Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844
Vol. VI (2011), No. 4 (December), pp. 603-614

Fuzzy Local Trend Transform based Fuzzy Time Series Forecasting Model

J. Dan, F. Dong, K. Hirota

Jingpei Dan, Fangyan Dong, Kaoru Hirota
Tokyo Institute of Technology
Japan, 226-8502 Yokohama, 4259 Nagatsuta, Midori-ku,
E-mail: {dan,tou,hirota}@hrt.dis.titech.ac.jp

Jingpei Dan
Chongqing University
P.R.China, 400044 Chongqing, 174 Shazhengjie, Shapingba
E-mail: danjingpei@hotmail.com

Abstract: A fuzzy local trend transform based fuzzy time series forecasting model is proposed to improve practicability and forecast accuracy by providing forecast of local trend variation based on the linguistic representation of ratios between any two consecutive points in original time series. Local trend variation satisfies a wide range of real applications for the forecast, the practicability is thereby improved. Specific values based on the forecasted local trend variations that reflect fluctuations in historical data are calculated accordingly to enhance the forecast accuracy. Compared with conventional models, the proposed model is validated by about 50% and 60% average improvement in terms of MLTE (mean local trend error) and RMSE (root mean squared error), respectively, for three typical forecasting applications. The MLTE results indicate that the proposed model outperforms conventional models significantly in reflecting fluctuations in historical data, and the improved RMSE results confirm an inherent enhancement of reflection of fluctuations in historical data and hence a better forecast accuracy. The potential applications of the proposed fuzzy local trend transform include time series clustering, classification, and indexing.
Keywords: time series forecasting, fuzzy time series, trend, transform.

1 Introduction

Based on Zadeh's works (see [1] and [2]), the concept of fuzzy time series and its models for forecasting have been proposed to solve the forecasting problems where the historical data are linguistic values (see [3]- [5]). Conventional fuzzy time series forecasting models that are based on fuzzy time series of original data are limited to forecasting specific values that do not reflect fluctuations in historical data. Local trend variations, however, are mainly concerned with real applications. For example, forecast of changing direction of stock price are more important for stock investors to make reasonable determinations than specific forecast values of stock price. In addition, the forecasted specific demand values are unreliable since historical data is distorted when it transfers along the supply chain due to the bullwhip effect, so local trend variations of demand that are not suffered from the bullwhip effect are more valuable and practical for supply chain managers. Therefore the practicability of conventional fuzzy time series forecasting methods suffers from the limitation of forecasting specific values. This study aims to improve the practicability and forecast accuracy by forecasting local trend variations that reflect fluctuations in historical data. It should be noted that the word "trend" usually refers to the long-term trend in statistics, whereas as used in this paper word "trend" means local trend variation in short term or during one period.

Recently, some trend involved fuzzy time series models have been proposed to improve forecasting. Huarng has proposed heuristic models by integrating problem-specific heuristic knowledge with Chen's model [6] to improve forecasting by reflecting the fluctuations in fuzzy time series [7]. A trend-weighted fuzzy time series model for forecasting Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) has been proposed in [8]. Chen and Wang have proposed a method to predict TAIEX based on fuzzy-trend logical relationship groups to improve forecast accuracy [9]. The interval rearranged method has been proposed to reflect fluctuations in historical data and improve forecast accuracy of fuzzy time series in [10]. Although all these methods are involved with trends, they are intrinsically conventional fuzzy time series forecasting methods since they are all based on original data and their forecasting targets are specific values. Fuzzy local trend transform is proposed to provide a different forecasting basis by transforming original data into a linguistic representation of local trend variations called fuzzy local trend time series. In contrast to conventional fuzzy time series forecasting models, local trend variations are forecasted based on the transformed fuzzy local trend time series in the proposed model. Forecast accuracy of specific values is hence enhanced by forecasted local trend variations that reflect fluctuations in historical data.

Three typical forecasting targets, enrollment forecasting, stock index forecasting, and inventory demand forecasting are used to validate the proposed model. To make an effective evaluation, forecasts are evaluated by two measures from different aspects. Root mean squared error (RMSE) is used to evaluate forecast accuracy of specific values while mean local trend error (MLTE) measure [11] is used to evaluate how accurately forecast reflects fluctuations in historical data. For enrollment forecasting, the proposed model outperforms typical fuzzy time series models in terms of RMSE and MLTE. Especially, comparing to Chen's model [6], Huarng's model [12], and Cheng et al.' model [13], MLTE results show an improvement of 73.3%, 55.6%, and 66.7%, respectively. For TAIEX forecasting, the proposed model gets the smallest RMSE result while the second rank in terms of MLTE compared to Chen's model [6], Yu's model [14], and Cheng's model [8]. For inventory demand forecasting, compared to Huarng and Yu' model [15], Cheng et al.' model [13], and Chen and Wang' model [9], the proposed model yields about 50%, 50%, and 33.3% improvement in terms of MLTE and 48.3%, 73.3%, and 49.1% improvement in terms of RMSE. The MLTE results demonstrate that the proposed model outperforms conventional fuzzy time series models in significantly reflecting fluctuations in historical data, and the improved RMSE results confirm an inherent enhancement of reflection of fluctuations in historical data and hence a better forecast accuracy.

