International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 2 Feb 2021, pp: 726-734 www.ijaem.net ISSN: 2395-5252

A partitioning method in high-order fuzzy time series model using hedge algebras for forecasting enrolments

Nghiem Van Tinh[1*], Le Thi Luong[2], Bui Thi Thi[1]
[1]*Thai Nguyen University of Technology, Thai Nguyen University, ThaiNguyen, VietNam*
[2]*Industrial Economic Technology College, ThaiNguyen, VietNam*
**Corresponding Author: Nghiem Van Tinh*

Submitted: 10-02-2021 Revised: 24-02-2021 Accepted: 27-02-2021

ABSTRACT: Many the one - factor fuzzy time series (FTS) models which have been widely applied in recent years and can be divided into two classes, which are called first - order and high - order models. In the forecasting field, many researchers suggest that the high-order FTS model improve the forecasting accuracy. One of the important parts of obtaining high forecasting accuracy in fuzzy time series is that the length of interval is very significant. In this study, a novel forecasting model which combines Hedge algebra (HA) in the determining of interval length with the high - order fuzzy time series model for improving the forecasting accuracy. The proposed method was employed to forecast the enrolments of the University of Alabama to show the considerable outperforming results.

KEYWORDS: Fuzzy time series, Hedge algebra, high – order, forecasting, enrolments.

I. INTRODUCTION

During last few years, forecasting using fuzzy time series is applied to deal with various problems for helping people to make decisions, such as forecasting university enrolments [1-14], stock index [14-23], temperature prediction [20-23], crop productions [24] and car road accidents [25-28]. In the conventional time series models, the recorded values of a special dynamic process are represented by crisp numerical values. But, in a fuzzy time series model, the recorded values of a special dynamic process are represented by linguistic values. Based on the fuzzy time series theory, first forecasting model was introduced by Song and Chissom [1, 2], which were used to forecast the time series values based on linguistic values. They presented the fuzzy time series model by means of fuzzy relational equations involving max–min composition operation, and applied the model for forecasting the enrolments in the University of Alabama. Unluckily, their models had many drawbacks such as huge computation when the fuzzy rule matrix is large and lack of persuasiveness in determining the universe of discourse and the length of intervals. Therefore, Chen [3] had proposed the first-order fuzzy time series model by using simple arithmetic calculations instead of max-min composition operations in approach [1, 2] for better forecasting accuracy. Later, many studies have provided some improvements in framework of Chen [3] in terms of following issues:

- determining of lengths of intervals;
- fuzzification;
- fuzzy logical relationships;
- defuzzification techniques.

To further enhance forecasting accuracy of model, many researchers proposed various fuzzy time series models. For example, K. Huarng [14] presented an effective approach which can properly adjust the lengths of intervals. He pointed out that the different lengths of intervals in the universe of discourse can affect the forecasting result and a proper choice of the length of each interval can greatly improve the forecasting accuracy rate. The studies mentioned above can be categorized under the name of first-order fuzzy time series model. Since first-order fuzzy time series models have a simple structure, they can generally be insufficient to explain more complex relationships. For this reason, researchers in articles [4, 13, 21] have proposed computational methods of forecasting based on the high-order FLRs to overcome the drawback of fuzzy first-order forecasting models [1-3]. In addition, in [29], authors introduced a new approach, which uses a feed-forward neural network for defining fuzzy relations and is based on a high - order fuzzy time series forecasting model.

A completely different way from fuzzy approach, HA considers to be a sound option. Just recently, several related research works have been presented. For example, in paper [30], authors have presented a forecasting model based on the theory of hedge algebra [31] for forecasting university enrolments. 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, Research in [32] proposed a HA-based forecasting model to obtain unequal – size intervals in universe of discourse (UoD) by mapping the semantics of linguistic terms into fuzziness intervals. However, two these research works only focus on building the first-order model for forecasting the number of students annually at the University of Alabama.

Based on analysing of the above-mentioned research works showed that the length of intervals and the order of fuzzy relationships are two factors influencing strongly forecasting accuracy of model. The purpose of this study is to suggest a new high – order FTS model which uses HA to partitionthe UoD into intervals with unequal – length. The proposed model is examined on real-world datasets as the historical enrolment dataset of University of Alabama [3]. The examined results point out that the proposed forecasting model outperforms the some of the recent FTS models in terms of forecasting accuracy rate.

