International Journal of Innovative Computing, Information and Control
Volume **12**, Number **2**, April **2016**

 ISSN 1349-4198
pp. **459–466**

APPLICATIONS OF TAIEX AND ENROLLMENT FORECASTING USING AN EFFICIENT IMPROVED FUZZY TIME SERIES MODEL

Shye-Chorng Kuo[1], Chin-Ling Lee[2] and Cheng-Jian Lin[3]

[1]Department of Multimedia Animation and Application
Nan Kai University of Technology
No. 568, Zhongzheng Rd., Caotun Township, Nantou County 54243, Taiwan
t127@nkut.edu.tw

[2]Department of International Business
National Taichung University of Science and Technology
No. 129, Sec. 3, Sanmin Road, Taichung City 404, Taiwan
merrylee@nutc.edu.tw

[3]Department of Computer Science and Information Engineering
National Chin-Yi University of Technology
No. 57, Sec. 2, Zhongshan Rd., Taiping Dist., Taichung 41170, Taiwan
cjlin@ncut.edu.tw

Received October 2015; revised February 2016

Abstract. *In this study, an improved fuzzy time series model is proposed for solving forecasting problems. The proposed forecasting model is a hybrid of the traditional fuzzy time series and the ratio value. First, the linguistic variable analysis of fuzzy theory in the traditional fuzzy time series is used to observe the uncertain data which can be modeled as fuzzified variables. According to the ratio value in fuzzy logical relationship, the subsequent predicted data will be determined to increase or decrease. Finally, a fuzzy rule table is established for performing prediction. To verify the efficacy of the proposed forecasting model, the university enrollment and the Taiwan Capitalization Weighted Stock Index (TAIEX) forecasting problems are used as experiments. Experimental results show that the proposed model can achieve a better forecasting accuracy than other models.*
Keywords: Fuzzy time series, Forecast, Fuzzy logic relationship, Weighted stock index, University enrollment

1. **Introduction.** In recent years, the fuzzy time series has been widely used and is effectively applied to various fields [1-3]. The fuzzy time series model with fuzzified data and fuzzy logical relationship was proposed by Song and Chissom [4,5]. They adopted the max-min composition operator. The advantages of the fuzzy time series are the time-invariant characteristic and the unlimited exact numerical value. That is, the historical data with the fuzzy logical relationship are used to construct the fuzzy prediction model. The traditional time series model is difficult to perform. Although the forecasting model [4,5] got so much attention from researchers, it required a complex and time-consuming calculation process.

For this reason, Chen [6] proposed a simplified computation method to construct the fuzzy time series model which was more convenient and with high accuracy to overcome this disadvantage, but it still has some defects for precision.

Recently, some researchers [7-17] proposed many fuzzy time series models to increase the accuracy in prediction. Tsaur [7] combined the fuzzy time series with the Markov model to predict the trend of the exchange rate. As there existed the unknown factors or unpredictable problems in exchange rate, the Markov process can solve these problems well. Yu [8] used a hybrid of fuzzy time series and weighted model, which assigned appropriate weights to time series and fuzzy relationship. This hybrid method created the weighted matrix to calculate the predicted values for improving the forecasting capabilities. Huarng [9] found that the length of interval has a significant effect on prediction results and thus suggested averaging the lengths instead of randomly assigning the intervals to improve forecasting performance. In addition, since the forecasting results using the first-order fuzzy time series were not good enough, Chen [10] adopted the higher-order fuzzy time series model to increase the accuracy in prediction. However, this method needs more adjustable parameters. Yu and Huarng [11] took the different impact factors into account in the bivariate model and it was superior to the univariate model. Huarng and Yu [12] used the type-2 fuzzy set model to improve forecasting performance. Evolutionary algorithm [13] was introduced to improve the effectiveness in forecasting problems.

According to the above-mentioned the existing fuzzy time series methods have some drawbacks, such as determining the increasing or decreasing of next forecasting data and degrading the forecasting accuracy. In the study, an efficient forecasting model based on improved fuzzy time series is proposed. The proposed forecasting method is a hybrid of the traditional fuzzy time series and the ratio value to increase the reference information and reduce the prediction error. Each ratio value corresponds to one rule and then these rules constitute the groups. Finally, the enrollments of the University of Alabama and the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) are used to verify the effectiveness of the proposed model in terms of forecasting accuracy. Experimental results show that the proposed model has the smallest prediction error compared with other models.

