J. Korean Soc. Math. Educ. Ser. B: Pure Appl. Math. ISSN 1226-0657
Volume 19, Number 1 (February 2012), Pages 7–21

FINANCIAL TIME SERIES FORECASTING USING FUZZY REARRANGED INTERVALS

Hye-young Jung [a], Jin Hee Yoon [b] and Seung Hoe Choi [c,*]

Abstract. The fuzzy time series is introduced by Song and Chissom([8]) to construct a pattern for time series with vague or linguistic value. Many methods using the interval and fuzzy logical relationship related with historical data have been suggested to enhance the forecasting accuracy. But they do not fully reflect the fluctuation of historical data. Therefore, we propose the interval rearranged method to reflect the fluctuation of historical data and to improve the forecasting accuracy of fuzzy time series. Using the well-known enrollment, the proposed method is discussed and the forecasting accuracy is evaluated. Empirical studies show that the proposed method in forecasting accuracy is superior to existing methods and it fully reflects the fluctuation of historical data.

1. Introduction

Time series analysis has been a very popular method with many successful applications. In time series analysis, the involved experimental data are assumed to be precise. However, in many practical situations we encounter data which are not only random but vague due to ambiguous information or linguistic structure as well, for example 'about 10', 'greater than 10', 'more or less between 5 and 10', or 'fair', 'good', 'excellent', etc. Moreover, sometimes in the context of economic systems, such as stock market price, foreign exchange rate or market sales price of particular commodities, the historical data record officially, closing values, rough average values, maximum values or minimum values, instead of the precise series with variation during the time period. And usually we use only closing values when we analyze financial daily data. In these situations, sometimes closing values cannot be the representative value of one day, especially when closing value is the highest or lowest value of the day. Also, if we use only closing value for data analysis, we might lose

Received by the editors November 11, 2011. Revised January 13, 2012. Accepted Jan. 26, 2012.
2000 *Mathematics Subject Classification.* 37M10, 62M10, 62P20, 91B84.
Key words and phrases. fuzzy time series, forecasting, rearranged interval method, fluctuation.
*Corresponding author.

lots of data information of the day. So, it is better for us to use fuzzy data which include variation of the day than using general time series models such as ARIMA, ARCH and GARCH in these cases. Furthermore, if the number of observed data is fewer than 50, it is not valid to apply the general time series models. An appropriate way of solving such problems is by the use of fuzzy time series model using the concept of fuzzy set.

The fuzzy time series was introduced by Song and Chissom([8]) to forecast the enrollments of the University of Alabama. Since then, the new methods of the fuzzy time series to reduce the forecasting error have been studied by many authors. Since the fuzzy time series is a forecasting method, reducing the forecasting error is very important. Chen([1]) proposed a method for forecasting the enrollments of the University of Alabama based on the simplified arithmetic operations. Huarng([3]) proposed a heuristic model for forecasting the enrollments of the University of Alabama and the Taiwan Futures Exchange(TAIFEX) based on heuristic knowledge. In [7], Liu suggested a improved fuzzy time series model with a trapezoidal fuzzy number based on the Chen's method. Lee et al.([6]) proposed a model for forecasting the temperature and the TAIFEX using the high-order fuzzy logical relationships and genetic simulated annealing techniques. Cheng et al.([2]) applied a fuzzy time series model based on the adaptive expectation model for forecasting the TAIEX and the enrollments of the University of Alabama.

However, the models mentioned above have some drawbacks: first, the heuristic model requires an assumption that there are heuristic knowledge showing the increase or decrease in historical data for the next period or the other variables to forecast the historical data. Second, high-order model is not easy to calculate. Third, the models mentioned above do not fully reflect the fluctuation of the historical data.

In this paper, we propose a new method to resolve these drawbacks. First, our method provides an easy calculation and does not need heuristic knowledge or other variables. Second, to better reflect the fluctuation of historical data, we consider the frequency of historical data in each interval and variation of fuzzy logical relationship. If the interval has the high frequency of historical data then the fuzzy time series model does not reflect the fluctuation of historical data well. Also, if forecasts have the same values even though the historical data have a great difference, then the forecasting accuracy is reduced. Thus, our method based on the frequency of historical data in the interval and variation in the fuzzy logical relationship group would provide better the forecasting accuracy and reflect actual data's fluctuation better than the existing methods.

The rest of this paper is organized as follows. In Section 2, we review some definitions for fuzzy time series. In Section 3, our new method is introduced. In Section 4, we present our new method to forecast the enrollments of the University of Alabama and TAIFEX. Also, we introduce the new target data KRW/USD average exchange rate per month(KUEXR). In Section 5, compares the forecasting accuracy with the previous methods' results. The conclusions are discussed in Section 6.

2. Fuzzy Time Series

In this section, we give some definitions for fuzzy time series proposed by Song and Chissom.

