GDP}_i = \tilde{\beta}_1(i) + \tilde{\beta}_2(i)I_i + \tilde{\beta}_3(i)WP_i, \quad i = 1, 2, \ldots, 18. \tag{27}$$

The Gaussian kernel (22) is used in the proposed method, and we set $f(\lambda) = \lambda$ in Xu's distance (6) which leads to $\Delta(L,f) = 1/6$. The optimal value of the smoothing parameter selected by the fuzzified cross-validation procedure is

Table 2
The carbon monoxide concentration data and the fitted values of $\widetilde{CO}$ obtained by fitting the fuzzy varying coefficient model with the proposed method and by fitting its linear counterpart with the Xu–Li and Coppi et al. methods.

Data					Fitted values of $\widetilde{CO}$		
U	RH	R	WS	$\widetilde{CO}$	Our method	The Xu–Li method	The Coppi et al. method
1	82.27	0.90	2.87	$(1.15,\ 0.70)_T$	$(1.228,\ 1.875)_T$	$(1.190,\ 1.139)_T$	$(1.194,\ 1.075)_T$
2	88.70	0.90	1.12	$(2.98,\ 2.36)_T$	$(3.126,\ 2.019)_T$	$(3.112,\ 1.939)_T$	$(3.119,\ 1.841)_T$
3	82.51	0.02	0.85	$(3.92,\ 1.97)_T$	$(3.783,\ 1.876)_T$	$(3.702,\ 2.044)_T$	$(3.701,\ 2.072)_T$
4	79.46	0.00	0.45	$(4.65,\ 2.07)_T$	$(4.017,\ 1.804)_T$	$(3.931,\ 2.093)_T$	$(3.927,\ 2.162)_T$
5	68.85	0.00	0.91	$(3.98,\ 2.13)_T$	$(3.089,\ 1.561)_T$	$(2.981,\ 1.609)_T$	$(2.970,\ 1.782)_T$
6	79.39	0.02	1.07	$(3.35,\ 1.97)_T$	$(3.402,\ 1.798)_T$	$(3.345,\ 1.872)_T$	$(3.342,\ 1.929)_T$
7	88.87	1.30	0.69	$(3.13,\ 1.99)_T$	$(3.236,\ 2.009)_T$	$(3.218,\ 2.001)_T$	$(3.226,\ 1.883)_T$
8	88.92	0.09	0.40	$(4.15,\ 2.41)_T$	$(4.400,\ 2.007)_T$	$(4.382,\ 2.379)_T$	$(4.385,\ 2.344)_T$
9	83.52	0.00	0.83	$(3.96,\ 2.27)_T$	$(3.801,\ 1.882)_T$	$(3.786,\ 2.086)_T$	$(3.785,\ 2.106)_T$
10	87.51	0.05	2.09	$(4.07,\ 2.18)_T$	$(2.813,\ 1.968)_T$	$(2.795,\ 1.763)_T$	$(2.798,\ 1.713)_T$
11	86.04	0.04	0.90	$(3.30,\ 1.98)_T$	$(3.816,\ 1.931)_T$	$(3.818,\ 2.130)_T$	$(3.819,\ 2.119)_T$
12	91.77	0.93	1.22	$(4.02,\ 2.55)_T$	$(3.183,\ 2.055)_T$	$(3.152,\ 1.992)_T$	$(3.161,\ 1.857)_T$
13	83.62	0.02	3.00	$(2.06,\ 2.12)_T$	$(1.799,\ 1.869)_T$	$(1.790,\ 1.337)_T$	$(1.791,\ 1.312)_T$
14	73.10	0.02	1.75	$(1.37,\ 0.66)_T$	$(2.357,\ 1.630)_T$	$(2.410,\ 1.445)_T$	$(2.402,\ 1.556)_T$
15	80.38	0.09	0.73	$(3.35,\ 1.73)_T$	$(3.597,\ 1.789)_T$	$(3.654,\ 2.004)_T$	$(3.651,\ 2.053)_T$
16	87.98	0.24	1.87	$(1.45,\ 1.05)_T$	$(2.877,\ 1.953)_T$	$(2.879,\ 1.810)_T$	$(2.883,\ 1.747)_T$
17	90.13	0.02	2.39	$(2.74,\ 1.41)_T$	$(2.678,\ 1.997)_T$	$(2.673,\ 1.747)_T$	$(2.678,\ 1.665)_T$
18	64.95	0.00	1.25	$(2.44,\ 1.62)_T$	$(2.311,\ 1.436)_T$	$(2.476,\ 1.373)_T$	$(2.462,\ 1.579)_T$
19	80.16	0.01	1.02	$(2.79,\ 1.86)_T$	$(3.335,\ 1.769)_T$	$(3.437,\ 1.916)_T$	$(3.434,\ 1.966)_T$
20	86.14	0.44	0.70	$(3.31,\ 1.44)_T$	$(3.657,\ 1.896)_T$	$(3.710,\ 2.111)_T$	$(3.712,\ 2.077)_T$
21	89.12	0.01	1.17	$(4.02,\ 2.46)_T$	$(3.688,\ 1.958)_T$	$(3.747,\ 2.139)_T$	$(3.750,\ 2.092)_T$

