Improved Fuzzy Time Series Forecasting Model Based on Optimal Lengths of Intervals Using Hedge Algebras and Particle Swarm Optimization

Nghiem Van Tinh\textsuperscript{1,}\textsuperscript{*}, Nguyen Cong Dieu\textsuperscript{2}, Nguyen Tien Duy\textsuperscript{1}, Tran Thi Thanh\textsuperscript{1}

\textsuperscript{1}Faculty of Electronics, Thai Nguyen - University of Technology, Thai Nguyen, 250000, Vietnam

\textsuperscript{2}Thang Long University, Hanoi, 100000, Vietnam

ARTICLE INFO

Article history:
Received: 05 December, 2020
Accepted: 12 February, 2021
Online: 28 February, 2021

Keywords:
FTS, FRGs
Hedge algebras
PSO
Enrolments and Car road accidents

ABSTRACT

Recently, numerous scholars have suggested fuzzy time series (FTS) models to forecast many different fields. One of the vital issues for high accurate forecasting in FTS model is method of partitioning in Universe of discourse (UoD). In this research, we propose a novel FTS model, which is established by using hedge algebra (HA) and particle swarm optimization (PSO) for forecasting the different problems. In this model, HA is considered an algebraic structure for partitioning the UoD into unequal - size intervals based on optimal parameters which is determined by PSO. After making the intervals with unequal - length, the data values of times series on each interval are symbolized by fuzzy sets and then, these fuzzy sets are utilized to make fuzzy relation groups. Lastly, we keep using the PSO to adjust the size of each interval with view to reaching the better accurate prediction rate. The effectiveness of the proposed method is demonstrated on two datasets which are often applied in many studies in literature as enrolments data of the University of Alabama and Car road accident data in Belgium. The obtained results show that the proposed model produces higher accuracy forecasting when compared with the some of the recent FTS prediction models for all orders of model.

1. Introduction

The time series forecasting problem is an attractive and vital research issue. This forecasting problem has been often handled by using a variety of methods like mathematical statistics, artificial neural networks, etc. The downsides of the traditional time series forecasting models are that they extensively dependent on historical data or require having the linearity assumption and cannot solve prediction problems in which the values of time series are linguistic terms. To overcome these difficulties, the authors in [1, 2] first produced the concepts of FTS, which have the ability to deal with vague and incomplete data sets by utilizing the fuzzy set theory [3]. They have proposed the two FTS forecasting models to implement on university enrolments of Alabama with a forecasting schema consisting of main five steps: (1) defining UoD, (2) Partitioning of the UoD into intervals, (3) determining the fuzzy sets and fuzzifying the time series, (4) Establishing fuzzy logical relationship, and (5) forecasting and defuzzifying the forecasting values.

However, their approaches take a lot of time to build forecasting model because of using the complex max– min operations in fuzzy relationship matrix and lack of persuasiveness in partitioning the UoD. These limitations led research [3] to develop a new FTS forecasting model using simple arithmetic operations to replace the complex matrix operations [1, 2] in the determination of fuzzy relationship matrix and defuzzification output values. In addition, research works [5, 6] found out the importance of assigning weights to deal with the issue of recurrent fuzzy relationship and to reflect the difference in their importance. Expansion of the framework [3] into a high-order FTS schema [7], and the influence of the length of intervals in article [8] come with the development from the one-factor FTS models into two-factor FTS model [9]. These forecasted approaches are the basis for the strong development of many FTS models in the next time periods. Recently, many authors have applied different advanced data mining techniques in each stage of FTS model with view to enhancing forecasting accuracy. Study in article [10] used the automatic clustering technique for partitioning the UoD into unequal - size intervals at the fuzzification stage in their forecasting model. Some other researches apply soft computing. 

\footnote{Corresponding Author: Nghiem Van Tinh, Email: nghiemvantinh@tnut.edu.vn}
techniques (especially evolutionary computing, clustering techniques) for adjusting and selecting intervals with unequal-size, can be found as genetic algorithm [11, 12], simulated annealing [13], PSO [14-22], K-mean [23, 24], fuzzy C-means [25, 26]. Just recently, a completely different way from fuzzy approach, several works with regards to HA have been published. In [27], the authors have presented a forecasting method based on the theory of hedge algebra [28] for forecasting university enrolments, to be a typical option. In which, the hedge algebra was used to construct linguistic domains and variables instead of performing data fuzzification and defuzzification in the fuzzy approach. In addition, researches in [29, 30] proposed the HA-based forecasting models to obtain unequal-length intervals in the UoD by mapping the semantics of linguistic variables into fuzziness intervals. However, two these research works only focus on building the first-order forecasting model to apply the number of students annually at the University of Alabama. In addition, their forecasting models have not yet applied the optimal techniques, so the obtained forecasting results are not really good enough.

Based on analyzing of the aforementioned research works showed that determining of the length of interval and the order of fuzzy relationships affect strongly forecasting performance of the model. To avoid the above-mentioned limitations and promote the advantage of combination with methods of partitioning in the UoD. The purpose of this study is to suggest a new partition method which uses PSO algorithm to optimize parameters of HA in the FTS prediction model. Therefore, we develop a novel hybrid prediction model using method of fuzzy relation group [17], integrating with HA and PSO algorithm in the identification of optimal intervals with view to enhancing the forecasting performance of the proposed model. For making it become reality, HA has been used to divide the UoD into intervals with unequal-size based on the parameters optimized by PSO. After obtaining the intervals, the time series data is put into the intervals by fuzzy sets and used them to create the FLRs, group of FLRs. Later, all information in FLR groups are utilized to produce the final prediction results based on the our defuzzification principle [31]. Finally, to enhance the accuracy of the model, we continue applying PSO algorithm to readjust the initial interval lengths which are obtained by fuzzy parameters of HA into intervals with the more proper length. Our forecasting model is examined on two following real-world datasets: 1) the historical enrolment dataset of University of Alabama [3], 2) the dataset of car road accident [32]. The examined results point out that our forecasting model outperforms the some of the recent FTS models in terms of prediction accuracy rate.

The next content of this paper introduces brief fundamental theories related to FTS model such as, fuzzy time series, Hedge Algebras and PSO algorithm. A method using PSO technique which has never been applied before in the selecting optimal parameters of HA and optimal length of intervals simultaneously, is presented in Section 3. Section 4 discusses the forecasting performance by comparing the obtained results of the proposed model with ones of the previous models. The last section gives conclusions and directions for future work.

2. The Fundamental Theories and Algorithms

In this section, we briefly introduce general knowledge related to FTS which is proposed in [1, 2] and improved by study [3]. In addition to, the hedge algebras [28] and PSO algorithm [33] is also reviewed.

2.1. Fuzzy time series

The concepts of FTS were proposed in [1, 2], in which the historical time series data are given in the form of fuzzy sets [3].

Assume that \(Y(t_a)\) (\(t_a = \ldots, 0, 1, 2, \ldots\) ) a real subset \(R\) (\(Y(t_a) \subseteq R\)), regarded as the UoD on which the fuzzy sets \(f_i(t_a)\) \(i = 1, 2, \ldots\) are defined. If \(F(t)\) including the collection of \(f_i(t), f_2(t), \ldots\), then \(F(t)\) is namely a FTS which is defined on \(Y(t)\) [1, 2].

