Fuzzy Optim Decis Making (2007) 6:63–80
DOI 10.1007/s10700-006-0025-9

An improved fuzzy time series forecasting method using trapezoidal fuzzy numbers

Hao-Tien Liu

Abstract One of the major drawbacks of the existing fuzzy time series forecasting models is the fact that they only provide a single-point forecasted value just like the output of the traditional time series methods. Hence, they cannot provide a decision analyst more useful information. The aim of this present research is to design an improved fuzzy time series forecasting method in which the forecasted value will be a trapezoidal fuzzy number instead of a single-point value. Furthermore, the proposed method may also increase the forecasting accuracy. Two numerical data sets were used to illustrate the proposed method and compare the forecasting accuracy with three fuzzy time series methods. The results of the comparison indicate that the proposed method can generate forecasting values that are more accurate.

Keywords Fuzzy sets · Time series · Forecast · Trapezoidal fuzzy number

1 Introduction

The time series forecast has been a widely used forecasting method. Although time series forecast can deal with many forecasting problems, it cannot solve forecasting problems in which the historical data are vague, imprecise, or are in linguistic terms. To address this problem, Song and Chissom (1993a,b, 1994) presented the definitions of fuzzy time series by using fuzzy relational equations and approximate reasoning. Since then, a number of researchers have built on their research and developed different fuzzy forecasting methods (Chen, 1996,

H.-T. Liu (✉)
Department of Industrial Engineering and Management, I-Shou University, No. 1, Sec. 1, Syuecheng Rd., Dashu Township, Kaohsiung County 840, Taiwan, ROC
e-mail: htliu@isu.edu.tw

2002; Hwang, Chen, & Lee, 1998; Chen & Hwang, 2000; Huarng, 2001a,b; Lee & Chou, 2004).

Generally, the existing fuzzy forecasting methods can be classified into two types: time-variant and time-invariant. In time-variant models (Song & Chissom, 1994; Hwang, et al., 1998; Chen & Hwang, 2000) fuzzy composition operations are used, such as $F(t) = F(t-1) \circ R_w(t, t-1)$ or $F(t) = F(t-1) \circ O^w(t)$, to calculate the forecasted values. On the other hand, time-invariant forecasting models (Song & Chissom, 1993a; Chen, 1996,2002; Huarng, 2001a,b; Lee & Chou, 2004) often form fuzzy logical relationships, such as $A_i \rightarrow A_m$ or $A_i, \ldots, A_k \rightarrow A_m$, based on historical data, and group them as heuristic rules to derive the forecasted values.

However, the drawback of both time-variant and time-invariant forecasting lies in the fact that their forecasting value is a single-point value. In some way, the forecasting results are just like the output of the traditional time series forecasting methods. Nevertheless, the single-point value cannot provide a decision analyst more useful information.

To resolve this problem, the present study intends to develop an improved fuzzy time series method based on Chen's method (1996) because it provides an efficient forecasting algorithm and generates better forecasting results. The present research can achieve two major goals. The first goal is to provide the forecasting values with a trapezoidal fuzzy number instead of a single-point value. By doing so, the decision analyst can gather the information about the possible forecasted ranges under different degrees of confidence. The second goal is to revise Chen's algorithm to improve the accuracy in forecasting values. Two numerical examples were employed to effectively compare the proposed method with three fuzzy time series methods (Chen, 1996; Hwang et al., 1998; Lee & Chou, 2004) as well as to illustrate the proposed method and evaluate its forecasting performance.

The rest of this paper is arranged as follows: Section 2 briefly introduces the definitions of the fuzzy time series and reviews the major steps of Chen's method (1996). Section 3 discusses the concepts and the detailed steps of the improved fuzzy time series forecasting method while Sect. 4 illustrates the research steps of the proposed method by using two numerical examples. Section 5 compares the forecasting accuracy between the proposed method and three fuzzy time series methods and the last section concludes the present research.

2 Fuzzy time series and Chen's method

Song and Chissom (1993a,b, 1994) defined their fuzzy time series by means of discrete fuzzy sets. The discrete fuzzy sets can be defined as follows:

Let U be the universe of discourse, where $U = \{u_1, u_2, \ldots, u_n\}$. A fuzzy set $\tilde{A}_i$ of U is defined by

$$\tilde{A}_i = \mu_{\tilde{A}_i}(\mu_1)/u_1 + \mu_{\tilde{A}_i}(\mu_2)/u_2 + \cdots + \mu_{\tilde{A}_i}(\mu_n)/u_n \tag{1}$$

where $\mu_{\tilde{A}_i}$ is the membership function of $\tilde{A}_i$, $\mu_{\tilde{A}_i} : U \to [0,1]$. $\mu_{\tilde{A}_i}(u_i)$ denotes the membership value of u_i in $\tilde{A}_i$, $\mu_{\tilde{A}_i}(\mu_i) \in [0, 1]$ and $1 \leqq i \leqq n$.

Next, Song and Chissom (1993a, b, 1994) presented the definitions of the fuzzy time series. The definitions are described as follows:

Definition 1 $Y(t)(t = \ldots, 0, 1, 2, \ldots)$, is a subset of R. Let $Y(t)$ be the universe of discourse defined by the fuzzy set $\mu_i(t)$. If $F(t)$ consists of $\mu_i(t)(i = 1, 2, \ldots)$, $F(t)$ is called a fuzzy time series on $Y(t)$.

Definition 2 If there exists a fuzzy relationship $R(t-1, t)$, such that $F(t) = F(t-1) \circ R(t-1, t)$, where $\circ$ is an arithmetic operator, then $F(t)$ is said to be caused by $F(t-1)$. The relationship between $F(t)$ and $F(t-1)$ can be denoted by $F(t-1) \to F(t)$.

