Iranian Journal of Fuzzy Systems

Volume 19, Number 2, (2022), pp. 63-81

Original paper

University of Sistan and Baluchestan

Fuzzy time series model using weighted least square estimation

G. Hesamian[1] and M. G. Akbari[2]

[1] *Department of Statistics, Payame Noor University Tehran 19395-3697, Iran.*
[2] *Department of Statistics, University of Birjand, Birjand, Iran.*

gh.hesamian@pnu.ac.ir, g_z_akbari@birjand.ac.ir

Abstract

The conventional fuzzy least-squares time series models show undesirable performance when the fuzzy data set involves the outliers. By introducing a strategy to detect the outliers, this paper introduced a method for reducing the influence of outliers on the future predictions. For this purpose, according to the weighted square distance error, an estimation procedure was suggested for determining the exact coefficients in the presence of outliers. The parameters of the fuzzy time series model were then estimated using an iterative algorithm. In order to identify the potential outliers of the fuzzy data, a quality control chart was employed based on the center of gravity criterion of fuzzy data. The defuzzification method was also employed to examine the performance of the proposed method via some scatter plots. Several common goodness-of-fit criteria used in traditional time series models were also extended to compare the performance of the proposed fuzzy time series method to an existing method. The effectiveness of the proposed method was illustrated through two numerical examples including a simulation study. The results clearly indicated that the proposed model performs well in terms of the both scatter plot criteria and goodness-of-fit evaluations in cases where the potential outliers exist among the fuzzy data.

Keywords: Fuzzy semi-parametric time series, weight, kernel method, optimal bandwidth, outliers, control chart, goodness-of-fit measure.

1 Introduction

The time series prediction method has found a wide range of applications in business, finance, computer science, engineering, medicine, physics, chemistry, and many other interdisciplinary fields. Statistical time series methods focus on a variety of (exact) data sources. In the real world, however, many concepts can not be expressed by precise values. A wide range of data might be vague and imprecise (fuzzy) either by their nature, e.g. environmental data or due to non-ideal measuring [33]. Many of these data and concepts (such as monthly energy consumption, weakly Co_2 emission constraints or annual global land-ocean temperature) can be expressed as sharp numbers or clear expressions of natural language. For such cases, fuzzy time series models based on imprecise information have gained considerable attention in the past decades. Many researchers have studied the conventional time series models for exact data to predict the future based on imprecise past relationships occurring in real-life applications. Soft computing techniques used in this context are mainly a combination of fuzzy sets, artificial neural networks, rough sets, and evolutionary computation. Such approaches have been widely employed in real-life applications like enrollment, stock index prices, temperature, financial prediction, and electricity load (for a comprehensive review on these methods, see [1, 13, 25, 29, 30, 37, 41, 43, 46]). Concerning non-fuzzy time series data, these methods first provide some fuzzy time series prediction values followed by their defuzzification to exact prediction values as the final goal. Other methods relied on fuzzy quantity time series data and fuzzy prediction. In these methods, it was assumed that the available time series data can be reported by fuzzy data rather than exact values. The goal is to predict the future as fuzzy quantities. In this regard, Tseng and Tzeng [59] proposed a fuzzy seasonal ARIMA for non-fuzzy data in the cases where the future situations are predicted as fuzzy

Corresponding Author: G. Hesamian
Received: January 2021; Revised: June 2021; Accepted: August 2021.
https://doi.org/10.22111/IJFS.2022.6788

values. They suggested a fuzzy confidence region for future prediction. Moreover, Hesamian and Akbari [31] proposed a statistical time series model based on fuzzy data. They suggested a semi-parametric time series model with fuzzy data, non-fuzzy coefficients, and fuzzy smooth functions. Zarei et al. [65] applied a specific version of the Hesamian and Akbari model for triangular fuzzy data and the different distance measures for fuzzy data.

Noteworthy, as discussed above, the traditional time series models fail to address forecasting problems with vague or ambiguous observations represented by fuzzy data. Such a shortcoming could be overcome by the fuzzy-based time series models. It should be mentioned that in the method of Hesamian and Akbari, a semiparametric time series model was proposed with fuzzy data. This method relies on the least square error technique that is frequently used in linear regression analysis. However, in the presence of outliers, the least-square methods may offer poor predictions. In such cases, the robust methods could present a higher degree of accuracy compared to the least-square technique, as they use a weighting mechanism to weigh down the outliers [38, 61]. In this paper, the unknown nonparametric part of the model was first improved using a novel generalized subtraction criterion. Such an arithmetic operator was then employed to solve a fuzzy equation in connection with the fuzzy nonparametric part of the model. The advantage of the method of Hesamian and Akbari is that it uses a modified general subtraction of two fuzzy numbers. Unlike Hesamian and Akbari's method, the other advantage of such an operator is that it always exists and is a fuzzy number. Then, a simple weighted least-square target was suggested to reduce the effect of outliers on the prediction. The performance of the improved method was then examined in terms of some common goodness-of-fit criteria. For practical reasons, the proposed method was further evaluated through a simulated study. The numerical results clearly indicated that the proposed method provides sufficiently accurate results compared to other methods especially in the presence of outliers in the data set.

The rest of this paper is organized as follows: Section 2 reviews some concepts including α-values of fuzzy numbers and fuzzy time series. In Section 3, a robust estimation is suggested for a fuzzy time series model with exact coefficient and fuzzy smooth function. Section 4 illustrates two numerical examples to evaluate the effectiveness and performance of the proposed robust method in terms of some common performance measures. Finally, the major contributions of this paper are summarized in Section 5.

2 Fuzzy numbers

This section reviews some basic definitions of fuzzy numbers which are used in the proposed method.
A fuzzy set $\widetilde{A}$ of $\mathbb{R}$ (the real line) is defined by its membership function $\widetilde{A} : \mathbb{R} \to [0,1]$. For each $\alpha \in (0,1]$, the subset $\{x \in \mathbb{X} \mid \widetilde{A}(x) \geq \alpha\}$ is called the α-cut of $\widetilde{A}$ and is denoted by $\widetilde{A}[\alpha]$. In addition, a fuzzy set $\widetilde{A}$ of $\mathbb{R}$ is called a fuzzy number (**FN**) if 1- there is a unique $x^*_{\widetilde{A}} \in \mathbb{R}$ so that $\widetilde{A}[1] = \{x^*_{\widetilde{A}}\}$ and 2- $\{\widetilde{A}[\alpha]\}_{\alpha\in[0,1]}$ are a sequence of nested closed intervals in $\mathbb{R}$. This interval is denoted by $\widetilde{A}[\alpha] = [\widetilde{A}^{L[\alpha]}, \widetilde{A}^{U[\alpha]}]$, where $\widetilde{A}^{L}[\alpha] = \inf\{x : x \in \widetilde{A}[\alpha]\}$ and $\widetilde{A}^{U}[\alpha] = \sup\{x : x \in \widetilde{A}[\alpha]\}$. The set of all fuzzy numbers is denoted by $\mathbb{F}(\mathbb{R})$. It is worth noting that fuzzy numbers are the approximate assessments, given by experts and accepted by decision-makers when the access to more accurate values is either impossible or unnecessary. To simplify representation and handling fuzzy numbers, several authors have captured the information contained in a (unimodal) fuzzy number using a functional parametric form known as LR-fuzzy number $\widetilde{A} = (a; l_a, r_a)_{LR}$ [16]. The membership function of a LR-fuzzy number $\widetilde{A}$ is defined as:

$$\widetilde{A}(x) = \begin{cases} L\big(\frac{a-x}{l_a}\big), & x \leq a, \\ R\big(\frac{x-a}{r_a}\big), & x > a, \end{cases} \tag{1}$$

where $a \in \mathbb{R}$, $l_a > 0$ and $r_a > 0$ are called the mean value, left and right spreads of $\widetilde{A}$, respectively. The shape function L (or R) is a decreasing function for $\mathbb{R}^+ \to [0,1]$ such that

1. $L(1) = R(1) = 0$,
2. $L(0) = R(0) = 1$, and
3. $L(x)$ and $R(x)$ are continuous and monotone decreasing functions on $[0,1]$.

An LR-number has been applied in various problems as a general model function of imprecision. In this paper, we employed the most commonly used LR-fuzzy numbers (with $L(x) = R(x) = \max\{0, 1-x\}$) called triangular fuzzy numbers (**TFN**s), to handle the imprecision in data set during the numerical evaluation. The membership function of

a triangular fuzzy number, denoted by $\widetilde{A} = (a; l_a, r_a)_T$, is given by:

$$\widetilde{A}(x) = \begin{cases} \frac{x-(a-l_a)}{l_a}, & a - l_a \leq x \leq a, \\ \frac{a+r_a-x}{r_a}, & a < x \leq a + r_a, \\ 0, & x \in \mathbb{R} - (a - l_a, a + r_a). \end{cases} \tag{2}$$

Definition 2.1. [32] *For a given $\widetilde{A} \in \mathbb{F}(\mathbb{R})$, the mapping $\widetilde{A}_\alpha : [0,1] \to \mathbb{R}$ is called the α-values of $\widetilde{A}$ and defined as follows:*

$$\widetilde{A}_\alpha = \begin{cases} \widetilde{A}^L[2\alpha], & \alpha \in [0, 0.5], \\ \widetilde{A}^U[2(1-\alpha)], & \alpha \in (0.5, 1], \end{cases} \tag{3}$$

where $\widetilde{A}^L[\alpha]$ and $\widetilde{A}^U[\alpha]$ denote the lower and upper limits of α-cuts of $\widetilde{A}$.