The rest of the paper is organized as follows: in section 2, fuzzy time series and fuzzy c-means clustering are briefly reviewed. The proposed fuzzy local trend transform based fuzzy time series forecasting model is elaborated in section 3. Empirical analyses on three forecasting targets to demonstrate the proposed model are illustrated in section 4.

2 Fuzzy Time Series and Fuzzy C-means Clustering: a Brief Review

The proposed model is based on conventional fuzzy time series forecasting model and the fuzzy c-means clustering method is applied in the proposed fuzzy local trend transform, so fuzzy time series and fuzzy c-means clustering are briefly reviewed by adjusting the notations.

2.1 Fuzzy time series

Let U be the universe of discourse, where $U = \{u_1, u_2, \cdots, u_b\}$. A fuzzy set A_i of U is defined as $A_i = \sum_{j=1}^{b} f_{A_i}(u_j)/u_j$, where f_{A_i} is the membership function of the fuzzy set A_i; $f_{A_i} : U \to [0, 1]$, u_a is a generic element of fuzzy set A_i and $f_{A_i}(u_a)$ is the degree of belongingness of u_a to A_i; $f_{A_i}(u_a) \in [0, 1]$ and $a \in [1, b]$. In [3], the general definitions of fuzzy time series are given as follows:

Definition 1. Let a subset of real numbers $Y(t)(t = \cdots, 0, 1, 2, \cdots)$ be the universe of discourse by which fuzzy sets $f_j(t)$ are defined. If $F(t)$ is a collection of $f_1(t), f_2(t), \cdots$, then $F(t)$ is called a fuzzy time series defined on $Y(t)$.

Definition 2. If fuzzy time series relationships assume that $F(t)$ is caused only by $F(t-1)$, then the relationship can be expressed as: $F(t) = F(t-1) * R(t, t-1)$, which is the fuzzy relationship between $F(t)$ and $F(t-1)$, where $*$ represents as an operator. (Note that the operator can be either max-min [4], min-max [5], or arithmetic operator [6].)

To sum up, let $F(t-1) = A_i$, and $F(t) = A_j$, the fuzzy logical relationship between $F(t)$ and $F(t-1)$ can be denoted as $A_i \to A_j$, where A_i refers to the left-hand side and A_j refers to the right-hand side of the FLR. Furthermore, these fuzzy logical relationships can be grouped to establish different fuzzy relationship. These groups are called fuzzy logical relationship groups (FLRGs). On account of its simplicity, FLR method is chosen by most researchers.

The procedure for forecasting using conventional fuzzy time series models has four main steps: (1) Define universe of discourse and intervals; (2) Define fuzzy sets and fuzzify observations in the original time series; (3) Establish fuzzy relationships; (4) Forecast and defuzzify the outcome. Assume that the current state of $F(t)$ is A_i, $F_{def}(t+1)$ can be forecasted and defuzzified by the following rules:

Rule 1: If there is $A_i \to A_j$ in fuzzy logical relationship groups, then $F(t+1) = A_j$ and defuzzified as $F_{def}(t+1) = center_j$, where $center_j$ is the center of cluster j to which A_j belongs.

Rule 2: If there is $A_i \to \#$ in fuzzy logical relationship groups, then $F(t+1) = A_i$ and defuzzified as $F_{def}(t+1) = center_i$, where $\#$ represents null value and $center_i$ is the center of cluster i to which A_i belongs.

Rule 3: If there is $A_i A_1, A_2, \cdots, A_j$ in fuzzy logical relationship groups, then the forecast at $t+1$ is calculated as $F_{def}(t+1) = (center_1 + center_2 + \cdots + center_j)/j$, where c_j is the center of cluster to which A_j belongs.

2.2 Fuzzy c-means clustering

Fuzzy c-means (FCM) clustering is a method of clustering which allows one piece of data to belong to two or more clusters, and is frequently used in pattern recognition [16].

The FCM is based on minimization of the objective function

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m \|x_i - c_j\|^2, 1 \le m \le \infty, \tag{1}$$

where m is any real number greater than 1, u_{ij} is the degree of membership of x_i in the cluster j, x_i is the i-th of d-dimensional measured data, c_j is the d-dimension center of the cluster, and $\| * \|$ is any norm expressing the similarity between any measured data and the center.

Fuzzy partitioning is carried out through an iterative optimization of the objective function (1), with the update of membership $u_{ij} = 1/\sum_{k=1}^{C} (\|x_i - c_j\|/\|x_i - c_k\|)^{2/m-1}$ and the cluster centers $c_j = \sum_{i=1}^{N} u_{ij}^m (x_i / \sum_{i=1}^{N} u_{ij}^m)$.

