Journal of Computer Science and Cybernetics, V.35, N.3 (2019), 267–292
DOI 10.15625/1813-9663/35/3/13496

A NEW HYBRID FUZZY TIME SERIES FORECASTING MODEL BASED ON COMBINING FUZZY C-MEANS CLUSTERING AND PARTICLE SWAM OPTIMIZATION

NGHIEM VAN TINH[1,*], NGUYEN CONG DIEU[2]

[1] *Thai Nguyen University of Technology, Thai Nguyen University*

[2] *Thang Long University, Ha Noi, Viet Nam*

[*] *nghiemvantinh@tnut.edu.vn*

Abstract. Fuzzy time series (FTS) model is one of the effective tools that can be used to identify factors in order to solve the complex process and uncertainty. Nowadays, it has been widely used in many forecasting problems. However, establishing effective fuzzy relationships groups, finding proper length of each interval, and building defuzzification rule are three issues that exist in FTS model. Therefore, in this paper, a novel FTS forecasting model based on fuzzy C-means (FCM) clustering and particle swarm optimization (PSO) was developed to enhance the forecasting accuracy. Firstly, the FCM clustering is used to divide the historical data into intervals with different lengths. After generating interval, the historical data is fuzzified into fuzzy sets. Then, fuzzy relationship groups were established according to chronological order of the fuzzy sets on the right-hand side of the fuzzy logical relationships with the aim to serve for calculating the forecasting output. Finally, the proposed model combined with PSO algorithm has been applied to optimize interval lengths in the universe of discourse for achieving the best predictive accuracy. The proposed model is applied to forecast three numerical datasets (enrollments data of the University of Alabama, the Taiwan futures exchange(TAIFEX) data and yearly deaths in car road accidents in Belgium). Computational results indicate that the forecasting accuracy of proposed model is better than that of other existing models for both first - order and high - order fuzzy logical relationship.

Keywords. Enrollments; Forecasting; FTS; Time - Variant Fuzzy Relationship Groups; PSO; FCM.

1. INTRODUCTION

Advance forecasting of events in our daily life like temperature, stock market, population growth, car fatalities, economy growth and crop productions are main scientific issues in the forecasting field. To make a forecast for these kinds of problems with 100% accuracy may not be possible, but obtaining results with the smallest forecasting error is possible. Previously, many classical forecasting models were developed to resolve different problems such as regression analysis, moving average, exponential moving average and ARIMA model. These approaches require having the linearity assumption and needing a large amount of historical data. The FTS forecasting models which were proposed by Song and Chrissom [32, 33] even don't need a limitation of the observations and the linearity assumption either. To forecast the enrollments of the University of Alabama, their approaches apply the max-min operations to handle uncertainty and imprecise data. However, the limitations in their scheme are not convincing to determine the length of intervals and whenever the fuzzy logical relation matrix becomes larger, more amount of computation they face. To overcome those drawbacks and be more accurate in forecasting, the first-order FTS approach suggested by Chen [6] uses simple arithmetic calculations rather than max-min composition operations [32]. Since then, fuzzy time series model is more discovered by many researchers. They presented various improvements from Chen's model [6] in terms of determining the lengths of intervals including the static length of intervals [7, 17, 18, 37, 38] and dynamic length of intervals [3, 4, 9, 14, 22, 26, 27, 35], constructing fuzzy relationship groups [4, 9, 10, 11, 15, 16, 22, 23, 26, 36] and defuzzication process [23, 30, 31, 35]. Specifically, Huarng [16] suggested an effective computational method to determine the appropriate intervals. He stated that the result of forecasting model is greatly influenced by different lengths of intervals in the universe of discourse. Other research works [3, 5, 7, 4, 9, 14, 15, 24, 25] offered different computational approaches in forecasting based on high-order FTS models to defeat the downsides of first-order forecasting models [6, 17]. Singh [31] introduced a new forecasting model for objective of decreasing amount of computations of fuzzy relational matrices or finding out a suitable defuzzification process for prediction enrollments of University of Alabama and crop production.

Recently, many authors have hybridized the intelligent computation with various FTS models to deal with complicated problems in forecasting. For example, Lee et al. [25] reviewed the high order FTS model for forecasting the temperature and the TAIFEX based on genetic algorithm. Furthermore, they also applied simulated annealing technique [24] in determining the length of each interval to achieve better forecasting accuracy. By introducing genetic algorithm(GA) for partitioning intervals in the universe of discourse, Chen & Chung introduced two first-order [4] and high - order forecasting models for forecasting the enrollments of University of Alabama. Moreover, to receive optimal intervals and avoid the harmful results of the mutation operation in GA. Eren Bas et al. [1] proposed a new GA called MGA for forecasting "killed in car accidents" in Belgium and the enrollments in the University of Alabama. At present, the application of PSO in selecting the proper intervals in FTS forecasting model has attracted many attentions of researcher. They demonstrate that suitable selection of intervals by using PSO also increases the performance of forecasting model, as can be seen in the works [5, 11, 16, 22, 23, 28, 39, 40]. Specifically, Kuo et al. proposed a novel forecasting model by hybridizing PSO with FTS model to improve forecasting accuracy. Kuo et al. [23] also based on PSO to suggest a new model for forecasting TAIFEX by proposing new defuzzification rule. Hsu et al. [15] provided a new two-factor high-order model for forecasting temperature and TAIFEX. With the same goal of using PSO in selection of appropriate intervals, Park et al. [28] considered a two-factor high-order FTS model combined with PSO to achieve more appropriate forecasting results. Huang et al. [16] presented the hybrid forecasting model which combined PSO and the refinement in the forecasting output rule for forecasting enrollments . In addition, Dieu N.C & Tinh N.V [11] introduced the time-variant fuzzy relationship groups concept (TV-FRGs) and combined it with PSO in finding optimal intervals to get better forecasting results. Except for this study, the forecasting model [36] is also based on PSO and TV-FRGs, but extended in the two cases of first- order and high- order FRGs to forecast stock market indices of TAIFEX and enrollments. Chen and Bui [8] use the PSO technique not only to bring optimal intervals but also to obtain optimal weight vectors. They proposed the forecasting model which used optimal partition of intervals and optimal weight vectors to predict the TAIFEX and the NTD/USD exchange rates. Cheng et al. [10] produced a FTS model to predict the TAIFEX based on use the PSO for obtaining the appropriate lengths of intervals and the K-means algorithm for partitioning the subscripts of the fuzzy sets into cluster center of each cluster. One another of the methods for determining the optimal intervals can be mentioned as clustering techniques which have been advanced for minimizing error in forecasting. The methods such as Rough Fuzzy C- means [3], automatic clustering [9], fuzzy C-means [13, 39], K-means [34, 35] are introduced in recent works. Some other FTS models use neural network for forecasting oil demand [29] and adaptive neuro-fuzzy inference systems to forecast the daily temperature of Taipei [30].

As already mentioned in researches above, determining the appropriate length of intervals, establishing fuzzy relationships and making the forecasting rules are considered to be challenging tasks and critically influence the accuracy of forecasting model. In spite of significant achievements in using the length of each interval as well as discovering forecasting output rules, these problems still raise attention of researchers. Up to now, there are still rather many ways to determine the length of intervals in the universe of discourse and calculate crisp output values from fuzzified values. Therefore, the objective of this study is to propose a new hybrid forecasting FTS model using high-order TV-FRGs [11], combining FCM clustering with PSO for selecting optimal length of intervals and refinement of forecasting values by new defuzzification rules. To verify effectiveness of the proposed model, three following real-world data sets are used for experimenting: (1) dataset of enrollments at the University of Alabama [6]; (2) Historical data of the TAIFEX [25] in Taipei, Taiwan; and (3) car road accident data in Belgium [1]. The experimental study shows that the performance of proposed model is better than those of any existing models. The remaining content of this paper is organized as follows.

In Section 2, the basic concepts of FTS and algorithms are briefly introduced. Section 3 presents a hybrid FTS forecasting model which combines with the FCM and PSO algorithm. Section 4 makes a comparison of forecasting results of the proposed model with the existing models from three real life data sets. Conclusion and future work are discussed in Section 5.

2. BASIC CONCEPTS OF FTS AND ALGORITHMS

2.1. Basic concepts of FTS

The idea of FTS was first introduced and defined by Song and Chissom [33, 34]. Let $U = \{u_1, u_2, ..., u_n\}$ be an universe of discourse; a fuzzy set A of U can be defined as

$$A = \{f_A(u_1)/u_1 + f_A(u_2)/u_2 + ... + f_A(u_n)/u_n\},$$

where f_A is a membership function of a given set $A : U \to [0, 1]$, $f_A(u_i)$ indicates the grade of membership of u_i in the fuzzy set A. $f_A(u_i) \in [0, 1]$ and $1 \le i \le n$. The basic definitions of FTS are as below.

Definition 1. (Fuzzy time series [32, 33]) Let $Y(t)$, $(t = 0, 1, 2, ...)$ a subset of R, be the universe of discourse on which fuzzy sets $f_i(t)$, $(i = 1, 2, ...)$ are defined and if $F(t)$ is a collection of $f_1(t), f_2(t), \cdots$ then $F(t)$ is called a FTS definition on $Y(t)$, $(t = 0, 1, 2, ...)$.

Definition 2. (Fuzzy logical relationships(FLRs) [32, 33]) The relationship between $F(t)$ and $F(t-1)$ can be presented as $F(t-1) \rightarrow F(t)$. If let $A_i = F(t)$ and $A_j = F(t-1)$, the relationship between $F(t)$ and $F(t-1)$ is represented by FLR $A_i \rightarrow A_j$, where A_i and A_j refer to the left - hand side and the right-hand side of FTS.

Definition 3. (m - order fuzzy logical relationships [33]) Let $F(t)$ be a FTS. If $F(t)$ is caused by $F(t-1), F(t-2), \cdots, F(t-m+1), F(t-m)$ then this fuzzy logical relationship is represented by $F(t-m), \cdots, F(t-2), F(t-1) \rightarrow F(t)$ and is called an m - order FTS.