The rest of this paper is organized as follows. Section 2 reviews the concept of the fuzzy time series. In Section 3, the proposed method is illustrated. Section 4 describes the forecasting results of enrollment and TAIEX using the various methods. The last, conclusions are discussed in Section 5.

2. Review of the Fuzzy Time Series.

In Song and Chissom [4], the fuzzy time series is defined as below. Let U be the universe of discourse, where $U = \{u_1, u_2, \cdots, u_n\}$, and fuzzy set A_i $(i = 1, 2, \cdots, n)$ in the universe of discourse U be described as follows:

$$A_i = \frac{f_A(u_1)}{u_1} + \frac{f_A(u_2)}{u_2} + \cdots + \frac{f_A(u_n)}{u_n} \tag{1}$$

where f_A represents the membership function of A, $f_A : U \to [0, 1]$, $f_{A_i}(u_i)$ represents the grade of membership of u_i in the fuzzy set A and $f_{A_i}(u_i) \in [0, 1]$, $1 \le i \le n$.

Let $Z(t)$ $(t = \cdots, 0, 1, 2 \cdots)$, a subset of R, be the universe of discourse on which fuzzy set $f_i(t)$ $(i = 1, 2, \cdots)$ are defined. Let $F(t)$ be a collection of $f_i(t)$ $(i = 1, 2, \cdots)$. Then $F(t)$ is called a fuzzy time series on $Z(t)$.

Assume there is existence of a relationship $R(t, t-1)$ and let $F(t) = F(t-1) \circ R(t, t-1)$, where $F(t)$ is caused by $F(t-1)$, the operator $\circ$ represents the Max-Min composition operation, and $R(t, t-1)$ is called fuzzy relation between $F(t)$ and $F(t-1)$.

Assuming $F(t-1) = A_i$ and $F(t) = A_j$. According to the previous description ($F(t)$ is caused by $F(t-1)$), we obtain $A_i \to A_j$, where $A_i \to A_j$ is called fuzzy logical relationship. However, fuzzy time series is not considered to accurately predict whether subsequent predicted data will increase or decrease according to ratio value in the fuzzy logical relationship.

3. The Improved Fuzzy Time Series.

In this section, the improved fuzzy time series is proposed and described as follows.

First, define a universe of discourse U, $U = [D_{\min} - D_1, D_{\max} + D_2]$, where $D_{\min}$ and $D_{\max}$ denote the historical data of the minimum and maximum values, respectively. D_1 and D_2 denote two appropriate values. According to the numbers of the interval, the length of each interval is calculated. In general, the number of the interval is pre-defined. If the number of the interval is larger, the forecasting accuracy increases and a time-consuming calculation process is required. On the contrary, the forecasting accuracy will decrease. Define the linguistic value of fuzzy set A_i and correspond to representatives of each interval.

$$\begin{aligned} A_1 &= \frac{1}{u_1} + \frac{0.5}{u_2} + \frac{0}{u_3} + \frac{0}{u_4} + \cdots + \frac{0}{u_n} \\ A_2 &= \frac{0.5}{u_1} + \frac{1}{u_2} + \frac{0.5}{u_3} + \frac{0}{u_4} + \cdots + \frac{0}{u_n} \\ &\vdots \\ A_{n-1} &= \frac{0}{u_1} + \frac{0}{u_2} + \cdots + \frac{0.5}{u_{n-2}} + \frac{1}{u_{n-1}} + \frac{0.5}{u_n} \\ A_n &= \frac{0}{u_1} + \frac{0}{u_2} + \cdots + \frac{0}{u_{n-2}} + \frac{0.5}{u_{n-1}} + \frac{1}{u_n} \end{aligned} \tag{2}$$

A historical datum belongs to interval u_j and the maximum membership value of fuzzy set A_j occurs at u_j, then the historical datum is fuzzified into fuzzy set A_j, where $1 \leq j \leq n$. The 1 or 0.5 represents the degree of membership function in the fuzzy set.