Definition 3 Suppose $F(t)$ is calculated by $F(t-1)$ only, and $F(t) = F(t-1) \circ R(t-1, t)$. For any t, if $R(t-1, t)$ is independent of t, then $F(t)$ is considered a time-invariant fuzzy time series. Otherwise, $F(t)$ is time-variant.

Definition 4 Suppose $F(t-1) = \tilde{A}_i$ and $F(t) = \tilde{A}_j$, a fuzzy logical relationship can be defined as

$$\tilde{A}_i \to \tilde{A}_j,$$

where $\tilde{A}_i$ and $\tilde{A}_j$ are called the left-hand side and the right-hand side of the fuzzy logical relationship, respectively.

Chen (1996) revised the time-invariant models in Song and Chissom (1993a, b) to simplify the calculations. In addition, Chen's method can generate more precise forecasting results than those of Song and Chissom (1993a, b). Chen's method consists of the following major steps:

Step 1: Define the universe of discourse U.
Step 2: Divide U into several equal-length intervals.
Step 3: Define the fuzzy sets on U and fuzzify the historical data.
Step 4: Derive the fuzzy logical relationships based on the historical data.
Step 5: Classify the derived fuzzy logical relationships into groups.
Step 6: Utilize three defuzzification rules to calculate the forecasted values.

3 An improved fuzzy time series forecasting method

The aim of the present research is to develop an improved fuzzy time series method that can both provide the forecasting values in terms of trapezoidal fuzzy numbers and generate more accurate forecasting results at the same time. As mentioned in Sect. 1, we chose Chen's method as a foundation to develop the proposed method. Several modifications between the proposed method and Chen's method are listed below:

1. Use a more advanced method to determine the number of equal-length intervals.
2. Use trapezoidal fuzzy numbers to define the fuzzy sets in fuzzy time series.
3. Apply the arithmetic operations of trapezoidal fuzzy numbers to compute the forecasted values.

First, the number and the length of intervals are assigned subjectively in Chen's method. However, Huarng (2001b) argued that the different number of intervals could affect the accuracy of the forecasting results. To resolve this problem, Huarng designed an average-based length method that can effectively determine the appropriate interval length in order to improve the forecasting results. Hence, the first modification is to employ the average-based length method to determine the appropriate length and number of intervals.

Second, the current fuzzy time series models (Song & Chissom, 1993a, b, 1994; Chen, 1996, 2002; Hwang et al., 1998; Chen & Hwang, 2000; Huarng, 2001a, b; Lee & Chou, 2004) utilize discrete fuzzy sets to define their fuzzy time series. Their discrete fuzzy sets are defined as follows:

Assume there are m intervals, which are $u_1 = [d_1, d_2], u_2 = [d_2, d_3]$, $u_3 = [d_3, d_4], u_4 = [d_4, d_5], \ldots, u_{m-3} = [d_{m-3}, d_{m-2}]$, $u_{m-2} = [d_{m-2}, d_{m-1}]$, $u_{m-1} = [d_{m-1}, d_m]$, and $u_m = [d_m, d_{m+1}]$. Thus, the fuzzy sets $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_m$ are defined by

$$
\begin{aligned}
\tilde{A}_1 &= 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + \cdots + 0/u_m, \\
\tilde{A}_2 &= 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + 0/u_5 + \cdots + 0/u_m, \\
\tilde{A}_3 &= 0/u_1 + 0.5/u_2 + 1/u_3 + 0.5/u_4 + 0/u_5 + \cdots + 0/u_m, \\
&\vdots \qquad\qquad\qquad\qquad \vdots \\
\tilde{A}_{m-1} &= 0/u_1 + 0/u_2 + \cdots + 0/u_{m-3} + 0.5/u_{m-2} + 1/u_{m-1} + 0.5/u_m, \\
\tilde{A}_m &= 0/u_1 + 0/u_2 + \cdots + 0/u_{m-3} + 0/u_{m-2} + 0.5/u_{m-1} + 1/u_m.
\end{aligned}
$$

The present study attempts to replace the above discrete fuzzy sets with trapezoidal fuzzy numbers. A trapezoidal fuzzy number $\tilde{A}$ can be defined as $\tilde{A} = (a, b, c, d)$ with its membership function (Fig. 1).

$$
\mu_{\tilde{A}}(x) = \begin{cases} 0\,, x < a \\ \frac{x-a}{b-a},\ a \le x \le b \\ 1\,,\ b \le x \le c \\ \frac{d-x}{d-c}, c \le x \le d \\ 0\,, x > d \end{cases} . \tag{2}
$$

According to the above definition, the discrete fuzzy sets can be replaced with the following trapezoidal fuzzy numbers:

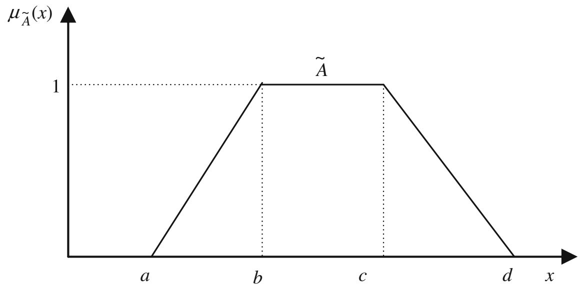

Fig. 1 A trapezoidal fuzzy number $\tilde{A}$

$$
\begin{aligned}
\tilde{A}_1 &= (d_0, d_1, d_2, d_3),\\
\tilde{A}_2 &= (d_1, d_2, d_3, d_4),\\
\tilde{A}_3 &= (d_2, d_3, d_4, d_5),\\
\vdots &\qquad \vdots\\
\tilde{A}_{m-1} &= (d_{m-2}, d_{m-1}, d_m, d_{m+1}),\\
\tilde{A}_m &= (d_{m-1}, d_m, d_{m+1}, d_{m+2}).
\end{aligned}
$$