The rest of this paper is organized as follows: In Section 2, the basic concepts of FTS and Hedge algebra are briefly introduced. Section 3 presents a forecasting model using the HA and high – order FTS. Section 4 gives the forecasting results makes a comparison of forecasting results of the proposed model with the existing models. Conclusions are discussed in Section 5.

II. BASIC CONCEPTS OF FTS AND HA

2.1. Fuzzy time series

Song and Chissom [1, 2] firstly introduced the definition of FTS and constructed its model by means of fuzzy relational equations, in which the values of historical time series data are presented by fuzzy sets. Let $U = \{u_1, u_2, \ldots, u_n\}$ be an universe of discourse; a fuzzy set A of U can be defined as $A = \{\mu_A(u_1)/u_1 +, \mu_A(u_2)/u_2 \ldots + \mu_A(u_n)/u_n\}$, where $\mu_A : U \rightarrow [0,1]$ is the membership function of A, $\mu_A(u_i)$ indicates the degree of membership of u_i in the fuzzy set A, $f_A(u_i) \in [0, 1]$, and $1 \leq i \leq n$. The basic definitions related to FTS are summarized as below:

Definition 1: Fuzzy time series

Let $Y(t)(t = \ldots, 0, 1, 2 \ldots)$, a subset R^1 ($Y(t) \subseteq R^1$), be the universe of discourse in which the fuzzy sets $f_i(t)$ $(i = 1,2 \ldots)$ are defined. Assume that F(t) consists a collection of $f_1(t), f_2(t), \ldots$, then F(t)is called a FTS defined on Y(t).

Definition 2: Fuzzy logical relationship (FLR)

Support F(t) is instigated only by F(t-1), as F(t-1)→ F(t). Then it can be expressed as F(t-1)∗ R(t-1, 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 fuzzy logical relationship $A_i \rightarrow A_j$, where A_i and A_j are called the current state and the next state of fuzzy relationship, respectively. Furthermore, these fuzzy logical relationships can be grouped to establish different fuzzy relationship.

Definition 3: n- order fuzzy logical relationship

Let F(t) be a fuzzy time series. If F(t) derived by more fuzzy sets F(t-1), F(t-2),…, F(t-n+1), F(t-n), then fuzzy relationship between them can be represented as $A_{kn}, A_{k(n-1)}, \ldots, A_{k1} \rightarrow A_i$; Where F(t-n) = A_{kn}, …, F(t-2) =A_{k2}, F(t-1) =A_{k1}, F(t) = A_i. This relationship is called the n- order fuzzy time series model.

2.2. The concepts of Hedge Algebras [31]

In this study, the concept of HA is employed as basis to partition data of time series into initial intervals with unequal-length. Assume that there is a set of linguistic values as the domain of the linguistic of variable X which are sorted by the following words: X = {Very small < small < little small < Little low < medium < Little large < large < Very large}. Each of linguistic variable $\mathcal{X}$ is represented by an algebraic structure as $\mathcal{AX} = (X, G, C, H, \leq)$ and called HA, where X is the set of terms in $\mathcal{X}$; ≤ denotes a natural semantically ordering relation on X; $G = \{c^-, c^+\}$, $c^- \leq c^+$is the set of generating elements. $C = \{0, w, 1\}$is a set of constants, with $(\mathbf{0} \leq c^- \leq \mathbf{W} \leq c^+ \leq \mathbf{1})$; $H = H^- \cup H^+$, with$H^- = \{h_{-q} \geq \cdots \geq h_{-2} \geq h_{-1}\}$ is the set of all negative hedges of X, $\forall h \in H^-$ then $hc^+ \leq c^+$ and $H^+ = \{h_1 \leq h_2 \leq \cdots \leq h_p\}$ is the set of all positive ones of X, $\forall h \in H^+$ then $hc^+ \geq c^+$.Some basic definitions of HA are given as follows:

Definition 4: Let $\mathcal{AX} = (X, G, C, H, \leq)$ be a HA. An fm: X → [0,1] is said to be a fuzziness measure of terms in X if:

(1). $fm(c^-) + fm(c^+) = 1$ and $\sum_{h \in H} fm(hu) = fm(u)$, for ∀u∈X; in this case fm is called complete;

(2). For the constants **0**, **W** and **1**, $fm(\mathbf{0}) = fm(\mathbf{W}) = fm(\mathbf{1}) = 0$;

(3). For ∀x, y ∈ X, ∀h ∈H, $\frac{fm(hx)}{fm(x)} = \frac{fm(hy)}{fm(y)}$, that is this proportion does not depend on specific elements and, therefore, it is called fuzziness measure of the hedge h and denoted by μ(h). The properties of fm(x) and μ(h) are offered as follows:

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 2 Feb 2021, pp: 726-734 www.ijaem.net ISSN: 2395-5252

Proposition 1. The fm(x) represents the fuzziness measurement on X, the following statements hold.