Example 2.2. *If $\widetilde{A} = (a; l_a, r_a)_T$ is a triangular fuzzy number, then:*

$$\widetilde{A}_\alpha = \begin{cases} (a - l_a) + 2l_a\alpha, & 0.0 \leq \alpha \leq 0.50, \\ a + r_a - 2r_a(1-\alpha), & 0.50 < \alpha \leq 1.0. \end{cases}$$

Notation 2.3. *Since $\widetilde{A}_\alpha$ is a decreasing function of α, the relationship between α-values and α-cuts can be expressed as:*

$$\widetilde{A}[\alpha] = [\widetilde{A}^L[\alpha], \widetilde{A}^U[\alpha]] = [\widetilde{A}_{\alpha/2}, \widetilde{A}_{1-\alpha/2}]. \tag{4}$$

Therefore, having a sequence of α-values $\{\widetilde{A}_\alpha\}_{\alpha\in[0,1]}$, the membership function of $\widetilde{A}$ can be evaluated as follows:

$$\widetilde{A}(x) = \sup\{\alpha \in [0,1] : x \in [\widetilde{A}_{\alpha/2}, \widetilde{A}_{1-\alpha/2}]\}, \ x \in \mathbb{R}.$$

In addition, for all $\widetilde{A}, \widetilde{B} \in \mathbb{F}(\mathbb{R})$, $\lambda \in \mathbb{R}$ and $\alpha \in [0,1]$, the addition and scalar multiplication operations between $\widetilde{A}$ and $\widetilde{B}$ can be evaluated as follows:

$$(\widetilde{A} \oplus \widetilde{B})_\alpha = \widetilde{A}_\alpha + \widetilde{B}_\alpha,$$
$$(\lambda \otimes \widetilde{A})_\alpha = \lambda\widetilde{A}_\alpha I(\lambda \geq 0) + \lambda\widetilde{A}_{1-\alpha} I(\lambda < 0),$$

where $I(.)$ is the indicator function. Such arithmetic operations will be applied to suggest a fuzzy multivariate regression model in the next section.

It should be noted that in the fuzzy arithmetic, the conventional addition and subtraction operators are not invertible operations, that is, for any nonconstant fuzzy numbers $\widetilde{A}$, $\widetilde{B}$ and $\widetilde{C}$:

1. $\widetilde{A} = \widetilde{C} \oplus \widetilde{B}$ does not imply $\widetilde{A} \ominus \widetilde{B} = \widetilde{C}$.
2. $(\widetilde{A} \oplus \widetilde{B}) \ominus \widetilde{B} \neq \widetilde{A}$.

To overcome such drawback, this paper used a novel generalized subtraction that always exists. Here, we recall a recent well-established generalized subtraction operator introduced by [32].

Definition 2.4. *The generalized subtraction between two fuzzy numbers of $\widetilde{A}$ and $\widetilde{B}$ is defined as a **FN** $\widetilde{A} \ominus_G \widetilde{B}$ with α-values:*

$$(\widetilde{A} \ominus_G \widetilde{B})_\alpha = \begin{cases} \inf_{\beta\in[\alpha,1-\alpha]}(\widetilde{A}_\beta - \widetilde{B}_\beta), & 0.0 \leq \alpha \leq 0.50, \\ \sup_{\beta\in[1-\alpha,\alpha]}(\widetilde{A}_\beta - \widetilde{B}_\beta), & 0.50 \leq \alpha \leq 1.0. \end{cases} \tag{5}$$

Example 2.5. *For two **TFNs** $\widetilde{A} = (a; l_a, r_a)_T$ and $\widetilde{B} = (b; l_b, r_b)_T$, we have*

$$(\widetilde{A} \ominus_G \widetilde{B})_\alpha = \begin{cases} \inf_{\beta\in[\alpha,1-\alpha]}(a - (1-2\beta)l_a - b + (1-2\beta)l_b), & 0.0 \leq \alpha \leq 0.50, \\ \sup_{\beta\in[1-\alpha,\alpha]}(a - (1-2\beta)r_a - b + (1-2\beta)r_b), & 0.50 \leq \alpha \leq 1.0, \end{cases}$$

$$= \begin{cases} a - b + \inf_{\beta\in[\alpha,1-\alpha]}((l_b - l_a)(1-2\beta)), & 0.0 \leq \alpha \leq 0.50, \\ a - b + \sup_{\beta\in[1-\alpha,\alpha]}((r_b - r_a)(1-2\beta)), & 0.50 \leq \alpha \leq 1.0, \end{cases}$$

$$= \begin{cases} a-b-(l_b-l_a)I(l_b \geq l_a)(1-2\alpha)+(l_b-l_a)I(l_b < l_a)(1-2\alpha), & 0.0 \leq \alpha \leq 0.50, \\ a-b-(r_b-r_a)I(r_b \geq r_a)(1-2\alpha)+(r_b-r_a)I(r_b < r_a)(1-2\alpha), & 0.50 \leq \alpha \leq 1.0, \end{cases}$$

$$= \begin{cases} a-b-(1-2\alpha)|l_a-l_b|, & 0.0 \leq \alpha \leq 0.50, \\ a-b-(1-2\alpha)|r_a-r_b|, & 0.50 \leq \alpha \leq 1.0, \end{cases}$$

which is the α-values of $(a-b;|l_a-l_b|,|r_a-r_b|)_T$. Therefore, $\widetilde{A} \ominus_G \widetilde{B} = (a-b;|l_a-l_b|,|r_a-r_b|)_T$.

Lemma 2.6. [32] *For three **FN**s of $\widetilde{A}$, $\widetilde{B}$ and $\widetilde{C}$, $\ominus_G$ meets the following properties:*

1) $\widetilde{A} \ominus_G \widetilde{A} = \widetilde{0}$ *where* $\widetilde{0}(x) = \begin{cases} 1 & x = 0, \\ 0 & x \neq 0. \end{cases}$

2) $\widetilde{A} \ominus_G \widetilde{B} = (-1) \otimes (\widetilde{B} \ominus_G \widetilde{A}) = ((-1) \otimes \widetilde{B}) \ominus_G ((-1) \otimes \widetilde{A})$.

3) $(\widetilde{A} \oplus \widetilde{B}) \ominus_G \widetilde{B} = \widetilde{A}$.

4) $\widetilde{A} \ominus_G \widetilde{B} = \widetilde{B} \ominus_G \widetilde{A}$ *iff* $\widetilde{A} = \widetilde{B}$.

5) $(\widetilde{A} \oplus \widetilde{C}) \ominus_G (\widetilde{B} \oplus \widetilde{C}) = \widetilde{A} \ominus_G \widetilde{B}$.

6) $I\{0\} \ominus_G (\widetilde{A} \ominus \widetilde{B}) = \widetilde{B} \ominus_G \widetilde{A}$.

7) $k \otimes (\widetilde{A} \ominus_G \widetilde{B}) = (k \otimes \widetilde{A}) \ominus_G (k \otimes \widetilde{B})$, *for* $k \in \mathbb{R}$.

Remark 2.7. *It should be noted that Akbari and Hesamian employed a generalized Hukuhara subtraction between $\widetilde{A} = (a; l_a, r_a)_{LR}$ and $\widetilde{B} = (b; l_b, r_b)_{LR}$ as*

$$\widetilde{A} \ominus_H \widetilde{B} = (a-b; \max\{l_a-l_b, 0\}, \max\{r_a-r_b, 0\})_{LR}. \tag{6}$$

The main advantage of $\ominus_G$ over $\ominus_H$ is that $\widetilde{A} \ominus_G \widetilde{B}$ is always an LR-fuzzy number while $\widetilde{A} \ominus_H \widetilde{B}$ may be evaluated as a non-fuzzy number $a-b$. Therefore, the possible information on the left and right spreads of $\widetilde{A} \ominus_H \widetilde{B}$ can not be involved in a fuzzy statistical model. This is why we used $\ominus_G$ instead of $\ominus_H$ in this paper.

Definition 2.8. *An L_p distance measure between two fuzzy numbers **FN**s $\widetilde{A}$ and $\widetilde{B}$ is defined as*

$$d_p(\widetilde{A}, \widetilde{B}) = \Big(\int_0^1 f(\alpha) |\widetilde{A}_\alpha - \widetilde{B}_\alpha|^p d\alpha \Big)^{1/p}, \tag{7}$$

where

$$f(\alpha) = \begin{cases} 4\alpha & 0 \leq \alpha \leq 0.5, \\ 4(1-\alpha) & 0.5 \leq \alpha \leq 1. \end{cases}$$

Any three **FN**s, $\widetilde{A}$, $\widetilde{B}$ and $\widetilde{C}$, satisfy the following conditions:

- $d_p(\widetilde{A}, \widetilde{B}) = 0$ if and only if $\widetilde{A} = \widetilde{B}$,
- $d_p(\widetilde{A}, \widetilde{B}) = d_p(\widetilde{B}, \widetilde{A})$,
- $d_p(\widetilde{A}, \widetilde{C}) \leq \big(d_p(\widetilde{A}, \widetilde{B}) + d_p(\widetilde{B}, \widetilde{C})\big)$.

