American Journal of Information Science and Computer Engineering
Vol. 5, No. 1, 2019, pp. 8-16
http://www.aiscience.org/journal/ajisce
ISSN: 2381-7488 (Print); ISSN: 2381-7496 (Online)

A Novel Forecasting Model Combining Fuzzy Time Series with Harmony Search

Nghiem Van Tinh*

Faculty of Electronics, Thai Nguyen University of Technology, Thai Nguyen University, Thai Nguyen, Viet Nam

Abstract

Fuzzy time series (FTS) models have been proposed to model time series data and have been extended to model numerical data in solving nonlinear and complexity problems. Many factors are believed to affect forecasting accuracy of FTS model. The design of forecasting rules and the lengths of intervals for observations are considered two of them. Therefore, how to cover both issues simultaneously is important for the improvement of forecasting results. In this paper, the development of forecasting model using the harmony search algorithm and fuzzy time series is introduced. Firstly, a forecasting model is constructed from the fuzzy logical relationship and calculate forecasting output by new defuzzification rule. 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 a design to increase forecasting accuracy. In the part of empirical analysis, numerical dataset of Gas online price of Viet Nam is utilized to illustrate the forecasting process and numbers of enrolment of Alabama University is used to compare between proposed model and it's several counterparts. The application results accentuate the superiority of the proposed model compared to the other models based on high -order FTS.

Keywords

Forecasting, Fuzzy Time Series, Harmony Search, Enrolments, Gas Online Price

Received: December 31, 2018 / Accepted: February 6, 2019 / Published online: March 5, 2019

1. Introduction

The forecasting of future events of time series has always influence people from the ancient times. Therefore, many researchers have been proposed different forecasting models to deal with various domain problems such as: enrolments forecasting [1-10], crop production prediction [6, 11], stock markets prediction [5, 12-14] and temperature forecasting [14, 15]. There is the matter of fact that the traditional forecasting models such as regression analysis, moving aver-age, autoregressive moving average and ARIMA model cannot deal with the forecasting problems in which the historical data are represented by linguistic values. Fuzzy set theory was firstly presented by Zadeh to handle problems with linguistic values. The concepts of fuzzy sets have been successfully applied to time series by Song and Chissom. They introduced both the time-invariant fuzzy time series [7] and the time-variant time series [8] 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 [2] proposed the first-order fuzzy time series model by using simple arithmetic calculations instead of max-min composition operations [7-9] 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 [3] proposed a new forecast model based on the high-order fuzzy time series to forecast the enrolments of University of Alabama. Yu [10] presented a new model which can refine the lengths of

* Corresponding author
E-mail address: nghiemvantinh@tnut.edu.vn

intervals during the formulation of fuzzy relationships and hence capture the fuzzy relationships more appropriately. Both the stock index and enrolment are used as the targets in the empirical analysis. Chen & Chung [1, 4] presented the first-order and high-order fuzzy time series model to deal with forecasting problems based on genetic algorithms. Singh [6, 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. Recently, Particle swarm optimization technique has been successfully applied in many applications. Based on Chen's model [2], Kuo et al. [16] developed a new hybrid forecasting model which combined fuzzy time series with PSO algorithm to find the proper length of each interval. Following, they continued to present a new forecast method to solve the TAIFEX forecasting problem based on fuzzy time series and PSO algorithm [17]. Some other authors who proposed some methods for the temperature prediction and the TAIFEX forecasting, based on two-factor fuzzy logical relationships [15] and use them in which combine with PSO algorithm in fuzzy time series. In addition, some study works using other hybrid techniques can be found such as: Pritpal and Bhogeswar [18] presented a new model based on hybridization of fuzzy time series theory with artificial neural network (ANN). Matarnenh et al. [19] use feed forward artificial neural network and fuzzy logic for weather forecasting achieve better results. Article in [20] introduced a high-order multi-variable model based on FTS to deal with the problem of intervals and their lengths. Sun et al.,[21] design a multivariate FTS model with multiple factors. In order to simplify the calculation and refine the evolved rules, this model integrates rough set theory into the model.