Fuzzify the each historical data to obtain the fuzzy set. The ratio value (R) of the fuzzy logical relationship is calculated by the data on time t and $t - 1$.

$$R = \left(\frac{x(t)}{x(t-1)}\right) \tag{3}$$

If the fuzzy logical relationships are duplicate, they will be deleted. On the contrary, the fuzzy logical relationships will be retained. Through the non-duplicate fuzzy logical relationship, the fuzzy logical relationships with the same left-hand side are classified into the same group. During the traditional forecasting process, the increasing or decreasing of next forecasting data is unknown. Therefore, we adopt the ratio value to determine the forecasting data. If all the ratio values in the corresponding fuzzy logical relationship group are greater than 1.0, this means that the next forecasting data is greater than the current data. On the contrary, if all the ratio values in the corresponding fuzzy logical relationship group are less than 1.0, this means that the next forecasting data is less than the current data. If some ratio values in the corresponding fuzzy logical relationship group are greater than 1 and the rest are less than 1.0, the average ratio value is calculated as the center value which divides the interval into two parts. One part is less than the average ratio value and the other is greater than the average ratio value.

Therefore, the individual ratio values of the two parts are recorded. Since the fuzzy logical relationship group consists of the same left-hand side A_i of the fuzzy logical relationship, the ratio values are accumulated and then the average ratio value (AR) is obtained and shown as follows:

$$AR = \frac{\sum_{k=1}^{r} R_k}{r} \tag{4}$$

where r is the number of the fuzzy logical relationship and R_k represents the ratio value of the kth fuzzy logical relationship.

The above-mentioned three kinds of information consist of fuzzy set, fuzzy logical relationship group, and the average ratio value (AR). The one condition is to determine whether all ratio values in the fuzzy logical relationship group are greater/less than 1. If the result of this condition is positive, the average ratio value is calculated as the center value which divides the interval into two parts. Finally, the forecasting is calculated and shown as follows:

$$F(t) = Y(t-1) * AR \tag{5}$$

where $F(t)$ represents the forecasting result at time t, $Y(t-1)$ denotes the historical data of time $t-1$, and AR is the average ratio value in the fuzzy logical relationship group.

Algorithm *Pseudo-code for the proposed IFTS*

1. $U \leftarrow [D_{\min} - D_1, D_{\max} + D_2]$
2. $U_n \leftarrow U/$number of interval (n)
3. Define the fuzzy set A_i, where $1 \leq i \leq n$
4. Fuzzify each historical data to obtain the fuzzy set A_i.
5. Calculate ratio value of the fuzzy logical relationship (FLR) by the data on time t and $t-1$.
   Let the FLR be $A_i \rightarrow A_j$, where A_i and A_j correspond to $x(t-1)$ and $x(t)$, respectively.
6. IF the FLR is duplicate
       It will be deleted.
   Else
       It will be retained.
7. The non-duplicate FLRs with the same left-hand side A_i are classified into the same group.
8. Determine the trend of the next forecasting data.
   IF all the ratio values in the corresponding FLR group are greater than 1.0
       The next forecasting data is greater than the current data.
   Else IF all the ratio values in the corresponding FLR group are less than 1.0
       The next forecasting data is less than the current data.
   Else
       The average ratio value is used as the center value which divides the interval into two parts. One part is less than the average ratio value and the other is greater than the average ratio value.
       For $k = 1$ to r
           sum = sum + R_k
       end
       $AR = \text{sum}/r$
9. The forecasting value is calculated using the $Y(t-1) * AR$.

4. **Applications of Enrollment and TAIEX Forecasting.**

In this section, we apply the proposed model to forecasting the enrollments of the University of Alabama and Taiwan Capitalization Weighted Stock Index (TAIEX).

