Adaptive Time-Variant Model Optimization for Fuzzy-Time-Series Forecasting

Khalil Khiabani, Saeed Reza Aghabozorgi

***Abstract*-Fuzzy time series forecasting model is one of the tools that can be used to identify factors in order to solve the complex process and uncertainty, nowadays widely used in forecasting problems, but having appropriate universe of discourse and interval length are two subjects that exist in the Fuzzy time series. Recently Adaptive Time-Variant Model for fuzzy time series (ATVF) has been proposed with a computational method and an adaptive selection of analysis windows. In this paper, first we have introduced particle swarm optimization algorithm which is used for interval lengths improvement for ATVF model, another challenge that ATVF model confront with it is universe of discourse and this problem is solved using K-means clustering algorithm. Two models are applied to predict three data bases (the Enrolment of University of Alabama, Taiwan Futures Exchange (TAIFEX) and Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX)). The experimental results show that the proposed methods gets good forecasting results as compared to other existing fuzzy-time-series forecasting models.**

***Index Terms*-Fuzzy time series, Adaptive, Forecasting, Particle Swarm Optimization, K-means Clustering.**

I. INTRODUCTION

Forecasting activities are engaged in many applications forecasting such as computer networks [1] where proposed a Fast Handoff Scheme based on a Fuzzy Logic Predictive Control (FLPC), which allows to skip the channel scanning process stated in 802.11 standard, greatly reducing the handoff time. The forecasting problem of time series data, a series of data ordered in time sequence segmented by fixed time intervals, is an interesting and important research topic. In various disciplines it has been commonly tackled by using a variety of approaches such as statistics, artificial neural networks, etc. Traditional time series forecasting models are usually extensively dependent on historical data, which can be incomplete, imprecise and ambiguous. If these uncertainties were widespread in real-world data, they could hinder forecasting accuracy, thus limiting the applicability of forecasting models.

Obviously, we need to investigate some intelligent forecasting paradigm to solve the forecasting problems.

Zadeh proposed the fuzzy set theory first and then got fruitful achievements both in theory and applications [2]. Song and Chissom introduced a new forecast model based on the concept of fuzzy time series [3]. They use the time variant fuzzy time series model and the time-invariant fuzzy time series model based on the fuzzy set theory for forecasting the enrollments of the University Alabama. Chen improved the fuzzy time series model by max–min composition operations [4]. Tan et al. discussed about a model to automate the process of optimizing the BP network in prediction of time series trends has been developed [5]. Hassan et al. presented a hybrid fuzzy time series model. Based on ARIMA and Interval Type-2 Fuzzy Inference System (IT2-FIS) the proposed model will improve the forecasting result by handling the measurement and parametric uncertainties of ARIMA model using Fuzzy approach [6].

Chen and Chung used genetic algorithms to adjust each interval length of first-order and high-order forecasting models [7], [8]. Li et al. applied fuzzy c-means clustering to interval partitioning in [9]. Kuo et al. proposed an improved method of particle swarm optimization to find the proper content of the interval length [10]. Elaal et al. introduced multivariate-factors fuzzy time series forecasting model based on fuzzy clustering to handle real-world multivariate forecasting problems [11]. Khiabani et al. presented new method of incorporation of the adaptive time-variant fuzzy time series forecasting model (ATVF) with PSO algorithm to make interval length better for Alabama University enrollments forecasting [12]. ATVF model automatically adapts the time order of fuzzy time series based on the accuracy of prediction in the training phase and uses heuristic rules to determine prediction values [13].

We propose two methods for improving ATVF. Firstly, we have presented the combination of the ATVF model and particle swarm optimization (PSO) algorithm to optimize interval lengths. Secondly, ATVF model applies k-means (KM) clustering to deal with interval partitioning, which takes the nature of data points into account and produces unequal-sized intervals. The proposed models are better than existing fuzzy-time series forecasting models for the Enrolment of University of Alabama, Taiwan Futures Exchange (TAIFEX) and Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX). The rest of this paper is organized as follows. Section II procedure of forecasting TAIFEX using the ATVF model will be shown. Section III describes PSO Algorithm. Section IV explains k-means clustering. Section V describes the ATVF-PSO and ATVF-KM proposed models. Section VI demonstrates the experimental results. Section VII is the conclusion.

II. The procedure of forecasting TAIFEX using the ATVF model

The Enrolment of university of Alabama problem is used to introduce the forecasting procedure based on the adaptive time-variant fuzzy time series forecasting model [13]. The brief concept of the adaptive time-variant fuzzy

Manuscript received October 6, 2014; revised March 15, 2015. This work was supported in part by Mashhad Branch, Islamic Azad University, Young Researchers Club, Mashhad, Iran.

Khalil Khiabani is with the Mashhad Branch, Islamic Azad University, Young Researchers Club, Mashhad, Iran
(e-mail: *kiankhiabani@gmail.com*).

Saeed Reza Aghabozorgi is with the Department of Information Systems, FSKTM, University of Malaya, KualaLlumpur, Malaysia (e-mail: *saeed@um.edu.my*).

time series is also introduced in this section. The procedure of forecasting the TAIFEX based on the adaptive time variant fuzzy time series is described as follows:

Step 1. Fuzzify all historical data.

Historical data of Taiwan Futures Exchange (TAIFEX) are listed in Table I. Let Y(t) be the historical data on date t ($1998/8/3 \ll t \ll 1998/9/29$). Let the universe of discourse on Y(t) be $[D_{min} - U_{min}, D_{max} + U_{max}]$, where D_{min} and D_{max} denote the minimum and maximum historical data, respectively; and U_{min} and U_{max} denote the buffers to adjust the lower bound and the upper bound of the universe of discourse, respectively. According to Table I, it is obvious that D_{min}=6200 and D_{max}= 7560. For convenience to demo the forecasting example here, we set U_{min} = 0 and U_{max} = 40, thus the universe of discourse on Y(t)=[6200, 7600].

The universe of discourse is then cut into predefined number of intervals. The experimental results show that different total number of intervals and different lengths of all intervals are two main factors influencing the forecasted accuracy very much. For convenience here, we let the number of intervals be seven and the lengths of all intervals are equal. The seven intervals are I_1= (6200, 6400], I_2 = (6400, 6600], I_3= (6600, 6800], I_4= (6800, 7000], I_5 = (7000, 7200], I_6 = (7200, 7400] and I_7 = (7400, 7600].

There are seven linguistic values which are A_1="worst", A_2="bad", A_3="a little bad", A_4="average", A_5="good", A_6="very good", and A_7="excellent" to represent different regions in the universe of discourse on Y(t) respectively. Each A_i ($1 \le i \le 7$) denotes a fuzzy set, and its definition is described in Eq. (6), where the symbol '+' denotes the set union operator.

$$A_i = {}^{I_1}/_{u_1} + {}^{I_2}/_{u_2} + {}^{I_3}/_{u_3} + {}^{I_4}/_{u_4} + {}^{I_5}/_{u_5} + {}^{I_6}/_{u_6} + {}^{I_7}/_{u_7} \quad (6)$$

In Eq. (6), the symbol u_j ($1 \le j \le 7$) is a real number ($0 \le u_j \le 1$) and denotes the membership degree that I_j belongs to A_i ($1 \le i \le 7$). In other words, A_i denotes a fuzzy set = $\{I_1, I_2, I_3, I_4, I_5, I_6, I_7\}$ with different membership degree = $\{u_1, u_2, u_3, u_4, u_5, u_6, u_7\}$. The detailed definitions of all fuzzy sets are described in Eq. (7).

$$\begin{aligned}
A_1 &= {}^{I_1}/_{1} + {}^{I_2}/_{0.5} + {}^{I_3}/_{0} + {}^{I_4}/_{0} + {}^{I_5}/_{0} + {}^{I_6}/_{0} + {}^{I_7}/_{0} \\
A_2 &= {}^{I_1}/_{0.5} + {}^{I_2}/_{1} + {}^{I_3}/_{0.5} + {}^{I_4}/_{0} + {}^{I_5}/_{0} + {}^{I_6}/_{0} + {}^{I_7}/_{0} \\
A_3 &= {}^{I_1}/_{0} + {}^{I_2}/_{0.5} + {}^{I_3}/_{1} + {}^{I_4}/_{0.5} + {}^{I_5}/_{0} + {}^{I_6}/_{0} + {}^{I_7}/_{0} \\
A_4 &= {}^{I_1}/_{0} + {}^{I_2}/_{0} + {}^{I_3}/_{0.5} + {}^{I_4}/_{1} + {}^{I_5}/_{0.5} + {}^{I_6}/_{0} + {}^{I_7}/_{0} \\
A_5 &= {}^{I_1}/_{0} + {}^{I_2}/_{0} + {}^{I_3}/_{0} + {}^{I_4}/_{0.5} + {}^{I_5}/_{1} + {}^{I_6}/_{0.5} + {}^{I_7}/_{0} \\
A_6 &= {}^{I_1}/_{0} + {}^{I_2}/_{0} + {}^{I_3}/_{0} + {}^{I_4}/_{0} + {}^{I_5}/_{0.5} + {}^{I_6}/_{1} + {}^{I_7}/_{0.5} \\
A_7 &= {}^{I_1}/_{0} + {}^{I_2}/_{0} + {}^{I_3}/_{0} + {}^{I_4}/_{0} + {}^{I_5}/_{0} + {}^{I_6}/_{0.5} + {}^{I_7}/_{1}
\end{aligned} \quad (7)$$

In Eq. (7), for example, A_1 means the linguistic value ''worst" and denotes a fuzzy set = $\{I_1, I_2, I_3, I_4, I_5, I_6, I_7\}$ consisting of seven members with different membership degree = {1,0.5,0,0,0,0,0}. The remaining fuzzy sets can be described similarly to A_1 as mentioned above.

