

# A Novel Forecasting Method Based on $F$ -Transform and Fuzzy Time Series

Woo-Joo Lee<sup>1</sup> · Hye-Young Jung<sup>2</sup> · Jin Hee Yoon<sup>3</sup> · Seung Hoe Choi<sup>4</sup> 

Received: 28 March 2016/Revised: 4 April 2017/Accepted: 7 July 2017/Published online: 31 July 2017  
© The Author(s) 2017. This article is an open access publication

**Abstract** The main goal of time series analysis is to establish forecasting model based on past observations and to reduce forecasting error. To achieve these goals, the present paper proposes a new forecasting algorithm based on the fuzzy transform ( $F$ -transform) and the fuzzy logical relationships. First, the  $F$ -transform is performed based on partitioning of the universe, and the fuzzy logical relationships are employed to forecast. Two experimental applications are used to illustrate and verify the proposed algorithm. The accuracies are evaluated on the basis of average forecasting error percentage and index of agreement to compare the proposed algorithm with other existing methods.

**Keywords** Time series · Forecasting · Fuzzy transform · Fuzzy logical relationship

---

✉ Seung Hoe Choi  
shchoi@kau.ac.kr

Woo-Joo Lee  
bobspace@yonsei.ac.kr

Hye-Young Jung  
hyjunglove@snu.ac.kr

Jin Hee Yoon  
jin9135@sejong.ac.kr

<sup>1</sup> Department of Mathematics, Yonsei University, Seoul, South Korea

<sup>2</sup> Faculty of Liberal Education, Seoul National University, Seoul, South Korea

<sup>3</sup> School of Mathematics and Statistics, Sejong University, Seoul, South Korea

<sup>4</sup> School of Liberal Arts and Science, Korea Aerospace University, Goyang, Republic of Korea

## 1 Introduction

One of the main goals of time series analysis is to construct forecasting model that is used to predict future values based on past historical observations. The traditional time series models such as ARIMA and ARCH can not deal with forecasting problems with vague or ambiguous observations represented by linguistic concept. An appropriate way of solving such problem is by the use of time series model based on the fuzzy set theory. In addition, the time series models based on the fuzzy set theory can be applied to small sample data which is not easy to handle in traditional data analysis. To date, many forecasting models based on the fuzzy set theory [45] have been established by many authors and used to reduce forecasting error.

In this paper, we propose a new algorithm to forecast time series which is based on the fuzzy transform ( $F$ -transform) and fuzzy logical relationship. The  $F$ -transform introduced by Perfilieva [25] has been studied and found useful in many applications in function approximation, image processing [22, 27], numerical solutions of partial differential equations [36], data analysis [29] or neural network approaches [37]. The  $F$ -transform converts original data into weighted mean values where the weights are given by the basic functions which are membership functions to identify fuzzy sets. This is a novel method to find an approximation of given data or function.

In a time series analysis and forecast, the use of  $F$ -transform and inverse  $F$ -transform has been reported in many studies [22, 23, 36–38, 41]. In some studies, the  $F$ -transform was used to extract a low-frequency trend component [23, 36, 37], whereas it was used for the modeling of an autoregression function in [22]. Also, the inverse  $F$ -transform was used as a technical indicator in a stock market instead of the commonly used simple and

exponential moving averages in other study [4]. Novák et al. [23] have used inverse  $F$ -transform in combination with perception-based logical deduction, and Štěpnička et al. [36, 37] have provided forecasts of the future  $F$ -transform components.

Song and Chissom [32, 33] firstly introduced fuzzy time series which is a new algorithm of time series forecasting based on fuzzy logic. Many authors have suggested methods to modify and improve the model of Song and Chissom [2, 4, 9–13, 32–35, 40]. Some authors have applied their models to temperature forecasting [3, 18, 19], and others have used stock index forecasting [9–11, 16, 18, 19, 43, 44] to verify their models. Lee and Hong [20] have applied into electric power load forecasting.

Our proposed algorithm using fuzzy transform as defuzzified value of fuzzy sets corresponding to the partitioned intervals of domain is constructed by fuzzy logical relationships based on fuzzy time series. Generally, the fuzzy time series has used the non-overlapped membership functions. However, the proposed fuzzy time series algorithm based on the  $F$ -transform allows overlapping membership functions. The weighted sum of defuzzified values based on overlapping membership functions may play an important role in reducing the forecasting error. The proposed algorithm is applied to two well-known data sets: the enrollments of the University of Alabama and the number of patents granted in Taiwan, to show that it is superior to existing methods.

The rest of this paper is structured as follows: In Sect. 2, some preliminary concepts which are required to develop the main results are presented. In Sect. 3, the procedure of a forecasting method based on the  $F$ -transform applying fuzzy logical relationship is proposed. In Sect. 4, the experimental applications with enrollment data [32] and patents granted data of Taiwan [21] are provided. Section 5 concludes this study.

## 2 Preliminaries

A fuzzy subset of  $\mathbb{R}$  is a map, the so-called membership function, from  $\mathbb{R}$  into  $[0, 1]$ . Thus, a fuzzy subset  $A$  is identified by its membership function  $\mu_A(x)$ . For any  $\alpha \in (0, 1]$ , the crisp set  $A_\alpha = \{x \in \mathbb{R} : \mu_A(x) \geq \alpha\}$  is called the  $\alpha$ -cut or  $\alpha$ -level set of  $A$ . The set of all fuzzy numbers will be denoted by  $\mathcal{F}_c(\mathbb{R})$ . As a special case, we often use the following parametric class of fuzzy numbers, the so-called  $LR$ -fuzzy numbers:

$$\mu_A(x) = \begin{cases} L((m-x)/l) & \text{if } x \leq m, \\ R((x-m)/r) & \text{if } x > m \end{cases} \quad \text{for } x \in \mathbb{R},$$

where  $L, R : \mathbb{R}^+ \rightarrow [0, 1]$  are fixed left-continuous and non-increasing functions with  $R(0) = L(0) = 1$  and  $R(1) = L(1) = 0$ .  $L$  and  $R$  are called left and right shape functions of  $A$ ,  $m$  the mode of  $A$  and  $l, r > 0$  are left, right spread of  $A$ . For any fuzzy set  $A$ , the function  $\mu_A(x) : \mathbb{R} \rightarrow [0, 1]$  represents the membership function for which  $\mu_A(x)$  indicates the degree of membership of  $x \in \mathbb{R}$  that belongs to the fuzzy set  $A$ .

