



A new explainable robust high-order intuitionistic fuzzy time-series method

Cem Kocak¹ · Erol Egrioglu^{2,3} · Eren Bas²

Accepted: 23 July 2021
© Crown 2021

Abstract

Fuzzy time series, based on type-1 fuzzy sets, continue to have a wide range of use in the literature. These methods use only membership values to determine the fuzzy relations. However, intuitionistic fuzzy time series models use both membership values and non-membership values. So it can be considered that the use of intuitionistic fuzzy time forecasting models will be able to increase the forecasting performance because the intuitionistic fuzzy sets have more information than fuzzy sets. Therefore, intuitionistic fuzzy time series models have started to employ for forecasting the real-life series in the fuzzy time series literature. A novel, explainable, robust high-order intuitionistic fuzzy time series forecasting method is proposed based on a newly defined model. In the proposed method, the intuitionistic fuzzy c-means algorithm is used for the fuzzification of observations, and a robust regression method employed for determining fuzzy relations. With the use of robust regression in determining the fuzzy relationships, all inputs of the proposed method can be explainable and they can be tested and commented on statistically. Applications of this study are made by using energy data of Primary Energy Consumption between the years 1965 and 2016 for 23 countries in the region of Europe-Eurasia. The forecasting performance of the proposed method is compared with the performance of some selected benchmarks, and the obtained results are discussed.

Keyword Intuitionistic fuzzy time series · Intuitionistic fuzzy sets · Forecasting · Robust regression · Intuitionistic fuzzy c-means · Explainable artificial intelligence · Principal component analysis · Energy data forecasting

1 Introduction

Forecasting techniques continue to be largely used in the time series analysis at present. One of the most popular forecasting techniques is the autoregressive integrated moving average (ARIMA) (Kadilar 2005). The approach of Box-Jenkins (1976) is also frequently used as a basic approach in time series literature to obtain the best ARIMA model. In addition, exponential smoothing methods have

also been used in the literature for time series forecasting. Indeed, exponential smoothing models are special cases of ARIMA models. In classical time series, exponential smoothing methods are generally described under three headings as the simple exponential smoothing, Holt's exponential smoothing and Winter's exponential smoothing (Kadilar 2005). However, there are also new exponential smoothing methods proposed in the literature. One of these methods is ATA exponential smoothing method proposed by Yapar (2016). ATA exponential smoothing method (Yapar 2016; Yapar et al. 2017a, b) is an updated and effective time series analysis method proposed as an alternative to exponential smoothing methods in the literature.

Conventional forecasting techniques need assumptions like normality, stability and linearity because these methods are based on probability theory. However, many time series in real life do not satisfy any assumptions, and so these time series need to be solved via using some non-probabilistic methods in the literature. In the literature, non-probabilistic time series forecasting methods are

Communicated by Mu-Yen Chen.

✉ Erol Egrioglu
erol.egrioglu@giresun.edu.tr; e.egrioglu@lancaster.ac.uk

¹ Department of Statistics, Hitit University Faculty of Health Sciences/Nursing, Çorum, Turkey

² Faculty of Arts and Science, Department of Statistics, Giresun University, 28200 Giresun, Turkey

³ Department of Management Science, Lancaster University Management School, Marketing and Forecasting Research Center, Bailrigg LA1 4YX, UK

generally separated into three parts as fuzzy time series, artificial neural networks and other computational methods. Additionally, these non-probabilistic methods could be used together. For instance, the analysis of a fuzzy time series method includes an artificial neural network or a computational method. In this regard, the content of this study is generally about a new non-probabilistic method. The first fuzzy time series approach was proposed by Song and Chissom (1993a). In Song and Chissom (1993a), the analysis of a fuzzy time series consists generally of three phases. These phases are fuzzification, the determination of fuzzy relations and defuzzification, respectively. In the fuzzification phase, dividing the universe of discourse proposed by Song and Chissom (1993a, b), was used largely in literature. Additionally, there are some studies about being selected for the optimal interval length in the literature because determining the interval lengths affects the forecasting performance significantly in the fuzzification phase. In the literature, the first study of these studies was proposed by Huarng (2001) to determine the optimal interval length. Huarng (2001) proposed two different approaches based on distribution and average. Huarng and Yu (2006) also proposed a ratio-based method to determine the optimal interval length. The other studies to determine the optimal intervals were given by Yolcu et al. (2009), Egrioglu et al. (2010) and Egrioglu et al. (2011). In fuzzy time series literature, the centralization technique is mainly used in the defuzzification phase that is the final phase to solve the fuzzy time series. The determining of fuzzy relations is very important in the fuzzy time series methods. Therefore, this stage is effective to improve the forecasting performance of any fuzzy time series method. Song and Chissom (1993b) used the complex matrix operations to determine fuzzy relations as being based on the fuzzy set theory of Zadeh (1965). Then, Chen (1996) proposed a new method based on fuzzy relation tables instead of the complex matrix operations of Song and Chissom (1993b). Thereby, many fuzzy time series methods based on fuzzy relation tables have been proposed so far. In fuzzy time series literature, some methods that use fuzzy relation tables to determine fuzzy relations were given by Chen (2002), Chen and Chen (2011), Chen and Tanuwijaya (2011), Kocak (2013), Cheng et al. (2016) and Kocak (2017). Artificial neural networks (ANNs) are intensely used in fuzzy time series. Aladag et al. (2009) used the ANNs to determine fuzzy relations for the high-order fuzzy time series model in the fuzzy time series literature. Egrioglu et al. (2009a) also used the ANNs for the high-order multivariate time series method. Yu and Huarng (2010) proposed a new method based on the artificial neural network to define fuzzy relations. In fuzzy time series literature, some methods that use ANNs to determine fuzzy relations were given by Egrioglu et al. (2009b), Aladag

et al. (2010), Aladag (2013), Kocak (2015), Bas et al. (2015), Kocak (2013) and Bas et al. (2018). Also, many fuzzy time series methods have been proposed based on particle swarm optimization (PSO). In the fuzzy time series literature, Kuo et al. (2009) developed a new method named hybrid particle swarm optimization (MPSO) for solving the historical data of enrollments of the University of Alabama. Similarly, Hsu et al. (2010) used the modified PSO method and proposed a new method named modified turbulent particle swarm optimization (MTPSO) for forecasting temperature. Some methods that use PSO were also given by Kuo et al. (2010), Park et al. (2010), Huang et al. (2011), Aladag et al. (2012), Chen and Kao (2013), Singh and Borah (2014) and Cagcag Yolcu and Lam (2017). Besides, some studies use fuzzy c-means (FCM) to obtain membership values. Egrioglu et al. (2013) and Cagcag Yolcu and Alpaslan (2018) proposed different models based on FCM and PSO. There are also non-probabilistic and non-parametric time series forecasting methods based on the methods of machine learning, fuzzy inference system, artificial neural network, artificial intelligence, besides fuzzy time series methods. Some of these studies were given by Yolcu et al. (2013), Aladag et al. (2014), Zhou et al. (2016), Egrioglu et al. (2017) and Tak (2021). Yolcu et al. (2013) proposed a new artificial neural network based on PSO for solving linear and nonlinear time series. Aladag et al. (2014) proposed a new artificial neural network time series forecasting method based on the median neuron model. Zhou et al. (2016) predicted some financial time series by using a dendritic neuron model. Tak (2018) proposed a meta fuzzy functions approach for time series forecasting. Tak et al. (2018) proposed a recurrent type-1 fuzzy functions approach for time series forecasting. Egrioglu et al. (2019a) proposed a median-Pi ANN model for time series analysis. Tak (2020a) proposed a type-1 possibilistic fuzzy forecasting functions method for time series forecasting. Fuzzy time series uses only membership values for solving fuzzy time series. However, intuitionistic fuzzy time series based on intuitionistic fuzzy sets use both membership and non-membership values. Therefore, more information is used for the forecasting process. The first article about intuitionistic fuzzy time series (IFTS) was proposed by Zheng et al. (2013a). For this reason, the subject of IFTS can be thought a new and an updated field for time series literature. Fundamental definitions of IFTS were initially given by Zheng et al. (2013a). Furthermore, Zheng et al. (2013b) used the intuitionistic fuzzy c-means algorithm proposed by Chaira (2011) to determine membership and non-membership values. Lei et al. (2016) proposed a high-order intuitionistic fuzzy time series model by using an unequal partition of the universe of discourse based on the fuzzy clustering algorithm. Kumar and Gangwar (2016) proposed a method to determine

intuitionistic fuzzy relations. Fan et al. (2016) proposed a long-term intuitionistic fuzzy time series forecasting model. Egrioglu et al. (2019b) proposed an intuitionistic fuzzy time series forecasting method. Also, some other studies about intuitionistic fuzzy time series were proposed by Zheng et al. (2014), Wang et al. (2014), Joshi et al. (2016), Hu et al. (2017), Fan et al. (2017), Abhishekh et al. (2018) and Tak (2020b). A novel, explainable, robust

2 New definitions for intuitionistic fuzzy time series

Many definitions of fuzzy time series are based on Song and Chissom’s definition that is the first definition of fuzzy time series in the literature. However, in recent years, new definitions of fuzzy time series have been made, and many researchers have preferred to use new definitions in their studies. New definitions were explained in the study of Egrioglu et al. (2019b). The following new definitions are introduced below.