Definition 4. (Fuzzy relationship groups (FRGs) [6]) The fuzzy logical relationships having the same left- hand side can be further grouped into a FRG. Assume there are exists FLRs as follows: $A_i \rightarrow A_{k1}$, $A_i \rightarrow A_{k2}, \cdots, A_i \rightarrow A_{km}$; these FLRs can be put into the same FRG as $A_i \rightarrow A_{\mathrm{k1}}$, $A_{\mathrm{k2}}, \cdots, A_{\mathrm{km}}$.

Definition 5. (Time-variant fuzzy relationship groups(TV-FRGs) [11]) The fuzzy logical relationship is determined by the relationship $F(t-1) \rightarrow F(t)$. Let $F(t) = A_i(t)$ and $F(t-1) = A_j(t-1)$, the FLR between $F(t-1)$ and $F(t)$ can be denoted as $A_j(t-1) \rightarrow A_i(t)$. Also at time t, we have the following fuzzy logical relationships $A_j(t-1) \rightarrow A_i(t)$; $A_j(t1-1) \rightarrow A_{\mathrm{i1}}(t1)$;...; $A_j(tp-1) \rightarrow A_{\mathrm{ip}}(tp)$ with $t1, t2, .., tp \leq t$. It is noted that $A_i(t1)$ and $A_i(t2)$ are the same fuzzy A_i but appear at different times $t1$ and $t2$, respectively. It means that if these FLRs occur before $A_j(t-1) \rightarrow A_i(t)$, we can group the FLRs having the same left - hand side into a FRG as $A_j(t-1) \rightarrow A_{i1}(t1), A_{i2}(t2), A_{in}(tn), A_i(t)$. It is called first- order TV-FRGs.

2.2. Algorithms

2.2.1. Fuzzy *C*- means clustering

Fuzzy C-Means is a method of clustering proposed by Bezdek [2]. The basic idea of the fuzzy C-means clustering is described as follows. From a raw data set of input vectors $X = \{x_1, x_2, ..., x_n\}$, the FCM employs fuzzy partitioning such that a data object can belong to two or more clusters with different membership grades between 0 and 1. It is based on the minimization of the following objective function

$$J(U, V) = \sum_{i=1}^{C} \sum_{j=1}^{n} u_{ij}^m d_{ij}^2(x_j, v_i), \tag{1}$$

where, m is fuzziness parameter which is a weighting exponent on each fuzzy membership, C is the number of clusters $(2 \leq C \leq n), n$ is the number of objects in the data set X, v_i is the prototype of the center of cluster i, u_{ij} is the grade of membership of x_j belonging to cluster i and $d_{ij}^2(x_j, v_i)$ or d_{ij} is the distance between object x_j and cluster center v_i, U is the membership function matrix, V is the cluster center vector. The FCM focused on minimizing $J(U, V)$, subject to the constrains on U by Eq. (2) as follows

$$u_{ij} \in [0, 1]; \quad \sum_{j=1}^{n} u_{ij} = 1; \quad \sum_{j=1}^{n} u_{ij} \leq n. \tag{2}$$

Algorithmic steps for Fuzzy C-Means clustering is presented as follows

Step 1. Fix the number of clusters C, initialize the cluster center matrix $V(0)$ by using a random generator from the original dataset. Record the cluster centers set $t = 0, m = 2$, and decided by ϵ, where ϵ is a small positive constant (e.g., $\epsilon = 0.0001$).

Step 2. Initialize the membership matrix $U(0)$ by using Eq. (3)

$$u_{ij}(t) = \frac{1}{\sum\limits_{k=1}^{C} \left(\frac{d_{ij}(t)}{d_{kj}(t)}\right)^{\frac{2}{m-1}}}, \tag{3}$$

where $d_{ij} = \|x_j - v_i\|^2$ is the distance between object x_j and cluster center v_i. If $d_{ij}(t) = 0$ then $u_{ij} = 1$ and $u_{rj} = 0$ $(r \neq j)$.

Step 3. Increase $t = t + 1$. Compute the new cluster center matrix V_{ij} using Eq. (4)

$$v_i(t+1) = \frac{\sum\limits_{j=1}^{n} u_{ij}^m(t) \times x_j}{\sum\limits_{j=1}^{n} u_{ij}^m(t)}. \tag{4}$$

Step 4. Compute the new membership matrix U_{ij} by using Eq. (3).

Step 5. If max $\{|u_{ij}(t+1) - u_{ij}(t)|\} \leq \epsilon$ then stop, otherwise go to Step 3 and continue to iterative optimization.

2.2.2. Particle swarm optimization

PSO algorithm is an intelligent optimization algorithm, which was firstly proposed by Eberhart and Kannedy [21] for finding the global optimal solution. In PSO, a set of particles which is called a swarm; each particle indicates a potential solution and always moves through the search space (d-dimensional space) for searching the optimal solution. In the movement process of particles (i.e, N particles), all particles have fitness values to evaluate their performance. Each particle id $(i = 1, \cdots, N)$ has a position vector $X_{id} = [x_{i,1}, x_{i,2}, \cdots, x_{i,d}]$ and a velocity vector $V_{id} = [v_{i,1}, v_{i,2}, \cdots, v_{i,d}]$ to indicate its current state in the search space. The position of the best particle of total number of particles found so far is saved and each particle retains its personal best position which has passed previously. The position X_{id} and the velocity V_{id} are updated by the best position $P_{best\_id} = [p_{id,1}, p_{id,2}, \cdots, p_{id,n}]$ encountered by the particle so far and the best position $G_{best} = \min(P^t_{best\_id})$ found by the whole population of particles according to formulas of velocity and position as follows

$$V_{id}^{t+1} = \omega^t \times V_{id}^t + C_1 \times \text{Rand}() \times (P_{best\_id} - X_{id}^t) + C_2 \times \text{Rand}() \times (G_{best} - X_{id}^t), \tag{5}$$

$$X_{id}^{t+1} = X_{id}^t + V_{id}^{t+1}, \tag{6}$$

$$\omega^t = \omega_{\max} - \frac{t \times (\omega_{\max} - \omega_{\min})}{\text{iter\_max}}. \tag{7}$$

In this paper, we combine the standard PSO [21] with Constrained Particle Swarm Optimization CPSO [12] by using the following Eq. (8) to replace Eq. (5) as follows

$$V_{id}^{t+1} = K \times [\omega^t \times V_{id}^t + C_1 \times \text{Rand}() \times (P_{best\_id} - X_{id}^t) + C_2 \times \text{Rand}() \times (G_{best} - X_{id}^t)], \tag{8}$$

$$K = \frac{2}{|2 - \varphi - \sqrt{(\varphi^2 - 4 \times \varphi)}|}. \quad (9)$$

The new position of the particle id is changed by adding a velocity to the current position as follows

$$X_{id}^{t+1} = X_{id}^{t} + V_{id}^{t+1}, \quad (10)$$

where X_{id}^{t} is the current position of the particle id at time step t; V_{id}^{t} is the velocity of the particle id at time step t, and is limited to [-$V_{\max}, V_{\max}$], where $V_{\max}$ is a constant predefined by user; ω is the time-varying inertia weight, which is the same as the ones presented in [22]; iter_max is the total number of iterations; c_1 and c_2 are two learning factors which control the influence of the cognitive and social components, respectively, $c_1 = c_2 = 2.05$ which are the same as the ones presented in [12], such that $\phi = c_1 + c_2 = 4.1$ and the constriction factor K= 0.7298.

Algorithm 1 briefly summarizes steps of the PSO algorithm for minimizing a fitness function (f) value.

Algorithm 1. A briefly description of the PSO
- Input: Population of N particles, the maximum number of iterations(iter_max)
- Output: G_best value

1. Initialize: Set $K = 0.7298$, $\omega_{\min}$, $\omega_{\max}$, $V_{\max}$
for each particle id, $(1 \le i \le N\ )$ **do**
- Random positions x_{id}, Random velocities v_{id} in d dimensional space
- Set $P_{best\_id}^{i} = x_{id}$;
if $f(P_{best\_id}^{i}) \le f(G_{best})$ **then** $G_{best} = P_{best\_id}^{i}$; end if
end **for**

2. **while** $(t \le$ iter_max$)$ **do**
2.1. **for each** particle id, $(1 \le i \le N\ )$ **do**
• calculate the fitness value of particle id: $f(x_{id})$
- if $f(x_{id}^{t+1}) < f(P_{best\_id}^{t})$ **then** $P_{best\_id}^{t+1} = f(x_{id}^{t+1})$
- if $f(x_{id}^{t+1}) > f(P_{best\_id}^{t})$ **then** $P_{best\_id}^{t+1} = f(P_{best\_id}^{t})$
end **for**
2.2. Update the $f(G_{best})$ position of all particles according to the fitness value.
2.3. **for each** particle id, $(1 \le i \le N)$ **do**
• update the velocity vector using Eq. (8)
• update the position vector using Eq. (10)
end **for**
• Update ω^t according to Eq. (7)
end **while**
return G_{best} value and corresponding position

3. A PROPOSED FTS FORECASTING MODEL BASED ON FCM AND PSO

In this section, a novel FTS forecasting model is suggested by incorporating FCM with PSO to increase forecasting accuracy. The outline of proposed model is presented in Figure 1, which consists of three stages; the first stage is to partition the historical data into intervals based on FCM algorithm in Subsection 3.1, the second stage is to build the FTS forecasting model which is presented details in Subsection 3.2 and uses PSO algorithm for finding optimal lengths of intervals in the third stage which is introduced Subsection 3.3. To handle these stages, all historical enrollments data [6] are utilized for illustrating forecasted process. The three stages of proposed model are described as follows.

