IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 19, Issue 6, Ver. III (Nov.- Dec. 2017), PP 26-35
www.iosrjournals.org

Handing Forecasting Problems Based on Combining Fuzzy Logical Relationships and Particle Swam Optimization Algorithm

Nghiem Van Tinh
Thai Nguyen University of Technology, Thai Nguyen University, ThaiNguyen, VietNam
nghiemvantinh@tnut.edu.vn

Abstract :*In recently year, most of the forecasting models are based on concept of fuzzy time series according to the historical data for forecasting the future. Generally, a fuzzy time series forecasting model include three main stages such a fuzzification of crisp time series observations, the identification fuzzy logical relationships and the defuzzification. In these stages, many researchers have used different algorithms to build models with aim to improve forecasting accuracy. In this paper, a hybrid model for forecasting of rice production based on fuzzy time series and Particle Swarm Optimization (PSO) algorithm is presented. Firstly, the historical data needs to be expressed by linguistic values in the fuzzified stage. Second, determining fuzzy logical relationships and fuzzy logical relationship groups is obtained from fuzzified historical data. Third, calculate the forecasting value for the fuzzy relation groups by the forecasted defuzzification rules. Finally, improving the forecasting performance of the model by using the optimal PSO algorithm. To illustrate the forecasting process and the effectiveness of the proposed model, the historical data of average rice production of Viet Nam is examined. The simulation result shows that the proposed model also gets a higher average forecasting accuracy than the existing methods when handle forecasting the enrollments of the University of Alabama based on the first -order and high -order FTS.*

Keywords -*Enrollments, rice production, forecasting, fuzzy time series, fuzzy relationship groups, particle swarm optimization.*

Date of Submission: 25-10-2017 Date of acceptance: 07-12-2017

I. Introduction

For planning the future forecasting plays an important role. During the last decade, various approaches have been developed for forecasting data of dynamic and non-linear in nature. Fuzzy theory [1] has been successfully employed to forecast. The fuzzy time series forecasting models based on fuzzy set theory have been widely applied to diverse fields such as enrolments forecasting[2] - [9] , crop productions prediction [10], stock markets [11] and temperature prediction [12] . The fuzzy time series and the corresponding forecast model was introduced by Song and Chissom in 1993. They introduced both the time-invariant fuzzy time series [2] and the time-variant time series [3] model which use the max–min operations to forecast the enrolments of the University of Alabama. Unfortunately, their method has 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, Ref. [5] proposed the first-order fuzzy time series model by using simple arithmetic calculations instead of max-min composition operations [3] for better forecasting accuracy. Thereafter, the fuzzy time series methods received increasing attention in many forecasting applications. To achieve better forecasting accuracy, Ref. [6] presented an effective approach which can properly adjust the lengths of intervals. Chen in [7] presented a new forecasting model based on the high-order fuzzy logical relationship groups to forecast the enrolments of the University of Alabama. Singh [10] developed a simplified and robust computational method for the forecasting rules based on one and various parameters as fuzzy logical relationships. Lee et al. in [12] 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 [5] , Kuo et al. in [13] introduced a new hybrid forecasting model which combined fuzzy time series with PSO algorithm to find the proper length of each interval. Then, to improve previous model in [13] . They continued to present a new forecast method to solve the TAIFEX forecasting problem based on fuzzy time series and PSO [14] . Huang et al. in [15] proposed a new hybrid forecasting model based on two computational methods, fuzzy time series and PSO for forecasting enrolments by considering more local information of latest fuzzy logical relationship in current state of fuzzy logical relationship group to find the forecasting value in FTS. Some other authors, proposed some methods for the temperature prediction and the TAIFEX forecasting, based on two-factor fuzzy logical relationships and PSO as shown in [16] and [17] . In Addition, other hybrid techniques such as: Chen and Kao [18] proposed a new method for forecasting the TAIEX, based on fuzzy time series, particle swarm optimization techniques and support vector machines. Pritpal and Bhogeswar[19] presented a new model based on hybridization of fuzzy time series theory with artificial neural network (ANN). Cheng and Li [20] proposed an enhanced HMM-based forecasting model by developing a novel fuzzy smoothing method to overcome the problem of rule redundancy and achieve better results.

The above-mentioned researches showed that the lengths of intervals and fuzzy logical relationship 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 PSO is presented to forecast rice production for each year on basis of historical time series of rice data in Viet Nam. Firstly, the proposed method fuzzifies the historical data into fuzzy sets to create the first-order and the high – order fuzzy logical relationship groups. Secondly, the PSO algorithm for the optimal lengths of intervals is developed by searching the space of the universe of discourse. The case study with the data of rice production of Viet Nam shows that the performance of proposed model is better than those of any existing models.