If there exists fuzzy logical relationship (FLR) between \(F(t-1)\) and \(F(t)\), namely \((t-1, t)\), such that they can be expressed as \(F(t) = F(t-1) \ast R(t-1, t)\) or \(F(t-1) \rightarrow F(t)\); Where \(R(t-1, t)\) is the first-order fuzzy relationship between \(F(t)\) and \(F(t-1)\) and \(\ast\) represents the max-min composition operator. Here \(F(t)\) and \(F(t-1)\) are fuzzy sets. If, let \(A_1 = F(t)\) and \(A_2 = F(t-1)\), the relationship between \(F(t)\) and \(F(t-1)\) is replaced by \(A_1 \rightarrow A_2\), where \(A_1\) and \(A_2\) are called the current state and the next state of fuzzy relationship, respectively [1, 2, 4].

Let \(F(t)\) be a fuzzy time series. If \(F(t)\) is derived by more fuzzy sets \(F(t-1), F(t-2), \ldots, F(t-\beta + 1), F(t-\beta)\), then fuzzy relationship between them can be represented as \(F(t-\beta), F(t-2), F(t-1) \rightarrow F(t)\). This relationship is called the \(\beta\) order FTS model [1, 2, 7].

Suppose that \(F(t)\) is derived from \(F(t-1)\), then the relationship can be denoted as \(F(t-1) \rightarrow F(t)\). If, let \(A_1 = F(t)\) and \(A_2 = F(t-1)\). The FLR of them can be replaced as \(A_1 \rightarrow A_2\). In addition, at the time \(t\), there are also exist fuzzy relationships as \(A_j(t_1 - 1) \rightarrow A_{j1}(t_1), \ldots, A_n(t - 1) \rightarrow A_{ni}(tn)\) with \(t1, t2, \ldots, tn \leq t\). It is noted that \(A_i(t1), A_i(t2), \ldots, A_i(tn)\) having the same fuzzy set \(A_i\), but look at different times \(t1, t2, \ldots, tn\). If these FLRs appear before \(A_j(t-1) \rightarrow A_j(t)\), they can be grouped into a fuzzy relationship group as follows: \(A_j(t - 1) \rightarrow A_{j1}(t1), A_{j2}(t2), \ldots, A_{jn}(tn), A_i(t)\), and it is called TV-FRG [17].

2.2. Some basis concepts of Hedge Algebras

Hedge Algebras are introduced by N.C. Ho in 1990. It is considered as a new approach to solve forecasting problems in which it is used to quantify the linguistic variables. Each of linguistic variable \(X\) is represented by an algebraic structure, which is built on the inherent semantic order of the linguistic terms [28] and defined as follows:

**Definition 1:** The linguistic variable \(X\) is a set including 5 components \(\mathbb{A}X = (X, G, C, H, \leq)\) and called HA;

In which, \(X\) is the basic set in \(\mathbb{A}X\); \(\leq\) is a natural semantically ordering relation on \(X\); \(G = \{c^-; c^+\}\); \(c^- \leq c^+\) is the set of generating elements (eg., \(\leq H\) : \(C = \{0, w, 1\}\) is a set of constants, with \((0 \leq c^- \leq W \leq c^+ \leq 1)\); \(H = H^- \cup H^+\), with \(H^- = \{h_i : -q \leq i \geq -1\}\) denotes the set of all negative hedges of \(X\), and \(H^+ = \{h_i : 1 \leq i \leq p\}\) denotes the set of all positive hedges;

**Definition 2:** Let \(\mathbb{A}X = (X, G, C, H, \leq)\) be a HA. The function \(fm : X \rightarrow [0, 1]\) is named to be a fuzziness measure of elements in \(X\), if:
1) \( f_m(c^-) + f_m(c^+) = 1 \) and \( \sum_{h \in H} f_m(hx) = f_m(x) \) for all \( x \in X \);
2) \( f(x) = 0 \) with \( \forall x \in X \) so that \( f(0) = f(W) = f(1) = 0 \);
3) For all \( x, y \in X \) and \( \forall h \in H \), \( f_m(hx) = f_m(hy) \), this equation does not depend on \( x \) and \( y \) and it is called fuzziness measure of the hedge \( h \) and namely by \( \mu(h) \). The properties of \( f_m(x) \) and \( \mu(h) \) are given as below:

**Proposition 1.** The \( f_m \) denotes the fuzziness measurement on \( X \), the following statements hold.

With \( x \in X, x = h_yh_{n-1} \ldots h_1c, h_j \in H, c \in G \)

1) \( f_m(hx) = \mu(h) f_m(x) \) for all \( x \in X \)
2) \( \sum_{q < p, i = 0} f_m(h_i c) = f_m(c) \)
3) \( \sum_{q < p, i = 0} f_m(h_i x) = f_m(x) \)
4) \( f_m(x) = f_m(h_yh_{n-1} \ldots h_1c) = \mu(h_n) \mu(h_{n-1}) \ldots \mu(h_1) f_m(c) \)
5) \( \sum_{i=1}^{n} \mu(h_i) = \alpha \) and \( \sum_{i=1}^{n} \mu(h_i) = \beta \), with \( \alpha, \beta > 0 \) and \( \alpha + \beta = 1 \)

\( f_m(h_y c^-) f_m(h_{n-1} c^-) \ldots f_m(h_1 c^-) = f_m(h_y c^+) f_m(h_{n-1} c^+) \ldots f_m(h_1 c^+) \)

**Figure 1:** The order of fuzziness measure of elements \( x \in X, h_j \in H, c \in G \).

### 2.3. Particle Swarm Optimization

PSO is an evolutionary computation algorithm which is introduced by article [33] for searching the global optimum solution. It is developed by work [14] for applying in the forecasting field. Each particle in the swarm represents a potential solution to the global optimization problem. When particles move from this position to other position in q-dimensional space, all particles (i.e. N particles) have fitness values which are estimated according to fuzzy function. In the moving process of particles. The position of the best particle among all particles found so far is saved and each particle maintains its individual best position which has passed previously. Each individual particle \( k_d (1 \leq k_d \leq N) \) is composed of three components: its position \( P_{kd} = [P_{kd,1}, P_{kd,2}, ..., P_{kd,q}] \), the velocity vector \( V_{kd} = [V_{kd,1}, V_{kd,2}, ..., V_{kd,q}] \) and the best position that it has individually found so far \( P_{best, kd} = [P_{kd,1}^{best}, P_{kd,2}^{best}, ..., P_{kd,q}^{best}] \). Then the best position global \( G_{best} = \min(P_{best, kd}) \) found by the overall best out of all the particles in the swarm. The briefly summarizes steps of the standard PSO algorithm in Algorithm 1 as follows:

**Algorithm 1:** The PSO algorithm

**Initialize:** learning factors \( C_i = C_2 ; \omega_{\text{max}}, \omega_{\text{min}} ; \) random positions \( P_{kd,i} ; \) random velocities \( V_{kd,i} \) in q-dimensional space (\( i = 1, 2, ..., q \));

- Positions of each \( k_d \)th (\( k_d = 1, 2, ..., N \)) particle’s position are randomly determined: \( P_{kd} = [P_{kd,1}, P_{kd,2}, ..., P_{kd,q}] \)

where: \( p_{kd,i} \) denotes \( i \)th position of \( k_d \)th particle; \( N \) is the number of particles in a swarm
- Velocities are randomly determined: \( V_{kd} = [V_{kd,1}, V_{kd,2}, ..., V_{kd,q}] \)
- Let \( P_{best, kd} = P_{kd} \)

while \( (t \leq \text{iter}) \) do // iter is maximal iteration number

for each particle \( k_d \) in swarm do

- Calculate the fitness value of particle \( k_d \): \( f(x_{kd}) \)
- Update the personal best position of particle \( k_d \)

\[
P_{t+1}^{best, kd} = \begin{cases} P_{t+1, kd} & \text{if } f(P_{t+1, kd}) > f(P_{t, kd}) \\ P_{t, kd} & \text{otherwise} \end{cases}
\]