It should be noted that $f(\alpha)$ also modifies the distance between the two **FN**s of $\widetilde{A}$ and $\widetilde{B}$, since it focuses on the values near the centers rather than the tails. We apply the distance measure d in order to estimate the unknown regression coefficients based on a weighted target in the next section. In a special case where $\widetilde{A} = (a; l_a, r_a)_T$ and $\widetilde{B} = (b; l_b, r_b)_T$ are two **TFN**s, one finds that

$$d_2^2(\widetilde{A}, \widetilde{B}) = 0.25(l_a-l_b)^2 + 0.667(a-b-l_a+l_b)(l_a-l_b) + 0.5(a-b-l_a+l_b)^2 + \\ 0.916(r_a-r_b)^2 + 1.334(a-b-l_a+l_b)(r_a-r_b) + 0.5(a-b-r_a+r_b)^2.$$

Furthermore, a common defuzzification method known as "center of gravity" is employed to examine the performance of the proposed weighed fuzzy time series model using some common scatter plots. For a **TFN** of $\widetilde{A} = (a; l_a, r_a)_T$, the center of gravity of $\widetilde{A}$ is evaluated as [31]:

$$M_{\widetilde{A}} = a + (r_a - l_a)/3. \tag{8}$$

3 A weighted least square error for fuzzy time series data

Weighted regression approach is one of the common techniques to reduce the effect of outlier(s) on the prediction of the future values. In this section, a weighted fuzzy time series model with exact coefficients, fuzzy data and fuzzy smooth function is suggested using a robust methodology in case where the outliers exist. First, we say that $\widetilde{X}_1, \widetilde{X}_2, \ldots, \widetilde{X}_T$ is a fuzzy time series if for every $\alpha \in [0,1]$, $(\widetilde{X}_1)_\alpha, (\widetilde{X}_2)_\alpha, \ldots, (\widetilde{X}_T)_\alpha$ is a sequence of ordinary random variables indexed by time t. A fuzzy semi-parametric time series model is defined as follows [31]:

$$\widetilde{X}_i = \oplus_{l=1}^{p}(\theta_l \otimes \widetilde{X}_{i-l}) \oplus \widetilde{f}(t_i) \oplus \widetilde{\epsilon}_i, \quad i = p+1, p+2, \ldots, T, \tag{9}$$

where

(I) $\widetilde{X}_{i-l} = (X_{i-l}; l_{X_{i-l}}, r_{X_{i-l}})_{LR} \otimes I(\theta_l \geq 0) \oplus (X_{i-l}; r_{X_{i-l}}, l_{X_{i-l}})_{RL} \otimes I(\theta_l < 0)$,

(II) $\widetilde{f}(t_i) = (f(t_i); l_{f(t_i)}, r_{f(t_i)})_{LR}$,

(III) θ_l showing unknown real-valued coefficients to be estimated,

(IV) $\widetilde{\epsilon}_i = (\epsilon_i; l_{\epsilon_i}, r_{\epsilon_i})_{LR}$ representing fuzzy errors,

(V) t_i indicate design points.

To estimate the nonparametric function $\widetilde{f}(.)$ and parametric coefficients $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)$, a two-stage procedure was introduced. First, by adding $\ominus_G(\oplus_{l=1}^{p}(\theta_l \otimes \widetilde{X}_{i-l}))$ on the both sides of Eq. (9), a fuzzy nonlinear model was obtained as follows:

$$\widetilde{X}_i \ominus_G (\oplus_{l=1}^{p}(\theta_l \otimes \widetilde{X}_{i-l})) = \widetilde{f}(t_i) \oplus \widetilde{\epsilon}_i, \tag{10}$$

where $\ominus_G$ denotes the proposed generalized subtraction between two fuzzy numbers. Therefore, the known fuzzy smooth function $\widetilde{f}$ can be estimated using a common kernel smoothing technique [6, 61] as follows:

$$\widetilde{\widehat{f}}_{\boldsymbol{\theta}}(t_i) = \oplus_{j=p+1}^{T} w_j(t_i; h) \otimes \big(\widetilde{X}_j \ominus_G (\oplus_{l=1}^{p}(\theta_l \otimes \widetilde{X}_{j-l}))\big), \tag{11}$$

where

$$w_j(t_i; h) = \frac{K\left(\dfrac{t_j - t_i}{h}\right)}{\sum_{j=p+1}^{T} K\left(\dfrac{t_j - t_i}{h}\right)}, \tag{12}$$

and $h > 0$ is the bandwidth of kernel $K(.)$ which satisfies the Mercer's conditions [19, 61]. Three popular kernels including Epanechnikov, triweight and Gaussian kernels (Table 1) were employed to examine their effects on the model performance and weight functions. At the second stage, $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)$ can be estimated using the least square error methods. However, as discussed in the introduction section, the least square methods may lead to poor predictions in the presence of outliers. Therefore, instead of the sum of square error technique used by Hesamian and Akbari, a weighted-variant of least square estimation through the use of kernel function and bandwidth is applied here to provide an optimized least square. In this regard, the weighted sum of square error criterion is introduced in this section to estimate the non-fuzzy coefficients $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)$:

$$WSSE = \sum_{i=p+1}^{T} w_i e_i, \tag{13}$$

where

(I) $e_i = d_2^2(\widetilde{X}_i, \widetilde{X}_i^*)$ where d_2^2 is the square error distance between two fuzzy numbers $\widetilde{A}$ and $\widetilde{B}$ introduced in Section 2,

(II) $\widetilde{X}_i^* = \oplus_{l=1}^{p}(\theta_l \otimes \widetilde{X}_{i-l}) \oplus \widetilde{\widehat{f}}_{\boldsymbol{\theta}}(t_i)$ where $\widetilde{\widehat{f}}_{\boldsymbol{\theta}}(t_i)$ given in Eq. (11), and

(III) $w_i = \frac{1}{1+d_2^2(\widetilde{X}_i, \widetilde{X}_i^*)}$ is the weight of i^{th} residual.

Remark 3.1. *For a specified $i^* \in \{1, 2, \ldots, T\}$, it should be noted that*

(I) $w_{i^*} \in (0, 1]$, and

(II) $w_{i^*} = 1$ if and only if $\widetilde{X}_{i^*} = \widetilde{X}^*_{i^*}$.

Therefore, according to the weighting scheme w_i, those time series data with high distance residuals (e_i) get low weights near zero.

Remark 3.2. *It should be noted that if all fuzzy time series data are reported by **TFNs** $\widetilde{X}_i = (X_i; l_{X_i}, r_{X_i})_T$, then, according to Example 2.5 and arithmetic operations of fuzzy numbers, we get*

$$\widetilde{\widehat{f}}_{\boldsymbol{\theta}}(t_i) = (X_i - \sum_{l=1}^{p} \theta_l X_{i-l}; |l_{X_i} - \sum_{l=1}^{p} (s_l \theta_l l_{X_{i-l}} - (1 - s_l)\theta_l r_{X_{i-l}})|, |r_{X_i} - \sum_{l=1}^{p} (s_l \theta_l r_{X_{i-l}} - (1 - s_l)\theta_l l_{X_{i-l}})|)_T,$$

and

$$\widetilde{X}^*_i = (X_i; \sum_{l=1}^{p} (s_l \theta_l l_{X_{i-l}} - (1 - s_l)\theta_l r_{X_{i-l}}) + |l_{X_i} - \sum_{l=1}^{p} (s_l \theta_l l_{X_{i-l}} - (1 - s_l)\theta_l r_{X_{i-l}})|,$$

$$\sum_{l=1}^{p} (s_l \theta_l r_{X_{i-l}} - (1 - s_l)\theta_l l_{X_{i-l}}) + |r_{X_i} - \sum_{l=1}^{p} (s_l \theta_l r_{X_{i-l}} - (1 - s_l)\theta_l l_{X_{i-l}})|)_T,$$

where $s_l = I(\theta_l \geq 0)$.

Table 1: Some common kernel functions

Epanechnikov	$K(y) = \begin{cases} \frac{3}{4}(1 - y^2), & \lvert y\rvert \leq 1, \\ 0, & \lvert y\rvert > 1. \end{cases}$
Triweight	$K(y) = \begin{cases} \frac{35}{32}(1 - y^2)^3, & \lvert y\rvert \leq 1, \\ 0, & \lvert y\rvert > 1. \end{cases}$
Biweight	$K(y) = \begin{cases} \frac{15}{16}(1 - y^2)^2, & \lvert y\rvert \leq 1, \\ 0, & \lvert y\rvert > 1. \end{cases}$

Thus the unknown parameter $\boldsymbol{\theta}$ can be then evaluated using the following target function:

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta} \in \mathbb{R}^p} WSSE(\boldsymbol{\theta}). \tag{14}$$

Therefore, the proposed target function provides a simple optimization procedure to evaluate the unknown exact coefficients in the presence of outliers.

Remark 3.3. *It should be pointed out that various performance measures have been proposed in literature to estimate and compare the forecasting accuracy in different models [10]. In this paper, some commonly used performance measures were extended to estimate and compare the forecasting accuracy of different models:*

1. *Mean Forecast Error:*

$$MFE = \frac{\sum_{i=p+1}^{T} d_2^2(\widetilde{\widehat{X}}_i, \widetilde{X}_i)}{T - p}. \tag{15}$$

2. *Mean Absolute Error:*

$$MAE = \frac{\sum_{i=p+1}^{T} \int |\widetilde{\widehat{X}}_i(x) - \widetilde{X}_i(x)| dx}{T - p}. \tag{16}$$

3. *Mean Absolute Scaled Error:*

$$MASE = \frac{\sum_{i=p+1}^{T} q_i^p}{T - p}, \tag{17}$$

where

$$q_i^p = \frac{\int |\widetilde{\widehat{X}}_i(x) - \widetilde{X}_i(x)|dx}{\frac{1}{T-p}\sum_{i=p+1}^{T}\int |\widetilde{\widehat{X}}_i(x) - \widetilde{X}_{i-1}(x)|dx}. \tag{18}$$

In order to evaluate the MASE, the raw fuzzy data was divided into two parts: training set and test set. The observations in the training set were used for constructing the desired model. Then a small subpart of the training set is kept for the validation purposes which is known as the validation set.

