International Journal of Scientific Engineering and Science
Volume 3, Issue 1, pp. 6-12, 2019. ISSN (Online): 2456-7361

Enhanced Accuracy of High – Order Fuzzy Time Series Forecasting Model Based on Harmony Search Algorithm

Nghiem Van Tinh, Bui Thi Thi
Thai Nguyen University of Technology, Thai Nguyen University, ThaiNguyen, VietNam
Email: nghiemvantinh@tnut.edu.vn, buithithi@tnut.edu.vn

Abstract— *In recent years, many fuzzy time series models have already been used to solve nonlinear and complexity issues. However, first-order fuzzy time series models have proven to be insufficient for solving these problems. For this reason, many researchers have been proposed high-order fuzzy time series model to improve the forecasting accuracy. From this viewpoint. This paper presents a high-order forecasting model based on fuzzy time series (FTS) and harmony search algorithm to overcome the drawbacks above. Firstly, a forecasting model is constructed from the high – order fuzzy logical relationship. Following, the harmony search algorithm is combined with FTS model to adjust the lengths of each interval and find optimal interval in the universe of discourse with an intend to increase forecasting accuracy. To illustrate the forecasting process and the usefulness of the proposed model, two numerical datasets like average rice production of Viet Nam and the enrolments of the University of Alabama are used. The application results accentuate the superiority of the proposed model over the other models for forecasting the enrolments of the University of Alabama.*

Keywords— *Forecasting, FTS, harmony search algorithm, enrolments, rice production.*

I. INTRODUCTION

Future forecasting of time series events has attracted people from the beginning of times. They were using some forecasting models to deal with various problems, such as academic enrollments [1] - [9] , crop production [10] , [11] , stock markets [12] - [14] and temperature prediction [14] , [15] . The tradition forecasting methods cannot deal with forecasting problems in which the historical data needs to be represented by linguistic values. Fuzzy set theory was firstly presented by Zadeh [16] to handle problems with linguistic values. The concepts of fuzzy sets have been successfully adopted to time series by Song and Chissom [1] . They introduced both the time-invariant fuzzy time series [1] and the time-variant time series [2] model which use the max–min operations to forecast the enrolments of the University of Alabama. Unfortunately, their method needs max–min composition operations to deal with fuzzy rules. It takes a lot of computation time when fuzzy rule matrix is big. Therefore, Chen [4] proposed the first-order fuzzy time series model by using simple arithmetic calculations instead of max-min composition operations [3] for better forecasting accuracy. After that, fuzzy time series has been widely studied for improving accuracy of forecasting in many applications. Huarng [5] presented effective approaches which can properly adjust the lengths of intervals to get better forecasting accuracy. Chen [6] proposed a new forecast model based on the high-order fuzzy time series to forecast the enrollments of University of Alabama. Yu [7] presented a new model which can refine the lengths of intervals during the formulation of fuzzy relationships and hence capture the fuzzy relationships more appropriately. Both the stock index and enrollment are used as the targets in the empirical analysis. Chen & Chung [8] , [9] presented the first-order and high-order fuzzy time series model to deal with forecasting problems based on genetic algorithms. Singh [10] , [11] presented simplified and robust computational methods for the forecasting rules based on one and various parameters as fuzzy relationships, respectively. Lee et al. [14] presented a method for forecasting the temperature and the TAIFEX based on the high-order fuzzy logical relation groups and genetic algorithm. They also used genetic algorithm and simulated annealing in it. Recently, Particle swarm optimization technique has been successfully applied in many applications. Based on Chen's model [4] , Kuo et al. [17] developed a new hybrid forecasting model which combined fuzzy time series with PSO algorithm to find the proper length of each interval. Following, to improve previous model [17] . They continued to present a new forecast method to solve the TAIFEX forecasting problem based on fuzzy time series and PSO algorithm [17] , [18] . Some other authors, propose some methods for the temperature prediction and the TAIFEX forecasting, based on two-factor fuzzy logical relationships [19] and use them in which combine with PSO algorithm in fuzzy time series [20] . In Addition, other hybrid techniques such as: Pritpal and Bhogeswar [21] presented a new model based on hybridization of fuzzy time series theory with artificial neural network (ANN). Matarneh et al. [22] use feed forward artificial neural network and fuzzy logic for weather forecasting achieve better results.

The above mentioned researches showed that the lengths of intervals and creating forecasting rules are two important issues considered to be serious influencing the forecasting accuracy and applied to different problems. However, most of the models were implemented for forecasting of other historical data and not rice production. In this paper, a forecasting model based on the fuzzy logical relationship groups and Harmony search algorithm is presented to forecast rice production for each year on basis of historical time series of rice data in Viet Nam. Firstly, the forecasting model is presented to find forecasting values based on fuzzy time series. Then, the root mean square error (RMSE) value is

International Journal of Scientific Engineering and Science
Volume 3, Issue 1, pp. 6-12, 2019. ISSN (Online): 2456-7361

applied to estimate the forecasting accuracy. Finally, a new hybrid forecasting model based on combined FTS and HS is developed to adjust the length of each interval in the universe of discourse by minimizing RMSE value. The case study with the data of rice production of Viet Nam and the enrolment data at the University of Alabama show that the performance of proposed model is better than those of any existing models based on the high – order FTS.