The rest of this paper is organized as follows. In Section 2, a brief review of the basic concepts of FTS and algorithms are introduced. In Section 3, we designed model to forecast rice production based on the FTS and PSO algorithm. Section 4 evaluates the forecasting performance of the proposed method with the existing methods on the enrolment data of the University of Alabama. Finally, Section 5 provides some conclusions.

II. Basic Concepts of Fuzzy Time Series and Algorithms

2.1. Basic concepts of fuzzy time series

Conventional time series refer to real values, but fuzzy time series are structured by fuzzy sets [1] . 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 [2] - [4]

Let $Y(t)(t = .., 0, 1, 2 ..)$, a subset of R, be the universe of discourse on which fuzzy sets $f_i(t)$ $(i = 1,2 \ldots)$ 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**)** [3]

If F(t) is caused by F(t-1) only, the relationship between F(t) and F(t-1) can be expressed by $F(t-1) \rightarrow F(t)$. Chen [3] suggested that when the maximum degree of membership of F(t) belongs toA_i, F(t) is considered A_j. Hence, the relationship between F(t) and F(t -1) is denoted by fuzzy logical relationship $A_i \rightarrow A_j$ 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*: **m - order fuzzy logical relationship [7]

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

Definition 4: Fuzzy Logical Relationship Group (FLRG)[5]

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

$A_i \rightarrow A_{k1}$, $A_i \rightarrow A_{k2}$, …….

So, based on [5] , these fuzzy logical relationship can be grouped into the same FRG as : $A_i \rightarrow A_{k1}, A_{k2} \ldots$

2.2. Particle swarm optimization algorithm (PSO)

PSO is a population-based search algorithm and is initialized with a population of random solutions, called particles. Each particle in PSO is also associated with a velocity. Particles fly through the search space with velocities which are dynamically adjusted according to their historical behaviours. Therefore, the particles have the tendency to fly towards the better and better search area over the course of search process. PSO was first introduced by Eberhart and Kannedy in 1995 to simulate birds searching food in an area. The development of PSO algorithm [13] ,was also inspired by the social behaviour of animals, such as fish schooling, birds flocking and the swarm theory. The PSO algorithm applies a cooperative particle swarm to find the best solution from all feasible solutions. Each particle is randomly initialized and then allowed to move in the virtual searching space. At each step of optimization, each particle evaluates its own fitness and the fitness of its neighbouring particles. The particles change its state according to the three principles: weight inertia i.e. ω , it's most optimist position i.e. P_{best_id} , swarm's most optimist position i.e. G_{best} ; and converges to the most optimal position in the entire solution space by continuous change in the personal best and global best position. A moving particle, indexed by *id*, adjusts its candidate solution according to the following formulas:

$$V_{id}^{k+1} = \omega^k * V_{id}^k + C_1 * \mathrm{Rand}(\,) * (P_{best\_id} - x_{id}^k) + C_2 * \mathrm{Rand}(\,) * (G_{best} - x_{id}^k) \quad (1)$$

$$x_{id}^{k+1} = x_{id}^k + x_{id}^{k+1} \quad (2)$$

$$\omega^k = \omega_{max} - \frac{k*(\omega_{max} - \omega_{min})}{iter\_max} \quad (3)$$

The P_{best_id} for i^{th} particle is presented as $P_i = [p_i^1, p_i^2, \ldots, p_i^d]$ and calculated as;

$$P_{best\_id}^{k+1} = f(x) = \begin{cases} P_{best_{id}}^{k+1}, & if\ fitness(x_{id}^{k+1}) > P_{best_{id}}^k \\ fitness(x_{id}^{k+1}), & if\ fitness(x_{id}^{k+1}) \le P_{best_{id}}^k \end{cases} \quad (4)$$

The G_{best} at k^{th}iteration is computed as: $G_{best} = \min(P_{best_{id}}^k)$ (5)

During iterations, movement of the particles is influenced by its local neighborhood or the global neighborhood. That is, the portion of the adjustment to the velocity is influenced by the individual's previous best position is considered as the cognition component, whereas the portion influenced by the best in the neighborhood is termed as the social component. With the addition of the inertia factor ω, brought in for balancing the global and the local search ability of the particles in PSO.

where,

- ✓ X_{id}^k is the current position of a particle id in k^{th} iteration;
- ✓ V_{id}^k is the velocity of the particle id in k^{th}iteration, and is limited to $[-V_{max}, V_{max}]$ whereV_{max} is a constant pre-defined by user.
- ✓ P_{best_id} is the position of the particle id that experiences the best fitness value.
- ✓ G_{best} is the best one of all personal best positions of all particles within the swarm.
- ✓ Rand() is the function can generate a random real number between 0 and 1 under normal distribution.
- ✓ C_1 and C_2 are acceleration values which represent the selfcondence coefficient and the social coefficient, respectively.
- ✓ ω is the inertia weight factor accoding to Eq. (3).

