Forecasting for Coronavirus Disease Spread in Vietnam Using Fuzzy Time Series Model and Particle Swarm Optimization

Nghiem Van Tinh
Thai Nguyen University of Technology, Thai Nguyen University
Thai Nguyen, Vietnam

***Abstract*—The current event in world is virus COVID-19; the spread of this virus can put all countries in situation of incapacity of how manage and face. In order to help people to make decisions in dealing this epidemic matter. This research suggests a forecasting model based on fuzzy time series(FTS) and Particle swarm optimization (PSO) for predicting the number of confirmed cases of COVID-19 in Vietnam. In the proposed model, the historical time series data set of COVID-19 is fuzzified based on fuzzy time series theory. Each fuzzified time series values are then used to create the FLRs. Based on these FLRs, a new defuzzification technique is introduced to calculate the output value of forecasting model. Then, the PSO algorithm is used for optimizing interval lengths in the universe of discourse with an intent to improve forcasting accuracy of model. To evaluate performance of the proposed model, a numerical dataset of COVID-19 in Vietnam utilized to illustrate for forecasting process. The experimental results bring a significant meaning for the future and is a reference that can help people make decisions in term** of disease outbreaks.

Keywords — Forecasting, forecasting rules, fuzzy time series, fuzzy relationship, particle swarm optimization. COVID-19 in Vietnam

I. INTRODUCTION

The World Health Organization (WHO) has declared the epidemic of the new coronavirus to be a pandemic as health authorities around the world continue to struggle to contain the disease, which was first detected in the central city of Wuhan. The virus, which causes a respiratory disease called COVID-19, has spread following exponentially in the number of such cases around the globe. As of 8th April 2020, the total confirmed cases of 1.676.213 out of which 87.685 have died. It can very well be observed that most countries have reported cases of new Coronavirus. As of now, it has affected 210 countries with COVID-19, including Vietnam. Considering Vietnam, after first emerging in late January 2020, the number remained constant until the beginning of March, when the 17th case returned from London, England, it grew exponentially. As of 7th April, the number of total cases has reached 245 cases. Given the current rate of growth, where can the cases expect to reach in the next 15 days, if no specific precaution is taken. In this context we are really worry of the coming days, and months, everybody now over world, and especially in Vietnam country, ask the question: what's the trend of this virus? under this serious question. We will try to give a response by using the forecasting model based on Fuzzy Time Series.

The fuzzy time series forecasting models based on fuzzy set theory [1] 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 [9] 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], [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 aforementioned 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 the number of confirmed cases of COVID-19. In this paper, a forecasting model based on Fuzzy time series and PSO is presented to forecast the confirmed cases of the COVID-19 in Vietnam from the 4 march 2020 to 8 April 2020. Firstly, the proposed method fuzzifies the historical data of the COVID-19 into fuzzy sets to construct fuzzy relationship. Then, from these fuzzy relations, forecasting results are obtained by new defuzzification technique. Finally, the PSO algorithm for the optimal lengths of intervals is developed by searching the space of the universe of discourse. For model verification, a numerical dataset of COVID-19 in Vietnam utilized to illustrate for forecasting process.

The rest of this paper is organized as follows. Basic definitions of fuzzy time series and PSO algorithm are given in the succeeding section. The model to forecast the confirmed cases of the COVID-19 in Vietnam based on the FTS and PSO are presented in Section 3. The results obtained from the implementation of the proposed method are presented in Section 4. Finally, conclusions are presented in Section 5.

II. FUZZY TIME SERIES AND PSO ALGORITHM

Fuzzy time series was firstly put forward by Song and Chissom [2 - 5], [7] Fuzzy time series can be shortly defined as time series whose observations are fuzzy sets. Let $\mathrm{U} = \{\mathrm{u}_1, \mathrm{u}_2, \dots, \mathrm{u}_\mathrm{n}\}$ be an universal set; a fuzzy set A_i of U is defined as $A_i = \{ \mathrm{f_A}(\mathrm{u}_1)/\mathrm{u}_1 +, \mathrm{f_A}(\mathrm{u}_2)/\mathrm{u}_2 \dots + \mathrm{f_A}(\mathrm{u_n})/\mathrm{u_n} \}$, where f_A is a membership function of a given set A, f_A : U $\rightarrow$ [0,1], $\mathrm{f_A}(\mathrm{u_i})$ indicates the grade of membership of ui in the fuzzy set A, $\mathrm{f_A}(\mathrm{u_i}) \in [0, 1]$, and 1≤ i ≤ n. General definitions of FTS and PSO algorithm [13] are given as follows:

Definition 1: Fuzzy time series

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)* be a collection of $f_i(t)$) (i = 1, 2…). Then, *F(t)* is called a fuzzy time series on *Y(t)* (t . . ., 0, 1,2, . . .).

Definition 2: Fuzzy relationship (FR) [2,3]

If there exists a fuzzy relationship R(t-1,t), such that F(t) = F(t-1)$*$R(t-1,t), where "$*$" is an arithmetic operator, then F(t) is said to be caused by F(t-1). The relationship between F(t) and F(t-1) can be denoted by F(t-1)→ F(t). Let A_i = F(t) and A_j = F(t-1), the relationship between F(t) and F(t -1) is denoted by fuzzy logical relationship $A_i \rightarrow A_j$ where A_i and A_j 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 $F(t-1),\ F(t-2), \dots, F(t-m+1)\ F(t-m)$ then this fuzzy relationship is represented by $F(t-m), \dots, F(t-2),\ F(t-1) \rightarrow F(t)$ and is called an m - order fuzzy time series.

PSO algorithm is introduced Kennedy and Eberhart [13] for dealing with optimization problems, where a set of potential solutions is represented by a swarm of particles and each particle is move through the search space for search the optimal solution. When particles moving, the position of the best particle among all particles found so far is recorded and each particle keeps its personal best position which has passed previously. The particles change its state according to the three principles: weight inertia i.e. ω , it's most optimist position i.e. $\mathrm{P_{best\_id}}$, swarm's most optimist position i.e. $\mathrm{G_{best}}$; and converges to the most optimal position in the entire solution space by continuous change in the personal best and global best position. Each element v_{id}^{k} in the velocity vector $\mathrm{V}_{id}^{k} = [\mathrm{v}_{id}^{1}, \mathrm{v}_{id}^{2}, \dots, \mathrm{v}_{id}^{n}]$ and each element x_{id}^{k} in the position vetor $\mathrm{X}_{id}^{k} = [\mathrm{x}_{id}^{1}, \mathrm{x}_{id}^{2}, \dots, \mathrm{x}_{id}^{n}]$ of particle id are calculated as follows:

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

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

The G_{best} at k^{th} iteration is computed as:

$$G_{best} = \min(P_{best_{id}}^{k}) \; ; \quad (3)$$

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 $[-V_{max}, V_{max}]$ where V_{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.

III. Forecasting Model Based on Weighted Fuzzy Relation Matrices and PSO

In this section, a forecasting model based on article [13] by combining the fuzzy logical relationship group with PSO algorithm is introduced. First, The original historical COVID-19 dataset 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.

A. Forecasting model using FTS.

The daily confirmed cases of COVID-2019 from March 4th, 2020 to April 8th, 2020 are used to illustrate the first - order fuzzy time series forecasting process. It was collected from the WHO website (https: //www.who.int/emergencies/diseases/novel - coronavirus-2019/situation-reports/) or https://ncov.moh.gov.vn and which listed in Table 1. The step-wise procedure of the proposed model is detailed as follows:

TABLE I: THE ORIGINAL HISTORICAL COVID-19 DATASET IN VIETNAM

Stt	Date(D/M/Y)	Confirmed Cases
1	4/3/2020	16
2	5/3/2020	17
3	6/3/2020	17
4	7/3/2020	20
5	8/3/2020	21
---	------	---
31	02/04/2020	227
32	03/04/2020	233
33	04/04/2020	239
34	05/04/2020	241
35	06/04/2020	245
36	07/04/2020	245
37	08/04/2020	251

Step 1: *Define the universe of discourse U*

Assume Y(t) be the historical data of enrolments at year t. 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 integers to tune the lower bound and upper bound of the U. From the historical data are shown in Table 1, we obtain $I_{min} = 16$ và $I_{max} = 251$. Thus, the universe of discourse is defined as U= $[I_{min} - N_1, I_{max} + N_2]$= [16, 260.02] with $N_1 = 0$ and $N_2 = 9.02$.