Third, Chen (1996) developed three heuristic rules to calculate the forecasted values. These three heuristic rules use the midpoints of intervals to derive the forecasted values. To maintain the complete forecasting information, the present study intends to replace the midpoints of intervals with the trapezoidal fuzzy numbers. Specifically, we can apply the addition operation and the scalar multiplication operation of the trapezoidal fuzzy numbers to compute the forecasted values. These two arithmetic operations are listed as follows:

Assume $\tilde{A} = (a_1, b_1, c_1, d_1), \tilde{B} = (a_2, b_2, c_2, d_2)$, and $S > 0$. Thus,

$$\tilde{A} \oplus \tilde{B} = (a_1 + a_2, b_1 + b_2, c_1 + c_2, d_1 + d_2), \tag{3}$$

$$S \cdot \tilde{A} = (Sa_1, Sb_1, Sc_1, Sd_1). \tag{4}$$

According to the above three modifications, the detailed research steps of the proposed method can be described as follows:

Step 1: *Collect the historical data Dv_t.*

Step 2: *Define the universe of discourse U.* Find the maximum $D_{\max}$ and the minimum $D_{\min}$ among all Dv_t. For easy partitioning of U, two small numbers D_1 and D_2 are assigned. The universe of discourse U is then defined by:

$$U = [D_{\min} - D_1, D_{\max} + D_2]. \tag{5}$$

Step 3: *Determine the appropriate length of interval l.* Here, the average-based length method (Huarng, 2001b) can be applied to determine the appropriate l. The length of interval l is computed by the following steps:

Table 1 Base mapping table

Range	Base
0.1–1.0	0.1
1.1–10	1
11–100	10
101–1000	100
1001–10000	1,000

1. Calculate all the absolute differences between the values Dv_{t-1} and Dv_t as the first differences, and then compute the average of the first differences.
2. Take one-half of the average as the length.
3. Find the located range of the length and determine the base (Table 1).
4. According to the assigned base, round the length as the appropriate l.

Step 4: *Define fuzzy numbers.* The number of intervals (fuzzy numbers), m, is computed by

$$m = (D_{\max} + D_2 - D_{\min} + D_1)/l. \tag{6}$$

Thus, there are m intervals and m fuzzy numbers, which are $u_1, u_2, \ldots, u_m$, and $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_m$, respectively. Assume that the m intervals are $u_1 = [d_1, d_2], u_2 = [d_2, d_3], u_3 = [d_3, d_4], \ldots, u_{m-2} = [d_{m-2}, d_{m-1}]$, $u_{m-1} = [d_{m-1}, d_m]$, and $u_m = [d_m, d_{m+1}]$. The fuzzy numbers,$\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_m$, can be defined as follows:

$$\begin{aligned}
\tilde{A}_1 &= (d_0, d_1, d_2, d_3),\\
\tilde{A}_2 &= (d_1, d_2, d_3, d_4),\\
\vdots & \qquad \vdots\\
\tilde{A}_{m-1} &= (d_{m-2}, d_{m-1}, d_m, d_{m+1}),\\
\tilde{A}_m &= (d_{m-1}, d_m, d_{m+1}, d_{m+2}).
\end{aligned}$$

Step 5: *Fuzzify the historical data.* If the value of Dv_t is located in the range of u_j, then it belongs to fuzzy number $\tilde{A}_j$. All Dv_t must be classified into the corresponding fuzzy numbers.

Step 6: *Generate the fuzzy logical relationships.* For all fuzzified data, derive the fuzzy logical relationships based on Definition 4. The fuzzy logical relationship is like $\tilde{A}_j \rightarrow \tilde{A}_k$, which denotes that "if the Dv_{t-1} value of time t-1 is $\tilde{A}_j$, then that of time t is $\tilde{A}_k$."

Step 7: *Establish the fuzzy logical relationship groups*. The derived fuzzy logical relationships can be arranged into fuzzy logical relationship groups based on the same fuzzy numbers on the left-hand sides of the fuzzy logical relationships.

The fuzzy logical relationship groups are like the following:

$$\begin{aligned}\tilde{A}_j &\to \tilde{A}_{k1},\\ \tilde{A}_j &\to \tilde{A}_{k2},\\ &\vdots\\ \tilde{A}_j &\to \tilde{A}_{kp}.\end{aligned}$$

Step 8: *Calculate the forecasted outputs.* The forecasted value at time t, Fv_t, is determined by the following three heuristic rules. Assume the fuzzy number of Dv_{t-1} at time t-1 is $\tilde{A}_j$.

Rule 1: If the fuzzy logical relationship group of $\tilde{A}_j$ is empty; $\tilde{A}_j \to \phi$, then the value of Fv_t is $\tilde{A}_j$, which is $(d_{j-1}, d_j, d_{j+1}, d_{j+2})$.
Rule 2: If the fuzzy logical relationship group of $\tilde{A}_j$ is one-to-one; $\tilde{A}_j \to \tilde{A}_k$, then the value of Fv_t is $\tilde{A}_k$, which is $(d_{k-1}, d_k, d_{k+1}, d_{k+2})$.
Rule 3: If the fuzzy logical relationship group of $\tilde{A}_j$ is one-to-many; $\tilde{A}_j \to \tilde{A}_{k1}$, $\tilde{A}_j \to \tilde{A}_{k2}, \ldots, \tilde{A}_j \to \tilde{A}_{kp}$, and then the value of Fv_t is calculated as follows:

$$\begin{aligned}Fv_t &= \frac{\tilde{A}_{k1} + \tilde{A}_{k2} + \cdots + \tilde{A}_{kp}}{p}\\ &= \left(\frac{d_{k1-1} + \cdots + d_{kp-1}}{p}, \frac{d_{k1} + \cdots + d_{kp}}{p}, \frac{d_{k1+1} + \cdots + d_{kp+1}}{p},\right.\\ &\qquad \left.\frac{d_{k1+2} + \cdots + d_{kp+2}}{p}\right). \end{aligned} \tag{7}$$

where $\tilde{A}_{k1} = (d_{k1-1}, d_{k1}, d_{k1+1}, d_{k1+2}), \tilde{A}_{k2} = (d_{k2-1}, d_{k2}, d_{k2+1}, d_{k2+2}), \ldots$, and $\tilde{A}_{kp} = (d_{kp-1}, d_{kp}, d_{kp+1}, d_{kp+2})$.

4 Numerical examples

The current research illustrates its forecasting method with two numercial examples: the enrollments of the University of Alabama and the number of patents granted in Taiwan.

4.1 Enrollments of the University of Alabama

Step 1: The example, adopted from Chen (1996, 2002), Hwang et al. (1998), Lee and Chou (2004), concerns the number of student enrollments at the University of Alabama from year 1971 to 1992, as shown in Table 2.

Table 2 Enrollments at the University of Alabama

Year	Enrollment
1971	13055
1972	13563
1973	13867
1974	14696
1975	15460
1976	15311
1977	15603
1978	15861
1979	16807
1980	16919
1981	16388
1982	15433
1983	15497
1984	15145
1985	15163
1986	15984
1987	16859
1988	18150
1989	18970
1990	19328
1991	19337
1992	18876

Step 2: Table 2 shows that the maximum and the minimum number of the enrollments are 19,337 (D_{max}) and 13,055 (D_{min}), respectively. For easy computation, let D_1= 5 and D_2= 13. The universe of discourse U is defined as follows:

$$U = [13055 - 5, 19337 + 13] = [13050, 19350].$$

Step 3: The appropriate length of interval l can be computed as follows:

1. Based on Table 2, we can calculate the average of the first differences, which is 510.3.
2. Take one-half of 510.3 as the length, which is 255.15.
3. Since the length 255.15 is located at the range [101, 1000] in Table 1, the base is assigned to be 100.
4. According to the base 100, the length 255.15 is rounded off to 300, which is the appropriate length of interval l.

Step 4: Use Eq. 6 to calculate the number of intervals (fuzzy numbers) as follows:

$$m = \frac{19350 - 13050}{300} = 21.$$

Thus, there are 21 intervals, which are u_1 = [13050, 13350], u_2 = [13350, 13650], u_3 = [13650, 13950], …, u_{19} = [18450, 18750], u_{20} = [18750, 19050], and u_{21} = [19050, 19350].

The fuzzy numbers can be then defined by

$$\begin{aligned}
\tilde{A}_1 &= (12750, 13050, 13350, 13650),\\
\tilde{A}_2 &= (13050, 13350, 13650, 13950),\\
\tilde{A}_3 &= (13350, 13650, 13950, 14250),\\
&\vdots \qquad\qquad \vdots\\
\tilde{A}_{19} &= (18150, 18450, 18750, 19050),\\
\tilde{A}_{20} &= (18450, 18750, 19050, 19350),\\
\tilde{A}_{21} &= (18750, 19050, 19350, 19650).
\end{aligned}$$

Step 5: Fuzzify the enrollments. For example, the enrollment in year 1971 is 13055, which is located at the range of u_1= [13050, 13350]. Thus, the corresponding fuzzy number of year 1971 is assigned as $\tilde{A}_1$.

Table 3 lists the corresponding fuzzy number for the enrollment of each year.

Step 6: According to Table 3, we can derive the fuzzy logical relationships as shown in Table 4. Notice that the repeated relationships are counted only once.

Step 7: Based on the same fuzzy numbers on the left-hand side of the fuzzy logical relationships in Table 4, 13 fuzzy logical relationship groups are generated as shown in Table 5.

Step 8: According to Tables 3 and 5, we can calculate the forecasted enrollments. For instance, the forecasted enrollments of years 1972 and 1976 can be illustrated below:

[1972]: The fuzzified enrollment of year 1971 in Table 3 is $\tilde{A}_1$, and from Table 5, we can find that there is one fuzzy logical relationship in group 1.

$$\tilde{A}_1 \rightarrow \tilde{A}_2.$$

According to *Rule* 2, the forecasted enrollment of year 1972 is $\tilde{A}_2$. Thus, Fv_{1972}= (13050, 13350, 13650, 13950).

[1976]: Because the fuzzified enrollment of 1975 in Table 3 is $\tilde{A}_9$, and from Table 5, we can find that there are three fuzzy logical relationships in group 7.

$$\tilde{A}_9 \rightarrow \tilde{A}_7, \qquad \tilde{A}_9 \rightarrow \tilde{A}_8, \qquad \tilde{A}_9 \rightarrow \tilde{A}_{10}.$$

