Forecasting Model for Enrolment Combining Weighted Fuzzy Time Series and Fourier Series Transform

Nghiem Van Tinh
Thai Nguyen University of Technology - Thai Nguyen University, Thai Nguyen, Vietnam

Nguyen Tien Duy
Thai Nguyen University of Technology - Thai Nguyen University, Thai Nguyen, Vietnam

***Abstract*—Fuzzy time series (FTS) methods was first introduced by Song and Chissom (1993, 1994) based on the fuzzy set theory proposed by Zadeh (1965). Over the earlier few years, some methods have been presented based on fuzzy time series to forecast real problems, such as forecasting stock market, temperature prediction, forecasting enrolments, disease diagnosing, etc. Traditionally, time series forecasting problems are being solved using a class of Autoregressive moving average models. Being linear statistical models, they cannot build relationship among the nonlinear variables. Calculating the parameters for multi-variables is another issue faced by them. The strong relationship among these variables may result in large errors. Furthermore, a model cannot be estimated correctly if the historical data is less. Therefore, this paper, we propose a new fuzzy forecasting model to overcome the drawbacks of the traditional forecasting models that aim increasing the forecasting accuracy. In our studies, a hybrid forecasting model based on aggregated FTS and Fourier series analysis. Firstly, we propose weighted models to tackle two issues in fuzzy time series forecasting, namely, recurrence and weighting. Then, the using Fourier series to modify the residuals of the weighted FTS for improving the forecasting performance. By using the enrolment data at the University of Alabama from 1971s to 1992s as the forecasting target, the empirical results show that the proposed model outperforms one of the conventional FTS models**

Keywords — Fuzzy time series(FTS), fuzzy logical relationship groups (FLRG), forecasting, Fourier series, enrollments.

I. INTRODUCTION

Fuzzy time series procedures, which have attracted the attention of many researchers in recent years, have a quite wide area of use, such as information technology, economy, environmental sciences and hydrology and it play an important role in our daily life. Therefore, many more forecasting models have been developed to deal with various problems in order to help people to make decisions, such as crop forecast [6], [7] academic enrolments [1], [10], the temperature prediction [13], stock markets [14], etc. There is the matter of fact that the traditional forecasting methods cannot deal with the forecasting problems in which the historical data are represented by linguistic values. Ref. [1], [2] proposed the time-invariant FTS and the time-variant FTS model which use the max–min operations to forecast the enrolments of the University of Alabama. However, the main drawback of these methods is enormous computation load. Then, Ref. [3] proposed the first-order FTS model by introducing a more efficient arithmetic method. After that, FTS has been widely studied to improve the accuracy of forecasting in many applications. Ref. [4] considered the trend of the enrolment in the past years and presented another forecasting model based on the first-order FTS. He pointed out that the effective length of the intervals in the universe of discourse can affect the forecasting accuracy rate. In other words, the choice of the length of intervals can improve the forecasting results. Ref. [5] presented a heuristic model for fuzzy forecasting by integrating Chen's fuzzy forecasting method [3]. At the same time, Ref.[8] proposed several forecast models based on the high-order fuzzy time series to deal with the enrolments forecasting problem. In [9], the length of intervals for the FTS model was adjusted to forecast the Taiwan Stock Exchange (TAIEX).

Clearly, fuzzy time series model has been applied to environmental sciences, business and engineering; however, fuzzy time series model is still necessary to overcome its drawbacks. In order to cope with these drawbacks in the fuzzy time series model, residual analysis becomes quite important in order to reuse some possible useful information [19].

In this paper, we proposed a hybrid forecasting model combining the weighted fuzzy relationship groups and Fourier series technique. In case study, we applied the proposed method to forecast the enrolments of the University of Alabama. Computational results show that the proposed model outperforms other existing methods.

The rest of this paper is organized as follows. A brief review of the theory of fuzzy time series is described in Section 2. In Section 3, a novel forecasting model base on the weighted fuzzy time series and Fourier series transform. Experiments are presented in Section 4, and some concluding remarks are given in Section 5.

II. FUZZY TIME SERIES

In this section, we provide briefly some definitions of fuzzy time series.