The rest of this paper is organized as follows. In Section 2, a brief review of the basic concepts of FTS and Harmony search algorithms are introduced. Section 3, first gives the details of fuzzy time series forecasting model to forecast rice production and then combines with the HS algorithm to find the effective lengths of intervals in the universe of discourse during training phase. Section 4 evaluates the forecasting performance of the proposed method with the existing methods based on the enrolments data of the University of Alabama. Finally, Section 5 provides some conclusions.

II. Basic Concepts of Fuzzy Time Series and Algorithms

A. *Basic Concepts of Fuzzy Time Series*

Conventional time series refer to real values, but fuzzy time series are structured by fuzzy sets [16]. Let $U = \{u_1, u_2, \ldots, u_n\}$ be an universal set; a fuzzy set A_i of U is defined as $A_i = \{f_A(u_1)/u_1 +, f_A(u_2)/u_2 \ldots + f_A(u_n)/u_n\}$, where f_A is a membership function of a given set A, f_A :U→[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 $1 \leq i \leq n$.

General definitions of FTS are given as follows:

Definition 1: Fuzzy time series [1] , [2]

Let $Y(t)$ $(t = \ldots, 0, 1, 2 \ldots)$, a subset of R, be the universe of discourse on which fuzzy sets $f_i(t)$ $(i = 1, 2 \ldots)$ are defined and if $F(t)$ is a collection of $f_1(t), f_2(t), \ldots$, then $F(t)$ is called a fuzzy time series on $Y(t)$ (t . . ., 0, 1,2 . ..). Here, $F(t)$ is viewed as a linguistic variable and $f_i(t)$ represents possible linguistic values of $F(t)$.

Definition 2: Fuzzy logic relationship(FLR**)** [2] , [3]

If F(t) is caused by F(t-1) only, the relationship between F(t) and F(t-1) can be expressed by $F(t-1) \to F(t)$. According to [2] suggested that when the maximum degree of membership of $F(t)$ belongs to A_i, $F(t)$ is considered A_j. Hence, the relationship between F(t) and F(t -1) is denoted by fuzzy logical relationship $A_i \to A_j$ where Ai and Aj refer to the current state or the left - hand side and the next state or the right-hand side of fuzzy time series.

Definition 3: γ - order fuzzy logical relationships [6]

Let $F(t)$ be a fuzzy time series. If $F(t)$ is caused by $F(t-1), F(t-2), \ldots, F(t-\gamma+1)$ $F(t-m)$ then this fuzzy relationship is represented by by $F(t-\gamma), \ldots, F(t-2), F(t-1) \to F(t)$ and is called an m -order fuzzy time series.

Definition 4: Fuzzy relationship group (FRG) [4]

Fuzzy logical relationships, which have the same left-hand sides, can be grouped together into fuzzy logical relationship groups. Suppose there are relationships such as follows:

$$A_i \to A_{k1}, A_i \to A_{k2}, \ldots\ldots$$

In previous study was proposed by Chen [1], the repeated fuzzy relations were simply ignored when fuzzy relationships were established. So, these fuzzy logical relationship can be grouped into the same FRG as: $A_i \to A_{k1}, A_{k2} \ldots$

B. *Harmony Search Algorithm (HS)*

The HS algorithm, proposed by Geem et al., is a phenomenon-mimicking algorithm inspired by the improvisation process of musicians [23] , which is depicted as Fig. 1. Compared with other heuristic optimization algorithms, it behaves with excellent effectiveness and robustness and presents lots of advantages when applied to optimization problems [24] . The HS algorithm optimization procedures consisting of Steps 1–5 is described in Algorithm 1.

Fig. 1. Algorithm from music phenomenon.

Algorithm 1: The HS algorithm

Step 1. Initialize the optimization problem and algorithm parameters:

Minimize f(x), with $x_i \in X_i$, *i= 1, 2, …N*

where *f*(*x*) is the objective function; *x* is the set of each design variable x_i; X_i is the set of the possible range of values for each design variable; *N* is the number of design variables. In addition, the HS algorithm parameters including harmony memory size (HMS), harmony memory considering rate (HMCR), pitch adjusting rate (PAR), the lower bounds (Lb) and upper bounds (Ub) for each decision variable and termination criterion should also be specified in this step.

Step 2. Initialize the Harmony Memory (HM).

The HM is a location storing all the solution vectors. In this step, the HM matrix is filled with randomly generated solution vectors and sorted by the values of the objective function *f*(*x*).

Step 3. Improvise a new harmony from the HM.

A new harmony vector is generated based on three rules: memory consideration, pitch adjustment and random selection.

Step 4. Update the HM.

On condition that the new harmony vector showed better fitness function than the worst harmony in the HM, the new harmony is included in the HM and the existing worst harmony is excluded from the HM.

Step 5. Repeat steps 3 and 4 until the termination criterion is satisfied.