In this paper, we assume that a fuzzy set contains seven members (or intervals). On the contrary, an interval belongs to all fuzzy sets with different membership degrees. For example, I_1 belongs to A_1 and A_2 with membership degrees 1 and 0.5 respectively, and other fuzzy sets with membership degree 0. Meanwhile, I_2 belongs to A_1, A_2 and A_3 with membership degrees 0.5, 1 and 0.5 respectively, and other fuzzy sets with membership degree 0.

In order to fuzzify all historical data, it is 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. For example, the linguistic values ''worst" and ''bad" are assigned to intervals I_1 and I_2 respectively because I_1 and I_2 belongs to A_1 and A_2 both with the highest membership degree 1.

The following step is to fuzzify all historical data. The way to fuzzify a historical data is to find the interval to which it belongs and then assign the corresponding linguistic value to it. For example, the historical data on date 1998/8/5 is 7487, and it belongs to interval I_7 because 7487 is within (7400, 7600], so we then assign the linguistic value ''excellent" (i.e. the fuzzy set A_7) corresponding to interval I_7 to it. Table I lists the results of fuzzification, where all historical data are fuzzified to the corresponding fuzzy sets.

Let Y(t) be a historical data time series on date t. The purpose of Step 1 is to get a fuzzy time series *F(t)* on Y(t). Each element of Y(t) is an integer with respect to the actual data. But each element of *F(t)* is a linguistic value (i.e. a fuzzy set) with respect to the corresponding element of Y(t). For example, in Table I, Y(1998/8/5)=7487 and *F(1998/8/5)*= A_7; Y(1998/8/10)=7365 and *F(1998/8/10)*= A_6

Step2. Find out all fuzzy relationships

After the fuzzy time series F(t) is created, all fuzzy relationships can be established under different orders. To establish a win-order (win ≥ 1) fuzzy relationship, we should find out any relationship which has the type

$$\big(F(t-\text{win}), F(t-\text{win}+1), \dots, F(t-2), F(t-1)\big) \rightarrow F(t)$$

,where

$$F\big(F(t-\text{win}), F(t-\text{win}+1), \dots, F(t-2), F(t-1)\big)$$

and $F(t)$ are called the current state and the next state, respectively. Then a win-order fuzzy relationship is got by replacing the corresponding linguistic values. For example, $F(1998/8/3) \rightarrow F(1998/8/4)$ is a relationship; and a fuzzy relationship $A_7 \rightarrow A_7$ is obtained by replacing *F(1998/8/3)* and *F(1998/8/4)* to A_7 and A_7, respectively. The complete first-order fuzzy relationships are listed in Table II, where there are eight groups and each member of the same group has the same current state. The first 7 groups are called the trained patterns, and the last one is called the untrained pattern. For the untrained pattern, group 8 has the fuzzy relationship $A_4 \rightarrow \#$ as it is created by the relationship $F(1998/9/29) \rightarrow F(1998/9/30)$ Since the linguistic value of $F(1998/9/30)$ is unknown within the historical data, and this unknown next state is denoted by the symbol '#'.

or when we suppose win=3, a fuzzy relationship $(A_7, A_7, A_6) \rightarrow A_6$ is got as $F(1998/8/6)$, $F(1998/8/7)$, $F(1998/8/10) \rightarrow F(1998/8/11)$.

In ATVF model was proposed to improve forecasting using an adaptive algorithm to automatically adjust the order fuzzy relationships (analysis window) to find out all fuzzy relationships in the training phase of historical data. The training phase representing historical data will learn to predict the unknown data in the future. Two orders fuzzy relationships such as, **win1** and **win2** are selected to represent data analysis windows of sizes 1 and 2 as initial window sizes, and **pred** represents the future data estimated, for example, according to Table I, historical data of TAIFEX are:

$$\underbrace{\overbrace{\overbrace{\widehat{A_7}}^{win1}, A_7}^{win2}, \overset{pred}{\widehat{A_7}}, A_7, A_7, A_7, A_6, A_6, A_6, A_6, A_6, A_6, A_6, \cdots}_{\text{training phase}}, \underbrace{\cdots}_{\text{future}}$$

It is supposed that the prediction accuracy of win2 is higher than that of win1 and that the analysis windows of sizes 2 and 3 are automatically selected to predict the next time series value. The step is depicted as follows:

$$\underbrace{\overbrace{A_7, \overbrace{\widehat{A_7}, \widehat{A_7}}^{win1}}^{win2}, \overset{pred}{\widehat{A_7}}, A_7, A_7, A_6, A_6, A_6, A_6, A_6, A_6, \cdots}_{\text{training phase}}, \underbrace{\cdots}_{\text{future}}$$

The supposed current state is given as follows:

$$\underbrace{\cdots, \overbrace{\overbrace{A_4, A_3\ A_3, A_3, A_2, A_2}^{win1}, A_2}^{win2}, \overset{pred}{\widehat{A_1}}, A_2, A_3, A_3, A_4, A_3, A_3, \cdots}_{\text{training phase}}, \underbrace{\cdots}_{\text{future}}$$

If the prediction accuracy of win1 is higher than that of win2, the analysis windows of sizes 4 and 5 are automatically selected to predict the next time-series value. The step is depicted as follows.

$$\underbrace{\cdots, \overbrace{\overbrace{A_4, A_4\ A_3, A_3, A_2}^{win1}, A_2}^{win2}, \overset{pred}{\widehat{A_2}}, A_1, A_2, A_3, A_3, A_4, A_3, A_3, \cdots}_{\text{training phase}}, \underbrace{\cdots}_{\text{future}}$$

step3. Forecast the training or the testing data based on the forecast rules.

Based on Algorithms 1–2, the TAIFEX from 1998/8/3 to 1998/9/30 are forecasted by the presented method. The TAIFEX observations from 1998/8/3 to 1998/9/29 are used in the training phase, while the observations from 1998/9/29 to 1998/9/30 are used in the testing phase. Based on the historical data of the past years, only the enrollment of the next year can be forecasted. For example, the historical data from 1998/8/3 to 1998/9/29 is used for forecasting 1998/9/30. The TAIFEX forecasting of 1998/9/30 is based on the database from 1998/8/3 to 1998/9/29.

1) Training Phase:

[1998/8/5] Select analysis window sizes 1 and 2 as initial values and flag n = 1. Fuzzy relationship of 1998/8/4 and 1998/8/5: $A_7 \rightarrow A_7$. According to Algorithm 1, For1 = 7500, and For2 = 7488. Because $PA_1 < PA_2$ (PA_i denotes the predictive accuracy of the analysis window size i), the forecasting value of 1998/8/5 is **7488**. Window sizes 2 and 3 are selected for forecasting the TAIFEX in 1998/8/6, and flag n = 2.

[1998/8/6] Fuzzy relationship $A_7 \rightarrow A_7$. For1 = 7464.9, andFor2 = 7466.7. Because $PA_1 > PA_2$, Forecast =**7464.9**. The analysis window sizes of 1998/8/7 are 1 and 2. Flag n = 1.

[1998/8/7] Fuzzy relationship $A_7 \rightarrow A_6$. For1 = 7300, and For2 = 7300. There is no different between For1 and For2, the predictive accuracy of the small window size is the higher value. So, Forecast =**7300**. The analysis window sizes of 1998/8/10 are 2 and 3. Flag n = 2.

2) *Testing Phase:*

Algorithm 2 is used for forecasting, the results are given as follows.