In [1], [2] Song and Chissom proposed the definition of fuzzy time series based on fuzzy sets, Let $U=\{u_1,u_2,\ldots,u_n\}$ be an universal set; a fuzzy set A of U

is defined as $A=\{ f_A(u_1)/u_1+\ldots+f_A(u_n)/u_n \}$, where f_A is a membership function of a given set A, $f_A: U \to [0,1]$, $f_A(u_i)$ indicates the grade of membership of u_i in the fuzzy set A, $f_A(u_i) \in [0, 1]$, and $1 \le i \le n$. General definitions of fuzzy time series are given as follows:

Definition 1: Fuzzy time series

Let $Y(t)$ $(t = ..., 0, 1, 2 ...)$, a subset of R, be the universe of discourse on which fuzzy sets $f_i(t)$ (i = 1,2…) are defined and if $F(t)$ be a collection of $f_i(t)$) (i = 1, 2…). Then, $F(t)$ is called a fuzzy time series on $Y(t)$ (t . . ., 0, 1,2, . . .).

Definition 2: Fuzzy logic relationship

If there exists a fuzzy relationship R(t-1,t), such that F(t) = F(t-1)$*$R(t-1,t), where "$*$" is an max - min 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)→ F(t). Let A_i = F(t) and A_j = F(t-1), the relationship between F(t) and F(t -1) is denoted by fuzzy logical relationship $A_i \to A_j$ where A_i and A_j refer to the current state or the left hand side and the next state or the right-hand side of fuzzy time series.

Definition 3: λ- order fuzzy time series

Let $F(t)$ be a fuzzy time series. If F(t) is caused by F(t-1), F(t-2),…, F(t-λ+1) F(t-λ) then this fuzzy relationship is represented by by F(t-λ), …, F(t-2), F(t-1)→ F(t) and is called an $\boldsymbol{\lambda}$- order fuzzy time series.

Definition 4: Fuzzy Relationship Group (FLRG)

Fuzzy logical relationships in the training datasets with the same fuzzy set on the left-hand-side can be further grouped into a fuzzy logical relationship groups. Suppose there are relationships such that

$$A_i \to A_j$$
$$A_i \to A_k$$
…….

So, these fuzzy logical relationships can be grouped into the same FLRG as : $A_i \to A_j, A_k \ldots$

III. Forecasting Model Based On Combined Weighted Fuzzy Time Series and Fourier Series Analysis

An improved hybrid model for forecasting the enrolments of University of Alabama combining the weighted FTS and Fourier series. At first, we present the weighted forecasting model based FTS in Subsection A. Base on the obtained forecasting results, we can adjust them by using Fourier series analysis to increase forecasting accuracy in Subsection B.

A. Weighted FTS for enrolment forecasting

To verify the effectiveness of the proposed model, all historical enrolments in Table 1 (*the enrolment data at the University of Alabama from 1971s to 1992s)* are used to illustrate for forecasting process. The step-wise procedure of the forecasting model is presented as following:

Step 1*: Defining the universe of discourse and intervals for observations.*

Assume Y(t) be the historical data of enrolments at year t($1971 \le t \le 1992$). The universe of discourse is defined as $U = [D_{min}, D_{max}]$. In order to ensure the forecasting values bounded in the universe of discourse U, we set $D_{min} = I_{min} - N_1$ and $D_{max} = I_{max} + N_2$; where I_{min}, I_{max} are the minimum and maximum data of $Y(t)$; N_1 and N_2 are two proper positive integers to tune the lower bound and upper bound of the U. Based on I_{min} and I_{max}, we define the universal discourse U. From the historical data shown in Table 1, we obtain $I_{min} = 13055$ và $I_{max} = 19337$. Thus, the universe of discourse is defined as U= [$I_{min} - N_1, I_{max} + N_2$]= [13000,20000] with $N_1 = 55$ and $N_2 = 663$.

Now, divide the universe of discourse into n equal lengths of intervals $u_1, u_2, \ldots, u_n$. Each interval u_i of time series data set can be calculated as follows:

$$u_i = (D_{min} + (i-1)\frac{D_{max}-D_{min}}{n}, D_{min} + i\frac{D_{max}-D_{min}}{n}] \quad (1)$$

Compared to the previous models in [2], [3], [4]. Based on Eq.(1), we cut U into seven intervals, $u_1, u_2, \ldots, u_7$ ($1 \le i \le n$, $n = 7$) respectively. Thus, the seven intervals are: u_1 = (13000, 14000], u_2 = (14000,15000], …, u_6 = (18000,19000], u_7 = (19000, 20000].

