AN IMPROVED FORECASTING MODEL BASED ON THE WEIGHTED FUZZY RELATIONSHIP MATRIX COMBINED WITH A PSO ADAPTATION FOR ENROLLMENTS

YAO-LIN HUANG¹, SHI-JINN HORNG¹, TZONG-WANG KAO², RAY-SHINE RUN³
JUI-LIN LAI³, RONG-JIAN CHEN³, I-HONG KUO⁴
AND MUHAMMAD KHURRAM KHAN⁵

¹Department of Computer Science and Information Engineering
National Taiwan University of Science and Technology
No. 43, Sec. 4, Keelung Rd., Taipei 106, Taiwan
ylhave@gmail.com; horngsj@yahoo.com.tw

²Department of Electronic Engineering
Technology and Science Institute of Northern Taiwan
No. 2, Xueyuan Rd., Peitou, Taipei 112, Taiwan
tkao@tsint.edu.tw

³Department of Electronic Engineering
National United University
No. 1, Lien-Kung, Miao-Li 360, Taiwan
run5116@ms16.hinet.net; {jillai; rjchen}@nuu.edu.tw

⁴Department of Information Management
St. Mary’s Medicine Nursing and Management College
No. 100, Sec. 2, San-shing Rd., Da-yin Village, Yilan County 266, Taiwan
yihonguo@smc.edu.tw

⁵Center of Excellence in Information Assurance
King Saud University Kingdom of Saudi Arabia
mkhurram@ksu.edu.sa

Received February 2010; revised August 2010

ABSTRACT. Most fuzzy forecasting approaches are based on modeling fuzzy relations according to the past data. In this paper, an improved forecasting model which combines weighted fuzzy relationship matrices and particle swarm optimization is presented for enrollments. First, the weighted fuzzy relationship matrices are more effective to capture fuzzy relations on time series data than fuzzy logical relationship rules. Second, the particle swarm optimization for the optimized lengths of intervals is developed to adjust interval lengths by searching the space of the universe of discourse. To verify the effectiveness of the proposed model, the empirical data for the enrollments of the University of Alabama are illustrated, and the experimental results show that the proposed model outperforms those of previous forecasting models for both the training and testing phases with various orders and different interval lengths. These results are very encouraging for future work on the development of fuzzy time series and particle swarm optimization in forecasting real-world applications.

Keywords: Fuzzy time series, Particle swarm optimization, Fuzzy logical relationship, Fuzzy relationship matrix

1. Introduction. In order to advance the decision-making process concerning future requirements, many researchers have focused on real-world problems to deal with various time series data, such as weather news [1,2], crop productions [3,4], stock markets [2,5-7] and academic enrollments [8-14]. However, conventional forecasting methods only
refer to real numbers and fail to solve forecasting problems in which the historical data are available in linguistic values. Fuzzy set theory was first presented by Zadeh [15] to deal with linguistic values. In 1993, the concepts of fuzzy sets [16] were successfully applied to time series forecasting by Song and Chissom [17]. They introduced fuzzy time-invariant and time-variant models [8,9] to forecast the enrollments of the University of Alabama. However, their methods performed max-min composition operations to handle the fuzzy process, which required a lot of computation time when a fuzzy relationship matrix (FRM) is large. In 1996, Chen [11] proposed an efficient model which used the first-order fuzzy logical relationships (FLR) to simplify the computational complexity of the forecasting process. Chen’s model employed simple arithmetic calculations instead of max-min composition operations for better forecasting accuracy. Since then, the fuzzy time series methods have received increasing attention in many forecasting applications.

To achieve better forecasting accuracy, Huarng [18] presented effective approaches which can properly adjust the lengths of intervals. Huarng [19] also proposed heuristic models to improve the forecasting accuracy based on fuzzy time series. Hwang et al. [20] presented a time-variant forecasting model in which fuzzy relations were defined by FRMs. Chen [21] presented a new forecasting model based on the high-order FLRs to forecast the enrollments of the University of Alabama. Yu [22] proposed a new model which refined the lengths of intervals during the formulation of FLRs, thus capturing the fuzzy relations more effectively. The stock index and enrollments were both used by Yu as targets for empirical analysis. Based on genetic algorithms, Chen et al. [12,13] presented the first-order and high-order fuzzy time series models to deal with forecasting problems. Singh [3,4] presented simplified and robust computational methods to create forecasting rules based on single and various parameters, respectively. Li [14] proposed a novel deterministic forecasting approach which can effectively partition interval lengths. Singh [23] proposed a 3rd-order fuzzy time series model which used a time-variant difference parameter on the current state to forecast the next state. Chen et al. [24] presented a high-order fuzzy time series method using multi-period adaptation and adaptive expectation models. Recently, particle swarm optimization (PSO) has been successfully applied in many applications. Lai et al. [25] introduced a PSO algorithm to automatically determine the proper number of features to classify spam e-mails. Cui et al. [26] proposed a novel Lévy velocity threshold to improve the performance of PSO. Cui et al. [27] introduced chaotic sequences to expanding PDPSO for solving high-dimensional problems. Based on Chen’s model [11], Kuo et al. [29] introduced a new hybrid forecasting model which combined fuzzy time series with PSO algorithm to find the proper length of each interval.

The literature review above shows that the lengths of intervals and fuzzy relations are two critical factors for forecasting accuracy. Therefore, an improved forecasting model (termed FMPSO) which combined the weighted FRMs and PSO algorithm was proposed to reconcile these factors. First, the FRMs with a decreasing weight (DW) scheme were more effective to capture fuzzy relations on time series data than FLR rules. Second, the PSO algorithm for the optimized lengths of intervals was developed to adjust the interval lengths by searching the space of the universe of discourse.

For verification of model’s effectiveness, the empirical data of enrollments from the University of Alabama are used in the training and testing phases. The mean square error (MSE) is employed to estimate the forecasting accuracy, and then applied to PSO algorithm as fitness value. The experimental results show that the proposed model outperforms those of previous forecasting models for both the training and testing phases with various orders and different interval lengths.
The rest of this paper is organized as follows. Section 2 briefly reviews the concepts of fuzzy time series and PSO algorithm. In Section 3, an improved forecasting model based on the weighted FRMs and PSO algorithm is detailed. Section 4 illustrates the proposed model by the empirical example of enrollments. Section 5 evaluates the forecasting performance of the proposed method as compared with the existing methods on the enrollments of the University of Alabama. Finally, some conclusions are discussed in Section 6.

2. Preliminaries. In view of making our exposition self-contained, the basic concepts of fuzzy time series and PSO algorithm are summarized and reproduced briefly.

2.1. Fuzzy time series. Fuzzy set theory was first developed by Zadeh [15,30–32] to deal with uncertainty using linguistic values. Song and Chissom [8,9,17] successfully modeled a fuzzy forecast by adjusting time series data to the fuzzy sets. Instead of complicated max-min composition operations, Chen [11] improved the fuzzy forecasting method by using simple arithmetic operations. Let $U$ be the universe of discourse, where $U = \{u_1, u_2, \cdots, u_n\}$. A fuzzy set $A$ of $U$ is defined as follows

$$A = \mu_A(u_1)/u_1 + \mu_A(u_2)/u_2 + \cdots + \mu_A(u_n)/u_n$$

where $\mu_A$ is the membership function of the fuzzy set $A$, $\mu_A: U \rightarrow [0, 1]$, $\mu_A(u_i)$ is the degree of membership of $u_i$ in $A$, $\mu_A(u_i) \in [0, 1]$, and $1 \leq i \leq n$. Some definitions of the fuzzy time series by Song and Chissom [8,9,17] and Chen [12,21] follow.

