A HYBRID PSO-SA SCHEME FOR IMPROVING THE ACCURACY OF FUZZY TIME SERIES FORECASTING MODELS

PHAM DINH PHONG¹, NGUYEN DUC DU¹*, PHAM HOANG HIEP², TRAN XUAN THANH³

¹Faculty of Information Technology, University of Transport and Communications, Ha Noi, Viet Nam
²HUS High School For Gifted Students, VNU Ha Noi - University of Science, Viet Nam
³Faculty of Information Technology, East Asia University of Technology, Bac Ninh, Viet Nam

Abstract. Forecasting methods based on fuzzy time series have been examined intensively during the last few years. Three main factors which affect the accuracy of those forecasting methods are the length of intervals, the way of establishing fuzzy logical relationship groups, and defuzzification techniques. Many researchers focus on studying the methods of optimizing the length of intervals to improve forecasting accuracies by utilizing various optimization techniques. In line with that research trend, this paper proposes a hybrid algorithm combining particle swarm optimization with the simulated annealing technique (PSO-SA) to optimize the length of intervals to improve forecasting accuracies. The experimental results on the datasets of the “enrolments of the University of Alabama,” “killed in car road accidents in Belgium,” and the “spot gold in Turkey” have shown that the proposed forecasting model is more effective than their counterparts.

Keywords. Fuzzy time series; Particle swarm optimization; Simulated annealing.

1. INTRODUCTION

Time series (TS) modeling and forecasting have attracted the research community’s attention over the last few years. Some TS forecasting models based on the probabilistic approach, such as ARMA, MA, and ARIMA [1], etc., have been proposed. Those models have good forecasting results on the large observations (greater than 50) and cannot forecast the TS whose values are linguistic terms such as “slow”, “medium”, “quick”, “very quick”, and so on.

In 1993, Song and Chissom proposed the fuzzy time series forecasting model (FTS-FM), in which the values of demand variables are linguistic values, and applied it to the forecasting problem of the “enrolments of the University of Alabama” (EUA) [3, 4]. That model uses the min-max composition operation in fuzzy relations leading to a large amount of computational time. Chen enhanced that model using simple fuzzy reasoning and defuzzification methods [5]. Yu proposed weighted fuzzy time series models for forecasting TAIEX [6] by assigning

*Corresponding author.
E-mail addresses: phongpd@utc.edu.vn (P.D. Phong); nducdu@utc.edu.vn (N.D. Du); phamhoanghiep03092004@gmail.com (P.H. Hiep); thanhtx@eaut.edu.vn (T.X. Thanh).

© 2022 Vietnam Academy of Science & Technology
weights to fuzzy relationships to resolve the recurrence of fuzzy relationships and reflect their
different importance. Those forecasting models include three main phases: 1) Fuzzify the
universe of discourse (UD) of TS using fuzzy sets; 2) Establish fuzzy logical relationship
groups (FLRGs) for fuzzy reasoning. 3) Forecast to get fuzzy outputs and then defuzzify the
fuzzy outputs to get crisp data. Since then, there have been many studies to improve the
effectiveness of FTS-FMs and applied them to many forecasting problems in the real world.

The accuracy of FTS-FMs depends much on the three phases described above. The first
factor is the length of intervals. In [2-4], Song and Chen partitioned the UD of historical
data of the EUA into seven equal-length intervals without expressing any reason. Huarng [7]
recognized that the interval length greatly influences the accuracy of FTS-FMs. The interval
length should be neither too small nor too large. The interval length is too small, resulting
in meaningless fuzzy time series (FTS). Whereas the interval length is too large, there is no
fluctuation in FTS. Then, he introduced two approaches to determine the effective lengths
of intervals: average- and distribution-based length. Later, Chen [8], Bas [9], and Lee [10,
11] applied the genetic algorithm to adjust the length of intervals. A genetic algorithm
integrated with an automatic clustering technique to tune interval length was proposed by
Wang [12]. Kuo et al. proposed two hybrid forecasting models, HPSO [13] and NPSO [14],
based on the integration between FTS and PSO. The difference between HPSO and NPSO
is only the forecasting rule. More information on each next state of FLRGs is considered
in the NPSO model, so it is more accurate than the HPSO. A new model aggregating both
the global information and the local information was proposed by Huang et al. [15]. Some
other methods of partitioning the UD based on information granules [16, 17], ant colony
optimization and auto-regression [18], fuzzy clustering [19, 20], rough-fuzzy [21, 22], etc.,
were proposed.

The second factor which affects forecasting accuracy is establishing FLRGs for forecast-
ing. Chen introduced the high-order model in [23]. Two-factors high-order models were
introduced by Lee [11] and Wang [12]. Yu introduced refinement relation [24] and weighted
scheme [6] models. A method to construct the FLRs into FLRGs by using the K-means
clustering algorithm was proposed by Cheng et al. [25]. We can easily see that in most of
the mentioned research, the time-invariant FLRGs are constructed for forecasting. It means
that all FLRs with the same left-hand side (LHS) are grouped into an FLRG regardless of
whether they occurred in the past or future. The concept of time-variant FLRG was applied
by Tinh and Dieu in [20]. Phong proposed linguistic time series with linguistic forecasting
rules in [26].

The third factor affecting forecasting accuracy is the defuzzification technique. Chen
proposed the average of the mid-points of fuzzy intervals corresponding to fuzzy sets of the
right-hand side (RHS) of FLRG as the crisp forecasted value of the current forecasting time
[5]. Yu granted the weights in chronological order to fuzzy sets in the RHS of FLRGs [6]. In
[14], Kuo used the information of sub-intervals of each fuzzy interval of the next state in the
RHS to compute crisp forecasted values. A new model aggregating the global information
of FLRs with the local information of the latest fuzzy fluctuation to find forecasting value
was proposed by Huang et al. [15]. A quite efficient defuzzification technique based on the
proportions of intervals was proposed by Chen et al. in [27].

There are numerous optimization techniques applied to solve FTS forecasting problems.
The intensively examined one is particle swarm optimization (PSO). However, PSO is ef-
cient for global search but weak for local search. Therefore, it is easily trapped into the local optimums and becomes premature convergence. This paper applies a hybrid algorithm combining PSO with the simulated annealing (SA) technique to optimize the length of intervals of the UD to enhance the forecasting accuracies of the FTS-FMs. The SA technique helps PSO jump out of the local optimums to continue its searching process. The experimental results on three datasets of the EUA, the “killed in car road accidents in Belgium” (CAB), and the “spot gold in Turkey” (SGT) show that our proposed forecasting model has better-forecasted accuracy than the existing ones.

The organization of the paper is as follows: Section 2 is some basic concepts of FTS and PSO. The proposed FTS-FM is presented in Section 3. Section 4 shows the experimental results and discussion. Conclusions are written in Section 5.

2. PRELIMINARIES

2.1. Some basic concepts

Song and Chissom introduced FTS-FM in 1993 [2–4]. In 1996, Chen enhanced that model using a simple defuzzification technique. Those FTS-FMs were developed based on the following concepts.

Definition 1 [2-4]. Let \( T(t) \) \((t = \ldots, 0, 1, 2,\ldots)\) be a subset of \( R^1 \), where \( t \) is the temporal variable. \( T(t) \) is the UD on which the fuzzy sets \( f_i(t), i = 1, 2, \ldots \) are defined. If \( F(t) \) is a series of fuzzy sets \( f_i(t) \) \((i = 1, 2,\ldots)\), then \( F(t) \) is called a fuzzy time series on \( T(t) \).