According to *Rule* 3, the forecasted enrollment of year 1976 is computed as follows:

$$Fv_{1976} = \frac{\tilde{A}_7 + \tilde{A}_8 + \tilde{A}_{10}}{3}$$

Table 3 Corresponding fuzzy numbers of the enrollments

Year	Enrollment	Fuzzy number
1971	13055	$\tilde{A}_1$
1972	13563	$\tilde{A}_2$
1973	13867	$\tilde{A}_3$
1974	14696	$\tilde{A}_6$
1975	15460	$\tilde{A}_9$
1976	15311	$\tilde{A}_8$
1977	15603	$\tilde{A}_9$
1978	15861	$\tilde{A}_{10}$
1979	16807	$\tilde{A}_{13}$
1980	16919	$\tilde{A}_{13}$
1981	16388	$\tilde{A}_{12}$
1982	15433	$\tilde{A}_8$
1983	15497	$\tilde{A}_9$
1984	15145	$\tilde{A}_7$
1985	15163	$\tilde{A}_8$
1986	15984	$\tilde{A}_{10}$
1987	16859	$\tilde{A}_{13}$
1988	18150	$\tilde{A}_{17}$
1989	18970	$\tilde{A}_{20}$
1990	19328	$\tilde{A}_{21}$
1991	19337	$\tilde{A}_{21}$
1992	18876	$\tilde{A}_{20}$

Table 4 Fuzzy logical relationships of the enrollments

$\tilde{A}_1 \to \tilde{A}_2$	$\tilde{A}_2 \to \tilde{A}_3$	$\tilde{A}_3 \to \tilde{A}_6$	$\tilde{A}_6 \to \tilde{A}_9$
$\tilde{A}_9 \to \tilde{A}_8$	$\tilde{A}_8 \to \tilde{A}_9$	$\tilde{A}_9 \to \tilde{A}_{10}$	$\tilde{A}_{10} \to \tilde{A}_{13}$
$\tilde{A}_{13} \to \tilde{A}_{13}$	$\tilde{A}_{13} \to \tilde{A}_{12}$	$\tilde{A}_{12} \to \tilde{A}_8$	$\tilde{A}_9 \to \tilde{A}_7$
$\tilde{A}_7 \to \tilde{A}_8$	$\tilde{A}_8 \to \tilde{A}_{10}$	$\tilde{A}_{13} \to \tilde{A}_{17}$	$\tilde{A}_{17} \to \tilde{A}_{20}$
$\tilde{A}_{20} \to \tilde{A}_{21}$	$\tilde{A}_{21} \to \tilde{A}_{21}$	$\tilde{A}_{21} \to \tilde{A}_{20}$	

$$
\begin{aligned}
&= \left(\frac{14550 + 14850 + 15450}{3}, \frac{14850 + 15150 + 15750}{3},\right. \\
&\quad \left.\frac{15150 + 15450 + 16050}{3}, \frac{15450 + 15750 + 16350}{3}\right) \\
&= (14950, 15250, 15550, 15850).
\end{aligned}
$$

In a similar way, we can calculate the forecasted enrollments for the other years. Table 6 shows the forecasted enrollments from years 1972 to 1992.

In Table 6, let us assume that a decision analyst is interested to find out the possible forecasted interval under α degree of confidence ($0 \leqq \alpha \leqq 1$). Here, we can apply the α-cut concept of the fuzzy number to obtain the forcasted interval. Take year 1976 as an example when the forecasted enrollment is (14950, 15250, 15550, 15850). Suppose the decision analyst wants to know the forecasted interval under 0.8 degree of confidence. By using Eq. 2, we can compute for the

Table 5 Fuzzy logical relationship groups

Group	Fuzzy logical relationships
1	$\tilde{A}_1 \to \tilde{A}_2$
2	$\tilde{A}_2 \to \tilde{A}_3$
3	$\tilde{A}_3 \to \tilde{A}_6$
4	$\tilde{A}_6 \to \tilde{A}_9$
5	$\tilde{A}_7 \to \tilde{A}_8$
6	$\tilde{A}_8 \to \tilde{A}_9, \quad \tilde{A}_8 \to \tilde{A}_{10}$
7	$\tilde{A}_9 \to \tilde{A}_7, \quad \tilde{A}_9 \to \tilde{A}_8, \quad \tilde{A}_9 \to \tilde{A}_{10}$
8	$\tilde{A}_{10} \to \tilde{A}_{13}$
9	$\tilde{A}_{12} \to \tilde{A}_8$
10	$\tilde{A}_{13} \to \tilde{A}_{12}, \quad \tilde{A}_{13} \to \tilde{A}_{13}, \quad \tilde{A}_{13} \to \tilde{A}_{17}$
11	$\tilde{A}_{17} \to \tilde{A}_{20}$
12	$\tilde{A}_{20} \to \tilde{A}_{21}$
13	$\tilde{A}_{21} \to \tilde{A}_{20}, \quad \tilde{A}_{21} \to \tilde{A}_{21}$

Table 6 Forecasted enrollment of each year

Year	Enrollment	Forecasted enrollment
1971	13055	
1972	13563	(13050, 13350, 13650, 13950)
1973	13867	(13350, 13650, 13950, 14250)
1974	14696	(14250, 14550, 14850, 15150)
1975	15460	(15150, 15450, 15750, 16050)
1976	15311	(14950, 15250, 15550, 15850)
1977	15603	(15300, 15600, 15900, 16200)
1978	15861	(14950, 15250, 15550, 15850)
1979	16807	(16350, 16650, 16950, 17250)
1980	16919	(16650, 16950, 17250, 17550)
1981	16388	(16650, 16950, 17250, 17550)
1982	15433	(14850, 15150, 15450, 15750)
1983	15497	(15300, 15600, 15900, 16200)
1984	15145	(14950, 15250, 15550, 15850)
1985	15163	(14850, 15150, 15450, 15750)
1986	15984	(15300, 15600, 15900, 16200)
1987	16859	(16350, 16650, 16950, 17250)
1988	18150	(16650, 16950, 17250, 17550)
1989	18970	(18450, 18750, 19050, 19350)
1990	19328	(18750, 19050, 19350, 19650)
1991	19337	(18600, 18900, 19200, 19500)
1992	18876	(18600, 18900, 19200, 19500)

forecasted interval, $[F_L, F_R]$, as follows:

$$0.8 = \frac{F_L - 14950}{15250 - 14950}, \qquad 0.8 = \frac{15850 - F_R}{15850 - 15550},$$
$$\Rightarrow F_L = 15190, \qquad F_R = 15610.$$

Therefore, the forecasted interval for year 1976 under 0.8 degree of confidence is [15190, 15610]. Table 7 shows all possible forecasted intervals for year 1976 from $\alpha = 0$ to 1.