Definition 2.1. Let $Y(t)$ $(t = \ldots, 0, 1, 2, 3, \ldots)$, a subset of real numbers, be the universe of discourse by which fuzzy sets $\mu_i(t)$ are defined. If $F(t)$ is a collection of $\mu_1(t), \mu_2(t), \cdots$, then $F(t)$ is a fuzzy time series defined as $Y(t)$.

Definition 2.2. When $F(t)$ and $F(t - 1)$ are fuzzy sets, if there exists a fuzzy logical relationship, $R(t - 1, t)$, such that $F(t) = F(t - 1) \circ R(t, t - 1)$, where the symbol \( \circ \) represents an operation, then $F(t)$ is said to be caused by $F(t - 1)$. If $F(t)$ is caused by $F(t - 1)$ only, the first-order fuzzy relationship is represented by $F(t - 1) \rightarrow F(t)$, where $F(t - 1)$ is the current state and $F(t)$ is the next state, respectively.

Definition 2.3. Let $F(t-n), F(t-n+1), \cdots$ and $F(t)$ be the fuzzy sets in a time series. If $F(t)$ is caused by $F(t-n), F(t-n+1), \cdots$ and $F(t-1)$, then the nth-order fuzzy logical relationship is represented as follows: $F(t-n), F(t-n+1), \cdots, F(t-1) \rightarrow F(t)$, where $F(t-n), F(t-n+1), \cdots$ and $F(t-1)$ refer to the current state and $F(t)$ refers to the next state.

Hwang et al. [20] presented a new forecasting model in which fuzzy relations were represented as the FRMs. In Hwang’s model, the fuzzy set $F(t)$ was defined by the variation of the enrollments between years $t$ and $t - 1$. In this paper, raw historical data are used to create the fuzzy sets. Therefore, the fuzzy set $F(t)$ is defined by the enrollment of year $t$. A window base denotes how many past years’ data will be used in the FRMs. Let $R(t)$ be the FRM of a window base $w$, $O^w(t)$ be the operation matrix formed by the fuzzy sets of $w$ past years, and $C(t)$ be the criterion matrix formed by the fuzzy set of the current year. In the training phase, the matrices $O^w(t)$ and $C(t)$ are defined as follows

$$O^w(t) = \begin{bmatrix} F(t-1) \\ F(t-2) \\ \vdots \\ F(t-w) \end{bmatrix}$$


and

\[ C(t) = F(t) \]  

(3)

Based on Formulas (2) and (3), \( R(t) \) can be calculated as follows

\[ R(t) = O^w(t) \otimes C(t) \]  

(4)

where the symbol ‘\( \otimes \)’ is the product composition operation. By performing the transpose and max composition operations, \( F(t) \) can be derived from \( R(t) \) as

\[ F(t) = (\max (R(t)^T))^T \]  

(5)

2.2. Particle swarm optimization. PSO was first introduced by Eberhart and Kennedy [33] in 1995. It belongs to a population-based evolutionary algorithm that can efficiently search a near optimal or optimal solution for optimization problems. Most population-based approaches are motivated by evolution as seen in nature. The development of PSO algorithm [25-29,34-36] was inspired by the social behavior of animals, such as fish schooling, birds flocking and the swarm theory. The PSO approach applies a cooperative particle swarm to find the best solution from all feasible solutions. Each particle is randomly initialized and then allowed to fly in the virtual searching space. At each step of optimization, each particle evaluates its own fitness and the fitness of its neighboring particles. Each particle can remember its own best solution, which results in the best fitness, as well as see the candidate solution for the best performing particle in its neighborhood.

A moving particle, indexed by \( id \), adjusts its candidate solution according to the following formulas

\[ V_{id} = \omega \times V_{id} + c_1 \times \text{Rand}(\cdot) \times (P_{id} - X_{id}) + c_2 \times \text{Rand}(\cdot) \times (P_{gd} - X_{id}) \]  

(6)

\[ X_{id} = X_{id} + V_{id} \]  

(7)

where \( \omega \) denotes the inertia weight factor; \( X_{id} \) is the current position of a particle; \( P_{id} \) is the position of a particle that experiences the best fitness value; \( P_{gd} \) is the position of all particles that experience a global best fitness value; \( c_1 \) and \( c_2 \) are acceleration values which represent the self-confidence coefficient and the social coefficient, respectively; \( \text{Rand}(\cdot) \) denotes a random function which can yield a random value in the range of 0 to 1; and \( V_{id} \) denotes the velocity of a particle. If \( V_{id} \) exceeds \( V_{\text{max}} \), then \( V_{id} \) is limited to \( V_{\text{max}} \), where \( V_{\text{max}} \) is a constant which determines the resolution of the searching regions between the present position and the target position. This study applied a standard PSO in which the value of \( \omega \) decreases linearly during the whole running generations, and \( c_1 \) and \( c_2 \) are two constants. A brief description of a standard PSO [36] is summarized in Figure 1.

\begin{verbatim}
1. initialize positions and velocities of all particles
2. while terminating condition (maximum iterations or minimum error criteria) is not reached do
3.   for each particle do
4.     evaluate the fitness
5.     update local best position and global best position by the fitness
6.     modify the velocity
7.     update the position
8.   end for
9. end while
\end{verbatim}

**Figure 1.** Standard PSO algorithm
3. Proposed Model on Time Series Data. Based on Hwang’s method [20], a new forecasting model which combined the weighted FRMs and PSO algorithm is introduced. In the proposed model, three key aspects have been applied to approach the lengths of intervals and fuzzy relations on time series data to increase forecasting accuracy. First, raw historical data are used instead of the variations of historical data in our forecasting model. Second, the FRMs are derived from the corresponding FLRs; and then the DW scheme assigns the largest weights to the latest past fuzzy set of a FRM for capturing efficient fuzzy relations. Third, the PSO algorithm is developed to adjust the interval lengths to obtain the optimal partition. A detailed explanation of the proposed model follows.

3.1. Fuzzy forecasting method by the weighted FRMs.

**Step 1.** Define the universe of discourse $U$. Let $Y(t)$ be the historical data at time $t$. The minimum and maximum data of $Y(t)$ are $Y_{\text{min}}$ and $Y_{\text{max}}$, respectively. The universe of discourse is defined as $U = [Y_{\text{min}} - \Delta_{\text{min}}, Y_{\text{max}} + \Delta_{\text{max}}]$, where $\Delta_{\text{min}}$ and $\Delta_{\text{max}}$ are proper positive integers to adjust the lower and upper bounds of $U$, respectively.

**Step 2.** Partition $U$ into intervals. Divide $U$ into $m$ continuous intervals $u_1, u_2, \ldots$ and $u_m$. For $U = [b_0, b_m]$, $m$ intervals are $u_1 = (b_0, b_1]$, $u_2 = (b_1, b_2]$, and $u_m = (b_{m-1}, b_m]$.

**Step 3.** Define the fuzzy sets. Each fuzzy set represents a linguistic value. The number of fuzzy sets is equal to the number of intervals. By Formula (1), all fuzzy sets are defined by the membership functions of the intervals with the form $A_i = \{\delta_j/u_j\}$ as follows

$$A_i = \sum_{j=1}^{m} \frac{\delta_j}{u_j} = \frac{\delta_1}{u_1} + \frac{\delta_2}{u_2} + \cdots + \frac{\delta_m}{u_m}$$

(8)

where $1 \leq i, j \leq m$, $0 \leq \delta_j \leq 1$, $m$ is the number of intervals, and $u_j$ is the $j$th number of intervals of $U$.