A briefly description of the standard PSO is summarized in the following algorithm 1.

Algorithm 1: Standard PSO algorithm

1. Initialize the particles' positions X_{id} and velocities V_{id}

2. **While** the stop condition (the optimal solution is found or the maximum moving step is reached) is not satisfied **do**

   2.1**. For** particle i, (1≤ i ≤ NumberOfParticles) **do**
   - ✓ Calculate the fitness value of particle i

     if fitness better than previous P_{best_id} then
   - ✓ Set fitness value is new P_{best_id} according to (4)

     end if

   end **for**

   2.2. Update the global best position of all particles G_{best} according to (5).

   2.3. F**or** particle i, (1≤ i ≤ NumberOfParticles) **do**
   - ✓ Move particle i to another position according to (1) and (2)

   end **fo**r end **while**

III. A Forecasting Model Based on the Fuzzy Time Series and PSO Algorithm

Based on Kuo et al. in [13] , a forecasting model which combined the fuzzy logical relationship group and PSO algorithm is introduced. First, original historical data are used instead of the variations of historical data in our forecasting model. Second, the FLRGs are derived from the fuzzified historical data and calculate the forecasting output based on the fuzzy sets on the right-hand side of the FLRGs. Third, the PSO algorithm is applied to adjust the interval lengths to increase forecasting accuracy. A detailed explanation of the proposed model is expressed as follows.

1.1. Forcasting model based on the first – order FLRGs

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 between 1990-2010 is listed in Table 1 in which it taken from the site www.gso.gov.vn, more precisely from https://www.gso.gov.vn/default.aspx?tabid=717 are used to illustrate the first - order fuzzy time series forecasting process. The step-wise procedure of the proposed model is detailed as follows:

Table 1: The annual data of the average rice production (thousand ton/ year) of Viet Nam

Year	Actualrice data	Year	Actualrice data	Year	Actualrice 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 $D_{min} = I_{min} - N_1$ 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 historical rice data are shown in Table 1, 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 [5] , [13] , we cut U into seven intervals, $u_1, u_2, \ldots, u_7$, respectively. The length of each interval is $L = \frac{D_{max} - D_{min}}{7} = \frac{40000 - 19000}{7} = 3000$. Thus, the seven intervals are defined as follows:

u_i= (D_{min} + (i-1)*L, D_{min} + i *L], 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 in (6) and (7) as follows.

$$A_i = \sum_{j=1}^{7} \frac{a_{ij}}{u_j} \quad (6)$$

$$a_{ij} = \begin{matrix} 1 & j == i \\ 0.5 & \text{if } j == i-1 \text{ or } j == i+1 \\ 0 & \text{otherwise} \end{matrix} \quad (7)$$

where $a_{ij} \in [0,1]$, $1 \leq i \leq 7, 1 \leq j \leq 7$ and $u_j$is the j-th interval of U. The value of a_{ij}indicates the grade of membership of u_j in the fuzzy set A_i

Step 4: Fuzzy all historical data of rice production

In order to fuzzify all historical data, it's necessary to assign a corresponding linguistic value to each interval first. 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 (6) and (7), 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 2.

Table 2. The results of fuzzification for rice production data.

Year	**Actual data**	**Fuzzy sets**	**Membership grades**
1990	19225.1	A1	A1= $1/u_1 + 0.5/u_2 + 0/u_3$
1991	19621.9	A1	A1= $1/u_1 + 0.5/u_2 + 0/u_3$
1992	21590.4	A1	A1= $1/u_1 + 0.5/u_2 + 0/u_3$
1993	22836.5	A2	A2= $0.5/u_1 + 1/u_2 + 0.5/u_3$
1994	23528.2	A2	A2= $0.5/u_1 + 1/u_2 + 0.5/u_3$
1995	24963.7	A2	A2= $0.5/u_1 + 1/u_2 + 0.5/u_3$
1996	26192	A3	A3= $0/u_1 + 0.5/u_2 + 1/u_3 + 0.5/u_4$
1997	27288.7	A3	A3= $0/u_1 + 0.5/u_2 + 1/u_3 + 0.5/u_4$