International Journal of Scientific Engineering and Science
Volume 3, Issue 1, pp. 6-12, 2019. ISSN (Online): 2456-7361

III. A FORECASTING MODEL BASED ON THE FUZZY TIME SERIES AND HARMONY SEARCH

A. *Forecasted Model Based on the High–Order FRGs*

In the section, to verify the effectiveness of the proposed model, the annual data to represent the average rice production (*thousand ton/ year*) of Viet Nam from 1990 to 2010 is shown in Table I, in which it taken from the site www.gso.gov.vn, more precisely from https://www.gso.gov.vn/default.aspx?tabid=717 is used to illustrate the first - order fuzzy time series forecasting process. The step-wise procedure of the proposed model is detailed as follows:

TABLE I. The annual data of the average rice production (***thousand ton/ year***) of Viet Nam.

Year	Actual rice data	Year	Actual rice data	Year	Actual rice data
1990	19225.1	1997	27523.9	2004	36148.9
1991	19621.9	1998	29145.5	2005	35832.9
1992	21590.4	1999	31393.8	2006	35849.5
1993	22836.5	2000	32529.5	2007	35942.7
1994	23528.2	2001	32108.4	2008	38729.8
1995	24963.7	2002	34447.2	2009	38950.2
1996	26396.7	2003	34568.8	2010	39988.9

Step 1: Define the universe of discourse U

Assume Y(t) be the historical data of rice production at year t ($1990 \leq t \leq 2010$). The university of discourse is defined as $U = [D_{min}, D_{max}]$. In order to ensure the forecasting values bounded in the universe of discourse U, we set $D_{min} = I_{min} - N_1$; $D_{max} = I_{max} + N_2$; where I_{min}, I_{max} are the minimum and maximum data of $Y(t)$; N_1 and N_2 are two proper positive to tune the lower bound and upper bound of the U. From the historical rice production data are shown in Table I, we obtain $I_{min} = 19225.1$ và $I_{max} = 39988.9$. Thus, the universe of discourse is defined as U= $[I_{min} - N_1, I_{max} + N_2]$= [19000, 40000] with $N_1 = 225.1$ and $N_2 = 11.1$

Step 2: Partition U into equal length intervals

Divide U into equal length intervals. Compared to the previous models in [4] , [17] , we divide U into seven intervals, $u_1, u_2, \ldots, u_7$, respectively. The length of each interval is $d = \frac{D_{max} - D_{min}}{7} = \frac{40000-19000}{7} = 3000$. Thus, the seven intervals are defined as follows:

u_i = (D_{min}+ (i-1)*d, D_{min} + i *d], with ($1 \leq i \leq 7$) gets seven intervals as:

u_1 = (19000, 22000], u_2 = (22000, 25000], …, u_6 = (34000, 37000], u_7 = (37000, 40000].

Step 3: Define the fuzzy sets for observation of rice production

Each interval in Step 2 represents a linguistic variable of "rice production". For seven intervals, there are seven linguistic values which are A_1 = "very poor rice production", A_2="poor rice production", A_3 = "above poor rice production", A_4 = "average rice production", A_5 = "above average rice production", A_6 = "good rice production", and A_7= "very good rice production" to represent different regions in the universe of discourse on U, respectively. Each linguistic variable represents a fuzzy set A_i and its definitions is described Eq.(1) as follows.

$$\begin{aligned} A_1 &= \frac{1}{u_1} + \frac{0.5}{u_2} + \frac{0}{u_3} + \cdots + \frac{0}{u_7} \\ A_2 &= \frac{0.5}{u_1} + \frac{1}{u_2} + \frac{0.5}{u_3} + \cdots + \frac{0}{u_7} \\ &\text{---------------------------------} \\ A_7 &= \frac{0}{u_1} + \frac{0}{u_2} + \cdots + \frac{0.5}{u_6} + \frac{1}{u_7} \end{aligned} \quad (1)$$

For simplicity, the membership values of fuzzy set A_i either are 0, 0.5 or 1. The value 0, 0.5 and 1 indicate the grade of membership of u_j ($1 \leq j \leq 7$). in the fuzzy set A_i ($1 \leq i \leq 7$). Where, where the symbol '+' denotes fuzzy set union, the symbol '/' denotes the membership of u_j which belongs to A_i.

Step 4: Fuzzify all historical data of rice production

To fuzzify all historical data, it's firstly necessary to assign a corresponding linguistic value to each interval. The simplest way is to assign the linguistic value with respect to the corresponding fuzzy set that each interval belongs to with the highest membership degree**.** For example, the historical rice data of year 1990 is 19225.1, and it belongs to interval u_1 because 19225.1 is within (19000, 22000]. So, we then assign the linguistic value ''very poor rice production" (eg. the fuzzy set A_1) corresponding to interval u_1 to it. Consider two time serials data $Y(t)$ and $F(t)$ at year t, where $Y(t)$ is actual data and $F(t)$ is the fuzzy set of $Y(t)$. According to formula (1), the fuzzy set A_1 has the maximum membership value at the interval u_1. Therefore, the historical data time series on date $Y(1990)$ is fuzzified to A_1. The completed fuzzified results of rice production are listed in Table II.