Definition 1 (New Definition: High-Order Single-Variable Intuitionistic Fuzzy Time Series Forecasting Model)

Let IF_t be an intuitionistic fuzzy time series. A_1, A_2, \dots, A_c are intuitionistic fuzzy sets on a universal set. $\mu_{A_j}(t), v_{A_j}(t)$ are membership and non-membership values of t th observation to j th intuitionistic fuzzy set. The high-order single-variable intuitionistic fuzzy time series model is given in Eq. (1).

$$X_t = G(\mu_{A_1}(t-1), \mu_{A_2}(t-1), \dots, \mu_{A_c}(t-1), \dots, \mu_{A_1}(t-p), \mu_{A_2}(t-p), \dots, \mu_{A_c}(t-p), v_{A_1}(t-1), v_{A_2}(t-1), \dots, v_{A_c}(t-1), \dots, v_{A_1}(t-p), v_{A_2}(t-p), \dots, v_{A_c}(t-p)) + \varepsilon_t \tag{1}$$

In Eq. (1), G is a linear or nonlinear function, ε_t is an error term with zero mean, $\mu_{A_1}(t-1), \dots, \mu_{A_c}(t-p)$ and $v_{A_2}(t-1), \dots, v_{A_c}(t-p)$ are lagged membership values and non-membership values, respectively, obtained from X_{t-1}, \dots, X_{t-p} . Equation (1) expresses that any time series are affected by lagged membership values and non-membership values.

Definition 2 (New Definition: High-Order Single-Variable Intuitionistic Fuzzy Time Series Forecasting Model based on the principal components analysis and the robust regression analysis)

Let IF_t be an intuitionistic fuzzy time series. A_1, A_2, \dots, A_c are intuitionistic fuzzy sets on a universal set. $\mu_{A_j}(t), v_{A_j}(t)$ are membership and non-membership values of t th observation to j th intuitionistic fuzzy set for

$j = 1, 2, \dots, c$. The high-order single-variable intuitionistic fuzzy time series model based on the principal components analysis is given in Eq. (2).

$$X_t = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \dots + \beta_q z_q + \varepsilon_t \quad \text{for } q \leq cp \tag{2}$$

where $z_i, i = 1, 2, \dots, q$ are principal components and are given in Eq. (3).

$$z_i = f_i(\mu_{A_1}(t-1), \mu_{A_2}(t-1), \dots, \mu_{A_c}(t-1), \dots, \mu_{A_1}(t-p), \mu_{A_2}(t-p), \dots, \mu_{A_c}(t-p), v_{A_1}(t-1), v_{A_2}(t-1), \dots, v_{A_c}(t-1), \dots, v_{A_1}(t-p), v_{A_2}(t-p), \dots, v_{A_c}(t-p)) = 1, 2, \dots, q \tag{3}$$

In Eq. (3), p is the number of lagged variables, $\mu_{A_1}(t-1), \dots, \mu_{A_c}(t-p)$ and $v_{A_1}(t-1), \dots, v_{A_c}(t-p)$ are lagged membership values and non-membership values obtained from X_{t-1}, \dots, X_{t-p} , q is the number of the important principal component for $q \leq cp$, f_i is i th linear function obtained by using the correlation matrix of the principal component analysis, z_i is the i th z-score calculated via the principal component analysis, $\beta_0, \beta_1, \beta_2, \dots, \beta_q$ are regression coefficients obtained via the robust regression method and ε_t is the error term with zero means.

3 Intuitionistic fuzzy C-means algorithm

Chaira (2011) proposed a clustering algorithm (IFCM) for intuitionistic fuzzy sets. The algorithms have been commonly used in intuitionistic fuzzy time-series studies (Zheng 2013a, b; Egrioglu et al. 2019b) in the literature. The algorithm of intuitionistic fuzzy c-means clustering was given in Algorithm 1.

Algorithm 1 The intuitionistic fuzzy c-means clustering algorithm.

Step 1. Calculate the membership values ($u_{ik}, i = 1, 2, \dots, c; k = 1, 2, \dots, n$) by using Eq. (6).

$$u_{ik} = \frac{r_{ik}}{\sum_{i=1}^c r_{ik}} \tag{4}$$

Suppose $x_k(k = 1, 2, \dots, n)$ is a time series with n observations. In Eq. (4), c is the number of clusters determined by the researcher, and $r_{ik}(i = 1, 2, \dots, c; k = 1, 2, \dots, n)$ are random values generated from the uniform distribution that has the parameters (0, 1).

Step 2. Calculate the hesitation degrees ($\pi_{ik}, i = 1, 2, \dots, c; k = 1, 2, \dots, n$) and intuitionistic fuzzy membership values ($u_{ik}^*, i = 1, 2, \dots, c; k = 1, 2, \dots, n$) by using Eqs. (5) and (6), respectively. Thereafter, save the intuitionistic membership values (u_{ik}^*) into a matrix named U_{old} .

$$\pi_{ik} = 1 - u_{ik} - (1 - u_{ik}^\alpha)^{1/\alpha}, \alpha > 0 \tag{5}$$

$$u_{ik}^* = u_{ik} + \pi_{ik} \tag{6}$$

Step 3. Calculate centres of the clusters ($v_i^*, i = 1, 2, \dots, c$) of the intuitionistic membership values (u_{ik}^*) by using Eq. (7).

$$v_i^* = \frac{\sum_{k=1}^n (u_{ik}^*)^f x_k}{\sum_{k=1}^n (u_{ik}^*)^f}; i = 1, 2, \dots, c \tag{7}$$

In Eq. (7), $x_k (k = 1, 2, \dots, n)$ are observations of the time series the and f is the fuzziness index.

Step 4. Update the membership values ($u_{ik}, i = 1, 2, \dots, c; k = 1, 2, \dots, n$) by using Eqs. (8)–(9).

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ik}}{d_{jk}}\right)^{2/(f-1)}}; i = 1, 2, \dots, c; k = 1, 2, \dots, n \tag{8}$$

$$d_{ik} = \sqrt{(x_k - v_i^*)^2} \tag{9}$$

In Eq. (8), $d_{ik} (i = 1, 2, \dots, c; k = 1, 2, \dots, n)$ is the Euclidean distance measure between i th cluster centre and k th observation.

Step 5. Go to Step 2 and update the hesitation degrees ($\pi_{ik}, i = 1, 2, \dots, c; k = 1, 2, \dots, n$) and fuzzy membership values ($u_{ik}^*, i = 1, 2, \dots, c; k = 1, 2, \dots, n$) by using u_{ik} values obtained in Step 4. Thereafter, save the new intuitionistic membership values (u_{ik}^*) into a matrix named U_{new} .

Step 6. Calculate non-membership values ($v_{ik}, i = 1, 2, \dots, c; k = 1, 2, \dots, n$) by using Eq. (10). Thereafter, save the new intuitionistic membership values (u_{ik}^*) into a matrix named V .

$$u_{ik}^* + \pi_{ik} + v_{ik} = 1 \tag{10}$$

In Eq. (10), u_{ik}^* and π_{ik} are intuitionistic membership values and hesitation degrees obtained in Step 5.

Step 7. Check the stopping criteria given in Eq. (11). If the condition is satisfied, stop the algorithm, otherwise make $U_{\text{old}} = U_{\text{new}}$ and go to Step 3. ε is a small positive number and $\|\cdot\|_2$ is the L_2 norm.

$$\|U_{\text{new}} - U_{\text{old}}\|_2 < \varepsilon \tag{11}$$

In Eq. (11), ε is a small positive number and $\|\cdot\|_2$ is the L_2 norm.

4 The robust regression

The robust regression is commonly used instead of the least-square method in the literature because robust estimators are not affected by outliers too much. Robust regression techniques can be given as the L1 technique of

Edgeworth (1887), M-Estimates of Mallows (1975), and S-Estimates of Rousseeuw (1984). Robust regression algorithms based on these fundamental robust techniques have been proposed in the literature. One of these algorithms is the iteratively reweighted least-squares technique (James et al. 1988). In this study, robust regression estimates were used iteratively reweighted the least-squares technique in defining the fuzzy relation phase. Accordingly, the algorithm of the iteratively reweighted least squares is given in Algorithm 2.

Algorithm 2 Iteratively reweighted least-squares algorithm.