End for

- Update the global best position \( G_{best} \) according to the fitness value.

for each particle \( k_d \) in swarm do

- Update the velocity: \( V_{t+1}^{kd} = \omega^t \cdot V_{t, kd} + \alpha \cdot R1 \cdot (P_{t+1}^{best, kd} - P_{kd}) + \beta \cdot R2 \cdot (V_{t, kd} - G_{best}) \)

End for

Update inertia weight \( \omega : \omega^t = \omega_{\text{max}} - \frac{t \cdot (\omega_{\text{max}} - \omega_{\min})}{\text{iter}} \)

End while

### 3. The FTS Proposed Model Using HA and PSO

The aim of this section is to present a new FTS forecasting model based on the advantage of using PSO to get optimal parameters of HA and optimal intervals in the UoD simultaneously. Firstly, PSO is selected in the proposed model to optimize parameters of HA like fuzziness measure of the hedges and fuzziness measure of primary generator for attaining initial intervals in the UoD. Then, we continue to apply PSO algorithm to readjust the initial interval lengths in fuzzy time series obtained by HA into optimal intervals with view to obtaining the better forecasting accuracy rate. Finally, from these optimal obtained intervals, we produce the forecasting results of model by defining fuzzy sets, fuzzy historical data on each divided interval, determining the FLRs, establishing fuzzy relationship groups and calculating the forecasting values from the defuzzification method [31]. The step-by-step of the our model is given as follows.

**Step 1:** Define UoD of historical time series

Assume that \( U = [d_{\text{min}} - n_1, d_{\text{max}} + n_2] \) is UoD. For defining U, the minimal value \( d_{\text{min}} \) and the maximal value \( d_{\text{max}} \) in the time series data is determined; \( n_1 \) and \( n_2 \) are two proper positive numbers, respectively to let the U cover the noise of the testing data. Then, partition UoD into several adjoining intervals based on optimal parameters of HA obtained by PSO algorithm.

**Step 2:** Call the proposed algorithm “Optimizing parameters of HA based on PSO algorithm” to obtain the initial partition of the intervals. This algorithm is introduced in the next part:
and μ(little), in which fm(Low) + fm(High) = 1 and μ(little) + μ(very) = 1. From the optimal parameters obtained, q initial adjoining intervals with different lengths which are defined as: 

\[
u_i = [d_{\text{min}} - n_i, x_i], \quad u_2 = [p_1, p_2], \ldots, \quad u_k = [p_{q-1}, d_{\text{max}} + n_2] ,
\]

respectively.

**Step 3:** From the intervals obtained in Step 2, Call the algorithm “PSO-based optimal lengths of intervals finding algorithm” to achieve the unequal intervals with optimal length. This step applies PSO to adjust the initial intervals which are obtained from HA.

**Step 4:** From the optimal intervals achieved in Step 3, define the fuzzy sets \(A_i\) as follows:

- From q intervals, there are q fuzzy sets to represent various intervals on U.
- Determine the fuzzy sets \(A_i (1 \leq i \leq q)\) as follows:

\[
A_1 = \frac{a_{11}}{u_1} + \frac{a_{12}}{u_2} + \ldots + \frac{a_{1q}}{u_q} \\
A_2 = \frac{a_{21}}{u_1} + \frac{a_{22}}{u_2} + \ldots + \frac{a_{2q}}{u_q} \\
\vdots \\
A_q = \frac{a_{q1}}{u_1} + \frac{a_{q2}}{u_2} + \ldots + \frac{a_{qq}}{u_q}
\]

Where, \(a_{ij} \in [0,1] (1 \leq i \leq q, 1 \leq j \leq q)\), \(u_i\) is the \(i\)th interval of the UoD. The value of \(a_{ij}\) denotes the degree of membership of \(u_i\) in the fuzzy set \(A_i\) which are defined by the triangular membership function with three values 0, 0.5 and 1.

**Step 5:** Fuzzy the historical time series data into fuzzy sets \(A_i\).

Fuzzy time series is generated by converting each historical set data into a fuzzy set. If a time series datum depends on interval \(u_1\) and the maximal membership value of fuzzy set \(A_i\) occurs at \(u_i\), then the historical datum is considered as fuzzy set \(A_i\). In this way, all historical data of time series is fuzzified into \(A_i\).

**Step 6:** Define all \(\beta\)-order fuzzy logical relations (\(\beta \geq 1\)).

The fuzzy logical relationship can be construct by two or several consecutive fuzzified values, respectively. To create an \(\beta\)-order FLR, we need to explore any relationship - type as \(F(t - \beta), F(t - \beta + 1), \ldots, F(t - 1) \rightarrow F(t)\), in which \(F(t), F(t - \beta), F(t - \beta + 1), \ldots, F(t - 1)\) and \(F(t)\) are called the “current state” and the “next state” of the fuzzy logical relationship, respectively. Then, it has been obtained by replacing the corresponding fuzzy sets \(A_i\).

**Step 7:** Establish all \(\beta\)-order fuzzy relationship groups (FRGs)

We use fuzzy relationship group [17] to form FRGs in this study. To clarify this, we consider three first-order FLRs at three different times \(t-2, t-1\) and \(t\) as follows: \(A_1 \rightarrow A_i; A_1 \rightarrow A_k\) and \(A_1 \rightarrow A_j\), respectively. Suppose that we want to forecast the value of time series data of time \(t\), the appearance of the fuzzy sets on the right-hand side of FLRs having the same left-hand side is considered to form into together G1 as \(A_1 \rightarrow A_i, A_k\). The same way, if forecasting time \(t\), the FLRs which have the same right-hand side are grouped into a group G2 as \(A_1 \rightarrow A_j, A_k, A_j\).

**Step 8:** Calculate and defuzzify the forecasted output values

To defuzzify the fuzzified time series data which are based on the established FRGs, we apply two rules which are introduced in papers [31, 14] to calculate the forecasting output values at time \(t\) as follows:

**Rule 1:** Applying for computing output values in the training stage

Our defuzzified principle in article [31] is employed to calculate value based on information of each group. For each group in the training stage, we divide each corresponding interval with regards to the fuzzy sets in the next state of the TV- FRGs into three sub-intervals with equal-length as calculated in (4)

\[
\text{Forecasted output} = \frac{1}{2n} \sum_{i=1}^{n} (sub_{ik} + Value_{lu_{ik}})
\]

where, \(n\) denotes the total number of fuzzy sets on the left-hand side of TV-FRG.

- **sub_{ik}** denotes the medium value of one of three sub-intervals (\(1 \leq k \leq 3\)) with regards to \(i\)-th fuzzy set in the next state of FRG that the real data at forecasting time falls into this sub-interval.

- **Value_{lu_{ik}}** denotes the one of two values belongs to lower or upper bound of one of three sub-intervals which has the real data at forecasting time ranges from \(L_{ik}\) to \(U_{ik}\) of sub-interval. If the real data value at forecasting time minor the mid-point value of sub-interval \(u_{ik}\), then **Value_{lu_{ik}}** is allocated as the lower bound of sub-interval \(u_{ik}\); else **Value_{lu_{ik}}** is allocated as the upper bound of sub-interval \(u_{ik}\).