TABLE I: Historical enrolments of the University of Alabama

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

Step 2: *Define the fuzzy sets for each observations*

Assume that there are n intervals $u_1, u_1, u_1, \ldots, u_n$ for data set obtained in Step 1. For n intervals, there are n linguistic values which are $A_1, A_2, A_3, \ldots, A_{n-1}$ and A_n to represent different regions in the universe of discourse, respectively. Each linguistic variable represents a fuzzy set A_i $(1 \le i \le n)$ and its definition is described in (2) .

$$A_i = \sum_{j=1}^{n} \frac{a_{ij}}{u_j}; \quad (2)$$

where $a_{ij} \in [0,1]$, $1 \le i \le n$, $1 \le j \le n$ and u_j is the j-th interval. The value of a_{ij} indicates the grade of membership of u_j in the fuzzy set A_i and it is shown as following:

$$a_{ij} = \begin{matrix} 1 & \text{if } j == i \\ 0.5 & \text{if } j == i-1 \text{ or } j == i+1 \\ 0 & \text{otherwise} \end{matrix} \quad (3)$$

From Eq.(2) and Eq.(3), each fuzzy set A_i ($1 \le i \le 7$) is defined as follows:

$$A_1 = \frac{1}{u_1} + \frac{0.5}{u_2} + \frac{0}{u_3} + \cdots. + \frac{0}{u_6} + \frac{0}{u_7}$$

$$A_2 = \frac{0.5}{u_1} + \frac{1}{u_2} + \frac{0.5}{u_3} + \cdots . + \frac{0}{u_6} + \frac{0}{u_7}$$
....
$$A_7 = \frac{0}{u_1} + \frac{0}{u_2} + \frac{0}{u_3} + \cdots . + \frac{0.5}{u_6} + \frac{1}{u_7}$$

***Step 3**: Fuzzify variations of the historical data (or observations)*

In order to fuzzify all historical data, it's necessary to assign a corresponding linguistic value to each interval first. The simplest way is to assign the linguistic value with respect to the corresponding fuzzy set that each interval belongs to with the highest membership degree. As in [2, 3], a historical data is fuzzified to A_i if the maximal degree of membership of that datum is in A_i.

Fuzzify(AD_t) = A_i if $F_{AD(t)}(A_i) = max[F_{AD(t)}(A_k)]$ for all k, where k =1,…,AD_t is the Actual data at time t; and $F_{AD(t)}(A_k)]$ is the degree of membership of AD_t under A_k. For example, from Table 1, we can see that the actual data of year 1971 is 13055, where 13055 falls in the interval u = [13055, 14000]. Therefore, the enrolment of year 1971 (i.e., 13055) is fuzzified into A_1. The results of fuzzification are listed in Table 2, where all historical data are fuzzified to be fuzzy sets.

TABLE II: FUZZIFIED ENROLMENTS OF THE UNIVERSITY OF ALABAMA

Year	Actual data	Fuzzy set	Year	Actual data	Fuzzy set
1971	13055	A1	1982	15433	A3
1972	13563	A1	1983	15497	A3
1973	13867	A1	1984	15145	A3
1974	14696	A2	1985	15163	A3
1975	15460	A3	1986	15984	A3
1976	15311	A3	1987	16859	A4
1977	15603	A3	1988	18150	A6
1978	15861	A3	1989	18970	A6
1979	16807	A4	1990	19328	A7
1980	16919	A4	1991	19337	A7
1981	16388	A4	1992	18876	A6

***Step 4**: Establishing all fuzzy logical relationships*

Relationships are identified from the fuzzified historical data obtained in Step 3. If the fuzzified enrolments of years t and t - 1 are A_i and A_j, respectively, then construct the first – order fuzzy logical relationship "$A_i \rightarrow A_j$", where A_i and A_j are called the fuzzy set on the left-hand side and fuzzy set on the right-hand side of fuzzy logical relationships, respectively. From Table 2, we can obtain fuzzy relationships are shown in Table 3 as follows:

TABLE III: THE FIRST-ORDER FUZZY LOGICAL RELATIONSHIP

No	Fuzzy relations	No	Fuzzy relations
1	A1 -> A1	11	A4 -> A3
2	A1 -> A1	12	A3 -> A3
3	A1 -> A2	13	A3 -> A3
4	A2 -> A3	14	A3 -> A3
5	A3 -> A3	15	A3 -> A3
6	A3 -> A3	16	A3 -> A4
7	A3 -> A3	17	A4 -> A6
8	A3 -> A4	18	A6 -> A6
9	A4 -> A4	19	A6 -> A7
10	A4 -> A4	20	A7 -> A7
		21	A7 -> A6

***Step 5**: Construct the fuzzy logical relationship groups*

By Chen [3], all the fuzzy relationship having the same fuzzy set on the left-hand side or the same current state can be put together into one fuzzy relationship group. However, the repeated FLRs are counted only once [1],[2],[3].

Suppose there are relationships such that

$A_1 \rightarrow A_1$; $A_1 \rightarrow A_2$; $A_1 \rightarrow A_1$;..........

We can be grouped into a relationship group as follows: $A_1 \rightarrow A_1, A_2 \ldots$

In the weighted models [9], the recurrence of each fuzzy logical relation should be taken into account.

Suppose there are FLRs in chronological order as follows:

(t= 1) $A_1 \rightarrow A_1$
(t=2) $A_1 \rightarrow A_2$
(t=3) $A_1 \rightarrow A_1$

From this viewpoint, we can be grouped into a relationship group as follows: : $A_1 \rightarrow A_1, A_2, A_1, \ldots$ then assign different weights for each FLR.

Based on this recurrence fuzzy logical relation and from Table 3, we get the first – order fuzzy logical relation groups are shown in Table 4.

TABLE IV: THE FIRST-ORDER FUZZY LOGICAL RELATIONSHIP GROUPS

No	Relationships
1	A1 –> A1, A1, A2
2	A2 –> A3
3	A3 –> A3, A3, A3, A4, A3, A3, A3, A3, A4
4	A4 –> A4, A4, A3, A6
5	A6 –> A6, A7
6	A7 –> A7, A6

***Step 6**: Compute the forecasting results.*

Calculate the forecasted output at time t by using the following principles:

Rule 1: If the fuzzified enrolment of year t-1 is A_j and there is only one fuzzy logical relationship in the fuzzy logical relationship group whose current state is A_j, shown as follows: $A_j \rightarrow A_k$;then the forecasted enrolment of year t forecasted = m_k

where m_k is the midpoint of the interval u_k and the maximum membership value of the fuzzy set A_k occurs at the interval u_k

Rule 2: If the fuzzified enrolment of year t -1 is A_j and there are the following fuzzy logical relationship group whose current state is A_j , shown as follows:

$$A_j \rightarrow A_{i1}(x_1), A_{i2}(x_2), \ldots, A_{ip}(x_p)$$

then the forecasted enrolment of year t is calculated as

follows: $forecasted = \frac{x_1 m_{i1} + x_2 m_{i2} + \cdots + x_p m_{ip}}{x_1 + x_2 + \cdots + x_p}$; $p \leq n$

where $m_{i1}, m_{i2}, \ldots and\ m_{ip}$ are the middle values of the intervals u_1 , u_2 and u_p respectively, and the maximum membership values of A_1, A_2 , ... ,A_p occur at intervals u_1 , u_2,…, up , respectively; $x_1, x_2, \ldots$ and x_p notes the number of fuzzy logical relationships "$A_j \rightarrow A_{ik}$", (1≤ k ≤p) in the fuzzy logical relationship group.

Rule 3: If the fuzzified enrolment of year t is A_j and there is a fuzzy logical relationship in the fuzzy logical relationship group whose current state is A_j, shown as follows: $A_j \rightarrow$ #

where the symbol "#" denotes an unknown value, then the forecasted enrollment of year t + 1 is m_j, where m_j is the midpoint of the interval u_j and the maximum membership value of the fuzzy set A_j, occurs at u_j.

From these rules, we can obtain these forecasting results are listed in Table 5.

TABLE V: FORECASTED ENROLMENTS OF UNIVERSITY OF ALABAMA BASED ON THE FIRST – ORDER FTS MODEL.