Step 1. Calculate initial regression coefficient estimates $\hat{\beta}$ of β via the least-square method for linear regression model given in Eq. (12) and calculate standard deviation estimate $\hat{\sigma}$ of σ by using Eq. (13).

$$Y = X_{n*(p+1)}\beta_{(p+1)*1} + \varepsilon_{n*1} \tag{12}$$

$$\hat{\sigma} = \frac{\text{Median}(|e_i|)}{0.6745}; i = 1, 2, \dots, n \tag{13}$$

In Eq. (12), n is the observation number, p is the independent variable number, Y is the dependent variable vector, and X is the independent variable matrix that values of its first column are 1, ε is an error parameter vector, e_i is an estimate of ε_i for i th observation obtained via the least and 0.6745 is the constant used to make estimates for normal distribution.

Step 2. Calculate residuals ($e_i; i = 1, 2, \dots, n$) by using Eq. (14).

$$e_i = y_i - x_i' \hat{\beta} \tag{14}$$

In Eq. (14), y_i is i th observation value of the dependent variable and x_i' is a row vector that is the transpose of the first column of X matrix.

Step 3. Calculate weight values ($w_i; i = 1, 2, \dots, n$) via Eq. (15) by using bisquare weight function (W) given in Eq. (16).

$$w_i = W\left(\frac{e_i}{\hat{\sigma}}\right) \tag{15}$$

$$W(e) = \begin{cases} \left[1 - \left(\frac{e}{4.685}\right)^2\right]^2, & \text{for } |e| \leq 4.685 \\ 0, & \text{for } |e| > 4.685 \end{cases} \tag{16}$$

In Eq. (15), $\hat{\sigma}$ is standard deviation estimates and 4.685 is the tuning constant for the bisquare distribution.

Step 4. Calculate $\hat{\beta}_{\text{new}}$ by using Eq. (17).

$$\sum_{i=1}^n w_i x_i (y_i - x_i' \hat{\beta}_{\text{new}}) = 0 \tag{17}$$

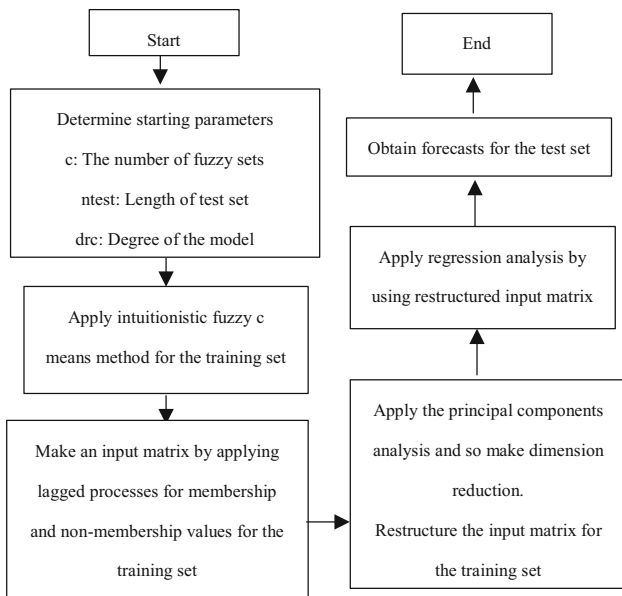


Fig. 1 Flowchart of the proposed method

Step 5. Calculate $e_{i(new)} (i = 1, 2, \dots, n)$ by using Eq. (18).

$$e_{i(new)} = y_i - x_i \hat{\beta}_{new} \tag{18}$$

Step 6. Check the stopping criteria given in Eq. (19). If the condition is satisfied, stop the algorithm, otherwise make $\hat{\beta} = \hat{\beta}_{new}$ and go to Step 2.

$$\max_i (|e_i - e_{i(new)}|) < \varepsilon \tag{19}$$

In Eq. (19), ε is a small positive number and $|\cdot|$ is the absolute value.

5 The proposed method

In this study, a new intuitionistic, high-order fuzzy time series forecasting method was proposed. The proposed method generally uses Eq. (3) given in Definition 1 and completely uses Eqs. (4)–(5) given in Definition 2. Thereby, both membership and non-membership are used in the intuitionistic fuzzy time series approach, unlike the fuzzy time series approach that uses only non-membership values. In the proposed method, the intuitionistic fuzzy c-means clustering method is used in the fuzzification phase. Furthermore, the principal component analysis is used to eliminate the multicollinearity problem and reduce dimension because there are many inputs before the phase of determining fuzzy relationships of the proposed method. The proposed method is based on the robust regression. For this reason, the iteratively reweighted least-squares technique is a robust regression method that is used to define

functional relationships in the stage of determining the fuzzy relationships. Also, the proposed method does not need the defuzzification step because the outputs of the proposed method are real forecasts. The flowchart and algorithm of the proposed method is given in Fig. 1 and Algorithm 3, respectively.

Algorithm 3 Algorithm of the proposed method.

Step 1. Determine $X_t, X_t^{test}, X_t^{train}, n, n_{test}, n_{train}, c$ and $m (1, 2, 3, \dots)$ are data and parameters used in the application of the proposed method. Here, X_t is the time series, X_t^{test} is the test set of X_t, X_t^{train} is the training test of X_t, n is the length of X_t, n_{test} (the last part of %5, %10 or %15 of n) is the length of X_t^{test}, n_{train} is the length of $X_t^{train}, c (2, 3, \dots)$ is the number of fuzzy sets and $m (1, 2, 3, \dots)$ is the order of the model.

Step 2. Calculate membership and non-membership for the training time series (X_t^{train}) by using the intuitionistic fuzzy c-means (IFCM) clustering method via Algorithm 1 given in Sect. 3.

Here, the matrix of memberships named $U(U_{ij}; i = 1, 2, \dots, n_{train}; j = 1, 2, \dots, c)$, the matrix of non-membership named $V(V_{ij}; i = 1, 2, \dots, n_{train}; j = 1, 2, \dots, c)$ and the vector of centres of clusters named $v^*(v_i^*; i = 1, 2, \dots, c)$ are obtained. Also, the intuitionistic fuzzy time series matrix ($IF_t; i = 1, 2, \dots, c; j = 1, 2, \dots, (2c + 1)$) is constituted by combining of X_t^{train}, U and V , respectively.

For example, Let $X_t = \{8, 10, 11, 12, 16, 13, 14\}, X_t^{test} = \{14\}, X_t^{train} = \{8, 10, 11, 12, 16, 13\}, c = 3, m = 2$ and $n_{test} = 1$ and $n_{train} = 6$. Also, let $\mu_{A_i}(t), \nu_{A_i}(t)$ are membership and non-membership values of t th observation to i th intuitionistic fuzzy set. U, V, v^* and IF_t calculated via Algorithm 1 are given below for $X_t^{train} = \{8, 10, 11, 12, 16, 13\}$.

$$U = \{\mu_{A_1}(t), \mu_{A_2}(t), \mu_{A_3}(t)\} = \begin{bmatrix} 0.0314 & 0.9630 & 0.1021 \\ 0.0482 & 0.8408 & 0.3144 \\ 0.0574 & 0.2494 & 0.8778 \\ 0.0017 & 0.0025 & 0.9988 \\ 0.9985 & 0.0015 & 0.0037 \\ 0.1878 & 0.1056 & 0.8907 \end{bmatrix} \tag{20}$$

$$V = \{\nu_{A_1}(t), \nu_{A_2}(t), \nu_{A_3}(t)\} = \begin{bmatrix} 0.9513 & 0.0030 & 0.8527 \\ 0.9270 & 0.0765 & 0.5897 \\ 0.9139 & 0.6664 & 0.0504 \\ 0.9971 & 0.9957 & 0.0012 \\ 0.0015 & 0.9974 & 0.9937 \\ 0.7420 & 0.8480 & 0.0419 \end{bmatrix} \tag{21}$$

$$v^* = \begin{bmatrix} 15.8612 \\ 8.9699 \\ 11.9189 \end{bmatrix} \quad (22)$$

$$\begin{aligned} \text{IF}_t &= \{U, V\} \\ &= \begin{bmatrix} 0.0314 & 0.9630 & 0.1021 & 0.9513 & 0.0030 & 0.8527 \\ 0.0482 & 0.8408 & 0.3144 & 0.9270 & 0.0765 & 0.5897 \\ 0.0574 & 0.2494 & 0.8778 & 0.9139 & 0.6664 & 0.0504 \\ 0.0017 & 0.0025 & 0.9988 & 0.9971 & 0.9957 & 0.0012 \\ 0.9985 & 0.0015 & 0.0037 & 0.0015 & 0.9974 & 0.9937 \\ 0.1878 & 0.1056 & 0.8907 & 0.7420 & 0.8480 & 0.0419 \end{bmatrix} \end{aligned} \quad (23)$$

PCA is applied to Z_{input} . And so, PCC and PC matrix are calculated. Also, important principal components are selected from principal components by taking into consideration the criterion that the explained variance ratio needs to be at least 0.85. In this way, IPC matrix is obtained. Lastly, the new input matrix named P_{input} is obtained by putting a vector that consists of 1 value to the first column of IPC.