Let $U = \{u_1, u_2, \cdots, u_n\}$ be the universe of discourse containing all values of a time series. Define a fuzzy set A_i of U:

$$A_i = \mu_{A_i}(u_1)/u_1 + \mu_{A_i}(u_2)/u_2 + \cdots + \mu_{A_i}(u_n)/u_n,$$

where μ_{A_i} is the membership function of fuzzy set A_i and $1 \leq i \leq n$.

Definition 1. If $F(t)$ consists of fuzzy set A_i, then $\{F(t) : t = \cdots, 0, 1, 2, \cdots\}$ is called a *fuzzy time series.*

Definition 2. Let $F(t)$ be a fuzzy time series. If there exists a fuzzy relationship $R(t, t-1)$ such that $F(t) = F(t-1) \circ R(t, t-1)$ where $\circ$ represents an fuzzy operator, then $F(t)$ is said to be *caused* by $F(t-1)$.

Definition 3. Assume that $F(t)$ is a fuzzy time series and $F(t)$ is caused by $F(t-1), F(t-2), \cdots$, and $F(t-n)$. Then $F(t)$ is called the *n-th order fuzzy time series.*

Definition 4. When $F(t-1) = A_i$ and $F(t) = A_j$, the fuzzy logical relationship between $F(t-1)$ and $F(t)$ is denoted by $A_i \to A_j$. The fuzzy logical relationship "$A_i \to A_j$" denotes that if the fuzzified historical data of time $t-1$ is A_i, then the fuzzified historical data of time t is A_j. Note that the repeated fuzzy logical relationships are removed ([1]).

Definition 5. Fuzzy logical relationships can be grouped together into fuzzy logical relationship groups. Suppose there are fuzzy logical relationships such that : $A_i \to A_{j1}$, $A_i \to A_{j2}, \cdots$, and $A_i \to A_{jk}$. Then the fuzzy logical relationship group is denoted by $A_i \to A_{j1}, A_{j2}, \cdots, A_{jk}$.

3. Fuzzy Time Series Using Rearranged Intervals

In this section, we propose a new method to improve forecasting accuracy, based on partitioning the intervals and fuzzy relationships. The algorithm for the proposed method is as follows.

Step 1: Define the universe of discourse U.

Let $x_{(1)}$ and $x_{(n)}$ be the maximum and the minimum value of historical data $\{x_t : t = 1, 2, \cdots, n\}$. Then the universe of discourse U is defined by $U = [x_{(1)} - c, x_{(n)} + d]$, where c and d are two proper positive numbers.

Step 2: Partition the universe of discourse into several intervals.

We consider a method splitting up the universe set using the intervals. First, we use the intervals $v_j = [e_{j-1}, e_j] (j = 1, \cdots, m_1)$ with equal length(l) to divide the set U into some subsets. Here,

$$e_j = \begin{cases} x_{(1)} - c & \text{if } \ j = 0, \\ x_{(1)} - c + j \cdot l & \text{if } \ 1 \le j \le m_1 - 1, \\ x_{(n)} + d & \text{if } \ j = m_1. \end{cases} \tag{3.1}$$

Next, if the historical data x_t is an left endpoint of the interval v_{j+1}, we divide the union of two interval v_j and v_{j+1} into several intervals $v^{(i)}_{j,j+1} (i = 1, \cdots, m_2)$ satisfying the following conditions:

(i) Both intervals $v^{(i)}_{j,j+1} = [e^{(i-1)}_{j,j+1}, e^{(i)}_{j,j+1}]$ has same length (l_1) and the historical data x_t is not an endpoint of the interval $v^{(i)}_{j,j+1}$ for each i.

(ii) The endpoint $e^{(i)}_{j,j+1}$ equals to

$$\begin{cases} e_{j-1} & \text{if } \ i = 0, \\ e_{j-1} + i \cdot l_1 & \text{if } \ 1 \le i \le m_2 - 1, \\ e_{j+1} & \text{if } \ i = m_2. \end{cases} \tag{3.2}$$

We represent the set of intervals given in (1) and (2) to $w_1, w_2, \cdots, w_m$ in a sequential order.

Finally, we consider the frequency of historical data belong to the interval $w_j (j = 1, \cdots, m)$. If a certain interval w_j contains more historical data than that other intervals hold, we do not use information can be inferred from the historical data to construct the fuzzy time series. This causes a decrease in the accuracy of the fuzzy time series constructed by the fuzzy logic. We can improve an accuracy of fuzzy time series by adjusting the frequency of the historical data belong to the interval. Let f_i be the frequency of historical data belongs to the interval w_i and f_{max} be a maximum frequency subjectively determined by the researcher. If the frequency of the interval w_k is greater than f_{max}, then we repeat to divide w_k into several sub-intervals $w_k^{(l)}(l = 1, \cdots, m_3)$ until $w_k^{(l)}$ have the frequency below f_{max}.