Fuzzy partitions which are needed in  $F$ -transform are introduced in several studies [8, 22, 25]. Here, we introduce some basic concepts defined in [25, 26, 28].

Let  $c_1 < \dots < c_k$  be fixed nodes within  $[a, b]$ , such that  $c_1 = a, c_k = b$  and  $k \geq 2$ . We say that fuzzy sets  $A_1, \dots, A_k$ , identified with their membership functions  $\mu_{A_1}(x), \dots, \mu_{A_k}(x)$  defined on  $[a, b]$ , constitute a fuzzy partition of  $[a, b]$  if they fulfill the following conditions for  $l = 1, \dots, k$ :

1.  $\mu_{A_l} : [a, b] \rightarrow [0, 1], \mu_{A_l}(c_l) = 1$ ;
2.  $\mu_{A_l}(x) = 0$  if  $x \notin (c_{l-1}, c_{l+1})$  where for uniformity of the notation, we put  $c_0 = a$  and  $c_k = b$ ;
3.  $\mu_{A_l}(x)$  is continuous;
4. For all  $x \in [a, b]$ ,

$$\sum_{l=1}^k \mu_{A_l}(x) > 0.$$

The membership functions  $\mu_{A_1}, \dots, \mu_{A_k}$  are called basic functions.

The example of fuzzy sets  $A_1, \dots, A_k$  with symmetric triangular membership functions on the interval  $[a, b]$  is given for  $(l = 1, \dots, k)$  below:

$$\mu_{A_l}(x) = \begin{cases} 1 - \left| \frac{x - c_l}{h_l} \right| & x \in [c_{l-1}, c_{l+1}], \\ 0 & \text{otherwise,} \end{cases}$$

where  $h_l = c_{l+1} - c_l, c_0 = a$  and  $c_{k+1} = b$ .

Let a discrete function  $f : X \rightarrow \mathbb{R}$  be given at a finite set of points  $X = \{x_t : t = 1, \dots, n\} \subseteq [a, b]$ . The  $F$ -transform of a discrete function  $f$  with respect to  $A_1, \dots, A_k$  defines the numerical vector  $F_k[f] = [F_1, F_2, \dots, F_k]$ , where each  $F_l$  is given by

$$F_l = \frac{\sum_{t=1}^n f(x_t) \mu_{A_l}(x_t)}{\sum_{t=1}^n \mu_{A_l}(x_t)}, \quad l = 1, \dots, k. \quad (1)$$

The  $F_l$  are weighted mean values of  $f$ , where the weights are determined by the membership values. The  $F_l$  are called components of the discrete  $F$ -transform.

Let  $F_k[f] = [F_1, \dots, F_k]$  be the  $F$ -transform of  $f$  with respect to  $A_1, \dots, A_k$ . Then the function

$$f_{F,k}(x_t) = \frac{\sum_{l=1}^k \mu_{A_l}(x_t) F_l}{\sum_{l=1}^k \mu_{A_l}(x_t)}, \quad t = 1, \dots, n. \quad (2)$$

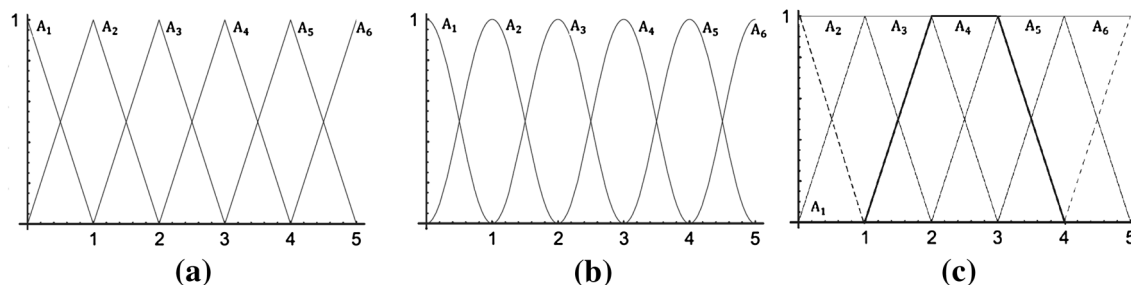


Fig. 1 Examples of basic functions. a Triangular function, b sinusoidal function, c trapezoidal function

is called the inverse  $F$ -transform of  $f$ . The inverse  $F$ -transform  $f_{F,k}$  can approximate  $f$  with an arbitrary precision. For various properties of the  $F$ -transform and detailed proofs, see [28].

Next, we give some definitions of fuzzy time series proposed by Song and Chissom [32–34] and present a well-known fuzzy time series algorithm.

Let  $y_t (t = \dots, 0, 1, 2, \dots)$ , a subset of  $\mathbb{R}$ , be the universe of discourse on which fuzzy sets  $A_i (i = 1, 2, \dots)$  are defined and let  $F(t)$  be a collection of  $A_i$ . Then,  $F(t)$  is called a fuzzy time series on  $y_t (t = \dots, 0, 1, 2, \dots)$ . If for any time  $t$ ,  $F(t) = F(t - 1)$  and  $F(t)$  only has finite elements, then  $F(t)$  is called a time-invariant fuzzy time series. Otherwise, it is called a time-variant fuzzy time series.

Suppose  $F(t - 1) = A_i$  and  $F(t) = A_j$ , a fuzzy logical relationship is defined as  $A_i \rightarrow A_j$ . Fuzzy logical relationships can be further grouped together into fuzzy logical relationship groups according to the same left-hand sides of the fuzzy logical relationships. For example, there are fuzzy logical relationships:  $A_i \rightarrow A_{j1}$ ,  $A_i \rightarrow A_{j2}, \dots$ , and  $A_i \rightarrow A_{jk}$ . These fuzzy logical relationships can be grouped into a fuzzy logical relationship group as follows:  $A_i \rightarrow A_{j1}, A_{j2}, \dots, A_{jk}$ .