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 models were implemented for forecasting of other historical data and not Gas online price. In this paper, a forecasting model based on the fuzzy logical relationship groups and Harmony search algorithm is presented to forecast Gas online price 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 applied to esti-mate the forecasting accuracy. Finally, a new hybrid fore-casting model based on combined FTS and HS is developed to adjust the length of each interval in the universe of dis-course by minimizing RMSE value. The case study with the data of Gas online price 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. This paper is organized as follows:

After section introduction, a brief review of the basic concepts of FTS and Harmony search algorithms are introduced in Section 2. Section 3, describes the details of FTS forecasting model for forecasting Gas online price and then combines with the HS algorithm to search the effective lengths of intervals in the universe of discourse. Section 4 compares the forecasting results of the proposed method with its counterparts from the enrolments data of the University of Alabama. Finally, Section 5 offers some conclusions.

2. Fuzzy Time Series Concepts and Harmony Search

2.1. Basic Concepts of Fuzzy Time Series

Conventional time series refer to real values, but fuzzy time series are structured by fuzzy sets. Let U = $\{u_1, u_2, \dots, u_n\}$ be an universal set; a fuzzy set A of U is defined as A = $\{ f_A(u_1)/u_1 +, f_A(u_2)/u_2 \dots + f_A(u_n)/u_n \}$, where fA is the membership function of A, f_A:U→[0,1], $f_A(u_i)$ is the degree of membership of the element ui in the fuzzy set A. Here, the symbol ''+'' indicates the operation of union and the symbol ''/'' indicates the separator rather than the commonly used summation and division in algebra, respectively. The definition of FTS is briefly reviewed as follows:

Definition 1: Fuzzy time series [7, 8]

Let Y(t) (t =..,0,1,2..), a subset of R, be the universe of discourse on which fuzzy sets f_i(t) (i = 1,2…) are defined and if F(t) is a collection of f_1(t), f_2(t),…, 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) [7, 8]

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) →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_k . Hence, the relationship between F(t) and F(t -1) is denoted by fuzzy logical relationship $A_i \rightarrow A_k$ where Ai and A_k 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: α - order fuzzy logical relationships [3, 8]

Let F(t) be a fuzzy time series. If F(t) is caused by F(t-1), F(t-2),…, F(t-α+1), F(t-α) then this fuzzy relationship is represented by F(t-α),…, F(t-2), F(t-1)→ F(t) and is called an m - order fuzzy time series.

Definition 4: Fuzzy relationship group (FRG) [2]

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:

$$A_i \rightarrow A_{k1},\ A_i \rightarrow A_{k2}, \ldots, A_i \rightarrow A_{km}$$

In previous study was proposed by Chen [4], 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: $A_i \rightarrow A_{k1}, A_{k2}, \ldots, A_{km}$

2.2. Harmony Search Algorithm (HS)

Using a fundamental harmony search approach, an n-dimensional real vector, called harmony, shows each solution [22]. First, a starting set of harmony vectors is randomly generated and assigned in a harmony memory (HM). The next step is a new harmony generation that considers all of the present harmonies in the HM. This goal can be achieved by deploying a harmony memory consideration rate (HMCR) aligned with a pitch adjustment rule. The procedure continues with comparing newly generated harmonies to the existing harmonies. The replacement condition is satisfied if the newly generated harmony has a better state [22]. The algorithm iterates until it satisfies a defined termination criterion condition. Compared with other heuristic optimization algorithms, it behaves with excellent effectiveness and robustness and presents lots of advantages when applied to optimization problems [23]. The five fundamental steps of the algorithm are described as follows:

Step 1. Initialize the optimization problem and algorithm parameters:

$$\text{To minimize the objective function } f(x) \tag{1}$$

subject to $x_i \in X_i$, 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.

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).