4. *Mean Absolute Percentage Error:*

$$MAPE = \frac{\sum_{i=p+1}^{T}\int |\frac{\widetilde{\widehat{X}}_i(x) - \widetilde{X}_i(x)}{\widetilde{X}_i(x)}|dx \times 100\%}{T-p}. \tag{19}$$

5. *Similarity measure* [31]*:*

$$S = \frac{1}{T-p}\sum_{i=p+1}^{T} S_{UI}(\widetilde{\widehat{X}}_i, \widetilde{X}_i), \tag{20}$$

$$S_{UI}(\widetilde{\widehat{X}}_i, \widetilde{X}_i) = \frac{Card(\widetilde{\widehat{X}}_i \cap \widetilde{X}_i)}{Card(\widetilde{\widehat{X}}_i \cup \widetilde{X}_i)}, \tag{21}$$

where $\cap$ and $\cup$ denote the intersection and union operators in the space of fuzzy numbers, respectively, and $Card(\widetilde{A})$ denotes the cardinal number of $\widetilde{A}$. Note that $S \in [0,1]$. Therefore, it is reasonable that the values of S greater than 0.5 indicate the degree of closeness between the outputs and the corresponding estimations.

For a good forecast, MFE, MFE, MASE and MAPE should be close to zero as much as possible. In addition, to examine the relationship between $\widetilde{X}s$ and $\widetilde{\widehat{X}}s$, beside the mentioned goodness-of-fit measures, we also applied the center of gravity to convert $\widetilde{X}s$ and $\widetilde{\widehat{X}}s$ into the exact values of $M_{\widetilde{X}}s$ and $M_{\widetilde{\widehat{X}}}s$. Then, according to the conventional regression models, the relationship between the defuzzified values of $M_{\widetilde{X}}s$ ($M_{\widetilde{\widehat{X}}}s$) were investigated according to the plots.

3.1 Algorithm used for selection of model components

Table 2: Values of $C(K)$ for some specific kernels.

Kernel	Epanechnikov	Biweight	Triweight
$C(K)$	2.34	2.78	3.15

To estimate the unknown components of the proposed *FSPRTSM* based on the fuzzy time series data $\widetilde{X}_1$, ..., $\widetilde{X}_T$ and specified kernel function K, the vector of exact parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)^\top$ and bandwidth h should be simultaneously estimated. In this regard, a hybrid optimization algorithm is required to estimate such parameters. For this purpose, a hybrid iterative method is extended to solve the underlying complex optimization system where the i^{th} approximation is derived from the previous ones [35]. In this regard, we suggest the following iterative algorithm to find the exact coefficients, fuzzy smooth function and optimal values for the bandwidth and autoregressive order as follows:
Step 1: Select a kernel function $K(\cdot)$.
Step 2: Let $p = 1$.
Step 3: To obtain the initial values for the bandwidth parameters h_p, in this paper, we extend Silverman Rule-of-Thumb [55], but based on the given fuzzy time series $\widetilde{X}_1, \widetilde{X}_2, \ldots, \widetilde{X}_T$ as $h_{opt} = \widehat{\sigma}C(K)T^{-1/5}$ where $C(K)$ is the constant evaluated from Table 2 and

$$\widehat{\sigma} = \sqrt{\frac{1}{T-1}d_2^2(\widetilde{X}_t, \overline{\widetilde{X}})},$$

is the standard deviation of the fuzzy data and $\overline{\widetilde{X}} = (\frac{1}{T-1}) \oplus_{t=1}^{T} \widetilde{X}_t$.

Step 4: According to the initial values of $\widehat{h}_p^{(0)}$, the optimal value of $\widehat{\boldsymbol{\theta}}^{(0)p} = (\widehat{\theta}_1^{(0)p}, \ldots, \widehat{\theta}_p^{(0)p})^\top$ can be evaluated as:

$$\widehat{\boldsymbol{\theta}}^{(0)p} = \arg\min_{\boldsymbol{\theta} \in \mathbb{R}^p} WSSE(\boldsymbol{\theta}, \widehat{h}_p^{(0)}). \tag{22}$$

Step 5: Using the optimal values of $\widehat{\boldsymbol{\theta}}_p^{(0)}$, the optimal value of $\widehat{h}_p^{'(1)}$ is selected according to the "generalized cross-validation criterion" (GCV)) [28] as follows:

$$\widehat{h}_p^{(1)} = arg\min_{h} GCV(\widehat{\boldsymbol{\theta}}^{(0)p}, h), \tag{23}$$

where

$$GCV(\widehat{\boldsymbol{\theta}}^{(0)p}, h) = \frac{1}{T-p} \sum_{t=p+1}^{T} \frac{d_2^2(\widetilde{X}_i, \widetilde{\widehat{X}}_i)}{(1 - tr(W_h)/(T-p))^2}, \tag{24}$$

where $tr(W_h)$ shows the trace of the matrix $W_h = [w^{h_l}(s,t)]$ with

$$w^h(t,s) = \frac{\sum_{i=1}^{p} K(\frac{X_{t-i} - X_{s-i}}{h_l})}{\sum_{j=p+1}^{T} \sum_{i=1}^{p} K(\frac{X_{j-i} - X_{s-i}}{h})}. \tag{25}$$

Step 6: Using the optimal value of $\widehat{h}_p^{(1)}$, the optimal value of $\widehat{\boldsymbol{\theta}}^{(1)p} = (\widehat{\theta}_1^{(1)p}, \ldots, \widehat{\theta}_p^{(1)p})^\top$ can be evaluated as

$$\widehat{\boldsymbol{\theta}}_p^{(1)} = \arg\min_{\boldsymbol{\theta} \in \mathbb{R}^p} WSSE(\boldsymbol{\theta}, \widehat{h}_p^{(1)}). \tag{26}$$

Step 7: For a small positive number of ε near zero, if

$$\max_{i=1}^{p} \left\{ |\widehat{\theta}_i^{(1)p} - \widehat{\theta}_i^{(0)p}| \right\} < \varepsilon,$$

then $\widehat{h}_p^{(1)}$, $\widehat{h}_p^{'(1)}$ and $\widehat{\boldsymbol{\theta}}_p^{(1)}$ are the optimal solutions; otherwise, return to Step 6 and repeat the algorithm until

$$\max_{i=1}^{p} \left\{ |\widehat{\theta}_i^{(i)p} - \widehat{\theta}_i^{(i-1)p}| \right\} < \varepsilon.$$

Step 8: Let $p = p + 1$ and return to Step 3 and repeat the algorithm until $p_{opt} = \max_{p \in \{1,2,3,\ldots\}} S_p^{(i)}$ where $S_p^{(i)}$ stands for the similarity measure in the i^{th} iteration. Therefore, $\hat{\boldsymbol{\theta}}_{p_{opt}}^{(i)}(\widehat{h}_{p_{opt}}^{(i)}, \widehat{h}_{p_{opt}}^{'(i)})$ is the final optimal regression coefficient.

Remark 3.4. *To examine the relationship between $\widetilde{X}s$ and $\widetilde{\widehat{X}}s$, we employed the center of gravity to convert $\widetilde{X}s$ and $\widetilde{\widehat{X}}s$ into the exact values $M_{\widetilde{X}}s$ and $M_{\widetilde{\widehat{X}}}s$. Then, according to the conventional regression models, the relationship between the defuzzified values of $M_{\widetilde{X}}s$ ($M_{\widetilde{\widehat{X}}}s$) was investigated according to the plots. However, the outliers may strongly affect the estimations of parameters in a statistical model. A series of powerful general methods have been developed for detecting and investigating the outliers* [38]*. Here, a common criterion was extended to identify possible influential fuzzy outliers. Inspired by* [15]*, the Cook's distance was extended based on the fuzzy data and distance measure d as follows:*

$$D_i = \frac{\sum_{j=p+1}^{T} d_2^2(\widetilde{\widehat{X}}_j, \widetilde{\widehat{X}}(j))}{(k+1)MSE}, \tag{27}$$

where $\widetilde{\widehat{X}}(j)$ denotes the regression estimate when the j^{th} data point is removed and $MSE = \frac{1}{n} \sum_{j=p+1}^{T} d_2^2(\widetilde{X}_j, \widetilde{\widehat{X}}_j)$. To identify the potential outliers, it is enough to look at the distribution of the Cook's distance and see whether there are any conspicuously large values relative to the others. For this purpose, this paper suggested the conventional $\overline{D}$-chart [40] *to measure the variations of the Cook's distances. It is worth noting that the upper and lower control limits of $\overline{D}$-chart are $LCL = \overline{D} - 3S_D$ and $UCL = \overline{D} + 3S_D$.*

Remark 3.5. *Compare to Hesamian and Akbari method, it should be noted that the proposed method used a modified generalized subtraction criterion. Unlike the Hesamian and Akbari, the modified generalized subtraction criterion always leads to a fuzzy number with positive spreads. In addition, it modified the least-square method in Hesamian and Akbari approach by employing a weighted method to control the effect of outliers on predictions. Note that Zarei et al.* [65] *applied a method similar to that of Hesamian and Akbari to estimate the parameters of the model and the fuzzy smoothing function based on a special distance, the kernel for triangular fuzzy numbers. In this regard, we only compared the proposed method with Hesamian and Akbari methods in our computational procedure.*

4 Numerical examples

In this section, the feasibility and effectiveness of the proposed *FSPRTSM* were examined via two numerical examples. Moreover, the proposed *FSPRTSM* was also examined with an existing fuzzy time series model (*FSPTSM*) introduced by [31]. In order to conduct a competitive study, the measures of MFE, MAE, $MAPE$, $MASE_p$ (MASE based on the whole data) and S were applied to calculate the goodness-of-fit criteria. In addition, the popular kernels including triweight, Epanechnikov and biweight were applied to examine the model's performance.

Table 3: Performance measures of proposed method in comparison with Hesamian and Akbari in presence of outliers in Example 4.1.