**Step 4.** Fuzzify all historical data. Find out the interval $u_j$ to which the historical data $Y(t)$ belongs, and then assign the fuzzy set $A_i$ which has the maximum degree of membership at $u_j$. The fuzzy set $F(t)$ for $Y(t)$ is expressed as

$$F(t) = A_i$$

(9)

**Step 5.** Create fuzzy logical relationships. By Definitions 2.2 and 2.3, all FLRs with various orders can be created based on the fuzzy time series in the training phase. Two consecutive fuzzy sets are used to create a first-order FLR. Let $F(t-1) = A_i$ and $F(t) = A_j$, then a first-order FLR for $F(t-1)$ and $F(t)$ is denoted by the following formula

$$A_i \rightarrow A_j$$

(10)

where “$A_i$” and “$A_j$” refer to the current state and the next state, respectively. Similarly, a $w$-th order FRM requires $w$ consecutive fuzzy sets in time series. Let $F(t-w) = A_{iw}$, $F(t-w+1) = A_{i(w-1)}, \cdots, F(t-2) = A_{i2}$, $F(t-1) = A_{i1}$ and $F(t) = A_j$, then a $w$-th order FRM is denoted as

$$A_{iw}, A_{i(w-1)}, \cdots, A_{i2}, A_{i1} \rightarrow A_j$$

(11)

where “$A_{iw}$, $A_{i(w-1)}, \cdots, A_{i2}, A_{i1}$” is the current state and “$A_j$” is the next state. In the testing phase, the symbol ‘#’ is used to represent the unknown next state.

**Step 6.** Establish the weighted FRMs. Choose an integer number $w$ ($w \geq 1$). The window base of a FRM is set to be $w$ for a $w$-th order FLR. The current state and next state of a FLR are converted to the operation matrix and the criterion matrix, respectively. The DW scheme assigns the largest weights to the latest past fuzzy set of a FRM. Let $O^w(t)$ be a operation matrix, $C(t)$ be a criterion matrix, $D^w(t)$ be a weighted operation
matrix, and \( R(t) \) be a FRM. The operation matrix is obtained by the fuzzy sets of the current state as follows

\[
O^w(t) = \begin{bmatrix}
F(t-1) \\
F(t-2) \\
\vdots \\
F(t\!-\!w)
\end{bmatrix} = \begin{bmatrix}
o_{11} & o_{12} & \cdots & o_{1m} \\
o_{21} & o_{22} & \cdots & o_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
o_{w1} & o_{w2} & \cdots & o_{wm}
\end{bmatrix}
\] (12)

where \( o_{ij} \) refers to the membership value of the fuzzy set \( F(t-i) \), \( 1 \leq i \leq w \) and \( 1 \leq j \leq m \). Based on the fuzzy sets of an operation matrix, the weighted operation matrix is calculated by the DW scheme which assigns the weights from \( w \) to 1 decreasingly. In the case of Formula (12), we assign \( w, w-1, \cdots \) and 1 to \( A_{11}, A_{12}, \cdots \) and \( A_{1w} \), respectively. Thus, the weighted operation matrix derived form \( O^w(t) \) is given as

\[
D^w(t) = \begin{bmatrix}
o_{11} \times w & o_{12} \times w & \cdots & o_{1m} \times w \\
o_{21} \times (w-1) & o_{22} \times (w-1) & \cdots & o_{2m} \times (w-1) \\
\vdots & \vdots & \ddots & \vdots \\
o_{w1} \times 1 & o_{w2} \times 1 & \cdots & o_{wm} \times 1
\end{bmatrix}
\] (13)

In the training phase, the next state of a FLR is known. Therefore, the criterion matrix is defined by the fuzzy set of the next state.

\[
C(t) = F(t) = \begin{bmatrix}
c_1 & c_2 & \cdots & c_m
\end{bmatrix}
\] (14)

where \( c_j \) refers to the membership value of the interval \( u_j \) in the fuzzy set \( F(t) \). In the testing phase, to deal with the unknown next state of a FLR, the criterion matrix is defined by the last fuzzy set of the current state. This leads

\[
C(t) = F(t-1) = \begin{bmatrix}
o_{11} & o_{12} & \cdots & o_{1m}
\end{bmatrix}
\] (15)

The weighted FRM is computed by the product composition operation of the weighted operation matrix and the criterion matrix. For example, the weighted FRM in the training phase is calculated as follows

\[
R(t) = D^w(t) \odot C(t)
\]

\[
= \begin{bmatrix}
c_1 \times o_{11} \times w & c_2 \times o_{12} \times w & \cdots & c_m \times o_{1m} \times w \\
c_1 \times o_{21} \times (w-1) & c_2 \times o_{22} \times (w-1) & \cdots & c_m \times o_{2m} \times (w-1) \\
\vdots & \vdots & \ddots & \vdots \\
c_1 \times o_{w1} \times 1 & c_2 \times o_{w2} \times 1 & \cdots & c_m \times o_{wm} \times 1
\end{bmatrix}
\] (16)

\[
= \begin{bmatrix}
r_{11} & r_{12} & \cdots & r_{1m} \\
r_{21} & r_{22} & \cdots & r_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
r_{w1} & r_{w2} & \cdots & r_{wm}
\end{bmatrix}
\]

**Step 7.** Calculate the forecasted outputs. Use the transpose and max composition operations in the weighted FRM as shown in Formula (16). The raw membership of the forecasted output \( F'(t) \) is obtained as follows

\[
F'(t) = \begin{bmatrix}
p'_1 & p'_2 & \cdots & p'_m
\end{bmatrix} = \left( \text{max} \left( R(t)^T \right) \right)^T = \begin{bmatrix}
\text{max} \left( r_{11}, r_{21}, \cdots, r_{w1} \right) \\
\text{max} \left( r_{12}, r_{22}, \cdots, r_{w2} \right) \\
\vdots \\
\text{max} \left( r_{1m}, r_{2m}, \cdots, r_{wm} \right)
\end{bmatrix}^T
\] (17)
Then, the raw membership of the forecasted output is normalized as follows

$$F(t) = \left[ p_1 \ p_2 \ \cdots \ p_m \right] = \left[ \frac{p'_1}{\sum_{i=1}^{m} p'_i} \ \frac{p'_2}{\sum_{i=1}^{m} p'_i} \ \cdots \ \frac{p'_m}{\sum_{i=1}^{m} p'_i} \right]$$  \hspace{1cm} (18)

Let $M$ denote a row matrix with the size of $1 \times m$. Each element of $M$ contains a midpoint as follows

$$M = \left[ m_1 \ m_2 \ \cdots \ m_m \right] = \left[ \frac{b_1 - b_0}{2} \ \frac{b_2 - b_1}{2} \ \cdots \ \frac{b_m - b_{m-1}}{2} \right]$$  \hspace{1cm} (19)

The forecasted value $fv$ is calculated by the product of the normalized membership matrix in Formula (18) and the midpoint matrix in Formula (19) as follows

$$fv = F(t) \cdot M^T = p_1 \times m_1 + p_2 \times m_2 + \cdots + p_m \times m_m$$  \hspace{1cm} (20)

**Step 8.** Compute forecasting accuracy by MSE value. The MSE value is employed as an evaluation criterion to verify the forecasting accuracy.

$$\text{MSE} = \frac{\sum_{t=w+1}^{N} (fv_t - av_t)^2}{N - w}$$  \hspace{1cm} (21)

where $N$ denotes total number of time series data, $w$ denotes the window basis, $fv_t$ refers to the forecasted value at time $t$, and $av_t$ refers to the actual value at time $t$.