Definition 2 [5]. Fuzzy logical relationship (FLR). At the times \( t - 1 \) and \( t \), if there exists a fuzzy relationship \( R(t - 1, t) \) between \( F(t - 1) \) and \( F(t) \) such that

\[
F(t) = F(t - 1) \ast R(t - 1, t),
\]

where \( \ast \) is an operator, \( F(t) \) is said to be inferred from \( F(t - 1) \). The relationship between \( F(t - 1) \) and \( F(t) \) is defined as \( F(t - 1) \rightarrow F(t) \). If \( F(t - 1) = X_i \) and \( F(t) = X_j \), the logical relationship between \( F(t - 1) \) and \( F(t) \) is denoted by \( X_i \rightarrow X_j \), where \( X_i \) is the LHS (current state) and \( X_j \) is the RHS (next state) of the fuzzy relation.

Definition 3 [5]. The FLRs with the same LHS are grouped together to form fuzzy logical relationship groups. For example, there are FLRs \( X_i \rightarrow X_j_1, X_i \rightarrow X_j_2, \ldots, X_i \rightarrow X_j_n \) that can be put into a group denoted by \( X_i \rightarrow X_j_1, X_j_2, \ldots, X_j_n \).

2.2. Fuzzy time series forecasting models

2.2.1. The FTS-FM of Chen

In the FTS-FM of Chen, FLRGs are established, and simple arithmetic operations are used to compute crisp forecasted values instead of complex min-max composition operations in FLRs. Hereafter is a brief description of the forecasting model of Chen [5]:

Step 1. Specify the UD of TS and partition it into equal-length intervals \( u_1, u_2, \ldots, u_n \).

Step 2. Design fuzzy sets on UD.

Step 3. Fuzzify historical data.

Step 4. Generate FLRs, then establish FLRGs.
Step 5. Forecast to get fuzzy forecasted values. Then, defuzzify fuzzy forecasted values to get crisp ones using defuzzification principles as follows:

**Principle 1.** If there exists FLRG $X_i \rightarrow X_j$ and the midpoint of $u_j$ is $v_j$, the crisp value of forecasting time is $v_j$.

**Principle 2.** If there exists FLRG $X_i \rightarrow X_{j1}, X_{j2}, ..., X_{jk}$, where $X_i$ is the fuzzy set of a time, say $t$, and the midpoints of $u_{j1}, u_{j2}, ..., u_{jk}$ are $v_{j1}, v_{j2}, ..., v_{jk}$, respectively, the crisp forecasted value of time $t + 1$ is specified as

$$\text{CFV}_{t+1} = \frac{v_{j1} + v_{j2} + \ldots + v_{jk}}{k}.$$  \hspace{1cm} (1)

**Principle 3.** If there exists FLRG $X_i \rightarrow \emptyset$, where the notion of empty set denotes that the RHS of this FLRG is empty, and $v_i$ is the midpoint $u_i$, the crisp forecasted value is $v_i$.

2.2.2. The FTS-FM of Yu

Unlike the FTS-FM of Chen, a fuzzy set can be repeated on the RHS of an FLRG of the FTS-FM of Yu. Therefore, fuzzy sets in the RHS of FLRGs are granted weights in chronological order to reflect their different importance. The defuzzification principle 2 of FTS-FM of Chen described above is modified as follows: if there exists FLRG $X_i \rightarrow X_{j1}, X_{j2}, ..., X_{jk}$, where $X_i$ is the fuzzy set of a time, say $t$, and the midpoints of $u_{j1}, u_{j2}, ..., u_{jk}$ are $v_{j1}, v_{j2}, ..., v_{jk}$, respectively, the crisp forecasted value of time $t + 1$ is specified as

$$\text{CFV}_{t+1} = \frac{1 \times v_{j1} + 2 \times v_{j2} + \ldots + k \times v_{jk}}{1 + 2 + \ldots + k}.$$ \hspace{1cm} (2)

2.3. Particle swarm optimization

Particle swarm optimization (PSO), which was introduced by Kennedy and Eberhart in 1995 [28, 29], mimics the behavior of birds flying to find food sources. Hereafter is a brief description of basic PSO.

Assume that we have a swarm $S = \{x_1, x_2, \ldots, x_N\}$, where $x_i$ is a particle having its position $Y^t_i$ at cycle $t$ computed as

$$Y^{t+1}_i = Y^t_i + V^{t+1}_i,$$ \hspace{1cm} (3)

where $V^{t+1}_i$ is the velocity of particle $x_i$ at cycle $t + 1$, which is computed as

$$V^{t+1}_i = \omega V^t_i + c_1 r_1 (P^t_g - Y^t_i) + c_2 r_2 (P^t_l - Y^t_i),$$ \hspace{1cm} (4)

where $P^t_g$ and $P^t_l$ are the global and local solutions that are found up to cycle $t$, respectively; $c_1$ and $c_2$ are self-cognitive and social cognitive factors; $r_1$ and $r_2$ are two random numbers which uniformly distribute in $[0, 1]$; $\omega$ is inertia weight. Hereafter is the basic PSO procedure:

- **Step 1**: Randomly initialize a swarm $S$ with its vector of velocity $V$ and vector of position $Y$, the iterative variable $t$, and the number of cycles $G_{\text{max}}$.

- **Step 2**: Calculate the value of objective function $f(Y^t_i)$ of particle $x_i$.

- **Step 3**: Compare the value of $f(Y^t_i)$ with the one of $f(P^t_l)$. Update $P^t_l$ if $f(Y^t_i)$ is better.
Step 4: Update the global best position $P^t_g$.
Step 5: Update $V^t_i$ and $Y^t_i$ by Eq. (4) and Eq. (3), respectively.
Step 6: Terminate if $t > G_{\text{max}}$. Otherwise, increase variable $t$ and go to Step 2.

2.4. Simulated annealing algorithm

Simulated annealing (SA) [30] is an algorithm operating based on the process of metal cooling in metal annealing. SA begins at a high temperature ($T_0$) when the metal is in a molten state. The temperature of metal begins to gradually decrease to the ambient temperature ($T_{\text{min}}$) after removing the heating source. At this temperature, the energy of metal reaches the minimal value, and the metal is in a solid state.

The brief description of SA with minimizing energy $E$ is as follows:

Step 1. Initialize an energy state $E_j$ with the cooling rate $\alpha \in [0, 1]$, $T = T_0$, where $T_0$ is the initial temperature.
Step 2. Calculate the energy change between the present state $E_i$ and the previous one $E_j$ of the configuration

$$\Delta E = E_i - E_j.$$

Step 3. If $\Delta E < 0$, new state $E_i$ is accepted (up-hill). Otherwise, the new state $E_i$ is accepted (down-hill) with probability $P = e^{-\frac{\Delta E}{k_B T}}$, where $k_B$ is the Boltzman constant.
Step 4. Terminate if reaching the termination condition. Otherwise, decrease the temperature $T = \alpha T$ and go to Step 2.

3. THE PROPOSED FUZZY TIME SERIES FORECASTING MODEL

In PSO, particles inside the swarm are considered solutions to the problems and explored throughout the solution space to seek the best solutions. Therefore, PSO is very effective in global search but weak in local search. In fact, particles easily get stuck in local optimums, and it is difficult for them to jump out to continue their searching process because of the update mechanism of the velocity equation. Whereas SA has the ability to jump out of local optimums to continue the search process with the help of the "Metropolis law." In [31], PSO-SA is applied efficiently to classification problems. This section presents the proposed FTS-FM with the application of PSO-SA to improve forecasting results.