Step 2: Divide U into equal length intervals.

Compared to the previous models in [2,5], we cut U into seven intervals, $u_1, u_2, \ldots, u_7$, respectively. The length of each interval is $l = \frac{D_{max} - D_{min}}{7}$. Thus, the seven intervals are: $u_1 = (D_{min}, D_{min}+l\,]$, $u_2 = (D_{min} + l, D_{min} + 2l]$, …, $u_7 = (D_{min} + 6l, D_{max}]$. From dataset in Table 1, we get 7 intervals as follows:

u_1 = [16, 50.86], u_2 = (50.86, 85.72], …, u_6 = (190.13.6, 225.16], u_7 = (225.16, 260.02].

Step 3: Define the fuzzy sets

Each interval in Step 2 represents a linguistic variable of “COVID-19”. For seven intervals, there are seven linguistic values [13] which are A_1 = “not many”, A_2=“not too many”, A_3 =“ many”, A_4 = “many many”, A_5 = “very many”, A_6 = “too many”, and A_7=“ too many many” 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 according to Eqs.(4) & (5) as follows.

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

For simplicity, the membership values of fuzzy set A_i either are 0, 0.5 or 1, where1 ≤ *i* ≤ 7. The value 0, 0.5 and 1 indicate the grade of membership of u_j in the fuzzy set A_i according to Eq.(5).

$$a_{ij} = \begin{matrix} 1 & \text{if } j == i \\ 0.5 & \text{if } j == i-1 \text{ or } j == i+1 \\ 0 & \text{otherwise} \end{matrix} \quad (5)$$

Step 4: Fuzzy all historical data of COVID-19

In order to fuzzify all historical data in Table 1, 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 COVID-19 dataset on 4 march, 2020 is 16, and it belongs to interval u_1 = [16, 48.71]. So, we then assign the linguistic value "not may" (eg. the fuzzy set A_1) corresponding to interval u_1 to it. Consider two time serials data $Y(t)$ and $F(t)$ at day t, where $Y(t)$ is actual data and $F(t)$ is the fuzzy set of $Y(t)$. According to Eqs. (4) and (5), the fuzzy set A_1 has the maximum membership value at the interval u_1. Therefore, the historical data time series on 4 march, 2020 is fuzzified to A_1. The completed fuzzified results of COVID-19 are listed in Table 2.

TABLE II. THE RESULTS OF FUZZIFICATION FOR COVID-19 DATA.

Year	Confirmed Cases	Fuzzy sets
4/3/2020	16	A1
5/3/2020	17	A1
6/3/2020	17	A1
-----	----	----
06/04/2020	245	A7
07/04/2020	245	A7
08/04/2020	251	A7

Step 5: Generate all m - order fuzzy relationships (FR)

Relationships are identified from the fuzzified historical data obtained in Step 4. If the fuzzified historical data of date t and t - 1 are A_i and A_k, respectively, then construct the first – order fuzzy logical relationship $A_i \rightarrow A_k$, where A_i and A_k are called the fuzzy set on the left-hand side and fuzzy set on the right-hand side of fuzzy logical relationships, respectively.

Similarly for m - order fuzzy relationship, we should find out any relationship which has the $F(t-m), F(t-m+1), \ldots, F(t-1) \rightarrow F(t)$, where $F(t-m), F(t-m+1), \ldots, F(t-1)$ 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.

For example, supposed $m = 3$, a fuzzy relationship $A_1, A_1, A_1 \rightarrow A_1$ is got as F(4/3/2020), F(5/3/2020). F(6/3/2020) → F(7/3/2020). So, from Table 2 we get third - order fuzzy relationships are shown in Table 3.