For example, PCC, PC, IPC and P_{input} matrices obtained by applying PCA to X_{input} in Eq. (25) are given below.

$$\text{PCC}^T = \begin{bmatrix} 0.3408 & -0.2341 & -0.2191 & -0.2722 & -0.3262 & 0.3051 & -0.3375 & 0.2447 & 0.2465 & 0.2765 & 0.3362 & -0.2889 \\ -0.1459 & -0.3597 & 0.3900 & 0.3239 & -0.1962 & 0.2564 & 0.1551 & 0.3544 & -0.3568 & -0.3169 & 0.1599 & -0.2918 \\ -0.3145 & 0.4443 & -0.1921 & 0.1181 & -0.3364 & 0.2708 & 0.3473 & -0.3256 & 0.2305 & -0.1198 & 0.3505 & -0.2223 \end{bmatrix} \quad (25)$$

Step 3. Apply lag process for U and V separately. And so, obtain input matrix (X_{input}) that includes lagged variables of membership and non-membership under Eq. (5) given in Definition 2.

For example, X_{input} matrix obtained by using U and V in Eqs. (20) and (21) is given in Eq. (24).

$$\text{PC} = \begin{bmatrix} -2.0114 & -2.2301 & 0.3612 \\ -1.7560 & 0.4749 & -0.6783 \\ -0.2017 & 2.5459 & 0.4131 \\ 3.9691 & -0.7907 & -0.0960 \end{bmatrix} \quad (26)$$

$$X_{\text{input}} = \begin{bmatrix} 0.0482 & 0.8408 & 0.3144 & 0.0314 & 0.9630 & 0.1021 & 0.9270 & 0.0765 & 0.5897 & 0.9513 & 0.0030 & 0.8527 \\ 0.0574 & 0.2494 & 0.8778 & 0.0482 & 0.8408 & 0.3144 & 0.9139 & 0.6664 & 0.0504 & 0.9270 & 0.0765 & 0.5897 \\ 0.0017 & 0.0025 & 0.9988 & 0.0574 & 0.2494 & 0.8778 & 0.9971 & 0.9957 & 0.0012 & 0.9139 & 0.6664 & 0.0504 \\ 0.9985 & 0.0015 & 0.0037 & 0.0017 & 0.0025 & 0.9988 & 0.0015 & 0.9974 & 0.9937 & 0.9971 & 0.9957 & 0.0012 \end{bmatrix} \quad (24)$$

Step 4. Principal component analysis (PCA) is applied to X_{input} matrix obtained in *Step 3*. And so, obtain the matrix of principal component coefficients (PCC), the matrix of principal component (PC), the matrix of important principal component (IPC) and the new input matrix (P_{input}), respectively.

Firstly, elements of X_{input} matrix are transformed to the standard normal matrix named Z_{input} that means and standard deviations of each column of Z_{input} are 0 and 1. Then,

$$\text{IPC} = \begin{bmatrix} -2.0114 & -2.2301 \\ -1.7560 & 0.4749 \\ -0.2017 & 2.5459 \\ 3.9691 & -0.7907 \end{bmatrix} \quad (27)$$

$$P_{\text{input}} = \begin{bmatrix} 1 & -2.0114 & -2.2301 \\ 1 & -1.7560 & 0.4749 \\ 1 & -0.2017 & 2.5459 \\ 1 & 3.9691 & -0.7907 \end{bmatrix} \quad (28)$$

In Eq. (25), “^T” indicates transpose operation. The explained variance ratio of two important principal components is %97.86. So, it can be said that the aims of dimension reduction and obtaining orthogonal inputs are accomplished via PCA.

Step 5. Define fuzzy relations by using robust Regression.

Robust regression coefficient estimates ($\hat{\beta}$) is calculated via iteratively reweighted least-squares algorithm given in *Algorithm 2* by taking $X = P_{input}$ and $Y = X_t^{train}$ in step 1 of *Algorithm 2*. Then, predictions of the training set (\hat{X}_t^{train}) can be calculated by using Eq. (29), if any researcher aims to predict.

$$\hat{X}_t^{train} = P_{input} \hat{\beta} \tag{29}$$

For example, $\hat{\beta}$ obtained via *Algorithm 2* by taking $X = P_{input}$ in (29) and $Y = X_t^{train} = [-, -, 11, 12, 16, 13]$ is given below.

$$\hat{\beta} = \begin{bmatrix} 13.000 \\ 0.2257 \\ 0.9445 \end{bmatrix} \tag{30}$$

Step 6. Apply determining membership values for time series (X_t) to obtain new memberships (U_{new}) and new non-membership (V_{new}) of X_t by using centres of clusters (v^*) obtained in *Step 1*.

For example, U_{new} and V_{new} for $X_t = \{8, 10, 11, 12, 16, 13, 14\}$ by using v^* calculated for $X_t^{train} = \{8, 10, 11, 12, 16, 13\}$ are given below.

$$U_{new} = \begin{bmatrix} 0.0314 & 0.9630 & 0.1021 \\ 0.0482 & 0.8408 & 0.3144 \\ 0.0574 & 0.2494 & 0.8778 \\ 0.0017 & 0.0025 & 0.9988 \\ 0.9985 & 0.0015 & 0.0037 \\ 0.1878 & 0.1056 & 0.8907 \\ 0.6269 & 0.1226 & 0.5278 \end{bmatrix} \tag{31}$$

$$V_{new} = \begin{bmatrix} 0.9513 & 0.0030 & 0.8527 \\ 0.9270 & 0.0765 & 0.5897 \\ 0.9139 & 0.6664 & 0.0504 \\ 0.9971 & 0.9957 & 0.0012 \\ 0.0015 & 0.9974 & 0.9937 \\ 0.7420 & 0.8480 & 0.0419 \\ 0.2570 & 0.8256 & 0.3573 \end{bmatrix} \tag{32}$$

Step 7. Apply lag process for U_{new} and V_{new} obtained in Step 6 and obtain new input matrix (X_{input}^{new}), the new standardized input matrix (Z_{input}^{new}) and standardized input matrix of the test set (Z_{input}^{test}) that includes z scores of lagged variables of memberships and non-memberships.

Firstly, the new input matrix (X_{input}^{new}) that includes lagged variables of memberships and non-memberships is obtained by using U_{new} and V_{new} . Then, elements of X_{input}^{new} matrix is transformed to standard normal values, and so, Z_{input}^{new} matrix has obtained that means and standard deviations of each column of Z_{input}^{new} are 0 and 1. Lastly, Z_{input}^{test} that includes the last some rows of Z_{input}^{new} is obtained.

For example, X_{input}^{new} and matrix obtained by using U_{new} and V_{new} obtained for $X_t = \{8, 10, 11, 12, 16, 13, 14\}$, $c = 3$ and $m = 2$ is given below:

$$X_{input}^{new} = \begin{bmatrix} 0.0482 & 0.8408 & 0.3144 & 0.0314 & 0.9630 & 0.1021 & 0.9270 & 0.0765 & 0.5897 & 0.9513 & 0.0030 & 0.8527 \\ 0.0574 & 0.2494 & 0.8778 & 0.0482 & 0.8408 & 0.3144 & 0.9139 & 0.6664 & 0.0504 & 0.9270 & 0.0765 & 0.5897 \\ 0.0017 & 0.0025 & 0.9988 & 0.0574 & 0.2494 & 0.8778 & 0.9971 & 0.9957 & 0.0012 & 0.9139 & 0.6664 & 0.0504 \\ 0.9985 & 0.0015 & 0.0037 & 0.0017 & 0.0025 & 0.9988 & 0.0015 & 0.9974 & 0.9937 & 0.9971 & 0.9957 & 0.0012 \\ 0.1878 & 0.1056 & 0.8907 & 0.9985 & 0.0015 & 0.0037 & 0.7420 & 0.8480 & 0.0419 & 0.0015 & 0.9974 & 0.9937 \end{bmatrix} \tag{33}$$

$$Z_{input}^{new} = \begin{bmatrix} -0.502 & 1.712 & -0.696 & -0.454 & 1.196 & -0.788 & 0.513 & -1.673 & 0.577 & 0.455 & -1.126 & 0.781 \\ -0.480 & 0.027 & 0.600 & -0.415 & 0.931 & -0.320 & 0.482 & -0.132 & -0.674 & 0.398 & -0.974 & 0.203 \\ -0.613 & -0.677 & 0.878 & -0.394 & -0.352 & 0.923 & 0.684 & 0.729 & -0.758 & 0.367 & 0.245 & -0.983 \\ 1.764 & -0.680 & -1.411 & -0.523 & -0.887 & 1.190 & -1.742 & 0.733 & 1.494 & 0.563 & 0.926 & -1.092 \\ -0.169 & -0.383 & 0.629 & 1.787 & -0.889 & -1.005 & 0.063 & 0.343 & -0.666 & -1.784 & 0.930 & 1.091 \end{bmatrix} \tag{34}$$

$$Z_{input}^{test} = [-0.169 \quad -0.383 \quad 0.629 \quad 1.787 \quad -0.889 \quad -1.005 \quad 0.063 \quad 0.343 \quad -0.666 \quad -1.784 \quad 0.930 \quad 1.091] \tag{35}$$

In Eq. (35), Z_{input}^{test} is the last row of Z_{input}^{new} given in Eq. (34) because $n_{test} = 1$.