Step 5: Define all γ – order fuzzy logical relationships.

Based on Definition 2. To establish a γ - order fuzzy relationship, we should find out any relationship which has the $F(t-\gamma), F(t-\gamma+1), \ldots, F(t-1) \rightarrow F(t)$, where $F(t-\gamma), F(t-\gamma+1), \ldots, F(t-1)$ and $F(t)$ are called the current state and the next state of fuzzy logical relationship, respectively. Then a γ - order fuzzy logic relationship in the training phase is got by replacing the corresponding linguistic values.

TABLE II. The results of fuzzification for rice production data.

Year	Actual data	Fuzzy set	Membership value
1990	19225.1	A1	[**1** 0.5 0 0 0 0 0]
1991	19621.9	A1	[**1** 0.5 0 0 0 0 0]
1992	21590.4	A1	[**1** 0.5 0 0 0 0 0]
1993	22836.5	A2	[0.5 **1** 0.5 0 0 0 0]
1994	23528.2	A2	[0.5 **1** 0.5 0 0 0 0]
1995	24963.7	A2	[0.5 **1** 0.5 0 0 0 0]
1996	26396.7	A3	[0 0.5 **1** 0.5 0 0 0]
1997	27523.9	A3	[0 0.5 **1** 0.5 0 0 0]
1998	29145.5	A4	[0 0 0.5 **1** 0.5 0 0]
1999	31393.8	A5	[0 0 0 0.5 **1** 0.5 0]
2000	32529.5	A5	[0 0 0 0.5 **1** 0.5 0]
2001	32108.4	A5	[0 0 0 0.5 **1** 0.5 0]
2002	34447.2	A6	[0 0 0 0 0.5 **1** 0.5]
2003	34568.8	A6	[0 0 0 0 0.5 **1** 0.5]
2004	36148.9	A6	[0 0 0 0 0.5 **1** 0.5]
2005	35832.9	A6	[0 0 0 0 0.5 **1** 0.5]
2006	35849.5	A6	[0 0 0 0 0.5 **1** 0.5]
2007	35942.7	A6	[0 0 0 0 0.5 **1** 0.5]
2008	38729.8	A7	[0 0 0 0 0 0.5 **1**]
2009	38950.2	A7	[0 0 0 0 0 0.5 **1**]
2010	39988.9	A7	[0 0 0 0 0 0.5 **1**]

For example, supposed $\gamma = 3$ from Table III, a fuzzy relation $A_1, A_1 \rightarrow A_1$ is got as $F(1990), F(1991) \rightarrow F(1992)$.

So on, we get the 3rd-order fuzzy relationships are shown in Table III, where there are 19 relationships; the first 18 relationships are called the trained patterns, and the last one , is called the untrained pattern (in the testing phase). For the untrained pattern, relation 17 has the fuzzy relation A7, A7, A7 → # as it is created by the relation $F(2008), F(2009), F(2010) \rightarrow F(2011)$, since the linguistic value of $F(2011)$is unknown within the historical data, and this unknown next state is denoted by the symbol '#'

TABLE III. The 3rd- order fuzzy logical relationships.

No	Fuzzy relations	No	Fuzzy relations
1	A1, A1, A1→ A2	11	A5, A5, A6 → A6
2	A1, A1, A2 → A2	12	A5, A6, A6 → A6
3	A1, A2, A2 → A2	13	A6, A6, A6 → A6
4	A2, A2, A2 → A3	14	A6, A6, A6 → A6
5	A2, A2, A3 → A3	15	A6, A6, A6 → A6
6	A2, A3, A3 → A4	16	A6, A6, A6 → A7
7	A3, A3, A4 → A5	17	A6, A6, A7 → A7
8	A3, A4, A5 → A5	18	A6, A7, A7 → A7
9	A4, A5, A5 → A5		A7, A7, A7 → #
10	A5, A5, A5 → A6		

Step 6: Establish all γ – order fuzzy logical relationships groups

Based on [4] all the fuzzy relationships having the same fuzzy set on the left-hand side or the same current state can be put together into one fuzzy relationship group. The fuzzy logical relationship as the same are counted only once. Thus, from Table III and based on Definition 4, we can obtain seven 3rd – order fuzzy logical relationship groups shown in Table IV.

TABLE IV. The complete result of the 3rd - order fuzzy relationship groups.

No group	Fuzzy relation groups	No group	Fuzzy relation groups
1	A1, A1, A1 → A2	9	A4, A5, A5 → A5
2	A1, A1, A2 → A2	10	A5, A5, A5 → A6
3	A1, A2, A2 → A2	11	A5, A5, A6 → A6
4	A2, A2, A2 → A3	12	A5, A6, A6 → A6
5	A2, A2, A3 → A3	13	A6, A6, A6 → A6, A6, A6, A7
6	A2, A3, A3 → A4	14	A6, A6, A7 → A7
7	A3, A3, A4 → A5	15	A6, A7, A7 → A7
8	A3, A4, A5 → A5	16	A7, A7, A7 → #

Step 7. Calculate and defuzzify the forecasted values

In order to calculate the forecast value for all high – order fuzzy relationship groups, we use [25] for the trained patterns in the training phase and use [17] for the untrained patterns in the testing phase.