TABLE III. THE THIRD - ORDER FUZZY RELATIONSHIPS

No	Fuzzy relationships
1	A1, A1, A1 -> A1
2	A1, A1, A1 -> A1
--	--------------------
28	A6, A7, A7 –> A7
29	A7, A7, A7 -> A7
30	A7, A7, A7 -> A7
31	A7, A7, A7 -> A7
32	A7, A7, A7 -> A7

Step 6: Establish forecasting rules and compute the forecasted output values

Suppose there a fuzzy relation $A_i \rightarrow A_k$; where A_i, the current state, is the fuzzified COVID-19 data of date t and A_k, the next state, is fuzzified COVID-19 data of date t + 1

To obtain the forecasted results, we present a new defuzzification technique to compute the forecasted values for all fuzzy relationships in training phase.

The proposed model employs the historical data of dates t − 2, t − 1, t for framing rules to implement on fuzzy relation, $A_i \rightarrow A_k$. In the proposed algorithm, new set of difference parameters is computed for each forecast rather than using the same set of difference parameters. The proposed method for forecasting is mentioned as computational algorithms for generating the relations between the historical data of dates t − 2, t − 1, t for forecasting the COVID-19 Confirmed Case date t + 1, then calculate forecasting value for date t + 1 forwards (i.e., F (7/3/2020)) and onwards. The step-wise procedure of the proposed algorithm is detailed as follows:

<u>The Forecasted Computational algorithm</u>

In put: Time series
Output: - fuzzy relation for date t to t + 1: $A_i \rightarrow A_k$
- Forecasted value: F_k
Begin
For t = 3 to n (end of time series data)
for s = t − 2 to t + 1
$D_s = \|\| (R_s - R_{s-1}) \| - \|(R_{s-1} - R_{s-2})\|\|$
$P1_s = R_s + D_s/2$
$P2_s = R_s - D_s/2$
$Q1_s = R_s + D_s$
$Q2_s = R_s - D_s$
Next s;
For j = t − 2
If ($P1_j$>= Lo_A_k && $P1_j$ <= Up_u_k)
{ $W_1 = P1_j$; a = 1}
Else { W_1 = 0; a = 0}
end if
For j = t - 1
If ($P2_j$>= Lo_A_k && $P2_j$ <= Up_u_k)
{ $W_2 = P2_j$; b = 1}
	else { W_2 = 0; b = 0}
	end if
	For j = t
	If ($Q1_j$>= Lo_A_k && $Q1_j$ <= Up_u_k)
	{ $W_3 = Q1_j$; c = 1}
	else { W_3 = 0; c = 0}
	end if
	For j = t+1
	If ($Q2_j$>= Lo_A_k && $Q2_j$ <= Up_u_k)
	{ $W_4 = Q2_j$; d = 1}
	Else { W_4 = 0; d = 0}
	end if
	w = W_1 + W_2 + W_3 + W_4
	If B = 0 Then F_k = Mid_A_k
	else F_k = (w+Mid_ A_k)/(a+b+c+d+t+1)
	end if
	Next t
	End begin.

Where;

- A_k denotes corresponding interval u_k for which highest membership in A_k takes place in this interval.
- Lo_A_k is the lower bound of interval u_k
- Up_u_k is the upper bound of interval u_k
- leg (A_k) is the length of the interval u_k whose membership in A_k is maximum
- Mid_A_k is the midpoint value of the interval u_k having the highest value in A_k
- R_k is the actual data of date t
- R_{k-1} is the actual data of date t − 1
- R_{k-2} is the actual data of date t − 2

- F_k is the crisp forecasted value of the date t + 1

The forecasting performance can be assessed by comparing the difference between the forecasted values and the actual values. Mean squared error (MSE) is common tool to measure the accuracy in fuzzy time series forecasting. Lower the MSE better the forecasting method. The MSE is defined as follows:

$$MSE = \frac{1}{n}\sum_{i=m}^{n}(F_i - R_i)^2 \qquad (6)$$

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

B. Forecasting model combining the FTS and PSO.

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 model for forecasting the confirmed cases COVID-19 in Vietnam 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 (6). In proposed 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 p_i where $1 \leq i \leq$ n-1 and $p_i \leq p_{i+1}$. Based on these n-1 elements, define the n intervals as $u_1 = [p_0, p_1]$, $u_2 = [p_1, p_2]$,..., $u_i = [p_{i-1}, p_i]$,…, $u_{n-1} = [p_{n-2}, p_{n-1}]$ 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 p_i ($1 \leq i \leq n-1$) arranges in an ascending order. The complete steps of the proposed method are presented in Algorithm 1.