3.1. Using FCM algorithm for generating intervals from a raw time series data

In this section, FCM clustering algorithm is applied to classify the collected data into clusters and adjusted these clusters into contiguous intervals. All historical enrollments data [6] from 1971s to 1992s are utilized to present in the stage of generating intervals. The algorithm composed of two main steps is introduced as follows:

Step 1. Apply the FCM clustering algorithm to partition the historical data into C clusters.

For simplicity we partition enrollments dataset into 7 clusters as shown in the second column 2 of Table 1. Similarly, we can change the number of clusters C from 5 to 21.

Step 2. Adjust the clusters into intervals.

In this step, we adjust the clusters into intervals based on cluster centers as follows:

Suppose that V_i and V_{i+1} are adjacent cluster centers and each cluster $Cluster_i$ is assigned as an interval $interval_i$, then the upper bound $Interval\_UB_i$ of $interval_i$ and the lower bound $Interval\_LB_{i+1}$ of $interval_{i+1}$ can be calculated according Eqs. (11) and (12) as below

$$Inteval\_UB_i = \frac{V_i + V_{i+1}}{2}, \tag{11}$$

$$Interval\_LB_{i+1} = Interval\_UB_i, \tag{12}$$

where $i = 1, \cdots, C-1$. Because of lacking intervals before the first interval and lacking intervals after the last interval, the lower bound $Interval\_LB_1$ of the first interval and the upper bound $Interval\_UB_C$ of the last interval can be computed according to Eqs. (13) and (14) as below.

Table 1. The completed result of clusters from the enrollments dataset

STT	Data in cluster	Cluster center (V_i)
1	{13055, 13563}	13309
2	{13867}	13867
3	{14696}	14696
4	{15145, 15163, 15311, 15433, 15460, 15497, 15603 }	15373.14
5	{15861, 16807, 16388, 15984 }	16260
6	{16919, 16859 }	16889
7	{18150, 18970, 19328, 19337, 18876 }	18932.2

$$Interval\_LB_1 = V_1 - (Interval\_UB_1 - V_1), \tag{13}$$

$$Interval\_UB_C = V_C + (V_C - Interval\_LB_C). \tag{14}$$

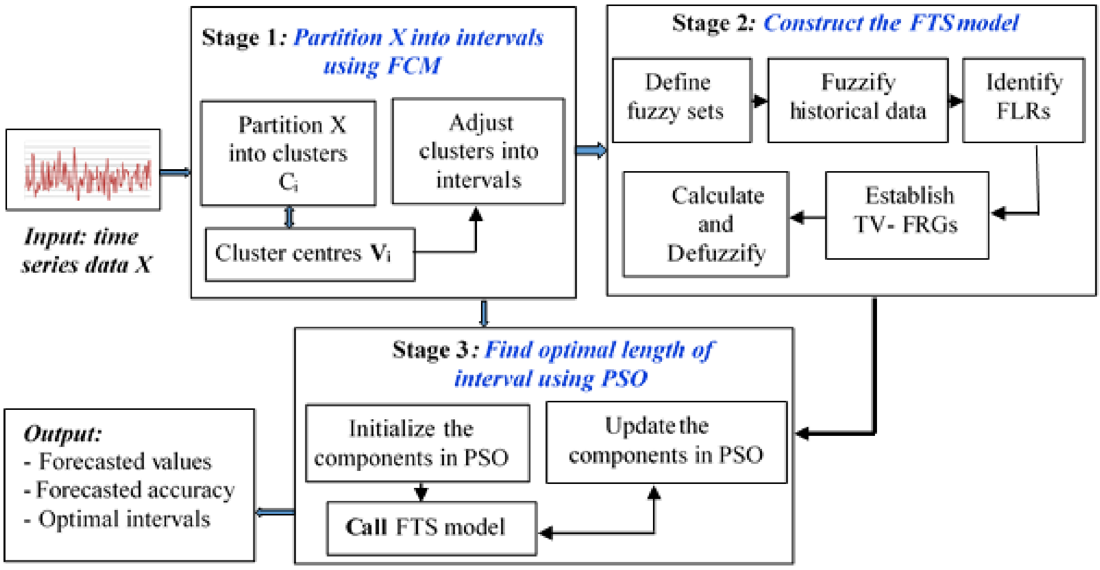

Figure 1. Flowchart of the proposed FTS forecasting model

Compute midpoint value Mid_value_i of the interval $Interval_i$ as follows

$$Mid\_value_i = \frac{Interval\_LB_i + Interval\_UB_i}{2}, \tag{15}$$

where $Interval\_LB_i$ and $Interval\_UB_i$ are the lower bound and the upper bound of the interval $Interval_i$, respectively. Based on the rules in Step 2, we obtain 7 intervals corresponding to the clusters in Step 1, named u_i $(1 \le i \le 7)$ and compute midpoint values of the intervals as listed in Table 2.

Table 2. The completed results of intervals

No	Interval	Mid_value
1	u_1= [13030, 13588)	13309
2	u_2= [13588, 14281.5)	13934.75
3	u_3= [14281.5, 15034.57)	14658.04
4	u_4= [15034.57, 15816.57)	15425.57
5	u_5= [15816.57, 16574.5)	16195.54
6	u_6= [16574.5, 17910.6)	17242.55
7	u_7= [17910.6, 19953.8)	18932.2

3.2. Establish FTS forecasted model based on the first order and high order TV-FRGs

The details of next steps of the forecasting model are established as follows:

Step 3. Determine linguistic terms for each of interval obtained in Step 2.

After creating the intervals in Step 2, linguistic terms are defined for each interval which the historical data is distributed among these intervals. For seven intervals, we get seven linguistic values which are the same as the ones in [6] i.e., $\{A_1, A_2, A_3, A_4, A_5, A_6, A_7\}$ which can be represented by fuzzy sets A_i, as below

$$A_i = \frac{a_{i1}}{u_1} + \frac{a_{i2}}{u_2} + \frac{a_{i3}}{u_3} + ... + \frac{a_{i7}}{u_7}, \tag{16}$$

where $a_{ij} \in [0, 1]$ is the membership grade of u_j belonging to A_i, which is defined by Eq. (17), the symbol '+' denotes the set union operator and the symbol '/' denotes the membership of u_j which belongs to A_i.

$$a_{ij} = \begin{cases} 1 & \text{if } i == j \\ 0.5 & \text{if } j == i-1 \text{ or } j = i+1 \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

From Eq. (16), each fuzzy set contains 7 intervals, and each interval belongs to all fuzzy sets with different grade of membership values presented in Eq. (17)). For instance, u_1 corresponds to linguistic variables A_1 and A_2 with degree of membership values 1 and 0.5 respectively, and remaining fuzzy sets with membership grade 0. The descriptions of remaining intervals, i.e., $u_2, u_3, \cdots, u_7$ can be explained in a similar way.

Step 4. Fuzzify all historical data.

Each of interval contains one or more historical data value of time series. To fuzzy all historical data, the common way is to map historical data into a fuzzy set which has the highest membership value in the interval containing this historical data. For instance, the historical data of year 1973 is 13867, and it belongs to interval u_2 = [13588, 14281.5). So, we allocate the linguistic value A_2 corresponding to interval u_2 to it. According to Eq.(16), the fuzzy set A_2 with the highest membership value occurs at interval u_2. Hence, the fuzzified value for year 1973 is considered as A_2. With a similar explanation for remaining years, we can obtain the results of fuzzification of enrollments data for all years which are shown in Table 3.

Table 3. The results of fuzzification for enrollments data under seven intervals

Year	Actual data	Fuzzy sets	Maximum membership value	Linguistic value
1971	13055	A_1	[1 0.5 0 0 0 0 0]	not many
1972	13563	A_1	[1 0.5 0 0 0 0 0]	not many
1973	13867	A_2	[0.5 1 0.5 0 0 0 0]	not too many
—	—	—	—	—
1991	19337	A_7	[0 0 0 0 0 0.5 1]	too many many
1992	18876	A_7	[0 0 0 0 0 0.5 1]	too many many

Step 5. Create all m^{th}- order fuzzy logical relationships ($m \geq 2$).

The m^{th}- order FLR is constructed based on two or many consecutive fuzzy sets in time series. After transforming historical data into fuzzy sets, then m^{th}- order FLRs can be created based on Definition 3. That means, we need to find any relationship which has the type $F(t-m),\ F(t-m+1), ...,\ F(t-1) \rightarrow F(t)$, where $F(t-m),\ F(t-m+1), \cdots,\ F(t-1)$

and $F(t)$ are called the left-hand side and the right-hand side of FLR, respectively. Then, the m^{th}- order FLR is obtained by substituting the corresponding fuzzy sets as follows: $A_{im},\ A_{i(m-1)}, \cdots,\ A_{i2},\ A_{i1} \rightarrow A_k$. For instance, suppose $m = 1$, we need to point out all first-order FLRs having the form $F(t-1) \rightarrow F(t)$. Based on Table 3, a fuzzy logical relationship $A_1 \rightarrow A_2$ is created by substituting the historical data of $F(1972)$ and $F(1973)$ with fuzzy set as A_1 and A_2, respectively. From this viewpoint, all first-order FLRs of historical time series are shown in column 2 of Table 4. Similarly, we can generate high-order fuzzy logical relationships. Suppose that there is a 2^{nd}- order FLR which is expressed as $F(1972), F(1973) \rightarrow F(1974)$. Based on Table 3, F(1972) = A_1, $F(1973) = A_2$ and $F(1974)$ = A_3 are obtained, then a 2^{nd} FLR $A_1,\ A_2 \rightarrow A_3$ is created by substituting the historical data of $F(1972),\ F(1973)$ and $F(1974)$ to A_1, A_2 and A_3, respectively. By a similar manner, we can establish the 2^{nd} FLRs from the fuzzified data values, which are shown in column 4 of Table 4, where, the symbol # within the last relationship is used to represent the unknown linguistic value.