This iteration will stop when $Max_{ij}|u_{ij}^{k+1} - u_{ij}^{k}| < \varepsilon$, where ε is a termination criterion between 0 and 1, whereas k is the iteration step. This procedure converges to a local minimum or a saddle point of J_m.

3 Proposed Fuzzy Local Trend Transform based Fuzzy Time Series Forecasting Model

In literature, trends are widely represented by using absolute variations, slopes, or relative variations between two consecutive points in literature. In section 3.1, trends or local trend variations in study are defined as relative variations, or ratios. The algorithm of fuzzy local trend transform based on local trend variations is then elaborated in section 3.2. The fuzzy time series forecasting model based on fuzzy local trend transform is presented in section 3.3.

3.1 Local trend variations

To address the limitation of absolute variations [17] and slopes([18]- [20])for representing local trend variations, relative variations, which are the ratios between two consecutive data points in a given historical time series, are adopted to indicate local trend variations in this study.

Assuming that for any given time series $P(t)$, $t = 1, 2, \cdots, n$, $n \in N$, local trend variation between time t and $t-1$ is defined as

$$r_t = (P(t) - P(t-1))/P(t-1), t = 2, 3, \cdots, n, \tag{2}$$

then the time series of local trend variations for $P(t)$ is defined as $T(t) = r_t$, $t = 2, \cdots, n$.

The reasons for forecasting based on local trend variations instead of original time series data are explained as follows. First, the forecasts based on original time series data may not reflect the fluctuations in historical data properly. In most previous studies, the forecasts are equal at some consecutive points which indicated that forecasting based on original time series are not appropriate for reflecting fluctuations [15].

Second, original time series data varies dramatically in different contexts while the time series of local trend variations varies slightly.

Finally, forecasting based on local trend variations are more suitable for reflecting fluctuations in historical data since directions and variation degrees of local trend variations can be indicated by signs and magnitudes of ratios easily. Ratios are preferable in terms of demonstrating the differences in various contexts, ratios-based lengths of intervals are hence adopted to improve fuzzy time series in [15].

Forecasting based on local trend variations are therefore considered more suitable for reflecting fluctuations in historical data and forecast accuracy should be further improved inherently.

3.2 Fuzzy local trend transform

To forecast local trend variations, original time series is represented by linguistic local trend variations of original data firstly, the algorithm of fuzzy local trend transform is hence proposed as follows:

Step 1 : Obtain the local trend variation time series $T(t)$ by calculating ratios between each two consecutive data points in the original time series $P(t)$ in terms of equation (2).

Step 2: Divide $T(t)$ into three basic clusters in terms of local trend changing direction, i.e., decreasing cluster T_d, unchanged cluster T_u, and increasing cluster T_i. It is easy to determine T_d and T_i in terms of sign of ratios after T_u is determined. Assume that the interval for unchanged

Table 1: Parameter ŚÁ for determining the interval of unchanged cluster

$Max(\|T(t)\|)(\times 10^{-2})$	$Max(\|T(t)\|) \leq 1$	$Max(\|T(t)\|) \leq 10$	$Max(\|T(t)\|) \leq 20$	$\cdots$
$\alpha(\times 10^{-2})$	0.01	0.1	0.2	$\cdots$

cluster is $[-\alpha, \alpha]$, α is determined by $Max(|T(t)|)$ according to Table 1 since it is possible that the definition of unchanged cluster varies from problem to problem. The observations in $T(t)$, of which the values are greater than α, are then assigned to T_i, while the observations of which the values less than $-\alpha$ are assigned to T_d.

Step 3: Divide T_i and T_d into c_i and c_d clusters by applying FCM, respectively. Assume that the number of observations in T_i and T_d are n_i and n_d, respectively, then the number of clusters for T_i, T_d and $T(t)$ are predefined by users as c_i $(2 \leq c_i \leq n_i)$, c_d $(2 \leq c_d \leq n_d)$, and c $(c = c_d + c_i + 1)$, respectively. T_i and T_d are then divided into c_i and c_d clusters by FCM, respectively, as described in section 2.2. Consequently, the cluster centers and memberships with respect to the clusters are obtained.

Step 4: Fuzzify the local trend time series $T(t)$ as fuzzy local trend time series $F_T(t)$. First, the clusters of $T(t)$ are achieved by combining the clustering results obtained in Step 2 and Step 3. Then the linguistic terms A_i $(i = 1, 2, \cdots, c)$ are defined corresponding to the clusters. $T(t)$ is finally fuzzified into $F_T(t)$ by assigning A_i to $T(t)$ when the maximum membership of $T(t)$ occurs at the cluster to which A_i belongs.

3.3 Fuzzy local trend transform based fuzzy time series forecasting model

Theoretically, the proposed fuzzy local trend transform can be integrated with any conventional fuzzy time series forecasting model. Because of simplicity, the proposed model is integrated with Chen's model [6] as stated in section 2.1. The proposed model differs from Chen' model in Step 1 and Step 4 as described in the following:

Step 1: Transform original time series into fuzzy local trend time series by the proposed fuzzy local trend transform as stated in section 3.2.