[1998/9/30] Fuzzy relationship $A_4 \rightarrow$ #. Because the flag of 1998/9/29 is 24 and the flag of 1998/9/28 is 23, according to Rule 2, the analysis window sizes are 24 and 25 to forecast 1998/9/30. According to Algorithm 2, For3 = 6747.5, and For4 = 6748.3. Because $E_{1998/9/29} < E_{1998/9/28}$, according to Rule 4, Forecast = max (For3 For4) = **6748.3.**

III. PARTICLE SWARM OPTIMIZATION

The particle swarm optimization (PSO) has been proposed by Kennedy and Eberhart [15], [16]. The PSO consists of a swarm of particles that search for the best position with respect to the corresponding best solution for an optimization problem in the virtual search space, just like the birds blocking or the fish grouping.

Table I
The results of fuzzification.

Date	Actual data	Fuzzy Set	Forecasted data
1998/8/3	7552	A_7	
1998/8/4	7560	A_7	
1998/8/5	7487	A_7	7488
1998/8/6	7462	A_7	7464.9
1998/8/7	7515	A_7	7300
1998/8/10	7365	A_6	7360
1998/8/11	7360	A_6	7355.38
1998/8/12	7330	A_6	7327.69
1998/8/13	7291	A_6	7291.69
1998/8/14	7320	A_6	7319.2
1998/8/15	7300	A_6	7300
1998/8/17	7219	A_6	7239.46
1998/8/18	7220	A_6	7244.68
1998/8/19	7285	A_6	7285.6
1998/8/20	7274	A_6	7283.11
1998/8/21	7225	A_6	6900
1998/8/24	6955	A_4	6937.36
1998/8/25	6949	A_4	6700
1998/8/26	6790	A_3	6848.14
1998/8/27	6835	A_4	6756.7
1998/8/28	6695	A_3	6695.21
1998/8/29	6728	A_3	6544.25
1998/8/31	6566	A_2	6538.88
1998/9/1	6409	A_2	6443.91
1998/9/2	6430	A_2	6354.2
1998/9/3	6200	A_1	6500
1998/9/4	6403.2	A_2	6700
1998/9/5	6697.5	A_3	6697.6
1998/9/7	6722.3	A_3	6837.42
1998/9/8	6859.4	A_4	6752.57
1998/9/9	6769.6	A_3	6754.31
1998/9/10	6709.75	A_3	6709.32
1998/9/11	6726.5	A_3	6712.16
1998/9/14	6774.55	A_3	6747.77
1998/9/15	6762	A_3	6854.28
1998/9/16	6952.75	A_4	6933.1
1998/9/17	6906	A_4	6905.71
1998/9/18	6842	A_4	7100
1998/9/19	7039	A_5	6944.44
1998/9/21	6861	A_4	6881.65
1998/9/22	6926	A_4	6909.98
1998/9/23	6852	A_4	6874.62
1998/9/24	6890	A_4	6890.9
1998/9/25	6871	A_4	6882.72
1998/9/28	6840	A_4	6870.1
1998/9/29	6806	A_4	6752.4
1998/9/30	N/A	#	6748.3

Any particle maintains its personal best position that it has passed until now when it moves to another position. The moving method of a particle is described as discussed below:

$$v_i^{t+1} = \omega^t \times v_i^t + c_1 \times r_1 \times (P_{best} - x_i^t) + c_2 \times r_2 \times (G_{best} - x_i^t) \quad (8)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (9)$$

Table II
The complete first-order fuzzy relationships.

Group label	Fuzzy relationships
1	$A_1 \rightarrow A_2$
2	$A_2 \rightarrow A_1$, $A_2 \rightarrow A_2$, $A_2 \rightarrow A_3$
3	$A_3 \rightarrow A_2$, $A_3 \rightarrow A_3$, $A_3 \rightarrow A_4$
4	$A_4 \rightarrow A_3$, $A_4 \rightarrow A_4$, $A_4 \rightarrow A_5$
5	$A_5 \rightarrow A_4$
6	$A_6 \rightarrow A_4$, $A_6 \rightarrow A_6$
7	$A_7 \rightarrow A_6$, $A_7 \rightarrow A_7$
8	$A_4 \rightarrow \#$

In Eq. (8) and (9), the symbol v refers to the velocity of the particle i, and is limited to [$-v_{min}$, v_{max}] where v_{max} is a pre-defined by user. The symbol ω denotes the inertial weight coefficient. The symbols C_1 and C_2 denote the self-confidence coefficient and the social confidence coefficient, respectively. In a standard PSO, the value of ω decreases linearly during the whole running procedure, and C_1 and C_2 are constants [15]. The symbol Rand ( ) denotes a function can generate a random real number between 0 and 1 under normal distribution. The symbols x_i and P_{best} denote the current position and the personal best position of the particle i, respectively. The symbol P_{gbest} denotes the best one of all personal best positions of all particles within the swarm. The whole running procedure of the standard PSO is described in Algorithm 3.

Algorithm 3. Standard PSO algorithm

1: initialize all particles' positions and velocity
2: **while** the stop condition (the optimal solution is found or the maximal moving steps are reached) is not satisfied **do**
3: **for** all particles i **do**
4: move it to another position according to Eqs. (8) and (9)
5: **end for**
6: **end while**

IV. K-MEANS CLUSTERING

K-means (McQueen) is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem [20]. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. These centroids should be placed in a cunning way because of different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed and an early group age is done. At this point we need to re-calculate k new centroids as bary centres of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid. A loop has been generated. As a result of this loop we may notice that the k centroids change their location step by step until no more changes are done. In other words centroids do not move any more. Finally, this algorithm aims at minimizing an objective function, in this case a squared error function. The objective function is:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2 \quad (10)$$

Where $\left\| x_i^{(j)} - c_j \right\|$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster center c_j is an indicator of the distance of the n data points from their respective cluster centers.

Agorithm 4. Standard K-means Clustering algorithm

1: Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
2: Assign each object to the group that has the centroid.
3: When all objects have been assigned, recalculate the positions of the K centroids.
4: Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into which the groups from metric to be minimized can be calculated.

Although it can be proved that the procedure will always terminate, the k-means algorithm does not necessarily find the most optimal configuration, corresponding to the global objective function minimum. The algorithm 4 is also significantly sensitive to the initial randomly selected cluster centers. The k-means algorithm can be run multiple times to reduce this effect. K-means is a simple algorithm that has been adapted to many problem domains.

V. PROPOSED MODELS

a. ATVF-PSO MODEL

A new forecast model, named ATVF-PSO, consisting of the adaptive fuzzy time series and the particle swarm optimization, is proposed in this paper. In the ATVF-PSO model, for the training phase, the use of the particle swarm optimization is to train all fuzzy forecast rules under all training data. Once all fuzzy forecast rules have been well trained, for the testing phase, we can use the ATVF-PSO model to forecast the new testing data. The detailed descriptions of the ATVF-PSO model are given in the following. Let the number of the intervals be n, the lower bound and the upper bound of the universe of discourse on historical data Y(t) be b_0 and b_n respectively.

b_1	b_2	...	b_i	...	b_{n-1}

Fig. 1. The graphical particle representation.

A particle is a vector consisting of n-1 elements (i.e. b_1, b_2, . . ., b_i, . . ., b_{n-2} and b_{n-1}, where $1 \le i \le n-1$ and $b_{i-1} < b_i$); based on these n-1 elements, define the n intervals as $I_1 = (b_0,b_1]$, $I_2 = (b_1,b_2]$, . . ., $I_i = (b_{i-1},b_i]$, . . ., $I_{n-1} = (b_{n-2},b_{n-1}]$ and $I_n = (b_{n-1},b_n]$, respectively. If a particle moves to another position, the elements of the corresponding new vector need to be sorted first to ensure that each element b_i ($1 \le i \le n-1$) arranges in an ascending order. The graphical particle representation is given in Fig. 1. The ascending model exploits the intervals denoted by each particle to create an independent group of fuzzy forecast rules to forecast tall historical training data and get the forecasted accuracy for each particle. The mean square error (MSE) value is used to represent the forecasted accuracy of a particle for the training phase. The lower the MSE value is the better for the forecasted accuracy. The MSE function is defined in Eq (11), where the symbol $N_{forecasted}$ denotes the number of the forecasted data, the symbol FD_i denotes the ith forecasted data and the symbol TD_i denotes the corresponding historical training data with respect to FD_i.

$$MSE = \sum_{i=1}^{n} \frac{(FD_i - TD_i)^2}{n} \tag{11}$$

$$RMSE = \sqrt{MSE} \tag{12}$$

For the training phase, the ATVF-PSO model moves all particles to another position, respectively, according to Eqs. (8) and (9) and repeat the steps mentioned above to evaluate the forecasted accuracy of all particles until the pre-defined stop condition (the optimal solution is found or the maximal moving steps are reached) is satisfied. If the stop condition is satisfied, then all fuzzy forecast rules trained by the best one of all personal best positions of all particles are chosen to be the final result. For the testing phase, the ATVF-PSO model uses all well trained fuzzy forecast rules to forecast the new testing data. The detailed procedure of the ATVF-PSO model for the training phase and the testing phase is described in Algorithms 5 and 6, respectively.