Table 4. The complete first - order and second - order fuzzy logical relationships

Year	1st-order FLR	1st-order F(t)	2nd-order FLR	2nd-order F(t)
1971	——	——-	——	——
1972	$A_1 \rightarrow A_1$	$F(1971) \rightarrow F(1972)$	——-	——
1973	$A_1 \rightarrow A_1$	$F(1972) \rightarrow F(1973)$	$A_1, A_1 \rightarrow A_2$	$F(1971), F(1972) \rightarrow F(1973)$
—-	——	——	——	——-
1992	$A_7 \rightarrow A_7$	$F(1991) \rightarrow F(1992)$	$A_7, A_7 \rightarrow A_7$	$F(1990), F(1991) \rightarrow F(1992)$
1993	$A_7 \rightarrow \#$	$F(1992) \rightarrow F(1993)$	$A_7, A_7 \rightarrow \#$	$F(1991), F(1992) \rightarrow F(1993)$

Step 6. Generate all m^{th}-order time-variant FRGs.

Each fuzzy relationship group may include one or more fuzzy logic relationships with the same left - hand side. In previous studies, the repeated FLR were simply ignored and it can be only counted one time [7, 6, 22] or the recurrent FLRs are taken into account but were not interested in chronological order [38] when fuzzy relationship groups were established. In this study, we rely on a concept of TV-FRGs [11] and it is mentioned in Definition 5 to create FRGs. In this approach, the TV-FRGs are determined by seeing the history of appearance of the fuzzy sets on the right-hand side of the FLRs. This means, only the fuzzy sets on the right - hand side appearing before the fuzzy sets on the left-hand side of the FLRs at forecasting time is grouped into a FRG. To explain this, two examples are described as below. Firstly, considering the three first -order FLRs at three different time functions, $F(t = 1972, 1973, 1974)$ in Table 4 as follows $F(t = 1972) : A_1 \rightarrow A_1; F(t = 1973) : A_1 \rightarrow A_2$; $F(t = 1974) : A_2 \rightarrow A_3$; where, there are two FLRs at time F(1972) and F(1973) with the same fuzzy set A_1 on the left hand side. If considering at forecasting time $t = 1992$, we obtain a first-order FRG (i.e., $G1$) as follows $A_1 \rightarrow A_1$. If considering at forecasting time $t = 1993$, before that there are two FLRs with the same on left - hand side, these FLRs can be grouped into a FRG as $G2 : A_1 \rightarrow A_1, A_2$. If we consider the forecasting time $t =$ 1994, then the group $G3$ is expressed as follows $A_2 \rightarrow A_3$. The column 3 of Table 5 shows the first-order FRGs, where there are 21 groups in training phase and one group in testing phase. Similarly, the second-order FRGs can be established and listed in column 5 of Table 5 including 20 groups in training phase and one group in testing phase.

Table 5. The complete first - order and second - order TV- FRGs

Year	1st-order FLR	1st-order F(t)	2nd-order FLR	2nd-order F(t)
1971	—-	——-	——	——-
1972	G1	$A_1 \to A_1$	——	——-
1973	G2	$A_1 \to A_1, A_2$	G1	$A_1, A_1 \to A_2$
1974	G3	$A_2 \to A_3$	G2	$A_1, A_2 \to A_3$
—-	—	———	——	———
1992	G21	$A_7 \to A_7, A_7, A_7, A_7$	G20	$A_7, A_7 \to A_7, A_7, A_7$
1993	G22	$A_7 \to \#$	G21	$A_7, A_7 \to \#$

Step 7. Defuzzify and calculate the forecasting output value for all TV-FRGs.

To defuzzify the fuzzified time series values and obtain the crisp output values. First, the new defuzzification rules is developed here to compute the forecasted value for all first - order and high - order time variant FRGs in training phase. Second, we use the master voting (MV) scheme [22] to calculate forecasted value for fuzzy relationship groups with the untrained pattern in testing phase. The forecasting principles is presented as follows:

Principle 1: Using for the first - order TV-FRGs.

For calculating forecasted value based on information of each group, we investigate all information which appear on the right-hand side of each FRG, which is called $Global_{inf}$, then combine with the local information of the same FRG which is presented as follows.

$$Forecasted\_value = 0.5 \times (Global\_inf + Local\_inf), \tag{18}$$

where: - $Global\_inf$ is the global information which can be determined based on all the fuzzy sets on the right-hand side of FRG.

- $Local\_inf$ is the local information which is determined by the fuzzy set appearing at forecasting time on the right-hand side and the latest past in the left - hand side of FRG. Suppose that there is a first - order FRG at forecasting time t is presented as: $A_{t-1} \to A_{t1}, A_{t2}, \cdots, A_{tn}$. Based on research [11], the value of $Global\_inf$ is calculated as follows

$$Global\_inf = \frac{1 \times m_{t1} + 2 \times m_{t2} + \cdots + n \times m_{tn}}{1 + 2 + \cdots + n}, \tag{19}$$

where $m_{t1}, m_{t2}, \cdots, m_{tn}$ are the midpoint values of intervals $u_1, u_2, \cdots, u_n$ with respect to n fuzzy sets existing on the right-hand side of FRG, respectively. By accounting into the variation of latest time on the left-hand side as a forecasting factor, the $Local\_inf$ value is expressed as follows

$$Local\_inf = Lb_{ti} + \frac{Ub_{ti} - Lb_{ti}}{2} \times \frac{m_{ti} - m_{t-1}}{m_{ti} + m_{t-1}}, \tag{20}$$

where A_{t-1} is the lastest fuzzy set on left-hand side of the firstorder FRG; A_{ti} $(1 \le i \le n)$ is the i^{th} fuzzy set in right - hand side of the first - order FRG. Here, m_{t-1} and m_{ti} are middle values of intervals u_{t-1} and u_{ti} with respect to A_{t-1} and A_{ti}. Lb_{ti}, Ub_{ti} denote the lower bound and upper bound value of interval $u_{ti} = [Lb_{ti}, Ub_{ti})$, t is forecasting time with respect to i^{th} fuzzy set on right - hand side of the first - order FRG.

For example, suppose that we want to forecast the enrollment of year 1973. Based on column 3 of Table 5, the first - order FRG (G2: $A_1 \rightarrow A_1,\ A_2$) is formed from two FLRs having next state respectively as $A_1 \rightarrow A_1,\ A_1 \rightarrow A_2$. The highest membership grade of the fuzzy sets A_1 and A_2 appear at intervals u_1 and u_2, respectively, where $u_1 = [Lb_{t1}, Ub_{t1})$ and $u_2 = [Lb_{t2}, Ub_{t2})$. From Table 2, u_1=[13030, 13588) and u_2=[13588, 14281.5). The midpoints of the intervals u_1 and u_2 are $m_{t1} = 13309$ and $m_{t2} = 13934.75$. From Eq. (19), the value of $Global\_inf = \dfrac{m_{t1} + 2 \times m_{t2}}{3} = 13726.2$. Based on Eq. (20), by setting $u_{t-1} = u_1,\ u_t = u_2$, then $Lb_{t2} = 13588,\ Ub_{t2} = 14281.5$ and the value of the $Local\_inf$ on the enrollment of year $t = 1973$ can be calculated as follows

$$Local\_inf = 13588 + \frac{14281.5 - 13588}{2} \times \frac{13934.75 - 13309}{13934.75 + 1330} = 13595.97.$$

From values of $Global\_inf$ and $Local\_inf$ obtained above, based on Eq. (18), the forecasting output value of year 1973 is calculated as $Forecasted\_value = 0.5 \times (13726.2 + 13595.97) =$ 13661.09.

Principle 2: Using the m^{th} order TV-FRGs ($m \geq 2$).

For getting the forecasted results of proposed model based on the high order TV-FRGs, we compute all forecasted values for these groups based on fuzzy sets on the right-hand side within the same group. The viewpoint of this rule is described as follows: For each high - order FRG, we partition each corresponding interval of each linguistic value on the right-hand side into four sub-intervals which have the same length, and compute forecasted output for each group according to Eq. (21).

$$Forecasted\_value = \frac{1}{2 \times n} \sum_{i=1}^{n} (Subm_{ik} + Val\_Lu_{ik}), \tag{21}$$

where n is the sum of fuzzy sets on the right-hand side of FRG; $Subm_{ik}$ is the midpoint value of one of four sub-intervals ($1 \leq k \leq 4$) with respect to i^{th} fuzzy set on the right-hand side of fuzzy relation group, in which the actual data at forecasting time belong to this sub-interval; Val_Lu_{ik} is one of two values belonging to the lower bound and upper bound value of one of four sub-intervals which has the actual data at forecasting time falling within sub-interval u_{ik} (i.e., $u_{ik} = [L_{ik}, U_{ik}]$.

- If the actual data at forecasting time is smaller than middle value of sub-interval u_{ik} Val_Lu_{ik} is assigned by the lower bound of sub-interval u_{ik}.
- If the actual data at forecasting time is larger than middle value of sub-interval u_{ik} Val_Lu_{ik} is assigned by the upper bound of sub-interval u_{ik}.

For instance, assume that we want to forecast the enrollment of year 1973. From column 5 of Table 5, it is seen that the second - order FRG (G1:$A_1, A_1 \rightarrow A_2$) is formed from a FLR with next state A_2 which occurs at year 1973, where the maximum membership grade of A_2 belongs to interval $u_{2.2} = [13588, 14281.5)$. Hence, we partition the interval u_2 into four sub-intervals which are $u_{2.1}$=[13588, 13761.38), $u_{2.2} = [13761.38, 13934.75)$, $u_{2.3} = [13934.75, 14108.13)$ and $u_{2.4}$=[14108.13,14281.5), respectively. The group G1 as $A_1, A_1 \rightarrow A_2$ achieve from relation F(1971), F(1972) $\rightarrow$ F(1973), where the historical data of year 1973 is 13867 and it is within sub-interval $u_{2.2}$ =[13761.38,13934.75) and then the middle value $subm_{2.2}$ of sub-interval $u_{2.2}$ is 13848.06. Then, we find out the value of Val_Lu_{ik} by comparing the historical data of year 1973 with the middle value of sub-interval $u_{2.2}$. From this viewpoint, we obtain the value of Val_Lu_{ik} ($Val\_Lu_{2.2}$) is 13934.75 (the historical data of year 1973 of 13867 is larger than middle value of sub-interval $u_{2.2}$). Finally, forecasted value of year 1973 can be calculated according to Eq. (21) as follows

$$Forecasted\_value = \frac{1}{2}(13848.06 + 13934.75) = 13891.4.$$

Principle 3: Calculate forecasting value in the testing phase.