1). fm(hx) = μ(h)fm(x), for every x∈X;

2). fm(c^-)+fm(c^+) = 1;

3). $\sum_{-q\le i\le p, i\ne 0} fm(h_i c) = fm(c)$, c∈{$c^-$, c^+};

4). $\sum_{-q\le i\le p, i\ne 0} fm(h_i x) = fm(x)$;

5). $\sum_{-q\le i\le -1} \mu(h_i) = \alpha$ and $\sum_{1\le i\le p} \mu(h_i) = \beta$, where α, β> 0 and α+ β = 1.

Definition 5. The fuzziness interval of the linguistic terms x ∈**X**, symbolized by ℑ(x), is a subinterval of [0,1], if |ℑ(x)| = fm(x) where |ℑ(x)| is the length of fm(x), and recursively determined by the length of x as follows:

1). If length of x is equal to 1 (l(x)=1), that mean x ∈ {c^-, c^+}, then |ℑ(c^-)| = fm(c^-), |ℑ(c+)|= fm(c^+) and ℑ(c^-) ≤ ℑ(c^+);

2). Assume that n is the length of x (l(x)=n) and fuzziness interval ℑ(x) has been defined with |ℑ(x)| = fm(x). The set {ℑ(h_jx)| j ∈ [-q^p]}, where [-q^p] = {j | -q ≤ j ≤ -1 or 1 ≤ j ≤ p}, is a partition of ℑ(x) and we have: for Sgn(h_px) = −1, ℑ(h_px) ≤ ℑ(h_{p-1}x) ≤ … ≤ ℑ(h_1x) ≤ ℑ(h_{-1}x) ≤ … ≤ ℑ(h_{-q}x); for Sgn(h_px) = +1, ℑ(h_{-q}x) ≤ ℑ(h_{-q+1}x) ≤ … ≤ ℑ(h_{-1}x) ≤ ℑ(h_1x) ≤ … ≤ ℑ(h_px).

III. A PROPOSED FORECASTING MODEL BASED ON HIGH-ORDER FTS AND HA

In this section, we introduce forecasting model using the high – order FTS and HA for forecasting enrolments of the University [3]. Initially, the HA is applied to partition the UoD into several intervals with unequal- sizes by quantitative mapping of linguistic terms into fuzzy intervals. Based on these newly obtained intervals, we defined fuzzy sets and fuzzy historical data on each divided interval. From these fuzzified values, we derive the FLRs and establish fuzzy relationship groups according to [13]. Later, all information on the right - hand side of these FRGs are used to get the final forecasting results. The proposed forecasting model can be given step-by-step as follows:

Step 1: Define the discourse of universe and subintervals.

Let $UoD = [D_{min} - N_1, D_{max} + N_2]$ is universe of discourse. Based on minimal and maximal values in the time series dataset, D_{min} and D_{max} variables are defined. Then, two arbitrary positive numbers which are N_1 and N_2 are are properly selected so that all the values of time series belong to UoD. From the enrolments time series [3], the universe of discourse is defined as UoD = [13000, 20000], where D_{min}, = 13055, D_{max} = 19337, D_1 = 55, D_2 = 663, LU = 7000.

Step 2: Partition UoD into different intervals based on HA

This step uses HA with structure which is presented in **Definition 4** as $\mathcal{AX} = (X, G, C, H, \le)$, where X is the set of terms of the linguistic variable "enrollments"{ X= dom(enrollments)}; ≤ denotes a natural semantically ordering relation on X; $G = \{c^-, c^+\}$ ={Low, High}, Low (**Lo**) ≤ High (**Hi**); C = {0, w, 1} a set of constants, with ($\mathbf{0} \le c^- \le \boldsymbol{W} \le c^+ \le \mathbf{1}$) and H ={ Very, Little }. In this paper, we choose the number of intervals equal to 7 and 14 corresponding to the number of linguistic terms used to quantify the time series values, as shown in Table 1. Calculate the value of each interval by performing the following sub-steps:

Step 2.1: The UoD is mapped to the domain [0, 1]