Method	Kernel		Results
FSPRTSM	triweight	$h_{opt} = 0.15, p = 2$	
		Coefficient estimation	$\widehat{\theta}_1 =, \quad \widehat{\theta}_2 =$
		MFE	650.54
		MAE	40.32
		$MAPE$	0.151
		$MASE_{20}$	30.55
		$MASE_{50}$	29.23
		$MASE_{70}$	36.22
		S	0.77
	biweight	$h_{opt} = 0.15, p = 2$	
		Coefficient estimation	$\widehat{\theta}_1 = 0.623, \quad \widehat{\theta}_2 = 0.182$
		MFE	670.25
		MAE	44.23
		$MAPE$	0.201
		$MASE_2$	30.26
		$MASE_{50}$	39.23
		$MASE_{70}$	38.30
		S	0.71
	epanechnikov	$h_{opt} = 0.20, p = 2$	
		Coefficient estimation	$\widehat{\theta}_1 = 0.623, \quad \widehat{\theta}_2 = 0.226$
		MFE	650.36
		MAE	45.56
		$MAPE$	0.205
		$MASE_{20}$	29.33
		$MASE_{50}$	36.57
		$MASE_{70}$	35.65
		S	0.68
FSPTSM	triweight	$h_{opt} = 0.175, p = 2$	
		Coefficient estimation	$\widehat{\theta}_1 = 0.652, \quad \widehat{\theta}_2 = 0.253$
		MFE	730.36
		MAE	77.23
		$MAPE$	0.35
		$MASE_{20}$	59.23
		$MASE_{50}$	63.23
		$MASE_{70}$	67.03
		S	0.58
	biweight	$h_{opt} = 0.18, p = 2$	
		Coefficient estimation	$\widehat{\theta}_1 = 0.732, \quad \widehat{\theta}_2 = 0.296$
		MFE	748.02
		MAE	78..03
		$MAPE$	0.402
		$MASE_{20}$	69.26
		$MASE_{50}$	68.16
		$MASE_{70}$	68.39
		S	0.52
	epanechnikov	$h_{opt} = 0.2, p = 2$	
		Coefficient estimation	$\widehat{\theta}_1 = 0.739, \quad \widehat{\theta}_2 = 0.281$
		MFE	750.36
		MAE	75.56
		$MAPE$	0.422
		$MASE_{20}$	71.33
		$MASE_{50}$	73.57
		$MASE_{70}$	72.85
		S	0.51

Example 4.1. *Consider a sample of simulated fuzzy data set with the size of 700 generated by the following fuzzy time series model:*

$$\widetilde{X}_i = \oplus_{l=1}^{2}(\theta_l \otimes \widetilde{X}_{i-l}) \oplus \widetilde{f}(t_i) \oplus \widetilde{\epsilon}_i, \quad i = 3, 4, \ldots, 700,$$

where

1. $\theta_1 = 0.6, \theta_2 = 0.2$,

2. $\widetilde{f}(t_i) = (f(t_i); l_{f(t_i)}, r_{f(t_i)})_T$ *where* $f(t_i) = \sqrt{t_i(t_i-1)}\sin(\frac{0.2\pi}{t_i+0.05})$ *and* $l_{f(t_i)}, r_{f(t_i)}$ *are the random variables taken from* $U(0, 0.2)$,

3. $\widetilde{X}_i = (X_i; l_{X_i}, r_{X_i})_T$, $i = 1, 2$ *are the initial fuzzy values where* $X_i \sim Levey(0, 1)$ *(with density function of* $f(x) = (1/\sqrt{2\pi})x^{-1.5}e^{-0.5/x}$, $x > 0$*) and* l_{X_i}, r_{X_i} *are the random samples from* $U(0, 2)$, *respectively,*

4. $\widetilde{\epsilon}_i = (\epsilon_i; l_{\epsilon_i}, r_{\epsilon_i})_T$ *where* $\epsilon_i \sim Laplace(0, 4))$ *(with density function of* $f(x) = (1/8)e^{-|x|/4}$, $x \in \mathbb{R}$*),* l_{ϵ_i} *and* r_{ϵ_i} *are the random variables taken from* $U(0, 0.01)$.

5. $t_i = \frac{i}{700}$.

The simulated data set is plotted in Fig. 1, which shows that the data set may include some potential outliers. In order to clearly detect the outliers, $\overline{D}$*-chart is plotted in Fig. 2 for the Cook's distances (*D_i *values). It can be observed that some of* D_i *values are excluded from the specified lower and upper control limits, confirming the presence of outliers in the data set. For this case, the goodness-of-fit values corresponding to FSPRTSM are summarized in Table 3 based on some specific kernels. The performance of the proposed method was also examined for some sample sizes of training data and test data. It can be seen that the best results are relevant to triweight kernel in terms of the performance measures of* $S = 0.77$, $MFE = 650.54$, $MAE = 40.32$, $MAPE = 0.151$, $MASE_{20} = 30.55$, $MASE_{50} = 29.23$ *and* $MASE_{70} = 36.22$. *In order to compare our method with that of [31], the results of FSPTSM are also summarized in Table 3 for some specific kernels. Regarding each applied kernel and goodness-of-fit criterion, it is observed that the proposed method outperformed the Hesamian and Akbari's method. The reason is that unlike the Hesamian and Akbari's method, the proposed method relies on a weighted least square which controls the effect of outliers on the prediction. Therefore, the proposed fuzzy time series method provided more accurate results for the prediction when compared to the Hesamian and Akbari's method in the presence of outliers. However, the best results are relevant to triweight kernel in terms of all performance measures including* S, MEF, MAE, $MAPE$, $MASE_{20}$, $MASE_{50}$ *and* $MASE_{70}$. *The performance of the proposed FSPRTSM was also compared to FSPTSM for triweight kernel (the model with the best performance) as also shown in Fig. 3. It can be observed that in our method, the values of* $M_{\widetilde{\widetilde{X}}}$ *are closer to* $M_{\widetilde{X}}$ *compared to the Hesamian and Akbari's method. Therefore, it can be concluded that the proposed FSPRTSM gives more accurate results than FSPTSM in this example.*

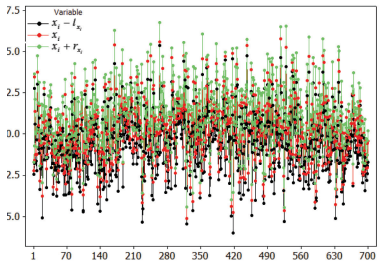


Figure 1: Plot of $\widetilde{X}_t$ versus t in Example 4.1.

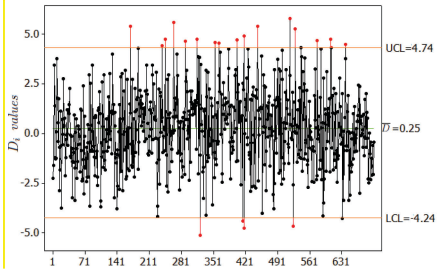


Figure 2: $\overline{D}$-chart for detecting outliers in Example 4.1.

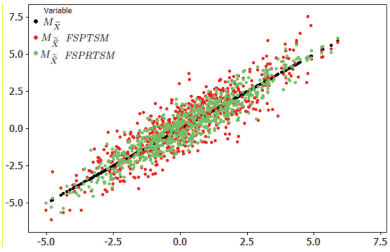


Figure 3: Comparison of $M_{\widetilde{X}}$ and $M_{\widetilde{X}}$ for *FSPRTSM* and *FSPTSM* in Example 4.1.

Table 4: Cook's distances for outliers in Example 4.2.

No.	42	46	47	48	49	51
D_i	0.142940	0.223297	0.491524	0.519409	0.184230	0.115417

Table 5: Data set in Example 4.2.

No.	$\widetilde{X}_i$	No.	$\widetilde{X}_i$
1	$(17967; 249.99)_T$	2	$(18983; 452.07)_T$
3	$(20150; 263.27)_T$	4	$(21525; 415.69)_T$
5	$(22692; 387.98)_T$	6	$(24833; 494.21)_T$
7	$(25575; 483.82)_T$	8	$(23533; 569.84)_T$
9	$(25217; 378.16)_T$	10	$(27550; 782.31)_T$
11	$(32708; 999.97)_T$	12	$(36108; 591.21)_T$
13	$(39427; 794.43)_T$	14	$(44283; 800.21)_T$
15	$(48983; 1238.42)_T$	16	$(55792; 1431.83)_T$
17	$(62750; 990.16)_T$	18	$(64708; 1074.45)_T$
19	$(68825; 1194.54)_T$	20	$(69300; 1311.74)_T$
21	$(75458; 1232.07)_T$	22	$(80017; 1034.61)_T$
23	$(84275; 1353.31)_T$	24	$(92233; 1666.23)_T$
25	$(104675; 3760.86)_T$	26	$(113350; 2391.38)_T$
27	$(119945; 2161.02)_T$	28	$(122500; 2749.92)_T$
29	$(120017; 1132.18)_T$	30	$(121300; 2543.81)_T$
31	$(126142; 1978.00)_T$	32	$(130408; 1774.20)_T$
33	$(133433; 1807.68)_T$	34	$(139767; 2235.50)_T$
35	$(145050; 1461.85)_T$	36	$(151975; 1699.14)_T$
37	$(159842; 2991.25)_T$	38	$(166908; 2994.72)_T$
39	$(172727; 2811.70)_T$	40	$(185025; 3768.37)_T$
41	$(191383; 3747.00)_T$	42	$(217817; 4203.11)_T$
43	$(234208; 3841.69)_T$	44	$(243067; 4464.07)_T$
45	$(243742; 5602.03)_T$	46	$(230408; 5604.34)_T$
47	$(214500; 3361.33)_T$	48	$(221242; 5846.25)_T$
49	$(224317; 4860.13)_T$	50	$(242108; 5918.42)_T$
51	$(265092; 5086.46)_T$	52	$(283775; 7859.47)_T$
53	$(297258; 4763.14)_T$	54	$(303613; 8264.19)_T$

Example 4.2. *The data set given in Table 5 denotes the annual average house prices in the United States from 1963 to 2016 and collected from "https:// fred.stlouisfed.org/ series/ HOSMEDUSM052N". The house prices are reported by fuzzy data. First, we want to detect the outliers. Based on the $\overline{D}$-chart for the Cook's distances of the fuzzy time series, the observations with numbers* $42, 46, 47, 48, 49$ *and* 51 *are potential to be outliers whose corresponding Cook's*

distances are summarized in Table 4. As it was discussed in Example 4.1, the best results were assigned to the triweight kernel. Therefore, the goodness-of-fit values corresponding to FSPRTSM and FSPTSM are summarized in Table 6 for the triweigh kernel. The estimated weights for each residual are listed in Table 7. As it is expected, the outliers of $42, 46, 47, 48, 49$ *and* 51 *get low weights compared to the other residuals. Comparing the goodness-of-fit measurers of two methods, the proposed FSPRTSM gives better results when the outliers are included in the data set. The performance of the proposed FSPRTSM relative to FSPTSM can also be investigated in Fig. 5. As can be seen, the* $M_{\widetilde{\widehat{X}}}$ *values in our method are closer to* $M_{\widetilde{X}}$ *than FSPTSM. This concludes that the proposed FSPRTSM is more efficient than FSPTSM when the outliers exist in the data set.*

Table 6: Performance measures of proposed *FSPRTSM* in comparison with *FSPTSM* in presence of outliers in Example 4.2.