For testing phase, we calculate forecasted value for a group of fuzzy relationship which has the unidentified linguistic value on the right-hand side based on the master vote scheme [22], and the forecasting value is estimated based on Eq. (22), where the symbol w_h is the highest votes predefined by user for each other problem, m is the order of the FLRs, the symbols $M_{t1}, M_{t2}, \cdots, M_{ti}, \cdots$ are the middle values of the corresponding intervals which are related to the latest fuzzy set and other fuzzy sets on the left-hand side of fuzzy logical relationship group, respectively with the maximum membership values of $A_{t1}, A_{t2}, \cdots, A_{ti}, \cdots$ and u_{tm} occur at intervals $u_{t1}, u_{t2}, \cdots, u_{ti}, \cdots$ and u_{tm}, respectively

$$Forecasted\_value = \frac{m_{t1} \times w_h + m_{t2} + \cdots + m_{ti} + \cdots + m_{tm}}{w_h + (m-1)}. \tag{22}$$

For instance, assume that we want to forecast the enrollment of year 1993 by using first-order fuzzy relationship. As shown in column 3 of Table 5, the group G22 has a first order fuzzy logical relationship as $A_7 \rightarrow \#$ which is created by the fuzzy relationship $F(1992) \rightarrow F(1993)$; since the linguistic value of $F(1993)$ is unknown within the historical data, and this unknown right-hand side state is symbolized by the sign #. Then, the forecasted enrollment of year 1993 is calculated by Eq. (22). Similarly, we can forecast the enrollment of year 1993 by using high-order fuzzy logical relationships. Based on the three forecasted rules above and from Table 3 and Table 5, we complete forecasted results for the enrollments in the period from 1971 to 1992 based on first-order and high order TV-FRGs under seven intervals as shown in Table 6.

Table 6. The complete forecasted output values based on the first order and high - order FTS

Year	Actual data	Fuzzy sets	1st -order forecasted value	2nd-order forecasted value
1971	13055	A_1	Not forecasted	Not forecasted
1972	13563	A_1	13169.5	Not forecasted
1973	13867	A_2	13661.09	13891.4
—	—	—	—	—
1992	18876	A_7	18421.6	19147.62
1993	N/A	N/A	18932.2	18932.2
MSE			**140045.4**	**49873.7**

To verify the forecasting accuracy of proposed model, two evaluation indices are used, the mean square error (MSE) and the root mean square error (MAPE). The formulas of both indices are listed as follows:

$$MSE = \frac{1}{n}\sum_{i=m}^{n}(F_i - R_i)^2, \tag{23}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=m}^{n}(F_i - R_i)^2}, \tag{24}$$

where R_i, F_i denotes actual data and forecasting value at year i, respectively; n is number of the forecasted data; m is order of the fuzzy logical relationships.

3.3. A hybrid FTS forecasting model based on combining the FCM and PSO algorithm

The goal of this subsection is that we present the hybrid FTS forecasting model by combining FCM algorithm for partition data set into the unequal lengths of intervals with Algorithm 1 in Subsection 2.2.2. The main purpose of this algorithm is to adjust the initial intervals length with an intent to obtain the optimal intervals that do not increase the number of intervals in the model. The detailed descriptions of the hybrid forecasting model are given as follows. In proposed model, each particle represents the partitioning of historical time series data into intervals. The number of intervals are determined by FCM (e.g., n intervals). Let the lower bound and upper bound of the universe of discourse U be x_0 and x_n, respectively. Each particle denotes a vector consisting of $n-1$ elements are $x_1, x_2, ..., x_{n-2}$ and x_{n-1}, where $(1 \leq i \leq n-1)$ and $x_i \leq x_{i+1}$. From these $n-1$ elements, define the n intervals as $u_1 = [x_0, x_1], u_2 = [x_1, x_2], \cdots, u_i = [x_{i-1}, x_i], \cdots$ and $u_n = [x_{n-1}, x_n]$, respectively. When a particle moves from one position to another position, the elements of the corresponding new array need to be sorted to ensure that each element x_i arranges in an ascending order such that $x_1 \leq x_2 \leq \cdots \leq x_{n-1}$. In the processing of the training phase, the hybrid forecasting model permits each particle to move from current position to other position by Eqs. (8) and (10), and repeat the steps until the stopping criterion is satisfied. If the stopping criterion is satisfied, then all the FRGs obtained by the global best position (Gbest) among all personal best positions (Pbest) of all particles which used to forecast the new testing data in testing phase. Here, the function MSE (23) is used to evaluate the forecasting accuracy of each particle. The complete steps of the proposed model are presented in Algorithm 2.

Algorithm 2: The FCM-FTS-PSO algorithm
1. **Input**: Historical time series data
2. **Output**: The forecasting results and the MSE value (MSE = Gbest = min(Pbest))
Begin
3. **Select** the initial set of intervals by applying FCM algorithm and use forecasting steps in Subsection 3.2 to get the initial forecasting accuracy (MSE).
4. Initialize: a population of N particles

- The initial position X_{id} of all particles be limited by: x_0+Rand( )$\times(x_n - x_0)$; where, x_0 and x_n are the lower bound and upper bound of the universe of discourse U which is created by FCM; the intervals created by particle 1 are identical to the one created by FCM in Subsection 3.1.

The velocity V_{id} of all particles be exceeded by

$$v_{\min} + \mathrm{Rand}() \times (v_{\max} - v_{\min});\ v_{\min} = -v_{\max}$$

- The initial personal best positions are set as the initial positions of all particles and find Gbest

5. **Repeat**
5.1. **for** particle id, $(1 \leq i \leq N)$ **do**

- Define linguistic terms according to all intervals defined by the current position of particle id based on Step 3 in Subsection 3.2
- Fuzzify all historical data according to the linguistic terms defined above by Step 4 in Subsection 3.2
- Create all m- order fuzzy logical relationships by Step 5 in Subsection 3.2
- Build all m- order time -variant fuzzy relationship groups by Step 6 in Subsection 3.2
- Forecast and defuzzify output values by Step 7 in Subsection 3.2
- Calculate the MSE values for particle id based on Eqs. (23) and (24)
- The new Pbest of particle id is saved according to the MSE values.

end **for**
5.2. The new Gbest of all particles is saved according to the MSE values
6. **for** particle id, $(1 \leq i \leq N)$ **do**

- The particle id is moved to another position according to Eqs. (8) and (10) end **for**
- Change ω according to Eq. (7)
  until (the stop condition (the maximal moving steps or minimum MSE criteria are reached) is true);

End.

4. EXPERIMENTAL RESULTS

4.1. Setup parameters for forecasting problems

In this study, the performance of the proposed model is evaluated based on three different data sets consisting of enrollments data of University of Alabama [6], Taiwan futures exchange dataset (TAIFEX) [25] and vehicle road accidents dataset [1]. These datasets are utilized to illustrate the proposed model's application in one-step-ahead prediction and the forecasting results got from the proposed model are compared to other forecasting models. For implementing the forecasting model on these datasets, we have coded the proposed model by the C sharp programming language on an Intel Core i7 PC with 8GB RAM. In the proposed model we use parameters of PSO, but there are no common principle to determine these parameter values. For ease of comparison with other forecast models using PSO. In the proposed model, we choose the maximum number of iterations (the stop condition of the optimal algorithm) is 150. Like the previous articles [16, 22, 23, 28] the maximum number of iterations have been generally defined intuitively due to the data in most of the applications and is usually set within range from 100 to 500 to achieve the best solution. This has been demonstrated through experimental results in articles such as: the model [22] set number of iterations to 100, the model [23] has number of iterations of 100, and the models in [28] use number of iterations is 500. Therefore, the parameters of PSO used in this research were intuitively determined like in other studies available in the literature. The parameter values of proposed model are determined for each dataset which are listed in Table 7. With the parameters describled in Table 7 the proposed model runs 30 times for each experiment, and takes the best value as the forecasting output value.

(1) The enrollments data of University of Alabama

The enrollments dataset contains 22 observations during the period from 1971 to 1992, see Figure 2(a). This data set has been selected to simulate with the great amount of study works published in the literatures [1, 3, 4, 6, 7, 9, 8, 11, 16, 18, 22, 26, 27, 32, 35]. The results of them will be utilized for comparing with that of the proposed model in this paper.

(2) The TAIFEX time series dataset

The dataset including daily values of the Taiwan futures exchange between August 3, 1998 and September 30, 1998, which has 47 observations is shown in Figure 2(b). This dataset is handled in the literatures [23, 24, 28, 25, 36]. In this study, the historical observations of the TAIFEX between 8/3/1998 and 9/23/1998 are used as the training data set. The last five observations between 9/24/1998 and 9/30/1998 are used as the testing dataset.

(3) The vehicle road accidents dataset in Belgium

The dataset of "killed in car road accidents" consists of 31 observations from 1974 to 2004 that were taken from National Institute of Statistics, Belgium. The plot of yearly deaths in car road accidents is shown in Figure 2(c). This dataset is published in the previuos works [1, 19, 20, 39] .These results are also referred to campare with that of the proposed model in this study.