3.1. A hybrid PSO-SA algorithm

The hybrid PSO-SA algorithm is a combination of PSO with SA, so-called PSO-SA. It makes use of the global search and local search made by PSO and SA, respectively. Hereafter is a brief description of PSO-SA:

Step 1: Initialize a random swarm with $n$ particles and all necessary variables, including cycle step $t$, initial temperature $T_0$, and cooling rate $\alpha$. Evaluate the objective value of each particle.
Step 2: For each particle $x_i$ in the swarm.
Step 2.1: Compute particle’s velocity $V_i^{t+1}$ by formula (4).
Step 2.2: Compute new particle position $Y_i^{t+1}$ by formula (3).
Step 2.3: Evaluate objective values of particle \( x_i \).

Step 2.4: Compare the fitness values at the new position \( y_i^{t+1} \) and the old one \( y_i^t \). If the objective value at \( y_i^{t+1} \) is better than the one at \( y_i^t \), meaning that the new position is better, then accept \( y_i^{t+1} \) as the new position of \( x_i \). Otherwise, compute the surplus of objective functions \( \Delta F \) between \( y_i^{t+1} \) and \( y_i^t \) as the following formula

\[
\Delta F = (\text{fitness}_{i}^{t+1} - \text{fitness}_{i}^{t}).
\]  

Step 2.5: Generate a random number \( \sigma \in [0, 1] \). Accept new position if \( \sigma > e^{-\left(\frac{\Delta F}{T_t}\right)} \) or the number of rejects exceeds 100. Go to Step 2.6 in case the new position is accepted. Otherwise, go to Step 2.1.

Step 2.6: Update the local best position \( P_i^t \) of all particles and the global best position \( P_g^t \).

Step 3: If the termination condition is satisfied, terminate and output the best solution. Otherwise, modify annealing temperature \( T^{k+1} = \alpha T^k \), \( t = t + 1 \), and jump to Step 2.

The SA algorithm does the exploitation by repeating from Step 2.1 to Step 2.5 until the appearance of a better position (uphill) or accepting a worse position with the probability \( P = e^{-\left(\frac{\Delta F}{T_t}\right)} \) (accepting downhill). By accepting the downhill, the SA takes the opportunity to help PSO jump out of that local optimum to continue the search in other locations.

As the structure organization, the PSO-SA algorithm should take longer to run because it tries to do exploitation to get a better solution. The running time analysis is mentioned in Sub-section 4.4.

3.2. The proposed fuzzy time series forecasting model

In this subsection, a new FTS-FM is proposed in which the PSO-SA algorithm is applied to optimize the interval length to enhance forecasting accuracy. Each interval can be defined by its start and end points, forming a split point set. Therefore, it is necessary to determine the split points so that they form an interval set that minimizes the value of the mean square error (MSE) function used as the objective function, calculated by the formula (12).

Suppose the number of intervals of the UD is \( n \). Then, split the UD \( U = [p_0, p_n] \) into \( n \) intervals with the split points \( p_1, p_2, \ldots, p_{n-2}, p_{n-1} \). Therefore, we have interval set \( u_1 = [p_0, p_1], u_2 = [p_1, p_2], \ldots, u_n = [p_{n-1}, p_n] \). Each particle position in PSO-SA is represented by a vector with \( n - 1 \) elements \( Y_i = [p_1, p_2, \ldots, p_{n-2}, p_{n-1}] \), where each \( p_i \) \((i = 1, \ldots, n - 1)\) is a split point. PSO-SA will find \( Y_i \), which generates minimal MSE value.

The proposed FTS-FM in detail is as follows.

Step 1. Apply PSO-SA to optimize the length of intervals \( u_1, u_2, \ldots, u_n \) of \( U \).

Let \( D_{\text{min}} \) and \( D_{\text{max}} \) be the minimal and maximal values of the UD, respectively, and they are defined as \( D_{\text{min}} = F_{\text{min}} - N_l \) and \( D_{\text{max}} = F_{\text{max}} + N_h \), where \( F_{\text{min}} \) and \( F_{\text{max}} \) are the minimal and maximal values of the historical data, respectively, \( N_l \) and \( N_h \) are two positive integers used to adjust the lower and upper bounds of \( U \) so that \( U \) should cover all values that occur in the future. PSO-SA is applied to find the optimal interval set of \( U \), where vector \( Y_i \) represents the position of particle \( x_i \).

Step 2. Design the fuzzy sets on \( U \).

As with the other models, each interval of \( U \) is assigned a fuzzy set \( X_i \) associated with a linguistic label. The fuzzy set \( X_i \) is defined as follows
where $0 \leq a_{ij} \leq 1$, $1 \leq i, j \leq n$, is the grade of membership of $u_j$ to $X_i$. For simplicity, $a_{ij}$ just takes three different membership values of 0, 0.5, and 1. For example, $u_2$ belongs to $X_1$, $X_2$, and $X_3$ with the membership degrees of 0.5, 1, and 0.5, respectively, and belongs to the rest with the membership degree of 0. The symbol + means the set union operator and the division operator indicates the membership degree of $u_j$ ($1 \leq j \leq n$) to $X_i$, respectively.

**Step 3.** Fuzzy all historical data.

In this step, all actual data of TS is fuzzified by converting it into fuzzy data. Each actual data is assigned a fuzzy set with the largest membership degree. For example, the historical data of the EUA is partitioned into seven equal-length intervals. The intervals from 1 to 7 are assigned the linguistic labels, namely $X_1$, $X_2$, ..., and $X_7$, respectively. The fuzzified data is shown in Table 1.

**Table 1.** The fuzzified data of the EUA in case the historical data is partitioned into seven equal-length intervals.

| Year | Actual data | Fuzzy set | Year | Actual data | Fuzzy set |
|------|-------------|-----------|------|-------------|-----------|
| 1971 | 13055       | $X_1$     | 1982 | 15433       | $X_3$     |
| 1972 | 13563       | $X_1$     | 1983 | 15497       | $X_3$     |
| 1973 | 13867       | $X_1$     | 1984 | 15145       | $X_3$     |
| 1974 | 14696       | $X_2$     | 1985 | 15163       | $X_3$     |
| 1975 | 15460       | $X_3$     | 1986 | 15984       | $X_3$     |
| 1976 | 15311       | $X_3$     | 1987 | 16859       | $X_4$     |
| 1977 | 15603       | $X_3$     | 1988 | 18150       | $X_6$     |
| 1978 | 15861       | $X_3$     | 1989 | 18970       | $X_6$     |
| 1979 | 16807       | $X_4$     | 1990 | 19328       | $X_7$     |
| 1980 | 16919       | $X_4$     | 1991 | 19337       | $X_7$     |
| 1981 | 16388       | $X_4$     | 1992 | 18876       | $X_6$     |

**Step 4.** Create FLRs and establish FLRGs.

FLRs are created based on the above Definition 2 in such a way that an FLR is created by a fuzzy set associated with time $t$ on the LHS and a fuzzy set associated with time $t + 1$ on the RHS. For an example of the EUA in Step 3, the created FLRs are shown in Table 2. The procedure for establishing FLRGs is as follows:

- For $t = 1$ to $N - 1$ do begin // $N$ is the number of historical data
  - Establish the FLR $X_t \rightarrow X_{t+1}$: // $X_t$ and $X_{t+1}$ are the fuzzy sets at the time $t$ and $t + 1$, respectively.
  - Find the FLRG whose LHS is $X_t$. If found then append $X_{t+1}$ to the end of its RHS, otherwise create a new FLRG whose LHS and RHS are $X_t$ and $X_{t+1}$, respectively;

- End;

Once FLRs are created, FLRGs are established by grouping all FLRs with the same LHSs. Chen’s forecasting model is applied, so a fuzzy set cannot be repeated in the RHS of an LLRG. All FLRGs in case of seven equal-length intervals are shown in Table 3.