Year	Actual	Fuzzified	Results
1971	13055	A1	Not forecasted
1972	13563	A1	14000
1973	13867	A1	14000
1974	14696	A2	14000
1975	15460	A3	15500
1976	15311	A3	15788.9
1977	15603	A3	15788.9
1978	15861	A3	15788.9
1979	16807	A4	15788.9
1980	16919	A4	17000
1981	16388	A4	17000
1982	15433	A3	17000
1983	15497	A3	15788.9
1984	15145	A3	15788.9
1985	15163	A3	15788.9
1986	15984	A3	15788.9
1987	16859	A4	15788.9
1988	18150	A6	17000
1989	18970	A6	19166.7
1990	19328	A7	19166.7
1991	19337	A7	18833.3
1992	18876	A6	18833.3
1993	N/A	#	19166.7

B. Adjusted the forecasting results by Fourier series

In order to improve the accuracy of forecasting models, the Fourier series has been successfully applied in modifying the residuals in fuzzy time series model which reduces the forecasting values.

The procedure to obtain the modified residuals from fuzzy time series forecasting model with Fourier series is as the following.

Suppose an original series (*the historical data series at time t*) with n entries is $\mathrm{Y_t} = \{\mathrm{x_1(t)}, \mathrm{x_2(t)}, \dots, \mathrm{x_n(t)}\}$ and its predicted series *(forecasting value at time t)*, under weighted fuzzy time series is $\widehat{\mathrm{Y_t}} = \{\widehat{\mathrm{x_1}}(\mathrm{t}), \widehat{\mathrm{x_2}}(\mathrm{t}), \dots, \widehat{\mathrm{x_n}}(\mathrm{t})\}$ then, its residual series is defined as: $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_t]$, with $t = \overline{1, n}$.

The analysis process for the residual series are derived as follows:

$$\boldsymbol{\varepsilon}_\mathbf{t} = \mathbf{Y_t} - \widehat{\mathbf{Y}}_\mathbf{t} \qquad (4)$$

Now, let's consider a sub-series $\boldsymbol{\varepsilon}^*$ as:

$$\boldsymbol{\varepsilon}^* = [\varepsilon_2, \varepsilon_3, \dots, \varepsilon_t \dots, \varepsilon_t], \text{ with } t = \overline{2, n} \qquad (5)$$

Based on Table 6, we get the residual series between actual value and forecasting value as follows:

$$\varepsilon_t = (-437, -113, \dots, 503, 42.7)_{1x21}$$

And

$$\boldsymbol{\varepsilon}^* = (-113, 696, \dots, 503, 42.7)_{1x20}$$

Following, Fourier series transform can be used to latch the implied periodic phenomenon in the residual series. Then, using Fourier modification technique in residual analysis, we can rise forecasting performance from the considered input data set. The estimated residual series can be modeled by Fourier series transform as follows:

$$\widehat{\boldsymbol{\varepsilon}}_\mathbf{t} = \frac{1}{2} a_0 + \sum_{i=1}^{d} [a_i \cos(\frac{2\pi i}{n-1} t) + b_i \sin(\frac{2\pi i}{n-1} t)], \text{ with } t = \overline{1, n} \qquad (6)$$

Where: $d = \frac{n-1}{2}$ is called the minimum deployment frequency of Fourier series [19] and only take integer number. And therefore, the residual sub-series is rewritten as:

$$\boldsymbol{\varepsilon}^* = \boldsymbol{P} * \boldsymbol{C} \qquad (7)$$

Where:

$$C = (a_0, a_1, b_1 \dots a_d, b_d)$$

$$\mathrm{P} = \left(\left[\tfrac{1}{2}\right]_{(n-1)x1} \mathrm{p_1} \dots \mathrm{P_k} \dots \mathrm{P_d} \right), \text{ with } P_k \text{ is}$$

determined by Eq.(8) as follows:

$$\boldsymbol{P_k} = \begin{bmatrix} \cos\left(\frac{2\pi * 2 * k}{n-1}\right) & \sin\left(\frac{2\pi * 2 * k}{n-1}\right) \\ \cos\left(\frac{2\pi * 3 * k}{n-1}\right) & \sin\left(\frac{2\pi * 3 * k}{n-1}\right) \\ \vdots & \vdots \\ \cos\left(\frac{2\pi * n * k}{n-1}\right) & \sin\left(\frac{2\pi * n * k}{n-1}\right) \end{bmatrix}$$