An example of illustrating the ATVF-PSO model for the training phase is given in the following. In this example, let the number of intervals and particles be 7 and 5 respectively, and the ATVF-PSO model uses the PSO to train all adaptive win-order fuzzy forecast rules under all historical training data (i.e. actual data (Y(t)) in Table I, $(1998/8/3 \ll t \ll 1998/9/29)$. The symbols b_0 and b_7 denotes the lower bound and the upper bound of the universe of discourse on Y(t). Let b_0 be 6200 and b_7 be 7600 respectively; in other words, the universe of discourse on Y(t)= (6200, 7600]. Based on Y(t)= (6200, 7600], for Eqs. (8) and (9), we can let the range of x_i be limited to (6200, 7600], the range of v_i be limited to [-350, 350], the values of C_1 and C_2 be 2, and the value of ω be 1.4 (ω linearly decreases its value to the lower bound, 0.4, through the whole procedure) respectively. Initially, the randomized positions and velocities of all particles are listed in Tables III and IV, respectively.

Algorithm 5. The ATVF-PSO algorithm for training phase

1: initialize all particles' positions and velocity.
2: **while** the stop condition (the optimal solution is found or the maximal moving steps are reached) is not satisfied **do**
3: **for** all particles i **do**
4: fuzzify all historical training data according to all intervals defined by the current position of particle i
5: find out all win-order (window size) fuzzy relationships and forecast all historical training data according to all forecast rules based on training phase of ATVF algorithm.
6: calculate the MSE value for particle i based on Eq. (11)
7: update the personal best position of particle i, if the existing MSE value of the best position is larger than the one calculated in Step 6.
8: **end for**
9: **for** all particles i **do**
10: move particle i to another position according to Eqs. (8) and (9)
11: **end for**
12: **end while**

Algorithm 6. The ATVF-PSO algorithm for testing phase

The appropriate interval length and time order (window size) are determined in Algorithm 5 and are used in testing phase of ATVF algorithm (deffuzifier), So untrained data are estimated.

In Table III, each particle defines an independent group of 7 intervals which are I_1= $(b_0, b_1]$, I_2= $(b_1, b_2]$, I_3= $(b_2, b_3]$, I_4= $(b_3, b_4]$, I_5= $(b_4, b_5]$, I_6= $(b_5, b_6]$, I_7= $(b_6, b_7]$, respectively. For example, the intervals of the initial position of particle 1 listed in Table III are then I_1= (6200, 6562.1], I_2= (6562.1, 6683.9], I_3= (6683.9, 6746.8], I_4= (6746.8, 7172.8], I5= (7172.8, 7221.4], I_6= (7221.4, 7570], I_7= (7570, 7600], respectively. We assume the 7 intervals created by particle 1 are identical to the one used in the forecasting example mentioned in Section II. So we follow the whole forecasting procedure described in Section II with respect to the corresponding steps in algorithm 5 for the training phase, and the forecasted results listed in Table VIII.

The MSE value for particle 1 is calculated in Eq. (13) based on Eq. (11), where $N_{forecasted}$ denotes the number of the forecasted data (i.e. 44), FD_i $(1 \le i \le 44)$ denotes the forecasted data on $Y(1998/8/4 + i)$ and TDi denotes the corresponding historical training data (i.e. $Y(1998/8/4 + i)$ with respect to FD_i . Note that FD_{47} denotes the forecasted data on the new testing data (i.e. $Y(1998/9/30)$, thus it is not adopted to calculate the MSE value for particle 1 after all particles have got their own MSE values, every particle updates its own personal best position so far according to the MSE value.

$$MSE = \frac{\sum_{i=1}^{N_{forecasted}} (FD_i - TD_i)^2}{N_{forecasted}} = \frac{\sum_{i=1}^{44} (FD_i - TD_i)^2}{44}$$

$$\frac{(7483.48 - 7487)^2 + (7461.17 - 7462)^2 + \cdots + (6810.67 - 6806)^2}{44} = 3345.84 \tag{13}$$

Table III
The randomized initial positions of all particles.

	b1	b2	b3	b4	b5	b6	MSE
Particle 1	6562.1	6683.9	6746.8	7172.8	7221.4	7570	6139.32
Particle 2	6508.7	6703.5	6760.9	7213.1	7241.4	7570	6369.06
Particle 3	6597.2	6626.9	6642.2	7138.1	7330.6	7570	8428.1
Particle 4	6516.2	6555.5	6739.3	7176.4	7242.2	7570	3586.09
Particle 5	6604.8	6624.1	6637.5	7208.2	7312.7	7570	3345.84

Table IV
The randomized initial velocities of all particles.

	V_1	V_2	V_3	V_4	V_5	V_6
Particle 1	-100	55	39	64	-36	74
Particle 2	-1	-84	-20	28	66	-48
Particle 3	19	60	90	-29	-91	-86
Particle 4	22	12	44	74	10	-14
Particle 5	23	76	50	-65	23	-9

Note that the initial personal best positions are set as the initial positions of all particles. The personal best positions of all particles so far are listed in Table V. The global best position (i.e. Pgbest in Eq. (8)) is created by particle 5 as its MSE value is the least among all 5 particles so far. Now the ATVF-PSO model moves all particles to the second positions according to Eqs. (8) and (9). After the first iteration, the second positions and the corresponding new MSE values of all particles are listed in Table VI.

In Table VI, for example, the second position of particle 1 is calculated in Eqs. (14) and (15) based on Eqs. (8) and (9) respectively. By comparing the MSE values listed in Table VI with those listed in Table V, it is clear that particle 1, particle 2, particle 3 and particle 4, reached a better position than their own personal best positions so far.

Table V
The initial personal best positions of all particles. The global best position is created by particle 5 as its MSE is the least among all particles.

	b1	b2	b3	b4	b5	b6	MSE
Particle 1	6562.1	6683.9	6746.8	7172.8	7221.4	7570	6139.32
Particle 2	6508.7	6703.5	6760.9	7213.1	7241.4	7570	6369.06
Particle 3	6597.2	6626.9	6642.2	7138.1	7330.6	7570	8428.1
Particle 4	6516.2	6555.5	6739.3	7176.4	7242.2	7570	3586.09
Particle 5	6604.8	6624.1	6637.5	7208.2	7312.7	7570	3345.84

Table VI
The second positions of all particles.

	b1	b2	b3	b4	b5	b6	MSE
Particle 1	6241.7	6511.4	6638.7	7206.4	7518.9	7580.9	2817.5
Particle 2	6250.4	6589.4	6763.7	7095.2	7570	7583	5838.8
Particle 3	6459.3	6482.2	6509.6	7248.2	7750	7592.1	3647.3
Particle 4	6519.3	6601	6721.8	7192.9	7303.4	7523	9540
Particle 5	6624.2	6632.6	6578.9	7218.4	7346.8	7552.3	6668.6

Table VII
The personal best positions of all particles. The global best position is created by particle 2 as its MSE is the least for all particles.

	b1	b2	b3	b4	b5	b6	MSE
Particle 1	6233.9	6692.5	6802.6	6988	7557.2	7570	11694.7
Particle 2	6253.2	6877.6	6721.8	6950.8	7538.5	7581	881.9
Particle 3	6254.2	6703	6799.8	7542	7570	7586	9747.2
Particle 4	6290	6507.5	6960.7	6754.9	7570	7592.3	1582.9
Particle 5	6230	6695.1	6987.7	6890	7220.4	7557.2	11388.4

$$
\begin{aligned}
v_{1,1} &= 1.4 \times -100 + 2 \times Rand() \times (6562.1 - 6562.1) \times Rand() \times (6604.8 - 6562.1) \approx -320.4 \\
v_{1,2} &= 1.4 \times 55 + 2 \times Rand() \times (6683.9 - 6683.9) + 2 \times Rand() \times (6624.1 - 6683.9) \approx -172.5 \\
v_{1,3} &= 1.4 \times 39 + 2 \times Rand() \times (6746.8 - 6746.8) + 2 \times Rand() \times (6637.5 - 6746.8) \approx -108.1 \\
v_{1,4} &= 1.4 \times 64 + 2 \times Rand() \times (7172.8 - 7172.8) + 2 \times Rand() \times (7208.2 - 7172.8) \approx 33.6 \\
v_{1,5} &= 1.4 \times -36 + 2 \times Rand() \times (7221.4 - 7221.4) + 2 \times Rand() \times (7312.7 - 7221.4) \approx 297.5 \\
v_{1,6} &= 1.4 \times 74 + 2 \times Rand() \times (7570 - 7570) + 2 \times Rand() \times (7570 - 7570) \approx 10.9
\end{aligned} \tag{14}
$$

$$
\begin{aligned}
X_{1,1} &= 6562.1 + v_{1,1} = 6562.1 + (-320.4) \approx 6241.7 \\
X_{1,2} &= 6683.9 + v_{1,2} = 6683.9 + (-172.5) \approx 6511.4 \\
X_{1,3} &= 6746.8 + v_{1,3} = 6746.8 + (-108.1) \approx 6638.7 \\
X_{1,4} &= 7172.8 + v_{1,4} = 7172.8 + 33.6 \approx 7206.4 \\
X_{1,5} &= 7221.4 + v_{1,5} = 7221.4 + 297.5 \approx 7518.9 \\
X_{1,6} &= 7570 + v_{1,6} = 7570 + 10.9 \approx 7580.9
\end{aligned} \tag{15}
$$

The new global best position is created by particle 2 as its MSE is the least for all particles now. The ATVF-PSO model repeats the above steps until the stop condition is satisfied. So, for the last iteration, the second positions of all particles can be determined from Table VII and Eqs. (8) and (9), similarly.