$h_0 = 3.2$. With the selected value of h, we calibrate model (27) with the restricted weighted least-squares procedure and report in Table 1 the fitted values of the response $\widetilde{GDP}$ at i ($i = 1, 2, \ldots, 18$) in the fifth column.

For comparison, the method in [11] and the method in [30] for fitting the linear counterpart of the model (27) are also used to analyze this dataset, and the fitted fuzzy outputs are listed in the sixth and seventh columns of Table 1.

Comparing the observed values of the response with its fitted values, we see that the proposed fuzzy varying coefficient model with its estimation procedure produces more accurate estimates of the response, especially for the center of the response. The quantity GOF is also calculated for the three methods; it is GOF $= 0.914$ for our method, GOF $= 1.385$ for the Xu–Li method and GOF $= 1.384$ for the Coppi et al. method. The smaller value of GOF for our method also demonstrates that the fuzzy varying coefficient model performs better than its linear counterpart in fitting this dataset.

Example 2. The dataset comes from [30] and consists of 21 observations of the atmospheric concentration of carbon monoxide (CO) and a set of meteorological variables collected in Rome, Italy, in the period from October 5 to 25, 1992. The concentration of CO was observed hourly for each of the 21 days and the meteorological variables were recorded daily in the same period. From the analysis in paper [30], three crisp meteorological variables, that is, relative humidity (RH), rain (R) and wind speed (WS), have a significant influence on the fuzzy response variable $\widetilde{CO}$ of the concentration of CO. Here, we take these three variables as the explanatory variables in our analysis.

In paper [30], however, the response $\widetilde{CO}$ was assumed to be a nonsymmetric triangular fuzzy number whose center was defined as the mean of the observations recorded each day, and the left and right spreads were the averaged deviations of the values that are smaller than the mean and those that are larger than the mean respectively. For the purpose of using the proposed fuzzy varying coefficient model, we redefine the response $\widetilde{CO}$ as a symmetric triangular fuzzy variable with its center being the same as that in paper [30] and its spread being the whole of the averaged absolute deviations of all observations in each day. The observations of the newly defined response can be computed from the data of the left spreads l_i and right spreads r_i of $\widetilde{CO}$ reported in Table 2 of paper [30] as follows. Firstly, let n_{i1} and n_{i2} be respectively the number of the observations of CO that are smaller than the mean and the number of the observations of CO that are larger than the mean on the ith day. Then, according to the definition of l_i and r_i in [30], we have

$$n_{i1}l_i = n_{i2}r_i, \quad i = 1, 2, \ldots, 21.$$

Therefore, the observations of the spread of the newly defined response $\widetilde{CO}$ are

$$s_i = \frac{n_{i1}l_i + n_{i2}r_i}{n_{i1} + n_{i2}} = \frac{2n_{i2}r_i}{n_{i2}r_i/l_i + n_{i2}} = \frac{2r_il_i}{r_i + l_i}, \quad i = 1, 2, \ldots, 21.$$

The observations of the explanatory variables and the newly defined response are shown in Table 2.