Assume that the number of enrollment value is less than 16000 is low, then we can establish the parameters of HA as follows$fm(low) = \frac{16000-13000}{20000-13000} = 0.428$, fm(high) = 1- $fm(low)$ = 0.572. With LU = 7000, mapping these values to U, we can obtain covfm(low) and covfm(high) by calculating as fm(low) x LU = 0.428 x 7000 = 2996, fm(high) x LU = 0.572 x 7000 = 4004, respectively. In this paper, we can choose μ(Little) = 0.4, μ(Very) = 1- 0.4 = 0.6. Based on μ(Little), μ(Very) value, the value of α , β is defined as 0.4, 0.6, respectively. For example, in case seven linguistic terms, Based on Proposition 1, the fuzziness measure of these linguistic terms in the domain [0,1] can be calculated as follows:$fm(VVLo)$ = 0.154, $fm(LVLo)$ = 0.1027, $fm(LLLo)$ = 0.068, $fm(VLLo)$ = 0.1027, $fm(VLHi)$ = 0.1373, $fm(LLHi)$ = 0.0915, $fm(VHi)$ = 0.3432.

Step 2.2: Generate the fuzzy interval of linguistic variable in the UoD

Based on step 2.1, The linguistic values of terms belong to fuzziness interval is calculated as follows:

$covfm(A_1)$ = μ(V) x μ(V) x $covfm$ (Lo) = 0.6*0.6*2996 = 1078.56;

$covfm(A_2)$ = μ(L) x μ(V) x $covfm$ (Lo) = 0.4*0.6*2996 = 719.04;

$covfm(A_3)$ = μ(L) x μ(L) x covfm(Lo) = 0.4*0.4*2996 = 479.36;

$covfm(A_4)$ = μ(Very) x μ($Little$) x covfm(Low) = 0.6*0.4*2996 = 719.04;

$covfm(A_5)$ = μ(V) x μ(L) x covfm(Hi) = 0.6*0.4*4004 = 960.96;

$covfm(A_6)$ = μ(L) x μ(L) x covfm(Hi) = 0.4*0.4*4004 = 640.64;

--;

$covfm(A_7)$ = μ(V) x covfm(Hi) = 2402.4

Mapping the value of linguistic terms to the domain of the universe of discourse UoD, we get the intervals corresponding to linguistic terms, which are listed in Table 2:

Table 1: The number of linguistic terms

Number of linguistic terms	Linguistic terms and their order
7	A_1= Very Very Low (VVLo) <A_2= Little Verry Low (LVLo) <A_3= Little Little Low (LLLo) <A_4=Very Little Low (VLLo) <A_5 =Very Little High (VLHi) <A_6= Little Little High (LLHi) <A_7= Very High (VHi)
14	A_1 = VVLo<A_2 = LLVLo<$A_3 = VLVLo$<$A_4 =$ $VLLLo$<$A_5 = LLLLo$<$A_6 = LVLLo$<$A_7 =$ $VVLLo$<$A_8 = VVLHi$<$A_9 = LVLHi$<$A_{10} =$ $LLLHi$<$A_{11} = VLLHi$<$A_{12} = VLVHi$<$A_{13} =$ $LLVHi < A_{14} = VVHi$

Table 2: The intervals obtained from HA

7 intervals	14 intervals
I_1= [13000, 14078.56)	I_1= [13000, 14078.56)
I_2= [14078.56, 14797.6)	I_2= [14078.56, 14366.18)
I_3= [14797.6, 15276.96)	I_3= [14366.18, 14797.6)
I_4= [15276.96, 15996)	---------------------------
I_5= [15996, 16956.96)	I_{12}= [17836.6, 18392.3)
I_6= [16956.96, 17597.6)	I_{13}= [18392.3, 18958.6)
I_5= [17597.6, 20000)	I_{14}= [18958.6, 20000)

Step 3: Define linguistic terms represented by the triangular fuzzy sets A_i

For seven intervals in Table 1, there are seven linguistic values to represent different regions in the UoD. Each linguistic value represents a triangular fuzzy set $A_i$and it is presented as follows.

$$A_1 = {a_{1,1}}/{I_1} + {a_{1,2}}/{I_2} + \cdots + {a_{1,7}}/{I_7}$$

$$A_1 = {a_{2,1}}/{I_1} + {a_{2,2}}/{I_2} + \cdots + {a_{2,7}}/{I_7}$$

....

$$A_7 = {a_{7,1}}/{I_1} + {a_{7,2}}/{I_2} + \ldots + {a_{7,7}}/{I_7}$$

Where, $a_{i,j} \in [0,1]$ ($1 \leq i \leq 7, 1 \leq j \leq 7$), $I_j$is the ^{j}th interval of the UoD. The value of $a_{i,j}$indicates the grade of membership of I_j in the fuzzy set A_i

Step 4: Fuzzify the historical time series data

For fuzzification of historical time series data, we choose the maximum degree of membership grades of $I_i (1 \leq i \leq 7)$. If the maximum membership value of one year's observation occurs at I_i, then the fuzzified value for that particular year is considered as A_i.