Step 8. Calculate the matrix of principal component (PC_{test}), the matrix of important principal component (IPC_{test}) and the new input matrix (P_{input}^{test}) for X_t^{test} by using Z_{input}^{test} obtained in Step 7.

Firstly, PC_{test} is calculated by using Eq. (36). In Eq. (36), PCC is the matrix of principal component coefficients calculated in Step 4. Also, important principal components are selected from principal components by taking into consideration the criterion that the explained variance ratio needs to be at least 0.85. In this way, IPC_{test} matrix is obtained. Lastly, the new input matrix (P_{input}^{test}) is obtained by putting a vector that consists of 1 value to the first column of IPC.

$$PC_{test} = Z_{input}^{test} * PCC \tag{36}$$

For example, PC_{test} , IPC_{test} and P_{input}^{test} calculated by using PCC and Z_{input}^{test} are given below for $X_t^{test} = \{14\}$.

$$PC_{test} = [-1.2061 \quad 1.6675 \quad 0.0540] \tag{37}$$

$$IPC_{test} = [-1.2061 \quad 1.6675] \tag{38}$$

$$P_{input}^{test} = [1 \quad -1.2061 \quad 1.6675] \tag{39}$$

Step 9. Calculate the forecasts of $X_t^{test}(\hat{X}_t^{test})$ by using $\hat{\beta}$ and P_{input}^{test} obtained Step 5 and Step 8 via Eq. (40).

$$\hat{X}_t^{test} = P_{input}^{test} * \hat{\beta} \tag{40}$$

For example, \hat{X}_t^{test} calculated by using $\hat{\beta}$ and P_{input}^{test} obtained in Eqs. (30) and (39) are given below for $X_t^{test} = \{14\}$.

$$\hat{X}_t^{test} = [14.3028] \tag{41}$$

Step 10. The root of mean square error (RMSE) and mean absolute percentage error (MAPE) values are calculated by using \hat{X}_t^{test} obtained in Step 9 via Eqs. (42) and (43).

$$RMSE = \sqrt{\frac{1}{n_{test}} \sum_{t=1}^{n_{test}} (X_t^{test} - \hat{X}_t^{test})^2} \tag{42}$$

$$MAPE = \frac{1}{n_{test}} \sum_{t=1}^{n_{test}} \left| \frac{X_t^{test} - \hat{X}_t^{test}}{X_t^{test}} \right| \tag{43}$$

For example, RMSE and MAPE calculated for $X_t^{test} = \{14\}$ by using \hat{X}_t^{test} are given in Eq. (43).

$$RMSE = 0.3028 \text{ and } MAPE = 0.0216 \tag{44}$$

6 Applications

Energy data taken from the ministry of energy and natural resources of Turkey consists of 69 variables of primary energy consumption (PEC). Each of 69 PEC time series has 52 yearly observations between the years 1965 and 2016 for each country in the world. These 69 countries are universally divided into six continents called North America, South-Central America, Europe-Eurasia, Middle East, Africa, and the Asia Pacific, respectively.

Initially, the energy data of the Europe-Eurasia continent that consists of PEC variables of 32 countries were selected to make applications. However, 23 variables of these 32 variables have complete data that includes 52 observations because PEC variables of the other nine countries have some missing observations. For this reason, the proposed method and some other forecasting methods in the literature are applied to the 24 PEC variables that include PEC variables of 24 countries in Europe-Eurasia and the PEC variable of the total Europe-Eurasia continent. So, it is aimed to compare the forecasting performances of the proposed method with some alternative forecasting methods for the entire Europe-Eurasia continent. The names of 24 countries are Austria, Belgium, Bulgaria, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Spain, Sweden, Switzerland, Turkey, United Kingdom and another Europe-Eurasia.

Some of the distributions of the 25 PEC time series used in the application are similar and the others are different. When the 24-time series graphs are viewed in general, it can be seen that the 24-time series may have characteristics

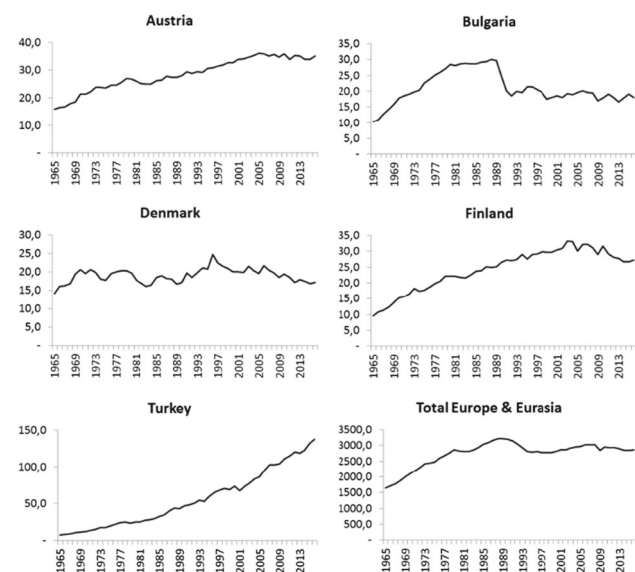


Fig. 2 Yearly PEC observations of some countries in the million tonnes of oil equivalent

such as linear increasing trend, curvilinear increasing trend, and stationary or random fluctuations. Some graphs of 24 PEC time series that differ from each other are given in Fig. 2.

The proposed method is compared with the following benchmarks.

DN-PSO Method: The Dendritic neuron (DN) method is an artificial intelligence model proposed by Zhou et al. (2016) to predict financial time series. Particle swarm optimization (PSO) is used in the training algorithm of the DN model and a Matlab m-file” is prepared for solving PEC time series with DN-PSO method.

ANN-PSO Method: This method is a multilayer perceptron artificial neural network (ANN) model trained by PSO used in Yolcu et al. (2013) for time series forecasting. A Matlab m-file” is prepared for solving PEC time series with the ANN-PSO method.

Median ANN Method: This method is a multilayer artificial neural network (ANN) model based on Median Neuron Model proposed by Aladag et al. (2014) for time series forecasting. A Matlab m-file” is prepared for solving PEC time series with the Median ANN method.

Median-Pi-ANN Method: The Median-Pi artificial neural network (Median-Pi-ANN) is a time series forecasting method proposed by Egrioglu et al. (2017). A Matlab m-file” is prepared for solving PEC time series with the Median-Pi-ANN method.

FTS-N: This method is a fuzzy time series network (FTS-N) model proposed by Bas et al. (2015) to forecast the linear and nonlinear time series. A Matlab m-file” is prepared for solving PEC time series with the FTS-N method.

SC-FTS: Song and Chissom (SC) (1993a) method is the first fuzzy time series (FTS) forecasting model in the fuzzy time series literature. A Matlab m-file” is prepared for solving PEC time series with the SC-FTS method.

C-FTS: This method is a fundamental fuzzy time series (FTS) model based on fuzzy logic relations proposed by Chen (2002) (C). A Matlab m-file” is prepared for solving PEC time series with the C-FTS method.

ARIMA Method: Autoregressive integrated moving average (ARIMA) method is a classic time series model that is often preferred by many researchers for time series analysis Kadilar (2005). Box-Jenkins’s (1976) approach is an effective method that provides to be obtained the best ARIMA model. A Matlab m-file” is prepared for solving PEC time series with Box-Jenkins’s (1976) approach.

H-ES Method: Holt’s exponential smoothing (H-ES) method is one of the exponential smoothing methods for time series analysis Kadilar (2005). A Matlab m-file” is prepared for solving PEC time series with the H-ES method.

A-ES Method: Ata’s exponential smoothing (A-ES) method proposed by Yapar (2016) is a new exponential smoothing method in the time series literature. A Matlab m-file” is prepared for solving PEC time series with the A-ES method.

How applications are made with the proposed method is given below.

- Each of the 24 PEC time series has 52 observations. The first 47 and the last five observations of the time series are used for the training set and the test set, respectively. In other words, the length of the test set is determined as 5 for all applications.
- The number of fuzzy sets (c) is tried between 3 and 10 increasing 1 unit.
- The degree of the model (m) is tried between 2 and 5 increasing 1 unit.
- The architecture that has minimum RMSE value is selected as the best architecture, and so, the best forecasts are obtained for the test set.
- 30 runs are executed using random values (rn), and so, some statistics are calculated for RMSE and MAPE values.

How applications are made with alternative method is given below.