We present the intervals obtained by this process as $u_i(i = 1, \cdots, p)$. Then the intervals u_i satisfies the following three conditions at the same time:

(1) The frequency of historical data belong to the interval u_i is below f_{max}.

(2) The endpoints of the interval u_i are not equal to the historical data.

(3) A center of the interval u_j is greater than that of the interval u_i if $i \leq j$.

Step 3: Define fuzzy set and fuzzify the crisp values based on the interval proposed in Step 2.

From the intervals $u_i(i = 1, \cdots, p)$, we define the fuzzy sets $A_1, A_2, \cdots, A_p$ are defined by

$$\begin{aligned}
A_1 &= 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + \cdots + 0/u_p, \\
A_2 &= 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + 0/u_5 + \cdots + 0/u_p, \\
&\vdots \qquad\qquad\qquad\qquad \vdots \\
A_i &= 0/u_1 + \cdots + 0.5/u_{i-1} + 1/u_i + 0.5/u_{i+1} + \cdots + 0/u_p, \\
&\vdots \qquad\qquad\qquad\qquad \vdots \\
A_p &= 0/u_1 + 0/u_2 + \cdots + 0/u_{p-2} + 0.5/u_{p-1} + 1/u_p.
\end{aligned}$$

The number that precedes slash symbol "/" is the membership degree of the fuzzy set. If the historical data x_t belongs to the interval u_i and there is a value represented by a fuzzy set A_i in which the maximum membership degree occurs at u_i, then x_t is fuzzified as the fuzzy set A_i. If we denote the fuzzy data at time t as A_i, denoted by $F(t) = A_i$, then time series consisting of crisp value $\{x_t : t = 1, 2, \cdots, n\}$ is translated to fuzzy time series $\{F(t) : t = 1, 2, \cdots, n\}$. Let c_i be a center of the interval u_i. The inequality $c_i < c_j$ for $i < j$ implies that $u_i < u_j$. Then we define the well-ordered fuzzy sets $A_i(i = 1, \cdots, p)$ from the well-ordered intervals $u_i(i = 1, \cdots, p)$. That is, $A_i < A_j$ if $u_i < u_j$.

Step 4: Establish the fuzzy logical relationships and the fuzzy logical relationship groups based on the well-ordered fuzzy set $\{A_i : i = 1, \cdots, p\}$.

Step 5: Rearrange the fuzzy set using the fuzzy logical relationship.

We denote by the number r_i as a maximum value of the set $\{|i - i_j| : F(t) = A_i, F(t + 1) = A_{i_j}, j = 1, \cdots, k\}$ generated by the fuzzy logical relationship about

A_i. In the next step, we are going to use an average of the modes of the fuzzy sets in order to obtain the forecasted value of the historical data x_t at the time t. Since the sample mean is insensitive to outliers, the forecasting accuracy may be affected by the outliers when the value r_i is large. Thus, it is reasonable that the researcher adjusts to the value r_i to obtain the high forecasting accuracy. For this, we let r_{max} be the permissible value designed by the researcher. If the fuzzy set given in Step 4 satisfies the relation $r_i \geq r_{max}$, we return back to Step 3 and then repeat the process of dividing the intervals until $r_i < r_{max}$ for each i. As like to Step 4, we present the fuzzy sets obtained by this process as $A_i(i = 1, \cdots, p_1)$ without loss of generality.

Step 6: Predict the historical data using the fuzzy sets.

From the fuzzy sets and the fuzzy logical relationships given in the previous step, we can obtain the forecasted value($\hat{x}_t$) of the historical data x_t at the time t. If $F(t-1) = A_i$ and the fuzzy logical relationship group is $A_i \to A_{i_1}, A_{i_2}, \cdots, A_{i_q}$, $q \geq 1$, then the forecasted value $\hat{x}_t$ is

$$\hat{x}_t = \frac{1}{q} \sum_{j=1}^{q} c_{i_j},$$

where c_{i_j} is the center of the interval u_{i_j}.

4. Empirical Studies

In this section, we present our method to forecast the enrollment of the University of Alabama and TAIFEX introduced in many literatures to show the efficiency of the fuzzy time series. Also, we introduce the new target data KUEXR.

4.1. Forecasting for enrollment The enrollment of the University of Alabama introduced by Song and Chissom([9]) was used in the previous studies of fuzzy time series. This is also used here to show the superiority of the proposed model.

Step 1: According to the actual enrollment of the University of Alabama shown in Table I, we can see that $x_{(1)} = 13055$ and $x_{(2)} = 19337$ are actual minimum and maximum enrollment, respectively. Now, we let two proper positive numbers $c = 55$ and $d = 163$. Then we define the universe of discourse $U = [13000, 20000]$.