Table 7 Forecasted intervals for year 1976 (from α= 0 to 1)

Degree of confidence	Interval estimate
$\alpha = 0$	[14950, 15850]
$\alpha = 0.1$	[14980, 15820]
$\alpha = 0.2$	[15010, 15790]
$\alpha = 0.3$	[15040, 15760]
$\alpha = 0.4$	[15070, 15730]
$\alpha = 0.5$	[15100, 15700]
$\alpha = 0.6$	[15130, 15670]
$\alpha = 0.7$	[15160, 15640]
$\alpha = 0.8$	[15190, 15610]
$\alpha = 0.9$	[15220, 15580]
$\alpha = 1.0$	[15250, 15550]

Table 8 Number of patents granted in Taiwan

Year	Patents granted
1980	6633
1981	6264
1982	7460
1983	7096
1984	8592
1985	9427
1986	10526
1987	10615
1988	12355
1989	19265
1990	22601
1991	27281
1992	21264
1993	22317
1994	19032
1995	29707
1996	29469
1997	29356
1998	25051
1999	29144
2000	38665

From Table 7, the decision analyst can find the possible forecasted intervals under different degrees of confidence. Unlike the existing fuzzy time series methods, the proposed method can provide the decision analyst various interval estimates instead of a single-point forecasted value offered by existing methods.

4.2 Number of patents granted in Taiwan

To further illustrate the proposed method, another numerical example, the number of patents granted in Taiwan, are employed from National Science Council (1995–2004).

Steps 1–2: The figures on the number of patents granted in Taiwan from years 1980 to 2000 are listed in Table 8.

Table 9 Corresponding fuzzy numbers of the patents granted

Year	Patents granted	Fuzzy number
1980	6633	$\tilde{A}_1$
1981	6264	$\tilde{A}_1$
1982	7460	$\tilde{A}_1$
1983	7096	$\tilde{A}_1$
1984	8592	$\tilde{A}_2$
1985	9427	$\tilde{A}_2$
1986	10526	$\tilde{A}_3$
1987	10615	$\tilde{A}_3$
1988	12355	$\tilde{A}_4$
1989	19265	$\tilde{A}_7$
1990	22601	$\tilde{A}_9$
1991	27281	$\tilde{A}_{11}$
1992	21264	$\tilde{A}_8$
1993	22317	$\tilde{A}_9$
1994	19032	$\tilde{A}_7$
1995	29707	$\tilde{A}_{12}$
1996	29469	$\tilde{A}_{12}$
1997	29356	$\tilde{A}_{12}$
1998	25051	$\tilde{A}_{10}$
1999	29144	$\tilde{A}_{12}$
2000	38665	$\tilde{A}_{17}$

Table 8 shows that the minimum and the maximum number of the patents granted are 6264 ($D_{\min}$) and 38,665 ($D_{\max}$), respectively. For easy computation, let $D_1 = 264$ and $D_2 = 1335$. The universe of discourse can be definded by

$$U = [6264 - 264, 38665 + 1335] = [6000, 40000].$$

Steps 3–4: According to Table 8, we can calculate the appropriate length of interval, which is 2000. The number of intervals (fuzzy numbers) is then computed as 17. Thus, there are 17 intervals, which are $u_1 = [6000, 8000]$, $u_2 = [8000, 10000], \ldots, u_{16} = [36000, 38000]$, and $u_{17} = [38000, 40000]$.

The fuzzy numbers are then defined by

$$\begin{aligned} \tilde{A}_1 &= (4000, 6000, 8000, 10000), \\ \tilde{A}_2 &= (6000, 8000, 10000, 12000), \\ \vdots & \qquad \vdots \\ \tilde{A}_{16} &= (34000, 36000, 38000, 40000), \\ \tilde{A}_{17} &= (36000, 38000, 40000, 42000). \end{aligned}$$

Step 5: Fuzzify the patents granted. Table 9 shows the corresponding fuzzy numbers of patents granted from years 1980 to 2000.

Steps 6–7: According to Table 9, we can derive the fuzzy logical relationships and then establish the following logical relationship groups (Table 10).

Step 8: Based on Tables 9 and 10, the forecasted patents granted are calculated and shown in Table 11.

Table 10 Fuzzy logical relationship groups

Group	Fuzzy logical relationships
1	$\tilde{A}_1 \to \tilde{A}_1, \quad \tilde{A}_1 \to \tilde{A}_2$
2	$\tilde{A}_2 \to \tilde{A}_2, \quad \tilde{A}_2 \to \tilde{A}_3$
3	$\tilde{A}_3 \to \tilde{A}_3, \quad \tilde{A}_3 \to \tilde{A}_4$
4	$\tilde{A}_4 \to \tilde{A}_7$
5	$\tilde{A}_7 \to \tilde{A}_9, \quad \tilde{A}_7 \to \tilde{A}_{12}$
6	$\tilde{A}_8 \to \tilde{A}_8$
7	$\tilde{A}_9 \to \tilde{A}_7, \quad \tilde{A}_9 \to \tilde{A}_{11}$
8	$\tilde{A}_{10} \to \tilde{A}_{12}$
9	$\tilde{A}_{11} \to \tilde{A}_8$
10	$\tilde{A}_{12} \to \tilde{A}_{10}, \quad \tilde{A}_{12} \to \tilde{A}_{12}, \quad \tilde{A}_{12} \to \tilde{A}_{17}$