For the training phase, based on [25] , we create all forecast output values for all high – order FRGs based on fuzzy sets on the right-hand or next state within the same group. For each group of 3rd – order fuzzy relationship in Table IV, we divide each corresponding interval of each next state into three sub-regions with equal size, and create a forecasted value for each group according to Eq. (2).

$$\text{forecasted}_{\text{output}} = \frac{1}{n}\sum_{k=1}^{n}\frac{(m_k + \text{submid}_k)}{2} \quad (2)$$

Where, n is the total number of next states within the same group.

✓ m_k is the midpoint of interval u_k corresponding to k-th fuzzy set on the right-hand side where the highest level of fuzzy set A_k takes place in these intervals, u_k.

✓ $submid_k$ is the midpoint of one of three sub-regions in which it has the historical data belong to this sub-interval corresponding to k-th fuzzy set on the right-hand side where the highest level of A_k takes place in this interval.

For the testing phase, we calculate forecasted value for a group which contains the unknown linguistic value based on [17] according to Eq.(3), where the symbol w_h means the highest votes predefined by user, m is the order of the fuzzy relationship, the symbols M_{t1} and M_{ti} denote the midpoints of the corresponding intervals of the latest past and other past linguistic values in the current state. From Table IV, it can be shown that group 16 has the fuzzy relationship A_7, A_7, A_7 , → # as it is created by the fuzzy relation $F(2008), F(2009), F(2010) \rightarrow F(2011)$; since the linguistic value of $F(2011)$ is unknown within the historical data, and this unknown next state is denoted by the symbol '#'.

$$Forecasted_{for\#} = \frac{(M_{t1} * w_h) + M_{t2} + \cdots + M_{ti} + \cdots + M_{t\gamma}}{w_h + (\gamma - 1)}; \text{ with } (1 \le i \le \gamma) \quad (3)$$

From forecasted Eqs.(2) & (3) above and based on IV &II, we complete forecasted results rice production of Viet Nam for all years the period from 1990 to 2011 based on 3rd- order FTS model with seven intervals are listed in Table V.

TABLE V. The complete forecasted outputs for rice production of Viet Nam based on the 3rd– order FTS model.

Year	Actual data	Fuzzy set	*Forecasted value*
1990	19225.1	A1	-----
1991	19621.9	A1	-----
1992	21590.4	A1	----
1993	22836.5	A2	23000
1994	23528.2	A2	23500
1995	24963.7	A2	24000
1996	26396.7	A3	26500
1997	27523.9	A3	27000
1998	29145.5	A4	29500
1999	31393.8	A5	32000
2000	32529.5	A5	32500
2001	32108.4	A5	32500
2002	34447.2	A6	35000
2003	34568.8	A6	35000
2004	36148.9	A6	36000
2005	35832.9	A6	36250
2006	35849.5	A6	36250
2007	35942.7	A6	36250
2008	38729.8	A7	36250
2009	38950.2	A7	38500
2010	39988.9	A7	39000
2011	N/A	N/A	38500

The performance of proposed model can be assessed by comparing the difference between the forecasted values and the actual values. The widely used indicators in time series models comparisons are the root mean square error (RMSE) according to Eq.(4) as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=\gamma}^{n}(F_t - R_t)^2} \quad (4)$$

Where, R_t denotes actual value at year t, F_t is forecasted value at year t, n is number of the forecasted data, γ is order of the fuzzy logical relationships

B. *Forecasting Model Combined the High – Order FTS and HS Algorithm*

To improve forecasted accuracy of the proposed, the effective lengths of intervals is main issue presented in this paper. A novel method for forecasting rice production is developed by HS algorithm to adjust the length each of intervals in the universe of discourse without increasing the number of intervals.

In proposed model, each Harmony exploits the intervals in the universe of discourse of historical data Y(t). Let the number of the intervals be n, the lower bound and the upper bound of the universe of discourse on historical data Y(t) be x_0 and x_n, respectively. Each harmony id is a vector consisting of n-1 elements x_k where $1 \le k \le n-1$ and $x_k \le x_{k+1}$. Based on these n-1 elements, define the n intervals as $u_1 = [x_0, x_1]$, $u_2 = [x_1, x_2]$,..., $u_i = [x_{k-1}, x_k]$,… and $u_n = [x_{n-1}, x_n]$, respectively. When a particle moves to a new position, the elements of the corresponding new vector need to be sorted to ensure that each element x_k $(1 \le k \le n-1)$ arranges in an ascending order.