Song and Chissom employed five main procedures in fuzzy time series as follows: (1) define and partition the universe of discourse; (2) define fuzzy sets and fuzzify the observed data; (3) establish the fuzzy logical relationships; (4) divide the derived fuzzy logical relationships into fuzzy logical relationship groups and forecast; and (5) defuzzify the forecasting results.

### 3 Fuzzy Time Series Forecasting

It is assumed in this paper that observed time series  $\{y_t\}$  is time invariant. In case  $\{y_t\}$  is time variant, we transform it to invariant time series applying proper time differences with proper time lags. The prediction procedure is proposed as follows:

1. Determine the universe of discourse  $U$ . Let  $D_{\max}$  and  $D_{\min}$  be the maximum and the minimum values of the observed time series data  $\{y_t : t = 1, 2, \dots, n\}$ . Then the universe of discourse  $U$  is defined by  $[D_{\min} - c_1, D_{\max} + c_2]$ , where  $c_1$  and  $c_2$  are the proper positive real numbers for simplifying these end points.
2. Partition the universe of discourse  $U$  into several subintervals and generate fuzzy sets  $\{A_1, A_2, \dots, A_k\}$  forming a fuzzy partition of  $U$ . A fuzzy partition for  $U$  is constructed by a decomposition  $\{a = c_1 < c_2 < \dots < c_k = b\}$  of  $U$  into  $k - 1$  subintervals  $I_j = [c_j, c_{j+1}]$ ,  $j = 1, \dots, k - 1$  and by fuzzy sets  $\{A_1, A_2, \dots, A_k\}$  identified by the membership functions (the basic functions)  $\mu_{A_1}(x), \dots, \mu_{A_k}(x)$  for  $x \in U$ .
3. Fuzzify the observed time series data  $\{y_t : t = 1, 2, \dots, n\}$ . Each of the data is fuzzified  $q$  times which is the number of overlapped membership functions of corresponding data. In Fig. 1a, if  $y_t$  is included in the subinterval  $I_j$  and membership function defined on  $I_j$  is a left shape function of  $\mu_{A_{j+1}}(x)$ , then we obtain  $A_{j+1}^1$  for its fuzzification. If membership function of defined on  $I_j$  is a right shape function of  $\mu_{A_j}(x)$ , then we obtain  $A_j^2$  for its fuzzification. In Fig. 1, the number of overlapped membership functions  $q = 2$  in (a), (b) and  $q = 3$  in (c) (Fig. 2).
4. Calculate  $F_1^h, F_2^h, \dots, F_k^h$  ( $h = 1, 2, \dots, q$ ) which are corresponding to  $A_1^h, A_2^h, \dots, A_k^h$  based on the following formula:

$$F_j^h = \frac{\sum_{t=1}^n y_t \mu_{A_j^h}(y_t)}{\sum_{t=1}^n \mu_{A_j^h}(y_t)}, \quad j = 1, \dots, k, \tag{3}$$

where  $\mu_{A_j^h}(y_t)$  is the membership degree of  $A_j^h$  at  $y_t$ .

5. Construction of fuzzy logical relationships. A fuzzy logical relationship is defined as the transition of the state at time  $t - 1$  to the state at time  $t$ . This is expressed as  $A_i^h \rightarrow A_j^h$  ( $h = 1, 2, \dots, p$ ), where  $A_i^h$  and  $A_j^h$  are the states at  $t - 1$  and  $t$ , respectively.
6. Calculate defuzzified predicted value  $m_j^h$  for each  $h$  at time  $t$ . At time  $t$ ,  $m_j^h$  ( $h = 1, \dots, p$ ) is determined by the

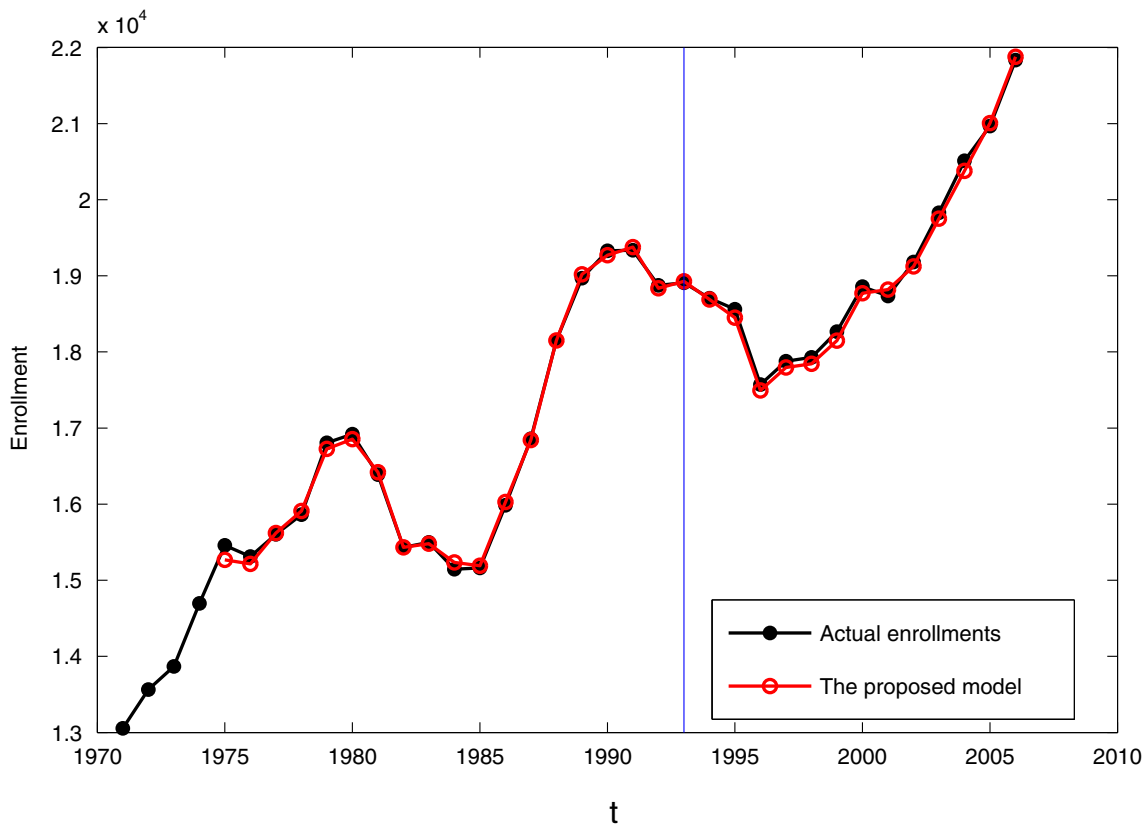


Fig. 2 Comparison of the predicted value and the actual value

following rules:

- (i) Rule 1: If the fuzzy logical relationship in the  $\lambda$ th order FLRGs is defined as

$$A_{i\lambda}^h, A_{i(\lambda-1)}^h, \dots, A_{i2}^h, A_{i1}^h \longrightarrow A_j^h$$

then the value of  $m_j^h$  is equal to  $F_j^h$  corresponding to  $A_j^h$ .

- (ii) Rule 2: If the fuzzy logical relationship in the FLRGs is shown as

$$A_{i\lambda}^h, A_{i(\lambda-1)}^h, \dots, A_{i2}^h, A_{i1}^h \longrightarrow A_{j1}^h, A_{j2}^h, \dots, A_{jr}^h (r \geq 2)$$

then the value of  $m_j^h$  is equal to average of  $F_{j1}^h, F_{j2}^h, \dots, F_{jr}^h$  corresponding to  $A_{j1}^h, A_{j2}^h, \dots, A_{jr}^h$ , respectively. Thus,  $m_j^h$  is given by

$$m_j^h = \frac{1}{r} \sum_{s=1}^r F_{js}^h \tag{4}$$

- (iii) Rule 3: If the fuzzy logical relationship in the FLRGs is shown as

$$A_{i\lambda}^h, A_{i(\lambda-1)}^h, \dots, A_{i2}^h, A_{i1}^h \longrightarrow \text{empty} \tag{5}$$

then the value of  $m_j^h$  is

$$m_j^h = \frac{\lambda F_{i1}^h + (\lambda - 1)F_{i2}^h + \dots + 1F_{i\lambda}^h}{\lambda + (\lambda - 1) + \dots + 1} \tag{6}$$

- 7. Calculate the forecasted value  $m_j$  at time  $t$  as follows:

$$m_j = \frac{\sum_{h=1}^p w_j^h m_j^h}{\sum_{h=1}^p w_j^h} \tag{7}$$

where  $w_j^h = \frac{\lambda \mu_{A_{i1}^h}(y_i) + (\lambda - 1) \mu_{A_{i2}^h}(y_i) + \dots + \mu_{A_{i\lambda}^h}(y_i)}{\lambda + (\lambda - 1) + \dots + 1}$  is the weighted value of  $m_j^h$  with respect to its membership degree.

Based on the seven steps presented above, the estimator for time series prediction can be obtained using  $F$ -transform.

The accuracy of the forecast can be evaluated by the basis of the index of agreement ( $D_1$ ) and the basis of the index of agreement ( $D_2$ ) suggested by Willmott [42]. These are computed as follows:

$$D_1(\%) = \frac{1}{N} \sum_{i=1}^N \left| \frac{P_i - O_i}{O_i} \right| \times 100, \tag{8}$$

$$D_2 = 1 - \left[ \frac{\sum (P_i - O_i)^2}{\sum (|P_i - \bar{O}| + |O_i - \bar{O}|)^2} \right], \quad 0 \leq D_2 \leq 1, \tag{9}$$

where  $N$  is the total number of data and  $O_i$ , and  $P_i$  are the observed and predicted loads, respectively.  $O_i$  is the mean value of the observed value. The metric  $D_2$  quantifies the relative contribution of systematic error to random error and has a value of 1 in a perfect model [42].

### 4 Applications

In this section, we compare the accuracies of the fuzzy time series predictions by the method using  $F$ -transform provided in Sect. 3 on two data sets which are frequently used in fuzzy time series analysis.

#### 4.1 Enrollment Data

The enrollment of the University of Alabama was introduced by Song and Chissom [32]. Many authors have used this data set which is provided in Table 1 to compare their works with other methods [5–7, 14, 24, 30, 31, 39]. Before we apply the whole procedure, we transform the data to time-invariant data using difference. The enrollment data from 1971 to 1992 are used as training data set to forecast the enrollment from 1993 to 2006 as follows:

1. Define the universe of discourse  $U$  as  $[D_{\min} - c_1, D_{\max} + c_2] = [-1000, 1400]$ , where  $D_{\min} = -955$  and  $D_{\max} = 1291$  with proper constants  $c_1$  and  $c_2$ .
2. Partition  $U$  into 8 intervals  $I_j = [c_j, c_{j+1}]$  where  $c_1 = -1000$  and  $c_{j+1} = -1000 + 300j$  ( $j = 1, \dots, 8$ ). Construct a fuzzy partition of an interval by fuzzy sets with symmetric triangular basic functions as follows:  $A_1 = [-1000, -700]$ ,  $A_2 = [-1000, -400]$ ,  $A_3 = [-700, -100]$ ,  $A_4 = [-400, 200]$ ,  $A_5 = [-100, 500]$ ,  $A_6 = [200, 800]$ ,  $A_7 = [500, 1100]$ ,  $A_8 = [800, 1400]$ ,  $A_9 = [1000, 1400]$ .
3. There are two triangular membership functions for each  $y_t$ ;  $y_t$  can be fuzzified into two different types. For example, at  $t = 1972$ , 508 has corresponding membership degree 0.027 of fuzzy set  $A_7 = [500, 1100]$ ; it is fuzzified into  $A_7^1$ . At the same time, 508 has corresponding membership degree 0.973 of fuzzy set  $A_6 = [200, 800]$ ; it is fuzzified into  $A_6^2$ . Similarly, at  $t = 1973$ , 304 has corresponding membership degree 0.347 of fuzzy set  $A_6 = [200, 800]$ ; it is fuzzified into  $A_6^1$ . At the same time, 304 has corresponding membership degree 0.653 of fuzzy set  $A_5$ ; it is fuzzified into  $A_5^2$ .
4. From (3), Table 2 shows  $F_1^h, \dots, F_9^h$  ( $h = 1, 2$ ) which are corresponding to fuzzy sets  $(A_1^h, \dots, A_9^h)$ . All values are rounded off to three decimal places.