**Step 5.** Forecast to get fuzzy outputs and then defuzzify the fuzzy outputs to get crisp data.
Table 2. The FLRs of the EUA in case the historical data is partitioned into seven equal-length intervals.

| Year | Actual data | Fuzzy set | F(t) | First-order relationships |
|------|-------------|-----------|------|--------------------------|
| 1971 | 13055       | $X_1$     | $F(1971) \rightarrow F(1972)$ | $X_1 \rightarrow X_1$ |
| 1972 | 13563       | $X_1$     | $F(1972) \rightarrow F(1973)$ | $X_1 \rightarrow X_1$ |
| 1974 | 14696       | $X_2$     | $F(1973) \rightarrow F(1974)$ | $X_1 \rightarrow X_2$ |
| 1975 | 15460       | $X_3$     | $F(1974) \rightarrow F(1975)$ | $X_2 \rightarrow X_3$ |
| 1976 | 15311       | $X_3$     | $F(1975) \rightarrow F(1976)$ | $X_3 \rightarrow X_3$ |
| 1977 | 15603       | $X_3$     | $F(1976) \rightarrow F(1977)$ | $X_3 \rightarrow X_3$ |
| 1978 | 15861       | $X_3$     | $F(1977) \rightarrow F(1978)$ | $X_3 \rightarrow X_3$ |
| 1979 | 16807       | $X_4$     | $F(1978) \rightarrow F(1979)$ | $X_3 \rightarrow X_4$ |
| 1980 | 16919       | $X_4$     | $F(1979) \rightarrow F(1980)$ | $X_4 \rightarrow X_4$ |
| 1981 | 16388       | $X_4$     | $F(1980) \rightarrow F(1981)$ | $X_4 \rightarrow X_4$ |
| 1982 | 15433       | $X_3$     | $F(1981) \rightarrow F(1982)$ | $X_4 \rightarrow X_3$ |
| 1983 | 15497       | $X_3$     | $F(1982) \rightarrow F(1983)$ | $X_3 \rightarrow X_3$ |
| 1984 | 15145       | $X_3$     | $F(1983) \rightarrow F(1984)$ | $X_3 \rightarrow X_3$ |
| 1985 | 15163       | $X_3$     | $F(1984) \rightarrow F(1985)$ | $X_3 \rightarrow X_3$ |
| 1986 | 15984       | $X_3$     | $F(1985) \rightarrow F(1986)$ | $X_3 \rightarrow X_3$ |
| 1987 | 16859       | $X_4$     | $F(1986) \rightarrow F(1987)$ | $X_3 \rightarrow X_4$ |
| 1988 | 18150       | $X_6$     | $F(1987) \rightarrow F(1988)$ | $X_4 \rightarrow X_6$ |
| 1989 | 18970       | $X_6$     | $F(1988) \rightarrow F(1989)$ | $X_6 \rightarrow X_6$ |
| 1990 | 19328       | $X_7$     | $F(1989) \rightarrow F(1990)$ | $X_6 \rightarrow X_7$ |
| 1991 | 19337       | $X_7$     | $F(1990) \rightarrow F(1991)$ | $X_7 \rightarrow X_7$ |
| 1992 | 18876       | $X_6$     | $F(1991) \rightarrow F(1992)$ | $X_7 \rightarrow X_6$ |

Table 3. The FLRGs of the EUA in case the historical data is partitioned into seven equal-length intervals.

| Group | FLRGs |
|-------|-------|
| Group 1 | $X_1 \rightarrow X_1, X_2$ |
| Group 2 | $X_2 \rightarrow X_3$ |
| Group 3 | $X_3 \rightarrow X_3, X_4$ |
| Group 4 | $X_4 \rightarrow X_4, X_3, X_6$ |
| Group 5 | $X_6 \rightarrow X_6, X_7$ |
| Group 6 | $X_7 \rightarrow X_7, X_6$ |

In this step, each fuzzy forecasted value is implicated by the RHS (next states) of FLRG associated with time $t$. Then, that fuzzy forecasted value is defuzzified to get a crisp forecasted value. Defuzzification techniques have a strong effect on forecasting accuracy. **Principle 1.** If there is FLRG $X_i \rightarrow X_{j1}, X_{j2}, ..., X_{jk}$ ($k \geq 1$), where $X_i$ is the fuzzy set of a time, say $t$, then the fuzzy forecasted value of time $t + 1$ is $X_{j1}, X_{j2}, ..., X_{jk}$, and it should be defuzzified to get its crisp forecasted value.

In [6], Yu used different weights in chronological order (Eq. 4)) to reflect the different importance of repeated fuzzy sets, and forecasting accuracy is improved in most of the
experimental cases. In [20], Tinh et al. applied Eq. (3) to first-order and applied Eq. (7) to high-order time-variant FTS.

\[ CFV_{t+1} = \frac{1}{k} \sum_{j=1}^{k} \text{subm}_{jl} \]  

(7)

where \(1 \leq j \leq k\), \(k\) is the number of next states; \(\text{subm}_{jl}\) is the midpoint of one of \(p\) equal sub-intervals within interval \(u_{jl}\) of the next state, which the actual datum of the forecasting time having the maximum value of membership function of \(X_{jl}\) falls into. This defuzzification technique does not reflect the different importance of repeated fuzzy sets leading to the forecasting accuracy is not good in some cases (e.g., first-order time series models). In [14], Kuo added the midpoint of intervals of the next state \(v_{jl}\) to equation (7) as follows

\[ CFV_{t+1} = \frac{1}{k} \sum_{j=1}^{k} \frac{\text{subm}_{jl} + v_{jl}}{2}. \]  

(8)

In [27], Chen proposed a new defuzzification formula based on proportions of intervals as follows

\[ CFV_{t+1} = \frac{1}{p} \left\{ \left[ p_t \times (u_{k1max} - m_{k1min}) + u_{k1min} \right] + \ldots + \left[ p_t \times (u_{kpmax} - m_{k1min}) + u_{kpmin} \right] \right\}, \]  

(9)

where \(f_t\) is the actual datum at time \(t\), \(p_t = (f_t - u_{jmin}) / (u_{jmax} - u_{jmin}) \in [0, 1]\), which \(f_t\) falls into, \(u_{jmin}\) and \(u_{jmax}\) are the lower and upper bounds of \(u_j\), respectively, and \(p\) is the number of next states.

In [20], Tinh and Dieu proposed a new defuzzification formula which was applied to high-order FTS-FM as follows

\[ CFV_{t+1} = \frac{1}{2 \times n} \sum_{j=1}^{n} (\text{subm}_{jl} + \text{LB}_{jl}), \]  

(10)

where \(\text{LB}_{jl}\) is the lower bound of one of \(p\) equal sub-intervals within interval \(u_{jl}\) of the next state, the actual datum of the forecasting time having the maximum value of membership function of \(X_{jl}\) falls into in case the actual value is less than \(\text{subm}_{jl}\). Otherwise, \(\text{LB}_{jl}\) is the upper bound of it. This defuzzification formula will be applied to our proposed first-order FTS-FM.

In this paper, we will apply all the above-mentioned defuzzification techniques to the proposed FTS-FMs to evaluate them to show the best.