- The methods of DN-PSO Zhou et al. (2016), ANN-PSO proposed by Yolcu et al. (2013), Median ANN proposed by Aladag et al. (2014), Median-Pi-ANN proposed by Egrioglu et al. (2017) and FTS-N proposed by Bas et al. (2015) were applied like those applications of the proposed method above-mentioned.
- In applications of the SC-FTS proposed by Song and Chissom (1993a) method and C-FTS method proposed by Chen (2002), the number of fuzzy sets is tried between 3 and 10 increasing 1 unit.
- The methods of the ARIMA method Box-Jenkins’s (1976), Holt’s exponential smoothing (Kadilar (2005)) and ATA method proposed by Yapar (2016) method were applied as classic time series methods.

In the applications, DN-PSO method proposed by Zhou et al. (2016), ANN-PSO method proposed by Yolcu et al. (2013), Median ANN method proposed by Aladag et al. (2014), Median-Pi-ANN method, FTS-N method and the proposed method were the methods that can be solved by trying different random initial parameters. For this reason, for each PEC time series, these methods were tried 30 times by using 30 different random seed values, and so, 30 different RMSE and MAPE values were obtained for each time series in the applications of this study. The mean and median of these RMSE values are given in Table 1. When Table 1 is examined, it is seen that the RMSE statistics of the proposed method are smaller than the RMSE statistics

Table 1 Means and medians of RMSE values of the methods for PEC variables of the countries in the Europe & Eurasia continent

Methods	DN-PSO method		ANN-PSO method		Median ANN method		Median-Pi-ANN method		FTS-N method		Proposed method	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Austria	3.386	3.774	0.929	0.894	1.906	1.109	0.889	0.830	1.339	1.322	1.103	1.157
Belgium	4.199	4.223	3.904	3.895	2.419	2.432	2.420	2.430	3.989	3.815	3.050	2.929
Bulgaria	1.625	1.471	1.377	1.358	2.312	2.503	1.080	1.083	1.638	1.695	0.744	0.681
Czech Republic	1.273	1.085	1.273	1.085	1.692	1.180	1.212	1.206	0.940	0.881	0.555	0.555
Denmark	1.241	1.201	0.870	0.857	1.345	1.373	1.121	1.046	1.273	1.312	0.942	0.830
Finland	2.042	2.007	1.694	1.675	0.536	0.499	1.249	1.232	1.476	1.543	1.449	1.470
France	6.501	5.792	9.903	10.33	4.651	4.687	8.097	7.863	8.044	7.160	3.405	3.108
Germany	9.665	9.205	9.639	8.804	9.010	7.862	8.141	8.191	8.775	7.616	9.702	5.568
Greece	2.502	2.134	4.155	4.440	1.030	0.977	2.464	2.326	1.449	1.206	1.580	1.540
Hungary	2.469	2.466	1.387	1.257	1.190	1.057	1.020	0.877	0.952	0.954	0.731	0.706
Ireland	0.741	0.699	1.185	1.226	0.772	0.635	1.626	0.718	1.210	1.176	0.339	0.339
Italy	6.145	5.500	11.87	11.80	6.597	6.889	9.876	9.225	8.003	8.268	6.390	6.210
Netherlands	3.894	3.936	4.116	3.985	3.514	3.128	3.602	3.603	4.574	3.839	3.317	3.114
Norway	4.742	4.721	3.961	4.047	5.825	5.225	3.068	3.103	1.950	1.732	4.507	4.839
Poland	5.421	5.536	4.998	4.537	4.434	3.184	2.437	2.388	3.680	3.628	2.048	1.923
Portugal	1.417	1.269	1.238	1.232	2.202	1.477	1.293	1.261	1.865	1.791	0.928	0.793
Romania	5.144	4.547	2.796	2.819	3.261	2.334	2.059	1.394	1.955	1.951	3.668	4.035
Slovakia	1.104	1.081	1.070	1.015	1.345	1.373	0.839	0.834	0.678	0.743	1.021	1.014
Spain	11.38	9.169	5.883	6.087	8.795	6.758	7.124	6.297	6.478	6.233	4.159	4.300
Sweden	2.468	2.641	1.729	1.734	1.472	1.478	1.405	1.439	2.013	1.859	1.196	1.294
Switzerland	1.937	1.992	1.106	1.109	1.156	1.157	1.135	1.102	1.527	1.573	0.982	1.011
Turkey	36.03	33.58	24.13	24.36	37.05	37.02	21.93	23.92	4.865	4.972	39.27	36.27
United Kingdom	11.37	10.76	7.866	7.702	12.56	12.85	12.64	12.51	8.074	7.909	5.039	5.040
Other	3.923	2.633	4.678	4.002	4.708	2.989	3.184	2.767	3.427	3.174	2.648	2.648
Europe & Eurasia continent	44.66	45.72	57.77	59.36	36.94	35.53	56.24	56.44	66.89	70.60	12.58	12.58

of the other methods for the vast majority of 25 PEC time series solved. Therefore, it can be said that the proposed method is largely better than other methods in forecasting the future.

In the applications, for each of 25 PEC time series, minimum RMSE values of all methods used in this study are given in Table 2. When Table 2 is examined, it is seen that minimum RMSE values of the proposed method are smaller than minimum RMSE values of the other methods for the vast majority of 25 PEC time series solved. Therefore, it can be said that the proposed method has a largely better forecasting performance than other methods.

In addition, success rates obtained by examining Tables 1 and 2 were given in Table 3. When Table 3 is examined, it is seen that the proposed method has the highest success rates with 56%, 64% and 44% for minimum RMSE means minimum RMSE medians and minimum RMSE values, respectively. For instance, according to Table 3, 16 RMSE medians ($16/25 = 64\%$) of 25 RMSE

medians of the proposed method are smaller than RMSE medians of the other methods. Other methods have quite lower success rates than the success rates of the proposed method. Therefore, it can be said that the proposed method has a largely better forecasting performance than other methods according to RMSE measurement. Similarly, the proposed method has also the highest success rates with 44%, 44% and 32% for minimum MAPE means, minimum MAPE medians and minimum MAPE values, respectively. So, it can be said that the proposed method has a largely better forecasting performance than other methods according to MAPE measurement. However, the success rates with scores of 36% and 32% of ANN-PSO method proposed by Yolcu et al. (2013) for MAPE means and median means, respectively, and the success rate 24% of FTS-N method proposed by Bas et al. (2015) for minimum MAPE values are remarkable high values. Thereby, as a result, it can be said that the proposed method is the best, ANN-PSO method Yolcu et al. (2013) is the second-best

Table 2 Minimum RMSE values of the methods for PEC variables of the countries in the Europe & Eurasia continent

Methods Countries	DN-PSO method	ANN-PSO method	Median ANN method	Median-Pi-ANN method	FTS-N method	SC-FTS method	C-FTS method	ARIMA method	H-ES method	A-ES method	Proposed method
Austria	0.966	0.680	1.226	0.671	0.942	0.872	0.982	1.054	1.008	1.015	0.673
Belgium	2.283	2.643	2.063	2.39	2.778	2.981	5.284	3.188	3.184	3.672	2.929
Bulgaria	0.712	0.857	0.911	1.006	1.230	0.988	1.165	1.057	1.179	1.092	0.665
Czech Republic	0.584	0.584	0.570	0.721	0.562	0.883	0.996	0.730	0.776	0.615	0.555
Denmark	0.636	0.550	0.920	0.828	0.846	0.699	0.699	0.664	0.812	1.102	0.828
Finland	0.749	0.259	0.381	1.201	0.665	1.241	2.083	0.824	1.012	1.607	1.360
France	3.549	4.174	2.976	3.771	3.245	4.400	4.400	5.445	6.136	5.051	2.947
Germany	5.451	4.884	7.514	4.478	4.654	5.504	5.504	8.202	8.139	7.154	5.415
Greece	0.820	1.282	0.854	1.470	0.684	2.085	3.493	0.868	0.783	1.129	1.380
Hungary	1.251	0.827	0.924	0.756	0.834	1.292	2.331	0.845	1.067	1.226	0.678
Ireland	0.508	0.764	0.460	0.348	0.560	0.423	0.791	0.395	0.446	0.470	0.331
Italy	3.477	7.304	3.669	7.537	5.109	6.707	22.31	5.090	6.831	5.978	5.433
Netherlands	2.209	2.819	2.045	2.142	2.731	3.221	3.229	2.505	3.699	4.352	3.066
Norway	3.667	3.385	3.427	2.234	0.790	3.038	4.878	1.862	1.944	1.530	2.971
Poland	1.794	2.834	2.199	2.098	2.350	1.659	1.659	2.631	2.638	2.651	1.136
Portugal	1.206	1.179	1.163	1.168	1.374	1.179	1.153	1.297	1.575	1.497	0.782
Romania	0.989	0.901	0.923	0.921	1.335	2.645	4.775	1.571	1.238	1.321	2.961
Slovakia	0.585	0.574	0.920	0.743	0.209	0.792	1.340	0.389	0.732	0.592	0.865
Spain	3.490	3.054	3.082	5.365	2.959	8.699	15.11	3.471	4.262	4.087	1.975
Sweden	1.369	1.229	1.244	1.077	1.178	1.174	1.174	1.443	1.904	1.639	0.832
Switzerland	1.062	0.984	1.067	0.918	1.089	1.008	1.276	1.207	1.193	1.311	0.812
Turkey	19.98	20.80	20.44	13.62	3.547	8.758	7.584	2.989	4.165	3.960	27.38
United Kingdom	9.043	5.566	9.76	11.44	2.526	6.878	6.690	5.808	5.920	5.809	5.033
Other	1.926	2.519	2.008	2.050	2.240	3.451	1.482	2.238	2.881	3.550	2.648
Europe & Eurasia continent	37.65	26.98	29.25	40.04	30.67	53.03	53.03	30.91	43.89	38.37	12.58