- ✓ The parameters of the harmony search algorithm that are supposed to be defined in this section are harmony memory size (HMS), i.e., the number of solution vectors or rows in the harmony memory matrix; harmony memory considering rate (HMCR); pitch adjusting rate (PAR); bandwidth (*BW*); and the number of iterations [23].
- ✓ The algorithm generates random solution vectors (harmonies) HMS times and puts them in the HM matrix, specified by Eq.(5):
- ✓ Improvise a new harmony: Once the harmony memory

matrix is initialized, the algorithm starts the first iteration by improvising a new harmony. $X = (x_1, x_2, \ldots, x_{n-1})$ is a new solution vector that is constructed based on three rules:

(1) **Memory** consideration with probability HMCR
(2) **Pitch** adjusting with probability PAR
(3) **Random** selection with probability (1-HMCR)

If the harmony search algorithm decides to adjust the pitch to the new selected value by considering the PAR parameter, X_i is altered by Eq.(11).

$$\text{HM} = \begin{bmatrix} x_1^1 & \cdots & x_N^1 \\ \vdots & \ddots & \vdots \\ x_1^{HMS} & \cdots & x_N^{HMs} \end{bmatrix}_{HMS \times N} \quad (5)$$

$$\mathbf{X}_i^{New} = \begin{Bmatrix} x_i(k) \in (x_i^1, x_i^2, \ldots, x_i^k)\ if\ P_{random} = 1 - HMCR \\ x_i(k) \in (x_i^1, x_i^2, \ldots, x_i^{HMS})\ if\ P_{memory} = HMCR \times (1 - PAR) \\ x_1 \pm (rand() \times Bw\ f\ P_{Pitch} = HMCR \times PAR \end{Bmatrix} \quad (6)$$

- ✓ Update the harmony memory: regarding the value of the objective function if the new harmony vector is better than the worst harmony in the HM, the new harmony is entered in the HM and the worst harmony is omitted from it. The complete steps of the proposed model are presented in Algorithm 2.

Algorithm 2: The FTS-HS algorithm

1. **Initialize** all parameters as follows:
   - ✓ HMS = 30, HMCR = 0.99, PAR = 0.5, BW = 1.
   - ✓ Lb = x_0 = 19000; Ub = x_n = 40000
   - ✓ maximum number of improvisations is **100**
   - ✓ Harmonies are Initialized $x_0 + Rand(\,) * (x_n - x_0)$;
2. **While** the stop condition (maximum number of improvisations or minimum RMSE criteria) is not satisfied **do**
   2.1. For Harmony id, $(1 \le id \le \text{N})$ **do**
   - ✓ Define linguistic values according to all intervals defined by Harmony id
   - ✓ Fuzzify all historical data by Step 4 in Subsection 3.1
   - ✓ Create all $\gamma -$ order fuzzy relationships by Step 5 in Subsection 3.1
   - ✓ Make all $\gamma -$ order fuzzy relationship groups by Step 6 in Subsection 3.1
   - ✓ Calculate forecasting values by Step 7 in Subsection 3.1
   - ✓ Compute the F(X) = RMSE values for Harmony id based on Eq.(4)
   - ✓ Update HM (5) id according to the RMSE values mentioned above.
   - ✓ Update the RMSE = $\text{F(X)}_{\text{HMS} \times 1}$values

   end for

end while

IV. Experimental Results

In this paper, we apply the proposed model to forecast the rice production of Viet Nam with the whole historical data the period from 1990 to 2010 is listed in Table I and we also the proposed model to handle other forecasting problems, such as the empirical data for the enrolments of University of Alabama [4] from 1971 to 1992 are used to perform comparative study in the training phase.

A. *Experimental results for Forecasting Rice Production of Viet Nam*

In this section, we apply the proposed method for forecasting the rice production from 1990 to 2010 are listed in Table I. Our proposed model is executed 20 independent runs for each order, and the best result of runs at each order is taken to be the final result. During simulation with parameters are expressed in algorithm 2, the number of intervals is kept fix for the proposed model. The forecasted accuracy of the proposed method is estimated using the RMSE (4). The forecasted results of proposed model under number of interval as 14 and various orders are listed in Table VI.

TABLE VI. The completed forecasting results for rice production data of Viet Nam under deferent orders of FTS.

Year	**Actual data**	***Forecasted results*** 2nd-order	3rd-order	4th-order	5th-order
1990	19225.1	-----	-----	-----	-----
1991	19621.9	-----	-----	------	------
1992	21590.4	21613	-----	-----	------
1993	22836.5	22872.3	22883.8	----	------
1994	23528.2	823428.6	23458.8	23502.3	-----
1995	24963.7	24963.3	24969.8	24992	24958.5
1996	26396.7	26459	26522.5	26392.8	26391
1997	27523.9	27536.3	27355.5	27480.8	27537.3
1998	29145.5	29174.3	29250.8	29225.8	29161.8
1999	31393.8	31367.3	31269.5	31259.8	31360
2000	32529.5	32493.5	32509.8	32548.3	32497.23
2001	32108.4	32220.5	32255.8	32220.3	32178.3
2002	34447.2	34441.5	34439.6	34309	34517.3
2003	34568.8	34441.5	34439.6	34567.5	345173.3
2004	36148.9	36190	36144.3	36141	36125.5
2005	35832.9	35675.6	35871.8	35877	35852.5
2006	35849.5	35675.75	35871.8	35877	35852.5
2007	35942.7	36190	35933.3	35877	35989

International Journal of Scientific Engineering and Science
Volume 3, Issue 1, pp. 6-12, 2019. ISSN (Online): 2456-7361

2008	38729.8	38727.8	38885.8	38668	38768.8
2009	38950.2	38727.8	38885.8	38949.5	38983.8
2010	39988.9	39893.8	39858	39837.5	39875.8
2011	N/A	39794.7	39680.9	39419.6	39323.6
RMSE		***108.49***	***95.57***	***74.19***	***45.86***

From Table VI, it can be seen that the performance of the proposed model is improved a lot with increasing number of orders of fuzzy relationship with the same number of interval. Particularly, the proposed model gets the lowest RMSE value of ***45.86*** with 5th-order fuzzy logical relationship. This means that the high – order FTS models are more suitable than first-order models in dealing with linguistic values.