Step 5: Buid all n–order (high-order)FLRs ($n \geq 2$). The fuzzy logical relationship can be construct by two or several consecutive fuzzified values, respectively. To create a n-order FLR, we need to explore any relationship - type as $F(t-n), F(t-n+1), \ldots, F(t-1) \rightarrow F(t)$, in which $F(t-n), F(t-n+1), \ldots, F(t-1)$ and $F(t)$ are called the "current state" and the "next state" of the fuzzy relationship, respectively. Then, it has been obtained by replacing the corresponding fuzzy sets A_i.

Step 6: Generate all high- order the FRGs

In this study, we rely on a framework [13] for generating the FRGs. To explain this, assume that there two 3rd-order FRs with the left-hand side in chronological order t-2, t-1and t respectively as follows $A_i, A_j, A_k \rightarrow A_p$; $A_i, A_j, A_k \rightarrow A_q$; $A_i, A_j, A_k \rightarrow A_u$. Suppose that we want to forecast the value of time series data at time t-1. Based on appearance history of the fuzzy sets on the right-hand side of the FRs, we get a 3rd-order FRG as G1: $A_i, A_j, A_k \rightarrow A_p, A_q$. By the same way, if considering forecasting time of t, the FLRs which

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 2 Feb 2021, pp: 726-734 www.ijaem.net ISSN: 2395-5252

have the same right-hand side are grouped into a group G2 as: A_i, A_j, $A_k \rightarrow A_p$, A_q, A_u.

Step 7: Calculate the forecasted output values

In order to calculate the forecasted values for all high – order TV-FRGs, we apply the approach in paper [13] for the trained patterns in the training phase. For each group obtained in Step 7, we divide each corresponding interval of each next state into q sub-regions with equal size, and create a forecasted value for each group according to formula (1) as below:

$$forecasted_{output} = \frac{1}{p}\sum_{j=1}^{p}\frac{(m_{kj}+subm_{kj})}{2} \quad (1)$$

- p is the total number of next states or the total number of fuzzy sets on the next state within the same group.
- m_{kj} ($1 \leq j \leq p$)is the midpoint of interval u_{kj} corresponding to j-th fuzzy set on the next statewhere the highest level of fuzzy set A_{kj} takes place in these intervals, u_{kj}.
- $subm_{kj}$ is the midpoint of one of q sub-regions (means the midpoint of k-th sub-interval in which the historical data belong to this sub-interval) corresponding to j-th fuzzy set on the right-hand side where the maximum grade of A_{kj} takes place in this interval.

Step 8: Evaluate the performance of model

Compute the mean square error (MSE) and root mean square error (RMSE) to check the forecasted accuracy rate. Lower values of MSE and RMSE indicate a better forecasting method. Their values are defined as follows:

$$MSE = \frac{1}{n}\sum_{i=\lambda}^{n}(F_i - R_i)^2 \quad (2)$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=n}^{m}(F_i - R_i)^2} \quad (3)$$

Where, Ri and Fi indicate the real and forecasted value at time i, respectively, m denotes the total number of years to be forecasted, n is the order of fuzzy relation.

IV. APPLICATION TO THE ENROLMENTS DATA

The proposed model is applied to the enrolment data of University of Alabama [3] which is presented in Fig.1.When obtaining forecasts for training set, all the enrolment observations in the dataset are used in creating fuzzy relations. The number of intervals used in the analysis varied between 7 and 14 as in the other studies.