4.1. **The enrollments of the University of Alabama.**

The data of historical enrollment of the University of Alabama is shown in Table 1. Table 2 shows the enrollment forecasting results using the improved fuzzy time series. We also compare the proposed forecasting model with Chen's method [6]. Figure 1 shows that the prediction results of the proposed model are better than Chen's method [6]. Recently, several researches [4-6]

TABLE 1. The historical enrollments and data fuzzification

Years	Historical Data
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

TABLE 2. The enrollment forecasting results using the improved fuzzy time series

Year	Actual enrollments	Forecasted value	Year	Actual enrollments	Forecasted value
1971	13055	-	1982	15433	15433.000000
1972	13563	13582.023523	1983	15497	15551.833674
1973	13867	14110.531217	1984	15145	15616.326472
1974	14696	14426.803538	1985	15163	15261.616082
1975	15460	15460.000000	1986	15984	15279.754681
1976	15311	15579.041573	1987	16859	16681.545350
1977	15603	15428.894277	1988	18150	18150.000000
1978	15861	16283.918424	1989	18970	18275.182635
1979	16807	16553.177602	1990	19328	19328.000000
1980	16919	16919.000000	1991	19337	19102.107281
1981	16388	17035.692287	1992	18876	19111.002095

adopt the mean square error to evaluate the forecasting model. The mean square error (MSE) is defined as follows:

$$MSE = \sum_{i=1}^{n} (A_i - F_i)^2 / n \tag{6}$$

where A_i and F_i represent the actual value and the forecasting value, and n denotes the number of data. In addition, in order to demonstrate the proposed method, we compare the performance of proposed model with other models [4-6,14,15]. Table 3 shows the comparison results using various models. In this table, we can find that the proposed model obtains the smallest MSE value compared with other models [4-6,14,15].

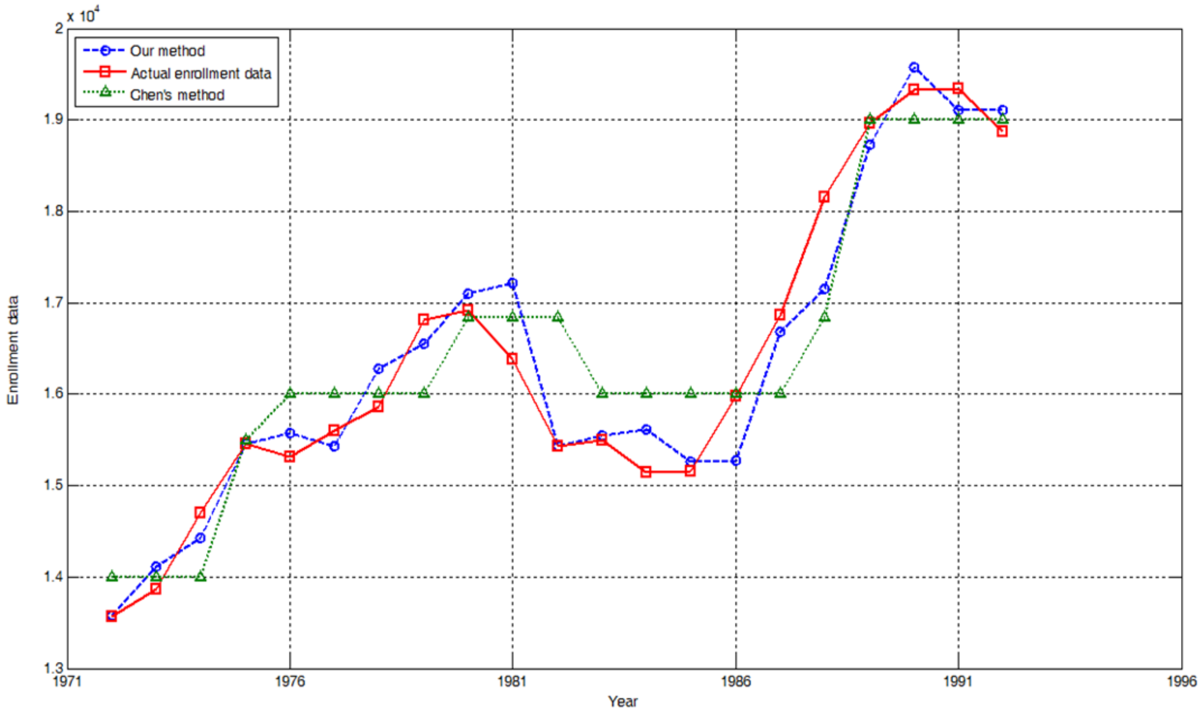

Figure 1. Enrollment forecasting results using the proposed model and Chen's method [6]