Step 2: The length of the intervals determined by the average-based length of intervals([3]) is 300. Since there are no enrollment in the interval $[19600, 20000]$, this interval is removed. Hence the U can be divided into equal-length intervals as follows: $w_i = [13000 + 300 \cdot (i-1), 13000 + 300 \cdot i]$, where $i = 1, 2, \cdots, 22$. We let

Table I

Enrollment and fuzzy enrollment

Year	Enrollment	Fuzzy enrollment based on step 2	Fuzzy enrollment based on step 5
1971	13055	A_1	A_1
1972	13563	A_2	A_2
1973	13847	A_3	A_3
1974	14696	A_6	A_6
1975	15460	A_{10}	A_{14}
1976	15311	A_9	A_{12}
1977	15603	A_{12}	A_{16}
1978	15861	A_{13}	A_{17}
1979	16807	A_{16}	A_{21}
1980	16919	A_{17}	A_{23}
1981	16388	A_{15}	A_{19}
1982	15433	A_{10}	A_{13}
1983	15497	A_{11}	A_{15}
1984	15145	A_8	A_9
1985	15163	A_8	A_{10}
1986	15984	A_{13}	A_{17}
1987	16859	A_{16}	A_{22}
1988	18150	A_{21}	A_{27}
1989	18970	A_{23}	A_{29}
1990	19328	A_{25}	A_{32}
1991	19337	A_{25}	A_{33}
1992	18876	A_{23}	A_{29}

$f_{max} = 3$. Since the interval w_8 and w_9 have the frequency over f_{max}, we repeat to divide into two sub-intervals of the equal length until the sub-intervals have the frequency below f_{max}.

(a) $w_8 = [15100, 15400]$: The enrollments in 1976, 1984, and 1985 belong to w_8. Thus, we divide w_8 into $w_8^{(1)} = [15100, 15250]$ and $w_8^{(2)} = [15250, 15400]$.

(b) $w_9 = [15400, 15700]$: The enrollments in 1795, 1977, 1982, and 1983 belong to w_9. First, we divide w_9 into $w_9^{(1)}$=[15400, 15550] and $w_9^{(2)} = [15550, 15700]$. But the enrollments in 1795, 1982 and 1983 belong to $w_9^{(1)} = [15400, 15550]$. Thus, we divide $w_9^{(1)} = [15400, 15550]$ into $w_9^{(1,1)} = [15400, 15475]$ and $w_9^{(1,2)} = [15475, 15550]$. Finally, we get the following intervals:

$$u_i = \begin{cases} [13000 + 300 \cdot (i-1), 13000 + 300 \cdot i] & 1 \le i \le 7, \\ [15100 + 150 \cdot (i-8), 15100 + 150 \cdot (i-7)] & 8 \le i \le 9, \\ [15400 + 75 \cdot (i-10), 15400 + 75 \cdot (i-9)] & 10 \le i \le 11, \\ [15550, 15700] & i = 12, \\ [15700 + 300 \cdot (i-13), 15700 + 300 \cdot (i-12)] & 13 \le i \le 25. \end{cases} \tag{4.1}$$

Step 3: Define fuzzy set based on interval in (3) and fuzzify the enrollment. Table I lists the enrollment at the University of Alabama and the corresponding fuzzy enrollment. Fuzzy sets given in Table I provide the fuzzy logical relationships in the left side of Table II.

Step 4: Establish the fuzzy logical relationships and groups. Since the fuzzy logical relationships are $A_1 \to A_2, A_2 \to A_3, A_3 \to A_6, A_6 \to A_{10}, A_{10} \to A_{11}, A_{11} \to A_{12}, \cdots, A_{25} \to A_{25}, A_{25} \to A_{23}$, the fuzzy logical relationship groups look as above Table II.

Step 5: In Table II, we find the fuzzy logical relationship groups having the high variation of fuzzy logical relationships and divide corresponding intervals into sub-intervals. In this paper we let $r_{max} = 2$. Since r_i is greater than r_{max}, we have to re-divide the intervals (3) given in Step 2.

(a) $A_8 \to A_8, A_{13}$: Since r_8 is greater than r_{max}, we have to divide u_8 into $u_8^{(1)} = [15100, 15175]$ and $u_8^{(2)} = [15175, 15250]$. But the all enrollments in the interval u_8 fall in the interval $u_8^{(1)} = [15100, 15175]$. Thus, we repeat to divide into two sub-intervals of equal length until the enrollments in the interval u_8 do not fall in the same sub-interval. We get the resulting intervals as follows: $[15100, 15137.5]$, $[15137.5, 15156.25]$, $[15156.25, 15175]$, $[15175, 15250]$.