HM matrix is filled with as many randomly generated solution vectors as the HMS and sorted by the values of the objective function f(x), as follows:

$$HM = \begin{bmatrix} x_1^1 & \cdots & x_N^1 \\ \vdots & \ddots & \vdots \\ x_1^{HMS} & \cdots & x_N^{HMs} \end{bmatrix}_{HMS \times N} \tag{2}$$

Step 3. Improvise a new harmony from the HM.

A new harmony vector is generated based on three rules: The memory consideration (HMCR), pitch adjustment (PAR) and random selection (1-HMCR) are presented according to formula (3) as follows:

$$x_i^{new} \leftarrow \begin{Bmatrix} x_i(k) \in \left(x_i^1, x_i^2, \ldots, x_i^k\right) \text{ if } P_{random} = 1 - HMCR \\ x_i(k) \in \left(x_i^1, x_i^2, \ldots, x_i^{HMS}\right) \text{ if } P_{memory} = HMCR \times (1 - PAR) \\ x_1 \pm rand(\,) \times Bw \text{ if } P_{Pitch} = HMCR \times PAR \end{Bmatrix} \tag{3}$$

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.

3. A Forecasting Model Combined the FTS with Harmony Search

3.1. Forecasted Model Based on the High-Order FTS

In the section, to verify the effectiveness of the proposed model based on high – order FTS, the monthly data to represent the Gas online price of Viet Nam from 1/4/2018 to 12/21/2018 is listed in Table 1 which it taken from the site https://vnexpress.net/kinh-doanh is used to illustrate the high - order fuzzy time series forecasting process. The step-wise procedure of the proposed model is detailed as follows:

Table 1. The monthly data of the Gas online price (*thousand / liter*) of Viet Nam.

Date	Gas online price	Date	Gasoline price
1/4/2018	18240	7/28/2018	19610
1/19/2018	18670	8/7/2018	19610
2/3/2018	18670	8/22/8/2018	19610
2/21/2018	18340	9/01/2018	19610
3/8/2018	18340	9/6/2018	19910
3/23/2018	18340	9/10/2018	19910
4/7/2018	18930	9/21/2018	20230
4/23/2018	18930	10/06/2018	20906
5/8/2018	19440	10/22/2018	20680
5/23/2018	19940	11/6/2018	19600
6/7/2018	19940	11/21/2018	18620
6/22/2018	19610	12/06/2018	17180
7/7/2018	19610	12/21/2018	18141

Step 1: Define the universe of discourse U

Assume Y(t) be the historical data of Gas online price on month t (1/4/2018 $\le t \le$ 12/21/2018). 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 $D_{min} = I_{min} - N_1$ and $D_{max} = I_{max} + N_2$; 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 Gas online price data is shown in Table 1, we obtain $I_{min} = 17180$ và $I_{max} = 20906$. Thus, the universe of discourse is defined as U= $[I_{min} - N_1, I_{max} + N_2]$= [17000, 21000] with $N_1 = 80$ and $N_2 = 18.806$

Step 2: Partition U into equal length intervals

For convenience of comparison with previous works in [2, 16], [14], we divide U into seven intervals with equal lengths, $u_1, u_2, \dots, u_7$, respectively. The length of each interval is $d = \frac{D_{max} - D_{min}}{7} = \frac{21000 - 17000}{7} = 571.43$. Thus, the seven intervals are defined as follows:

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

u1 = (17000, 17571.43], u2 = (17571.43, 18142.86], …, u6 = (19857.15, 20428.57], u7 = (20.428.57, 21000].