At last, the well trained fuzzy forecast appropriate interval length and win-order (window size) are created by the global best position that the ATVF-PSO model reaches so far, and are used to forecast the new testing data in the testing phase based on algorithm 6 which is mentioned in testing phase of section II.

b. ATVF-KM MODEL

In most of the fuzzy time series literature, the universe of discourse was partitioned into equal-width intervals, which forms the most popular unsupervised discretization method. However, there are two major problems with interval partitioning. First, some of the parameters used for partitioning were selected arbitrarily. The other problem is that equal-width interval partitioning may not produce good results in cases where the distribution of continuous values is not uniform. According to these problems, we have presented another new method that is named ATVF-KM, including of the adaptive fuzzy time series and k-means clustering. In this model, we applied from k-means clustering for adjusting the universe of discourse. In this method, centers of each cluster are acquired by k-means clustering in training phase, such as C_1, C_2, ..., C_n and all data are put into each cluster and this technique continually repeated while the comparison between centers of each cluster and new ones are not be changed then centers are middle of triangular membership of fuzzy logic which are used for range of linguistic variables of fuzzy time series (fuzzify).

In testing phase, ATVF could forecast more accurately by using from improved universe of discourse by investigated model. The detailed procedure of the ATVF-KM model for the training phase and the testing phase is described in algorithms 7 and 8, respectively.

Algorithm 7. The ATVF-KM algorithm for training phase

1: find centers of each cluster by using k-means clustering algorithm (part 1 of algorithm 4).
2: **while** centers of clusters changed **do**
3: least distance between each data and center of each cluster according to Eqs. (10).
4: fuzzify all historical training data based on putting data in the best cluster (having the least distance (J) between data and cluster in Eqs. (10)).
5: **end while**
6: find out all win-order (window size) fuzzy relationships and forecast all historical training data according to all forecast rules based on training phase of ATVF algorithm.

Algorithm 8. The ATVF-KM algorithm for testing phase

The appropriate universe of discourse and time order (window size) are determined in Algorithm 7 and are used in testing phase of ATVF algorithm (deffuzifier), So untrained data are estimated.

$$
\begin{aligned}
(x_1^{(1)} - c_1)^2 &= (7552 - 6887.28)^2 = 441850.4 \\
(x_1^{(2)} - c_2)^2 &= (7552 - 7365)^2 = 34969 \\
(x_1^{(3)} - c_3)^2 &= (7552 - 7462)^2 = 8100 \\
(x_1^{(4)} - c_4)^2 &= (7552 - 7487)^2 = 4225 \\
(x_1^{(6)} - c_6)^2 &= (7552 - 7552)^2 = 0 \\
(x_1^{(7)} - c_7)^2 &= (7552 - 7560)^2 = 64
\end{aligned} \tag{16}
$$

Least distance between $x_1^{(6)} = 7552$ and $C_6 = 7552$ is 0, So $x_1^{(6)}$ Belongs to fuzzy set A_6.

In testing phase, ATVF could forecast more accurately by using from improved universe of discourse by investigated model. The detailed procedure of the ATVF-KM model for the training phase and the testing phase is described in algorithms 7 and 8, respectively.

The first 7 canters for seven clusters can be selected randomly (i.e.C_1=6887.28, C_2=7365, C_3=7462, C_4=7487, C_5=7515, C_6=7552, $C7$=7560) which are the middle of triangular membership. Note that we should choose the range of C_i be limited to (6200, 7600]. We then compute distance measure between a data point $x_i^{(j)}$ and the cluster centre C_j (i.e. $(x_i^{(j)} - c_j)^2$). For example, for $Y(1998/8/3) = 7552$, Distance measure between 7552 and each cluster centre is calculated in Eqs. (16).

Therby, all historical data are fuzzified to the corresponding fuzzy sets based on algorithm 7 in the first iteration. Afterward average data are located in a cluster. In other words the averages data are housed in the same fuzzy sets are calculated and thus their cluster centers are then computed in Eqs. (17). Similarly, we fuzzify all historical training data which means that putting data in the best new clusters based on algorithm 7 in the second iteration The ATVF-KM model repeats the above steps until the centers of each cluster be the same with the previous ones.

$$C_1=\frac{\sum_{i=8}^{46}\left\|x_i^{(1)}-c_1\right\|^2}{39}=\frac{(7360+7330+\cdots+6787)}{39}=6759.82, \tag{17}$$

C_2=7289.9, C_3=7462, C_4=7487, C_5=7515, C_6=7552, C_7=7560

At last, the well trained fuzzy forecast universe of discourse and win-order (window size) are created by the best fuzzy sets that the ATVF-KM model reaches so far, and are used to forecast the new testing data in the testing phase based on algorithm 8 which is mentioned in testing phase of section II.

VI. EXPERIMENTAL RESULTS

Methods presented using MATLAB has been implemented and we use the Enrolment of University of Alabama, Taiwan Futures Exchange (TAIFEX) and Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) as data sets in this paper.

Experimental results for ATVF-PSO model and ATVF-KM are compared with those of existing methods. First, in ATVF-PSO, Let the number of particles be 30, the maximal number of move for each particle be 100, the value of inertial weight (i.e. ω) is linearly decreased from 1.4 to 0.4, the self-confidence coefficient (C_1) and the social confidence coefficient (C_2) both be random, the velocity be limited to [-350, 350], respectively. Second, in ATVF-KM, according to the parameters of K-means Clustering, for changing universe of discourse, we have used the number of different canters. We compare pointed models (ATVF-PSO and ATVF-KM) with other models like ATVF, HPSO model [10], The C96 model [4], the Chen model [14], the combination of GA and fuzzy time series model [8], the SC2 model [18] and HCL98 [19] for different databases (the Enrolment of University of Alabama, Taiwan Futures Exchange (TAIFEX) and Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX)).

A. *Enrollment Forecasting*

Enrolment of University of Alabama with different number of intervals and different order of fuzzy time series is estimated with different methods. Based on RMSE values (i.e. Eqs. (12)) shown in Table IX, we can see that ATVF-PSO model is more precise than ATVF-KM, ATVF model, HPSO model, C96 model and the Chen's model, [8], SC2 [18] and HCL98 [19] at all.

According to the historical data for the past years, we can forecast the new enrollment for the next year only. For example, the historical data of enrollments under years 1971–1989, is used to forecast the new enrollment of year 1990. To forecast the new enrollment of year 1991, it is then based on the enrollments under years 1971–1990 among the existing methods.

TableVIII

The results of fuzzification.