(b) $A_{10} \to A_9, A_{12}$: Since r_{10} is greater than r_{max}, the interval $u^{10} = [15400, 15475]$

Table II

Fuzzy logical relationship groups (FLRG)

FLRG based on step 3	FLRG based on step 5
$A_1 \to A_2$	$A_1 \to A_2$
$A_2 \to A_3$	$A_2 \to A_3$
$A_3 \to A_6$	$A_3 \to A_6$
$A_6 \to A_{10}$	$A_6 \to A_{14}$
$A_8 \to A_8, A_{13}$	$A_9 \to A_{10}$
$A_9 \to A_{12}$	$A_{10} \to A_{17}$
$A_{10} \to A_9, A_{11}$	$A_{12} \to A_{16}$
$A_{11} \to A_8$	$A_{13} \to A_{15}$
$A_{12} \to A_{13}$	$A_{14} \to A_{12}$
$A_{13} \to A_{16}$	$A_{15} \to A_9$
$A_{15} \to A_{10}$	$A_{16} \to A_{17}$
$A_{16} \to A_{17}, A_{21}$	$A_{17} \to A_{22}$
$A_{17} \to A_{15}$	$A_{19} \to A_{13}$
$A_{21} \to A_{23}$	$A_{21} \to A_{23}$
$A_{23} \to A_{25}$	$A_{22} \to A_{27}$
$A_{25} \to A_{23}, A_{25}$	$A_{23} \to A_{19}$, $A_{27} \to A_{29}$
	$A_{29} \to A_{32}$, $A_{32} \to A_{33}$
	$A_{33} \to A_{29}$

divides into the two sub-intervals $[15400, 15437.5]$ and $[15437.5, 15475]$.

(c) For $A_{16} \to A_{17}, A_{21}$ and $A_{25} \to A_{23}, A_{25}$, we apply the same method above.

Finally, we get the resulting intervals as follows: $u_i = [13000+300\cdot(i-1), 13000+300\cdot i](i = 1, 2, \cdots, 7)$, u_9=[15137.5, 15156.25], $u_{10} = [15156.25, 15175]$, u_{11}=[15175, 15250], u_{12}=[15250, 15400], u_{13}=[15400, 15437.5], u_{14}=[15437.5, 15475], u_{15}=[15475, 15550], u_{16}=[15550, 15700] u_{17}=[15700, 16000], u_{18}=[16000, 16300], u_{19}=[16300,16600], u_{20}=[16600,16750], u_{21}=[16750, 16825], u_{22}=[16825,16900], $u_{i+9} = [13000 + 300 \cdot (i-1), 13000 + 300 \cdot i]$ $(i = 14, 15, \cdots, 21)$, u_{31}=[19300, 19318.75], u_{32}=[19318.75, 19328.13], u_{33}=[19328.13, 19337.5], u_{34}=[19337.5, 19375], u_{35}=[19375, 19450], u_{36}=[19450, 19600].

Also, we establish the final fuzzy logical relationships and groups based on the intervals given above in Table II.

Step 6: Defuzzify the forecasting results using the "Centroid". The forecasted enrollments by proposed method and Chen's method([1]) are shown in Table III, respectively.

Table III

Forecasted enrollment

Year	Actual data	Chen's method	The proposed method
1971	13055		
1972	13563	14000	13450
1973	13867	14000	13750
1974	14696	14000	14650
1975	15460	15500	15456.2
1976	15311	16000	15325
1977	15603	16000	15625
1978	15861	16000	15887.5
1979	16807	16000	16787.5
1980	16919	16833	17050
1981	16388	16833	16450
1982	15433	16833	15418.8
1983	15497	16000	15512.5
1984	15145	16000	15146.8
1985	15163	16000	15165.6
1986	15984	16000	15962.5
1987	16859	16000	16862.5
1988	18150	16833	18250
1989	18970	19000	18850
1990	19328	19000	19323.4
1991	19337	19000	19332.8
1992	18876	19000	18850

Fig.I shows the actual and forecasted data. From Fig.I, we see that our method reflects the fluctuation of historical data.