Step 3: Define the fuzzy sets for observation of Gas online price

Each interval in Step 2 represents a linguistic variable of “Gas online price”. For seven intervals, there are seven linguistic values which are A_1= “ Very Low gas price”, A_2= “ Low gas price”, A_3 =“ Average gas price”, A_4 = “Quite High gas price”, A_5 = “High gas price”, A_6 = “ Very High gas price”, and A_7=“ Very Very High gas price” 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 in formula (4) as below.

$$\begin{aligned} A_1 &= \frac{1}{u_1} + \frac{0.5}{u_2} + \frac{0}{u_3} + \cdots + \frac{0}{u_7} \\ A_2 &= \frac{0.5}{u_1} + \frac{1}{u_2} + \frac{0.5}{u_3} + \cdots + \frac{0}{u_7} \\ &- - - - - - - - - - - - \\ A_7 &= \frac{0}{u_1} + \frac{0}{u_2} + \cdots + \frac{0.5}{u_6} + \frac{1}{u_7} \end{aligned} \tag{4}$$

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 $(1 \le j \le 7)$. in the fuzzy set A_i $(1 \le i \le 7)$.

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 Gas online price

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 data of date 1/4/2018 is 18240, and it belongs to interval u_3 because 18240 is within (18142.86, 18714.29]. So, we then assign the linguistic value ‘‘Average gas price” (eg. the fuzzy set A_3) corresponding to interval u_3 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 (4), the fuzzy set A_3 has the maximum membership value at the interval u_3. Therefore, the historical data time series on date Y(1/4/2018) is fuzzified to A_3. The completed fuzzified results of Gas online price are listed in Table 2.

Table 2. The results of fuzzification for Gas online price data.

Date	Gas online price	Fuzzy set	Membership of interval
1/4/2018	18240	A3	[0 0.5 1 0.5 0 0 0]
1/19/2018	18670	A3	[0 0.5 1 0.5 0 0 0]
2/3/2018	18670	A3	[0 0.5 1 0.5 0 0 0]
2/21/2018	18340	A3	[0 0.5 1 0.5 0 0 0]
----------	------	-----	--------------------
9/10/2018	19910	A6	[0 0 0 0 0.5 1 0.5]
9/21/2018	20230	A6	[0 0 0 0 0.5 1 0.5]
10/06/2018	20906	A7	[0 0 0 0 0 0.5 1]
10/22/2018	20680	A7	[0 0 0 0 0 0.5 1]

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

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

For example, suppose that there a 3nd - order fuzzy relation as F(1/4/2018), F(1/19/2018), F(2/3/2018) → F(2/21/2018) occurs at consecutive days in the time series ass 1/4/2018, 1/19/2018, 2/3/2018, 2/21/2018. This relation is replaced by the corresponding linguistic values which is A3, A3, A3 → A3, so on. From Table 2, we obtain all the 3nd-order fuzzy relationships are shown in Table 3, 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 20 has the fuzzy relation A3, A1, A2 → # as it is created by the relation F(11/21/2018), F(12/06/2018), F(12/21/2018) → F(12/22/2018), since the linguistic value of F(12/22/2018) is unknown within the historical data, and this unknown next state is denoted by the symbol ‘#‘

Table 3. The complete results for the 3rd- order fuzzy logical relationships.

No	Fuzzy relations	No	Fuzzy relations
1	A3, A3, A3 -> A3	11	A5, A5, A5 -> A6
2	A3, A3, A3 -> A4	12	A5, A5, A6 -> A6
3	A3, A3, A4 -> A4	13	A5, A6, A6 -> A6
4	A3, A4, A4 -> A5	14	A6, A6, A6 -> A7
5	A4, A4, A5 -> A6	15	A6, A6, A7 -> A7
6	A4, A5, A6 -> A6	16	A6, A7, A7 -> A5
7	A5, A6, A6 -> A5	17	A7, A7, A5 -> A3
8	A6, A6, A5 -> A5	18	A7, A5, A3 -> A1
9	A6, A5, A5 -> A5	19	A5, A3, A1 -> A2
10	A5, A5, A5 -> A5	20	A3, A1, A2 -> #

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

Based on [2] 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 3 and based on Definition 4, we can obtain seven 3rd – order fuzzy relationship groups shown in Table 4.

Table 4. The complete results for the 3rd - order fuzzy relationship groups.