**Table 1** Results of training set using 1st order FLR

Year	Enrollment	$y_t$	Fuzzy number <sup>1</sup>	$\mu_{A_t^1}$	Fuzzy number <sup>2</sup>	$\mu_{A_t^2}$
1971	13,055	–				
1972	13,452	508	$A_7^1$	0.027	$A_6^2$	0.973
1973	13,867	304	$A_6^1$	0.347	$A_5^2$	0.653
1974	14,696	829	$A_8^1$	0.097	$A_8^2$	0.903
1975	15,460	764	$A_7^1$	0.880	$A_6^2$	0.120
1976	15,311	–149	$A_4^1$	0.837	$A_3^2$	0.163
1977	15,603	292	$A_6^1$	0.307	$A_5^2$	0.693
1978	15,861	258	$A_6^1$	0.193	$A_5^2$	0.807
1979	16,807	946	$A_8^1$	0.487	$A_7^2$	0.513
1980	16,919	112	$A_5^1$	0.707	$A_4^2$	0.293
1981	16,388	–531	$A_3^1$	0.563	$A_2^2$	0.437
1982	15,433	–955	$A_2^1$	0.150	$A_1^2$	0.850
1983	15,497	64	$A_5^1$	0.547	$A_4^2$	0.453
1984	15,145	–352	$A_4^1$	0.160	$A_3^2$	0.840
1985	15,163	18	$A_5^1$	0.393	$A_4^2$	0.607
1986	15,984	821	$A_8^1$	0.070	$A_7^2$	0.930
1987	16,859	875	$A_8^1$	0.250	$A_7^2$	0.750
1988	18,150	1291	$A_9^1$	0.637	$A_8^2$	0.363
1989	18,970	820	$A_8^1$	0.067	$A_7^2$	0.933
1990	19,328	358	$A_6^1$	0.527	$A_5^2$	0.473
1991	19,337	9	$A_5^1$	0.363	$A_4^2$	0.637
1992	18,876	–461	$A_3^1$	0.797	$A_2^2$	0.203

**Table 2** Defuzzified values  $F_j^h$  of fuzzy sets  $A_j^h$

Fuzzy value	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$	$A_6$	$A_7$	$A_8$	$A_9$
$F_p^1$	–	–955	–490.00	–181.59	61.93	315.55	756.47	898.36	1291
$F_p^2$	–955	–508.76	–318.95	39.46	296.44	536.10	848.53	1291	–

**Table 3** Fuzzy logical relationships of first order

Index	Left FLR			Right FLR		
1	$A_2^1$	→	$\{A_5^1\}$	$A_1^2$	→	$\{A_4^2\}$
2	$A_3^1$	→	$\{A_2^1\}$	$A_2^2$	→	$\{A_1^2\}$
3	$A_4^1$	→	$\{A_6^1, A_5^1\}$	$A_3^2$	→	$\{A_5^2, A_4^2\}$
4	$A_5^1$	→	$\{A_3^1, A_4^1, A_8^1, A_3^1\}$	$A_4^2$	→	$\{A_2^2, A_3^2, A_7^2, A_2^2\}$
5	$A_6^1$	→	$\{A_8^1, A_6^2, A_8^1, A_5^1\}$	$A_5^2$	→	$\{A_7^2, A_5^2, A_7^2, A_4^2\}$
6	$A_7^1$	→	$\{A_6^1, A_4^1\}$	$A_6^2$	→	$\{A_5^2, A_3^2\}$
7	$A_8^1$	→	$\{A_7^1, A_5^1, A_8^6, A_9, A_6^1\}$	$A_7^2$	→	$\{A_6^2, A_4^2, A_7^2, A_8^2, A_5^2\}$
8	$A_9^1$	→	$\{A_8^1\}$	$A_8^2$	→	$\{A_7^2\}$

**Table 4** Results of training set using 1th order FLR

Year	$m_t^1$	$m_t^2$	$w_t^1$	$w_t^2$	$\hat{y}_t$	Forecasted values
1971						
1972						
1973	66.98	–11.26	0.027	0.973	–9.171	13,553.83
1974	543.55	508.24	0.347	0.653	520.482	14,387.48
1975	664.66	602.31	0.097	0.903	608.333	15,304.33
1976	66.98	–11.26	0.880	0.120	57.593	15,517.59
1977	188.74	167.95	0.837	0.163	185.346	15,496.35
1978	543.55	508.24	0.307	0.693	519.069	16,122.07
1979	543.55	508.24	0.193	0.807	515.067	16,376.07
1980	664.66	602.31	0.487	0.513	632.653	17,439.65
1981	–65.81	–121.99	0.707	0.293	–82.284	16,836.72
1982	–955.00	–955.00	0.563	0.437	–955.00	15,433.00
1983	61.93	39.46	0.150	0.850	42.827	15,475.83
1984	–65.81	–121.99	0.547	0.453	–91.273	15,405.73
1985	188.74	167.95	0.160	0.840	171.274	15,316.27
1986	–65.81	–121.99	0.393	0.607	–99.888	15,063.11
1987	664.66	602.31	0.070	0.930	606.670	16,590.67
1988	664.66	602.31	0.250	0.750	617.895	17,476.89
1989	898.36	848.55	0.637	0.363	880.257	19,030.26
1990	664.66	602.31	0.067	0.933	606.462	19,576.46
1991	543.55	508.24	0.527	0.473	526.838	19,854.84
1992	–65.81	–121.99	0.363	0.637	–101.573	19,235.43

5. Table 3 shows 1st order fuzzy logical relationship obtained from given fuzzy time series.

6. From Table 3, defuzzified values of predicted can be obtained. For example, left and right FLRs at  $t = 1973$  are  $A_7^1 \rightarrow \{A_6^1, A_4^1\}$ ,  $A_6^2 \rightarrow \{A_5^2, A_3^2\}$ , respectively. By