1998	28919.3	A4	$A4 = 0/u_1 + 0/u_2 + 0.5/u_3 + 1/u_4 + 0.5/u_5$
1999	31393.8	A5	$A5 = 0/u_1 + 0/u_2 + 0/u_3 + 0.5/u_4 + 1/u_5 + 0.5/u_6$
2000	32529.5	A5	$A5 = 0/u_1 + 0/u_2 + 0/u_3 + 0.5/u_4 + 1/u_5 + 0.5/u_6$
2001	32108.4	A5	$A5 = 0/u_1 + 0/u_2 + 0/u_3 + 0.5/u_4 + 1/u_5 + 0.5/u_6$
2002	34447.2	A6	$A6 = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5 + 1/u_6 + 0.5/u_7$
2003	34568.8	A6	$A6 = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5 + 1/u_6 + 0.5/u_7$
2004	36148.9	A6	$A6 = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5 + 1/u_6 + 0.5/u_7$
2005	35832.9	A6	$A6 = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5 + 1/u_6 + 0.5/u_7$
2006	35849.5	A6	$A6 = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5 + 1/u_6 + 0.5/u_7$
2007	35942.7	A6	$A6 = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5 + 1/u_6 + 0.5/u_7$
2008	38729.8	A7	$A7 = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0.5/u_6 + 1/u_7$
2009	38950.2	A7	$A7 = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0.5/u_6 + 1/u_7$
2010	39988.9	A7	$A7 = 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0.5/u_6 + 1/u_7$

Step 5.Create all m – order fuzzy relationships.

Based on Definition 3. To establish a m-order fuzzy relationship, we should find out any relationship which has the $\mathrm{F}(\mathrm{t}-m), \mathrm{F}(\mathrm{t}-m+1), \ldots, \mathrm{F}(\mathrm{t}-1) \rightarrow F(t)$, where $\mathrm{F}(\mathrm{t}-m), \mathrm{F}(\mathrm{t}-m+1), \ldots, \mathrm{F}(\mathrm{t}-1)$ and $F(t)$are called the current state and the next state, respectively. Then a m - order fuzzy relationship is got by replacing the corresponding linguistic values. The same linguistic values (fuzzy set) cannot appear more than once on the right hand side. For example, supposed $m = 1$, a fuzzy relationship $\mathrm{A}_1 \rightarrow \mathrm{A}_1$is got as $\mathrm{F}(1990) \rightarrow \mathrm{F}(1991)$. So, from Table 2. we get fist - order fuzzy logical relationships are shown in Table 3.

Table 3. The first - order fuzzy logical relationships

No	**Fuzzy relation**	**F(t)**	**No**	**Fuzzy relation**	**F(t)**
1	A1 -> A1	$\mathrm{F}(1990) \rightarrow \mathrm{F}(1991)$	7	A4 -> A5	$\mathrm{F}(1998) \rightarrow \mathrm{F}(1999)$
2	A1 -> A2	$\mathrm{F}(1992) \rightarrow \mathrm{F}(1993)$	8	A5 -> A5	$\mathrm{F}(1999) \rightarrow \mathrm{F}(2000)$
3	A2 -> A2	$\mathrm{F}(1993) \rightarrow \mathrm{F}(1994)$	9	A5 -> A6	$\mathrm{F}(2001) \rightarrow \mathrm{F}(2002)$
4	A2 -> A3	$\mathrm{F}(1995) \rightarrow \mathrm{F}(1996)$	10	A6 -> A6	$\mathrm{F}(2002) \rightarrow \mathrm{F}(2003)$
5	A3 -> A3	$\mathrm{F}(1996) \rightarrow \mathrm{F}(1997)$	11	A6 -> A7	$\mathrm{F}(2007) \rightarrow \mathrm{F}(2008)$
6	A3 -> A4	$\mathrm{F}(1997) \rightarrow \mathrm{F}(1998)$	12	A7 -> A7	$\mathrm{F}(2008) \rightarrow \mathrm{F}(2009)$

Step 6: Establish all m – order fuzzy logical relationshipsgroups

Based on [5] 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. Thus, from Table 3 and based on Definition 4, we can obtain seven fuzzy logical relationship groups shown in Table 4.

Table 4. The first - order fuzzy logical relationship groups

No	**FLRGs**	**No**	**FLRGs**	**No**	**FLRGs**
1	A1 → A1, A2	4	A4 → A5	6	A6 → A6, A7
2	A2 → A2, A3	5	A5 → A5, A6	7	A7 → A7
3	A3 → A3, A4				

Step 7.Calculate and defuzzify the forecasted outputs

First, we calculate the predicted value for each group of fuzzy relation by the proposed rules [5] , then based on the value of each of these groups to defuzzify the forecasted output for year i. For each FLRG in the training phase and the testing phase, we use the principles to compute it as follows.

Rule 1: If the fuzzified rice production of year i is A_i, and there is only one fuzzy logical relationship in the fuzzy logical relationship groups, which is shown as $A_i \to A_k$, the forecasted output of year i+1 is m_k ; where m_k is the midpoint of interval u_k.
Rule 2: If the fuzzified rice production of year i is A_i, and there are more one fuzzy logical relationship in the fuzzy logical relationship groups, which are shown as $A_i \to A_{k1}$, $A_i \to A_{k2}$, …,$A_i \to A_{kp}$; the forecasted output of year i+1 is $\frac{m_{k1}+m_{k2}+\cdots+m_{kp}}{p}$.
Where, $m_{k1}, m_{k2}, \dots, m_{kp}$ is the midpoint if interval $u_1, u_{2,}, \dots, u_p$, respectively .
Rule 3: If the fuzzified rice production of year i is A_i, and there is only one fuzzy logical relationship in which the right - hand side of fuzzy logical relationship group is empty, which is shown as $A_i \to \neq$; the forecasted output of year i+1 is m_i .
From above rules and based on Table 4, we obtain the forecasted value for the first – order fuzzy logical relationship groups are shown in Table 5.