***Algorithm 1*: The FTS and PSO algorithm**

1. initialize N particles and particles' positions X_i and velocities V_i

2. **while** the stop condition (maximum iterations) is not satisfied **do**

2.1. **for** particle i, ($1 \leq i \leq$ N) **do**

- ✓ Define fuzzy sets based on the current position of particle i by Step 3
- ✓ Fuzzify all historical data by Step 4
- ✓ Establish all $m-$ order fuzzy relationships by Step 5
- ✓ Calculate forecasting values by Step 6
- ✓ Compute the MSE values for particle i based on (6)
- ✓ 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, ($1 \leq i \leq$ N) **do**

- ✓ move particle i to another position according to (1) and (2)

end for

end while

IV. Experimental Results

In this paper, the COVID-19 dataset in Vietnam is used to evaluate the effectiveness of the proposed model. It contains the daily confirmed cases in Vietnam from 4 march 2020 to 8 April 2020, as shown in Table 1.The 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 Table 6. The forecasted accuracy of the proposed method is estimated using the MSE value (6). The forecasted results of proposed model under number of interval is 16 and various orders are listed in Table 7.

Table VI: Parameters used the proposed forecasting model

Number of particles N	30
Maximum number of iterations	100
The value of inertial weigh **ω** be linearly decreased	0.9 to 0.4
The coefficient $C_1 = C_2$	2
The velocity V_i be limited to	[-50,50]
The position X_i be limited to	[16, 260.02]

TABLE VII: The completed forecasting results the confirmed cases COVID-19 in Viet Nam based on different orders

Date	Actual data of Covid-19	Forecasted results					
		1st - order	2nd - order	3rd-order	4th-order	5th-order	6th-order
4/3/2020	16	-	-	-	-	-	
5/3/2020	17	21.9	-	-	-	-	
6/3/2020	17	21.9	22.2	-	-	-	
7/3/2020	20	21.9	22.2	23.7	-	-	
8/3/2020	21	21.9	22.2	23.7	27.8	-	
9/3/2020	30	21.9	34	23.7	27.8	29.6	
10/3/2020	31	36.4	37.1	31	27.8	29.6	36
11/3/2020	34	36.4	37.1	34.5	27.8	29.6	36
12/3/2020	34	36.4	37.1	38.2	36.3	35.3	36
--------	---	---	---	---	---	---	---

01/04/2020	218	220	218.7	217	222.3	219.3	219.7
02/04/2020	227	230	224.3	229	226.5	224.7	225.3
03/04/2020	233	230	234.3	232.5	230.7	236.7	236.3
04/04/2020	239	242.6	238.7	236	241.3	240	239
05/04/2020	241	242.6	243.8	242.5	241.3	240	239
06/04/2020	245	242.6	243.8	242.5	243.7	243.3	241.7
07/04/2020	245	251.7	249	249	243.7	243.3	241.7
08/04/2020	251	251.7	251	252	251	253.7	251.7
09/04/2020		250.6	251.4	252	251.7	252.2	250.2
MSE		**22.7**	**17.1**	**15.9**	**14.3**	**3.8**	**4.1**

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 MSE criteria in Eq.(6).

In addition, from the parameters are expressed in Table 6. The proposed model is also executed 15 runs according to the different number of intervals, and the best result of runs is taken to be the final result. The simulation results according to the intervals of proposed model are presented in Table 8. In Table 8, it can be seen that the accuracy of the proposed model is improved significantly. Particularly, the proposed model gets the smallest MSE value with number of interval equal to 16 and obtains the average MSE value of ***34.01***. These finding suggest that the proposed model is able to provide effective forecasting capability for the fuzzy time series model with different number of orders in the same number of interval.