Step 2: Establish fuzzy logical relationships and fuzzy logical relationship groups based on fuzzy local trend time series obtained in Step 1 as described in section 2.1.

Step 3: Forecast and defuzzify the possible outcomes of local trend variations, which is denoted as $T_{def}(t)$, $t = 1, 2, \cdots, n-1$, based on fuzzy logical relationship groups as described in section 2.1.

Step 4: Calculate specific values based on forecasted local trend variations obtained in Step 3 in terms of equation(3) that is defined as

$$P_{pre}(t) = P(t) \times T_{def}(t), t = 1, 2, \cdots, n-1, \tag{3}$$

where $P_{pre}(t)$ indicates predicted specific values.

4 Empirical Analyses on Forecasting based on Fuzzy Local Trend Time Series

To validate the proposed model, three applications are used in the empirical analyses, including student enrollment forecasting (the enrollments of the University of Alabama [4]), stock index forecasting (Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX)), and inventory demand forecasting [15]. The first two data sets are algebraic growth data and widely used to validate fuzzy time series models in many relevant studies while the inventory demand data set is exponential growth data and typical in supply chain management application. We compare the proposed model with typical fuzzy time series models in terms of two forecast accuracy measures. One is conventional measure RMSE (root mean squared error) that is commonly used to measure forecast accuracy of specific values based on quantitative error in fuzzy time series forecasting, while the other is MLTE (mean local trend error), which is proposed to measure how accurately forecasts reflect fluctuations in actual data based on local trend error [11]. It is more effective and proper for comparing the models by evaluating them from two different aspects than using one or more conventional measures that based on quantitative error. For a given time series $y_t, t = 1, \cdots, n$, the prediction of y_t is f_t, $t = 1, \cdots, n$, RMSE is defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(f_t - y_t)^2}, t = 1, 2, \cdots, n, n \in N, \tag{4}$$

ĄĄ while MLTE is defined as

$$MLTE = \frac{1}{n-1}\sum_{i=1}^{n-1} E_i \times 100\%, \quad E_i = \begin{cases} 1 & sign(y_{t+1} - y_t) \neq sign(f_{t+1} - f_t) \\ 0 & sign(y_{t+1} - y_t) = sign(f_{t+1} - f_t) \end{cases} \quad t = 1, 2, \cdots, n, n \in N, \tag{5}$$

where E_i indicates the number of local trend change errors, sign indicates the operator for outputting the sign of the operand. When the predicted local trend variation is inconsistent with the original one in the same interval, E_i is equal to 1, otherwise E_i equals to 0 [11].

The experiments are implemented using Matlab R2010a. In section 4.1, the forecasting algorithm is illustrated step by step with the example of enrollment forecasting. Performance of the proposed model for stock index and inventory demand forecasting are analyzed in section 4.2 and section 4.3, respectively.

4.1 Forecasting enrollments

The yearly data of student enrollments of the University of Alabama from 1971 to 1992 are commonly used in previous studies on fuzzy time series to validate fuzzy time series models. To make proper comparison with other models, the same data set is used in this study to validate the proposed model. The procedure of forecasting enrollments by the proposed model is as follows:

Step 1: Transform original time series into fuzzy local trend time series by the proposed fuzzy local trend transform.

(1) Obtain the local trend time series $T(t)$ as shown in the third column of Table 2 by calculating the local trend variations of the enrollments between each two consecutive years in the original time series $P(t)$ in terms of equation(2) .

(2) Divide $T(t)$ into decreasing cluster T_d, unchanged cluster T_u, and increasing cluster T_i. Since none of the observations in $T(t)$ is equal to 0 and maximum absolute value of the observations in $T(t)$ is 7.6675%, α is determined to be 0.1% according to Table 1. Consequently, $T_d = \{-5.8274, -3.1385, -2.3840, -2.2714, -0.9638\}$, $T_u = \{0.0466\}$, and $T_i = \{0.1189,$ $0.4147,$ $0.6664,$ $1.6535,$ $1.8872,$ $1.9071,$ $2.2414,$ $3.8912,$ $4.5179,$ $5.1987,$ $5.4145,$ $5.4742,$ $5.9643,$ $5.9782,$ $7.6576\}$.

(3) Divide T_i and T_d into c_i into c_d clusters by applying FCM, respectively. For FCM, the predefined number of clusters $c = c_d + c_i + 1$. To make fair comparison, the total number of clusters is predefined to be 7 in this study as the same as used in the previous studies. Assume that $c_d = c_i = 3$, the cluster centers of T_d and T_i and the membership grades of T_d and T_i are shown in Table 3 and Table 4, respectively. Naturally, the center of unchanged cluster T_i is 0%.