Table 7. Parameters of the proposed model are setup for forecasting enrollments, TAIFEX and car road accidents

Description for the parameters	Values of enrollments	Values of TAIFEX	Values of car road accident
Number of particles	30	30	30
The max iteration number is set	150	150	150
The inertial weigh limit from	1.4 to 0.4	1.4 to 0.4	1.4 to 0.4
The acceleration coefficient $C_1 = C_2$	2	2	2
The velocity in search range	[-100,100	[-50,50]	[-50, 50]
The position in search range	By FCM	By FCM	By FCM

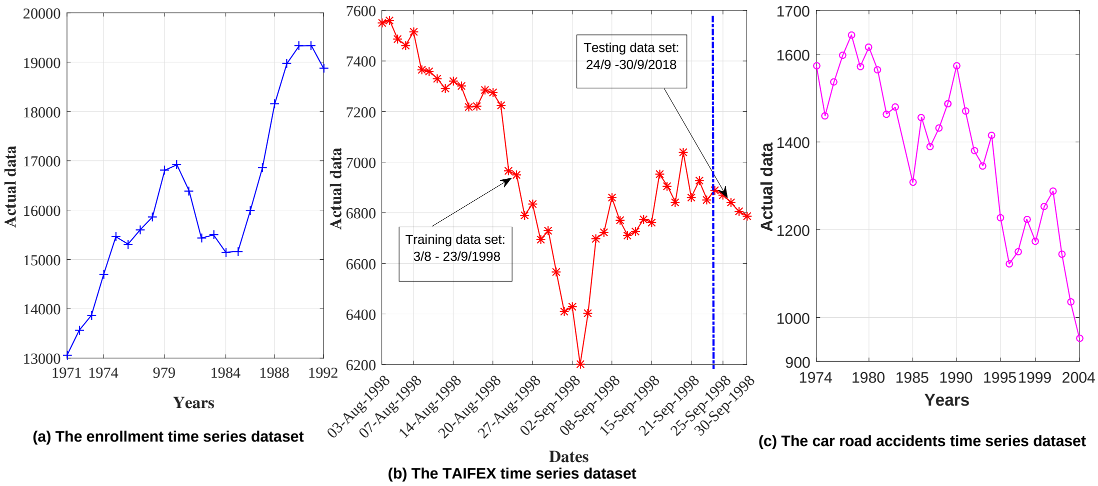

Figure 2. The value of change of historical time series

4.2. Forecasting enrollments of University of Alabama

In this subsection, the proposed forecasting model is applied for forecasting enrollments from yearly observations [6]. To show the performance of the proposed forecasting model based on the first order FTS under different number of intervals, four forecasting models presented in articles [4, 22, 26, 27] are selected for the purpose of comparison. Table 8 shows a comparison of the MSE and RMSE values for different forecasting models. To be easily visualized, Figure 3 depicts the trend of actual data compared to the trend of forecasted value between the proposed model and other models. From this figure, it can be seen that the curve of proposed model is closest to the actual data among five compared models. Based on forecasting results in Table 8, the proposed model gets the smallest MSE value of 4070 and RMSE value of 63.8 among all the compared models with different number of intervals. This can be seen that the proposed model gives the most accurate forecasting results for enrollments of University of Alabama. Differences between the proposed model and models mentioned above accord to the way that the fuzzy relationship group and methods of partitioning the universe of discourse are applied to the structuring process of model. Four forecasting models [4, 22, 26, 27] are constructed based on Chen's model to forecast different problems and perform various methods of interval partitioning such as, the unequal-sized intervals partitioning by using GA algorithm, by using PSO algorithm, the different intervals partitioning based on hedge algebras and intervals partitioning based on interval information granules to improve forecasting accuracy while the proposed model uses an approach that benefits from the concept of time-variant FRG [11] to establish the forecasting model and combine FCM clustering with PSO algorithm for finding optimal interval lengths with an intent to reach better forecasting accuracy.

Next, in order to test the accuracy in the proposed forecasting model according to various number of intervals, five FTS models in papers [4, 11, 16, 22, 36] are referred for comparing in terms of the MSE value . The MSE value is obtained from the proposed forecasting model, as listed in Table 9 is far smaller than that of all the existing forecasting models mentioned

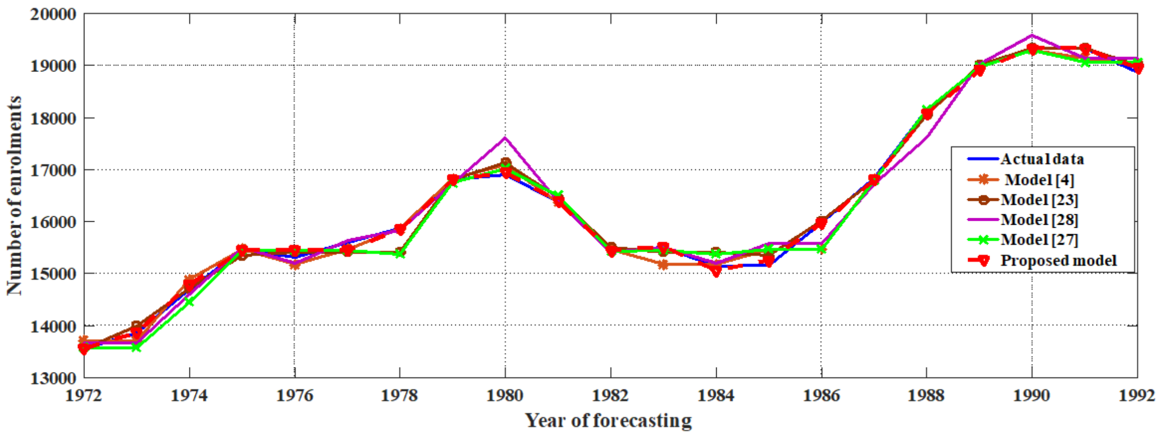

Figure 3. Flowchart of the proposed FTS forecasting model

Table 8. A comparison of the forecasting results between the proposed model and its counterparts based on the first - order FTS using 14 intervals

Year	Actual data	Model [4]	Model [23]	Model[28], h=17	Model [27]	Proposed model
1971	13055	——	——	——	——	——
1972	13563	13714	13555	13678	13582	13558.75
1973	13867	13714	13994	13678	13582	13868
1974	14696	14880	14711	14602	14457	14783.75
——	——	——	——	——	——	——
1990	19328	19300	19340	19574	19297	19325.5
1991	19337	19149	19340	19146	19059	19325.5
1992	18876	19014	19014	19146	19059	18960.835
MSE		35324	22965	65689	46699	**4070**
RMSE		187.9468	151.5421	256.2987	216.1	**63.8**

above based on first-order FLRs for all intervals. In Table 9, all forecasting models use fuzzy relationship group to service for computing the forecasting output values. But three models [4, 16, 22] are designed based on establishing FRGs from Chen's model [6]. The remaining three models such as the model [11], the model [36] and the proposed model all use TV - FRGs. In addition, the proposed model is different from the model [4] in the way that the optimization approaches are utilized. The former employs the PSO, while the latter utilizes the GA for obtaining the proper lengths of intervals, respectively. From Table 9, it is obvious that the optimal performance of the proposed model using PSO is better than the model [4] using GA. This conclusion is also remarked in previous papers. Comparing with four models presented in articles [11, 16, 22, 36], the proposed model is able to generate forecasting values with better accuracy than the three compared models. It can be easily seen that the combination of the FCM algorithm with the PSO in the proposed model yields more optimal interval lengths.
In addition, the forecasting results of the proposed model are also compared with each model

Table 9. A comparison of MSE value between the proposed model and the models [4, 12, 17, 23,37] based on first - order FTS with different number of intervals

forecasting models	Number of intervals						
	8	**9**	**10**	**11**	**12**	**13**	**14**
Model [4]	132963	96244	85486	55742	54248	42497	35324
Model [23]	119962	90527	60722	49257	34709	24687	22965
Model [17]	27435	24860	19698	19040	16995	11589	8224
Model [12]	34457	25855	20533	15625	14630	10004	7475
Model [37]	33983	25841	20322	15472	12588	7078	5396
Proposed model	28681	22076.4	14603	10243.7	8337.6	6096.4	4070

which is introduced in articles [4, 7, 16, 22, 31, 36] based on the various high - order FTS with different number of intervals. A comparison of these models is shown in Table 10, where four models, namely, the model [22], the model [16], the model [36] and the proposed model use 9th-order FLR and number of intervals is 14 for forecasting the enrollments.
Table 10 shows that proposed model bears the lowest MSE value of 5.08 and far exceeds compared to its counterparts. The major difference among all the high - order FTS models mentioned above is that the defuzzification rules is used to forecast output results and optimization technique is handled to get the proper intervals. The different parameters of the model [31] were used as fuzzy relation in forecasting years for calculating output value. Three forecasting models [4, 7, 22] apply Chen's [6] defuzzification rules for computing forecasting value. The model [16] gets the forecasting value by combining the global information of fuzzy logical relationships with the local information of latest fuzzy fluctuation. Meanwhile, the proposed model shows that the forecasting accuracy can be improved by considering more information of sub-intervals within all next states of all fuzzy relationships which has the actual data at forecasting time belonging to these sub-intervals. Among forecasting models above, there are three models using the PSO algorithm as the HPSO model [22], the AFPSO model [16] and the model [36], but the proposed model still obtains far lower MSE value from 9^{th} - order fuzzy logical relationship.