Date	Actual data	Fuzzy Set	Forecasted data
1998/8/3	7552	A_6	
1998/8/4	7560	A_6	
1998/8/5	7487	A_6	7483.48
1998/8/6	7462	A_6	7461.17
1998/8/7	7515	A_6	7505.47
1998/8/10	7365	A_6	7377.23
1998/8/11	7360	A_6	7372.58
1998/8/12	7330	A_6	7277.36
1998/8/13	7291	A_5	7354.05
1998/8/14	7320	A_6	7277.54
1998/8/15	7300	A_5	7282.6
1998/8/17	7219	A_5	7240.57
1998/8/18	7220	A_5	7239.1
1998/8/19	7285	A_5	7273.1
1998/8/20	7274	A_5	7267.1
1998/8/21	7225	A_5	7129.78
1998/8/24	6955	A_4	6953.71
1998/8/25	6949	A_4	6947.95
1998/8/26	6790	A_4	6795.31
1998/8/27	6835	A_4	6838.51
1998/8/28	6695	A_4	6720.37
1998/8/29	6728	A_4	6402.4
1998/8/31	6566	A_1	6544.9
1998/9/1	6409	A_1	6408.7
1998/9/2	6430	A_1	6428.89
1998/9/3	6200	A_1	6248.77
1998/9/4	6403.2	A_1	6922.85
1998/9/5	6697.5	A_4	6725.45
1998/9/7	6722.3	A_4	6740.86
1998/9/8	6859.4	A_4	6861.93
1998/9/9	6769.6	A_4	6775.73
1998/9/10	6709.75	A_4	6738.69
1998/9/11	6726.5	A_4	6746.16
1998/9/14	6774.55	A_4	6780.48
1998/9/15	6762	A_4	6774.37
1998/9/16	6952.75	A_4	6951.55
1998/9/17	6906	A_4	6906.67
1998/9/18	6842	A_4	6845.23
1998/9/19	7039	A_4	7030.06
1998/9/21	6861	A_4	6863.47
1998/9/22	6926	A_4	6925.87
1998/9/23	6852	A_4	6854.83
1998/9/24	6890	A_4	6891.3
1998/9/25	6871	A_4	6873.07
1998/9/28	6840	A_4	6843.31
1998/9/29	6806	A_4	6810.67

We then only show the comparisons of the forecasted results produced based on the number of different intervals and the vote Wh=15 (constant-defined by the user) in Tables (X-XII). Considering that ATVF-PSO model and ATVF-KM is more accurately than other models. Which means All forecasting models are well trained by historical training data to forecast the new testing data (i.e. the enrollments of years 1990, 1991, and 1992), and to use the RMSE values to evaluate the forecasted accuracy.

A comparison between the proposed method and Singh's method [17] in terms of forecasted accuracy results under different numbers of intervals are shown in Table XIII. It is clear that the proposed models show higher accuracy than Singh's method. (Note: the accuracy of three models rise steadily with increase of intervals.)

B. *TAIFEX Forecasting*

The training data and testing data of TAIFEX are years 1998/8/3-1998/9/25 and 1998/9/28-1998/9/30, respectively, For Taiwan Futures Exchange (TAIFEX) data base, as shown in Fig. 2.

Table IX
A comparison of the forecasted enrollments with different number of intervals and different order of fuzzy time series

Year	Actual data	C96	HPSO	Chen	[8]	SC2	HCL98	ATVF	ATVF-PSO	ATVF- KM
1971	13055	14000	13676.8	14129.9						
1972	13563	14000	13676.8	14129.9	13494					
1973	13867	15500	14766	14361.1	15078.6			14538.6	14681.5	14499.5
1974	14696	15500	15721.8	14944.4	15078.6	14286	14500	15650	14934.5	15625
1975	15460	15785.7	15721.8	15770.8	16189.9	14700	15361	15150.5	15590	15061.1
1976	15311	15785.7	15721.8	14750	15819.5	14800	16260	15513.8	15422.9	15509.2
1977	15603	15785.7	16376.7	15479.2	16318.9	15400	15511	15603	15603	15625
1978	15861	17500	16376.7	15479.2	15634.3	15500	16003	16879.1	15861	16705.8
1979	16807	16000	16376.7	16791.7	15634.3	15500	16261	16807	16807	16799.3
1980	16919	16000	16376.7	16791.7	15634.3	16800	17407	16288	16919	16286.5
1981	16388	15500	15482.9	16208.3	15634.3	16200	17119	15546.6	16388	15716.3
1982	15433	15785.7	15721.8	15479.2	15634.3	16400	16188	15500.2	15553.9	15433
1983	15497	15785.7	15482.9	15479.2	15634.3	16800	14833	15105.7	15497	15079.6
1984	15145	15785.7	15721.8	14750	15634.3	16400	15497	15105.4	15255.6	15107.7
1985	15163	15785.7	15721.8	14750	15634.3	15500	14745	16214.6	15276.5	15945.5
1986	15984	17500	16376.7	16208.3	15634.3	15500	15163	16868.2	15984	15951.4
1987	16859	18500	18518.4	16791.6	17946.1	15500	16384	15719.2	16859	16838.6
1988	18150	19500	18518.4	18250	19155.3	16800	17659	19168.1	18736.2	18956
1989	18970	19500	18518.4	19708.3	19155.3	19300	19150	19075.2	18970	19007.3
1990	19328	19500	18518.4	19125	19155.3	17800	19770	19173.2	19204.6	19226.5
1991	19337	19500	18518.4	19125	19155.3	19300	19928	19271.8	19212.5	19199.4
1992	18876	19500	18518.4	19125	19155.3	19600	19537	19034.1	18876	19072.7
RMSE		911.51	677.23	317.35	708.75	880.7	566.9	635.91	**230.4**	**476.20**

The two presented models gives more accurate TAIFEX fluctuation forecasts than other models under the same intervals. Table XIV shows that the proposed models outperform the models reported in ATVF model under different number of intervals and different order of fuzzy time series. Some experimental results of the TAIFEX forecasting models for the testing phase in the number of different intervals and the vote W_h=15 are listed in Tables (XV-XVII). The observations show that the ATVF-KM is often more suited model than ATVF-PSO model, On the other hand, in Tables (XV-XVII), it is clear that the ATVF-PSO model shows superiority over C96 model, [8], ATVF model, Chen's model and HPSO model based on the number of different intervals and the vote W_h=15 in testing phase. In Table XVIII, presented models outperform Singh's method.

Table X
A comparison of the forecasted results produced based on the number of intervals = 10 and the vote W_h=15

Year	Actual data	C96	[8]	HPSO	ATVF	ATVF-PSO	ATVF-KM
1990	19328	18168	18059	18326	19540	19226	19525
1991	19337	18909	18669	19212	19137	19182	19150
1992	18876	19609	19083	19203	18876	18876	18933
RMSE		773.66	576.66	484.16	168.03	**107.29**	**160.43**

Table XI
A comparison of the forecasted results produced based on the number of intervals = 14 and the vote W_h=15

Year	Actual data	C96	[8]	HPSO	ATVF	ATVF-PSO	ATVF-KM
1990	19328	18162	17862	18120	19288	19238	19287
1991	19337	18721	18633	19027	18766	19224	18811
1992	18876	19221	19085	19137	18835	18945	18836
RMSE		709	653.66	621.91	331.62	**92.35**	**305.7**

Table XII
A comparison of the forecasted results produced based on the number of intervals = 15 and the vote W_h=15

Year	Actual data	C96	Chen	[8]	HPSO	ATVF	ATVF-PSO	ATVF- KM
1990	19328	18288	18979	17903	18637	19328	19270	19328
1991	19337	18810	18979	18542	18873	18912	19276	18926
1992	18876	19277	18979	19087	19098	18834	18971	18834
RMSE		656	294.6	669.66	492.55	246.36	**73.34**	**238.82**

C. TAIEX Forecasting

The TAIEX of the period from 2000/11/02 to 2000/12/30 is also used for model validation. Table XIX compares the TAIEX forecasts of the ATVF-PSO model, different order of fuzzy time series. TAIEX of year 2000/12/28. All forecasting models are well trained by historical training data to forecast the new testing data (i.e. forecasted accuracy. Fig. 3 compares different TAIEX forecasting results among HPSO model, C96 model, Chen's model, ATVF model, the ATVF-PSO model and ATVF-KM model under the same number of intervals and shows that our methods outperforms the mentioned models. The historical data of TAIEX under years 2000/11/02-2000/12/27 is used to forecast the new testing data (i.e. the TAIEX of years 2000/12/28, 2000/12/29, and 2000/12/30), and to use the RMSE values to evaluate the forecasted accuracy. ATVF-KM and [8] with different number of intervals and The actual TAIEX results, forecasted by Chen's model, C96 model, [8], HPSO model, ATVF model and the proposed models with different numbers of intervals are compared in Table (XX-XXII). These show that the proposed models more accurately than the other models in the testing phase of

Table XIII
Comparisons Under Differnet Numbers of Intervals

Models	Number of Intervals				
	6	8	10	15	20
Singh	342.29	310.85	170.5	150.88	128.85
ATVF-PSO	258.3	168.3	121.33	110.29	86.26
ATVF-KM	272.42	213.54	144.45	120.48	100.54

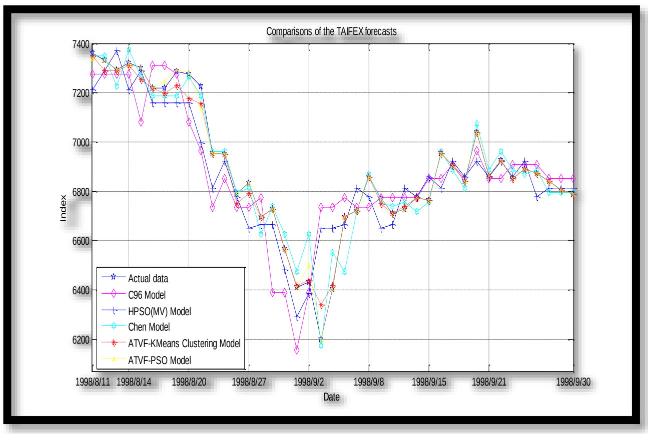

Fig. 2. Comparison of TAIFEX forecasting

the TAIEX forecasting. In Table XXIII, the proposed methods gets Smaller forecasting error rate than Singh's method for TAIEX forecasting.