Table 5. The complete forecasted values for all fuzzy logical relationship groups in Table 4.

No	FLRGs	Value	No	FLRGs	Value	No	FLRGs	Value
1	A1 → A1, A2	22000	4	A4 → A4, A5	32500	6	A6 → A6, A7	37000
2	A2 → A2, A3	25000	5	A5 → A5, A6	34000	7	A7 → A7	38500
3	A3 → A3, A4	28000						

Following, based on Table 5 and the data of Table 1, we complete forecasted results for rice production the period from 1990 to 2010 based on the first - order fuzzy time series with seven intervals are listed in Table 6.

Table 6. The complete forecasted outputs for rice production of Viet Nam based on the first– order FTS model

Year	Actual rice data	Fuzzy set	Forecasted rice value
1990	19225.1	A1	Not forecasted
1991	19621.9	A1	22000
1992	21590.4	A1	22000
1993	22836.5	A2	22000
1994	23528.2	A2	25000
1995	24963.7	A2	25000
1996	26192	A3	25000
1997	27288.7	A3	28000
1998	28919.3	A4	28000
1999	31393.8	A5	32500
2000	32529.5	A5	34000
2001	32108.4	A5	34000
2002	34447.2	A6	34000
2003	34568.8	A6	37000
2004	36148.9	A6	37000
2005	35832.9	A6	37000
2006	35849.5	A6	37000
2007	35942.7	A6	37000
2008	38729.8	A7	37000
2009	38950.2	A7	38500
2010	40005.6	A7	38500

The forecasting performance 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 mean squared error (MSE), mean error (ME) according to (8) and (9) as follows:

$$MSE = \frac{1}{n}\sum_{i=m}^{n}(F_i - R_i)^2 \quad (8)$$

$$ME = \frac{1}{n}\sum_{i=m}^{n} |F_i - R_i| \quad (9)$$

Where, $R_i$denotes actual value at year i, F_i is forecasted value at year i, n is number of the forecasted data, m is order of the fuzzy logical relationships

1.2. Forecasting model based on aggregated the FTS and PSO algorithm

To improve forecasted accuracy of the proposed, the effective lengths of intervals and fuzzy logical relationship groups which are two main issues presented in this paper. A novel method for forecasting rice

production is developed to adjust the length each of intervals in the universe of discourse without increasing the number of intervals by minimizing the MSE value (8).
In our model, each particle 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 p_0 and p_n, respectively. Each particle is a vector consisting of n-1 elements pi where 1≤i≤ n-1 and $p_i \leq p_{i+1}$. Based on these n-1 elements, define the n intervals as $u_1 = [p_0, p_1]$, $u_2 = [p_1, p_2]$,,..., $u_i = [p_{i-1}, p_i]$,… and $u_n = [p_{n-1}, p_n]$, 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 pi ($1 \leq i \leq n-1$) arranges in an ascending order. The complete steps of the proposed method are presented in Algorithm 3.

Algorithm 3: The FTS-PSO algorithm

1. Initialize all particles' positions X_{id} , velocities V_{id} and parameters of the proposed method. These parameters are:
   - ✓ Number of particles is **50**
   - ✓ Maximum number of iterations is **150**
   - ✓ The value of inertial weigh ω be linearly decreased from $\omega_{max} = \mathbf{0.9}$ to$\omega_{min} = \mathbf{0.4}$
   - ✓ The coefficient C1 equal C2 as **2**
   - ✓ The position of particle i be limited by: $x_{min} + Rand(\ ) * (x_{max} - x_{min})$; where x_{min} and x_{max} are lower and upper bounds of universal set, respectively.
   - ✓ The velocity of particle i be exceeded by $v_{min} + Rand(\ ) * (v_{max} - v_{min})$
2. **While** the stop condition (maximum iterations or minimum MSE criteria) is not satisfied **do**

   2.1. For particle i, (1≤ i ≤ NumberOfParticles) **do**
   - ✓ Define linguistic values according to all
   - ✓ intervals defined by the current position of particle i
   - ✓ Fuzzify all historical data by Step 4 in Subsection 3.1
   - ✓ Create all $m -$ order fuzzy relationships by Step 5 in Subsection 3.1
   - ✓ Make all $m -$ order fuzzy relationship groups by Step 6 in Subsection 3.1
   - ✓ Calculate forecasting values by Step 7 in Subsection 3.1
   - ✓ Compute the MSE values for particle i based on Eq. (8)
   - ✓ Update the personal best position of particle i according to the MSE values mentioned above.