Table 3. Comparison results of enrollment forecasting for various forecasting models

Model	*Mean Square Error (MSE)*
Song and Chissom 1993 [4]	*775687*
Song and Chissom 1994 [5]	*412499*
Chen [6]	*407507*
Sullivan and Woodall [14]	*386055*
Huarng [15]	*226611*
The proposed model	*151365*

4.2. **Taiwan capitalization weighted stock index.** The period of the historical data is from 2014/01/02 to 2014/12/31 and the number of data is 248. We compare the forecasting performance of our model with Chen's [6]. In this experiment, both the two models set as seven intervals. The comparison results of the two models are shown in Figure 2. In this figure, the prediction results of the proposed method are better than Chen's method [6]. In order to demonstrate TAIEX forecasting of the proposed method, we also compare the performance of proposed model with other models [4-6,14,15]. Table 4 shows that the proposed model obtains a lower MSE value than other models [4-6,14,15].

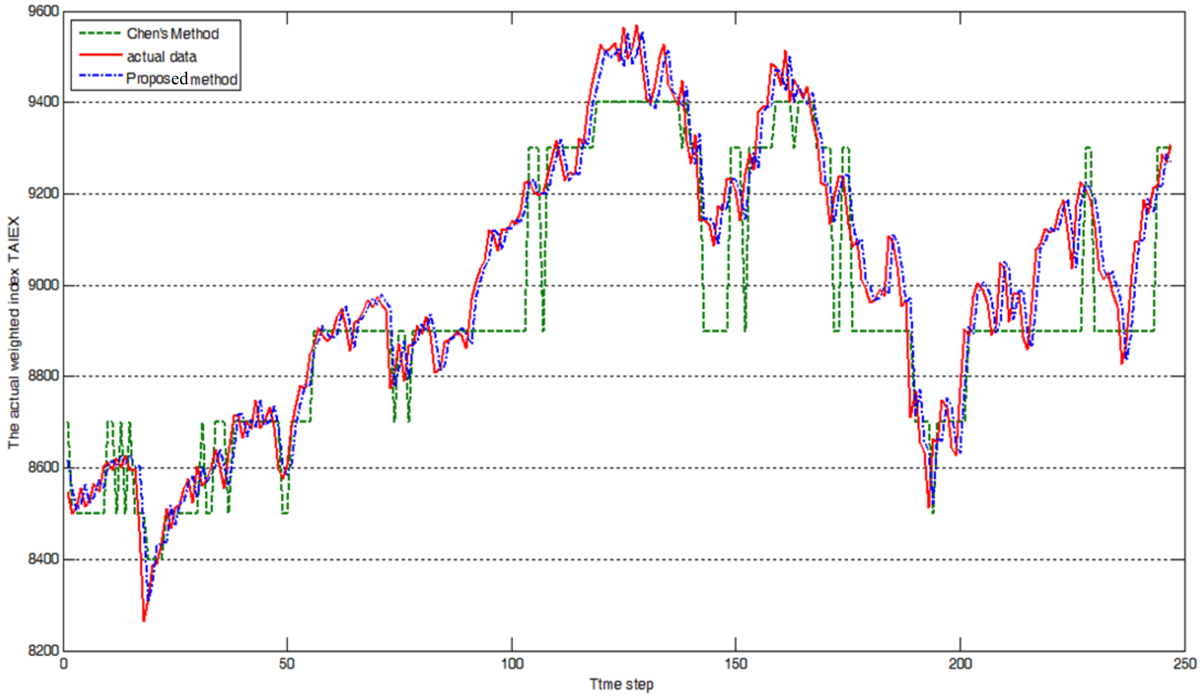

FIGURE 2. Forecasting results of TAIEX using the proposed model and Chen's method [6]