We still take the time orders from October 5 to 25 as the observations of the variable U in the fuzzy coefficients to build a fuzzy varying coefficient model between $\widetilde{CO}$ and the variables RH, R and WS as

$$\widetilde{CO}_i = \tilde{\beta}_1(i) + \tilde{\beta}_2(i)RH_i + \tilde{\beta}_3(i)R_i + \tilde{\beta}_4(i)WS_i, \quad i = 1, 2, \ldots, 21. \tag{28}$$

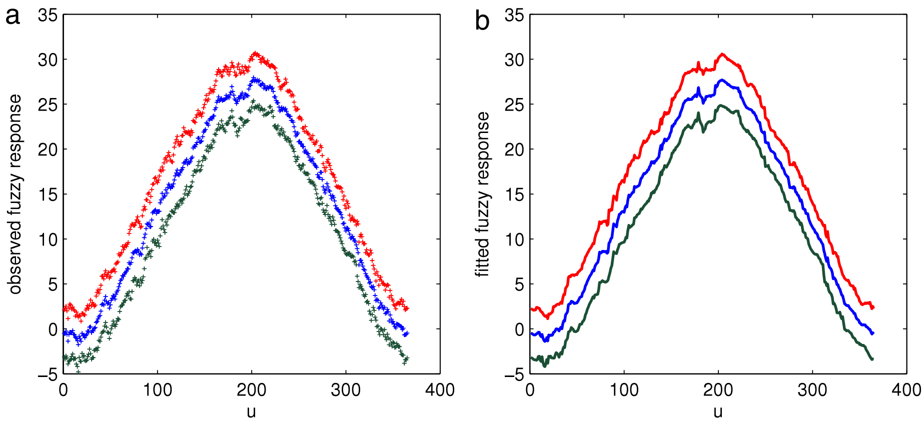

Fig. 1. (a) The center, lower and upper limits of the observed values of $\widetilde{AT}$ and (b) the fitted corresponding lines of $\widetilde{AT}$ obtained with our method.

With use of the Gaussian kernel and $f(\lambda) = \lambda$ in the distance (6), the selected optimal value of the smoothing parameter is $h_0 = 12.7$. The fitted values of the response $\widetilde{CO}$ calculated by the proposed method are listed in the sixth column of Table 2. Furthermore, we also listed in the seventh and eighth columns of Table 2 the fitted values of the fuzzy response obtained by fitting the corresponding fuzzy linear model with the Xu–Li method [11] and the Coppi et al. method [30].

From the results, we can see that the fitted values of $\widetilde{CO}$ obtained by the proposed method and those obtained by the Xu–Li method and the Coppi et al. method are all comparable. The values of GOF for the three methods are 0.634, 0.651 and 0.652 respectively, which have little difference among them.

The above results also indicate that a fuzzy linear model may be appropriate for the dataset. From the inferential point of view, because the fuzzy varying coefficient model is more flexible than its linear counterpart in tracking the trend of the response, it can also be used as an exploratory tool for judging whether a fuzzy linear model is appropriate to a given dataset by fitting a fuzzy varying coefficient model and its linear counterpart and comparing the fitting results. If there is no significant difference between the results, a fuzzy linear model may be chosen to fit the dataset.

Example 3. The dataset is graphically shown in Fig. 1(a) which consists of 365 observations of the average temperature ($\widetilde{AT}$) and sunlight time (ST) for each day of a year in Xi'an, China from January 1 to December 31. In this dataset, $\widetilde{AT}$ is assumed to be a triangular fuzzy variable. The observations of $\widetilde{AT}$ and ST are formulated as follows. With the observed data of the average temperature and sunlight time of each day from January 1, 1951, to December 31, 2000, the mean of the average temperatures collected on the same days of the 50 years is taken as the center value of $\widetilde{AT}$ on that day and the corresponding sample standard variance is used as the spread value of $\widetilde{AT}$. The value of the crisp variable ST is defined as the mean of the sunlight times on the same days of the 50 years. It is worth pointing out that the data on February 29 during the 50 years have been excluded.

Like in the above examples, the observations of the variable U in the regression coefficients are taken as the time orders from January 1 to December 31, and the following fuzzy varying coefficient model connecting $\widetilde{AT}$ and ST is considered:

$$\widetilde{AT}_i = \tilde{\beta}_1(i) + \tilde{\beta}_2(i)ST_i, \quad i = 1, 2, \dots, 365. \tag{29}$$

The proposed method with the Gaussian kernel (22) and $f(\lambda) = \lambda$ for the distance (6) is used to fit model (29). The optimal value of the smoothing parameter selected by the fuzzified cross-validation method is $h_0 = 2$. The fitted values of the center, lower and upper limits of the fuzzy response are shown in Fig. 1(b) and the value of GOF is 0.230, which indicates that the fuzzy varying coefficient model with its estimation method can produce a satisfactory fit of the fuzzy response.