TABLE VIII: THE FORECASTED RESULTS OF THE PROPOSED MODEL BASED ON THE FIRST – ORDER FUZZY TIME SERIES WITH DIFFERENT NUMBER OF INTERVALS.

	Number of intervals								
	10	11	12	13	14	15	16	17	Average
MSE	52.6	46.5	39.3	31	29.4	26.2	25.1	22	**34.01**

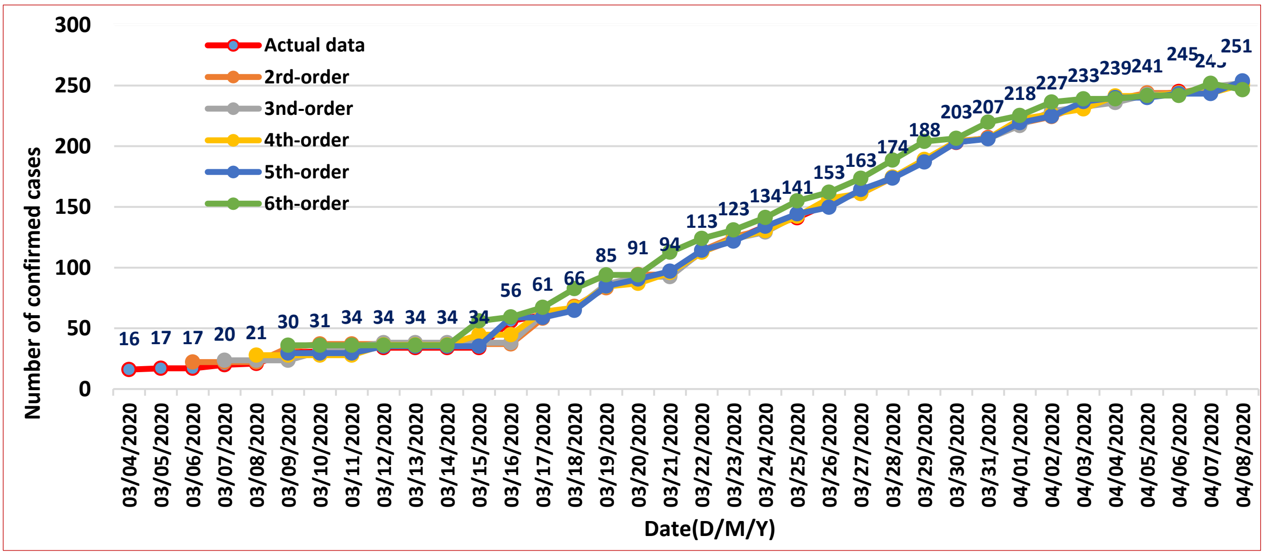

Fig. 1: The forecasting trend of proposed model under different number of orders with the same intervals

V. CONCLUSIONS

In this study, we briefly and speedily as corona virus spread over world, studied the trend of this virus in Vietnam. To check this hypothesis, we have introduced a hybrid forecasting model by combining fuzzy time series with PSO to forecast the daily confirmed cases in Vietnam from 4 march 2020 to 8 April 2020. The major contributions of this paper are illustrated in the following. First, fuzzy relationships are established from the historical data of COVID-19 after it is fuzzified. Then, the forecasted values obtained by a new forecasting technique is introduced in Step 6 of the model. The PSO algorithm for the optimized lengths of intervals is developed to adjust the interval lengths by searching the space of the universe. Finally, dataset of the daily COVID-19 confirmed cases in Vietnam were used to forecast, and the confirmed cases evaluation outcomes showed it's the best performance based on 5th - order fuzzy time series with number of interval equal to 16. According to the promising results achieved from the proposed model, it can be applied in different forecasting problems.

The proposed model was only established by one factor of COVID-19 confirmed cases data. In the further research, we will add the impact of factors such as population birth rate and natural mortality into forecasting model.

ACKNOWLEDGMENT

The author thanks the support of Scientific Council of Thai Nguyen University of Technology (TNUT) to 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.