**Rule 2:** Applying for calculating output value in the testing stage

In the testing stage, prediction value of each group which has the unknown linguistic value on the next state is estimated by the master vote scheme [14]. Assume there a \(\beta\) - order FRG which has type as \(A_{t-\beta}, A_{t-\beta+1}, A_{t} \rightarrow \#\), the prediction value is estimated according to (5) as follows:

\[
\text{Forecasted output} = \frac{(M_{t-1} \cdot w_h) + M_{t-2} + \ldots + M_{t-\beta}}{w_{h+1}}
\]

Where, the symbol \(w_h\) is the highest votes predefined by user; is the order of the FLRs; the symbols \(M_{t-1}, M_{t-2}, \ldots\) and \(M_{t-\beta}\) are the medium values corresponding to intervals in accordance to the latest fuzzy set and other fuzzy sets on the current state of FRG having the maximal membership values of \(A_{t-1}, A_{t-2}, \ldots\) and \(A_{t-\beta}\) occur at intervals \(u_{t-1}, u_{t-2}, \ldots, u_{t-\beta}\), respectively.

**Step 9:** Evaluate the performance of the proposed model

The forecasting performance of the proposed model estimated by two following criterions as: Mean Square Error (MSE) and Mean Absolute Percentage Error (RMSE).

\[
\text{MSE} = \frac{1}{n} \sum_{k=1}^{n} (F_k - R_k)^2
\]

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{k=1}^{n} (F_k - R_k)^2}
\]

Here, \(R_k\) and \(F_k\) are the actual and forecasted value at time \(k\), respectively, \(n\) is number of observations to be forecasted, \(\beta\) is the order of fuzzy relationship.

In the following, we propose a new approach for optimizing parameters of Hedge Algebras, called “Optimizing parameters of HA based on PSO”. This approach is utilized in Step 2 of the proposed model to get the optimal parameters of HA. Details of this approach are shown in Algorithm 2.

www.astesj.com
Algorithm 2: Optimizing parameters of HA based on PSO algorithm

Step 1: Initialize; Generate \( P \) particles in two-dimensional space

- Assume that the fuzziness interval of “Low” term denotes the first dimension and fuzziness intervals of “Little” hedge denotes the second dimension in the construct HA.
- Let \( kd \) be a particle including two elements \( l_{kd,1} \) and \( l_{kd,2} \), represented by the position vector \( L_{kd} = \{ I_{kd,1}, I_{kd,2} \} \); where \( l_{kd,1} \leq 0 \leq 1 \) and \( l_{kd,1} + l_{kd,2} = 1 \). These two elements act as the fuzziness intervals of \( fm(Low) \) and \( \mu(Little) \) of HA, as given in Figure 2(a). The velocity of each particle \( kd \) is denoted by the velocity vector \( V_{kd} = \{ v_{kd,1}, v_{kd,2} \} \) including two elements \( v_{kd,1} \) and \( v_{kd,2} \) as shown in Figure 2(b).

- In PSO, the initial position vector \( L_{kd} \) and the initial velocity \( V_{kd} \) of each particle \( kd \) are generated randomly in range \([0, 1]\), the personal best position vector \( P_{best,kd} = [p_{kd,1}, p_{kd,2}] \) of each particle \( kd \) denoting the best position which has the minimum objective value found so far. Initially, the personal best position vector \( P_{best,kd} \) of each particle \( kd \) like its initial position vector \( L_{kd} \).

![Figure 2: The graphical representation of particle kd](image)

Step 2: while \((t \leq \text{iter})\) do // \text{iter} is a predefined number of iterations

For each particle \( kd \) do the following steps:

Step 2.1: Calculate the objective value \( MSE_{kd} \) of each particle \( kd \), as given by following sub-steps:

- Step 2.1.1: Calculate \( q \) linguistic terms corresponding to \( q \) intervals

  From the position vector \( L_{kd} = \{ l_{kd,1}, l_{kd,2} \} \) or two parameters \( fm(Low) \) and \( \mu(Little) \) of HA, divide the \( U \) into \( q \) intervals as \( u_k = [d_{min} - n_1, p_1] \), \( u_2 = [p_1, p_2] \), \ldots, \( u_k = [p_{q-1}, d_{max} + n_2] \), respectively.

- Step 2.1.2: Based on the obtained intervals, define \( q \) fuzzy sets as \( A_1, A_2, \ldots, A_{q-1} \) and \( A_q \)

- Step 2.1.3: Fuzzy the historical time series data into fuzzy sets \( A_i \)

- Step 2.1.4: Establish fuzzy relationships based on the fuzzy sets \( A_i \) defined and fuzzified data

- Step 2.1.5: Establish all TV-FRGs based on FLRs defined

- Step 2.1.6: Defuzzify and calculate the forecasting output values

- Step 2.1.7: Compute the objective value \( MSE_{kd} \) of each particle based on formula (6)

Step 2.2: Update the private best position vector \( P_{best,kd} \) of each particle \( kd \), if \( MSE'(kd) \leq MSE^{-1}(kd) \); \( p_{kd,1} = l_{kd,1} \) and \( p_{kd,2} = l_{kd,2} \)

Step 2.3: Choose the best particle \( Gbest \) among all \( P \) particles, which has the minimum MSE value, set \( P_{Gbest} = [I_{Gbest,1}, I_{Gbest,2}] \) be the position vector of the Gbest

Step 2.4: Update the velocity \( V_{kd} \) and the position \( L_{kd} \) of each particle \( kd \) according to (1) and (2), respectively; update \( \omega \) in (3).

end for

Step 3: Check the stopping criterion

\[ \text{If} \ (t < \text{iter}) \ \text{then} \ \text{let} \ (t \rightarrow t+1) \ \text{and go to Step 2.1 else, print} \ \text{the results (the position vector} \ \ P_{Gbest} = [I_{Gbest,1}, I_{Gbest,2}] \ \text{be the optimal parameters of HA by letting} \ \ Fm(Low) = I_{Gbest,1} \ \text{and} \ \mu(Little) = I_{Gbest,2}). \]

end while

To sum up, the flowchart of the proposed algorithm “Optimizing parameters of HA based on PSO”, which is shown in Figure 3.

Next, we present algorithm “Finding optimal intervals using PSO” which has been called in Step 3 of the proposed model to find the best length of each interval with view to getting the better accurate forecasting.

Finding optimal intervals using PSO

In this algorithm, PSO is used to adjust the initial interval lengths which are determined by Algorithm 2. The briefly explanations of this algorithm are given as below: Each particle of PSO in q-dimensional space is used to represent the partitioning of time series data, where \( q \) is the number of intervals in the \( UoD \). Assume that the lower bound and upper bound of UoD be \( p_0 = (d_{min} - n_1) \) and \( p_q = (d_{max} + n_2) \), respectively. Each particle represents a vector including \( q \)-elements as \( p_{kd,1}, p_{kd,2}, \ldots, p_{kd,q-2} \) and \( p_{kd,q-1} \), where \( 1 \leq i \leq q - 1 \) and \( p_{kd,1} \leq p_{kd,i+1} \). From \( q \)-1 elements, attain the \( q \) adjoining intervals as \( u_1 = [p_0, p_{kd,1}], u_2 = [p_{kd,1}, p_{kd,2}], \ldots, u_q = [p_{kd,q-1}, p_{kd,q}] \), respectively. If particles in a swarm move to from current position to another, the elements of the new vector with regards to position of particles that need to be adjusted in an ascending order \( (p_{kd,1} \leq p_{kd,2} \leq \cdots \leq p_{kd,q-1}) \). In the training phase, position of each particle is changed by using (1) and (2), and repeated the steps until the repeated value (t) equal to the predefined number of iterations (iter). If (\( t = \text{iter} \)) then all the FRGs obtained by the \( Gbest \) among all personal best positions \( P_{best,kd} \) of all particles which used to forecast the new data in testing phase and presented in Algorithm 4. Here, the MSE function in (6) is used to represent the forecasting accuracy of each particle in the training stage. The steps of the algorithm “Finding optimal intervals using PSO” are shown in Algorithm 3 as follows:

Algorithm 3: Finding optimal intervals using PSO

Input: Historical time series data
Output: Optimal intervals and the MSE value

Initialize:

- \( P \) particles in \( q \)-dimensional space, the maximum iteration (iter)
- The initial position \( p_{kd} \) and the velocity \( v_{kd} \) of all particles, respectively. Where, the intervals in position vector is created by the particle 1 be the same as the one which are created from HA as \( u_1 = [p_{0}, p_{1,1}], u_2 = [p_{1,1}, p_{1,2}], \ldots, u_q = [p_{1,q-1}, p_{1,q}]. \)

- The initial personal best position vectors of the \( kd \)th particle is the same as its initial position vector at the beginning: let \( P_{best,kd} = p_{kd} \)

while \((t < \text{iter})\) do // \text{iter} is maximum iteration number

- for each particle \( kd \) (\( 1 \leq kd \leq P \)) do
Calculate the objective value $MSE_{kd}$ of each particle $kd$ by performing the steps from Step 4 to Step 8 above, such as: defining fuzzy sets, fuzzify time series data, determining all $\beta$-order fuzzy relations, establishing all $\beta$-order TV- FRGs, computing forecasted values.