Additionally, the linear counterpart of the model (29) is also used to fit the dataset with the methods in [11,30]. The resulting values of GOF are 8.430 and 8.429 respectively, which are much larger than that obtained by the proposed method. In order to graphically show the differences in fitting the response with the three methods, we calculate, for each of the methods, Xu's distance in (6) with $f(\lambda) = \lambda$ between the observed value and the fitted value of $\widetilde{AT}$ on each day, that is

$$d_i = \left((y_i - \hat{y}_i)^2 + \frac{1}{6}(s_i - \hat{s}_i)^2 \right)^{\frac{1}{2}}, \quad i = 1, 2, \dots, 365,$$

where $\hat{y}_i$ and $\hat{s}_i$ are the fitted center and spread of the response with one of the three methods. We draw the scatter plots of each d_i obtained by our method versus those obtained by the Xu–Li method and the Coppi et al. method respectively in Fig. 2(a) and (b). Note that the scales for the two axes are very different. The scatter plots show that the proposed fuzzy varying coefficient model is still powerful in fitting some complicated regression relationships, but the corresponding fuzzy linear model may completely fail to fit this dataset.

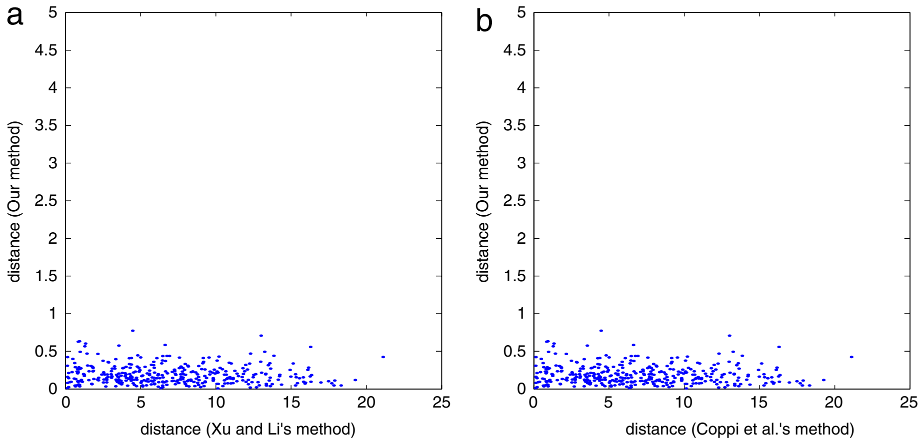

Fig. 2. The scatter plots of the distances between the observed and fitted values of $\widetilde{AT}$ obtained by the three methods.

4. Final remarks

In this paper, a fuzzy varying coefficient model with crisp explanatory variables and a symmetric fuzzy response is proposed and a restricted weighted least-squares procedure is suggested for fitting the model. By allowing the fuzzy regression coefficients in the fuzzy linear model to vary with a covariate, the model flexibility and adaptability can be greatly increased. Furthermore, this model can effectively avoid the problem of the curse of dimensionality encountered in the fuzzy nonparametric regression models. The analysis of some practical datasets has demonstrated that the proposed model with its estimation approach performs quite satisfactorily in predicting the fuzzy response even in the case where the regression relationship seriously deviates from the fuzzy linear model.

Acknowledgement

This research was supported by the National Natural Science Foundation of China (10971161).