### 3.2. Interval partition by PSO method.

The PSO method is used to minimize the MSE value by adjusting the interval lengths. Once all historical data are well trained by the PSO method, the interval partition which has the minimum MSE value is applied to forecast in the testing phase. In the FMPSO model, each particle explores the interval lengths in $U$. Let the number of intervals be $m$, and the lower and upper bounds of $U$ be $b_0$ and $b_m$, respectively. In Section 3.1, each particle defines a vector $b = \{b_1, b_2, \cdots, b_{m-1}\}$, where $b_i \leq b_{i+1}$. Then, vector $b$ divides $U$ into $m$ intervals which are $u_1 = (b_0, b_1]$, $u_2 = (b_1, b_2]$, $\cdots$, and $u_m = (b_{m-1}, b_m]$. When a particle moves to a new position, all elements in vector $b$ must be sorted in ascending order. The step-wise procedure of the FMPSO model in the training phase is illustrated in Figure 2.

| 1. | define the universe of discourse |
| 2. | define the fuzzy sets |
| 3. | initialize positions and velocities of all particles |
| 4. | **while** terminating condition (maximum iterations or minimum MSE criteria) is not reached **do** |
| 5. | **for** each particle **do** |
| 6. | divide $U$ into new intervals |
| 7. | fuzzify all historical data |
| 8. | create all FLRs |
| 9. | establish all FRMs |
| 10. | calculate the forecasted values |
| 11. | compute the MSE value |
| 12. | evaluate the fitness using the MSE value |
| 13. | update local best position and global best position by the fitness |
| 14. | modify the velocity |
| 15. | update the position |
| 16. | **end for** |
| 17. | **end while** |

**Figure 2.** FMPSO algorithm in the training phase
4. Application of Enrollment Forecasting. The empirical enrollments of the University of Alabama shown in Table 1 are adopted to illustrate our implementation. According to a step-by-step procedure in the previous section, the proposed model can be set out as follows.

4.1. Fuzzy forecasting method by the weighted FRMs on enrollments.

**Step 1.** Define the universe of discourse. Table 1 lists actual enrollments of the University of Alabama, where \( Y_{\text{min}} \) and \( Y_{\text{max}} \) are 13055 and 19337, respectively. Two proper numbers \( \Delta_{\text{min}} \) and \( \Delta_{\text{max}} \) are 55 and 663, respectively. Hence, the universe of discourse is defined as \( U = [13000, 20000] \).

**Table 1.** Enrollment data from the University of Alabama

<table>
<thead>
<tr>
<th>Year</th>
<th>Actual Enrollment</th>
<th>Year</th>
<th>Actual Enrollment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971</td>
<td>13055</td>
<td>1982</td>
<td>15433</td>
</tr>
<tr>
<td>1972</td>
<td>13563</td>
<td>1983</td>
<td>15497</td>
</tr>
<tr>
<td>1973</td>
<td>13867</td>
<td>1984</td>
<td>15145</td>
</tr>
<tr>
<td>1974</td>
<td>14696</td>
<td>1985</td>
<td>15163</td>
</tr>
<tr>
<td>1975</td>
<td>15460</td>
<td>1986</td>
<td>15984</td>
</tr>
<tr>
<td>1976</td>
<td>15311</td>
<td>1987</td>
<td>16859</td>
</tr>
<tr>
<td>1977</td>
<td>15603</td>
<td>1988</td>
<td>18150</td>
</tr>
<tr>
<td>1978</td>
<td>15861</td>
<td>1989</td>
<td>18970</td>
</tr>
<tr>
<td>1979</td>
<td>16807</td>
<td>1990</td>
<td>19328</td>
</tr>
<tr>
<td>1980</td>
<td>16919</td>
<td>1991</td>
<td>19337</td>
</tr>
<tr>
<td>1981</td>
<td>16388</td>
<td>1992</td>
<td>18876</td>
</tr>
</tbody>
</table>

Source: Song and Chissom (1993b), Song and Chissom (1994).

**Step 2.** Partition \( U \) into intervals. Following previous models [11,29], \( U \) is divided into seven intervals, \( u_1, u_2, \ldots, u_7 \), respectively. Initially, all intervals are of equal length: \( u_1 = (13000, 14000], u_2 = (14000, 15000], u_3 = (15000, 16000], u_4 = (16000, 17000], u_5 = (17000, 18000], u_6 = (18000, 19000] \) and \( u_7 = (19000, 20000] \).

**Step 3.** Define the fuzzy sets. For seven intervals in \( U \), each fuzzy set \( A_i \) \((1 \leq i \leq 7)\) is defined by the membership values of seven intervals.

\[
\begin{align*}
A_1 &= 1/u_1 + 0.5/u_2 + 0/u_3 + 0/u_4 + 0/u_5 + 0/u_6 + 0/u_7 \\
A_2 &= 0.5/u_1 + 1/u_2 + 0.5/u_3 + 0/u_4 + 0/u_5 + 0/u_6 + 0/u_7 \\
A_3 &= 0/u_1 + 0.5/u_2 + 1/u_3 + 0.5/u_4 + 0/u_5 + 0/u_6 + 0/u_7 \\
A_4 &= 0/u_1 + 0/u_2 + 0.5/u_3 + 1/u_4 + 0.5/u_5 + 0/u_6 + 0/u_7 \\
A_5 &= 0/u_1 + 0/u_2 + 0/u_3 + 0.5/u_4 + 1/u_5 + 0.5/u_6 + 0/u_7 \\
A_6 &= 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5 + 1/u_6 + 0.5/u_7 \\
A_7 &= 0/u_1 + 0/u_2 + 0/u_3 + 0/u_4 + 0.5/u_5 + 0.5/u_6 + 1/u_7 \\
\end{align*}
\]

Based on [8], seven linguistic values of enrollments 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.

**Step 4.** Fuzzify all historical data. Find out the interval to which each historical enrollment in Table 1 belongs. For example, the enrollment of 1976 is 15311, which falls within \((15000, 16000]\), and so it belongs to the interval \( u_3 \). According to formula (22), the fuzzy set \( A_3 \) has the maximum membership value at the interval \( u_3 \). Therefore, the historical data \( Y(1976) \) is fuzzified to \( A_3 \) with the linguistic value “many”, where
AN IMPROVED FORECASTING MODEL

$A_3 = [0 \ 0.5 \ 1 \ 0.5 \ 0 \ 0 \ 0]$. The completed fuzzified results of the University of Alabama are listed in Table 2.

**Step 5.** Create fuzzy logical relationships. Based on Formula (10) and Table 2, all first-order FLRs can be created in the training phase. For example, $F(1974) = A_2$ and $F(1975) = A_3$, and then the first-order FLR is $A_2 \rightarrow A_3$. In the case of high-order FLRs, based on formula (11) and Table 2, $F(1973) = A_1$, $F(1974) = A_2$, $F(1975) = A_3$ and $F(1976) = A_3$, and thus the 3rd-order FLR is $A_1, A_2, A_3 \rightarrow A_3$. In the testing phase, the linguistic value of $F(1993)$ is unknown. Hence, the 3rd-order of FLR is expressed as $A_7, A_7, A_6 \rightarrow \#$. Table 3 shows all the first-order and the 3rd-order FLRs.