Table 10. A comparison of the results obtained between the proposed model and its counterparts from high - order of the FTS with different number of intervals

Years	Actual data	Model [32]	Model [7]	Model [8]	Model [23]	Model [17]	Model [37]	Proposed model
1971	13055	N/A	N/A	N/A	N/A	N/A	N/A	N/A
—	—	—	—	—	—	—	—	—
1979	16807	16500	16500	16846	N/A	N/A	N/A	N/A
1980	16919	16361	16500	16846	16890	16920	16919	16920
1981	16388	16362	16500	16420	16395	16388	16390	16388
—	—	—	—	—	—	—	—	—
1991	19337	19487	19500	19334	19337	19335	19334	19332
1992	18876	18744	18500	18910	18882	18882	18872	18876
MSE		133700	86694	1101	234	173	9.23	5.08

Table 11. The comparison of the MSE value between the proposed model and its counterparts with various number of orders under 7 intervals

Models	Number orders of FLRs							
	2	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Model [7]	89093	86694	89376	94539	98215	104056	102179	102789
Model [8]	67834	31123	32009	24984	26980	26969	22387	18734
Model [23	67123	31644	23271	23534	23671	20651	17106	17971
Model [17]	19594	31189	20155	20366	22276	18482	14778	15251
Model [12]	19868	31307	23288	23552	23684	20669	17116	17987
Model [37]	8836.2	822.47	686.39	658.18	659.14	618.9	358.43	617.8
Proposed model	8551.81	600.32	447.67	387.12	495.62	370.6	319.86	463.46

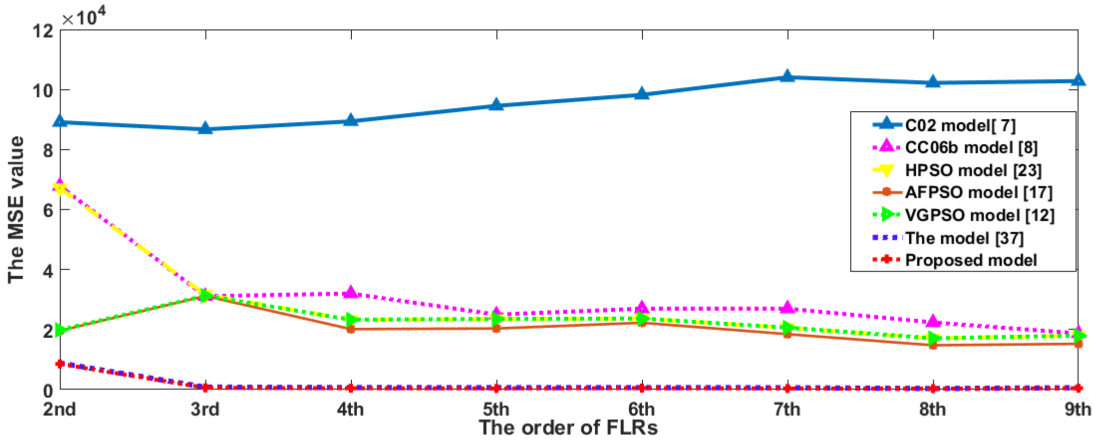

Figure 4. Flowchart of the proposed FTS forecasting model

For more detail, we also perform experiment for each order of proposed forecasting model under seven intervals to compare with the existing models [4, 7, 11, 16, 22, 36], as listed in Table 11. From Table 11, it is obvious that the forecasting error of the proposed model decreases significantly for all orders from 2 to 9. Particularly, the proposed model gets the lowest MSE value of 319.86 with 8^{th}-order fuzzy logical relationship. For easily visualizing, from curves in Figure 4, it can clearly see that the proposed model gives remarkably better forecasting accuracy compared with its counterparts based on high - order FTS. From the above analyses, it can be concluded that the proposed forecasting model outperforms any existing methods for forecasting the enrollments of the University of Alabama.

4.3. Forecasting TAIFEX

In this subsection, the proposed model is applied to forecast the TAIFEX [25] between 8/3/1998 and 9/30/1998. All historical data of TAIFEX are partitioned into two phases to implement comparison results of the proposed model with the existing models based on various orders and different intervals. The performance of the proposed model is evaluated using the MSE (21).

4.3.1. Experimental results in the training phase

In the training phase, the TAIFEX dataset between 8/3/1998 and 9/30/1998 is used and the simulated results of the proposed model are compared with the models as H01 [17], L08 [24], HPSO [22], MTPSO [15], THPSO [28] and NPSO [23]. During implementation of the proposed model, parameters in column 3 of Table 7 don't change and number of intervals are the same with ones of the compared models which is 16 intervals. A comparison of forecasting results in term of MSE are reported in Table 12.

From the experimental results listed in Table 12, it can be seen that the proposed model has the smallest MSE value among the eight compared models for forecasting TAIFEX. Specifically, the proposed model obtains the smallest MSE value of 5.1 among four models [23, 28, 15, 22] also using the PSO technique based on 5th - order FTS with the same number of intervals is 16. Furthermore, from Table 12, it can be concluded that the proposed model has far smaller MSE value than three models in [6, 17, 24] with different number of orders.

4.3.2. Experimental results in the testing phase

In order to verify the performance of forecasting for TAIFEX in the future, historical data of TAIFEX index is split in two parts for independent testing. The first part is used as a training dataset and the second part is used as a testing dataset. From historical data in the past few days, we can forecast the new TAIFEX index for the next day. In this paper the historical data of TAIFEX between March 8, 1998 and September 23, 1998 was used as a training dataset and the remaining data was used in the testing phase. To forecast for testing dataset, the highest votes(w_h) for MV scheme in model [22] are used as 3. Other parameters are taken similar to training set. For instance, for forecasting the new data of date 9/24/1998, the data under days from 8/3/1998 to 9/23/1998 are utilized as the training dataset. Similarly, a new data of date 9/25/1998 can be forecasting based on the data of dates between 8/3/1998 and 9/24/1998. A comparison of results for actual data and the forecasting results between the proposed model and the models [15, 22, 24] which use 16 intervals with the 3rd - order FTS. The results in Table 13 indicate that the proposed model is more precise than four compared models based on 3rd - order FTS and also gets the smallest MSE of 116.37.

4.4. Experimental results for forecasting the vehicle road accidents

In addition, the proposed model is also used for forecasting the vehicle road accidents in Belgium [1] from 1974 to 2004 and there is made a comparison of the forecasting results with the previous works [1, 19, 20, 39]. A comparison of the forecasting results using RMSE (24) is shown in Table 14. It is evident that the proposed method gets better forecasting results than the forecasting models above. More detailedly comparison, for the same number of interval of 13, the proposed model obtains the smallest RMSE value of 1.96 among two models [20, 39] using the 3rd - order FTS. Beside that, the proposed model also has far smaller RMSE value than model [19] and model [39] based on first - order FTS with different number of intervals. To sum up, demonstrations above show that the proposed model outperform the existing models based on both the first- order and high -order FTS model with different number of intervals in forecasting the vehicle road accidents.

Table 12. A comparison of the forecasting results of the proposed method with the existing models based on the high - order FTS under number of intervals = 16

Date	Actual data	H01b	L08	HPSO	MPTSO	THPSO	NPSO	Proposed model
8/3/1998	7552	N/A	N/A	N/A	N/A	N/A	N/A	N/A
8/4/1998	7560	7450	N/A	N/A	N/A	N/A	N/A	N/A
8/5/1998	7487	7450	N/A	N/A	N/A	N/A	N/A	N/A
8/6/1998	7462	7500	N/A	N/A	N/A	N/A	7452.54	N/A
8/7/1998	7515	7500	N/A	N/A	N/A	N/A	7331.62	N/A
8/10/1998	7565	7450	N/A	N/A	N/A	N/A	7285.63	7361.5
8/11/1998	7360	7300	N/A	N/A	N/A	N/A	7331.62	7361.5
8/12/1998	7330	7300	7329	7289.56	7325.28	7325	7291.67	7328.16
8/13/1998	7291	7300	7289.5	7320.77	7287.48	7287.5	7217.15	7290.41
—	—	—	—	—	—	—	—	—
9/29/1998	6806	6850	6796	6800.07	6781.01	6794.3	7331.62	6810.92
9/30/1998	6787	6750	6796	7289.56	6781.01	6794.3	7285.63	6789.25
MSE		5437.58	105.02	103.61	92.17	55.96	35.86	**5.1**

Table 13. A comparison of the MSE value for testing phase based on 3rd-order FTS under 16 intervals using $w_h = 3$.

Date	Actual data	Model [25]	Model [23]	Model [16]	Proposed model
9/24/1998	6890	6959.07	6861.0	6916.62	6886
9/25/1998	6871	6833.52	6897.8	6886.0	6874
9/28/1998	6840	6896.95	6912.8	6892.4	6852
9/29/1998	6806	6863.76	6858.4	6871.54	6825.88
9/30/1998	6787	6823.38	6800.5	6859.12	6791.2
MSE		2815.69	1957.42	2635.23	**116.37**

Table 14. A comparison of the forecasting results between proposed model and various models based on first - order and high - order FLRs

Year	Actual data	Model [20]	Model [21]	Model [1]	Model [40]	Proposed model	
						1st-order	3rd-order
1974	1574	—	—	—	—	—	—
1975	1460	1497	—	1458	—	1445	—
1976	1536	1497	—	1467	—	1548	—
1977	1597	1497	1497	1606	1594	1582	1597
1978	1644	1497	1497	1592	1643	1609	1642
—	—	—	—	—	—	—	—
2003	1035	995	997	1097	1036	1041	1039
2004	953	995	997	929	954	954	950
RMSE		83.12	46.78	37.66	19.2	**16.68**	**1.96**

5. CONCLUSION AND FUTURE WORK

In this study, a new FTS forecasting model which combines FCM and PSO algorithm is proposed for forecasting real-world time series. The advantages of the proposed model are that it combines the PSO and FCM to get the optimum partition of the intervals for increasing the forecasting accuracy rates. The time variant - fuzzy relationship groups were established to overcome the shortcomings of the conventional FTS model which also uses the fuzzy relationship groups. In addition to that the paper also proposes a new defuzzification method for calculating the forecasting output values, which has been the main contribution issue for improving forecasting accuracy of the proposed model. From the empirical study on three datasets of forecasting enrollments, TAIFEX forecasting and car road accidents forecasting, the experimental results show that the proposed model outperforms other existing forecasting models with various orders and different interval lengths. The detail of comparison was presented in Tables 8 - 14 and Figs. 3 - 4.