TABLE 4. Comparison results of TAXIEX forecasting for various forecasting models

Model	*Mean Square Error (MSE)*
Song and Chissom 1993 [4]	*8106*
Song and Chissom 1994 [5]	*7348*
Chen [6]	*7130*
Sullivan and Woodall [14]	*6854*
Huarng [15]	*6126*
The proposed model	*3630*

5. **Conclusions.** In this study, an improved fuzzy time series model is proposed for solving the forecasting problems. The proposed model combines the traditional fuzzy time series with the ratio value to construct the forecasting model. Each ratio value corresponds to one rule and then these rules constitute the groups. Finally, the rule table is established for forecasting the enrollment and TAIEX. Experimental results show that the proposed model has the smallest prediction error compared with other models.

REFERENCES

[1] C. H. Wang, Predicting tourism demand using fuzzy time-series and hybrid grey theory, *Tourism Management*, vol.25, no.3, pp.367-374, 2004.

[2] O. Duru, A fuzzy integrated logical forecasting model for dry bulk shipping index forecasting: An improved fuzzy time series approach, *Expert Systems with Applications*, vol.37, no.7, pp.5372-5380, 2010.

[3] V. Vamitha, M. Jeyanthi, S. Rajaram and T. Revathi, Temperature prediction using fuzzy time series and multivariate Markov chain, *International Journal of Fuzzy Mathematics and Systems*, vol.2, no.3, pp.217-230, 2012.

[4] Q. Song and B. S. Chissom, Forecasting enrollments with fuzzy time series – Part I, *Fuzzy Sets and Systems*, vol.54, no.1, pp.1-9, 1993.
[5] Q. Song and B. S. Chissom, Forecasting enrollments with fuzzy time series – Part II, *Fuzzy Sets and Systems*, vol.62, no.1, pp.1-8, 1994.
[6] S. M. Chen, Forecasting enrollments based on fuzzy time series, *Fuzzy Sets and Systems*, vol.81, no.3, pp.311-319, 1996.
[7] R.-C. Tsaur, A fuzzy time series-Markov chain model with an application to forecast the exchange rate between the Taiwan and US dollar, *International Journal of Innovative Computing, Information and Control*, vol.8, no.7(B), pp.4931-4942, 2012.
[8] H. K. Yu, Weighted fuzzy time series models for TAIEX forecasting, *Physica A*, vol.349, nos.3-4, pp.609-624, 2005.
[9] K. Huarng, Effective lengths of intervals to improve forecasting in fuzzy time series, *Fuzzy Sets Syst.*, vol.123, no.3, pp.387-394, 2001.
[10] S. M. Chen, Forecasting enrollments based on high-order fuzzy time series, *Cybern. Syst. Int. J.*, vol.33, no.1, pp.1-16, 2002.
[11] T. H. Yu and K. Huarng, A bivariate fuzzy time series model to forecast the TAIEX, *Expert Systems with Applications*, vol.34, no.4, pp.2945-2952, 2008.
[12] K. Huarng and H. K. Yu, A type 2 fuzzy time series model for stock index forecasting, *Physica A*, vol.353, no.1, pp.445-462, 2005.
[13] Q. Cai, D. F. Zhang, B. Wu and C. H. Leung, A novel stock forecasting model based on fuzzy time series and genetic algorithm, *Procedia Computer Science*, vol.18, pp.1155-1162, 2013.
[14] J. Sullivan and W. H. Woodall, A comparison of fuzzy forecasting and markov modeling, *Fuzzy Sets and Systems*, vol.64, no.3, pp.279-293, 1994.
[15] K. Huarng, Heuristic models of fuzzy time series for forecasting, *Fuzzy Sets and Systems*, vol.123, no.3, pp.369-386, 2011.
[16] Y. C. Cheng and S. T. Li, Fuzzy time series forecasting with a probabilistic smoothing hidden Markov model, *IEEE Trans. Fuzzy Syst.*, vol.20, no.2, pp.291-304, 2012.
[17] S. H. Cheng, S. M. Chen and W. S. Jian, A novel fuzzy time series forecasting method based on fuzzy logical relationships and similarity measures, *Proc. of IEEE International Conference on Systems, Man, and Cybernetics*, 2015.