**Principle 2.** If there is FLRG \(A_1 \to \emptyset\), Kuo’s master voting scheme [13] is applied to compute the crisp forecasted values of the testing patterns. This voting scheme lets us put the weight for the latest past linguistic value

\[ CFV_{t+1} = \frac{v_{i1} \times w + v_{i2} + \ldots + v_{i\lambda}}{w + (\lambda - 1)}, \]  

(11)

where \(w\) is the voting weight pre-specified by the user, \(\lambda\) is the order of FLR, and
\(v_l (1 \leq l \leq \lambda)\) are the mid-points of the corresponding intervals of the \(\lambda\) latest past fuzzy sets.

The mean square error (MSE) measure used to evaluate the forecasting models is defined as follows

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} (FD_i - RD_i)^2,
\]

where \(n\) is the number of forecasted data, \(FD_i\) and \(RD_i\) are the forecasted data and the historical training datum at the time \(i\), respectively. The smaller MSE value indicates the better solution. Besides, the root mean square error (RMSE) and the mean absolute percentage error (MAPE) are also used to evaluate the forecasting model and are defined as the formulas (13) and (14), respectively:

\[
RMSE = \sqrt{MSE}.
\]

\[
MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{FD_i - RD_i}{FD_i} \right| \times 100.
\]

4. EXPERIMENTAL RESULTS AND DISCUSSION

This section presents the experimental results of our proposed forecasting model on three datasets of the EUA, the CAB, and the SGT, as well as compares its experimental results with the ones of state-of-the-art forecasting models. The MSE value is used to evaluate the accuracy of the forecasting models.

In the first step of our proposed model, PSO-SA is applied to optimize the interval lengths of the UD by minimizing the MSE value. The diversity of the population is important. Therefore, in our experiments, the number of particles is 30, the number of cycles is 100, the Inertia coefficient \(\omega\) is 0.4, the self-cognitive factor \(c_1\) and the social cognitive factor \(c_2\) are 0.2, the cooling rate \(\alpha\) is 0.995, the initial temperature \(T_0\) is 120.

All experiments are implemented by C# and performed using an Intel Core i5-8250U 1.6GHz CPU with 8 GB of memory and running Microsoft Windows 10 64-bit. There are three runs for each experiment, so we get three MSE values, and the smallest one is chosen as forecasting performance.

4.1. Experimental results on the “enrolments of the University of Alabama”

The UD of the historical data of the EUA observed from 1971 to 1992 is defined so that it should cover all data that may occur in the future. Therefore, \(D_{min}\) is set to 13,000 and \(D_{max}\) is set to 20,000 leading to \(U = [13,000, 20,000]\). The number of intervals is 16, as in [12, 30].

First, to show the impact of the defuzzification technique on the accuracy of FTS-FMs, the experiments of the proposed FTS-FM are implemented and executed with different defuzzification techniques. Then, the experimental results are compared to show the best
defuzzification technique. Last, to show the efficiency of the application of PSO-SA in improving forecasting results, the experimental results of the proposed FTS-FM are compared with the ones of the existing forecasting models.

The proposed FTS-FM with the application of defuzzification techniques (1), (2), (7), (8), (9), and (10) are denoted by M1, M2, M7, M8, M9, and M10, respectively. Each experiment of FTS-FM is executed 15 times, so we receive 15 $MSE$ values. Then, the best one is selected as an experimental result. The experimental results of M1, M2, M7, M8, M9, and M10 are shown in Table 4. It is easy to see that the $MSE$ value of M10 is the smallest, indicating the best value. Therefore, when comparing by the $MSE$ values, we have the ranking order: M10 is the best FTS-FM, M9 is the second, M7 is the third, M8 is the fourth, M1 is the fifth, and M2 is the worst. We do further comparisons with the $MAPE$ values, and the ranking order is changed as M9 is the best, M7 is the second, M10 is the third, M8 is the fourth, M1 is the fifth, and M2 is the worst.

To show the efficiency of the application of PSO-SA, our proposed FTS-FMs with the application of different defuzzification techniques are compared with the FTS-FM proposed by Chen&Zou in [27], Uslu in [32], and linguistic time series (LTS) proposed by Phong in [33]. The comparison results are shown in Table 4 and visualized in Figure 1. In [27], the defuzzification technique (9) proposed by Chen&Zou is the same as M9. It is seen in Table 4 that the $MSE$ value of M9 is 6174, decreased by 73.96%, compared to the value of 23710 in Chen&Zou’s model. In addition, the $MAPE$ value of M9 is 0.25% smaller than that of Chen&Zou, which is 0.73%. The difference between M9 and Chen&Zou’s model is only the optimization algorithm. In [33], Phong et al. applied PSO to optimize the fuzziness parameter values of the LTS forecasting model. The $MSE$ of it is better than those of Chen&Zou and Uslu, but worse than ours. We receive the same result when comparing by the $MAPE$ values. Therefore, we can state that PSO-SA is the main factor that helps to improve the forecasting performance of the proposed forecasting models.

4.2. Experimental results on “killed in car road accidents in Belgium”

The UD of the historical data of CAB observed from 1974 to 2004 is defined as follows: $D_{min}$ is set to 900 and $D_{max}$ is set to 1700, so we have $U = [900, 1700]$, the number of intervals is 17, as in [12, 30]. M9 and M10 are applied to the forecasting problem of the CAB, and their forecasted results are compared with the existing ones. The forecasted values of M9 and M10 compared to those of Uslu [32] and Chen&Zou [27] are shown in Table 5. The $MSE$ value of M9 and M10 are 880 and 861, respectively, decreased, in turn, 14.06% and 15.92%, compared to the one of Chen&Zou. Similarly, the $MSE$ values of M9 and M10 decreased, in turn, 49.16% and 50.26%, compared to the one of Uslu. Besides, when comparing by the $MAPE$ values, the one of M9 is 1.39% smaller than those of M10, Uslu, and Chen&Zou, which are 1.49%, 2.29%, and 1.77%, respectively. These comparison results show that both our proposed FTS-FMs, M9 and M10, are better than those of Uslu and Chen&Zou. Recall that the difference between M9 and the model of Chen&Zou is just the optimization algorithm.
Table 4. The forecasted values of the Enrolments of Alabama of our proposed FTS-FM with different defuzzification formulas compared with the existing ones.

| Year | Actual data | M1  | M2  | M7  | M8  | M9  | M10 | Uslu | Chen&Zou | LTS |
|------|-------------|-----|-----|-----|-----|-----|-----|------|---------|-----|
| 1971 | 13055       | 13066| 13075| 13084| 13094| 13104| 13114| 13124| 13134| 13144|
| 1972 | 13563       | 13574| 13585| 13595| 13605| 13615| 13625| 13635| 13645| 13655|
| 1973 | 13563       | 13574| 13585| 13595| 13605| 13615| 13625| 13635| 13645| 13655|
| 1974 | 14696       | 14611| 14673| 14697| 14698| 14699| 14699| 14699| 14699| 14699|
| 1975 | 15460       | 15389| 15468| 15462| 15462| 15462| 15462| 15462| 15462| 15462|
| 1976 | 15311       | 15365| 15277| 15366| 15366| 15366| 15366| 15366| 15366| 15366|
| 1977 | 15903       | 15863| 15863| 15863| 15863| 15863| 15863| 15863| 15863| 15863|
| 1978 | 15864       | 15864| 15864| 15864| 15864| 15864| 15864| 15864| 15864| 15864|
| 1979 | 16807       | 16881| 16831| 16834| 16827| 16831| 16827| 16831| 16827| 16831|
| 1980 | 16019       | 16010| 17071| 17039| 16934| 16919| 16920| 17075| 17009| 16949|
| 1981 | 16388       | 16404| 16417| 16412| 16384| 16388| 16384| 16384| 16384| 16384|
| 1982 | 15433       | 15489| 15408| 15402| 15423| 15433| 15423| 15423| 15423| 15423|
| 1983 | 15497       | 15635| 15277| 15356| 15385| 15353| 15347| 15441| 15212| 15569|
| 1984 | 15145       | 15365| 15277| 15356| 15385| 15353| 15347| 15441| 15212| 15569|
| 1985 | 15163       | 15210| 15281| 15168| 15193| 15163| 15193| 15163| 15193| 15163|
| 1986 | 15984       | 15973| 15801| 16014| 15985| 15984| 15982| 15984| 15982| 15984|
| 1987 | 16859       | 16781| 16833| 16834| 16827| 16833| 16829| 16835| 16833| 16833|
| 1988 | 17150       | 17206| 17297| 17142| 17151| 17150| 17150| 17150| 17150| 17150|
| 1989 | 17970       | 19060| 19012| 18994| 18989| 18970| 18960| 18880| 18937| 18911|
| 1990 | 19328       | 19364| 19315| 19343| 19339| 19343| 19343| 19343| 19343| 19343|
| 1991 | 19337       | 19364| 19315| 19343| 19339| 19343| 19343| 19343| 19343| 19343|
| 1992 | 18876       | 19060| 19012| 18898| 18914| 18898| 18914| 18914| 18914| 18914|