(4) Fuzzify the local trend time series $T(t)$ into fuzzy local trend time series $F_T(t)$ as shown in the fourth column of Table 2. Combine the clustering results obtained in Step (2) and Step

Table 2: Forecasting local trend variations and specific values of the enrollments

Year	Actual enrollment	Local trend variation	Fuzzified local trend variation	Forecasted local trend variation ($\times 10^{-2}$)	Forecasted enrollment
1971	13055	-	-	-	-
1972	13563	3.8912	A6	-	-
1973	13867	2.2414	A5	0.1233	13580
1974	14696	5.9782	A7	1.1623	14028
1975	15460	5.1987	A6	3.8752	15266
1976	15311	-0.96378	A3	0.1233	15479
1977	15603	1.9071	A5	1.2224	15498
1978	15861	1.6535	A5	1.1623	15784
1979	16807	5.9643	A7	1.1623	16045
1980	16919	0.66639	A5	3.8752	17458
1981	16388	-3.1385	A2	1.1623	17116
1982	15433	-5.8274	A1	-2.3003	16011
1983	15497	0.4147	A5	1.2224	15622
1984	15145	-2.2714	A2	1.1623	15677
1985	15163	0.11885	A5	-2.3003	14797
1986	15984	5.4145	A7	1.1623	15339
1987	16859	5.4742	A7	3.8752	16603
1988	18150	7.6576	A7	3.8752	17512
1989	18970	4.5179	A6	3.8752	18853
1990	19328	1.8572	A5	0.1233	18993
1991	19337	0.046565	A4	1.1623	19553
1992	18876	-2.384	A2	-2.5770	18839
RMSE					438.18
MLTE					21.0526%

(3) to achieve the clusters of $T(t)$, the linguistic variables are then defined as shown in Table 5 according to the clusters. Each local trend variation in $T(t)$ is fuzzified by the linguistic variable to which the maximum membership belongs in terms of the results in Table 3 and Table 4.

Table 3: Membership grades of decreasing cluster for each linguistic variable

clusters	Cluster Centers	Linguistic variables	$T_d(1)$	$T_d(2)$	$T_d(3)$	$T_d(4)$	$T_d(5)$
1	-2.5770	A2	0.0000018	0.9	0.9785	0.9407	0.000055
2	-5.8231	A1	0.9999974	0.0393	0.0031	0.0070	0.000006
3	-0.9758	A3	0.0000008	0.0607	0.0184	0.0523	0.999939
		$F_T(t)$	A1	A2	A2	A2	A3

Step 2: The FLRs of $F_T(t)$ are established as shown in Table 6 according to Definition 3 in 2.1. Then, the FLRs are rearranged into FLRGs as shown in Table 7.

Step 3: The possible outcomes of local trend variations from 1973 to 1992 are forecasted and defuzzified as shown in the fifth column of Table 2.

Table 4: Membership grades of increasing cluster for each linguistic variable

clusters	Cluster Centers	Linguistic variables	$T_i(1)$	$T_i(2)$	$T_i(3)$	$T_i(4)$	$T_i(5)$	$T_i(6)$	$T_i(7)$	$T_i(8)$
1	4.3997	A6	0.0603	0.0387	0.0215	0.0238	0.0639	0.0684	0.1719	0.9139
2	1.2224	A5	0.9077	0.9416	0.9680	0.9667	0.9123	0.9063	0.7715	0.0332
3	6.0036	A7	0.0319	0.0197	0.0105	0.0095	0.0238	0.0253	0.0566	0.0530
		$F_T(t)$	A5	A5	A5	A5	A5	A5	A5	A6
			$T_i(9)$	$T_i(10)$	$T_i(11)$	$T_i(12)$	$T_i(13)$	$T_i(14)$	$T_i(15)$	
1	4.3997	A6	0.9924	0.4936	0.2486	0.1929	0.00063	0.00027	0.1947	
2	1.2224	A5	0.0013	0.0199	0.0146	0.0123	0.00007	0.00003	0.0499	
3	6.0036	A7	0.0063	0.4864	0.7371	0.7948	0.9993	0.9997	0.7554	
		$F_T(t)$	A6	A6	A7	A7	A7	A7	A7	

Step 4: The enrollment values from the year 1973 to 1992 are calculated based on the forecasted local trend variations obtained in Step 3 as shown in the last column of Table 2. The number of intervals of the universe of discourse affects forecasting results [21]. When analyzing the sensitivity of c, which varies in $\{7, 9, 11, 13\}$, the RMSE and MLTE results $\{502.401, 440.3566, 380.8953, 313.2307$

Table 5: Cluster centers of local trend variations of enrollments

Cluster Centers ($\times 10^{-2}$)	Linguistic variables
-5.8231	A1(big decrease)
-2.5770	A2(decrease)
-0.9758	A3(small decrease)
0	A4(almost unchanged)
1.2224	A5(small increase)
4.3997	A6(increase)
6.0036	A7(big increase)

and $\{23.1579\%, 24.2105\%, 20\%, 13.6842\%\}$ respectively. The average RMSE and MLTE results indicate that forecast error decrease as the number of clusters increases, in other words, the bigger the number of clusters the higher the forecast accuracy.