   end **for**

   2.2. Update the global best position of all particles according to the MSE values mentioned above.
3. **For** particle i, (1≤ i ≤ NumberOfParticles) **do**
   - ✓ move particle i to another position according to (1) and (2)

end for
   - ✓ update ω according to Eq. (3)

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 1 and we also the proposed model to handle other forecasting problems, such as the empirical data for the enrolments of University of Alabama [5] from 1971 to 1992 are used to perform comparative study in the training phase.

4.1 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 1. Our proposed model is executed 15 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 3, the number of intervals is kept fix for the proposed model. The forecasted accuracy of the proposed method is estimated using the ME value (9). The forecasted results of proposed model under number of interval is 14 and various orders are listed in Table 7.

Table 7: The completed forecasting results for rice production data of Viet Nam

Year	**Actual rice data**	**Forecasted rice value**				
		1st – order	2nd – order	3rd – order	4th – order	5th – order
1990	19225.1	-	-	-	-	-
1991	19621.9	20720.2	-	-	-	-
1992	21590.4	20720.2	21576.7	-	-	-
1993	22836.5	22934.9	23177.1	23193.1	-	-

1994	23528.2	22934.9	23177.1	23193.1	23574.3	-
1995	24963.7	25006.1	24937.4	24970.6	24988.5	24966
1996	26192	26266.2	26285.4	26163.4	26204.1	26133.4
1997	27288.7	27271.6	27195.1	27325.4	27209.5	27298.1
1998	28919.3	29126.8	29028.8	29166.4	29126.8	28876.8
1999	31393.8	31636.8	31224.2	31078	31106	31370.1
2000	32529.5	33395.4	32397.6	32449.5	32384.8	32343.9
2001	32108.4	31636.8	32397.6	32449.5	32384.8	32343.9
2002	34447.2	33395.4	34429.5	34396	34390.5	34493.6
2003	34568.8	35618.3	34429.5	34396	34390.5	34493.6
2004	36148.9	35618.3	36412.4	36135	35968	35950.4
2005	35832.9	35676.9	35828	35927.6	35968	35950.4
2006	35849.5	35809	35828	35927.6	35968	35950.4
2007	35942.7	35809	36412.4	35927.6	35968	35950.4
2008	38729.8	39140.3	38478.2	38852.2	38897.5	38846
2009	38950.2	39140.3	38478.2	38852.2	38897.5	38846
2010	39988.9	39140.3	39993.8	39958.4	39822.2	39992.3
2011	N/A	39140.3	39899	39828.2	39505.3	39446.2
ME		**449.66**	**168.73**	**134.7**	**127.1**	**82.9**

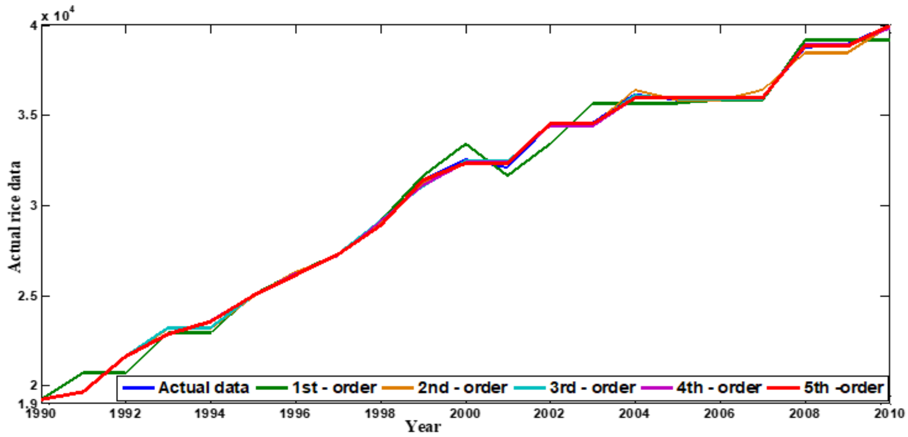

Fig. 1. The forecasting trend of proposed model under different number of orders in the same intervals

The forecasting results of proposed model based on the different orders are also depicted in Fig 1. From Fig. 1 shown that the performance of the proposed model is improving a lot with increasing number of orders in the same number of interval. All of these conclusion have also been shown in Table 7 with the ME criteria (9).