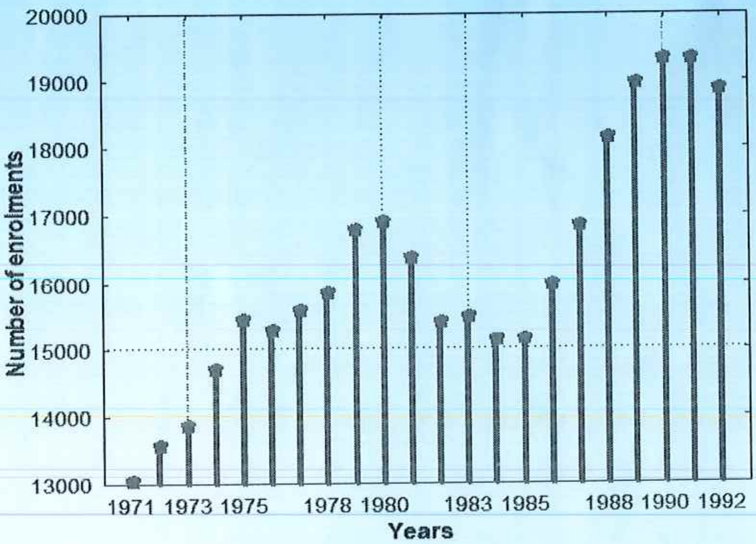

Fig.1: The actual data of enrolments

Table 3 gives results of the proposed method for 2nd-order through 7th-order fuzzy time series under seven intervals. The best forecasts are obtained from the proposed model with the order 4th and achieved the forecasting accuracy in term of RMSE equal to 152.4. In addition, our model is also tested on the different high - order FLRs with number of intervals of 14. The results obtained from this case are summarized in Table 4. From Table 4, it is obvious that the forecasting model gets the RMSE value of 85 when 4th - order FLR is examined which is the smallest among all orders of model.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 2 Feb 2021, pp: 726-734 www.ijaem.net ISSN: 2395-5252

Table 3: The results obtained from the proposed model using the different high -order under 7 intervals

Years	Actual data	Forecasted results					
		2nd- order	3rd- order	4th - order	5th - order	6th-order	7th-order
1971	13055						
1972	13563						
1973	13867	13719.1					
1974	14696	14138.6	14558				
1975	15460	15516.7	15516.7	15516.7			
1976	15311	15516.7	15516.7	15516.7	15516.7		
1977	15603	15636.2	15636.2	15636.2	15636.2	15636.2	
1978	15861	15696.2	15756.2	15756.2	15756.2	15756.2	15756.2
1979	16807	16009.7	16196.5	16636.7	16636.7	16636.7	16636.7
1980	16919	16636.7	16636.7	16636.7	16636.7	16636.7	16636.7
1981	16388	16476.2	16476.2	16476.2	16476.2	16476.2	16476.2
1982	15433	15996.5	15516.7	15516.7	15516.7	15516.7	15516.7
1983	15497	15516.7	15516.7	15516.7	15516.7	15516.7	15516.7
1984	15145	15786.6	15117.1	15117.1	15117.1	15117.1	15117.1
1985	15163	15117.1	15117.1	15117.1	15117.1	15117.1	15117.1
1986	15984	15756.2	15756.2	15756.2	15756.2	15756.2	15756.2
1987	16859	16636.7	16636.7	16636.7	16636.7	16636.7	16636.7
1988	18150	17517.6	18398.4	18398.4	18398.4	18398.4	18398.4
1989	18970	18798.9	18798.9	18798.9	18798.9	18798.9	18798.9
1990	19328	19199.4	19199.4	19199.4	19199.4	19199.4	19199.4
1991	19337	19199.4	19199.4	19199.4	19199.4	19199.4	19199.4
1992	18876	19065.9	18999.2	18798.9	18798.9	18798.9	18798.9
RMSE		350.9	203.9	152.4	156.2	152.6	157.4

Table 4: The forecasting results of the proposed model using the different high -order under 14 intervals