Table 3 Success rates of the methods

Methods	RMSE			MAPE		
	Mean	Median	Minimum	Mean	Median	Minimum
DN-PSO method	1 (%4)	1 (%4)	1 (%4)	0 (%0)	0 (%0)	1 (%4)
ANN-PSO method	1 (%4)	0 (%0)	3 (%12)	9 (%36)	8 (%32)	2 (%8)
Median ANN method	3 (%12)	2 (%8)	2 (%8)	3 (%12)	4 (%16)	3 (%12)
Median-Pi-ANN method	2 (%8)	3 (%12)	1 (%4)	3 (%12)	1 (%4)	3 (%12)
FTS-N method	4 (%16)	3 (%12)	5 (%20)	2 (%8)	2 (%8)	6 (%24)
SC-FTS method	–	–	0 (%0)	–	–	0 (%0)
C-FTS method	–	–	1 (%4)	–	–	1 (%4)
ARIMA method	–	–	1 (%4)	–	–	2 (%8)
H-ES method	–	–	0 (%0)	–	–	0 (%0)
A-ES method	–	–	0 (%0)	–	–	0 (%0)
Proposed method	14 (%56)	16 (%64)	11 (%44)	11 (%44)	11 (%44)	8 (%32)

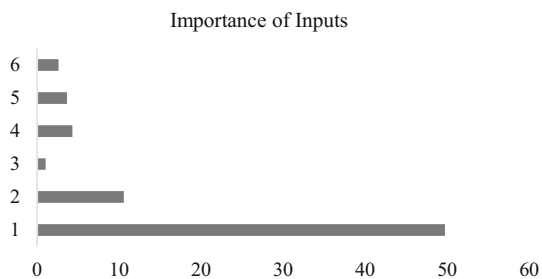


Fig. 3 The importance graph of model inputs

method and the FTS-N method Bas et al. (2015) is the third-best method according to MAPE in forecasting the future.

The optimal parameter values of the best method for each series are given in a Table in the supplementary file.

7 Explainable results of the proposed method

The inputs in the model can be tested and commented on. The inputs of the models are scores that are obtained from the principal component analysis. The contribution of each score to the forecasting model can be tested and commented on. An example for the Turkey time series is given below by using tables and graphs.

In the forecasting model of the Turkey time series, six PCA score variables are employed as inputs. It is seen that all inputs are significant in the model in Table 4. The importance of the inputs can be determined according to coefficients. The importance graph is given in Fig. 3.

8 Conclusions and discussion

The high-order definition of intuitionistic fuzzy time series definition of Egrioglu et al. (2019b) means that any time series is affected by lagged time series, lagged membership values and non-membership values. However, indeed,

Table 4 Explainable results for Turkey time series

No	Coefficients	Std. Dev	t-statistics	Prob
1	49.72	0.74	66.81	0.000000
2	- 10.55	0.29	- 36.74	0.000000
3	1.05	0.31	3.35	0.001790
4	- 4.28	0.34	- 12.78	0.000000
5	- 3.62	0.35	- 10.38	0.000000
6	- 2.59	0.43	- 6.07	0.000000

membership and non-membership values are obtained from the lagged time series. So, it can be thought that membership and non-membership values include, largely, the information of lagged time series. From this viewpoint, in this study, a new high-order fuzzy intuitionistic time series definition given Definition 1 is introduced. It is explained that any time series is affected by lagged membership values and non-membership values in Definition 1. For this reason, it can be said that the advantage of the proposed method that uses Definition 1 is simplifying the model, since the number of inputs in the model is reduced. In the literature, intuitionistic fuzzy time series models have been proposed in recent years, since the definitions of classical fuzzy time series in literature can be insufficient for many real-life series. The reason for this is fuzzy time series approaches use only membership values for forecasting any time series. However, both membership and non-membership values are used in the intuitionistic fuzzy time series methods. Thus, the intuitionistic fuzzy time series methods are the approaches to increase the forecasting performance because of being used of more information. Also in this study, similarly, it has been shown that the forecasting performance of the proposed method that has a new intuitionistic fuzzy time series approach has much better forecasting performance than the classical time series and the classical fuzzy time series methods. Therefore, in future studies, it could be suggested that researchers prefer to use the intuitionistic fuzzy time series approaches instead of the classical time series and the classical fuzzy time series approaches.

In this study, a new forecasting algorithm is proposed, and it is shown that the proposed method has better forecasting performance with the best success rates between 32 and 64% than many time series models in the literature. In the proposed method, the principal component analysis method was used to make the dimension reduction. Then, the robust regression is used in the stage of determination relationships of the proposed method. Therefore, it can be suggested that some regression methods or statistical methods are used in the stage of determination relationships of the intuitionistic fuzzy time series that will be made in the future. In this study, the forecasting performances of 11 different time series forecasting methods are compared with each other. Then, for each PEC time series in the Europe-Eurasia continent, the best methods and the parameters of the best methods are given in Table 2. So, researchers that have worked in the energy fields could decide which methods they will select by looking at Table 2, when they would like to solve any PEC time series in the Europe-Eurasia continent.

In fuzzy time series methods, inputs of the models cannot be tested or commented on. The proposed method is an explainable artificial intelligence method. The proposed

method is the first explainable fuzzy time series method in the literature. It is possible to test, order and comment on all inputs in the model because the fuzzy relations are modeled by using robust regression analysis. As a result of this study, it can be said that explainable artificial intelligence methods can be produced by combining statistical and machine learning methods. It can be expected that the studies about explainable fuzzy time series methods will be increased soon.

Declarations

Conflict of interest The author declares that there is no conflict of interest toward the publication of this manuscript.