4.2 Experimental results for forecasting enrollments

In order to verify the forecasting effectiveness of the proposed model under different number of intervals and first - order FLRGs , five FTS models in SC93b [3] , C96 [5] , H01b [6] CC06b [9] and HPSO [13] , are examined and compared. The forecasted accuracy of the proposed method is estimated using the MSE function (8). From the parameters are expressed in Algorithm 3. Our proposed model is executed 10 runs, and the best result of runs is taken to be the final result. A comparison of the forecasting accuracy of all models mentioned above and the proposed model, are listed in Table 8. where all models use the first-order FTS to forecast enrolments of university of Alabama in the training phase.

Table 8. A comparison of the forecasted results of the proposed method with the existing models based on the first – order fuzzy time series under number of interval is 14.

Year	**Actual data**	**SC93b**	**C96**	**H01b**	**CC06b**	**HPSO**	**Our model**
1971	13055	-	-	-	-	-	
1972	13563	14000	14000	14000	13714	13555	13715.6
1973	13867	14000	14000	14000	13714	13994	13715.6
1974	14696	14000	14000	14000	14880	14711	14768.4
1975	15460	15500	15500	15500	15467	15344	15330.4
1976	15311	16000	16000	15500	15172	15411	15437.1

DOI: 10.9790/0661-1906032635

1977	15603	16000	16000	16000	15467	15411	15437.1
1978	15861	16000	16000	16000	15861	15411	15437.1
1979	16807	16000	16000	16000	16831	16816	16806.4
1980	16919	16813	16833	17500	17106	17140	16918.1
1981	16388	16813	16833	16000	16380	16464	16416.8
1982	15433	16789	16833	16000	15464	15505	15502.8
1983	15497	16000	16000	16000	15172	15411	15437.1
1984	15145	16000	16000	15500	15172	15411	15437.1
1985	15163	16000	16000	16000	15467	15344	15330.4
1986	15984	16000	16000	16000	15467	16018	16040
1987	16859	16000	16000	16000	16831	16816	16806.4
1988	18150	16813	16833	17500	18055	18060	18148.8
1989	18970	19000	19000	19000	18998	19014	18943
1990	19328	19000	19000	19000	19300	19340	19304.9
1991	19337	19000	19000	19500	19149	19340	19304.9
1992	18876		19000	19000	19149	19014	18943
MSE		**423027**	**407507**	**226611**	**35324**	**22965**	**20318.3**

From Table 8, it can be seen that our model has the smallest MSE value of **20318.3** among all the forecasting model compared. This proves that the proposed model is more efficient than those of existing models based on first – order FTS.In addition, we also perform 10 more runs to be compared with various high-order models under seven intervals such as C02 model [7] , CC06b model [9] and HPSO model [13] . The detail of comparison is shown in Table 9. The forecasting trend is also depicted in Fig. 2 for clearer illustration.

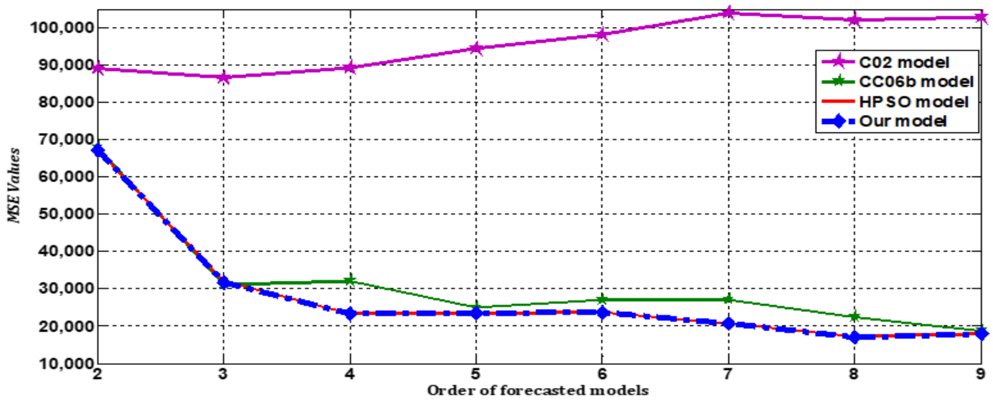

Fig. 2. A comparison of the MSE values for 7 intervals with different high-order FLRs

Table 9. A comparison of the MSE 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	7	8	9	Average
C02	89093	86694	89376	94539	98215	104056	102179	102789	95868
CC06b	67834	31123	32009	24984	26980	26969	22387	18734	31373
HPSO	67123	31644	23271	23534	23671	20651	17106	17971	28121
Our model	67104.9	31641	23270.8	23533.8	23662	20645	**17090.6**	17962	**28113.8**

In Table 9, it can be seen that the accuracy of the proposed model is improved significantly. Particularly, our model gets the lower MSE values than two models presented in C02 [5] and CC06b [9] . For the HPSO model [13] , the forecasting results of proposed model is similar to the HPSO model and obtains the average MSE value of ***28113.8***. 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 in the same number of interval.