**Table 2.** Fuzzy sets of enrollments

<table>
<thead>
<tr>
<th>Year</th>
<th>Actual Enrollment</th>
<th>Fuzzy set</th>
<th>Membership</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971</td>
<td>13055</td>
<td>$A_1$</td>
<td>([1 \ 0.5 \ 0 \ 0 \ 0 \ 0 \ 0])</td>
</tr>
<tr>
<td>1972</td>
<td>13563</td>
<td>$A_1$</td>
<td>([1 \ 0.5 \ 0 \ 0 \ 0 \ 0 \ 0])</td>
</tr>
<tr>
<td>1973</td>
<td>13867</td>
<td>$A_1$</td>
<td>([1 \ 0.5 \ 0 \ 0 \ 0 \ 0 \ 0])</td>
</tr>
<tr>
<td>1974</td>
<td>14696</td>
<td>$A_2$</td>
<td>([0.5 \ 1 \ 0.5 \ 0 \ 0 \ 0 \ 0])</td>
</tr>
<tr>
<td>1975</td>
<td>15460</td>
<td>$A_3$</td>
<td>([0 \ 0.5 \ 1 \ 0.5 \ 0 \ 0 \ 0])</td>
</tr>
<tr>
<td>1976</td>
<td>15311</td>
<td>$A_3$</td>
<td>([0 \ 0.5 \ 1 \ 0.5 \ 0 \ 0 \ 0])</td>
</tr>
<tr>
<td>1977</td>
<td>15603</td>
<td>$A_3$</td>
<td>([0 \ 0.5 \ 1 \ 0.5 \ 0 \ 0 \ 0])</td>
</tr>
<tr>
<td>1978</td>
<td>15861</td>
<td>$A_3$</td>
<td>([0 \ 0.5 \ 1 \ 0.5 \ 0 \ 0 \ 0])</td>
</tr>
<tr>
<td>1979</td>
<td>16807</td>
<td>$A_4$</td>
<td>([0 \ 0 \ 0.5 \ 1 \ 0.5 \ 0 \ 0])</td>
</tr>
<tr>
<td>1980</td>
<td>16919</td>
<td>$A_4$</td>
<td>([0 \ 0 \ 0.5 \ 1 \ 0.5 \ 0 \ 0])</td>
</tr>
<tr>
<td>1981</td>
<td>16388</td>
<td>$A_4$</td>
<td>([0 \ 0 \ 0.5 \ 1 \ 0.5 \ 0 \ 0])</td>
</tr>
<tr>
<td>1982</td>
<td>15433</td>
<td>$A_3$</td>
<td>([0 \ 0.5 \ 1 \ 0.5 \ 0 \ 0 \ 0])</td>
</tr>
<tr>
<td>1983</td>
<td>15497</td>
<td>$A_3$</td>
<td>([0 \ 0.5 \ 1 \ 0.5 \ 0 \ 0 \ 0])</td>
</tr>
<tr>
<td>1984</td>
<td>15145</td>
<td>$A_3$</td>
<td>([0 \ 0.5 \ 1 \ 0.5 \ 0 \ 0 \ 0])</td>
</tr>
<tr>
<td>1985</td>
<td>15163</td>
<td>$A_3$</td>
<td>([0 \ 0.5 \ 1 \ 0.5 \ 0 \ 0 \ 0])</td>
</tr>
<tr>
<td>1986</td>
<td>15984</td>
<td>$A_3$</td>
<td>([0 \ 0.5 \ 1 \ 0.5 \ 0 \ 0 \ 0])</td>
</tr>
<tr>
<td>1987</td>
<td>16859</td>
<td>$A_4$</td>
<td>([0 \ 0 \ 0.5 \ 1 \ 0.5 \ 0 \ 0])</td>
</tr>
<tr>
<td>1988</td>
<td>18150</td>
<td>$A_6$</td>
<td>([0 \ 0 \ 0 \ 0.5 \ 1 \ 0.5 \ 0])</td>
</tr>
<tr>
<td>1989</td>
<td>18970</td>
<td>$A_6$</td>
<td>([0 \ 0 \ 0 \ 0.5 \ 1 \ 0.5 \ 0])</td>
</tr>
<tr>
<td>1990</td>
<td>19328</td>
<td>$A_7$</td>
<td>([0 \ 0 \ 0 \ 0 \ 0.5 \ 1 \ 0])</td>
</tr>
<tr>
<td>1991</td>
<td>19337</td>
<td>$A_7$</td>
<td>([0 \ 0 \ 0 \ 0 \ 0.5 \ 1 \ 0])</td>
</tr>
<tr>
<td>1992</td>
<td>18876</td>
<td>$A_6$</td>
<td>([0 \ 0 \ 0 \ 0.5 \ 1 \ 0.5 \ 0])</td>
</tr>
</tbody>
</table>

**Step 6.** Establish the weighted FRMs. Take the 3rd-order FLRs to forecast enrollments. Thus, the window base of the FRMs is set to be 3. In the training phase, for example, forecasting begins in 1976, where the 3rd-order FLR is $A_1, A_2, A_3 \rightarrow A_3$. According to Formula (12) and Tables 2 and 3, the operation matrix $O^3(1976)$ is depicted as follows

$$O^3(1976) = \begin{bmatrix}
F(1975) \\
F(1974) \\
F(1973)
\end{bmatrix} =
\begin{bmatrix}
A_4 \\
A_2 \\
A_1
\end{bmatrix} =
\begin{bmatrix}
0 & 0.5 & 1 & 0.5 & 0 & 0 & 0 \\
0.5 & 1 & 0.5 & 0 & 0 & 0 & 0 \\
1 & 0.5 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

Based on the DW scheme, by decreasing the weights from 3 to 1, the weighted operation matrix $D^3(1976)$ is calculated as follows

$$D^3(1976) = \begin{bmatrix}
F(1975) \times 3 \\
F(1974) \times 2 \\
F(1973) \times 1
\end{bmatrix} =
\begin{bmatrix}
A_4 \times 3 \\
A_2 \times 2 \\
A_1 \times 1
\end{bmatrix} =
\begin{bmatrix}
0 & 1.5 & 3 & 1.5 & 0 & 0 & 0 \\
1 & 2 & 1 & 0 & 0 & 0 & 0 \\
1 & 0.5 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$
Based on the DW scheme, by decreasing the weights from 3 to 1, the weighted operation matrix $D^3(1993)$ is calculated as follows

$$
D^3(1993) = \begin{bmatrix}
\end{bmatrix} 
= \begin{bmatrix}
A_6 \\ A_7 \\ A_7
\end{bmatrix} = \begin{bmatrix}
0 & 0 & 0 & 0 & 0.5 & 1 & 0.5 \\
0 & 0 & 0 & 0 & 0 & 0.5 & 1 \\
0 & 0 & 0 & 0 & 0 & 0.5 & 1
\end{bmatrix}
$$
In this case, the next state of the FLR is unknown, and so the criterion matrix for the testing phase is set to $F(1992)$, as given below

$$C(1993) = F(1992) = A_6 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.5 & 1 & 0.5 \end{bmatrix}$$

The weighted FRM $R(1993)$ is computed by the product composition operation of $D^3(1993)$ and $C(1993)$ as follows

$$R(1993) = D^3(1993) \otimes C(1993) = \begin{bmatrix} 0 & 0 & 0 & 0.75 & 3 & 0.75 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 \end{bmatrix}$$

**Step 7.** Calculate the forecasted outputs. In the training phase, the raw membership of the forecasted output $F'(1976)$ is calculated by Formula (17) as follows

$$F'(1976) = (\max (R(1976)^T))^T = \begin{bmatrix} 0 & 1 & 3 & 0.75 & 0 & 0 & 0 \end{bmatrix}$$

where the membership values can be specified as: $p'_1 = 0$, $p'_2 = 1$, $p'_3 = 3$, $p'_4 = 0.75$, $p'_5 = 0$, $p'_6 = 0$ and $p'_7 = 0$. All membership values of $F'(1976)$ are summed up as follows

$$\sum_{i=1}^{7} p'_i = 0 + 1 + 3 + 0.75 + 0 + 0 + 0 = 4.75$$

Then, the raw membership of the forecasted output is normalized by Formula (18).

$$F(1976) = \begin{bmatrix} 0 & 1 & 3 & 0.75 & 0 & 0 & 0 \\ 4.75 & 4.75 & 4.75 & 4.75 & 4.75 & 4.75 & 4.75 \end{bmatrix} = \begin{bmatrix} 0 & 0.2105263 & 0.6315789 & 0.1578947 & 0 & 0 & 0 \end{bmatrix}$$

Table 4 shows all raw and normalized memberships of the forecasted outputs from 1974 to 1993.