References

- Abhishekh Gautam SS, Singh SR (2018) A score function-based method of forecasting using intuitionistic fuzzy time series. *New Math Nat Comput* 14(1):91–111
- Aladag CH (2013) Using multiplicative neuron model to establish fuzzy logic relationships. *Expert Syst Appl* 40(3):850–853
- Aladag CH, Basaran MA, Egrioglu E, Yolcu U, Uslu VR (2009) Forecasting in high order fuzzy time series by using neural networks to define fuzzy relations. *Expert Syst Appl* 36(3):4228–4231
- Aladag CH, Yolcu U, Egrioglu E (2010) A high order fuzzy time series forecasting model based on adaptive expectation and artificial neural networks. *Math Comput Simul* 81(4):875–882
- Aladag CH, Yolcu U, Egrioglu E, Dalar AZ (2012) A new time invariant fuzzy time series forecasting method based on particle swarm optimization. *Appl Soft Comput* 12(10):3291–3299
- Aladag CH, Yolcu U, Egrioglu E (2014) Robust multilayer neural network based on median neuron model. *Neural Comput Appl* 24(3–4):945–956
- Bas E, Egrioglu E, Aladag CH, Yolcu U (2015) Fuzzy-time-series network used to forecast linear and nonlinear time series. *Appl Intell* 43(2):343–355
- Bas E, Grosan C, Egrioglu E, Yolcu U (2018) High order fuzzy time series method based on pi-sigma neural network. *Eng Appl Artif Intell* 72:350–356
- Box GEP, Jenkins GM (1976) *Time series analysis: forecasting and control*. Holden Day Press, San Francisco
- Cagcag Yolcu O, Alpaslan F (2018) Prediction of TAIEEX based on hybrid fuzzy time series model with single optimization process. *Appl Soft Comput* 66:18–33
- Cagcag Yolcu O, Lam HK (2017) A combined robust fuzzy time series method for prediction of time series. *Neurocomputing* 247:87–101
- Chaira T (2011) A novel intuitionistic fuzzy C-means clustering algorithm and its application to medical images. *Appl Soft Comput* 11(2):1711–1717
- Chen SM (1996) Forecasting enrollments based on fuzzy time-series. *Fuzzy Sets Syst* 81(3):311–319
- Chen SM (2002) Forecasting enrollments based on high order fuzzy time series. *Cybern Syst* 33(1):1–16
- Chen SM, Chen CD (2011) Handling forecasting problems based on high-order fuzzy logical relationships. *Expert Syst Appl* 38(4):3857–3864
- Chen SM, Kao PY (2013) TAIEEX forecasting based on fuzzy time series particle swarm optimization techniques and support vector machines. *Appl Inf Sci* 247:62–71
- Chen SM, Tanuwijaya K (2011) Fuzzy forecasting on high-order relationships and automatic clustering techniques. *Expert Syst Appl* 38(12):15425–15437
- Cheng SH, Chen SM, Jian WS (2016) Fuzzy time series forecasting based on fuzzy logical relationships and similarity measures. *Inf Sci* 327:272–287
- Edgeworth FY (1887) On observations relating to several quantities. *Hermathena* 6(13):278–285
- Egrioglu E, Aladag CH, Yolcu U, Uslu VR, Basaran MA (2009a) A new approach based on artificial neural networks for high order multivariate fuzzy time series. *Expert Syst Appl* 36(7):10589–10594
- Egrioglu E, Uslu VR, Yolcu U, Basaran MA, Aladag CH (2009b) A new approach based on artificial neural networks for high order bivariate fuzzy time series. *Applications of Soft Computing*. Springer-Verlag, Berlin, pp 265–273
- Egrioglu E, Aladag CH, Yolcu U, Uslu VR, Basaran MA (2010) Finding an optimal interval length in high order fuzzy time series. *Expert Syst Appl* 37(7):5052–5055
- Egrioglu E, Aladag CH, Basaran MA, Uslu VR, Yolcu U (2011) A new approach based on the optimization of the length of intervals in fuzzy time series. *J Intell Fuzzy Syst* 22(1):15–19
- Egrioglu E, Aladag CH, Yolcu U (2013) Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks. *Expert Syst Appl* 40(3):854–857
- Egrioglu E, Yolcu U, Bas E (2019a) Intuitionistic high-order fuzzy time series forecasting method based on pi-sigma artificial neural networks trained by artificial bee colony. *Granular Comput* 4(4):639–654
- Egrioglu E, Yolcu U, Bas E, Dalar AZ (2019b) Median-Pi artificial neural network for forecasting. *Neural Comput Appl* 31(1):307–316
- Fan XS, Lei YJ, Lu YL, Wang YN (2016) Long-term intuitionistic fuzzy time series forecasting model based on DTW. *J Commun* 37(8):95–104
- Fan X, Lei Y, Wang Y (2017) Adaptive partition intuitionistic fuzzy time series forecasting model. *J Syst Eng Electron* 28(3):585–596
- Hsu LY, Horng SJ, Kao TW, Chen YH, Run RS, Chen RJ, Lai JL, Kuo IH (2010) Temperature prediction and TAIFEX forecasting based on fuzzy relationships and MTPSO techniques. *Expert Syst Appl* 37(4):2756–2770
- Hu D, Zan L, Chen X, Jie W (2017) Prediction of satellite clock errors based on deterministic intuitionistic fuzzy time series. In: *International conference on signal processing (ICSP) proceedings*, Art. no. 7877981, pp 1006–1009
- Huang YL, Horng SJ, He M, Fan P, Kuo IH (2011) A hybrid forecasting model for enrollments based on aggregated fuzzy time series and particle swarm optimization. *Expert Syst Appl* 38(7):8014–8023
- Huang K (2001) Effective length of intervals to improve forecasting in fuzzy time-series. *Fuzzy Sets Syst* 123(3):387–394
- Huang K, Yu TKK (2006) Ratio-based lengths of intervals to improve fuzzy time series forecasting. *IEEE Trans Syst Man Cybern-Part B: Cybern* 36(2):328–340
- James JO, Raymond LC, Ruppert D (1988) A note on computing robust regression estimates via iteratively reweighted least squares. *Am Stat* 42(2):152–1154
- Joshi BP, Pandey M, Kumar S (2016) Use of intuitionistic fuzzy time series in forecasting enrollments to an academic institution. *Adv Intell Syst Comput* 436:843–852
- Kadilar C (2005) *Introduction to time series analysis with SPSS applications*. Bizim Press, Istanbul

- Kocak C (2013) First-order ARMA type fuzzy time series method based on fuzzy logic relation tables. *Math Probl Eng*, Article ID 769125, 12
- Kocak C (2015) A new high order fuzzy ARMA time series forecasting method by using neural networks to define fuzzy relations. *Math Probl Eng*, Article ID 128097, pp 1–14
- Kocak C (2017) ARMA(p, q) type high order fuzzy time series forecast method based on fuzzy logic relations. *Appl Soft Comput* 58:92–103
- Kumar S, Gangwar SS (2016) Intuitionistic fuzzy time series: an approach for handling nondeterminism in time series forecasting. *IEEE Trans Fuzzy Syst* 24(6):1270–1281
- Kuo IH, Horng SJ, Kao TW, Lin TL, Lee CL, Pan Y (2009) An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization. *Expert Syst Appl* 36(3):6108–6117
- Kuo IH, Horng SJ, Chen YH, Run RS, Lin TL (2010) Forecasting TAIEX based on fuzzy time series and particle swarm optimization. *Expert Syst Appl* 37(2):1494–1502
- Lei Y, Lei Y, Fan X (2016) Multi-factor high-order intuitionistic fuzzy time series forecasting model. *J Syst Eng Electron* 27(5):1054–1062
- Mallows C (1975) On some topics in robustness. Technical Memorandum. Bell Telephone Laboratories, Murray Hill, NJ
- Park JI, Lee DJ, Song CK, Chun MG (2010) TAIEX and KOSPI 200 forecasting based on two-factors high-order fuzzy time series and particle swarm optimizations. *Expert Syst Appl* 37(2):959–967
- Rousseeuw PJ (1984) Least median of squares regression. *J Am Stat Assoc* 79(388):871–880
- Singh P, Borah B (2014) Forecasting stock index price based on M-Factors fuzzy time series and particle swarm optimization. *Int J Approx Reason* 55(3):812–833
- Song Q, Chissom BS (1993a) Fuzzy time series and its models. *Fuzzy Sets Syst* 54(3):269–277
- Song Q, Chissom BS (1993b) Forecasting enrollments with fuzzy time series—part I. *Fuzzy Sets Syst* 54(1):1–9
- Tak N (2018) Meta fuzzy functions: application of recurrent type-1 fuzzy functions. *Appl Soft Comput* 73:1–13
- Tak N (2020a) Type-1 possibilistic fuzzy forecasting functions. *J Comput Appl Math* 370:112653
- Tak N (2020b) Type-1 recurrent intuitionistic fuzzy functions for forecasting. *Expert Syst Appl* 140:112913
- Tak N (2021) Meta fuzzy functions-based feed-forward neural networks with a single hidden layer for forecasting. *J Stat Comput Simul*:1–17
- Tak N, Evren AA, Tez M, Egrioglu E (2018) Recurrent type-1 fuzzy functions approach for time series forecasting. *Appl Intell* 48(1):68–77
- Wang L, Liu X, Pedrycz W, Shao Y (2014) Determination of temporal information granules to improve forecasting in fuzzy time series. *Expert Syst Appl* 41(6):3134–3142
- Yapar G (2016) Modified simple exponential smoothing. *Hacettepe J Math Stat* 47(143):1–16
- Yapar G, Capar S, Selamlar HT, Yavuz I (2017a) Modified Holt's linear trend method. *Hacettepe J Math Stat* 47(2):1–10
- Yapar G, Yavuz I, Selamlar HT (2017b) Why and how does exponential smoothing fail? An in-depth comparison of ATA-simple and simple exponential smoothing. *Turk J Forecast* 1(1):30–39
- Yolcu U, Egrioglu E, Uslu VR, Basaran MA, Aladag CH (2009) A new approach for determining the length of intervals for fuzzy time series. *Appl Soft Comput* 9(2):647–651
- Yolcu U, Egrioglu E, Aladag CH (2013) A new linear & nonlinear artificial neural network model for time series forecasting. *Decis Support Syst* 54(3):1340–1347
- Yu THK, Huang KHA (2010) A neural network-based fuzzy time series model to improve forecasting. *Expert Syst Appl* 37(4):3366–3372
- Zadeh LA (1965) Information and control. *Fuzzy Sets* 8(3):338–353
- Zheng KQ, Lei YJ, Wang R, Wang Y (2013a) Prediction of IFTS based on deterministic transition. *J Appl Sci Electron Inf Eng* 31(2):204–211
- Zheng KQ, Lei YJ, Wang R, Wang YF (2013b) Modeling and application of IFTS. *Control Decis* 28(10):1525–1530
- Zheng KQ, Lei YJ, Wang R, Xing YQ (2014) Method of long-term IFTS forecasting based on parameter adaptation. *Syst Eng Electron* 36(1):99–104
- Zhou T, Gao S, Wang J, Chu C, Todo Z (2016) Financial time series prediction using a dendritic neuron model. *Knowl-Based Syst* 105:214–224
- Egrioglu E, Yolcu U, Bas E, Dalar AZ (2017) Median-Pi artificial neural network for forecasting. *Neural Comput Appl* 31(1):307–316

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.