Method		Results
FSPRTSM	h_{opt}, p	0.25,3
	Coefficient estimation	$\widehat{\theta}_1 = 0.0120,\ \widehat{\theta}_2 = 0.0799,\ \widehat{\theta}_3 = 0.0623$
	MFE	25372.72
	MAE	45183.73
	$MASE_2$	3408.71
	$MASE_{48}$	3839.79
	$MASE_{39}$	4582.81
	S	0.71
FSPTSM	h_{opt}, p	0.25,3
	Coefficient estimation	$\widehat{\theta}_1 = 0.00056,\ \widehat{\theta}_2 = 0.1129,\ \widehat{\theta}_3 = -0.0727$
	MFE	26018.01
	MAE	46139.71
	$MAPE$	0.268
	$MASE_2$	4210.02
	$MASE_{48}$	4692.36
	$MASE_{39}$	5166.48
	S	0.66

Table 7: Estimated weights in Example 4.2.

No.	w_i	No.	w_i	No.	w_i
1	0.109351	19	0.02843987	37	0.00431485
2	0.908621	20	0.00704139	38	0.00417109
3	0.263417	21	1	39	0.00085046
4	0.256738	22	0.0025701	40	0.0229669
5	0.0464713	23	0.0332899	41	0.276312
6	0.311965	24	0.0256306	42	**0.00023314**
7	0.0682574	25	0.00289482	43	0.248999
8	0.00710177	26	0.00256324	44	0.00710758
9	0.0430426	27	0.0892919	45	0.0536404
10	0.0144649	28	0.0259371	46	**0.00179771**
11	0.0160415	29	0.0475018	47	**0.000147954**
12	0.124338	30	0.122847	48	**0.00310744**
13	0.01516	31	0.00180854	49	**0.00032392**
14	0.0707164	32	0.00187089	50	0.0055931
15	0.0115311	33	0.00719297	51	**0.000464458**
16	0.028455	34	0.00729025	52	0.0027041
17	0.109351	35	0.00183896	53	0.00311214
18	0.908621	36	0.00556081	54	0.00395565

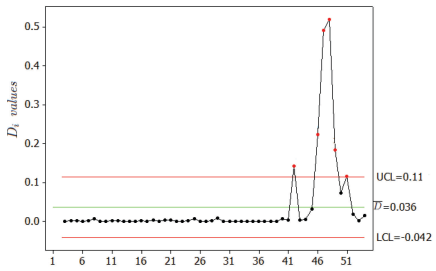


Figure 4: $\overline{D}$-chart for detecting outliers in Example 4.2.

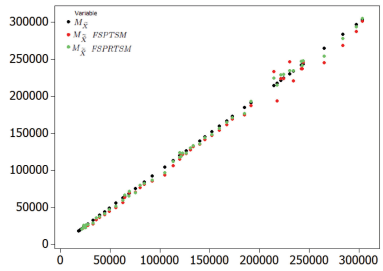


Figure 5: Comparison of the values of $M_{\widetilde{\widetilde{X}}}$ and $M_{\widetilde{X}}$ for *FSPRTSM* and *FSPTSM* in Example 4.2.

Example 4.3. *Consider a study to predict the global land-ocean temperature [39] according to the proposed FSPRTSM. For this purpose, the monthly global land-ocean temperature reports from 1880 to 2021 were represented by triangular fuzzy number as* $\widetilde{x}_t = (x_t; 0.01|x_t|)_L$ *where* x_t *values are shown in Fig. 8 and* $L(x) = 1 - x^2$*. Fig. 7 (for the Cook's distances) shows that the data set includes no outliers. The goodness-of-fit values corresponding to the proposed method are listed in Table 8 for the triweight kernel. The estimated values of weights for each residual are given in Table 9. The performance of the FSPTSM is also listed in Table 8. It can be seen that the proposed method exhibited a better performance according to all goodness-of-fit measures* $MFE, MAE, MAPE, MASE$ *and* S*. The plots of fuzzy data and corresponding estimates of the proposed FSPRTSM and FSPTSM are depicted in Fig. 6. Compared to the mentioned methods, the fuzzy predicted values in our method are closer to fuzzy observations which shows a better result in prediction.*

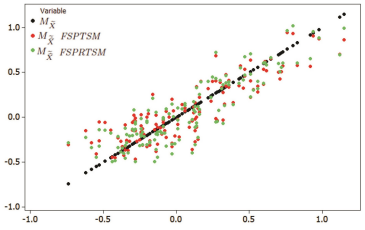


Figure 6: Comparison of $M_{\widetilde{\widetilde{X}}}$ and $M_{\widetilde{X}}$ for *FSPRTSM* and *FSPTSM* in Example 4.3.

Table 8: Performance measures of proposed *FSPRTSM* in comparison with *FSPTSM* in presence of outliers in Example 4.3.

Method		Results
FSPRTSM	$h_{opt} = 0.0.5, p = 3$	
	Coefficient estimation	$\widehat{\theta}_1 = 0.342,\ \widehat{\theta}_2 = 0.321,\ \widehat{\theta}_3 = 0.265$
	MFE	3.06
	MAE	4.13
	$MAPE$	0.0081
	$MASE_{110}$	3.50
	$MASE_{120}$	3.51
	$MASE_{130}$	4.82
	S	0.48
FSPTSM	$h_{opt} = 0.05, p = 3$	
	Coefficient estimation	$\widehat{\theta}_1 = 0.374,\ \widehat{\theta}_2 = 0.344,\ \widehat{\theta}_3 = 0.237$
	MFE	3.70
	MAE	4.92
	$MAPE$	0.015
	$MASE_{110}$	3.58
	$MASE_{120}$	3.57
	$MASE_{130}$	5.03
	S	0.43

Table 9: Estimated weights in Example 4.3.

No.	$\widehat{w}_i$	No.	$\widehat{w}_i$	No.	$\widehat{w}_i$	No.	$\widehat{w}_i$	No.	$\widehat{w}_i$	No.	$\widehat{w}_i$	No.	$\widehat{w}_i$	No.	$\widehat{w}_i$	No.	$\widehat{w}_i$
1	-	19	0.96	37	0.99	55	0.94	73	0.92	91	0.94	109	0.92	127	0.98		
2	-	20	0.99	38	0.81	56	0.98	74	0.96	92	0.99	110	0.94	128	0.89		
3	-	21	0.93	39	0.99	57	1	75	0.96	93	0.93	111	0.99	129	0.85		
4	0.93	22	0.99	40	0.99	58	0.99	76	0.98	94	0.90	112	0.99	130	0.99		
5	0.97	23	0.99	41	0.98	59	0.94	77	0.97	95	0.94	113	0.98	131	0.97		
6	0.91	24	0.99	42	0.97	60	0.99	78	0.99	96	0.97	114	0.99	132	0.99		
7	0.98	25	0.82	43	0.97	61	0.99	79	0.89	97	0.99	115	0.98	132	0.99		
8	0.93	26	0.99	44	0.99	62	0.92	80	0.98	98	0.97	116	0.95	133	0.98		
9	0.99	27	0.94	45	0.99	63	0.95	81	0.98	99	0.99	117	0.97	134	0.99		
10	0.91	28	0.99	46	0.97	64	0.94	82	0.99	100	0.99	118	0.99	135	0.96		
11	0.99	29	0.98	47	0.86	65	0.95	83	0.99	101	0.96	119	0.94	136	0.94		
12	0.97	30	0.92	48	0.98	66	0.99	84	0.99	102	0.92	120	0.98	137	0.85		
13	1	31	0.99	49	0.99	67	0.99	85	0.99	103	0.96	121	0.98	138	0.98		
14	0.83	32	0.98	50	0.84	68	0.85	86	0.98	104	0.95	122	0.99	139	0.96		
15	0.99	33	0.96	51	0.99	69	0.55	87	0.99	105	0.99	123	0.92	140	0.99		
16	0.99	34	0.99	52	0.93	70	0.99	88	0.99	106	0.98	124	0.95	141	0.92		
17	0.89	35	0.82	53	0.86	71	0.90	89	0.98	107	0.99	125	0.99	142	0.98		
18	0.97	36	0.99	54	0.94	72	0.96	90	0.99	108	0.99	126	0.99	-	-		

Table 10: Data set in Example 4.4.