No group	Fuzzy relation groups	No group	Fuzzy relation groups
1	A3, A3, A3 -> A3, A4	10	A5, A5, A6 -> A6
2	A3, A3, A4 -> A4	11	A6, A6, A6 -> A7
3	A3, A4, A4 -> A5	12	A6, A6, A7 -> A7
4	A4, A4, A5 -> A6	13	A6, A7, A7 -> A5
5	A4, A5, A6 -> A6	14	A7, A7, A5 -> A3
6	A5, A6, A6 -> A5, A6	15	A7, A5, A3 -> A1
7	A6, A6, A5 -> A5	16	A5, A3, A1 -> A2
8	A6, A5, A5 -> A5	17	A3, A1, A2 -> #
9	A5, A5, A5 -> A5, A6		

Step 7. Calculate and defuzzify the forecasted values

In order to calculate the forecast value for all high – order fuzzy relationship groups, we introduce forecasting rule to calculate output value for the trained patterns and use [16] for the untrained patterns in the testing phase.

For the training phase, 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 4, 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 formula (5).

$$\mathrm{Forecasted}_{\mathrm{Value}} = \frac{1}{\mathrm{n}}\sum_{\mathrm{k}=1}^{\mathrm{n}} \frac{(\mathrm{m}_k + \mathrm{submid}_\mathrm{k})}{2} \tag{5}$$

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

$\mathrm{m_k}$ is the midpoint of interval $\mathrm{u_k}$ corresponding to k-th fuzzy set on the right-hand side where the highest level of fuzzy set Ak takes place in these intervals, uk.

$submid_k$ 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 $\mathrm{A_k}$ takes place in this interval.

For the testing phase, we calculate forecasted value for a group which contains the unknown linguistic value according to formula (6), 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 4, it can be shown that group 17 has the fuzzy relationship → #, as it is created by the fuzzy relationship $\mathrm{F(11/21/2018), F(12/06/2018), F(12/21/2018 \rightarrow F(12/22/2018)}$; where, the linguistic value of $\mathrm{F(12/22/2018)}$ is unknown within the historical data, and this unknown next state is denoted by the symbol '#'.

$$\mathrm{Forecasted}_{\mathrm{for\#}} = \frac{(\mathrm{M_{t1}} * \mathrm{w_h}) + \mathrm{M_{t2}} + \cdots + \mathrm{M_{ti}} + \cdots + \mathrm{M_{t\alpha}}}{\mathrm{w_h} + (\alpha - 1)} \tag{6}$$

From formulas (5) and (6) above and based on Tables 4 and 2, we complete forecasted results Gas online price of Viet Nam for all dates between 1/4/2018 and 12/22/2018 based on 3nd-order FTS model with seven intervals are listed in Table 5.

Table 5. The complete forecasted outputs for Gas online price of Viet Nam based on the 3nd- order FTS model.

Date	Gas online price	Fuzzy set	Forecasted value
/4/2018	18240	A3	Not forecasted
1/19/2018	18670	A3	Not forecasted
2/3/2018	18670	A3	Not forecasted
2/21/2018	18340	A3	18714.5
---------	-----	----	-------
11/21/2018	18620	A3	18524
12/06/2018	17180	A1	17190.5
12/21/2018	18141	A2	17952.5
12/22/2018	N/A	N/A	17857.06

The performance of proposed model is assessed with help of the root mean square error (RMSE) to compare the difference between the forecasted values and the actual values. The RMSE is calculated according to formula (7) as follows:

$$\mathrm{R}MSE = \sqrt{\frac{1}{n}\sum_{t=\alpha}^{n}(F_t - R_t)^2} \tag{7}$$

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

3.2. 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 Gas online price 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 i is a vector consisting of n-1 elements x_k where 1≤ k ≤ n-1 and $x_k \leq x_{k+1}$. 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 harmony changes to a new position, the elements of the corresponding new vector need to be sorted to ensure that each element x_k $(1 \leq k \leq n-1)$ 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 [22]. The algorithm generates random solution vectors (harmonies) HMS times and puts them in the HM matrix, specified by formula (2). Improvise a new harmony: Once the harmony memory matrix is initialized, the algorithm starts the first iteration by improvising a new harmony. X = $(x_1, x_2,..., x_{n-1})$ 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)