V. Conclusion

In this paper, we have presented a hybrid forecasting model to forecast the rice production of Viet Nam and Actual enrollments of the University of Alabama are used to compare with those of existing models in the training phase. The main contributions of this paper are illustrated in the following. First, fuzzy logical relationship groups are established from the historical data of rice production after it is fuzzified. Second, the

PSO algorithm for the optimized lengths of intervals is developed to adjust the interval lengths by searching the space of the universe. Third, based on the performance comparison in Tables 8, 9 and Fig. 2, the author shows the proposed model outperforms previous forecasting models for the training phase with various orders.

Although this study shows the superior forecasting capability compared with the existing forecasting models, but the proposed model is only tested by two problems: enrolments data and rice production dataset. To continue considering the effectiveness of the forecasting model, there are two suggestions for future research: First, we 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, we 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 research.

References

[1] L. A. Zadeh, 1965, Fuzzy sets, Information and Control, vol.8, no.3, pp.338-353.
[2] Q. Song and B. S. Chissom, Fuzzy time series and its models, Fuzzy Sets and Systems, vol.54, no.3,(1993a) 269-277.
[3] Q. Song and B. S. Chissom, Forecasting enrolments with fuzzy time series - Part I, Fuzzy Sets and Systems, vol.54, no.1, (1993b) 1-9.
[4] Q. Song and B. S. Chissom, Forecasting enrolments with fuzzy time series - part II, Fuzzy Sets and System, vol. 62, (1994) 1-8.
[5] S.M. Chen, "Forecasting Enrolments based on Fuzzy Time Series," Fuzzy set and systems, vol. 81, (1996) 311-319.
[6] K. Huarng, Effective lengths of intervals to improve forecasting in fuzzy time series, Fuzzy Sets and Systems, vol.123, no.3, (2001b) 387-394.
[7] S. M. Chen, Forecasting enrolments based on high-order fuzzy time series, Cybernetics and Systems, vol.33, no.1, (2002) 1-16.
[8] H. K. Yu, A refined fuzzy time-series model for forecasting, Physical A: Statistical Mechanics and Its Applications, vol.346, no.3-4, (2005) 657-681.
[9] Chen, S.-M., & Chung, N.-Y. Forecasting enrolments of students by using fuzzy time series and genetic algorithms. International Journal of Intelligent Systems, 17, (2006b) 1–17.
[10] Singh, S. R. A simple method of forecasting based on fuzzy time series. Applied Mathematics and Computation, 186, (2007a) 330–339.
[11] Chen, T.-L., Cheng, C.-H., & Teoh, H.-J., High-order fuzzy time-series based on multi-period adaptation model for forecasting stock markets. Physical A: Statistical Mechanics and its Applications, 387, (2008) 876–888.
[12] 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, (2008b) 328–336.
[13] 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 , (2009) 6108–6117.
[14] I-H. Kuo, S.-J. Horng, Y.-H. Chen, R.-S. Run, T.-W. Kao, R.-J. C., J.-L. Lai, T.-L. Lin, Forecasting TAIFEX based on fuzzy time series and particle swarm optimization, Expert Systems with Applications 2(37), (2010) 1494–1502.
[15] Y.-L. Huang, S.-J. Horng, M. He, P. Fan, T.-W. Kao, M. K. Khan, J.-L. Lai, I-H. Kuo, A hybrid forecasting model for enrolments based on aggregated fuzzy time series and particle swarm optimization, Expert Systems with Applications 7(38), (2011) 8014–8023.
[16] Lee, L. W., Wang, L. H., Chen, S. M., &Leu, Y. H. Handling forecasting problems based on two-factors high-order fuzzy time series. IEEE Transactions on Fuzzy Systems, 14, (2006) 468–477.
[17] Ling-Yuan Hsu et al. Temperature prediction and TAIFEX forecasting based on fuzzy relationships and MTPSO techniques, Expert Syst. Appl.37, (2010) 2756–2770.
[18] S. Pritpal, B. Bhogeswar, High-order fuzzy-neuro expert system for time series forecasting, Knowl.-Based Syst. 46 (2013) 12–21.
[19] Y.C. Cheng, S.T. Li, Fuzzy time series forecasting with a probabilistic smoothing hidden Markov model, IEEE Trans. Fuzzy Syst. 20 (2) (2012) 291–304.
[20] L.Y. Wei, C.H. Cheng, H.H. Wu, A hybrid ANFIS based on n-period moving average model to forecast TAIEX stock, Appl. Soft Comput. 19 (2014) 86–92.

Nghiem Van Tinh "Handing Forecasting Problems Based on Combining Fuzzy Logical Relationships and Particle Swam Optimization Algorithm." IOSR Journal of Computer Engineering (IOSR-JCE) , vol. 19, no. 6, 2017, pp. 26-35.