According to Formula (19), seven midpoints are calculated by $m_1 = (13000 + 14000)/2$, $m_2 = (14000 + 15000)/2$, $m_3 = (15000 + 16000)/2$, $m_4 = (16000 + 17000)/2$, $m_5 = (17000 + 18000)/2$, $m_6 = (18000 + 19000)/2$ and $m_7 = (19000 + 20000)/2$. Let $M$ be a midpoint matrix with seven elements.

$$M = \begin{bmatrix} m_1 & m_2 & m_3 & m_4 & m_5 & m_6 & m_7 \\ 13500 & 14500 & 15500 & 16500 & 17500 & 18500 & 19500 \end{bmatrix} \quad (23)$$

Based on Formulas (20) and (23), the forecasted enrollment of year 1976 can be calculated as follows

$$fv = F'(1976) \cdot M^T = 14500 \times 0.2105263 + 15500 \times 0.6315789 + 16500 \times 0.1578947 = 15447.$$  

In the testing phase, the raw membership of the forecasted output $F'(1993)$ is computed by Formula (17) as given below

$$F'(1993) = (\max (R(1993)^T))^T = \begin{bmatrix} 0 & 0 & 0 & 0.75 & 3 & 1 \end{bmatrix}$$

where the membership values of $F'(1993)$ are specified as: $p'_1 = 0$, $p'_2 = 0$, $p'_3 = 0$, $p'_4 = 0$, $p'_5 = 0.75$, $p'_6 = 3$ and $p'_7 = 1$. All membership values of $F'(1993)$ are summed up as follows

$$\sum_{i=1}^{7} p'_i = 0 + 0 + 0 + 0.75 + 3 + 1 = 4.75$$

Then, the raw membership of the forecasted output is normalized by Formula (18).

$$F(1993) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.75 & 3 & 1 \\ 4.75 & 4.75 & 4.75 & 4.75 & 4.75 & 4.75 & 4.75 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.1578947 & 0.6315789 & 0.2105263 \end{bmatrix}$$
Table 4. Raw and normalized memberships of the forecasted outputs

<table>
<thead>
<tr>
<th>Year</th>
<th>Raw membership</th>
<th>Normalized membership</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971</td>
<td>[1.5 1.5 0 0 0 0 0]</td>
<td>[0.5 0.5 0 0 0 0 0]</td>
</tr>
<tr>
<td>1972</td>
<td>[0 1.5 1.5 0 0 0 0]</td>
<td>[0 0.5 0.5 0 0 0 0]</td>
</tr>
<tr>
<td>1973</td>
<td>[0 1 3 0.75 0 0 0]</td>
<td>[0.2105263 0.6315789 0.1578947 0 0 0]</td>
</tr>
<tr>
<td>1974</td>
<td>[0.75 3 0.75 0 0 0]</td>
<td>[0.1666667 0.6666667 0.1666667 0 0 0]</td>
</tr>
<tr>
<td>1975</td>
<td>[0.75 3 0.75 0 0 0]</td>
<td>[0.1666667 0.6666667 0.1666667 0 0 0]</td>
</tr>
<tr>
<td>1976</td>
<td>[0 1.5 1.5 0 0 0]</td>
<td>[0 0.5 0.5 0 0 0]</td>
</tr>
<tr>
<td>1977</td>
<td>[0 1.5 1.5 0 0 0]</td>
<td>[0 0.5 0.5 0 0 0]</td>
</tr>
<tr>
<td>1978</td>
<td>[0 1.5 1.5 0 0 0]</td>
<td>[0 0.5 0.5 0 0 0]</td>
</tr>
<tr>
<td>1979</td>
<td>[0 1.5 1.5 0 0 0]</td>
<td>[0 0.5 0.5 0 0 0]</td>
</tr>
<tr>
<td>1980</td>
<td>[0 1.5 1.5 0 0 0]</td>
<td>[0 0.5 0.5 0 0 0]</td>
</tr>
<tr>
<td>1981</td>
<td>[0 1.5 1.5 0 0 0]</td>
<td>[0 0.5 0.5 0 0 0]</td>
</tr>
<tr>
<td>1982</td>
<td>[0 1.5 1.5 0 0 0]</td>
<td>[0 0.5 0.5 0 0 0]</td>
</tr>
</tbody>
</table>

Based on Formulas (20) and (23), the forecasted enrollment of year 1993 is computed as follows

\[
f_{v} = F(1993) \cdot M^T = 17500 \times 1578947 + 18500 \times 0.6315789 + 19500 \times 0.2105263 = 18552.
\]

Table 5 shows all the results of the forecasted outputs in the training and testing phases.

Table 5. Enrollments of the forecasted outputs

<table>
<thead>
<tr>
<th>Year</th>
<th>Actual enrollment</th>
<th>Forecasted enrollment</th>
<th>Year</th>
<th>Forecasted enrollment</th>
<th>Actual enrollment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971</td>
<td>13055</td>
<td>14000</td>
<td>1983</td>
<td>15497</td>
<td>15553</td>
</tr>
<tr>
<td>1972</td>
<td>13563</td>
<td>14000</td>
<td>1984</td>
<td>15145</td>
<td>15500</td>
</tr>
<tr>
<td>1973</td>
<td>13867</td>
<td>15000</td>
<td>1985</td>
<td>15163</td>
<td>15500</td>
</tr>
<tr>
<td>1974</td>
<td>14696</td>
<td>15000</td>
<td>1986</td>
<td>15984</td>
<td>15500</td>
</tr>
<tr>
<td>1975</td>
<td>15460</td>
<td>15447</td>
<td>1987</td>
<td>15859</td>
<td>16000</td>
</tr>
<tr>
<td>1976</td>
<td>15311</td>
<td>15447</td>
<td>1988</td>
<td>18150</td>
<td>17500</td>
</tr>
<tr>
<td>1977</td>
<td>15603</td>
<td>15500</td>
<td>1989</td>
<td>18970</td>
<td>18500</td>
</tr>
<tr>
<td>1978</td>
<td>15861</td>
<td>15500</td>
<td>1990</td>
<td>19328</td>
<td>19000</td>
</tr>
<tr>
<td>1979</td>
<td>16807</td>
<td>16000</td>
<td>1991</td>
<td>19337</td>
<td>19250</td>
</tr>
<tr>
<td>1980</td>
<td>16919</td>
<td>16447</td>
<td>1992</td>
<td>18876</td>
<td>18885</td>
</tr>
<tr>
<td>1981</td>
<td>16388</td>
<td>16500</td>
<td>1993</td>
<td>19328</td>
<td>19000</td>
</tr>
<tr>
<td>1982</td>
<td>15433</td>
<td>16000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Step 8. Compute forecasting accuracy by MSE values. The MSE value of the forecasted enrollments from 1971 to 1992 is calculated by Formula (21) in the training phase. For example, the number of historical data $N$ is 22 and the window basis $w$ is set to be 3 for the 3rd-order FLRs. Based on Table 5, the MSE value of the forecasted enrollments is calculated as follows

$$\text{MSE} = \frac{\sum_{t=w+1}^{N} (f_{vt} - a_{vt})^2}{N - w} = \frac{\sum_{t=4}^{22} (f_{vt} - a_{vt})^2}{22 - 3}$$

$$= \frac{(14000 - 14696)^2 + (15000 - 15460)^2 + \cdots + (18885 - 18876)^2}{19}$$

$$= 212474$$

4.2. Interval partition by PSO method on enrollments. By using PSO method, each particle divides $U$ into new intervals to calculate a new MSE value until the terminating condition (maximum iterations or minimum MSE criteria) is reached. As the window base is 3 and the number of intervals is 7, the smallest MSE value is 25104.