No.	$\widetilde{X}_i$	No.	$\widetilde{X}_i$
1	$(1.7337; 0.8051)_T$	2	$(2.3302; 0.8051)_T$
3	$(1.3199; 0.7742)_T$	4	$(5.0507; 0.8948)_T$
5	$(1.4206; 1.0540)_T$	6	$(4.0273; 0.9331)_T$
7	$(2.8624; 1.0480)_T$	8	$(4.7107; 1.0647)_T$
9	$(4.1098; 1.1028)_T$	10	$(4.4843; 1.0881)_T$
11	$(1.3249; 1.0064)_T$	12	$(3.2249; 0.5503)_T$
13	$(3.2916; 0.5049)_T$	14	$(3.5508; 0.9268)_T$
15	$(2.9145; 1.1507)_T$	16	$(2.6085; 1.1335)_T$
17	$(3.0432; 0.4489)_T$	18	$(6.8010; 0.9588)_T$
19	$(4.9351; 0.8115)_T$	20	$(3.5672; 0.6054)_T$
21	$(3.8828; 1.1579)_T$	22	$(0.5183; 0.9652)_T$
23	$(3.6846; 0.8175)_T$	24	$(3.5117; 0.4248)_T$
25	$(2.8294; 0.7956)_T$	26	$(2.3836; 1.0101)_T$
27	$(3.9454; 0.7558)_T$	28	$(3.6012; 1.0266)_T$

Example 4.4. *The data set in Table 5 shows the Mauna Loa atmospheric CO_2 concentration as a time series*

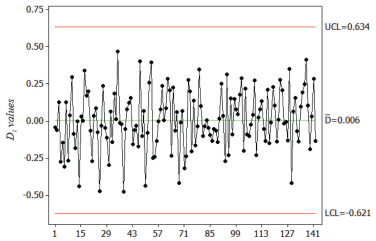


Figure 7: $\overline{D}$-chart for detecting outliers in Example 4.3.

with annual records for each months from 1959 to December 1995. The source of the data set (www.cotf. edu/ete/ modules/carbon/ CO2MAUNA) includes monthly atmospheric CO_2 concentration for all years. Inspired by Buckely [7], such a data set can be used for evaluating the annual CO_2 concentration as a ***TFN*** *as* $\widetilde{x}_t = (\overline{x}_t - t_{11,0.025}s_t/\sqrt{12}, \overline{x}_t, \overline{x}_t + t_{11,0.025}s_t/\sqrt{12})_T$ *where* $\overline{x}_t$ *and* s_t *denote the mean and standard deviation for atmospheric CO_2 concentration in the year of* $t \in \{1959, 1960, ..., 1995\}$*, respectively. The fuzzy data set is listed in Table 10. Fig. 8 (for the Cook's distances) shows that the data set includes a potential outlier. The results of the proposed method along with other time series models (for the triweght kernel) are summarized in Table 11. A comparison of different methods indicated that the proposed method led to the more accurate results in terms of all performance measures (MSM, $MAPE$, S and MFE). The performance of the proposed method can be also examined (see Fig. 9). Clearly,* $\widetilde{X}$ *values estimated by our method are closer to their corresponding* $\widetilde{\widetilde{X}}$ *values reflecting the superior performance of the proposed method in investigating the relationship between predictors and fuzzy responses in this example.*

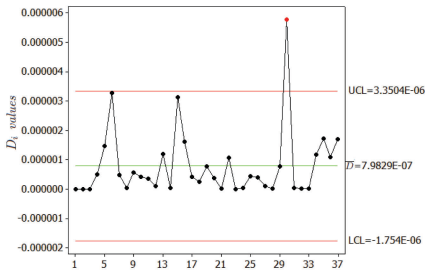


Figure 8: $\overline{D}$-chart for detecting outliers in Example 4.4.

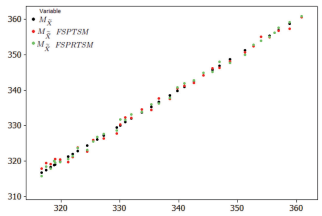


Figure 9: Comparison of $M_{\widetilde{\widetilde{X}}}$ and $M_{\widetilde{X}}$ for *FSPRTSM* and *FSPTSM* in Example 4.4.

Table 11: Performance measures of proposed *FSPRTSM* in comparison with *FSPTSM* in presence of outliers in Example 4.4.

Method		Results
FSPRTSM	$h_{opt}=0.15, p=1$	
	Coefficient estimation	$\widehat{\theta}_1=1.0629$
	MFE	2934.23
	MAE	3026.552
	$MAPE$	0.0924
	S	0.87
FSPTSM	$h_{opt}=0.20, p=1$	
	Coefficient estimation	$\widehat{\theta}_1=1.0985$
	MFE	2987.03
	MAE	30647.39
	$MAPE$	0.117
	S	0.81

Table 12: Estimated weights in Example 4.4.

No.	$\widehat{w}_i$	No.	$\widehat{w}_i$	No.	$\widehat{w}_i$
1	-	13	0.73240	25	0.89307
2	-	14	0.98814	26	0.90177
3	-	15	0.51682	27	0.97352
4	0.86090	16	0.67658	28	0.99479
5	0.68236	17	0.88898	29	0.83043
6	0.49233	18	0.93389	30	**0.39826**
7	0.87064	19	0.81609	31	0.99012
8	0.98730	20	0.90135	32	0.99605
9	0.85137	21	0.99583	33	0.99235
10	0.88491	22	0.76899	34	0.69834
11	0.89894	23	0.99781	35	0.78811
12	0.96759	24	0.99030	36	0.70416

5 Conclusion

Many of the statistical prediction models proposed in fuzzy environments rely on the square error technique. However, such methods may perform poorly in the presence of unusual data. However, the statistical weighting could resolve such drawbacks. This paper is thus aimed to introduce a novel weighted fuzzy semiparametric time model for the fuzzy data, exact coefficients and fuzzy smooth function. For this purpose, a novel generalized subtraction was utilized to evaluate the unknown fuzzy smooth function of the model. Compared to the most commonly used Hukuhara subtraction, the main advantages of such fuzzy arithmetic operators were also discussed. Then, a simple efficient weighted least square target was suggested to control the outliers for the prediction accuracy of the fuzzy time series model. The effectiveness and advantages of the modified time series model (for both cases where the data set contains outliers or not) were examined and compared in terms of some common goodness-of-fit criteria including the similarity measure-based criterion, MFE, MASE and MAPE. Specifically, the performance measure of MASE was developed in this paper to identify the effectiveness of the sample and out-of-sample data in the prediction procedure of the proposed method. Moreover, a common center of gravity criterion was employed to examine the outlier detection and performance of the proposed method by some scatter plots. The results of the numerical examples suggested that the proposed method outperformed the least-square methods in the presence of outliers. The results also revealed that the proposed method exhibited better predictions where the dataset includes no outliers. The proposed approach showed potential effectiveness for the fuzzy time series model in real applications such as hydrology, agriculture, industry, economics, and social studies. Although the triangular fuzzy numbers were used in the numerical evaluations, the proposed fuzzy time series model is still applicable to all kinds of fuzzy numbers. Extending the proposed model to other types of statistical time series models such as nonparametric or seasonal time series could be a potential topic for future studies.

Acknowledgements. The authors would like to thank the referees and editor for their remarks and criticisms that have led to improve the presentation of this work.

References

[1] S. S. G. Abhishekh, S. R. Singh, *A score function-based method of forecasting using intuitionistic fuzzy time series*, New Mathematics and Natural Computation, **14** (2018), 91-111.

[2] M. G. Akbari, G. Hesamian, *Linear model with exact inputs and interval-valued fuzzy outputs*, IEEE Transactions on Fuzzy Systems, **26** (2018), 518-530.

[3] S. Askari, N. Montazerin, M. H. F. Zarandi, *A clustering based forecasting algorithm for multivariate fuzzy time series using linear combinations of independent variable*, Applied Soft Computing, **35** (2015), 151-160.

[4] K. A'yun, A. M. N. Abadi, F. Y. Saptaningtyas, *Application of weighted fuzzy time series model to forecast trans jogja's passengers*, International Journal of Applied Physics and Mathematics, **5** (2015), 76-85.

[5] E. Bas, E. Egrioglu, C. H. Aladag, U. Yolcu, *Fuzzy time-series network used to forecast linear and nonlinear time series*, Applied Intelligence, **43** (2015), 343-355.

[6] D. Bosq, *Nonparametric statistics for stochastic process*, Springer, New York, 1996.

[7] J. J. Buckley, *Fuzzy statistics, studies in fuzziness and soft computing*, Springer-Verlag, Berlin, 2006.

[8] O. Cagcag, U. Yolcu, E. Egrioglu, C. H. Aladag, *A high order fuzzy time series forecasting method based on operation of intersection*, Applied Mathematical Modelling, **40** (2016), 8750-8765.

[9] Q. Cai, D. Zhang, W. Zheng, S. C. H. Leung, *A new fuzzy time series forecasting model combined with ant colony optimization and auto-regression*, Knowledge-Based Systems, **74** (2015), 61-68.

[10] L. J. Cao, E. H. T. Francis, *Support vector machine with adaptive parameters in financial time series forecasting*, IEEE Transactions on Neural Networks, **14** (2003), 1506-1518.

[11] F. T. S. Chan, A. Samvedi, S. H. Chung, *Fuzzy time series forecasting for supply chain disruptions*, Industrial Management and Data Systems, **115** (2015), 419-435.

[12] S. M. Chen, S. W. Chen, *Fuzzy forecasting based on two-factors second-order fuzzy-trend logical relationship groups and the probabilities of trends of fuzzy logical relationships*, Transactions on Cybernetics, **45** (2015), 405-417.

[13] C. H. Cheng, C. H. Chen, *Fuzzy time series model based on weighted association rule for financial market forecasting*, Expert System, **35** (2018), 23-30.

[14] S. H. Cheng, S. M. Chen, W. S. Jian, *Fuzzy time series forecasting based on fuzzy logical relationships and similarity measures*, Information Sciences, **327** (2016), 272-287.