- **Update** $P_{\text{best},kd}$ value of particle $kd$ by the MSE values

$$P_{\text{best},kd}^{t+1} = \begin{cases} 
P_{\text{best},kd}^t & \text{if } MSE(P_{kd}^{t+1}) > P_{\text{best},kd}^t \\
\frac{MSE(P_{kd}^{t+1})}{MSE(P_{\text{best},kd}^t)} & \text{if } MSE(P_{kd}^{t+1}) \leq P_{\text{best},kd}^t
\end{cases}$$

End for

- **Update** the global best position $G_{\text{best}}$ by the MSE value.

for each particle $kd$ ($1 \leq kd \leq P$) do

- **Update** the velocity: $V_{kd, i}^{t+1} = \omega_t \times V_{kd, i}^t + c_1 \times R1 \times (P_{\text{best}, kd} - P_{kd, i}^t) + c_2 \times R2 \times (G_{\text{best}} - P_{kd, i}^t)$

- **Update** the position: $P_{kd, i}^{t+1} = P_{kd, i}^t + V_{kd, i}^{t+1}$

End for

- **Update** inertia weight $\omega$: $\omega_t = \omega_{\text{max}} - \frac{\omega_{\text{max}} - \omega_{\text{min}}}{\text{iter}}$

End while

**Instance**: Explanation of the optimization process of the proposed model using Algorithm 3 on the enrolments data [4] is given as follows: the number of intervals and particles corresponding to $q = 7$ and $P = 4$, respectively. From the historical enrolments data, we define the UoD as $U = [13000, 20000]$, where lower bound $p_0 = 13000$ and upper bound $p_7 = 20000$, respectively. For finding the optimal solution, the parameters in PSO is defined as: the values of $p_{kd,i}$ fall within the range of $(13000, 20000)$, the values of $v_{kd,i}$ fall within the range of $[-100, 100]$, $(1 \leq i \leq 7, 1 \leq kd \leq 4)$, the values of $C_1$ and $C_2$ be 2, and the $\omega$ value ranges from 0.9 to 0.4 and maximum number of iterations be 2, respectively. The positions and velocities of all particles are initialized randomly and listed in Tables 1 and 2, respectively.

**Algorithm 4**: The forecasting algorithm in the testing stage

The optimal lengths of intervals and order of FLRs obtained in Algorithm 3 that are used to estimate untrained data in the testing stage based on the Principle 2 in the forecasting model.

In Table 1, we have given the 7 intervals for 4 particles which are $u_1 = [p_0, p_1], u_2 = [p_1, p_2], ..., u_7 = [p_6, p_7]$, respectively. Where, the initial position of particle 1 act as the intervals are created as the same the one which are obtained from HA in Algorithm 2 and listed as $u_1 = [13000, 14470.47), u_2 = [14470.47, 15149.8), u_3 = [15149.8, 15829.15), u_4 = [15829.15, 16143), u_5 = [16143, 16528.14), u_6 = [16528.14, 17361.8), u_7 = [17361.8, 20000)$. The MSE value of particles is considered according to equation (6). From the corresponding MSE values, every particle records its own personal best positions ($P_{\text{best}}$) so far.

Figure 3: Flow chart of the proposed algorithm “Optimizing parameters of HA based on PSO”

www.astesj.com
Firstly, the $P_{\text{best}}$ values are initialized to the same as the initial position of all particles. Table 3 presents the $P_{\text{best}}$ values of all particles so far and the global best position $\text{Gbest} = \min (P_{\text{best}})$ which is particle 1. After the first iteration, all particles change its positions based on (1) and (2). The second positions and the corresponding new MSE values of all particles are presented in Table 4.

Comparing the MSE values in Table 3 with the MSE values in Table 4, it can be seen that particle 2 and particle 3 in Table 4 attained a better position than their own $P_{\text{best}}$ values so far. Thus, the two particles are updated in Table 5. The new Gbest is obtained by particle 2, because of the its smallest MSE value. The proposed model is accomplished by repeating the steps in Algorithm 3 until the maximal number of iterations is reached. Finally, the proper lengths of intervals are achieved corresponding to Gbest value that the particle 2 gets so far. These obtained intervals are employed to forecast the final output results.

## 4. Experiments and analysis

In this paper, our forecasting model has been implemented on two datasets as enrolments data of University of Alabama [3] and number of deaths in car road accidents in Belgium [32]. These two datasets have been applied for forecasting with the huge amount of research works in the literature. Before implementing the proposed forecasting model, two time series datasets are briefly described. Then, the simulated results and analyses related to these datasets are given, respectively. Description of time series data and evaluation of the proposed model are discussed as follows.

### 4.1. Prepare data for experiments

#### 4.1.1. Time series description

This study, we focus on two time series datasets which are often used to demonstrate validity and performance of the FTS forecasting model. The statistical characteristics of two these time series are expressed as follows.

**(a) The enrolments dataset of University of Alabama**

This time series data consists of 22 values between 1971 and 1992, see Figure 4(a). This dataset has utilized to examined with the huge amount of research works which are presented in the articles [1, 2, 4 - 7, 10 -12, 14, 15, 17, 21 - 24, 29 - 31]. The obtained results among these works are choosed for comparing with our proposed model. Some of results among these studies are considered for comparing with that of the proposed model in this paper. The UoD of enrolments time series is determined as $\text{UoD} = [d_{\text{min}} - n_1, d_{\text{max}} + n_2] = [13000, 20000]$. 