Table 6: Fuzzy logical relationships of local trend variations of enrollments

A6→A5	A5→A7	A7→A6	A6→A3	A3→A5	
A5→A5	A5→A7	A7→A5	A5→A2	A2→A1	
A1→A5	A5→A2	A2→A5	A5→A7	A7→A7	
A7→A7	A7→A6	A6→A5	A5→A4	A4→A2	A2→#

Table 7: Fuzzy logical relationship groups of local trend variations of enrollments

Group 1: A1→A5	Group 2: A2→A1,A5,#	Group 3: A3→A5	Group 4: A4→A2
Group 5: A5→A2, A4, A5, A7	Group 6: A6→A3, A5	Group 7: A7→A5, A6, A7	

Table 8: Comparisons for enrollment forecasting under the same conditions

Year	Actual enrollment	Chen [6]	Huarng [12]	Cheng et al. [13]	The proposed model
1971	13055	-	-	-	-ĄĄ
1972	13563	14000	14000	14242	-ĄĄ
1973	13867	14000	14000	14242	13580
1974	14696	14000	14000	14242	14028
1975	15460	15500	15500	15474.3	15266
1976	15311	16000	15500	15474.3	15479
1977	15603	16000	16000	15474.3	15498
1978	15861	16000	16000	15474.3	15784
1979	16807	16000	16000	16146.5	16045
1980	16919	16833	17500	16988.3	17458
1981	16388	16833	16000	16988.3	17116
1982	15433	16833	16000	16146.5	16011
1983	15497	16000	16000	15474.3	15622
1984	15145	16000	15500	15474.3	15677
1985	15163	16000	16000	15474.3	14797
1986	15984	16000	16000	15474.3	15339
1987	16859	16000	16000	16146.5	16603
1988	18150	16833	17500	16988.3	17512
1989	18970	19000	19000	19144	18853
1990	19328	19000	19000	19144	18993
1991	19337	19000	19500	19144	19553
1992	18876	19000	19000	19144	18839
RMSE ĄĄ		646.79	477.91	466.17	438.18
MLTE		78.9474%	47.3684%	63.1579%	21.0526%

The proposed method is compared with typical models for enrollment forecasting, including Chen's model [6], Huarng's model [12], and Cheng et al.' model [13]. To make fair comparison, seven linguistic variables are defined in all the compared models and the RMSEs and MLTEs are computed with reference to the year 1973. As shown in Table 8, the comparative results show that the proposed method outperforms the other models in terms of RMSE and MLTE.

Especially, comparing to Chen's model, Huarng's model, and Cheng et al.' model, the proposed model makes about 73.3%, 55.6%, and 66.7% improvements in terms of MLTE, respectively.