[15] R. D. Cook, *Detection of influential observations in linear regression*, Technometrics, **19** (1977), 15-18.

[16] D. Dubois, P. Prade, *Operations on fuzzy numbers*, International Journal of Systems Science, **9** (1978), 613-626.

[17] R. Efendi, M. M. Deris, Z. Ismail, *Implementation of fuzzy time series in forecasting of the non-stationary data*, International Journal of Computational Intelligence and Applications, **15** (2016), 1-15.

[18] R. Efendi, Z. Ismail, M. M. Deris, *A new linguistic out-sample approach of fuzzy time series for daily forecasting of Malaysian electricity load demand*, Applied Soft Computing, **28** (2015), 422-430.

[19] S. Efromovich, *Nonparametric curve estimation: Methods, theory and applications*, New York, Springer, 1999.

[20] E. Egrioglu, C. H. Aladag, U. Yolcu, A. Z. Dalar, *A hybrid high order fuzzy time series forecasting approach based on PSO and ANNs methods*, American Journal of Intelligent Systems, **6** (2016), 22-29.

[21] E. Egrioglu, E. Bas, C. H. Aladag, U. Yolcu, *Probabilistic fuzzy time series method based on artificial neural network*, American Journal of Intelligent Systems, **6** (2016), 42-47.

[22] R. L. Eubank, J. D. Hart, P. Speckman, *Trigonometric series regression estimators with an application to partially linear models*, Journal of Multivariate Analysis, **32** (1990), 70-83.

[23] M. Fanyong, Z. Jinxian, Z. Qiang, *The shapley function for fuzzy games with fuzzy characteristic functions*, International Journal of Fuzzy Systems, **25** (2013), 23-35.

[24] C. Gang, W. Peng-fei, L. I. Jin-ling, *Optimization algorithm for fuzzy time series model based on autocorrelation function*, Control and Decision, **30** (2015), 1977-1802.

[25] S. S. Gautam, S. Singh, *A refined method of forecasting based on high-order intuitionistic fuzzy time series data*, Progress in Artificial Intelligence, **7** (2018), 339-350.

[26] H. Ghosh, S. Chowdhury, Prajneshu, *An improved fuzzy time-series method of forecasting based on L-R fuzzy sets and its application*, Journal of Applied Statistics, **43** (2016), 1128-1139, Doi:10.1080/02664763.2015.1092111.

[27] A. Ghotmar, P. G. Khot, P. Sharma, *Forecasting model based on fuzzy time series with first order diffencing*, International Journal of Mathematical Archive, **7** (2016), 75-79.

[28] G. Golub, M. Heath, G. Wahba, *Generalized cross-validation as a method for choosing a good ridge parameter*, Technometrics, **21** (1979), 215-223.

[29] H. Guan, Z. Dai, A. Zhao, J. He, *A novel stock forecasting model based on High-order-fuzzy-fluctuation trends and back propagation neural network*, PLoS ONE, **13**(2) (2018), Doi:10.1371/journal.pone.0192366.

[30] C. Gupta, G. Jain, D. K. Tayal, O. Castillo, *ClusFuDE: Forecasting low dimensional numerical data using an improved method based on automatic clustering, fuzzy relationships and differential evolution*, Engineering Applications of Artificial Intelligence, **71** (2018), 175-189.

[31] G. Hesamian, M. G. Akbari, *A semi-parametric model for time series based on fuzzy data*, IEEE Transactions on Fuzzy Systems, **26** (2018), 2953-2966.

[32] G. Hesamian, M. G. Akbari, *Fuzzy absolute error distance measure based on a generalised difference operation*, International Journal of Systems Science, Taylor and Francis Journals, **49**(11) (2018), 2454-2462.

[33] M. Hudec, D. Praenka, *Collecting and managing fuzzy data in statistical relational databases*, Statistical Journal of the IAOS, **32** (2016), 245-255.

[34] W. S. Jian, *Fuzzy forecasting based on fuzzy logical relationships, fuzzy trend logical relationship groups, k-means clustering algorithm, similarity measures and particle swarm optimization techniques*, Information Sciences, **327** (2016), 272-287.

[35] C. T. Kelley, *Iterative methods for optimization*, SIAM, (1999).

[36] K. H. Lee, *First course on fuzzy theory and applications*, Springer-Verlag, Berlin, 2005.

[37] R. Li, *Water quality forecasting of Haihe River based on improved fuzzy time series model*, Desalination and Water Treatment, **106** (2018), 285-291.

[38] R. A. Maronna, D. R. Martin, V. J. Yohai, *Robust statistics: Theory and methods*, New York, John Wiley and Sons, 2006.

[39] T. C. Mills, *Applied time series analysis: A practical guide to modelling and forecasting*, London: Academic Press, 2019.

[40] D. C. Montgomery, *Introduction to statistical quality control*, John Wiley and Sons, New York, 2009.

[41] V. Novák, *Detection of structural breaks in time series using fuzzy techniques*, The International Journal of Fuzzy Logic and Intelligent Systems, **18** (2018), 1-12.

[42] T. O. Olatayo, A. I. Taiwo, *Statistical modelling and prediction of rainfall time series data*, Global Journal of Computer Science and Technology, **14** (2014), 1-9.

[43] T. T. H. Phan, A. Bigand, E. P. Caillault, *A new fuzzy logic-based similarity measure applied to large gap imputation for uncorrelated multivariate time series*, Applied Computational Intelligence and Soft Computing, (2018), 1-15.

[44] W. Qiu, C. Zhang, Z. Ping, *Generalized fuzzy time series forecasting model enhanced with particle swarm optimization*, International Journal of u- and e- Service, Science and Technology, **8** (2015), 129-140.

[45] W. Qiu, P. Zhang, Y. Wang, *Fuzzy time series forecasting model based on automatic clustering techniques and generalized fuzzy logical relationship*, Mathematical Problems in Engineering, (2015), 1-8.

[46] N. F. Rahim, M. Othman, R. Sokkalingam, E. A. Kadir, *Forecasting crude palm oil prices using fuzzy rule-based time series method*, IEEE Access, **6** (2018), 32216-32224.

[47] P. M. Robinson, *Root-n-consistent semi-parametric regression*, Econometrica, **56** (1988), 931-954.

[48] A. Roy, *A novel multivariate fuzzy time series based forecasting algorithm incorporating the effect of clustering on prediction*, Soft Computing, **20** (2016), 1991-2019.

[49] S. Rubinstein, A. Goor, A. Rotshtein, *Time series forecasting of crude oil consumption using neuro-fuzzy inference*, Journal of Industrial and Intelligent Information, **3** (2015), 84-90.

[50] K. Sabzi, T. Allahviranloo, S. Abbasbandy, *A fuzzy generalized power series method under generalized Hukuhara differentiability for solving fuzzy Legendre differential equation*, Soft Computing, **24** (2020), 8763-8779.

[51] A. Sachdeva, V. Sharma, *A survey on stock forecasting model based on combined fuzzy genetic algorithm*, International Journal of Advanced Research in Computer Science and Software Engineering, **5** (2015), 397-401.

[52] H. J. Sadaei, R. Enayatifar, F. G. Guimaraes, M. Mahmud, Z. A. Alzamil, *Combining ARFIMA models and fuzzy time series for the forecast of long memory time series*, Neurocomputing, **175** (2016), 782-796.

[53] S. Sakhuja, V. Jain, S. Kumar, C. Chandra, S. K. Ghildayal, *Genetic algorithm based fuzzy time series tourism demand forecast model*, Industrial Management and Data Systems, **116** (2016), 1-24.

[54] S. Sharma, M. Chouhan, *A review: Fuzzy time series model for forecasting*, International Journal of Advanced Science and Technology, **2** (2014), 32-35.

[55] B. W. Silverman, *Density estimation*, Chapman and Hall, 1986.

[56] P. Singh, *A brief review of modeling approaches based on fuzzy time series*, International Journal of Machine Learning and Cybernetics, **8** (2017), 397-420.

[57] L. Stefanini, *A generalization of Hukuhara difference and division for interval and fuzzy arithmetic*, Fuzzy Sets and Systems, **161** (2010), 1564-1584.

[58] B. Sun, H. Guo, H. R. Karimi, Y. Ge, S. Xiong, *Prediction of stock index futures prices based on fuzzy sets and multivariate fuzzy time series*, Neurocomputing, **151** (2015), 1528-1536.

[59] F. M. Tseng, G. H. Tzeng, *A fuzzy seasonal ARIMA model for forecasting*, Fuzzy Sets and Systems, **126** (2002), 367-376.

[60] Y. Wang, Y. Lei, X. Fan, Y. Wang, *Intuitionistic fuzzy time series forecasting model based on intuitionistic fuzzy reasoning*, Mathematical Problems in Engineering, (2016), 1-12.

[61] L. Wasserman, *All of nonparametric statistics*, New York, Springer, 2007.

[62] U. Yolcu, *A new approach based on optimation of ratio for seasonal fuzzy time series*, Iranian Journal of Fuzzy Systems, **13** (2016), 19-36.

[63] U. Yolcu, E. Bas, *The forecasting of labour force participation and the unemployment rate in poland and turkey using fuzzy time series methods*, Comparative Economic Research, **19** (2016), 5-25.

[64] O. C. Yolcu, H. K. Lam, *A combined robust fuzzy time series method for prediction of time series*, Neurocomputing, **247** (2017), 87-101.

[65] R. Zarei, M. Gh. Akbari, J. Chachi, *Modeling autoregressive fuzzy time series data based on semi-parametricmethods*, Soft Computing, **24**(10) (2020), 7295-7304.

[66] K. Zhang, Z. Li, H. F. Wang, H. X. Wang, *Fuzzy time series prediction model and application based on fuzzy inverse*, International Journal of Signal Processing, Image Processing and Pattern Recognition, **8** (2015), 121-128.