$where$, $k = \overline{1, d}$

The parameters $\mathrm{a_0}, \mathrm{a_1}, \mathrm{b_1}, \dots, \mathrm{a_d}, \mathrm{b_d}$ are obtained by using the ordinary least squares method (OLS) which results in the Eq.(8) as following:

$$\mathbf{C} = (\mathbf{P^T P})^{-1} \mathbf{P^T} \boldsymbol{\varepsilon}^{*\mathbf{T}} \qquad (8)$$

From value of ε^* and value of matrix P, we calculate the parameters as following:

$$\mathrm{C} = (5.13, 36.87, -113, \dots, -5.30, -130)_{1x21}$$

Next, Once the parameters are calculated, the forecasting series residual $\widehat{\varepsilon(t)}$ is then easily achieved based on the Eq.(6). Therefore, based the predicted series $\widehat{\mathrm{Y_t}}$ obtained from weighted fuzzy time series model, the forecasting series $\widehat{\boldsymbol{\varepsilon}'_\mathbf{t}}$ of the modified model is determined by:

$$\widehat{\boldsymbol{\varepsilon}'_\mathbf{t}} = [\widehat{\boldsymbol{\varepsilon}_1}, \widehat{\boldsymbol{\varepsilon}_2}, \widehat{\boldsymbol{\varepsilon}_3}, \dots, \widehat{\boldsymbol{\varepsilon}_n}] \qquad (9)$$

Where , $\begin{cases} \widehat{\boldsymbol{\varepsilon}'_1} = \widehat{\mathrm{x_1}} \\ \widehat{\boldsymbol{\varepsilon}'_\mathbf{t}} = \widehat{\mathbf{Y}}_\mathbf{t} + \boldsymbol{\varepsilon}(\boldsymbol{t}), with\ t = \overline{2, n} \end{cases}$

From Eq.(9) and based on Table 5, we get the finally forecasting results in 5-th columns of Table 6 after adjusted residual series.

TABLE VI: FORECASTED ENROLMENTS OF UNIVERSITY OF ALABAMA BASED ON WEIGTH FTS MODEL AND FOURIER SEIRIES TRANSFORM.

Year	Actual	Fuzzified	Weigth FTS model	Modified model
1971	13055	A1	-	-
1972	13563	A1	14000	-
1973	13867	A1	14000	13864
1974	14696	A2	14000	14693
1975	15460	A3	15500	15457

1976	15311	A3	15788.9	15308
1977	15603	A3	15788.9	15600
1978	15861	A3	15788.9	15858
1979	16807	A4	15788.9	16804
1980	16919	A4	17000	16916
1981	16388	A4	17000	16385
1982	15433	A3	17000	15430
1983	15497	A3	15788.9	15494
1984	15145	A3	15788.9	15142
1985	15163	A3	15788.9	15160
1986	15984	A3	15788.9	15981
1987	16859	A4	15788.9	16856
1988	18150	A6	17000	18147
1989	18970	A6	19166.7	18967
1990	19328	A7	19166.7	19325
1991	19337	A7	18833.3	19334
1992	18876	A6	18833.3	18873
1993	N/A	#	19166.7	18965

IV. Experimental Results

Experimental results for our model will be compared with the existing methods, such as the **SCI** model [2], the **C96** model [3], the S.R.Singh model [7] and the **H01** model [5] by using the enrolment of Alabama University from 1972s to 1992s are listed in Table 7 .

To evaluate the forecasted performance of proposed method in the FTS, the mean square error (MSE) and the mean absolute percentage error (MAPE) are used as a comparison criterion to represent the forecasted accuracy. The MSE value and MAPE value are computed according to (10) and (11) as follows:

$$\mathrm{MSE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{F_i} - \mathrm{R_i})^2$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{F_i - R_i}{R_i}\right| * 100\%$$

Where, R_i notes actual data on year i, F_i forecasted value on year i, n is number of the forecasted data

Table 7 shows a comparison of MSE and MAPE of our method using the first-order FTS under different number of intervals, where MSE and MAPE are calculated according to (10) and (11) as follows:

$$MSE = \frac{\sum_{i=1}^{20}(F_i - R_i)^2}{N} = \frac{(13864-13867)^2+(14693-14696)^2 \ldots +(18873-18876)^2}{20} = 8.57$$

$$MAPE = \frac{1}{20}\sum_{i=1}^{20}\left|\frac{F_i - R_i}{R_i}\right| * 100\% = \frac{1}{20}\left(\frac{abs(13864-13563)}{13563} + \cdots + \frac{abs(18873-18876)}{18876}\right) = 0.0175\%$$

where N denotes the number of forecasted data, F_i denotes the forecasted value at time i and R_i denotes the actual value at time i.