4.2 Stock index forecasting

Table 9: Comparisons for TAIEX forecasting

Date	Actual Index	Actual local trend ($\times 10^{-2}$)	Chen [6]	Yu [14]	Cheng [8]	The proposed model	Forecasted local trend ($\times 10^{-2}$)
00/11/02	5,626.08	1.477	5300	5340	5463.85	5524.81	-0.35
00/11/03	5,796.08	2.933	5750	5721.67	5644.8	5611.08	-0.27
00/11/04	5,677.30	-2.0922	5450	5435	5797.8	5815.39	0.33
00/11/06	5,657.48	-0.35033	5750	5721.67	5690.9	5657.47	-0.35
00/11/07	5,877.77	3.7478	5750	5721.67	5673.06	5648.99	-0.15
00/11/08	6,067.94	3.134	5750	5760	5871.32	5897.36	0.33
00/11/09	6,089.55	0.35487	6075	6062.5	6042.47	6088.09	0.33
00/11/10	6,088.74	-0.013303	6075	6062.5	6061.92	6093.84	0.07
00/11/13	5,793.52	-5.0957	6075	6062.5	6061.19	6079.54	-0.15
00/11/14	5,772.51	-0.36397	5450	5435	5795.5	5685.36	-1.87
00/11/15	5,737.02	-0.61861	5450	5435	5776.59	5763.81	-0.15
00/11/16	5,454.13	-5.1867	5450	5435	5744.65	5728.37	-0.15
00/11/17	5,351.36	-1.9204	5300	5340	5409.92	5352.29	-1.87
00/11/18	5,167.35	-3.561	5350	5350	5317.42	5304.12	-0.88
00/11/20	4,845.21	-6.6486	5150	5150	5151.81	5070.94	-1.87
00/11/21	5,103.00	5.0517	4850	4850	4861.89	4953.18	2.23
00/11/22	5,130.61	0.53814	5150	5150	5093.9	5120.92	0.35
00/11/23	5,146.92	0.31689	5150	5150	5118.75	5134.17	0.07
00/11/24	5,419.99	5.0382	5150	5150	5213.56	5150.48	0.07
00/11/27	5,433.78	0.25378	5300	5340	5459.32	5439.04	0.35
00/11/28	5,362.26	-1.3338	5300	5340	5391.6	5425.62	-0.15
00/11/29	5,319.46	-0.80459	5350	5350	5327.23	5314.92	-0.88
00/11/30	5,256.93	-1.1895	5350	5350	5288.71	5311.49	-0.15
00/12/01	5,342.06	1.5936	5250	5250	5232.44	5210.46	-0.88
00/12/02	5,277.35	-1.2262	5350	5350	5309.05	5327.84	-0.27
00/12/04	5,174.02	-1.9971	5250	5250	5250.81	5230.78	-0.88
00/12/05	5,199.20	0.48431	5150	5150	5157.82	5128.29	-0.88
00/12/06	5,170.62	-0.55274	5150	5150	5180.48	5202.82	0.07
00/12/07	5,212.73	0.80783	5150	5150	5154.76	5162.82	-0.15
00/12/08	5,252.83	0.7634	5250	5250	5192.66	5216.33	0.07
00/12/11	5,284.41	0.59761	5250	5250	5228.75	5256.46	0.07
00/12/12	5,380.09	1.7784	5250	5250	5257.17	5288.08	0.07
00/12/13	5,384.36	0.079304	5350	5350	5343.28	5398.00	0.33
00/12/14	5,320.16	-1.2067	5350	5350	5347.12	5376.30	-0.15
00/12/15	5,224.74	-1.8263	5350	5350	5289.34	5273.20	-0.88
00/12/16	5,134.10	-1.7655	5250	5250	5203.46	5178.54	-0.88
00/12/18	5,055.20	-1.5608	5150	5150	5121.89	5088.74	-0.88
00/12/19	5,040.25	-0.29661	5450	5405	5050.88	5010.54	-0.88
00/12/20	4,947.89	-1.8667	5450	5405	5037.42	5032.72	-0.15
00/12/21	4,817.22	-2.7126	4950	4950	4954.3	4904.19	-0.88
00/12/22	4,811.22	-0.12471	4850	4850	4836.7	4800.37	-0.35
00/12/26	4,721.36	-1.9033	4850	4850	4831.3	4803.96	-0.15
00/12/27	4,614.63	-2.3129	4750	4750	4750.42	4679.69	-0.88
00/12/28	4,797.14	3.8046	4650	4650	4654.37	4598.48	-0.35
00/12/29	4,743.94	-1.1214	4750	4750	4818.62	4813.06	0.33
00/12/30	4,739.09	-0.10234	4750	4750	4770.74	4736.76	-0.15
RMSE			176.32	170.27	121.47	114.63	
MLTE			64.44%	62.22%	26.67%	31.11%	

The daily stock index, TAIEX (Taiwan Stock Exchange Capitalization Weighted Stock Index), which is the other widely used data set in fuzzy time series studies, is used to further validate the proposed model for out-sample forecasting. The TAIEX data during 2000/01/01 - 2000/10/31 are used as training data set, and the data during 2000/11/01 - 2000/12/31 are used as testing data set.

The proposed model is compared with Chen's model [6], Yu's model [14], and Cheng's model [8]. The comparison of the forecasting results is shown in Table 9. The proposed model gets the smallest RMSE result while the second rank in terms of MLTE among the compared models. Cheng's model gets the smallest MLTE since it incorporates trend-weighting into Chen's model for TAIEX forecasting. The proposed model, however, improves forecast accuracy of specific value about 6% compared to Cheng's model by reflecting fluctuations in historical data.

4.3 Inventory demand forecasting

Demand forecasting plays a very important role in supply chain management. To further validate the applicability of the proposed method for demand forecasting, an inventory demand data set [15] is used in this study. This data set has been used in several previous studies, so it is proper for fair comparison. Inventory demand data from 1 to 19 are used as training set while data from 20 to 24 are used as testing set.

Table 10: Comparisons for inventory demand forecasting

Time	Actual inventory demand	Huarng and Yu [15]	Cheng et al. [13]	Chen and Wang [9]	The proposed model
20	227	206	205.5290	209.945	215.5516
21	223	228	216.4187	224.055	231.8985
22	242	228	216.4187	224.055	243.5307
23	239	244	216.4187	234.11	246.5969
24	266	244	216.4187	244.33	261.0037
RMSE		15.1	28.7295	14.88	7.6845
MLTE		100%	100%	75%	50%

The proposed model is compared with Huarng and Yu' model [15], Cheng et al.' model [13], and Chen and Wang' model [9]. As shown in Table 10 comparing with Huarng and Yu' model, Cheng et al.' model, and Chen and Wang' model, the MLTE results are improved by the proposed model about 50%, 50%, and 33.3%, respectively. This indicates that the proposed model outperforms the comparative models significantly in reflecting fluctuations in historical data. Consequently, the forecast accuracy of specific value is improved by the proposed model about 48.3%, 73.3%, and 49.1%, respectively, in terms of RMSE, which indicates that forecast accuracy is inherently improved by the proposed model by reflecting fluctuations in historical data.