References

[1] L.A. Zadeh, Fuzzy sets, Information and Control 8 (1965) 338–353.
[2] H. Tanaka, S. Uejima, K. Asai, Linear regression analysis with fuzzy model, IEEE Transactions on Systems, Man and Cybernetics 12 (1982) 903–907.
[3] H. Tanaka, Fuzzy data analysis by possibilistic linear models, Fuzzy Sets and Systems 24 (1987) 363–375.
[4] H. Tanaka, J. Watada, Possibilistic linear systems and their application to the linear regression model, Fuzzy Sets and Systems 27 (1988) 275–289.
[5] D.A. Savic, W. Pedrycz, Evaluation of fuzzy linear regression models, Fuzzy Sets and Systems 39 (1991) 51–63.
[6] G. Peters, Fuzzy linear regression with fuzzy intervals, Fuzzy Sets and Systems 63 (1994) 45–55.
[7] P. Diamond, Fuzzy least squares, Information Sciences 46 (1988) 141–157.
[8] K.J. Kim, H. Moskowitz, M. Koksalan, Fuzzy versus statistical linear regression, European Journal of Operational Research 92 (1996) 417–434.
[9] M. Ma, M. Friedman, A. Kandel, General fuzzy least squares, Fuzzy Sets and Systems 88 (1997) 107–118.
[10] R.N. Xu, S-curve regression model in fuzzy environment, Fuzzy Sets and Systems 90 (1997) 317–326.
[11] R.N. Xu, C.L. Li, Multidimensional least-squares fitting with a fuzzy model, Fuzzy Sets and Systems 119 (2001) 215–223.
[12] C. Kao, C.L. Chyu, Least squares estimates in fuzzy regression analysis, European Journal of Operational Research 148 (2003) 426–435.
[13] M.S. Yang, H.H. Liu, Fuzzy least-squares algorithms for interactive fuzzy linear regression models, Fuzzy Sets and Systems 135 (2003) 305–316.
[14] G. González-Rodríguez, Á. Blanco, A. Colubi, M.A. Lubiano, Estimation of a simple linear regression model for fuzzy random variables, Fuzzy Sets and Systems 160 (2009) 357–370.
[15] H. Ishibuchi, H. Tanaka, H. Okada, An architecture of neural networks with interval weights and its application to fuzzy regression analysis, Fuzzy Sets and Systems 57 (1993) 27–39.
[16] C.B. Cheng, E.S. Lee, Fuzzy regression with radial basis function networks, Fuzzy Sets and Systems 119 (2001) 291–301.
[17] D. Zhang, L.F. Deng, K.Y. Cai, Fuzzy nonlinear regression with fuzzified radial basis function network, IEEE Transactions on Fuzzy Systems 13 (2005) 742–760.
[18] B. Ye, C.Z. Zhu, Y.J. Cao, Adaptive extended fuzzy basis function network, Neural Computing and Applications 16 (2007) 197–206.
[19] M.H.F. Zarandi, I.B. Türksen, J. Sobhani, A.A. Ramezanianpour, Fuzzy polynomial neural networks for approximation of the compressive strength of concrete, Applied Soft Computing 8 (2008) 488–498.
[20] D.H. Hong, C. Hwang, Support vector fuzzy regression machines, Fuzzy Sets and Systems 138 (2003) 271–281.
[21] D.H. Hong, C. Hwang, Extended fuzzy regression models using regularization method, Information Sciences 164 (2004) 31–46.
[22] D.H. Hong, C. Hwang, Ridge regression procedures for fuzzy models using triangular fuzzy numbers, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 12 (2004) 145–159.
[23] D.H. Hong, C. Hwang, C. Ahn, Ridge estimation for regression models with crisp inputs and Gaussian fuzzy output, Fuzzy Sets and Systems 142 (2004) 307–319.
[24] X. Jin, Z.D. Liu, R.F. Bie, G.X. Zhao, J.X. Ma, Support vector machines for regression and applications to software quality prediction, Lecture Notes in Computer Science 3944 (2006) 781–788.
[25] H. Dong, S.Y. Yang, D.H. Wu, Intelligent prediction method for small-batch producing quality based on fuzzy least squares SVM, Systems Engineering—Theory and Practice 27 (2007) 98–104.

[26] R. Khemchandani, Jayadeva, S. Chandra, Regularized least-squares fuzzy support vector regression for financial time series forecasting, Expert Systems with Applications 36 (2009) 132–138.
[27] C.B. Cheng, E.S. Lee, Nonparametric fuzzy regression—*k*-NN and kernel smoothing techniques, Computers and Mathematics with Applications 38 (1999) 239–251.
[28] N. Wang, W.X. Zhang, C.L. Mei, Fuzzy nonparametric regression based on local linear smoothing technique, Information Sciences 177 (2007) 3882–3900.
[29] P. D'Urso, Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data, Computational Statistics and Data Analysis 42 (2003) 47–72.
[30] R. Coppi, P. D'Urso, A. Giordani, A. Santoro, Least squares estimation of a linear regression model with LR fuzzy response, Computational Statistics and Data Analysis 51 (2006) 267–286.
[31] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning—I, Information Sciences 8 (1975) 199–249.
[32] R.N. Xu, A linear regression model in fuzzy environment, Advance in Modelling Simulation 27 (1991) 31–40.
[33] M.S. Waterman, A restricted least squares problem, Technometrics 16 (1974) 135–136.
[34] J.D. Hart, Nonparametric Smoothing and Lack-of-Fit Tests, Springer-Verlag, New York, 1997.
[35] H. Lee, H. Tanaka, Fuzzy approximations with non-symmetric fuzzy parameters in fuzzy regression analysis, Journal of the Operations Research Society of Japan 42 (1999) 98–112.