**Table 5** Forecasting results of high order

Order	1	2	3	4	5	6
MAPE	1.8770	0.7924	0.3105	0.2575	0.2281	0.2372
$d$	0.9874	0.9851	0.9996	0.9997	0.9998	0.9998

Rule 2, defuzzified values of predicted fuzzy numbers are:

$$m_{1973}^1 = \frac{1}{2}(F_4^1 + F_6^1) = 66.98,$$

$$m_{1973}^2 = \frac{1}{2}(F_3^2 + F_5^2) = -11.26$$

At  $t = 1983$ , left and right FLRs are  $A_2^1 \rightarrow \{A_3^1\}$ ,  $A_1^2 \rightarrow \{A_4^2\}$ . By Rule 1, defuzzified values of predicted fuzzy numbers are:

$$m_{1983}^1 = F_5^1 = 61.93, m_{1983}^2 = F_4^2 = 39.46$$

7. Forecasted values can be obtained as follows: For example, for  $1972 \rightarrow 1973$ , the weight of 1th order

**Table 9** Forecasting accuracies of patents granted data of Taiwan

Order	1	2	3	4	5	6
$D_1$	12.1900	5.8654	5.0228	5.1794	4.8438	4.9767
$D_2$	0.9643	0.9937	0.9943	0.9938	0.9945	0.9942

**Table 6** The forecasting results of different forecasting methods

Method	Chen et al. [5]	Jilani et al. [14]	Stevenson et al. [39]	Chou [6]
$D_1$	2.4452	0.8891	0.5713	2.6914
$D_2$	0.9785	0.9972	0.9981	0.9752
Method	Eğrioğlu [7]	Saxena et al. [31]	Rzayev et al. [30]	Nurhayadi [24]
$D_1$	2.2059	0.3407	1.4219	1.7899
$D_2$	0.9791	0.9992	0.9890	0.9874

**Table 7** Forecasted values and accuracies

Year	Enrollment	1st order	2nd order	3rd order	4th order	5th order	6th order
1993	18,909	18,933.36	18,930.12	18,929.10	18,927.59	18,927.12	18,926.62
1994	18,707	18,686.46	18,667.28	18,689.99	18,686.45	18,680.65	18,678.31
1995	18,561	18,437.89	18,418.88	18,448.81	18,443.69	18,435.62	18,430.39
1996	17,572	17,482.89	17,463.88	17,493.81	17,488.69	17,480.62	17,475.39
1997	17,877	17,782.19	17,765.45	17,796.76	17,792.27	17,784.90	17,779.95
1998	17,929	17,831.08	17,808.66	17,844.48	17,840.27	17,833.18	17,828.65
1999	18,267	18,135.73	18,112.73	18,148.12	18,143.99	18,137.14	18,132.78
2000	18,859	18,748.70	18,734.86	18,771.06	18,761.70	18,755.06	18,752.04
2001	18,735	18,798.30	18,779.83	18,820.20	18,810.52	18,803.62	18,800.59
2002	19,181	19,102.96	19,084.23	19,124.56	19,114.89	19,108.06	19,104.91
2003	19,828	19,719.42	19,708.95	19,751.63	19,740.02	19,733.95	19,731.10
2004	20,512	20,341.40	20,577.07	20,379.41	20,366.55	20,360.37	20,358.17
2005	20,969	20,965.26	20,879.63	21,007.49	20,992.40	20,987.56	20,985.54
2006	21,835	21,834.37	21,743.74	21,877.25	21,861.51	21,856.87	21,854.90
MAPE		0.4238	0.5006	0.3887	0.4004	0.4209	0.4329
d		0.9985	0.9982	0.9989	0.9988	0.9986	0.9985

**Table 8** Patents granted data and forecasted values

Year	Patents granted	Forecasted patents	Year	Patents granted	Forecasted patents	Year	Patents granted	Forecasted patents
1980	6633	–	1987	10,615	12,195.07	1994	19,032	18,293.39
1981	6264	–	1988	12,355	11,988.66	1995	29,707	29,151.18
1982	7460	–	1989	19,265	14,047.29	1996	29,469	29,804.23
1983	7096	9085.32	1990	22,601	23,447.91	1997	29,356	29,458.45
1984	8592	8434.99	1991	27,281	26,584.1	1998	25,051	25,463.35
1985	9427	10,243.53	1992	21,264	21,264	1999	29,144	29,003.28
1986	10,526	10,881.63	1993	22,317	22,696.34	2000	38,665	39,195.72



**Table 10** Comparison of accuracies of various forecasting methods

	Chen [2]	Hwang et al. [13]	Lee et al. [17]	Liu [21]	Jung et al. [15]	Proposed method
$D_1$	18.9356	13.6651	19.4385	10.4509	10.0458	5.0228
$D_2$	0.9497	0.9111	0.9377	0.9742	0.9746	0.9943

**Table 11** Forecasted values and accuracies of high orders of patents granted data

Year	1996	1997	1998	1999	2000	$D_1$	$D_2$
Patents granted	29,469	29,356	25,051	29,144	38,665		
Forecasted granted	29,909.59	30,044.01	26,759.01	28,230.71	38,905.71	2.8826	0.9882

FLR is  $w_{1973}^h = \mu_{A_i^h(508)}$ . So, we get  $w_{1973}^1 = 0.027$ ,  $w_{1973}^2 = 0.973$ . Hence, the forecasted value at  $t = 1973$  is

$$m_{1983} = \frac{0.027 * 66.98 + 0.973 * (-11.26)}{0.027 + 0.973} + 13,452 = 13,553.83.$$

Similarly, when  $1982 \rightarrow 1983$ ,  $w_{1983}^1 = 0.150$ ,  $w_{1983}^2 = 0.850$  are the weights. Hence, the forecasted value at  $t = 1983$  is

$$m_{1983} = \frac{0.150 * 61.93 + 0.850 * 39.46}{0.150 + 0.850} + 15,433 = 15,475.83.$$

For training set, we have used the data from 1971 to 1992. Table 3 shows 1st order FLR, and Table 4 shows the accuracies of the results. As it is shown in Table 5, generally higher order may improve the accuracy, but it may cause difficulties to find the forecasting value because of its complexity. Table 6 provides comparison with many method suggested by many authors such as Chen et al. [5].