TABLE VII: A comparison of the forecasted results of proposed model with the existing models based on the second-order fuzzy time series by adjusting Fourier seiries.

Year	Actual data	SCI	C96	H01	S.R.Singh	Our model
1971	13055	-	-	-		-
1972	13563	14000	14000	14000		-
1973	13867	14000	14000	14000		13864
1974	14696	14000	14000	14000	14500	14693
1975	15460	15500	15500	15500	15358	15457
1976	15311	16000	16000	15500	15500	15308
1977	15603	16000	16000	16000	15500	15600
1978	15861	16000	16000	16000	15500	15858
1979	16807	16000	16000	16000	16500	16804
1980	16919	16813	16833	17500	16500	16916
1981	16388	16813	16833	16000	16500	16385
1982	15433	16789	16833	16000	15581	15430
1983	15497	16000	16000	16000	15500	15494
1984	15145	16000	16000	15500	15500	15142
1985	15163	16000	16000	16000	15500	15160
1986	15984	16000	16000	16000	15500	15981
1987	16859	16000	16000	16000	16402	16856
1988	18150	16813	16833	17500	18500	18147
1989	18970	19000	19000	19000	18500	18967
1990	19328	19000	19000	19000	19471	19325
1991	19337	19000	19000	19500	19500	19334
1992	18876	19000	19000	19149	19651	18873
MSE		**423027**	**407507**	**226611**	**115972**	**8.57**
MAPE		**3.22%**	**3.11%**	**2.66%**	**1.71%**	**0.0175%**

From Table 7, we can see that the proposed method by modifying Fourier series with number of interval is seven which has a smaller MSE value of 8.57 and MAPE value of 0.0175% than SCI model [2], the C96 model [3], the H01 model [5], S.R Singh model [7]. To be clearly visualized, Fig. 1 depicts the trends for actual data and forecasted results of the C96 model, H01 model, S.R.Singh model with forecasting results of proposed method. From Fig. 1, It is obvious that the forecasting accuracy of the proposed model is more close than any existing models for the different-orders fuzzy time series model.

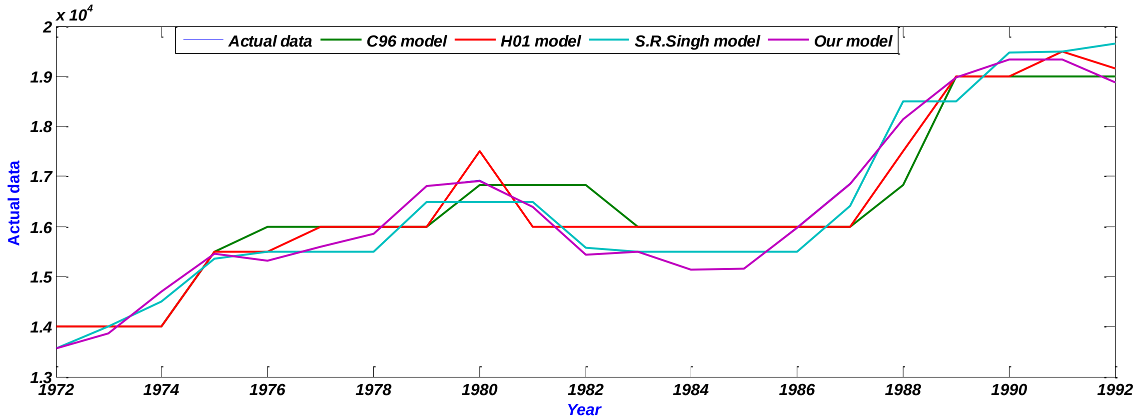

Fig. 1: The curves of the C96, H01, S.R.Singh models and our model for forecasting enrolments of University of Alabama

V. Conclusions

In this paper, we have presented a hybrid forecasted method to handle forecasting enrolments of the University of Alabama based on which combines the weighted fuzzy time series and Fourier series techniques. Firstly, we propose weighted models to tackle two issues in fuzzy time series forecasting, namely, recurrence and weighting. Based on forecasted result obtained. Then, we use Fourier series to modify the residuals of the weighted FTS for improving the forecasting performance. Next, we calculate forecasting output and compare forecasting accuracy with other existing models. Lastly, based on the performance comparison in Tables 7 and Fig. 1, it can show that our model outperforms previous forecasting models with various orders and the same interval length.