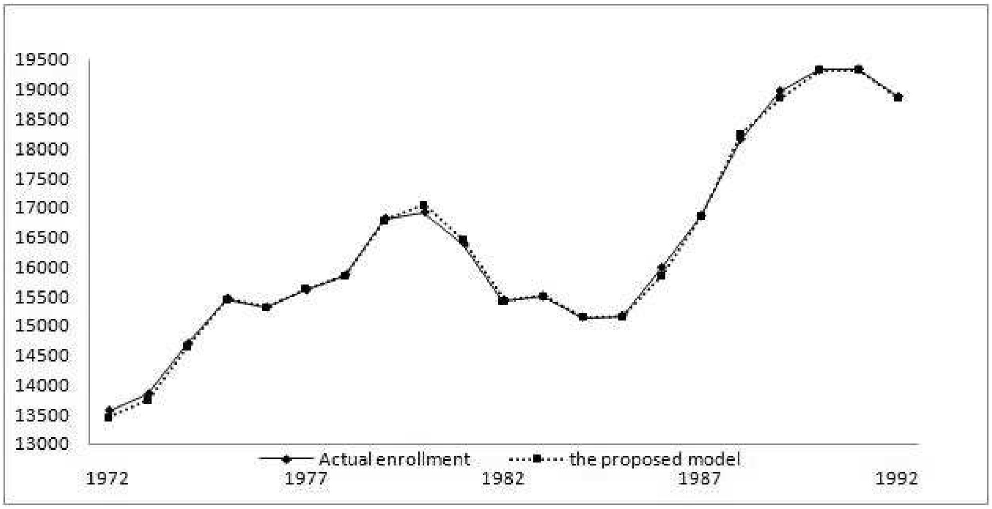

Fig.I. The actual and forecasted enrollment

4.2. Forecasting for TAIFEX TAIFEX data set introduced to Huarng([4]) has been used as the forecasting target for many authors studying the fuzzy time series. The proposed method gets higher the forecasting accuracy and easier calculation than the other methods using the TAIFEX.

Step 1: According to the actual TAIFEX shown in Table IV, we can see that $x_{(1)} = 6200$ and $x_{(2)} = 7560$ are actual minimum and maximum data, respectively. Now, we let two proper positive numbers $c = 50$ and $d = 40$. Then we define the universe of discourse $U = [6150, 7600]$.

Step 2: The length of the intervals using average-based length is 40 (Huarng ([3])). Since there are no actual data in the interval $[6150, 6190]$, this interval is removed. Hence the U can be divided into equal-length intervals as follows: $v_i = [6190 + 40 \cdot (i-1), 6190 + 40 \cdot i]$, for $i = 1, 2, 3, \cdots, 34$, and $v_{35} = [7550, 7600]$. We let $f_{max} = 3$. Then we obtain 53 intervals having the frequency of historical data below f_{max}.5

Step 3-5: We let $r_{max} = 3$. Then we obtain 73 intervals having r_i below r_{max}. Since the methods are similar to those in the *A. Forecasting for enrollment*, for the sake of simplicity, we omit the processes of establishing fuzzy sets and fuzzy logical relationship groups.

Step 6: Defuzzify the forecasting results using the"Centroid". The forecasted TAIFEX by proposed method and existing methods are shown in Table IV.

Table IV

Forecasted TAIFEX

date	actual data	Chen's method	Huarng's method	Lee et al.'s method([6])	The Proposed method
8/04	7560	7450	7450		7559.375
8/05	7487	7450	7450		7490
8/06	7462	7500	7500		7450
8/07	7515	7500	7500		7530
8/10	7365	7450	7450		7370
8/11	7360	7300	7300		7350
8/12	7330	7300	7300	7329	7350
8/13	7291	7300	7300	7289.5	7296.667
8/14	7320	7183.33	7183.33	7329	7330
8/15	7300	7300	7300	7289.5	7296.667
8/17	7219	7300	7300	7215	7218.751
8/18	7220	7183.33	7100	7215	7219.861
8/19	7285	7183.33	7300	7289.5	7285
8/20	7274	7183.33	7183.33	7289.5	7275
8/21	7225	7183.33	7100	7215	7225
8/24	6955	7183.33	7100	6949.5	6970
8/25	6949	6850	6850	6949.5	6921.667
8/26	6790	6850	6850	6796	6790
8/27	6835	6775	6775	6848	6833.333
8/28	6695	6850	6750	6698.5	6695
8/29	6728	6750	6750	6726	6725.625
8/31	6566	6775	6650	6569.5	6570
9/01	6409	6450	6450	6417	6410
9/02	6430	6450	6550	6417	6430
9/03	6200	6450	6350	6205	6210
9/04	6403.2	6450	6450	6417	6396.667
9/05	6697.5	6450	6550	6698.5	6695
9/07	6722.3	6450	6750	6726	6725.625
9/08	6859.4	6750	6850	6848	6857.5
9/09	6769.6	6775	6750	6763	6766.667
9/10	6709.75	6850	6650	6726	6705
9/11	6726.5	6775	6775	6726	6726.25
9/14	6774.55	6775	6775	6763	6773.333
9/15	6762	6775	6775	6763	6756.667
9/16	6952.75	6775	6850	6949.5	6970
9/17	6906	6775	6850	6904.5	6921.667
9/18	6842	6850	6850	6848	6842.222
9/19	7039	6850	6850	7064	7050
9/21	6861	6850	6850	6848	6865
9/22	6926	6850	6850	6904.5	6920
9/23	6852	6850	6850	6848	6852.5
9/24	6890	6850	6850	6904.5	6890
9/25	6871	6850	6850	6848	6873.333
9/28	6840	6850	6750	6848	6840
9/29	6806	6850	6850	6796	6816.667
9/30	6787	6850	6750	6796	6790