5 Conclusion

In contrast to conventional fuzzy time series forecasting models that are based on original fuzzy time series, a different forecasting basis, fuzzy local trend time series, which is the linguistic representation of local trend variations of original data, is provided by the proposed fuzzy local trend transform. Local trend variations, which are defined as ratios between any two consecutive data points in original time series, are thereby forecasted based on fuzzy local trend time series and the specific values are are calculated accordingly. Therefore the practicability and forecast accuracy are improved by reflecting fluctuations in historical data.

The proposed model is validated by using three typical forecasting targets. The results are evaluated by two measures from different aspects. Compared to conventional fuzzy time series models, the proposed model yields about 50% and 60% average improvement in terms of MLTE and RMSE, respectively for the three application areas. The MLTE results indicate that the proposed model outperforms conventional fuzzy time series models significantly in reflecting fluctuations in historical data, and the improved RMSE results confirm an inherent enhancement of reflection of fluctuations in historical data and hence a better forecast accuracy.

Theoretically, the proposed fuzzy local trend transform can be integrated with any fuzzy time series model. Optimization of the parameters predefined in the proposed fuzzy local trend transform is being considered. A linguistic representation of time series data based on the proposed fuzzy local trend transform for the application to time series data mining is ongoing.

Bibliography

[1] L. A. Zadeh, Outline of a New Approach to the Analysis of Complex Systems and Decision Processes, *IEEE Trans. Systems, Man,and Cybernet*, Vol.3, No.1, pp.28-44, 1973.

[2] L.A. Zadeh, The Concept of a Linguistic Variable and its Application to Approximate Reasoning, Part 1, *Information Sciences*, Vol.8, No.3, pp.199-249, 1975.

[3] Q. Song, B. S. Chissom, Fuzzy Time Series and its Model, *Fuzzy Sets and Systems*, Vol.54, No.3, pp.269-277, 1993.

[4] Q. Song, 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, 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] K. H. Huarng, Heuristic Models of Fuzzy Time Series for Forecasting, *Fuzzy Sets and Systems*, Vol. 123, No.3, pp.369-386, 2001.

[8] C. H. Cheng et al, Trend-weighted Fuzzy Time Series Model for TAIEX Forecasing, *Lecture Notes in Compter Science*, Part III, Vol.4234, pp.469-477, 2006.

[9] S. M. Chen, N. Y. Wang, Fuzzy Forecasting based on Fuzzy-trend Logical Relationship Groups, *IEEE Trans. Systems, Man, and Cybernetics-Part B: Cybernetics*, Vol.40, No.5, pp.1343 - 1358, 2010.

[10] H. Y. Jung et al, Fuzzy Time Series Reflecting the Fluctuation of Historical Data, *Seventh International Conference on Fuzzy Systems and Knowledge Discovery*, pp.473-477, 2010.

[11] J. Dan et al, Mean local trend error and fuzzy-inference-based multicriteria evaluation for supply chain demand forecasting, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 15, No. 2, pp.134-144, 2011.

[12] J. R. Hwang et al, Handling Forecasting Problems Using Fuzzy Time Series, *Fuzzy Sets and Systems*, Vol.100, No.1-3, pp.217-228, 1998.

[13] C. H. Cheng et al, Multi-attribute Fuzzy Time Series Method based on Fuzzy Clustering, *Expert Systems with Applications*, Vol.34, No.2, pp.1235-1242, 2008.

[14] H. K. Yu, A Refined Fuzzy Time Series Model for Forecasting, *Physica A*, Vol.346, No.3/4, pp.657-681, 2005.

[15] K. H. Huarng, H. K. Yu, Ratio-based Lengths of Intervals to Improve Fuzzy Time Series Forecasting, *IEEE Trans. Systems, Man, and Cybernetics-Part B: Cybernetics*, Vol.36, No.2, pp.328-340, 2006.

[16] J. C. Bezdek et al, FCM: The Fuzzy C-means Clustering Algorithm, *Computers & Geosciences*, Vol.10, No.2-3, pp.191-203, 1984.

[17] J. R. Hwang et al, Handling Forecasting Problems Using Fuzzy Time Series, *Fuzzy Sets and Systems*, Vol.100, No.1-3, pp.217-228, 1998.

[18] A. Udechukwu et al, Discovering All Frequent Trends in Time Series, *The 2004 winter international symposium on information and communication technologies*, pp.1-6, 2004.

[19] C. H. Chen et al, Mining Fuzzy Frequent Trends from Time Series, *Expert Systems with Applications*, Vol.36, No.2, Part 2, pp.4147-4153, 2009.

[20] I. Z. Batyrshin, L. B. Sheremetov, Perception-based Approach to Time Series Data Mining, *Applied Soft Computing*, Vol.8, No.3, pp.1211-1221, 2008.

[21] K. H. Huarng, Effective Lengths of Intervals to Impove Forecasting in Fuzzy Time Series, *Fuzzy Sets and Systems*, Vol.124, No.3, pp.387-397, 2001.