| MSE  | RMSE | MAPE |
|------|------|------|
| 110.86 | 83.93 | 0.31% |

Figure 1. The MSE values of the proposed FTS-FMs compared with those of Uslu, Chen&Zou, and LTS.
Table 5. The forecasted values of CAB of our proposed FTS-FM are compared with those of the existing FTS-FMs.

| Year | Actual data | Uslu | Chen&Zou | M9  | M10 |
|------|-------------|------|----------|-----|-----|
| 1974 | 1574        |      |          |     |     |
| 1975 | 1460        | 1506 | 1451     | 1466| 1465|
| 1976 | 1536        | 1483 | 1490     | 1468| 1470|
| 1977 | 1507        | 1598 | 1622     | 1586| 1583|
| 1978 | 1644        | 1584 | 1575     | 1593| 1592|
| 1979 | 1572        | 1584 | 1593     | 1593| 1592|
| 1980 | 1616        | 1506 | 1585     | 1616| 1614|
| 1981 | 1564        | 1584 | 1582     | 1593| 1592|
| 1982 | 1464        | 1506 | 1513     | 1464| 1465|
| 1983 | 1479        | 1453 | 1494     | 1468| 1470|
| 1984 | 1369        | 1375 | 1393     | 1369| 1371|
| 1985 | 1308        | 1383 | 1336     | 1370| 1312|
| 1986 | 1456        | 1454 | 1419     | 1436| 1435|
| 1987 | 1390        | 1453 | 1485     | 1468| 1470|
| 1988 | 1432        | 1383 | 1384     | 1370| 1433|
| 1989 | 1488        | 1509 | 1459     | 1468| 1470|
| 1990 | 1574        | 1598 | 1585     | 1586| 1587|
| 1991 | 1471        | 1506 | 1451     | 1466| 1465|
| 1992 | 1380        | 1375 | 1369     | 1369| 1371|
| 1993 | 1346        | 1383 | 1361     | 1370| 1312|
| 1994 | 1415        | 1383 | 1437     | 1436| 1435|
| 1995 | 1298        | 1231 | 1217     | 1228| 1229|
| 1996 | 1122        | 1135 | 1152     | 1148| 1147|
| 1997 | 1150        | 1180 | 1172     | 1150| 1093|
| 1998 | 1224        | 1245 | 1211     | 1239| 1238|
| 1999 | 1173        | 1135 | 1147     | 1148| 1147|
| 2000 | 1253        | 1245 | 1245     | 1239| 1238|
| 2001 | 1288        | 1284 | 1280     | 1288| 1288|
| 2002 | 1145        | 1143 | 1148     | 1145| 1143|
| 2003 | 1035        | 970  | 1028     | 1035| 1093|
| 2004 | 953         | 970  | 953      | 953 | 952 |

| MSE  | 1731 | 1024 | 880 | 861 |
|------|------|------|-----|-----|
| RMSE | 41.61| 32.0 | 29.66| 29.34|
| MAPE | 2.29%| 1.77%| 1.39%| 1.49%|

4.3. Experimental results on the “spot gold in Turkey”

To show the efficiency of our proposed FTS-FM in a wide variety of forecasting problems, it is applied to the forecasting problem of the SGT with the historical data observed from December 7th to November 10th. The minimum and maximum values of the historical data are 30,503 and 62,450, respectively. Therefore, the UD is determined as $U = [30000, 63000]$. The number of intervals is 16, as in [27, 32]. The forecasting results of M9 and M10 compared with the ones of Uslu and Chen&Zou are shown in Table 6. By a simple calculation, we see that the $MSE$ values of M9 and M10 are 848.51 and 840.66, respectively decreased, in
Table 6. The forecasted values of the SGT of our proposed FTS-FM are compared with those of the existing FTS-FMs.

| Date  | Actual spot gold | Uslu   | Chen&Zou | M9       | M10       |
|-------|------------------|--------|----------|----------|-----------|
| 7-Dec | 30503            | 32740.18 | 32341.38 | 34166.50 | 33271.26  |
| 8-Jan | 33132            | 34882.78 | 34479.36 | 34166.50 | 35136.85  |
| 8-Feb | 35201            | 37409.66 | 36055.47 | 38529.00 | 38578.87  |
| 8-Apr | 38300            | 39894.23 | 38203.34 | 37706.33 | 37717.49  |
| 8-May | 36142            | 37023.88 | 37406.67 | 37706.33 | 37717.49  |
| 8-Jun | 35837            | 37049.66 | 36749.36 | 36455.50 | 36447.30  |
| 8-Jul | 35704            | 37049.66 | 36522.85 | 36455.50 | 36447.30  |
| 8-Aug | 32955            | 32740.18 | 31805.51 | 32955.00 | 32865.10  |
| 8-Sep | 33277            | 34882.78 | 34335.42 | 34166.50 | 33271.26  |
| 8-Oct | 38295            | 37409.66 | 38120.71 | 38529.00 | 38578.87  |
| 8-Nov | 38677            | 37409.66 | 37402.31 | 37706.33 | 37717.49  |
| 8-Dec | 38295            | 37409.66 | 38120.71 | 38529.00 | 38578.87  |
| 9-Jan | 41985            | 43666.21 | 44515.67 | 43985.00 | 43836.54  |
| 9-Feb | 49931            | 49662.40 | 49800.77 | 49931.00 | 48984.80  |
| 9-Mar | 50823            | 51971.99 | 50622.66 | 50823.00 | 52317.31  |
| 9-Apr | 46167            | 45938.07 | 45869.80 | 46167.00 | 46186.90  |
| 9-May | 46716            | 46435.40 | 46548.24 | 47384.50 | 46239.92  |
| 9-Jun | 47337            | 46435.40 | 47067.02 | 47384.50 | 46239.92  |
| 9-Jul | 46088            | 46435.40 | 47653.83 | 47384.50 | 46239.92  |
| 9-Aug | 45839            | 46435.40 | 46473.59 | 47384.50 | 46239.92  |
| 9-Sep | 48053            | 46435.40 | 46238.30 | 47384.50 | 48984.80  |
| 9-Oct | 49592            | 49662.40 | 48330.41 | 49592.00 | 49599.79  |
| 9-Nov | 53693            | 51971.99 | 54338.06 | 53693.00 | 52317.31  |
| 9-Dec | 54553            | 54188.41 | 54509.96 | 54867.00 | 54362.80  |
| 10-Jan | 53922           | 54188.41 | 53663.01 | 54867.00 | 54362.80  |
| 10-Feb | 55031           | 54188.41 | 54183.79 | 54867.00 | 54362.80  |
| 10-Mar | 55031           | 54188.41 | 54471.07 | 54867.00 | 54362.80  |
| 10-Apr | 55181           | 54188.41 | 55887.68 | 54867.00 | 54362.80  |
| 10-May | 60300           | 60069.32 | 60030.78 | 61200.00 | 60532.56  |
| 10-Jun | 62100           | 60069.32 | 59888.46 | 61200.00 | 60532.56  |
| 10-Jul | 60500           | 59849.50 | 61610.89 | 61475.00 | 60532.56  |
| 10-Aug | 59200           | 60069.32 | 60079.84 | 61200.00 | 60532.56  |
| 10-Sep | 61250           | 60069.32 | 61520.74 | 61200.00 | 60532.56  |
| 10-Oct | 62450           | 62437.15 | 60797.52 | 61475.00 | 60532.56  |
| 10-Nov | 61600           | 59849.50 | 61945.80 | 61475.00 | 60532.56  |