B. *Experimental Results for Forecasting Enrolments*

In order to verify the forecasting effectiveness of the proposed model under different number of intervals and different high - order FLRGs , five FTS models, named C02 in [6] , CC06b [9] , S07s [25] , HPSO [17] are examined and compared. The forecasted accuracy of the proposed method is estimated by using the RMSE (4). A comparison of the forecasting accuracy with various orders and different number of intervals among the C02, CC06b, S07s, HPSO models and the proposed model are listed in Table VII.

TABLE VII. A comparison of the forecasted results of the proposed method with the existing models based on high – order of the fuzzy time series under different number of intervals.

Years	**Actual data**	**S07s**	**C02**	**CC06b**	HPSO	**Our model**
1971	13055	N/A	N/A	N/A	N/A	
1972	13563	N/A	N/A	N/A	N/A	
1973	13867	N/A	N/A	N/A	N/A	
1974	14696	N/A	N/A	N/A	N/A	
1975	15460	15500	N/A	N/A	N/A	
1976	15311	15468	15500	N/A	N/A	
1977	15603	15512	15500	N/A	N/A	
1978	15861	15582	15500	N/A	N/A	15877
1979	16807	16500	16500	16846	N/A	16836
1980	16919	16361	16500	16846	16890	16910
1981	16388	16362	16500	16420	16395	16385
1982	15433	15744	15500	15462	15434	15442
1983	15497	15560	15500	15462	15505	15482
1984	15145	15498	15500	15153	15153	15153
1985	15163	15306	15500	15153	15153	15153
1986	15984	15442	15500	15977	15971	15970
1987	16859	16558	16500	16846	16890	16836
1988	18150	17187	18500	18133	18124	18151
1989	18970	18475	18500	18910	18971	18957
1990	19328	19382	19500	19334	19337	19328
1991	19337	19487	19500	19334	19337	19328
1992	18876	18744	18500	18910	18882	18885
RMSE		365.65	294.44	33.18	15.3	13.53

From Table VI, it is obvious that our model has a MSE value is 13.53 which is the lowest among all forecasting models compared.

In addition, the proposed model is also compared with various high –order forecasting model under seven intervals such C02 [6] model, CC06 [9] model and HPSO [17] model. The details of comparison are shown in Table VIII. The forecasting trend according to order of model is depicted in Fig. 2 for clearer illustration.

TABLE VIII. A comparison of the RMSE value between our model and C02 model, CC06b model, HPSO model under different number of orders and the number of interval is 7.

Models	**Number of orders**				
	2	**3**	**4**	**5**	**6**
C02	298.48	294.44	298.96	307.47	313.39
CC06b	260.45	176.42	178.91	158.06	164.26
HPSO	259.08	177.89	152.55	153.41	153.85
Our model	206.52	75.27	59.26	60.43	56.06

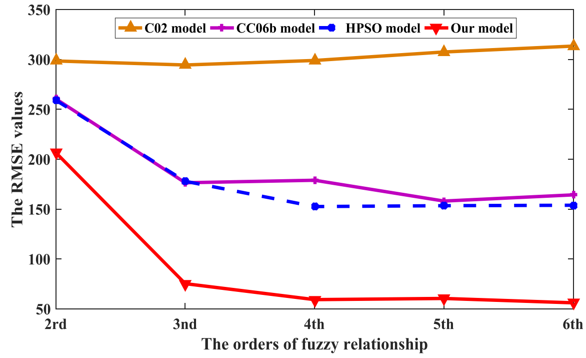

Fig. 2. A comparison of the RMSE values for 7 intervals with different high-order fuzzy relationships

From Table VII, it can be seen that the accuracy of the proposed model is improved significantly. Particularly, our model gets the lower RMSE values than three models presented in C02 [6] and CC06 [9] and HPSO [17] . These finding suggest that the proposed model is able to provide effective forecasting capability for the high – order FTS model with different number of orders for the same number of interval. The graphical comparison clearly shows that the forecasting accuracy of the proposed model is more precise than those of existing models with the different number of orders.