<table>
<thead>
<tr>
<th>P</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
<th>$P_4$</th>
<th>$P_5$</th>
<th>$P_6$</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14470.47</td>
<td>15149.8</td>
<td>15829.15</td>
<td>16143</td>
<td>16528.14</td>
<td>17361.8</td>
<td>103486.13</td>
</tr>
<tr>
<td>2</td>
<td>14188.81</td>
<td>14232.49</td>
<td>15853.02</td>
<td>16448.05</td>
<td>16905.66</td>
<td>18224.27</td>
<td>183494.58</td>
</tr>
<tr>
<td>3</td>
<td>13951.02</td>
<td>14390.22</td>
<td>15579.73</td>
<td>16510.86</td>
<td>17408.76</td>
<td>18386.16</td>
<td>145004.74</td>
</tr>
<tr>
<td>4</td>
<td>16098.76</td>
<td>17518.15</td>
<td>18332.49</td>
<td>18779.17</td>
<td>18995.24</td>
<td>19826.96</td>
<td>558449.1</td>
</tr>
</tbody>
</table>

Table 1: The randomized initial positions of all particles

<table>
<thead>
<tr>
<th>P</th>
<th>$v_1$</th>
<th>$v_2$</th>
<th>$v_3$</th>
<th>$v_4$</th>
<th>$v_5$</th>
<th>$v_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>32.16</td>
<td>-60.62</td>
<td>97.31</td>
<td>86.1</td>
<td>-51.3</td>
<td>-26.34</td>
</tr>
<tr>
<td>2</td>
<td>-81.93</td>
<td>88.23</td>
<td>-94.65</td>
<td>54.95</td>
<td>-1.92</td>
<td>26.4</td>
</tr>
<tr>
<td>3</td>
<td>-56.46</td>
<td>-12.85</td>
<td>33.18</td>
<td>74.45</td>
<td>94.08</td>
<td>-27.5</td>
</tr>
<tr>
<td>4</td>
<td>-71.06</td>
<td>-70.14</td>
<td>85.8</td>
<td>-92.89</td>
<td>-56.09</td>
<td>-53.16</td>
</tr>
</tbody>
</table>

Table 2: Randomly generated initial velocities of all particles

<table>
<thead>
<tr>
<th>P</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
<th>$P_4$</th>
<th>$P_5$</th>
<th>$P_6$</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14470.47</td>
<td>15149.8</td>
<td>15829.15</td>
<td>16143</td>
<td>16528.14</td>
<td>17361.8</td>
<td>103486.13</td>
</tr>
<tr>
<td>2</td>
<td>14188.81</td>
<td>14232.49</td>
<td>15853.02</td>
<td>16448.05</td>
<td>16905.66</td>
<td>18224.27</td>
<td>183494.58</td>
</tr>
<tr>
<td>3</td>
<td>13951.02</td>
<td>14390.22</td>
<td>15579.73</td>
<td>16510.86</td>
<td>17408.76</td>
<td>18386.16</td>
<td>145004.74</td>
</tr>
<tr>
<td>4</td>
<td>16098.76</td>
<td>17518.15</td>
<td>18332.49</td>
<td>18779.17</td>
<td>18995.24</td>
<td>19826.96</td>
<td>558449.1</td>
</tr>
</tbody>
</table>

Table 3: The initial Pbest of all particles; the Gbest value is created by particle 1.

<table>
<thead>
<tr>
<th>P</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
<th>$P_4$</th>
<th>$P_5$</th>
<th>$P_6$</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14499.41</td>
<td>15095.24</td>
<td>15916.73</td>
<td>16220.49</td>
<td>16448.05</td>
<td>16905.66</td>
<td>18224.27</td>
</tr>
<tr>
<td>2</td>
<td>14288.81</td>
<td>14332.49</td>
<td>15762.27</td>
<td>16348.05</td>
<td>16805.66</td>
<td>18124.27</td>
<td>99566.34</td>
</tr>
<tr>
<td>3</td>
<td>14051.02</td>
<td>14490.22</td>
<td>15679.73</td>
<td>16410.86</td>
<td>17308.76</td>
<td>18286.16</td>
<td>135650.18</td>
</tr>
<tr>
<td>4</td>
<td>15998.76</td>
<td>17418.15</td>
<td>18332.49</td>
<td>18779.17</td>
<td>18995.24</td>
<td>19726.96</td>
<td>600250.48</td>
</tr>
</tbody>
</table>

Table 4: The second positions of all particles

<table>
<thead>
<tr>
<th>P</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
<th>$P_4$</th>
<th>$P_5$</th>
<th>$P_6$</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14470.47</td>
<td>15149.8</td>
<td>15829.15</td>
<td>16143</td>
<td>16528.14</td>
<td>17361.8</td>
<td>103486.13</td>
</tr>
<tr>
<td>2</td>
<td>14188.81</td>
<td>14232.49</td>
<td>15853.02</td>
<td>16448.05</td>
<td>16905.66</td>
<td>18224.27</td>
<td>183494.58</td>
</tr>
<tr>
<td>3</td>
<td>13951.02</td>
<td>14390.22</td>
<td>15579.73</td>
<td>16510.86</td>
<td>17408.76</td>
<td>18386.16</td>
<td>145004.74</td>
</tr>
<tr>
<td>4</td>
<td>16098.76</td>
<td>17518.15</td>
<td>18332.49</td>
<td>18779.17</td>
<td>18995.24</td>
<td>19826.96</td>
<td>558449.1</td>
</tr>
</tbody>
</table>

Table 5: The second Pbest of all particles; the Gbest value is obtained by particle 2
In which, the minimal value and the maximal are $d_{\text{min}}=13055$ and $d_{\text{max}}=19337$, respectively, and two proper positive values $n_1$ and $n_2$ are set as 55 and 663, respectively.

(b) The dataset of car road accidents in Belgium

There are 31 observations about the car road accidents ranges from 1974 to 2004 that taken from National Institute of Statistics, Belgium. Figure 4(b) depicts the yearly deaths in car road accidents in Belgium. This time series data is investigated in the research works [6, 32, 40 - 42]. The obtained results from these works have been also selected to compare with our proposed model. In this time series, the minimal value and the maximal are $d_{\text{min}}=953$ and $d_{\text{max}}=1644$, and two proper positive values $n_1$ and $n_2$ are set as 3 and 6, respectively. Therefore, the UoD can be defined as $U = [d_{\text{min}} - n_1, d_{\text{max}} + n_2] = [950, 1650]$.

Table 6: The parameters of PSO are applied to the enrolments data and car road accidents data

<table>
<thead>
<tr>
<th>The parameters in PSO</th>
<th>Enrolments</th>
<th>Car road accidents</th>
</tr>
</thead>
<tbody>
<tr>
<td>The number of particles $N$</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>The max number of iterations $iter$</td>
<td>150</td>
<td>150</td>
</tr>
<tr>
<td>The inertial weight $\omega$ is decreased by $\omega_{\text{max}}=0.9$ to $\omega_{\text{min}}=0.4$</td>
<td>$\omega_{\text{max}}=0.9$ to $\omega_{\text{min}}=0.4$</td>
<td></td>
</tr>
<tr>
<td>The coefficient $C1 = C2$</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>The positions $P$ in search space $[13000, 20000]$</td>
<td>$[950, 1650]$</td>
<td></td>
</tr>
<tr>
<td>The velocities $V$ in search space [-100, 100]</td>
<td>[-100, 100]</td>
<td></td>
</tr>
<tr>
<td>Number of intervals $q$ Defined by HA</td>
<td>Defined by HA</td>
<td></td>
</tr>
<tr>
<td>The number of particles $N$</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

4.1.2. Setup parameters for experiments

For implementing the experiments, we use C# programming tool on an Intel Core i7 PC with 8GB RAM. From parameters of each time series data in Table 6, our forecasting model is tested 30 independent runs on each of dataset with various number of orders and intervals to make forecasting results. Then, the best result of among testing runs is recorded to compare with most well-known models in the same dataset with regards to MSE (6) and RMSE (7) functions. The selected parameters of the PSO are used in experiments for receiving optimal intervals and final forecasting results are placed in Table 6.

4.2. Application of forecasting and comparing results

In this section, we give out results of two experiments with regards to real-world time series datasets which are described in Section 4.1. Then, comparison of results between the proposed model and well-known FTS models in the literature are also presented.