Years	Actual data	Forecasted results					
		2nd- order	3rd- order	4th - order	5th - order	6th-order	7th-order
1971	13055						
1972	13563						
1973	13867	13719.1					
1974	14696	14186.5	14653.9				
1975	15460	15420.9	15420.9	15420.9			
1976	15311	15372.9	15372.9	15372.9	15372.9		
1977	15603	15708.1	15708.1	15608.1	15708.1	15708.1	
1978	15861	15852.1	15852.1	15852.1	15852.1	15852.1	15852.1
1979	16807	16764.9	16764.9	16764.9	16764.9	16764.9	16764.9
1980	16919	16828.9	16828.9	16828.9	16828.9	16828.9	16828.9
1981	16388	16380.1	16380.1	16380.1	16380.1	16380.1	16380.1
1982	15433	15420.9	15420.9	15420.9	15420.9	15420.9	15420.9
1983	15497	15468.9	15468.9	15468.9	15468.9	15468.9	15468.9
1984	15145	15428.6	15149	15149	15149	15149	15149
1985	15163	15181	15181	15181	15181	15181	15181
1986	15984	15852.1	15852.1	15852.1	15852.1	15852.1	15852.1
1987	16859	16828.9	16828.9	16828.9	16828.9	16828.9	16828.9
1988	18150	17453.5	18078	18078	18078	18078	18078
1989	18970	19167.2	19167.2	19007.2	19167.2	19167.2	19167.2
1990	19328	19375.2	19375.2	19375.2	19375.2	19375.2	19375.2
1991	19337	19375.2	19375.2	19375.2	19375.2	19375.2	19375.2

1992	18876	19271.2	19167.2	19167.2	19167.2	19167.2	19167.2
RMSE		233.64	97.4	85	102	104.4	104

To verify the superiority of the proposed model under the different number of orders with number of interval equal to 14, five existing forecasting models [29, 32, 33, 11, 10] are selected for comparing. A comparison of the forecasted results with regards to MSE (2) is presented in Table 5. Based on Table 5, it can be seen that the proposed model produces more precise results than the existing competing models under number of order equal to 4.

Table 5: The comparison of the forecasting accuracy between our model and existing models

Models	**orders**	**MSE**
[29]	1	66661
[32]	1	44057
[33]	3	60714
[11]	8	17106
[10]	8	14778
Our model	4	7225

V. CONCLUSION

In this study, we propose the high – order FTS model which uses HA to addressed two issues are considered to be important and greatly affect the forecasting accuracy that is the problem about determining of length of intervals and selecting orders of fuzzy relationship. The proposed model is employed to forecast the well-known data of the enrolments of the University of Alabama to show the considerable outperforming results. Also, the methods proposed by previous researchers and conventional time series methods are applied to the data for comparison. In the end of the comparison, it is obviously seen that the proposed method produces better forecasts than those of other methods. These results indicate that using the hedge algebra for determining length of intervals in UoD significantly enhance the forecasting accuracy in high - order FTS model.

ACKNOWLEDGMENT

The study described in this paper was supported by Thai Nguyen University of Technology (TNUT).

REFERENCES

[1]. Q. Song and B. S. Chissom, Forecasting enrolments with fuzzy time series - Part I, Fuzzy Sets and Systems, vol.54, no.1, 1-9, 1993.

[2]. Q. Song and B. S. Chissom, Fuzzy time series and its models, Fuzzy Sets and Systems, vol.54, no.3, 269-277, 1993.

[3]. S.-M. Chen, Forecasting Enrolments based on Fuzzy Time Series, Fuzzy set and systems, vol. 81, 311-319, 1996.

[4]. S.M. Chen, forecasting enrolments based on high-order fuzzy time series, Cybernetics and Systems, vol.33, no.1, 1-16, 2002.

[5]. S.M. Chen, N.Y. Chung, Forecasting enrolments using high-order fuzzy time series and genetic algorithms. International Journal of Intelligent Systems, 21, 485–501, 2006.

[6]. S.M. Chen, K. Tanuwijaya, Fuzzy forecasting based on high-order fuzzy logical relationships and automatic clustering techniques, Expert Systems with Applications 38,15425–15437, 2011

[7]. S.M. Chen, Bui Dang H.P. Fuzzy time series forecasting based on optimal partitions of intervals and optimal weighting vectors. Knowledge-Based Systems 118, 204–216, 2017.

[8]. M. Bose1, K. Mali, A novel data partitioning and rule selection technique for modelling high-order fuzzy time series. Applied Soft Computing,https://doi.org/10.1016/j.asoc.2017.11.011, 2017.

[9]. Egrioglu, E., Aladag, C. H., Yolcu, Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks. Expert Systems with. Applications, 40(3), 854–857, 2013.

[10]. Y. L. Huang. et al., A hybrid forecasting model for enrolments based on aggregated fuzzy time series and particle swarm optimization, Expert Systems with Applications 7(38), 8014–8023, 2011.

[11]. I-H. Kuo, et al., An improved method for forecasting enrolments based on fuzzy time series and particle swarm optimization, Expert systems with applications, 36, 6108–6117, 2009.