V. CONCLUSION

In this paper, a hybrid forecasting model based on combining fuzzy time series and Harmony search algorithm to forecast the rice production of Viet Nam and Actual enrollments of the University of Alabama. The main contributions of this paper is the applying HS algorithm to for optimizing lengths of intervals in the universe of discourse. From the performance comparison in Tables VII, VIII and Fig.2 the author shows the proposed model outperforms previous forecasting models for the training phase with various high - orders. Although this study shows the superior forecasting capability compared with the existing forecasting models, the proposed model is only tested by two problems: enrolments data and rice production dataset. To continue improving the effectiveness of the forecasting model, there are some suggestions for future research: Firstly, author can apply proposed model to deal with more complicated real-world problems for decision-making such as weather forecast, traffic accident prediction, pollution forecasting and etc. Secondly, author can combine the forecasted model with more intelligent algorithms to build a new forecasting model with the aim of achieving the best possible forecasting performance. That will be the future work of this study.

Volume 3, Issue 1, pp. 6-12, 2019. ISSN (Online): 2456-7361

REFERENCES

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

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

[3] Q. Song and B. S. Chissom, "Forecasting enrolments with fuzzy time series - part II," *Fuzzy Sets and System*, vol. 62, pp. 1-8, 1994.

[4] S. M. Chen, "Forecasting enrolments based on fuzzy time series," *Fuzzy Set and Systems*, vol. 81, pp. 311-319, 1996.

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

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

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

[8] S.-M. Chen and N.-Y. Chung, "Forecasting enrollments using high-order fuzzy time series and genetic algorithms: Research Articles," *International Journal of Intelligent Systems*, vol. 21, pp. 485–501, 2006.

[9] S.-M. Chen, and N.-Y. Chung, "Forecasting enrollments of students by using fuzzy time series and genetic algorithms," *International Journal of Intelligent Systems*, vol. 17, pp. 1–17, 2006.

[10] S. R. Singh, "A simple method of forecasting based on fuzzy time series," *Applied Mathematics and Computation*, vol. 186, pp. 330–339, 2007.

[11] S. R. Singh, "A simple method of forecasting based on fuzzy time series," *Applied Mathematics and Computation*, vol. 186, pp. 330–339, 2007.

[12] T.-L. Chen, C.-H. Cheng, and H.-J. Teoh, "High-order fuzzy time-series based on multi-period adaptation model for forecasting stock markets," *Physical A: Statistical Mechanics and its Applications*, vol. 387, pp. 876–888, 2008.

[13] Chu, H.-H., Chen, T.-L., Cheng, C.-H., & Huang, C.-C. . "Fuzzy dual-factor time- series for stock index forecasting," *Expert Systems with Applications*, vol. 36, pp. 165–171, 2009.

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

[15] Ling-Yuan Hsu et al. "Temperature prediction and TAIFEX forecasting based on fuzzy relationships and MTPSO techniques," *Expert Syst. Appl.*, vol. 37, pp. 2756–2770, 2010.

[16] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338-353, 1965.

[17] I. H. Kuo, et al., "An improved method for forecasting enrolments based on fuzzy time series and particle swarm optimization," *Expert Systems with Applications*, vol. 36, pp. 6108–6117, 2009.

[18] I-H. Kuo, S.-J. Horng, Y.-H. Chen, R.-S. Run, T.-W. Kao, R.-J. C., J.-L. Lai, and T.-L. Lin, "Forecasting TAIFEX based on fuzzy time series and particle swarm optimization," *Expert Systems with Applications*, vol. 2, no. 37, pp. 1494–1502, 2010.

[19] L. W. Lee, L. H. Wang, S. M. Chen, and Y. H. Leu, "Handling forecasting problems based on two-factors high-order fuzzy time series," *IEEE Transactions on Fuzzy Systems*, vol. 14, pp. 468–477, 2006.

[20] Ling-Yuan Hsu et al. "Temperature prediction and TAIFEX forecasting based on fuzzy relationships and MTPSO techniques," *Expert Syst. Appl*., vol. 37, 2010, pp. 2756–2770.

[21] S. Pritpal and B. Bhogeswar, "High-order fuzzy-neuro expert system for time series forecasting," *Knowl.-Based Syst*., vol. 46, pp. 12–21, 2013.

[22] L. Al - Mataneh, A. Senta, S. Bani –Ahmad, J. Alshaer and Al- qily. "Development of temperature based weather forecasting model using Neural network and fuzzy logic," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 9, no. 12, pp. 343-366, 2014.

[23] Z. W. Geem, J. H. Kim, and G. V. Loganathan, "A new heuristic optimization algorithm: Harmony search," *Simulation*, vol. 76, no. 2, pp. 60–68, 2001.

[24] M. Fesanghary, M. Mahdavi, M. Minary, and Y. Alizadeh, "Hybridizing harmony search algorithm with sequential quadratic programming for engineering optimization problems," *Comput. Methods Appl. Mech. Energy*, vol. 197, pp. 3080–3091, 2008.

[25] I-H. Kuo, S.-J. Horng, Y.-H. Chen, R.-S. Run, T.-W. Kao, R.-J. C., J.-L. Lai, and T.-L. Lin, "Forecasting TAIFEX based on fuzzy time series and particle swarm optimization," *Expert Systems with Applications*, vol. 2, no. 37, pp. 1494–1502, 2010.

[26] S. R. Singh, "A simple method of forecasting based on fuzzy time series," *Applied Mathematics and Computation*, vol. 186, pp. 330–339, 2007.