Table 11 Forecasted number of patents granted

Year	Patents granted	Forecasted patents
1980	6633	
1981	6264	(5000, 7000, 9000, 11000)
1982	7460	(5000, 7000, 9000, 11000)
1983	7096	(5000, 7000, 9000, 11000)
1984	8592	(5000, 7000, 9000, 11000)
1985	9427	(7000, 9000, 11000, 13000)
1986	10526	(7000, 9000, 11000, 13000)
1987	10615	(9000, 11000, 13000, 15000)
1988	12355	(9000, 11000, 13000, 15000)
1989	19265	(16000, 18000, 20000, 22000)
1990	22601	(23000, 25000, 27000, 29000)
1991	27281	(20000, 22000, 24000, 26000)
1992	21264	(18000, 20000, 22000, 24000)
1993	22317	(18000, 20000, 22000, 24000)
1994	19032	(20000, 22000, 24000, 26000)
1995	29707	(23000, 25000, 27000, 29000)
1996	29469	(28000, 30000, 32000, 34000)
1997	29356	(28000, 30000, 32000, 34000)
1998	25051	(28000, 30000, 32000, 34000)
1999	29144	(26000, 28000, 30000, 32000)
2000	38665	(28000, 30000, 32000, 34000)

In Table 11, suppose a decision analyst wants to find out the possible forecasted intervals for year 2000. By using Eq. 2, we can compute all possible forecasted intervals under different degrees of confidence, as shown in Table 12.

From Table 12, a decision analyst can easily evaluate the forecasted ranges while making a forecast.

5 Forecasting performance validation

As described in Sect. 1, the present study can both provide various interval estimates and increase the accuracy of the forecasting results. To evaluate the forecasting performance, three fuzzy time series methods (Chen, 1996; Hwang

Table 12 Forecasted intervals for year 2000 (from α= 0 to 1)

Degree of confidence	Interval estimate
$\alpha = 0$	[28000, 34000]
$\alpha = 0.1$	[28200, 33800]
$\alpha = 0.2$	[28400, 33600]
$\alpha = 0.3$	[28600, 33400]
$\alpha = 0.4$	[28800, 33200]
$\alpha = 0.5$	[29000, 33000]
$\alpha = 0.6$	[29200, 32800]
$\alpha = 0.7$	[29400, 32600]
$\alpha = 0.8$	[29600, 32400]
$\alpha = 0.9$	[29800, 32200]
$\alpha = 1.0$	[30000, 32000]

Table 13 Forecasted enrollments of the four methods

Year	Enroll-ment	Chen's method	Hwang et al.'s method (at $w=2$)	Hwang et al.'s method (at $w=3$)	Hwang et al.'s method (at $w=4$)	Lee and Chou's method	The proposed method
1971	13055						
1972	13563	14000				14025	13500
1973	13867	14000				14568	13800
1974	14696	15500	14267			14568	14700
1975	15460	16000	15296	15296		15654	15600
1976	15311	16000	16260	16260	16260	15654	15400
1977	15603	16000	15711	15711	15511	15654	15750
1978	15861	16000	15803	16003	16003	15654	15400
1979	16807	16833	16261	16261	16261	16197	16800
1980	16919	16833	17409	17407	17407	17283	17100
1981	16388	16833	17319	17119	17119	17283	17100
1982	15433	16833	16188	16188	16188	16197	15300
1983	15497	16000	14833	14833	14833	15654	15750
1984	15145	16000	15097	15297	15497	15654	15400
1985	15163	16000	14945	14745	14745	15654	15300
1986	15984	16000	14963	15163	15163	15654	15750
1987	16859	16000	16384	16384	16384	16197	16800
1988	18150	16833	17659	17659	17659	17283	17100
1989	18970	19000	19150	19150	19150	18369	18900
1990	19328	19000	19970	19770	19770	19454	19200
1991	19337	19000	19928	19928	19928	19454	19050
1992	18876	19000	19537	19537	19537	19454	19050

et al., 1998; Lee & Chou, 2004) are adopted for comparing of their forecasting results with those obtained by the proposed method. Two evaluation indices—the mean absolute percent error (MAPE) and the mean square error (MSE)—are selected to evaluate the forecasting accuracy. The formulas of both indices are provided below:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Dv_t - Fv_t}{Dv_t}\right|, \quad \text{MSE} = \frac{1}{n}\sum_{t=1}^{n}(Dv_t - Fv_t)^2. \tag{8}$$

Table 14 Forecasting performance of the four methods

vIndex	Chen's method	Hwang et al.'s method (at $w = 2$)	Hwang et al.'s method (at $w = 3$)	Hwang et al.'s method (at $w = 4$)	Lee and Chou's method	The proposed method
MAPE (%)	3.08	2.99	2.94	3.12	2.69	1.33
MSE	398051	333274	299634	321418	255227	108097