Table XIV
A comparison of the forecasted TAIFEX with different number of intervals and different order of fuzzy time series

Year	Actual data	ATVF	ATVF-PSO	ATVF-KM
1998/8/3	7552			
1998/8/4	7560			
1998/8/5	7487			
1998/8/6	7462	7537.5		
1998/8/7	7515	7231.54		
1998/8/10	7365	7512	7353.36	7362.19
1998/8/11	7360	7450.87	7347.31	7348.02
1998/8/12	7330	7385.89	7330	7287.5
1998/8/13	7291	7306.63	7291	7287.5
1998/8/14	7320	7315.96	7313.13	7311.77
1998/8/15	7300	7291	7300	7253
1998/8/17	7219	7307.48	7219	7219
1998/8/18	7220	7207.79	7228.65	7197.87
1998/8/19	7285	7190.77	7285	7230.37
1998/8/20	7274	7257.17	7274	7175.75
1998/8/21	7225	7285	7121.33	7155
1998/8/24	6955	7073.02	6955	6955
1998/8/25	6949	7055.53	6949	6949
1998/8/26	6790	6939.66	6790	6750.51
1998/8/27	6835	6842.04	6835	6792.75
1998/8/28	6695	6790	6695	6695
1998/8/29	6728	6721.78	6728	6728
1998/8/31	6566	6774.9	6566	6566
1998/9/1	6409	6840.36	6409	6418.42
1998/9/2	6430	6447.58	6459.59	6440.21
1998/9/3	6200	6409	6200	6339.5
1998/9/4	6403.2	6712.19	6403.2	6418.61
1998/9/5	6697.5	6618.46	6697.5	6697.5
1998/9/7	6722.3	6688.3	6722.3	6722.3
1998/9/8	6859.4	6690.29	6859.4	6859.4
1998/9/9	6769.6	6867.96	6769.6	6750.52
1998/9/10	6709.75	6813.88	6709.75	6709.75
1998/9/11	6726.5	6672.43	6726.5	6738.50
1998/9/14	6774.55	6773.97	6774.55	6774.55
1998/9/15	6762	6769.61	6762	6762
1998/9/16	6952.75	6782.17	6952.75	6952.75
1998/9/17	6906	6893.17	6906	6906
1998/9/18	6842	6939.19	6842	6842
1998/9/19	7039	6768.38	7039	7039
1998/9/21	6861	7007.14	6861	6861
1998/9/22	6926	6911.6	6926	6926
1998/9/23	6852	6891.9	6852	6852
1998/9/24	6890	6926	6890	6890
1998/9/25	6871	6888.53	6871	6871
1998/9/28	6840	6896.52	6840	6840
1998/9/29	6806	6799.53	6806	6806
1998/9/30	6787	6812.34	6787	6787
RMSE		131.96	**16.93**	**33.09**

Table XV
A comparison of the forecasted results produced based on the number of intervals = 8 and the vote W_h=15

Year	Actual data	C96	[8]	Chen	HPSO	ATVF	ATVF-PSO	ATVF-KM
98/9/28	6840	6850	6834.4	6683.64	6842.31	6818.29	6840	6845
98/9/29	6806	6850	6775.83	6786.82	6849.43	6823.48	6816.1	6812
98/9/30	6787	6850	6774.29	6786.82	6822.71	6823.99	6798.39	6792
RMSE		39	16.09	11.85	27.15	26.74	**8.80**	**7.33**

Table XVI
A comparison of the forecasted results produced based on the number of intervals = 12 and the vote W_h=15

Year	Actual data	C96	[8]	Chen	HPSO	ATVF	ATVF-PSO	ATVF-KM
98/9/28	6840	6913.22	6851.73	6835.35	6848.28	6840.22	6695	6840
98/9/29	6806	6919.35	6813.55	6839.03	6830.66	6843.66	6728	6806
98/9/30	6787	6820.77	6783.01	6797.15	6814.54	6833.7	6576.59	6799.31
RMSE		73.45	8.37	17.45	20.16	34.64	**6.29**	**7.1**

Table XVII
A comparison of the forecasted results produced based on the number of intervals = 15 and the vote W_h=15

Year	Actual data	C96	[8]	Chen	HPSO	ATVF	ATVF- PSO	ATVF-KM
98/9/28	6840	6907.44	6824.24	6824.87	6857.78	6858.88	6829.87	6828.95
98/9/29	6806	6912.81	6805.92	6815.35	6818.83	6847.78	6815.89	6794.04
98/9/30	6787	6826.55	6775.38	6797.15	6812.61	6840	6797.99	6787
RMSE		70.94	11.3	18.72	18.74	40.46	**10.34**	**9.4**

Table XVIII
Comparisons Under Differnet Numbers of Intervals

Models	Number of Intervals				
	6	8	10	15	20
Singh	150.22	135.48	120.53	100.88	128.85
ATVF-PSO	100.33	85.64	70.26	45.5	21.5
ATVF-KM	120.58	95.52	88.37	51.45	35.27

Table XIX
A comparison of the forecasted TAIEX with different number of intervals and different order of fuzzy time series

Year	Actual data	[8]	ATVF-PSO	ATVF-KM
2000/11/02	5626.08	5697.42		
2000/11/03	5796.08	5697.42		
2000/11/04	5677.3	5902.72		
2000/11/06	5657.48	5902.72	5657.48	
2000/11/07	5877.77	6078.83	5877.77	
2000/11/08	6067.94	6078.83	5886.85	6088.84
2000/11/09	6089.55	5902.72	5897.65	6089.14
2000/11/10	6088.74	5614.84	6088.74	5793.52
2000/11/13	5793.52	5697.42	5793.52	5992.19
2000/11/14	5572.51	5503.36	5572.51	5772.13
2000/11/15	5737.02	5360.63	5737.02	5498.57
2000/11/16	5454.13	5209.94	5454.13	5342.9
2000/11/17	5351.3	64950.41	5351.36	5209.97
2000/11/18	5167.35	5110.91	5167.35	4792.92
2000/11/19	4845.21	5110.91	4913.46	4792.92
2000/11/20	5103	5156.2	5141.25	4913.83
2000/11/21	5130.61	4950.4	5167.36	5193.42
2000/11/22	5146.92	5407.29	5182.12	5469.62
2000/11/23	5419.99	5360.63	5419.99	5461.30
2000/11/24	5433.78	5312.53	5433.78	5349.79
2000/11/27	5362.26	5268.27	5362.26	5340.83
2000/11/28	5319.46	5360.63	5319.46	5190.93
2000/11/29	5256.93	5268.27	5256.93	5304.75
2000/11/30	5342.06	5209.94	5342.06	5331.22
2000/12/01	5277.35	5209.94	5277.35	5228.58
2000/12/04	5147.02	5229.38	5174.02	5184.61
2000/12/05	5199.2	4950.41	5199.2	5199.20
2000/12/06	5170.62	4950.41	5170.62	5132.12
2000/12/07	5212.73	5312.53	5459.24	5172.42
2000/12/08	5252.83	5360.63	5479.29	5285.20
2000/12/11	5284.41	5407.29	5495.08	5327.26
2000/12/12	5380.09	5312.53	5380.09	5338.90
2000/12/13	5384.36	5209.94	5384.36	5344.91
2000/12/14	5320.16	5110.91	5320.16	5221.84
2000/12/15	5224.74	5110.91	5224.74	5212.81
2000/12/16	5134.1	4950.41	5134.1	4999.48
2000/12/18	5055.2	4950.41	5083.33	4983.74
2000/12/19	5040.25	4950.41	5068.93	4973.71
2000/12/20	4947.89	4950.41	5002.31	4890.46
2000/12/21	4817.22	4950.41	4946.20	4785.89
2000/12/22	4811.22	4950.41	4752.94	4811.22
2000/12/26	4721.36	4950.41	4721.36	4721.36
2000/12/27	4614.63	4950.41	4675.06	4656.63
2000/12/28	4797.14	4950.41	4734.28	4795.03
2000/12/29	4743.94	4950.41	4709.71	4743.94
2000/12/30	4739.09	4950.41	4724.51	4739.09
RMSE		188.67	**79**	**123**

VII. CONCLUSION

In this paper, we have presented two new hybrid forecast models (named ATVF-PSO and ATVF-KM) based on the particle swarm optimization and k-means clustering to improve the adaptive time variant model for fuzzy time series to forecast the Enrolments of the University of Alabama, Taiwan Futures Exchange (TAIFEX) and Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX)). These techniques adjust the length of each interval and universe of discourse adaptive time variant model for fuzzy time series for forecasting, respectively. The experimental results show two methods have better forecasting accuracy than previous ones. We will decide to use multi factor forecasting based on the described scheme in the further research.