| **MSE**  | **1030692** | **805291** | **719964** | **706706** |
| **RMSE** | **1.015.23** | **897.38** | **848.51** | **840.66** |
| **MAPE** | **1.80%**   | **1.55%**   | **1.33%**   | **1.29%**   |

turn, by 10.6% and 12.24%, compared to the one of Chen&Zou. Similarly, the MSE values of M9 and M10 decreased, in turn, by 30.15% and 31.43%, compared to the one of Uslu. It is also seen that M10 is better than M9. When comparing by the MAPE values, we can see that the one of M10 is the best, the one of M9 is the second, and those of them are better
than those of Uslu and Chen&Zou. These results state that our proposed FTS-FMs have the best forecasting performance compared with Uslu and Chen&Zou, and M10 is slightly better than M9.

4.4. The analysis of running time

In this subsection, the running time of the proposed forecasting models with the application of PSO-SA are compared with that of PSO. Both PSO-SA and PSO are executed with the number of cycles and particles are 100 and 30, respectively, and by a single-threaded program. The comparison results of training time between PSO-SA and PSO on the datasets of EUA, CAB, and SGT are shown in Table 7. We can see that PSO-SA takes longer to run on all datasets because it tries to do exploitation with the help of SA to get a better solution, but in turn, we get better forecasted results.

| Datasets | EUA | CAB | SGT |
|----------|-----|-----|-----|
| Algorithms | PSO | PSO-SA | PSO | PSO-SA | PSO | PSO-SA |
| Training time (second) | 5 | 7 | 4 | 12 | 4 | 8 |

4.5. The analysis of result variation

Table 8. The statistic of 15 execution times of the proposed forecasting models with the application of formula (10) on the datasets of EUA, CAB, and SGT

| Run | Dataset | EUA | CAB | SGT |
|-----|--------|-----|-----|-----|
| 1   | 20644  | 1469| 778213.8 |
| 2   | 17245  | 964 | 914344.4 |
| 3   | 16978  | 861 | 706706.0 |
| 4   | 5359   | 1093| 1057949.0 |
| 5   | 5406   | 1129| 1006000.0 |
| 6   | 16735  | 1211| 908568.4 |
| 7   | 17015  | 1354| 869023.8 |
| 8   | 13286  | 1308| 996662.1 |
| 9   | 10430  | 1184| 789337.4 |
| 10  | 8406   | 1221| 799327.3 |
| 11  | 13454  | 1040| 1083637.0 |
| 12  | 10438  | 860 | 865035.8 |
| 13  | 8447   | 1147| 934467.0 |
| 14  | 16047  | 1070| 745990.5 |
| 15  | 13738  | 969 | 778243.3 |
| Average | 12908.5| 1125.3| 876200.4 |
| Variance | 162434021.7| 1158805.6| 7.42253E+11 |
| Standard deviation | 12745.0| 1076.5| 861540.8 |
To illustrate the variation of the experimental results, we execute the proposed forecasting models with the application of formula (10) on the datasets of EUA, CAB, and SGT 15 times, and the $MSE$ values of the executions are shown in Table 8. It is easy to calculate that the difference between the smallest value and the average one of the datasets EUA, CAB, and SGT are 7549.53, 265.33, and 169494.39, respectively. The standard deviations of all datasets are less than the average values.

5. CONCLUSIONS

FTS-FM plays an essential role in the forecasting research field because of its numerous practical applications. Three main factors which have a strong effect on the forecasting accuracy of FTS-FMs are partitioning historical data, establishing fuzzy logical relationship groups, and defuzzification techniques. Among those factors, the study of the methods of optimizing the interval length of the UD to improve forecasting accuracy has attracted many researchers. This paper presents our proposed hybrid FTS-FMs combined with PSO-SA to optimize the length of intervals of the universe of discourse to improve forecasting accuracy. The characteristic of PSO-SA is that it makes use of the local search performed by SA and the global search performed by PSO to improve the search result. The experimental results on the datasets of the “enrolments of the University of Alabama,” “killed in car road accidents in Belgium,” and the “spot gold in Turkey” have shown that our proposed FTS-FM with the help of PSO-SA outperforms its counterparts. Furthermore, the experimental results also evaluate the influence and efficiency of different defuzzification techniques to show the best one. In specifically, in the experiment of the EUA dataset, when comparing by the $MSE$ values, the formula (10) is the best and when comparing by the $MAPE$ values, the formula (9) is the best. We get the same comparison result with the EUA dataset. However, the formula (10) is the best on SGT datasets when compared by both the $MSE$ and the $MAPE$ values.

ACKNOWLEDGMENT

This research is funded by University of Transport and Communications, Ha Noi, Viet Nam under grant number T2022-CN-002.

REFERENCES

[1] G.E.P. Box, G. Jenkins, and G.C. Reinsel, Time Series Analysis, Forecasting and Control, John Wiley & Sons, Inc., Hoboken, New Jersey, 2008. https://doi.org/10.1002/9781118619193
[2] Q. Song, B.S. Chissom, “Fuzzy Time Series, and its Model, Fuzzy set and systems,” vol. 54, no. 3, pp. 269-277, 1993. https://doi.org/10.1016/0165-0114(93)90372-O
[3] Q. Song, and B.S. Chissom, “Forecasting enrollments with fuzzy time series — Part I,” Fuzzy Set and Systems, vol. 54, no. 1, pp. 1-9, 1993. https://doi.org/10.1016/0165-0114(93)90355-L
[4] Q. Song, and B.S. Chissom, “Forecasting enrollments with fuzzy time series — Part II,” Fuzzy Set and Systems, vol.62, no. 1, pp. 1-8, 1994. https://doi.org/10.1016/0165-0114(94)90067-1
[5] S. M. Chen, Forecasting Enrollments based on Fuzzy Time Series, Fuzzy Sets and Systems, vol. 81, no. 3, pp. 311-319, 1996. https://doi.org/10.1016/0165-0114(95)00220-0
[6] H.K. Yu, “Weighted fuzzy time series models for TAIEX forecasting,” *Physica A: Statistical Mechanics and its Applications*, vol. 349, no. (3–4), pp. 609–624, 2005. https://doi.org/10.1016/j.physa.2004.11.006

[7] K. Huarng, “Effective lengths of intervals to improve forecasting in fuzzy time series,” *Fuzzy Sets and Systems*, vol. 123, no. 3, pp. 387–394, 2001. https://doi.org/10.1016/S0165-0114(00)00057-9