The forecasting model combines with HS are presented in algorithm 1 as below:

Algorithm 1: The FTS-HS algorithm

Input: *time series data*

Output: *the RMSE, forecasted value and optimum intervals*

1. Initialize all parameters as follows:

HMCR = 0.99, PAR = 0.5, BW = 1.

Lb =x_0= 17000; Ub = x_n = 21000;

maximum number of improvisations is 100;

Number of Harmonys HMS = 50,

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 i, (1≤ i ≤ HMS) do

i. Define linguistic values according to all intervals defined by Harmony i

ii. Fuzzify all historical data by Step 4 in Subsection 3.1

iii. Create all $\alpha -$ order fuzzy relationships by Step 5 in Subsection 3.1

iv. Make all $\alpha -$ order fuzzy relationships groups by Step 6 in Subsection 3.1

v. Calculate forecasting values by Step 7 in Subsection 3.1

vi. Compute the F(X) = RMSE values for Harmony i based on formula (7)

vii. Update HM (2) i according to the RMSE values mentioned above.

viii. Update the RMSE = $F(X)_{HMS\times1}$values

end for

end while

4. Experimental Results

In this study, we apply the proposed model to forecast the Gas online price in Viet Nam with the whole historical data between 1/04/2018 and 12/21/2018 is listed in Table 1 and we also the proposed model to handle other forecasting problems, such as the empirical data for the enrolments of University of Alabama [2] period between 1971 and 1992 are used to perform comparative study in the training phase.

4.1. Experimental Results for Forecasting Gas Online Price in Viet Nam

In this section, we apply the proposed method for forecasting the Gas online price from 1990 to 2010 are listed in Table 1. Our proposed model is executed 50 independent runs for each of 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 1, the number of intervals is kept fix for the proposed model. The forecasted accuracy of the proposed method is estimated using the RMSE (7). The forecasted results of proposed model under number of interval as 14 and various orders are listed in Table 6.

4.2. 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 presented in articles [3, 4, 6, 16, 24] which are examined and compared. The forecasted accuracy of the proposed method is estimated by using the RMSE (7). A comparison of the forecasting accuracy with various orders and different number of intervals among the models [3, 4, 6, 16, 24] and the proposed model are shown in Table 7. Table 7 shows that proposed model has a MSE value is 10.25 which is the lowest among all forecasting models compared.

In addition, the proposed model is also compared with it counterparts [3, 4, 16, 24] based on various high – order fuzzy relationships under seven. The details of comparison are shown in Table 8. The forecasting trend according to order of model is depicted in Figure 1 for clearer illustration.