From the FLR of the training set in Table 3, the corresponding fuzzy set of  $t = 1992$  are  $A_3^1, A_2^2$ . So we obtain  $m_{1993}^1 = 61.93$ ,  $m_{1993}^2 = 39.46$ , and  $w_{1993}^1 = 0.797$ ,  $w_{1993}^2 = 0.203$  at  $t = 1993$ . So, the 1th order forecasted value at  $t = 1993$  can be represented by

$$m_{1993} = \frac{0.797 * 61.93 + 0.203 * 39.46}{0.797 + 0.203} + 18,876 = 18,933.36.$$

Forecasted value ( $m_{1994}$ ) of  $t = 1994$  can be obtained after we add one more data item ( $m_{1993}$ ) of  $t = 1993$  to the training data set. Table 7 shows the forecasted values from 1993 to 2006.

#### 4.2 The Number of Patents Granted in Taiwan

Table 8 shows the number of patents granted in Taiwan from 1980 to 2000 introduced by Liu [21]. The data set is

transformed to be invariant time series by using first-order time difference because it is increasing.  $[-7000, 11, 000]$  is chosen to be the universe discourse based on minimum value  $-6017$  and maximum value  $10675$ . Here, we have used triangular fuzzy numbers for basic functions on 9 intervals ( $q = 9$ ). Forecasted data are provided in Table 8 with original data. Table 9 shows the accuracies after performing all 7 steps based on the orders of transform through time differences. 2nd order is chosen to forecast the data. Comparisons of the proposed method with other existing methods are provided in Table 10.

Next, we perform validations to confirm the superiority of the proposed method. The data from 1980 to 1995 have been used for training data set, and from 1996 to 2000 for testing data. The data from 1980 to 1995 are used to forecast the data at  $t = 1996$ . For forecasting the data at  $t = 1997$ , we add the data at  $t = 1996$  to training data set. Likewise, we update the training data set adding one data point for the next forecasting. The forecasted results and the accuracies are provided in Table 11, which shows very good performances with  $D_1 = 2.8826$  and  $D_2 = 0.9882$ .

### 5 Conclusions

In this paper, we propose a novel forecasting method based on  $F$ -transform and fuzzy time series to improve the forecasting accuracy. After applying  $F$ -transform to observed data, we construct a new fuzzy time series algorithm based on fuzzy logical relationships and  $F$ -transformed data. Through the real applications to enrollments of the University of Alabama and the number of patents granted in Taiwan, we show that forecasting accuracy can be further enhanced thorough applying  $F$ -transform. In addition, it is confirmed from these applications that the proposed method has better forecasting performances than existing methods. Future work involves extending the proposed method to handle two-factors fuzzy time series model and giving some results using various real applications to prove that it is correct.