Table 15 Forecasted patents of the four methods

vYear	Patents granted	Chen's method	Hwang et al.'s method (at w = 2)	Hwang et al.'s method (at w = 3)	Hwang et al.'s method (at w = 4)	Lee and Chou's method	The proposed method
1980	6633						
1981	6264	10700				8259	8000
1982	7460	10700				8259	8000
1983	7096	10700	8485			8259	8000
1984	8592	10700	8121	8121		8259	8000
1985	9427	10700	9617	9617	9617	13941	10000
1986	10526	10700	10452	10452	10452	13941	10000
1987	10615	10700	11551	11551	11551	13941	12000
1988	12355	10700	11640	11640	11640	13941	12000
1989	19265	17750	13380	13380	13380	13941	19000
1990	22601	27150	23140	23140	23140	25305	26000
1991	27281	22450	27901	26476	26476	22464	23000
1992	21264	28717	31156	31156	31156	30040	21000
1993	22317	22450	21264	21264	21264	22464	21000
1994	19032	22450	20492	24767	24767	22464	23000
1995	29707	27150	17207	14357	14357	25305	26000
1996	29469	27150	29707	29707	29707	30040	31000
1997	29356	28717	29469	27644	30494	30040	31000
1998	25051	28717	30381	30381	30381	30040	31000
1999	29144	28717	23226	23226	23226	22464	29000
2000	38665	28717	29144	31594	31594	30040	31000

Since the forecasted outputs of the present research are trapezoidal fuzzy numbers, we can use the centroid method to defuzzify these numbers and compare them with the forecasted values generated by three forecasting methods (Chen, 1996; Hwang et al., 1998; Lee & Chou, 2004). The formula of the centroid method is given below:

$$x^* = \frac{\int x \cdot \mu_{\tilde{A}}(x)\, \mathrm{d}x}{\int \mu_{\tilde{A}}(x)\, \mathrm{d}x}, \tag{9}$$

where x^* is the defuzzified value.

Table 16 Forecasting performance of the four methods

Index	Chen's method	Hwang et al.'s method (at $w = 2$)	Hwang et al.'s method (at $w = 3$)	Hwang et al.'s method (at $w = 4$)	Lee and Chou's method	The proposed method
MAPE (%)	18.94	13.67	15.19	15.68	19.44	10.45
MSE	14524448	25021153	30653762	32453514	18172556	8378697

Furthermore, the two numercial examples in Sect. 4 are employed to compare the forecasting accuracy between our method and three forecasting methods. Table 13 shows the forecasted enrollments obtained by these four methods.

Notice that the w value in Hwang et al. (1998) forecasts better between 2 and 4 in their own emprical analysis. Thus, in their approach, we just calculate the forecasted enrollments from $w = 2$ to 4.

To compare the forecasting accuracy, the MAPE values and the MSE values of these four methods are computed and shown in Table 14.

Table 14 indicates that the MAPE value (1.33%) of the proposed method is smaller than those of the methods (Chen, 1996; Hwang et al., 1998; Lee & Chou, 2004). Similarly, the MSE value (108097) of the proposed method is also smaller than those of the methods. Therefore, both evaluation indices indicate that the proposed method can decrease the forecasting error of the enrollments.

Next, we can calculate the forecasted patents granted by these four methods. The computation results are as shown in Table 15.

Moreover, the MAPE values and the MSE values of the four methods are computed and shown in Table 16.

Table 16 shows that both the MAPE value (10.45%) and the MSE value (8378697) of the proposed method are the smallest among the four methods. Hence, both evaluation indices indicate that the proposed method can decrease the forecasting error of the patents granted compared to those obtained from Chen (1996), Hwang et al. (1998), Lee and Chou (2004).

According to Tables 14 and 16, it is clear that the proposed method can provide better forecasting results than these three methods (Chen, 1996; Hwang et al., 1998; Lee & Chou, 2004).

6 Conclusion

In this paper, we have developed an improved fuzzy time series forecasting method can provide a decision analyst various forecasted intervals under different degrees of confidence and at the same time decrease the forecasting error and improve the forecasting accuracy. Furthermore, two numerical examples were employed to illustrate the proposed method and compare the forecasting accuracy with three fuzzy time series methods (Chen, 1996; Hwang et al., 1998;

Lee & Chou, 2004). Both Tables 14 and 16 indicate that the proposed method is preferrable to those in these three methods with regard to both MAPE and MSE values. The results therefore prove that the present study can provide a decision analyst a better decision-aid tool in making predictions.

Acknowledgements This research was partially supported by the National Science Council, Republic of China, under Grant NSC 91-2213-E-214-042.

References

Chen, S.M. (1996). Forecasting enrollments based on fuzzy time series. *Fuzzy Sets and Systems, 81*, 311–319.

Chen, S.M. (2002). Forecasting enrollments based on high-order fuzzy time-series. *Cybernetics and Systems, 33*, 1–16.

Chen, S.M., & Hwang, J.R. (2000). Temperature prediction using fuzzy time series. *IEEE Transaction on Systems, Man, & Cybernetics, 30*, 263–275.

Huarng, K.H. (2001a). Heuristic models of fuzzy time series for forecasting. *Fuzzy Sets and Systems, 123*, 369–386.

Huarng, K.H. (2001b). Effective lengths of intervals to improve forecasting in fuzzy time series. *Fuzzy Sets and Systems, 123*, 387–394.

Hwang, J.R., Chen, S.M., & Lee, C.H. (1998). Handling forecasting problems using fuzzy time series. *Fuzzy Sets and System, 100*, 217–228.

Lee, H.S., & Chou, M.T. (2004). Fuzzy forecasting based on fuzzy time series. *International Journal of Computer Mathematics, 81*, 781–789.

National Science Council (1995–2004). *Indicator of science and technology Republic of China*. R.O.C.: National Science Council.

Song, Q., & Chissom, B.S. (1993a). Fuzzy forecasting enrollments with fuzzy time series-Part 1. *Fuzzy Sets and Systems, 54*, 1–9.

Song, Q., & Chissom, B.S. (1993b). Fuzzy Time Series and Its Models. *Fuzzy Sets and Systems, 54*, 269–277.

Song, Q., & Chissom, B.S. (1994). Fuzzy forecasting enrollments with fuzzy time series-Part 2. *Fuzzy Sets and Systems, 62*, 1–8.