4.3. Forecasting for KRW/USD exchange rate

Table V

Forecasted KUEXR

date(year/ month)	actual data	Liu's method	The Proposed method
06/01	983.76		
06/02	969.96	969	969.8333
06/03	975.17	979	979
06/04	952.92	969	969.8333
06/05	941.21	959	950.25
06/06	955.28	939	945.6667
06/07	950.56	944	943.375
06/08	961.02	959	950.25
06/09	952.77	944	943.375
06/10	954.17	959	950.25
06/11	935.52	944	943.375
06/12	925.08	939	929
07/01	936.90	929	927.75
07/02	937.17	939	939.625
07/03	943.23	939	939.625
07/04	930.95	939	945.6667
07/05	927.39	929	929
07/06	928.16	929	927.75
07/07	918.45	929	927.75
07/08	934.92	929	928.5833
07/09	930.89	939	929
07/10	914.81	929	929
07/11	918.11	929	928.5833
07/12	930.76	929	928.5833
08/01	942.72	929	929
08/02	944.42	939	945.6667
08/03	982.51	959	979
08/04	987.24	969	969.8333
08/05	1038.21	1039	1039
08/06	1031.07	1029	1029
08/07	1018.18	1019	1019
08/08	1047.11	1049	1049
08/09	1136.64	1139	1137.125
08/10	1326.85	1202.333	1329
08/11	1400.81	1399	1399
08/12	1368.80	1369	1369
09/01	1354.68	1359	1359
09/02	1440.19	1439	1439
09/03	1453.35	1449	1449
09/04	1336.28	1339	1339
09/05	1255.62	1259	1259
09/06	1262.28	1209	1249
09/07	1262.96	1209	1249
09/08	1239.69	1209	1249
09/09	1215.00	1219	1219
09/10	1174.80	1179	1179
09/11	1163.18	1159	1161.5
09/12	1166.13	1182.333	1165.25
10/01	1138.77	1179	1138.375
10/02	1156.83	1202.333	1156.5
10/03	1136.11	1182.333	1135.25
10/04	1115.71	1202.333	1119
10/05	1168.41	1169	1167.75
10/06	1214.02	1179	1219

In this section, we investigate KRW/USD average exchange rate per month(KUEXR) to ensure that the proposed method is superior to the other methods.

Step 1: According to the actual KUEXR shown in Table V, we can see that $x_{(1)} = 914.81$ and $x_{(2)} = 1453.35$ are actual minimum and maximum data, respectively.

Then we define the universe of discourse $U = [914, 1454]$.

Step 2: The length of the intervals using average-based length is 10(Huarng ([3])). Also, we let $f_{max} = 4$. Then we obtain 33 intervals having the frequency of historical data below f_{max}.

Step 3-5: We let $r_{max} = 4$. Then we obtain 41 intervals having r_i below r_{max}. For the sake of simplicity, we omit the processes of establishing fuzzy sets and fuzzy logical relationship groups.

Step 6: Defuzzify the forecasting results using the"Centroid". The forecasted KUEXR by proposed method and Liu([7]) are shown in Table V.

Fig.II shows the actual and forecasted data. From Fig.II we see that our method reflects the fluctuation of historical data.

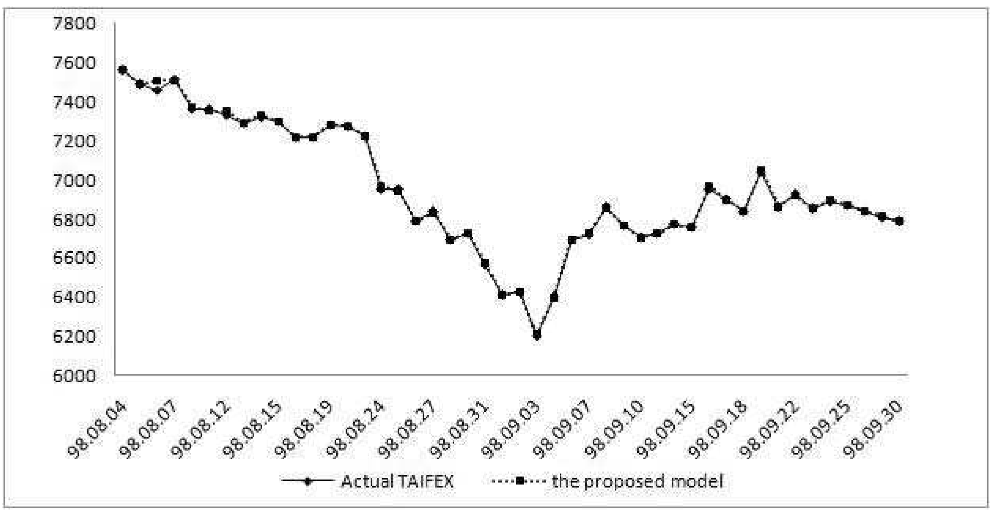

Fig.II. The actual and forecasted TAIFEX

5. The Forcasting Accuracy Validation

In this section, we compare a performance of proposed method with various fuzzy time series constructed by many authors.