5. Experimental Results. Actual enrollments of the University of Alabama are used to perform comparative study in the training and testing phases. Table 1 shows all historical enrollments from 1971 to 1992. In order to verify forecasting effectiveness, the proposed model is compared with those of corresponding models for various orders and different intervals.

5.1. Experimental results from the training phase. All historical enrollments from 1971 to 1992 are used as the training data in the training phase. Without loss of generality, the standard PSO model with the same experimental coefficients as the HPSO model [29] is applied. The coefficients of the FMPSO model are set as follows. The number of particles is 30, the maximal steps of movement for each particle are 100, the value of inertial weight $\omega$ is decreased linearly from 1.4 to 0.4 during the whole running generations, and the self confidence $c_1$ and the social confidence $c_2$ are both set to be 2. The velocity of each particle $V_{id}$ is limited to $[-100, 100]$. In our experiment, the FMPSO model performs 100 runs, and the best result is chosen to be the final result. All forecasting accuracies are evaluated by the MSE values.

To verify the forecasting accuracy of the proposed model for the first-order FLRs with different number of intervals, two fuzzy time series models in CC06a [13] and HPSO [29], are selected for purposes of comparison. Three fuzzy forecasting models are developed based on Chen’s method [11], and the window basis is set to be 1 for the first-order FLRs. A comparison of the forecasting results among these models is shown in Table 6. It is evident that the proposed model has smaller MSE values than the CC06a or HPSO models with different number of intervals. The major difference between the CC06a and FMPSO models is in the fuzzification and optimization methods used. The former performs the genetic algorithm on FLR rules, while the latter proceeds with the PSO algorithm on the FRMs to achieve the best interval lengths. As shown in Table 6, the PSO algorithm searches the virtual problem space more efficiently than the genetic algorithm. Although the HPSO model and the FMPSO model both use the PSO method, but the FMPSO model gets lower MSE values in forecasting. In addition, the FMPSO model obtains a lowest average MSE value of 24174 among three forecasting models. These findings suggest that the FMPSO model is able to provide effective forecasting capability for the first-order FLRs with different number of intervals.

To verify the forecasting effectiveness for high-order FLRs, the HCL98 [20], C02 [21], CC06b [12], and HPSO [29] models are used to compare with the proposed model. Table 7
Table 6. A comparison of the forecasting accuracy of the CC06a, HPSO and FMPSO models for the first-order FLRs with different number of intervals

<table>
<thead>
<tr>
<th>Model</th>
<th>Number of intervals</th>
<th>Average MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>CC06a [13]</td>
<td>132963</td>
<td>96244</td>
</tr>
<tr>
<td>HPSO [29]</td>
<td>119962</td>
<td>90527</td>
</tr>
<tr>
<td>FMPSO</td>
<td>30426(^\d)</td>
<td>24672(^\d)</td>
</tr>
</tbody>
</table>

\(^\d\) Best performance for compared models.

states that the proposed model has smaller MSE values than any of the models presented in HCL98 [20], C02 [21], CC06b [12] and HPSO [29]. The FMPSO model also gets the lowest MSE value of 8647 for the 9th-order FRMs (or window base 9) among all the compared models. From Figure 3, the graphical comparison clearly shows that the forecasting accuracy of the proposed model is more precise than those of existing models with different high-order FLRs. The average MSE value of the FMPSO model is 17781, which is smallest among all forecasting models compared.

Table 7. A comparison of the forecasting accuracy for seven intervals with different high-order FLRs

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>333171</td>
<td>89093</td>
<td>67834</td>
<td>67123</td>
<td>34469(^\d)</td>
</tr>
<tr>
<td>3</td>
<td>299634</td>
<td>86694</td>
<td>31123</td>
<td>31644</td>
<td>25104(^\d)</td>
</tr>
<tr>
<td>4</td>
<td>315489</td>
<td>89376</td>
<td>32009</td>
<td>23271</td>
<td>21234(^\d)</td>
</tr>
<tr>
<td>5</td>
<td>278919</td>
<td>94539</td>
<td>24948</td>
<td>23534</td>
<td>17200(^\d)</td>
</tr>
<tr>
<td>6</td>
<td>296950</td>
<td>98215</td>
<td>26980</td>
<td>23671</td>
<td>13936(^\d)</td>
</tr>
<tr>
<td>7</td>
<td>316720</td>
<td>104056</td>
<td>26969</td>
<td>20651</td>
<td>11314(^\d)</td>
</tr>
<tr>
<td>8</td>
<td>301228</td>
<td>102179</td>
<td>22387</td>
<td>17106</td>
<td>10347(^\d)</td>
</tr>
<tr>
<td>9</td>
<td>306485</td>
<td>102789</td>
<td>18734</td>
<td>17971</td>
<td>8647(^\d)</td>
</tr>
</tbody>
</table>

\(^\d\) Best performance for compared models.

To evaluate the performance of the proposed model with different number of intervals, five forecasting models in SC93b [8], C96 [11], H01H [19], CC06a [13] and HPSO [29] are selected for comparison. All forecasting models use the first-order FLRs (or window base 1) with different number of intervals. A comparison of all forecasting results is listed in Table 8. To be clearly visualized, Figure 4 depicts the trends for actual data and forecasted results. The smallest MSE value of 22340 is obtained from the proposed method. It is obvious that the forecasting accuracy of the proposed model is more precise than any existing models for the first-order FLRs with different number of intervals.

5.2. Experimental results from the testing phase. To verify the forecasting accuracy for future enrollments, the historical enrollments are divided into two parts for independent testing. One part is used as the training data and the other part is used as the testing data. In this experiment, the enrollments from 1990 to 1992 are used for testing individually. For example, to forecast a new enrollment of 1990, the enrollments of 1971-1989 are employed as the training data. Similarly, a new enrollment of 1991 can be forecasted based on the enrollments of 1971-1990. After the training data have been well trained by the proposed model and those of existing models, future enrollments could...
Figure 3. A comparison of the MSE values for 7 intervals with different high-order FLRs

Table 8. A comparison of the forecasted results for the first-order FLRs with different number of intervals