[8] S.M. Chen, and N.Y. Chung, “Forecasting enrollments using high-order fuzzy time series and genetic algorithms,” *International Journal of Intelligent Systems*, vol. 21, no. 5, pp. 485–501, 2006. https://doi.org/10.1002/int.20145

[9] E. Bas, V.R. Uslu, U. Yolcu, and E. Egrioglu, “A modified genetic algorithm for forecasting fuzzy time series,” *Applied Intelligence*, vol. 41, no. 2, pp. 453–463, 2014. https://doi.org/10.1007/s10489-014-0529-x

[10] L.W. Lee, L.H. Wang, and S.M. Chen, “Temperature prediction and TAIFEX forecasting based on fuzzy logical relationships and genetic algorithms,” *Expert Systems with Applications*, vol. 33, no. 5, pp. 539–550, 2007. https://doi.org/10.1016/j.eswa.2006.05.015

[11] L.W. Lee, L.H. Wang, and S.M. Chen, “Temperature prediction and TAIFEX forecasting based on high-order fuzzy logical relationships and genetic simulated annealing techniques,” *Expert Systems with Applications*, vol. 34, no. 1, pp. 328–336, 2008. https://doi.org/10.1016/j.eswa.2006.09.007

[12] N.Y. Wang and S.M. Chen, “Temperature prediction and TAIFEX forecasting based on automatic clustering techniques and two-factors high-order fuzzy time series,” *Expert Systems with Applications*, vol. 36, no. 2, pp. 2143–2154, 2009. https://doi.org/10.1016/j.eswa.2007.12.013

[13] I.H. Kuo, S.J. Horng, T.W. Kao, T.L. Lin, C.L. Lee, and Y. Pan, “An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization,” *Expert Systems with Applications*, vol. 36, no. 3, pp. 6108–6117, 2009. https://doi.org/10.1016/j.eswa.2008.07.043

[14] I.H. Kuo, S.J. Horng, Y.H. Chen, R.S. Run, T.W. Kao, R.J. Chen, J.L. Lai, and T.L. Lin, “Forecasting TAIFEX based on fuzzy time series and particle swarm optimization,” *Expert Systems with Applications*, vol. 37, no. 2, pp. 1494–1502, 2010. https://doi.org/10.1016/j.eswa.2009.06.102

[15] Y.L. Huang, S.J. Horng, M. He, P. Fan, T.W. Kao, M.K. Khan, J.L. Lai, and I-H. Kuo, “A hybrid forecasting model for enrollments based on aggregated fuzzy time series and particle swarm optimization,” *Expert Systems with Applications*, vol. 38, no. 7, pp. 8014–8023, 2011. https://doi.org/10.1016/j.eswa.2010.12.127

[16] W. Lu, J. Yang, X. Liu, and W. Pedrycz, “The modeling of time series based on fuzzy information granules,” *Expert Systems with Applications*, vol. 41, no. 8, pp. 3799–3808, 2014. https://doi.org/10.1016/j.eswa.2013.12.005

[17] L. Wang, X. Liu, W. Pedrycz, and Y. Shao, “Determination of temporal information granules to improve forecasting in fuzzy time series,” *Expert Systems with Applications*, vol. 41, no. 6, pp. 3134–3142, 2014. https://doi.org/10.1016/j.eswa.2013.10.046
[18] Q. Cai, D. Zhang, W. Zheng, and S.C.H. Leung, “A new fuzzy time series forecasting model combined with ant colony optimization and auto-regression,” *Knowledge-Based Systems*, vol. 74, pp. 61-68, 2015. https://doi.org/10.1016/j.knosys.2014.11.003

[19] C.H. Cheng, G.W. Cheng, and J.W. Wang, “Multi-attribute fuzzy time series method based on fuzzy clustering,” *Expert Systems with Applications*, vol. 34, no. 2, pp. 1235–1242, 2008. https://doi.org/10.1016/j.eswa.2006.12.013

[20] N.V. Tinh and N.C. Dieu, “A new hybrid fuzzy time series forecasting model based on combining fuzzy c-means clustering and particle swarm optimization,” *Journal of Computer Science and Cybernetics*, vol. 35, no. 3, pp. 267-292, 2019. https://doi.org/10.15625/1813-9663/35/3/13496

[21] M. Bose and K. Mali, “A novel data partitioning and rule selection technique for modeling high-order fuzzy time series,” *Applied Soft Computing*, vol. 63, pp. 87–96, 2018. https://doi.org/10.1016/j.asoc.2017.11.011

[22] M. Bose and K. Mali, “Designing fuzzy time series forecasting models: A survey,” *International Journal of Approximate Reasoning*, vol. 111, pp. 78–99, 2019. https://doi.org/10.1016/j.ijar.2019.05.002

[23] S.M. Chen, “Forecasting enrollments based on high-order fuzzy time series,” *Cybernetics and Systems An International Journal*, vol. 33, no. 1, pp. 1-16, 2002. https://doi.org/10.1080/019697202753306479

[24] H.K. Yu, “A refined fuzzy time-series model for forecasting,” *Physica A: Statistical Mechanics and its Applications*, vol. 346, no. 3–4, pp. 657–681, 2005. https://doi.org/10.1016/j.physa.2004.07.024

[25] S.H. Cheng, S.M. Chen, and W.S. Jian, “Fuzzy time series forecasting based on fuzzy logical relationships and similarity measures,” *Information Sciences*, vol. 327, pp. 272-287, 2016. https://doi.org/10.1016/j.ins.2015.08.024

[26] P.D. Phong, “A time series forecasting model, based on linguistic forecasting rules,” *Journal of Computer Science and Cybernetics*, vol. 37, no. 1, pp. 23-42, 2021.

[27] S. M. Chen, X.Y. Zou, and G.C. Gunawan, “Fuzzy time series forecasting based on proportions of intervals and particle swarm optimization techniques,” *Information Sciences*, vol. 500, pp. 127–139, 2019. https://doi.org/10.1016/j.ins.2019.05.047

[28] R. Eberhart and J. Kennedy, “A new optimizer using particle swarm theory,” in *MHS’95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, 1995, pp. 39-43. Doi: 10.1109/MHS.1995.494215.

[29] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Proceedings of ICNN’95 - International Conference on Neural Networks*, 1995, pp. 1942-1948, vol.4. Doi: 10.1109/ICNN.1995.488968.

[30] S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi, “Optimization by simulated annealing,” *Science*, vol. 220, no. 4598, pp. 671-680, 1983. Doi: 10.1126/science.220.4598.671.
[31] P.D. Phong, N.T. Thuy, T.X. Thanh, “A hybrid multi-objective PSO-SA algorithm for the fuzzy rule based classifier design problem with the order based semantics of linguistic terms,” *VNU Journal of Science: Computer Science and Communication Engineering*, vol. 30, no. 4, pp. 44-56, 2014. https://jcsce.vnu.edu.vn/index.php/jcsce/article/view/35

[32] V.R. Uslu, E. Bas, U. Yolcu, and E. Egrioglu, “A fuzzy time series approach based on weights determined by the number of recurrences of fuzzy relations,” *Swarm and Evolutionary Computation*, vol. 15, pp. 19–26, 2014. https://doi.org/10.1016/j.swevo.2013.10.004

[33] P.D. Phong, N.D. Hieu and M.V. Linh, “A hybrid linguistic time series forecasting model combined with particle swarm optimization,” in *2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, 2022, pp. 1-6. https://doi.org/10.1109/ICECET55527.2022.9873100

*Received on August 10, 2022
Accepted on October 08, 2022*