Expert systems with applications, 36, 6108–6117, 2009.
[12]. Nguyen Cong Dieu, Tran Quang Duy, Nghiem Van Tinh, Forecasting model based on fuzzy time series and particle swarm optimization, Journal of Mathematical Applications. vol.14, no.1, 65-81, 2016.
[13]. Nghiem Van Tinh, Nguyen Cong Dieu (2018) Handling Forecasting Problems Based on Combining High-Order Time-Variant Fuzzy Relationship Groups and Particle Swam Optimization. Technique International Journal of Computational Intelligence and Applications 17: 1-19.
[14]. K. Huarng, Effective lengths of intervals to improve forecasting in fuzzy time series, Fuzzy Sets and Systems, vol.123, no.3, 387-394, 2001.
[15]. K. Huarng, T.H.K. Yu, the application of neural networks to forecast fuzzy time series, Phys. A, Stat. Mech. Appl. 363(2), 481–491, 2006.
[16]. S.M. 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 391–392, 65–79, 2017.
[17]. S.H. Cheng. S.-M Chen, W.S. Jian, Fuzzy time series forecasting based on fuzzy logical relationships and similarity measures. Information Sciences, 327, 272–287, 2016
[18]. I-H. Kuo et al., Forecasting TAIFEX based on fuzzy time series and particle swarm optimization, Expert Systems with Applications 2(37), 1494–1502, 2010.
[19]. L.W. Lee, L.H. Wang, S.M. Chen, Y.H. Leu, Handling forecasting problems based on two-factors high-order fuzzy time series. IEEE Transactions on Fuzzy Systems, 14, 468–477, 2006.
[20]. Park J.I., Lee, D.J., Song C.K., Chun M.G. TAIFEX and KOSPI 200 forecasting based on two-factors high-order fuzzy time series and particle swarm optimization Expert Systems with Applications 37, 959–967, 2010.
[21]. Lee et al., Temperature prediction and TAIFEX forecasting based on high order fuzzy logical relationship and genetic simulated annealing techniques, Expert Systems with Applications, 34, 328–336, 2008.
[22]. Ling-Yuan Hsu et al., Temperature prediction and TAIFEX forecasting based on fuzzy relationships and MTPSO techniques, Expert Systems with Applications. 37, 2756–2770, 2010.
[23]. Lee, L.-W. Wang, L.-H., & Chen, S.-M, "Temperature prediction and TAIFEX forecasting based on high order fuzzy logical relationship and genetic simulated annealing techniques," Expert Systems with Applications, 34, 328–336, 2008.
[24]. R.S. Singh, A simple method of forecasting based on fuzzy time series. Applied Mathematics and Computation, 186, 330–339, 2007.
[25]. S M. Yusuf, M.B. Mu'azu, O. Akinsanmi, " a Novel Hybrid fuzzy time series Approach with Applications to Enrollments and Car Road Accident", International Journal of Computer Applications, 129(2):37 – 44, 2015.
[26]. E.Bas, V. Uslu. , U.Yolcu, E. Egrioglu, "A modified genetic algorithm for forecasting fuzzy time series", Appl Intell, 41, 453-463, 2014.
[27]. T A. Jilani, S.M. Burney, C. Ardil, "Multivariate high order FTS forecasting for car road accident", World Acad Sci Eng Technol, 25, 288 - 293, 2008.
[28]. T.A Jilani, A.M. Burney, Multivariate stochastic fuzzy forecasting models". Expert Syst Appl, 353, 691–700, 2008.
[29]. Aladag, C. H., Basaran, M. A., Egrioglu, E., Yolcu, U., &Uslu, V. R., Forecasting in high order fuzzy times series by using neural networks to define fuzzy relations. Expert Systems with Applications, 36, 4228–4231, 2009.
[30]. N.C. Ho, N.C. Dieu, V.N. Lan. "The application of hedge algebras in fuzzy time series forecasting", Journal of science and technology, 54(2), 161-177, 2016.
[31]. N.C. Ho., W. Wechler,"Hedge algebra: An algebraic approach to structures of sets of linguistic truth values", fuzzy Sets and Systems, 35, 281-293, 1990.
[32]. H. Tung, N.D. Thuan, V.M. Loc, "The partitioning method based on hedge algebras for fuzzy time series forecasting", Journal of Science and Technology, 54 (5), 571-583, 2016.
[33]. E. Egrioglu et.al., Finding an optimal interval length in high order fuzzy time series, Expert Systems with Applications 37, 5052–5055, 2010.