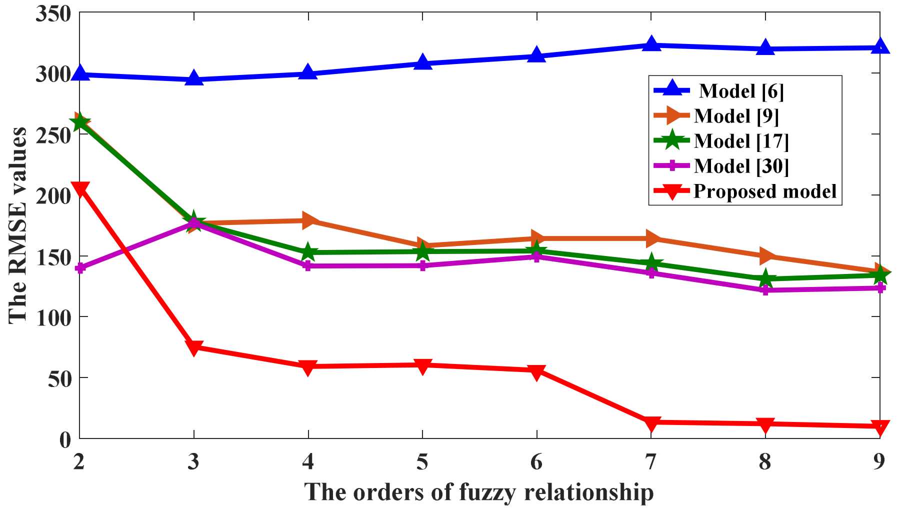

Figure 1. A comparison of the RMSE values under number of intervals of 7 with different high-order fuzzy logical relationships.

Table 6. The completed forecasting results for Gas online price data in Viet Nam under deferent orders of fuzzy relationships and number of interval is 14.

Date	Gas online price	Forecasted value					
		3nd-order	4th-order	5th-order	6th -order	7th-order	8th- order
2/21/2018	18340	18350.25	Not forecasted	Not forecasted	Not forecasted	Not forecasted	Not forecasted
3/8/2018	18340	18350.25	18348.25	Not forecasted	Not forecasted	Not forecasted	Not forecasted
3/23/2018	18340	18350.25	18348.25	18342	Not forecasted	Not forecasted	Not forecasted
4/7/2018	18930	18916.75	18924.75	18911.25	18929	Not forecasted	Not forecasted
4/23/2018	18930	18916.75	18924.75	18911.25	18929	18930	Not forecasted
5/8/2018	19440	19421	19434.75	19449	19438.7	19462.3	19438.5
--------	------	------	-------	------	------	------	-------
11/21/2018	18620	18623.25	18601.75	18615.75	18622.3	18616	18616.7
12/06/2018	17180	17171.25	17192.25	17180.25	17183.5	17181.5	17177.5
12/21/2018	18141	18132	18129.75	18120.25	18140.3	18144.3	18144.7
12/22/2018	N/A	18058.06	18129.94	18304.47	18423.5	18592.9	18661.7
RMSE		61.94	58.7	55.01	46.4	10.94	8.28

Table 7. A comparison of the forecasted results of the proposed model with its counterparts based on high – order FTS under different number of intervals.

Years	Actual data	[6]	[3]	[4]	[16]	[24]	Proposed model
1971	13055	N/A	N/A	N/A	N/A	N/A	N/A
1972	13563	N/A	N/A	N/A	N/A	N/A	N/A
1973	13867	N/A	N/A	N/A	N/A	N/A	N/A
1974	14696	N/A	N/A	N/A	N/A	N/A	N/A
1975	15460	15500	N/A	N/A	N/A	N/A	N/A
1976	15311	15468	15500	N/A	N/A	N/A	N/A
1977	15603	15512	15500	N/A	N/A	N/A	N/A
1978	15861	15582	15500	N/A	N/A	N/A	N/A
1979	16807	16500	16500	16846	N/A	N/A	N/A
1980	16919	16361	16500	16846	16890	16920	16919
1981	16388	16362	16500	16420	16395	16388	16394
-------	------	-------	------	------	-----	------	------
1990	19328	19382	19500	19334	19337	19338	19330
1991	19337	19487	19500	19334	19337	19335	19330
1992	18876	18744	18500	18910	18882	18882	18884
RMSE		365.65	294.44	33.18	15.3	13.15	10.25

Table 8. A comparison of the RMSE value between proposed model its counterparts under different number of orders with the number of interval is 7.

Models	Number of order of forecasting models							
	2	3	4	5	6	7	8	9
[3]	298.48	294.44	298.96	307.47	313.39	322.58	319.65	320.61
[4]	260.45	176.42	178.91	158.06	164.26	164.22	149.62	136.87
[16]	259.08	177.89	152.55	153.41	153.85	143.7	130.79	134.06
[24]	139.98	176.6	141.97	142.71	149.25	135.95	121.56	123.49
Our model	206.52	75.27	59.26	60.43	56.06	13.53	12.28	10.25

From Table 8 and Figure 1, it can be seen that the accuracy of the proposed model is improved significantly. Particularly, the proposed model gets the lower RMSE value than four models presented in articles [3, 4, 16, 24]. 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.