The goal of forecasting, which is the process of estimation in unknown situations, is to be as accurate as possible. For this purpose, we consider a performance measure providing forecasting error, which is the difference between the actual value(x_t) and the forecasted value($\hat{x}_t$). The Mean Square Error(MSE) and Mean Absolute Percentage Error(MAPE) are used to measure the forecasting accuracy:

$$\mathrm{MSE} = \frac{1}{n}\sum_{t=1}^{n}(x_t - \hat{x}_t)^2,$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\frac{|x_t - \hat{x}_t|}{|x_t|}.$$

Table VI shows that the proposed method's forecasting accuracy is superior to existing methods. The MSE of Liu's method is 108097 but that of the proposed method is 4287. Also, the MAPE of Liu's method is 1.33 but that of the proposed method is 0.0028.

Table VI

The forecasting accuracy for enrollment

Index	Chen's method	Huarng's method	Cheng's method	Liu's method	The proposed method
MSE	407507	226611	191844	108097	4287
MAPE(%)	3.11	2.45	2.09	1.33	0.0028

Now, we consider the forecasting accuracy for the TAIFEX given by Chen([1]), Huarng([4]) and Lee et al.([5,6]). From Table VII, we can see that the MSE and MAPE of the proposed method are obviously smaller than those of other methods.

Table VII

The forecasting accuracy for TAIFEX

Index	Chen's method	Huarng's method	Lee et al.'s method([5])	Lee et al.'s method([6])	The proposed method
MSE	9668.97	5389.3	1364.56	105.02	67.71
MAPE(%)	1.051	0.857	0.434	0.119	0.00077

Table VIII shows that MSE and MAPE of the proposed method for our new target data KUEXR is obviously smaller than those of the Liu's method ([7]). The MSE of Liu's method is 756.9311 but that of the proposed method is 50.50642.

Table VIII

The forecasting accuracy for KUEXR

Index	Liu's method	The proposed method
MSE	756.4894	50.50642
MAPE(%)	0.013582	0.005156

We considered frequency of historical data belongs to one interval(step 2), and distance of the fuzzy logical relationship(step 5) at the same time. So we could reflect the fluctuation of historical data, which resulted in great performance.

6. Conclusions

We proposed a new method of fuzzy time series for improving the forecasting accuracy. In this paper, we considered the frequency of historical data belong to some fuzzy set and the distance between fuzzy sets in the fuzzy logical relationship group. The frequency and distance played important roles to improve the forecasting accuracy and to reflect the fluctuation of historical data. Empirical studies illustrated that the proposed method's forecasting accuracy is superior to existing methods.

References

1. Chen, S.M.: Forecasting enrollments based on fuzzy time series. *Fuzzy Sets and Systems* **81** (1996), 311-319.
2. Cheng, C.H., Chen, T.L., Teoh, H.J. & Chiang, C.H.: Fuzzy time-series based on adaptive expectation model for TAIEX forecasting. *Expert Systems with Applications* **34** (2008), 1126-1132.
3. Huarng, K.: Effective lengths of intervals to improve forecasting in fuzzy time series. *Fuzzy Sets and Systems* **123** (2001a), 155-162.
4. ______: Heuristic models of fuzzy time series for forecasting. *Fuzzy Sets and System* **123** (2001b), 369-386.
5. Lee, L.W., Wang, L.H., Chen, S.M. & Leu, Y.H.: Handling forecasting problems based on two-factors high-order fuzzy time series. *IEEE Transactions on Fuzzy Systems* **14** (2006), 468-477.
6. Lee, L.W., Wang, L.H. & Chen, S.M.: Temperature prediction and TAIFEX forecasting based on high-order fuzzy logical relationships and genetic simulated annealing techniques. *Expert Systems with Applications* **34** (2008), 328-336.
7. Liu, H.T.: An improved fuzzy times forecasting method using trapezoidal fuzzy numbers. *Fuzzy Optimization Decision Making* **6** (2007), 63-80.
8. Song, Q. & Chissom, B.S.: Forecasting enrollments with fuzzy time series-part I. *Fuzzy Sets and Systems* **54** (1993a), 1-10.
9. ______: Forecasting enrollments with fuzzy time series-part II. *Fuzzy Sets and Systems* **62** (1993a), 1-8.

[a]Department of Mathematics, Yonsei University, Seoul 120-749, Korea
Email address: pinkjhy@hanmail.net

[b]School of Economics, Yonsei University, Seoul 120-749, Korea
Email address: jin9135@yonsei.ac.kr

[c]School of Liberal Arts and Science, Korea Aerospace University, Goyang 412-791, Korea
Email address: shchoi@kau.ac.kr