4.2.1. Applying for Experiment 1

Case (1): Forecasted results obtained by the first – order FTS

The forecasting results obtained from this experiment are compared with the ones of the current models [34-37, 27, 29] under the same number of intervals equal to 7. A comparison with regards to RMSE value between the proposed model and the different forecasting models are given in Table 7. Considering the Table 7, the results show that the proposed model has the smallest forecasting errors with regards to RMSE value equal to 188.8 among all its counterparts. There are significant differences between the proposed model and the compared models above. It is the way which determining of the fuzzy relationship group and method of partitioning the UoD are applied in the forecasting model. Three models in works [34-36] are constructed according to the framework [3] to forecast different problems and apply information granules for partitioning, respectively, whereas the proposed model uses hedge algebras for determining unequal-sized interval lengths. Comparing with two models in articles [27, 29]. These models apply the fuzzy relation groups [3] to structure the forecasting model, in which the proposed model uses the fuzzy relation groups that we have proposed in article [17] to build the forecasting model. Comparing the model [37], the proposed model employs the HA combining with PSO to select the optimal intervals, whereas the model [37] applies the maximum spanning tree based fuzzy clustering for dividing intervals with different lengths in the intuitionistic FTS model. In addition, the proposed model is also given to compare with other models which are presented in [6, 11, 14, 15, 17, 36, 38] under the number of intervals of 14. The forecasting results and MSE values between our model and other models are given in Table 8. Table 8 shows that our model has capable of more accurate forecasting and obtains the MSE value 5938.8 which is the smallest among all the existing models.
To authenticate the superiority of our prediction model based on the different high-order FLRs, the research works [7, 12, 14, 15, 17] are cited for comparing. The comparative results for all forecasting models under the number of intervals equal to 7 are shown in Table 9. From Table 9 shows that the proposed model outperforms in term of forecasting accuracy the other existing models under different high-order fuzzy relationships at all. In particular, our model has the smallest average MSE value of 1608.26 among all of compared models. Among all fuzzy relationships is done in the model, the proposed model obtains the lowest MSE value equal to 111.6 by 6th-order fuzzy relationships. The major difference between our model and the compared models is approach of forming FRGs and optimization method they used. In optimization method, the model [12] performs genetic algorithm but the models in articles [14, 15, 17] and the proposed model proceed the PSO algorithm to achieve the best intervals, respectively. Also using PSO to find suitable intervals, our model incorporates HA to partition the different initial intervals of the UoD instead of equal length intervals. In the determining of FRGs, our model is constructed from model [17], the remaining models in articles [7, 12, 14, 15] are created from structure [3]. From the above analysis, it is clearly seen that our model provides more convincing forecasted results when compared to five models considered above.

In addition, our proposed model has been also applied to compare with other existing models based on the different high-order FTS under number of intervals equal to 14. These compared models which are presented in papers [7, 39, 12, 14, 15]. The comparative results of the proposed model with its counterparts are placed in Table 10. Comparing model [39] with the proposed model, the proposed model provides the better MSE value. In addition, comparing the forecasting model in article [7] and the proposed model, both of them use the 5th-order fuzzy relationship but our model is much more superior in term of forecasting accuracy. When compared with remaining forecasting models in articles [12, 14, 15]. Although these models use the fuzzy logical relationship with number of orders is larger, but the results obtained from our model are also better than the existing competing models. In particular, from Table 10, our model obtains the forecasting error MSE of 16.9 which is the lowest among five compared models above. This can conclude that the proposed model not only provides superior forecasting results but also shows the best stability based on the various high-order FLRs, for all cases. To be clearly imagined, Figure 5 describes the trend in term of forecasting accuracy between our model and the previous models for different orders. Viewing these curves, it is clearly seen that forecasting accuracy of our model is more accurate than those of compared models under dissimilar high-order FLRs at all. To sum up, the comparisons above is enough to demonstrate the effectiveness of our model which outpace the previous models based on high-order model with unlike number of intervals in the forecasting the enrolments of University of Alabama.
Table 9: The results of the our model and the compared models with 7 intervals

<table>
<thead>
<tr>
<th>Orders</th>
<th>[7]</th>
<th>[12]</th>
<th>[14]</th>
<th>[15]</th>
<th>[17]</th>
<th>Proposed model</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>89093</td>
<td>67834</td>
<td>67123</td>
<td>19594</td>
<td>19868</td>
<td>10824.85</td>
</tr>
<tr>
<td>3</td>
<td>86694</td>
<td>31123</td>
<td>31644</td>
<td>31189</td>
<td>31307</td>
<td>533.2</td>
</tr>
<tr>
<td>4</td>
<td>89376</td>
<td>32009</td>
<td>23271</td>
<td>20155</td>
<td>23288</td>
<td>441.36</td>
</tr>
<tr>
<td>5</td>
<td>94539</td>
<td>24984</td>
<td>23534</td>
<td>20366</td>
<td>23552</td>
<td>289.6</td>
</tr>
<tr>
<td>6</td>
<td>98215</td>
<td>26980</td>
<td>23671</td>
<td>22276</td>
<td>23684</td>
<td>111.6</td>
</tr>
<tr>
<td>7</td>
<td>102179</td>
<td>22387</td>
<td>17106</td>
<td>14778</td>
<td>17116</td>
<td>251.6</td>
</tr>
<tr>
<td>8</td>
<td>102789</td>
<td>18734</td>
<td>17971</td>
<td>15251</td>
<td>17987</td>
<td>270.4</td>
</tr>
<tr>
<td>9</td>
<td>104056</td>
<td>26969</td>
<td>20651</td>
<td>18482</td>
<td>20669</td>
<td>143.46</td>
</tr>
<tr>
<td>Average MSE</td>
<td>95867.63</td>
<td>31377.5</td>
<td>28121.38</td>
<td>20261.38</td>
<td>22183</td>
<td>1608.26</td>
</tr>
</tbody>
</table>

Table 10: The obtained results of our model and the compared models with 14 intervals

<table>
<thead>
<tr>
<th>Years</th>
<th>Real data</th>
<th>[7]</th>
<th>[39]</th>
<th>[12]</th>
<th>[14]</th>
<th>[15]</th>
<th>Proposed model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971</td>
<td>13055</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1972</td>
<td>13563</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1973</td>
<td>13867</td>
<td></td>
<td>14934.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1974</td>
<td>14696</td>
<td></td>
<td>15590</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1975</td>
<td>15460</td>
<td></td>
<td>15422.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1976</td>
<td>15311</td>
<td>15500</td>
<td>15603</td>
<td></td>
<td></td>
<td></td>
<td>15311.38</td>
</tr>
<tr>
<td>1977</td>
<td>15603</td>
<td>15500</td>
<td>15861</td>
<td></td>
<td></td>
<td></td>
<td>15605.19</td>
</tr>
<tr>
<td>1978</td>
<td>15861</td>
<td>15500</td>
<td>16807</td>
<td></td>
<td></td>
<td></td>
<td>15856.73</td>
</tr>
<tr>
<td>1979</td>
<td>16807</td>
<td>16500</td>
<td>16919</td>
<td>16846</td>
<td></td>
<td></td>
<td>16808.78</td>
</tr>
<tr>
<td>1980</td>
<td>16919</td>
<td>16500</td>
<td>16388</td>
<td>16846</td>
<td>16890</td>
<td></td>
<td>16917.96</td>
</tr>
<tr>
<td>1981</td>
<td>16388</td>
<td>16500</td>
<td>15553.9</td>
<td>16420</td>
<td>16395</td>
<td>16388</td>
<td>16386.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>19337</td>
<td>19500</td>
<td>18876</td>
<td>19334</td>
<td>19337</td>
<td>19335</td>
<td>19332.48</td>
</tr>
<tr>
<td>1992</td>
<td>18876</td>
<td>18500</td>
<td>14934.5</td>
<td>18910</td>
<td>18882</td>
<td>18882</td>
<td>18875.09</td>
</tr>
<tr>
<td>MSE</td>
<td>86694</td>
<td>53084</td>
<td>1101</td>
<td>234</td>
<td>173</td>
<td></td>
<td>16.9</td>
</tr>
<tr>
<td>RMSE</td>
<td>294.44</td>
<td>230.4</td>
<td>33.18</td>
<td>15.3</td>
<td>13.15</td>
<td></td>
<td>4.11</td>
</tr>
</tbody>
</table>