Even though, the proposed method shows that the superior forecasting capability compared with existing forecasting models, there still remain some aspects which needs to be mentioned, such as the computational complexity when combining many methods in forecasting model and the forecasting of multi-factor problems. To continue evaluating the performance of the forecasting model and overcoming those weaknesses. There are two suggestions for future research as the proposed model need to combine with some more effective optimal techniques to deal with more complicated and multi-factor factors problems for decision-making such as: weather forecasting, monthly inflation, and so on. Moreover, we will study some methods for automatically determining the optimal order of the fuzzy logical relationship for forecasting real-world time series. The main contributions of this paper are summarized as below:

1) The appearance of fuzzy sets on the right - hand side of the fuzzy relationship group is considered in the process of determining the FRGs, which makes a more effective use of the historical data and become more reasonable in reality;

2) The forecasting accuracy of FTS model constructed on basis of unequal-sized intervals that are formed by combining FCM with PSO is prominently improved;

3) The information on the right - hand side of all fuzzy logical relationships are considered to calculate the forecasting output by the new defuzzification technique.

REFERENCES

[1] Bas E, Uslu V.R., Yolcu U, Egrioglu E., "A modified genetic algorithm for forecasting fuzzy time series," *Applied Intelligence,* vol. 41, no. 2, pp. 453–463, 2014.

[2] Bezdek J C., *Pattern Recognition with Fuzzy Objective Function Algorithms.* Plenum, press. 1981.

[3] Bosel M, Mali K., "A novel data partitioning and rule selection technique for modelling high-order fuzzy time series," *Applied Soft Computing,* vol. 67, pp. 87–96, February 2018. [Online]. Available: https://doi.org/10.1016/j.asoc.2017.11.011

[4] Chen S-M, Chung N.-Y., "Forecasting enrollments of students by using fuzzy time series and genetic algorithm," *International Journal of Information and Management Sciences,* vol. 21, no. 5, pp. 485-501, 2006.

[5] Chen S-M, Jian W.-S., "Fuzzy forecasting based on two-factors second-order fuzzy-trend logical relationship groups, similarity measures and PSO techniques," *Information Sciences,* volumes 391-392, pp. 65–79, June 2017.

[6] Chen S M., "Forecasting enrollments based on fuzzy time series," *Fuzzy Sets and Systems,* vol. 81, no. 3, pp. 311–319, August 1996.

[7] Chen S M., "Forecasting enrollments based on high-order fuzzy time series," *Journal Cybernetics and Systems An International Journal,* vol. 33, no. 1, pp. 1–16, 2002.

[8] Chen S M, Phuong H B D., "Fuzzy time series forecasting based on optimal partitions of intervals and optimal weighting vectors," *Knowledge-Based Systems,* vol. 118, pp. 204–216, February 2017.

[9] Luc Tri Tuyen, et al., "A normal-hidden Markov model in forecasting stock index," *Journal of Computer Science and Cybernetics*, vol. 28, no. 3, pp. 206–216, 2012.

[10] Cheng S H, Chen S-M, Jian W S., "Fuzzy time series forecasting based on fuzzy logical relationships and similarity measures," *Information Sciences*, vol. 327, pp. 272–287, 2016.

[11] Dieu N C, Tinh N V., "Fuzzy time series forecasting based on time depending fuzzy relationship groups and particle swarm optimization," *Proceedings of the 9th National Conference on Fundamental and Applied Information Technology Research (FAIR'9),* Can Tho, Viet Nam, 2016, pp. 125–133.

[12] Eberhart R C, Shi Y., "Comparing inertia weights and constriction factors in particle swarm optimization," *Proceedings of the 2000 IEEE Congress on Evolutionary Computation*, La Jolla California U. S. A, 2000, pp. 84–88.

[13] Egrioglu E, Aladag C H, Yolcu, "Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural network," *Expert Systems with Applications*, vol. 40, no. 3, pp. 854–857, 2013.

[14] Egrioglu E, Aladag C H, Yolcu U, Uslu V R, Basaran M A., "Finding an optimal inter-val length in high order fuzzy time series," *Expert Systems with Applications,* vol. 37, no. 7, pp. 5052–5055, 2010.

[15] Hsu L-Y, et al., "Temperature prediction and TAIFEX forecasting based on fuzzy relationships and MTPSO techniques," *Expert Systems with Applications,* vol. 37, no. 4, pp. 2756–2770, 2010.

[16] Huang Y L, et al., "A hybrid forecasting model for enrollments based on aggregated fuzzy time series and particle swarm optimiza-tion," *Expert Systems with Applications,* vol. 38, no. 7, pp. 8014–8023, 2011.

[17] Huarng K., "Effective lengths of intervals to improve forecasting in fuzzy time series," *Fuzzy Sets and Systems*, vol. 123, no. 3, pp. 387–394, 2001.

[18] Hwang J R, Chen S M, Lee C H., "Handling forecasting problems using fuzzy time series," *Fuzzy Sets and Systems,* vol. 100, no. 1–3, pp. 217–228, 1998.

[19] Jilani T A, Burney S. M. A., Ardil C. Multivariate high order FTS forecasting for car road accident. World Acad Sci Eng Technol. vol. 25, pp. 288 – 293, 2008.

[20] Jilani T A, Burney S M A., "Multivariate stochastic fuzzy forecasting models," *Expert Systems with Applications,* vol. 35, no. 3, pp. 691–700, 2008.

[21] Kennedy J, Eberhart R. Particle swarm optimization, in: Proceedings of the IEEE International Conference on Neural Networks,Perth, Australia:pp. vol.4, 1942–1948, 1995, http://dx.doi.org/10.1109/ICNN.

[22] Kuo I-H, et al., "An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization," *Expert Systems with Applications,* vol. 36, no. 3, part 2, pp. 6108–6117, 2009.

[23] Kuo I-H, et al., "Forecasting TAIFEX based on fuzzy time series and particle swarm optimization," *Expert Systems with Applica-tions,* vol. 37, no. 2, pp. 1494–1502, 2010.

[24] Lee L-W, Wang, L.-H., Chen, S.-M., "Temperature prediction and TAIFEX forecasting based on high order fuzzy logical relationhip and genetic simulated annealing techniques," *Expert Systems with Applications,* vol. 34, pp. 328–336, 2008.

[25] Lee L W, Wang L H, Chen S M, Leu Y H., "Handling forecasting problems based on two-factors high-order fuzzy time series," *IEEE Transactions on Fuzzy Systems,* vol. 14, no. 3, pp. 468–477, 2006.

[26] Loc V M, Nghia P T H. Context-aware approach to improve result of forecasting enrollment in fuzzy time series. International Journal of Emerging Technologies in Engineering Research (IJETER) vol.5, no.7, pp.28–33, 2017.

[27] Lu. W, XueyanChen., Xiao-dongLiua. W, JianhuaYang, "Using interval information granules to improve forecasting in fuzzy time series," *International Journal of Approximate Reasoning,* vol. 57, pp. 1–18, 2015.

[28] Park J I, Lee D.J., Song C.K., Chun M.G., "TAIFEX and KOSPI 200 forecasting based on two-factors high-order FTS and particle swarm optimization," *Expert Systems with Applications,* vol. 37, no. 2, pp. 959–967, 2010.

[29] Rubinstein S, Goor A, Rotshtein A., "Time series forecasting of crude oil consumption using neuro-fuzzy inference," *Journal of Industrial and Intelligent Information,* vol. 3, no. 2, June 2015.

[30] Singh P, Borah B., "An effective neural network and fuzzy time series based hybridized model to handle forecasting problems of two factors," *Knowledge and Information Systems,* vol. 38, no. 3, pp. 669–690, March 2014.

[31] Singh S R., "A simple method of forecasting based on fuzzy time series," *Applied Mathematics and Computation,* vol. 186, no. 1, pp. 330–339, 2007.

[32] Song Q, Chissom B S., "Forecasting enrollments with fuzzy time series - Part I," *Fuzzy Sets and Systems,* vol. 54, no. 1, pp. 1–9, 1993.

[33] Song Q, Chissom B S., "Fuzzy time series and its models," *Fuzzy Sets and Systems,* vol. 54, no. 3, pp. 269–277, 1993.

[34] Tian Z H, Wang P, He T Y., "Fuzzy time series based on K-means and particle swarm optimization algorithm," *International Conference on Man-Machine-Environment System Engineering. Man-Machine-Environement System Engineering.* 2017, pp. 181–189. [Online]. Available: https://link.springer.com/chapter/10.1007/978-981-10-2323-1_21

[35] Tinh N V, Dieu N C., "Novel forecasting model based on combining time-variant fuzzy relationship groups and K-means clustering technique," *Proceedings of the 9th National Conference on Fundamental and Applied Information Technology Research(FAIR10),* Can Tho, Viet Nam, 2017. Doi: 10.15625/vap.2017.0002

[36] Nghiem Van Tinh, Nguyen Cong Dieu, "A new hybrid fuzzy time series forecasting model combined the time -variant fuzzy logical relationship groups with particle swam optimization," *Computer Science and Engineering,* vol.7, no.2, pp.52–66, 2017.

[37] Yu H K., "A refined fuzzy time-series model for forecasting," *Physical A: Statistical Mechanics and its Applications,* vol. 346, no. 3–4, pp. 657–681, 2005.

[38] Yu H K., "Weighted fuzzy time series models for TAIEX forecasting," *Physical A: Statistical Mechanics and its Applications,* vol. 349, no. 3–4, pp. 609–624, 2005.

[39] Yusuf S M, Mu'azu M B, Akinsanmi.O., "A novel hybrid fuzzy time series approach with applications to enrollments and car road accident," *International Journal of Computer Applications,* vol. 129, no. 2, pp. 37–44, 2015.

[40] Pham Thi Minh Phuong, Pham Huy Thong, Le Hoang Son, "Theoretical analysis of picture fuzzy clustering," *Journal of Computer Science and Cybernetics,* vol. 34, no. 1, pp. 17–31, 2018.

Received on December 22, 2018
Revised on April 25, 2019