The proposed model was only tested by the forecasting enrollment problem, and it can actually be applied to other practical problems such as population forecast, and rice production forecasting in the further research.

References

[1] Q. Song, B.S. Chissom, “Forecasting Enrollments with Fuzzy Time Series – Part I,” Fuzzy set and system, vol. 54, pp. 1-9, 1993b.

[2] Q. Song, B.S. Chissom, “Forecasting Enrollments with Fuzzy Time Series – Part II,” Fuzzy set and system, vol. 62, pp. 1-8, 1994.

[3] S.M. Chen, “Forecasting Enrollments based on Fuzzy Time Series,” Fuzzy set and system, vol. 81, pp. 311-319. 1996.

[4] Hwang, J. R., Chen, S. M., & Lee, C. H. Handling forecasting problems using fuzzy time series. Fuzzy Sets and Systems, 100(1–3), 217–228, 1998.

[5] Huarng, K. Heuristic models of fuzzy time series for forecasting. Fuzzy Sets and Systems, 123, 369–386, 2001b .

[6] Singh, S. R. A simple method of forecasting based on fuzzy time series. Applied Mathematics and Computation, 186, 330–339, 2007a.

[7] Singh, S. R. A robust method of forecasting based on fuzzy time series. Applied Mathematics and Computation, 188, 472–484, 2007b.

[8] S. M. Chen, “Forecasting enrollments based on high-order fuzzy time series”, Cybernetics and Systems: An International Journal, vol. 33, pp. 1-16, 2002.

[9] H.K.. Yu “Weighted fuzzy time series models for TAIEX forecasting ”, Physica A, 349 , pp. 609–624, 2005.

[10] Chen, S.-M., Chung, N.-Y. Forecasting enrollments of students by using fuzzy time series and genetic algorithms. International Journal of Information and Management Sciences 17, 1–17, 2006a.

[11] Chen, S.M., Chung, N.Y. Forecasting enrollments using high-order fuzzy time series and genetic algorithms. International of Intelligent Systems 21, 485–501, 2006b.

[12] Huarng, K. theH. Effective lengths of intervals to improve forecasting fuzzy time series. Fuzzy Sets and Systems, 123(3), 387–394, 2001a.

[13] Lee, L.-W., Wang, L.-H., & Chen, S.-M. Temperature prediction and TAIFEX forecasting based on fuzzy logical relationships and genetic algorithms. Expert Systems with Applications, 33, 539–550, 2007.

[14] Jilani, T.A., Burney, S.M.A. A refined fuzzy time series model for stock market forecasting. Physica A 387, 2857–2862. 2008.

[15] Wang, N.-Y, & Chen, S.-M. Temperature prediction and TAIFEX forecasting based on automatic clustering techniques and two-factors high-order fuzzy time series. Expert Systems with Applications, 36, 2143–2154, 2009

[16] Kuo, I. H., Horng, S.-J., Kao, T.-W., Lin, T.-L., Lee, C.-L., & Pan. An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization. Expert Systems with applications, 36, 6108–6117, 2009a.

[17] S.-M. Chen, K. Tanuwijaya, “ Fuzzy forecasting based on high-order fuzzy logical relationships and automatic clustering techniques”, Expert Systems with Applications 38 ,15425–15437, 2011.

[18] Shyi-Ming Chen, Nai-Yi Wang Jeng-Shyang Pan. Forecasting enrollments using automatic clustering techniques and fuzzy logical relationships, Expert Systems with Applications 36 , 11070–11076, 2009.

[19] Y. L. Huang and Y. H. Lee, Accurately forecasting model for the Stochastic Volatility data in tourism demand, *Modern Econ.* 2, pp. 823-829, 2011.