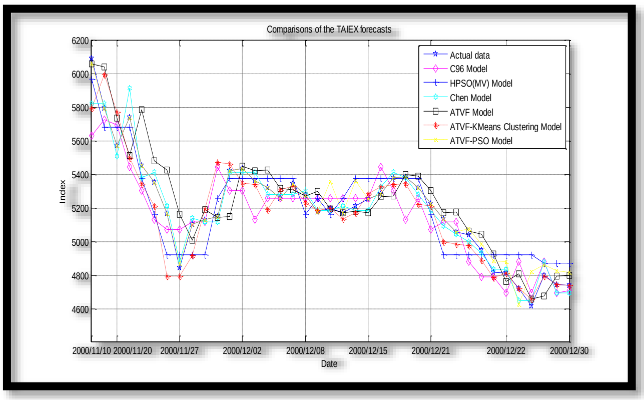

Fig. 3. Comparison of TAIFEX forecasting

Table XX
A comparison of the forecasted results produced based on the number of intervals = 7 and the vote W_h=15

Year	Actual data	C96	[8]	Chen	HPSO	ATVF	ATVF- PSO	ATVF-KM
00/12/28	4797.14	4719.7	4881.25	4733.4	4799.2	4707.89	4785.21	4745.78
00/12/29	4743.94	4707.1	4693.75	4912.1	4786.5	4688.47	4755.87	4727.26
00/12/30	4739.09	4707.1	4693.75	4866.3	4773.5	4751.81	4749.81	4724.83
RMSE		52.91	62.31	119.7	26.34	61.11	**11.54**	**32.25**

Table XXI
A comparison of the forecasted results produced based on the number of intervals = 12 and the vote W_h=15

Year	Actual data	C96	[8]	HPSO	ATVF	ATVF-PSO	ATVF-KM
00/12/28	4797.14	4669.8	4814	4809.5	4676.19	4797.14	4788.8
00/12/29	4743.94	4662.5	4752.3	4726.2	4792.89	4773.73	4758.85
00/12/30	4739.09	4780	4773.1	4762.2	4797.14	4769.51	4760.26
RMSE		83.28	19.72	17.71	82.952	**4.58**	**15.71**

Table XXII
A comparison of the forecasted results produced based on the number of intervals = 15 and the vote W_h=15

Year	Actual data	C96	[8]	Chen	HPSO	ATVF	ATVF-PSO	ATVF-KM
00/12/28	4797.14	4755.9	4760.71	4750.8	4847.3	4667.28	4815.58	4768.57
00/12/29	4743.94	4655.9	4760.71	4721.3	4741	4721.96	4773.13	4743.94
00/12/30	4739.09	4744.1	4760.71	4773.2	4768.1	4764.35	4768.28	4739.1
RMSE		44.8	26.3	34.34	27.39	77.44	**26.11**	**6.49**

Table XXIII
Comparisons Under Differnet Numbers of Intervals

Models	Number of Intervals				
	6	8	10	15	20
Singh	180.52	175.75	160.12	152.23	140.52
ATVF-PSO	130.14	115.42	105.78	96.26	79.5
ATVF-KM	160.87	151.26	142.58	129.38	120.73

REFERENCES

[1] Sepúlveda, R., Montiel-Ross, O., Manzanarez, J. C., & Quiroz, E. E. (2012). Fuzzy logic predictive algorithm for wireless-LAN fast inter-cell handoff. *Engineering Letters*, 20(1), 109-115.

[2] Zadeh, L. A. (1965). Fuzzy sets. Information and control, 8(3), 338-353.

[3] Song, Q., & Chissom, B. S. (1993). Fuzzy time series and its models. *Fuzzy sets and Systems*, 54(3), 269-277.

[4] Chen, S. M. (1996). Forecasting enrollments based on fuzzy time series. *Fuzzy sets and systems*, *81*(3), 311-319.

[5] Tan, J. Y. B., Bong, D. B. L., & Rigit, A. R. H. (2012). Time Series Prediction using Backpropagation Network Optimized by Hybrid K-means-Greedy Algorithm. *Engineering Letters*, *20*(3), 203-210.

[6] Hassan, S., Jaafar, J., Samir, B. B., & Jilani, T. A. (2012). A hybrid fuzzy time series model for forecasting. *Engineering Letters*, *20*(1), 88-93.

[7] Chen, S. M., & Chung, N. Y. (2006). Forecasting enrollments using high-order fuzzy time series and genetic algorithms. *International Journal of Intelligent Systems*, *21*(5), 485-501.

[8] Chen, S. M., & Chung, N. Y. (2006). Forecasting enrollments of students by using fuzzy time series and genetic algorithms. *International journal of information and management sciences*, *17*(3), 1-17.

[9] Li, S. T., Cheng, Y. C., & Lin, S. Y. (2008). A FCM-based deterministic forecasting model for fuzzy time series. *Computers & Mathematics with Applications*, *56*(12), 3052-3063.

[10] Kuo, I., Horng, S. J., Chen, Y. H., Run, R. S., Kao, T. W., Chen, R. J., & Lin, T. L. (2010). Forecasting TAIFEX based on fuzzy time series and particle swarm optimization. *Expert Systems with Applications*, *37*(2), 1494-1502.

[11] Abd-Elaal, A. K., Hefny, H. A., & Abd-Elwahab, A. H. (2013). Forecasting of egypt wheat imports using multivariate fuzzy time series model based on fuzzy clustering. *IAENG International Journal of Computer Science*, 40(4), 230-237.

[12] Khiabani, K., Yaghoobi, M., Mohamadzade Lary, A., & Khani, Y. An Improved Adaptive Time-Variant Model for Fuzzy-Time-Series Forecasting Enrollments based on Particle Swarm Optimization. *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering and Computer Science 2011*, WCECS 2011, 23-25 October, 2011, San Francisco, USA, pp325-330.

[13] Wong, W. K., Bai, E., & Chu, A. W. (2010). Adaptive time-variant models for fuzzy-time-series forecasting. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 40(6), 1531-1542.

[14] Chen, S. M., & Hsu, C. C. (2004). A new method to forecast enrollments using fuzzy time series. *International Journal of Applied Science and Engineering*, *2*(3), 234-244.

[15] Kennedy, J. F., & Eberhart, R. C. particle Swarm optimization (1995) proceedings of IEEE international conference of Neural network, perth.

[16] Kennedy, J. F., Kennedy, J., & Eberhart, R. C. (2001). *Swarm intelligence*. Morgan Kaufmann.

[17] Singh, S. R. (2008). A computational method of forecasting based on fuzzy time series. *Mathematics and Computers in Simulation*, *79*(3), 539-554.

[18] Song, Q., & Chissom, B. S. (1994). Forecasting enrollments with fuzzy time series—part II. *Fuzzy sets and systems*, *62*(1), 1-8.[26] Hwang, J. R., Chen, S. M., & Lee, C. H. (1998). Handling forecasting problems using fuzzy time series. *Fuzzy sets and systems*, 100(1), 217-228.

[19] Hwang, J. R., Chen, S. M., & Lee, C. H. (1998). Handling forecasting problems using fuzzy time series. *Fuzzy sets and systems*, 100(1), 217-228.

[20] MacQueen, J. (1967, June). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* Vol. 1, No. 14, pp. 281-297.

[21] Singh, S. R. (2007). A simple method of forecasting based on fuzzy time series. *Applied Mathematics and Computation*, *186*(1), 330-339.

[22] Duda, R. O., Hart, P. E., & Stork, D. G. (2000). Pattern classification and scene analysis. Part 1, Pattern classification.

[23] Lee, L. W., Wang, L. H., Chen, S. M., & Leu, Y. H. (2006). Handling forecasting problems based on two-factors high-order fuzzy time series. *Fuzzy Systems, IEEE Transactions on*, *14*(3), 468-477.

[24] Kuo, I., Horng, S. J., Kao, T. W., Lin, T. L., Lee, C. L., & Pan, Y. (2009). An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization. *Expert Systems with Applications*, *36*(3), 6108-6117.

[25] Aladag, C. H., Basaran, M. A., Egrioglu, E., Yolcu, U., & Uslu, V. R. (2009). Forecasting in high order fuzzy times series by using neural networks to define fuzzy relations. *Expert Systems with Applications*, *36*(3), 4228-4231.