Figure 5: The curves of the MSE values between the our model and the compared models

b) The obtained results in the testing stage

From enrolments time series from 1971 to 1992, to forecast the new enrolments for the next year with one head - step. Can be explained through the examples as below: the historical data of enrolments from year 1971 to 1989, is utilized to forecast the new enrolment of year 1990. In the same way, the enrolments data between 1971 and 1990 are used to forecast data of year 1991. Thereafter the enrolments data have been well trained by our model, the future enrolments values could be accomplished to compare to the future ones of the forecasting models proposed in articles [4, 11, 14, 39]. A comparative forecasting results produced by the 3rd - order FTS with different number of intervals and the highest vote $W_h=15$ [14] which are shown in Table 11 and 12. From these Tables show that our model obtains the smallest RMSEs value equal to 99.2 and 60.28 among five competing models, respectively.
Table 11: Comparative results of our model with other models under the number of intervals of 7 and which use vote \( W_h=15 \)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>19328</td>
<td>18168</td>
<td>18059</td>
<td>18326</td>
<td>19525</td>
<td>19226</td>
<td>19246.4</td>
</tr>
<tr>
<td>1991</td>
<td>19337</td>
<td>18909</td>
<td>18669</td>
<td>19212</td>
<td>19150</td>
<td>19182</td>
<td>19267.5</td>
</tr>
<tr>
<td>1992</td>
<td>18876</td>
<td>19609</td>
<td>19083</td>
<td>19203</td>
<td>18933</td>
<td>18876</td>
<td>19010.2</td>
</tr>
<tr>
<td>RMSE</td>
<td>773.66</td>
<td>576.66</td>
<td>484.16</td>
<td>160.43</td>
<td>107.29</td>
<td>99.2</td>
<td></td>
</tr>
</tbody>
</table>

Table 12: Comparative results of our model with other models under the number of intervals of 14 and which use vote \( W_h=15 \)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>19328</td>
<td>18162</td>
<td>17862</td>
<td>18120</td>
<td>19287</td>
<td>19238</td>
<td>19278.3</td>
</tr>
<tr>
<td>1991</td>
<td>19337</td>
<td>18721</td>
<td>18633</td>
<td>19027</td>
<td>18811</td>
<td>19224</td>
<td>19363.2</td>
</tr>
<tr>
<td>1992</td>
<td>18876</td>
<td>19221</td>
<td>19085</td>
<td>19137</td>
<td>18836</td>
<td>19224</td>
<td>18964</td>
</tr>
<tr>
<td>RMSE</td>
<td>709</td>
<td>653.66</td>
<td>621.91</td>
<td>305.7</td>
<td>92.35</td>
<td>60.28</td>
<td></td>
</tr>
</tbody>
</table>

Table 13: The obtained results between our model and the competing models using the different number of intervals and various orders

<table>
<thead>
<tr>
<th>Year</th>
<th>Real data</th>
<th>[40]</th>
<th>[41]</th>
<th>[42]</th>
<th>[32]</th>
<th>[6]</th>
<th>Proposed model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1st - order</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3rd - order</td>
</tr>
<tr>
<td>1974</td>
<td>1574</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>1574</td>
</tr>
<tr>
<td>1975</td>
<td>1460</td>
<td>1497</td>
<td>----</td>
<td>1458</td>
<td>----</td>
<td>1506</td>
<td>1463.02</td>
</tr>
<tr>
<td>1976</td>
<td>1536</td>
<td>1497</td>
<td>----</td>
<td>1467</td>
<td>----</td>
<td>1553</td>
<td>1529.42</td>
</tr>
<tr>
<td>1977</td>
<td>1597</td>
<td>1497</td>
<td>1497</td>
<td>1606</td>
<td>1594</td>
<td>1598</td>
<td>1593.24</td>
</tr>
<tr>
<td>1978</td>
<td>1644</td>
<td>1497</td>
<td>1497</td>
<td>1592</td>
<td>1643</td>
<td>1584</td>
<td>1643.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>1145</td>
<td>1095</td>
<td>1098</td>
<td>1114</td>
<td>1146</td>
<td>1143</td>
<td>1137.48</td>
</tr>
<tr>
<td>2003</td>
<td>1035</td>
<td>995</td>
<td>9710</td>
<td>1097</td>
<td>1036</td>
<td>970</td>
<td>1087.35</td>
</tr>
<tr>
<td>2004</td>
<td>953</td>
<td>995.9</td>
<td>7929</td>
<td>929</td>
<td>954</td>
<td>970</td>
<td>958.54</td>
</tr>
<tr>
<td>MSE</td>
<td>83.12</td>
<td>46.78</td>
<td>37.66</td>
<td>19.2</td>
<td>41.61</td>
<td>29.4</td>
<td>0.85</td>
</tr>
</tbody>
</table>

4.2.2. Applying for Experiment 2

In this section, our proposed model have been implemented for forecasting the car road accidents in Belgium [32] ranges period from 1974 to 2004. We also test 30 times and take the best forecasting result to compare with the results from the other five forecasting models in articles [40 - 42, 32, 6]. The comparative results based on the different number of intervals and the various orders of FTS are shown in Table 13. Comparing between models [40, 6] and our proposed model, our model achieves the far better MSE value of 29.4 in two compared models based on the first - order FTS with different number of intervals. Also, from Table 13, we can be seen that, comparing model [41] and model [32], our proposed model produces the far smaller MSE value of 0.85 in two considered competing models using the 3rd - order FTS with different number of intervals. To summarize, our model provides better forecasting results and higher accuracy than the models in [40 – 42, 32, 6] corresponding to the number of orders of FTS and number of intervals equal to 14 as shown in Table 13.

5. Conclusions and upcoming work

In this paper, a novel model for predicting enrolments and car road accident is developed. To remedy the downside of the conventional FTS model, the FTS proposed model combines hedge algebras and PSO is developed to resolve two issues which are considered to be important and greatly affect the forecasting accuracy is that the length of intervals and fuzzy relationship group. By utilizing the concept of time variant fuzzy relationship group, the proposed model has handled the more persuasive historical data and has been demonstrated to be more appropriate for real-world applications. In addition to that the parameters of HA are modified by PSO algorithm to get the initial intervals partitioning of the UoD. In data mining and finding of optimal solution, PSO is considered to accomplish better compared to other heuristic techniques with regards to success rate and solution quality. Furthermore, the forecasting efficiency of the proposed model is significantly improved in adjusting the lengths of intervals. The forecasting performance of the proposed model is demonstrated by forecasting the enrolments at University of Alabama and the car road accidents in Belgium. Details of the comparison in Tables 7-13 indicate that the proposed model achieves the lowest forecasting errors when compared with other forecasting models., for many cases. Also, from Figure 5, it can be observed that the amount of error rate in terms of MSE obtained by our high - order FTS model are smaller than all other models considered in this research. Even though our model shows that the greater forecasting capability when compared with some of recent ones based on the high-order FLRs. Determining of the high-order FLR spends a lot of computational time than first-order FLR. Therefore, development of new approaches that can automatically select out the optimal degree of the high-order FLRs is a worthy idea in FTS forecasting model. Those will be the work closely related to this research in the coming time.
Acknowledgment

This study was supported in the Science Council of Thai Nguyen University of technology (TNUT), Thai Nguyen, Vietnam.

References