5. Conclusion

In this study, a hybrid forecasting model based on aggregated FTS and Harmony search algorithm for forecasting the gas online price of Viet Nam and Actual enrolments 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 and calculating forecasted value by new defuzzification technique. The combination of optimum techniques always leads to the development of new architecture, which should be more advantageous and expert, providing robust, cost effective, and approximate solution, in comparison to conventional techniques. From the comparison results in Tables 7-8 and Figure 1 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 gas online price 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 an intend of achieving the best possible forecasting performance. That will be the future work of this research.

References

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

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

[3] Chen S M (2002). Forecasting enrolments based on high-order fuzzy time series. Cybernetics and Systems, 33 (1), 1-16.

[4] Chen S-M&Chung N-Y (2006b). Forecasting enrollments of students by using fuzzy time series and genetic algorithms. International Journal of Intelligent Systems, 17, 1-17.

[5] Huarng K (2001). Effective lengths of intervals to improve forecasting in fuzzy time series. Fuzzy Sets and Systems, 123 (3), 387-394.

[6] Singh S R (2007a). A simple method of forecasting based on fuzzy time series. Applied Mathematics and Computation, 186, 330–339.

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

[8] Song Q & Chissom. B S (1993b). Forecasting enrolments with fuzzy time series - Part I. Fuzzy Sets and Systems, 54 (1), 1-9.

[9] Song Q & S. Chissom. B (1994). Forecasting enrolments with fuzzy time series - part II. Fuzzy Sets and Systems, 62, 1-8.

[10] Yu H K (2005). A refined fuzzy time-series model for forecasting. Physical A: Statistical Mechanics and Its Applications, 346 (3-4), 657-681.

[11] Singh S R (2007b). A simple method of forecasting based on fuzzy time series. Applied Mathematics and Computation, 186, 330–339.

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

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

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

[15] al L-Y H e (2010). Temperature prediction and TAIFEX forecasting based on fuzzy relationships and MTPSO techniques. Expert Syst Appl, 37, 2756–2770.

[16] I. H. Kuo e a (2009). An improved method for forecasting enrolments based on fuzzy time series and particle swarm optimization. Expert Systems with Applications, 36, 6108–6117.

[17] Kuo. I-H, Horng. S-J, Che. Y-H, R-S R, Kao. T-W, R.-J. C. J-L L, et al. (2010). Forecasting TAIFEX based on fuzzy time series and particle swarm optimization. Expert Systems with Applications, 2 (37), 1494–1502.

[18] Pritpal. S&Bhogeswar B (2013). High-order fuzzy-neuro expert system for time series forecast-ing. Knowl-Based Syst, 46, 12-21.

[19] L. Al - Mataneh e a (2014). Development of temperature based weather forecasting model using Neural network and fuzzy logic. International Journal of Multimedia and Ubiquitous Engineering, 9 (12), 343-366.

[20] S. Askari&Montazerin N (2015). A high-order multi-variable fuzzy time series forecasting algorithm based on fuzzy clustering. Expert Systems with Applications, 42, 2121-2135.

[21] B. Sun, H. Guo, H. R. Karimi&Y. Ge S X (2015). Prediction of stock index futures prices based on fuzzy sets and multivariate fuzzy time series. Neuro computing, 151, 1528-1536.

[22] Geem Z W, Kim J H&Loganathan G V (2001). A new heuristic optimization algorithm: Harmony search. Simulation, 76, 60–66.

[23] Fesanghary M, Mahdavi M, Minary M&Alizadeh (2008). Y. Hybridizing harmony search algorithm with sequential quadratic programming for engineering optimization problems. Comput Methods Appl Mech Energy, 197, 3080–3091.

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