**Acknowledgements** This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2017R1D1A1B03034813).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Boukezzoula, R., Galichet, S., Foulloy, L.: MIN and MAX operator for fuzzy intervals and their potential use in aggregation operators. *IEEE Trans. Fuzzy Syst.* **15**, 1135–1144 (2007)
- Chen, S.M.: Forecasting enrollments based on fuzzy time series. *Fuzzy Sets Syst.* **81**, 311–319 (1996)
- Chen, S.M., Hwang, J.R.: Temperature prediction using fuzzy time series. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **30**, 263–275 (2000)
- Chen, S.M.: Forecasting enrollments based on high order fuzzy time series. *Cybern. Syst.* **33**, 1–16 (2002)
- Chen, S.M., Hsu, C.C.: A new method to forecast enrollments using fuzzy time series. *Int. J. Appl. Sci. Eng.* **2**, 234–244 (2004)
- Chou, M.T.: Long-term predictive value interval with the fuzzy time series. *J. Mar. Sci. Technol.* **19**, 509–513 (2011)
- Eğrioglu, E.: A New Time-Invariant Fuzzy Time Series Forecasting Method Based on Genetic Algorithm. *Advances in Fuzzy Systems*. Hindawi Publishing Corporation, Cairo (2012)
- Holcapeka, M., Tichy, T.: A smoothing filter based on fuzzy transform. *Fuzzy Sets Syst.* **180**, 69–97 (2011)
- Huarng, K.: Effective lengths of intervals to improve forecasting in fuzzy time series. *Fuzzy Sets Syst.* **123**, 155–162 (2001a)
- Huarng, K.: Heuristic models of fuzzy time series for forecasting. *Fuzzy Sets Syst.* **123**, 369–386 (2001b)
- Huarng, K., Yu, H.K.: A type 2 fuzzy time series model for stock index forecasting. *Phys. A Stat. Mech. Appl.* **353**, 445–462 (2005)
- Huarng, K., Yu, H.K.: Ratio-based lengths of intervals to improve fuzzy time series forecasting. *IEEE Trans. Syst. Man Cybern. B Cybern.* **36**, 328–340 (2006)
- Hwang, J.R., Chen, S.M., Lee, C.H.: Handling forecasting problems using fuzzy time series. *Fuzzy Sets Syst.* **100**, 217–228 (1998)
- Jilani, T., Burney, S., Ardil, C.: Multivariate high order fuzzy time series forecasting for car road accidents. *Proc. World Acad. Sci. Eng. Technol.* **25**, 288–293 (2007)
- Jung, H.Y., Choi, S.H.: Time series using fuzzy logic. *Commun. Korean Stat. Soc.* **15**, 517–530 (2008)
- Kuo, S.C., Chen, C.C., Li, S.T.: Evolutionary fuzzy relational modeling for fuzzy time series forecasting. *Int. J. Fuzzy Syst.* **17**, 1295–1306 (2015)
- Lee, H.S., Chou, M.T.: Fuzzy forecasting based on fuzzy time series. *Int. J. Comput. Math.* **81**, 781–789 (2004)
- Lee, L.W., Wang, L.H., Chen, S.M., Leu, Y.H.: Handling forecasting problems based on two factor high-order fuzzy time series. *IEEE Trans. Fuzzy Syst.* **14**, 468–477 (2006)
- Lee, L.W., Wang, L.H., Chen, S.M.: Temperature prediction and TAIFEX forecasting based on fuzzy logical relationships and genetic algorithms. *Expert Syst. Appl.* **33**, 539–550 (2007)
- Lee, W.J., Hong, J.A.: A hybrid dynamic and fuzzy time series model for mid-term power load forecasting. *Int. J. Electr. Power Energy Syst.* **64**, 1057–1062 (2015)
- Liu, H.T.: An improved fuzzy time series forecasting method using trapezoidal fuzzy numbers. *Fuzzy Optim. Decis. Mak.* **6**, 63–80 (2007)
- Di Martino, F., Loiab, V., Sessa, S.: Fuzzy transforms method in prediction data analysis. *Fuzzy Sets Syst.* **180**, 146–163 (2011)
- Novák, V., Štěpnička, V.M., Dvořák, A., Perfilieva, I., Pavliska, V.: Analysis of seasonal time series using fuzzy approach. *Int. J. Gen. Syst.* **39**, 305–328 (2010)
- Nurhayadi, Subanar, Abdurakhman, Abadi, A.M.: Fuzzy model translation for time series data in the extent of median error and its application. *Appl. Math. Sci.* **8**, 2113–2124 (2014)
- Perfilieva, I., Haldeeva, E.: Fuzzy transformation. In: *Proceedings of the Joint 9th IFSA World Congress and 20th NAFIPS International Conference*, pp. 127–130 (2001)
- Perfilieva, I.: Fuzzy transforms. *Lecture notes in computer science* **3135**, 63–81 (2004)
- Perfilieva, I., Valášek, R.: Fuzzy transforms in removing noise. In: Reusch, B. (ed.) *Computational Intelligence, Theory and Applications, Advances in Soft Computing*, pp. 221–230. Springer, Berlin (2005)
- Perfilieva, I.: Fuzzy transforms: theory and applications. *Fuzzy Sets Syst.* **157**, 993–1023 (2006)
- Perfilieva, I., Novák, V., Antonín, Dvořák, A.: Fuzzy transform in the analysis of data. *Int. J. Approx. Reason.* **48**, 36–46 (2008)
- Rzayev, R., Askerov, N., Agamaliyev, M.: Modifying a fuzzy time series models based on point-estimate of fuzzy outputs. *Int. J. Data Min. Intell. Inf. Technol. Appl.* **3**, 24–30 (2013)
- Saxena, P., Sharma, K., Easo, S.: Forecasting enrollments based on fuzzy time series with higher forecast accuracy rate. *Comput. Technol. Appl.* **3**, 957–961 (2012)
- Song, Q., Chissom, B.S.: Forecasting enrollments with fuzzy time series-part I. *Fuzzy Sets Syst.* **54**, 1–10 (1993a)
- Song, Q., Chissom, B.S.: Fuzzy time series and its models. *Fuzzy Sets Syst.* **54**, 269–277 (1993b)
- Song, Q., Chissom, B.S.: Forecasting enrollments with fuzzy time series-part II. *Fuzzy Sets Syst.* **62**, 1–8 (1994)
- Song, Q., Leland, R.P., Chissom, B.S.: A new fuzzy time-series model of fuzzy number observations. *Fuzzy Sets Syst.* **73**, 341–348 (1995)
- Štěpnička, M., Valášek, R.: Numerical solution of partial differential equations with help of fuzzy transform. In: *IEEE International Conference on Fuzzy Systems*, pp. 1104–1109 (2005)
- Štěpnička, M., Polakovic, O.: A neural network approach to the fuzzy transform. *Fuzzy Sets Syst.* **160**, 1037–1047 (2009)
- Štěpnička, M., Dvořák, A., Pavliska, V., Vavříčková, L.: Linguistic approach to time series modeling with the help of F-transform. *Fuzzy Sets Syst.* **180**, 164–184 (2011)
- Stevenson, M., Porter, J.E.: Fuzzy time series forecasting using percentage change as the universe of discourse. *Proc. World Acad. Sci. Eng. Technol.* **55**, 154–157 (2009)
- Sullivan, J., Woodal, W.H.: A comparison of fuzzy forecasting and Markov modeling. *Fuzzy Sets Syst.* **64**, 279–293 (1994)
- Troiano, L., Kriplani, P.: Supporting trading strategies by inverse fuzzy transform. *Fuzzy Sets Syst.* **180**, 121–145 (2011)
- Willmott, C.J.: Some comments on the evaluation of model performance. *Bull. Am. Meteorol. Soc.* **63**, 1309–1313 (1982)
- Xihao, S., Yimin, L.: Average-based fuzzy time series models for forecasting Shanghai compound index. *World J. Model. Simul.* **2**, 104–111 (2008)
- Yu, H.K.: Weighted fuzzy time series models for TAIFEX forecasting. *Phys. A* **349**, 609–624 (2005)
- Zadeh, L.A.: Fuzzy sets. *Inf. Control* **8**, 338–353 (1965)



**Woo-Joo Lee** obtained a Ph.D. in Mathematics majoring mathematical statistics from Yonsei University, Seoul, Korea. His main research interests are data mining, fuzzy time series and intelligent systems.

several journals. etc. Also, has been regularly working as an organizer and committee member of several international conferences.



**Seung Hoe Choi** obtained his Ph.D. degree in Mathematical Statistics from Yonsei University Korea in 1994. Since 1996, he is currently full professor of Korea Aerospace University. His main research interests are mathematical prediction method using the soft computing and statistical prediction in sports like soccer, baseball and basketball.



**Hye-Young Jung** received Ph.D. degree in mathematics from Yonsei university in 2014. Currently, she is an associate teaching professor at the Faculty of Liberal Education, Seoul National University. Her research interests include statistical methods, fuzzy theory and statistical genetics.



**Jin Hee Yoon** received B.S., M.S. and Ph.D. degree in Mathematics from Yonsei University. She is currently faculty of school of Mathematics and Statistics at Sejong University, Seoul, South Korea. Her research interests are fuzzy regression analysis, fuzzy time series and intelligent systems. She is a board member of KIIS (Korean institute of Intelligent Systems) and has been working as an editor, guest editor and editorial board member of