<table>
<thead>
<tr>
<th>Year</th>
<th>Actual data</th>
<th>SC93b</th>
<th>C96</th>
<th>H01H</th>
<th>CC06a</th>
<th>HPSO</th>
<th>FMPSO</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971</td>
<td>13055</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1972</td>
<td>13563</td>
<td>14000</td>
<td>14000</td>
<td>14000</td>
<td>13714</td>
<td>13555</td>
<td>13527</td>
</tr>
<tr>
<td>1973</td>
<td>13867</td>
<td>14000</td>
<td>14000</td>
<td>14000</td>
<td>13714</td>
<td>13994</td>
<td>13928</td>
</tr>
<tr>
<td>1974</td>
<td>14696</td>
<td>14000</td>
<td>14000</td>
<td>14000</td>
<td>14880</td>
<td>14711</td>
<td>14680</td>
</tr>
<tr>
<td>1975</td>
<td>15460</td>
<td>15500</td>
<td>15500</td>
<td>15500</td>
<td>15467</td>
<td>15344</td>
<td>15305</td>
</tr>
<tr>
<td>1976</td>
<td>15311</td>
<td>16000</td>
<td>16000</td>
<td>15500</td>
<td>15172</td>
<td>15411</td>
<td>15477</td>
</tr>
<tr>
<td>1977</td>
<td>15603</td>
<td>16000</td>
<td>16000</td>
<td>16000</td>
<td>15467</td>
<td>15411</td>
<td>15477</td>
</tr>
<tr>
<td>1978</td>
<td>15861</td>
<td>16000</td>
<td>16000</td>
<td>16000</td>
<td>15386</td>
<td>15411</td>
<td>15891</td>
</tr>
<tr>
<td>1979</td>
<td>16807</td>
<td>16000</td>
<td>16000</td>
<td>16000</td>
<td>16831</td>
<td>16816</td>
<td>16655</td>
</tr>
<tr>
<td>1980</td>
<td>16919</td>
<td>16813</td>
<td>16833</td>
<td>17500</td>
<td>17106</td>
<td>17140</td>
<td>17143</td>
</tr>
<tr>
<td>1981</td>
<td>16388</td>
<td>16813</td>
<td>16833</td>
<td>16000</td>
<td>16380</td>
<td>16464</td>
<td>16655</td>
</tr>
<tr>
<td>1982</td>
<td>15433</td>
<td>16789</td>
<td>16833</td>
<td>16000</td>
<td>15464</td>
<td>15505</td>
<td>15649</td>
</tr>
<tr>
<td>1983</td>
<td>15497</td>
<td>16000</td>
<td>16000</td>
<td>16000</td>
<td>15172</td>
<td>15411</td>
<td>15477</td>
</tr>
<tr>
<td>1984</td>
<td>15145</td>
<td>16000</td>
<td>16000</td>
<td>16000</td>
<td>15172</td>
<td>15411</td>
<td>15305</td>
</tr>
<tr>
<td>1985</td>
<td>15163</td>
<td>16000</td>
<td>16000</td>
<td>16000</td>
<td>15467</td>
<td>15344</td>
<td>15114</td>
</tr>
<tr>
<td>1986</td>
<td>15984</td>
<td>16000</td>
<td>16000</td>
<td>16000</td>
<td>15467</td>
<td>16018</td>
<td>15649</td>
</tr>
<tr>
<td>1987</td>
<td>16859</td>
<td>16000</td>
<td>16000</td>
<td>16000</td>
<td>16831</td>
<td>16816</td>
<td>16655</td>
</tr>
<tr>
<td>1988</td>
<td>18150</td>
<td>16813</td>
<td>16833</td>
<td>17500</td>
<td>18055</td>
<td>18060</td>
<td>18023</td>
</tr>
<tr>
<td>1989</td>
<td>18970</td>
<td>19000</td>
<td>19000</td>
<td>19000</td>
<td>18998</td>
<td>19014</td>
<td>18932</td>
</tr>
<tr>
<td>1990</td>
<td>19328</td>
<td>19000</td>
<td>19000</td>
<td>19000</td>
<td>19300</td>
<td>19340</td>
<td>19307</td>
</tr>
<tr>
<td>1991</td>
<td>19337</td>
<td>19000</td>
<td>19000</td>
<td>19500</td>
<td>19149</td>
<td>19340</td>
<td>19307</td>
</tr>
<tr>
<td>1992</td>
<td>18876</td>
<td>19000</td>
<td>19000</td>
<td>19000</td>
<td>19149</td>
<td>19014</td>
<td>18932</td>
</tr>
<tr>
<td>MSE</td>
<td>423027</td>
<td>407507</td>
<td>226611</td>
<td>35324</td>
<td>22965</td>
<td>22340</td>
<td></td>
</tr>
</tbody>
</table>

† Best performance for compared models.
be obtained to compare with testing data. Table 9 shows a comparison for actual enrollments and the forecasted results of the C96 [11], HPSO [29] and FMPSO models which use seven intervals with different order FLRs (orders 1-5).

**Table 9.** A comparison of actual enrollments and forecasted results for seven intervals in the testing phase

<table>
<thead>
<tr>
<th>Year</th>
<th>Actual Data</th>
<th>Order 1</th>
<th>Order 2</th>
<th>Order 3</th>
<th>Order 4</th>
<th>Order 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>C96</td>
<td>HPSO</td>
<td>FMPSO</td>
<td>HPSO</td>
<td>FMPSO</td>
</tr>
<tr>
<td>1990</td>
<td>19328</td>
<td>18500</td>
<td>18685</td>
<td>18856</td>
<td>18599</td>
<td>18868</td>
</tr>
<tr>
<td>1991</td>
<td>19337</td>
<td>19500</td>
<td>19138</td>
<td>19286</td>
<td>19246</td>
<td>19283</td>
</tr>
<tr>
<td>1992</td>
<td>18876</td>
<td>19000</td>
<td>19716</td>
<td>19295</td>
<td>19264</td>
<td>19297</td>
</tr>
</tbody>
</table>

Table 10 shows a comparison of the forecasted results of the HPSO and FMPSO models for seven intervals with different order FLRs (orders 1-9). The average MSE value of the FMPSO model is 117422, with a standard deviation (DEV) of 22807; for the HPSO model the average MSE value is 405233 with a standard DEV of 302990. It is obvious that the proposed model gets both a smaller standard DEV and an average MSE value than the HPSO model does.

**Table 10.** A comparison of the MSE values and the standard DEVs in the testing phase

<table>
<thead>
<tr>
<th>Model</th>
<th>Order 1</th>
<th>Order 2</th>
<th>Order 3</th>
<th>Order 4</th>
<th>Order 5</th>
<th>Order 6</th>
<th>Order 7</th>
<th>Order 8</th>
<th>Order 9</th>
<th>Average MSE</th>
<th>Standard DEV</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPSO</td>
<td>181017</td>
<td>230089</td>
<td>98607</td>
<td>148371</td>
<td>261209</td>
<td>425424</td>
<td>563948</td>
<td>778182</td>
<td>960252</td>
<td>405233</td>
<td>302990</td>
</tr>
<tr>
<td>FMPSO</td>
<td>133649</td>
<td>130586</td>
<td>128864</td>
<td>128246</td>
<td>120556</td>
<td>127039</td>
<td>121449</td>
<td>64182</td>
<td>96215</td>
<td>117422</td>
<td>22807</td>
</tr>
</tbody>
</table>

Best performance for compared models.

The trend of the curves in Figure 5 indicates the FMPSO model is still stable for the high-order FLRs. It confirms that the proposed model provides effective forecasting capability with different order FLRs. In addition, for seven intervals, the smallest MSE value of 64182 is obtained from the FMPSO model with the 8th-order FLRs (or window base 8).
6. **Conclusions.** The lengths of intervals and fuzzy relations are two critical factors that affect forecasting accuracy of time series data. In this paper, we have presented an improved forecasting model based on the weighted FRMs and PSO algorithm to approach these factors. Actual enrollments of the University of Alabama are used to verify the forecasting performance. The FMPSO model is compared with those of existing models for both the training and testing phases with various orders and different interval lengths.

The main contributions of this paper have been as follows. First, the FRMs with a decreasing weight (DW) scheme are more effective to capture fuzzy relations on time series data than FLR rules does. Second, the PSO algorithm for the optimized lengths of intervals is developed to adjust the interval lengths by searching the space of the universe of discourse. Third, based on the performance comparison in Tables 6-10 and Figures 3-5, the authors show the FMPSO model outperforms previous forecasting models for both the training and testing phases with various orders and different interval lengths.

Although this study shows the superior forecasting capability compared with existing forecasting models; but the FMPSO model is a new forecasting model and only tested by the enrollment data. To assess the effectiveness of the forecasting model, there are two suggestions for future research: (1) apply proposed model to deal with more complicated real-world problems for decision-making such as weather news, crop production, stock markets, and so on; and (2) use more intelligent methods (e.g., an ant colony or a neural network) to deal with forecasting problems.

**Acknowledgment.** This work was supported in part by the National Science Council under contract number NSC-98-2219-E-011-002, NSC-98-2221-E-011-133-MY3, NSC-98-2923-E-011-004-MY3 and NSC-99-2916-I-011-002-A1; and Yao-Lin Huang who is working in Department of Computer and Communication Engineering, De Lin Institute of Technology, New Taipei City, Taiwan.

**REFERENCES**


