A Particle Swarm Optimization Algorithm for Forecasting Based on Time Variant fuzzy Time Series

M.Mahnam & S.M.T.Fatemi Ghomi

ABSTRACT

Forecasting has always been a crucial challenge for managers and scientists to arrive at accurate decisions. Time-series analysis is an important tool for forecasting the future in terms of past history. Time-series methods are generally used when there is not much information about the generation process of the underlying variable and when other variables provide no clear explanation of the studied variable. A recent review of the literature on time series forecasting is considered by Gooijer and Hyndman [1]. Many techniques for time-series analysis have been developed assuming linear relationships among the series variables [2].

On the other hand, in many manufacturing and services industries, because of lacking sufficient historical data, traditional time series forecasting methods do not usually make reasonable judgments, so that there are large margins of errors between the predictive and actual values [3]. To solve this problem, Song and Chissom [4] first proposed the concept of fuzzy time series. The main property of the fuzzy time series is that the values of the demand variables are linguistic values.

There are two kinds of fuzzy time series: time invariant and time-invariant. If the relations are only between time and its prior time, it is a time-invariant fuzzy time series; otherwise, it is time-variant. In time-invariant fuzzy time series, several forecasting models have been proposed and implemented in different applications [5-15].

Recently, Liu [16] developed an integrated fuzzy time series forecasting system in which the forecasted value is a trapezoidal fuzzy number and applied it on two numerical data sets to compare to previous methods. Hsu et al. [17] proposed a modified turbulent particle swarm optimization method for the temperature prediction and the Taiwan Futures Exchange.
(TAIFEX) forecasting, based on the two-factor high order time-invariant series

One problem of these models is that they only group some heuristic rules that they have no response to trends and recent data are not different to older. So, the time-variant series are proposed by Song and Chissom [18] who considered a window base variable \( w \) controlling the relations to historical data. Hwang et al. [19] improved Song and Chissom’s method [4] by adding a heuristic function to get a better forecasting accuracy rate based on the variations of the historical data. These models are more sensitive to fluctuation of data, and should have more influence on prediction precision.

Chen and Hwang [20] further extended previous model to two-factor forecasting method, and applied it to forecast daily temperature. Lee et al. [21] presented genetic simulated annealing for forecasting the temperature and the TAIFEX based on two-factor high-order fuzzy time series. Kuo et al. [22] improved their method for forecasting enrollments based on the fuzzy time series and particle swarm optimization. Most existing methods in fuzzy time series assume that the intervals in the universe of discourse have the same length.

Huang and Yu [23] have shown that different lengths of intervals may affect the accuracy of forecast. Liu et al. [3] developed a time-variant fuzzy forecasting model based on a historic method for different length of intervals and window bases. Moreover, Liu and Wei [24] improved their method for seasonal time series.

In this paper, we propose a hybrid algorithm to deal with the forecasting problem based on time variant fuzzy time series and particle swarm optimization. The proposed algorithm determines the length of each interval in the universe of discourse and degree of membership values, simultaneously.

In order to evaluate the performance of the proposed algorithm, two well-known numerical data sets in the literature, i.e. the enrollment of the University of Alabama and the sales volume of products for a manufacturing company, are selected and compare the forecasting accuracy with three fuzzy time series methods presented by Hwang et al. [19], Lee and Chou [10], and Liu et al. [16].

The remainder of the paper is organized as follows. Section 2 describes the basic concepts and formulation of the problem in details. Section 3 explains the standard particle swarm optimization algorithm; and our hybrid approach is presented in section 4. Then, the experimental results are illustrated and analyzed in Section 5. Finally, Section 6 provides conclusion and suggestions for further study.

2. Fuzzy Time series

In this section, we briefly review the basic concepts of fuzzy time series and the proposed model in this study.

2-1. Basic Concepts

Fuzzy time series was first presented and defined by Song and Chissom [4]. A brief overview of the fuzzy time series definitions in the literature is included within the forecasting procedure is described as follows:

Let \( U \) be the universe of discourse, where \( U=\{u_1,u_2,\ldots,u_n\} \) and let \( A \) be a fuzzy set in the universe of discourse \( U \) defined as follows:

\[
A= f_1(u_1)/u_1 + f_2(u_2)/u_2 + \ldots + f_m(u_m)/u_m
\]

where \( f_i \) is the membership function of the fuzzy set \( A \) such that \( f_i: U \rightarrow [0, 1] \) and \( f_i \) (\( u_k \)) represents the grade of membership of \( u_k \).

Let \( X(t) \) (\( t = 0, 1, 2, \ldots \)) a subset of a real number, be the universe of discourse on which fuzzy sets \( f_j(t), j=1,2,\ldots,n \) are defined. \( F(t) \) is a collection of \( f_j(t) \), then \( F(t) \) is called a fuzzy time series on \( X(t) \) (\( t = 0, 1, 2,\ldots \)). Therefore, \( F(t) \) is a linguistic variable and \( f_j(t) \) as the possible linguistic value of \( F(t) \). If there exists a fuzzy relationship \( R(t-1,t) \), such that \( F(t) = F(t-1) \circ R(t-1,t) \) then \( F(t) \) is said to be caused by \( F(t-1) \); wherein \( \circ \) is an Max–Min composition operator. Considering a fuzzy logical relationship \( A \rightarrow A_j \), where \( A=F(t-1) \) and \( A_j=F(t), A, \) and \( A_j \) are the left and right-hand sides of the fuzzy logical relationship, respectively.

If \( R(t-1,t) \) is independent of \( t \), then \( F(t) \) is considered as a time-invariant fuzzy time series; otherwise, \( F(t) \) is a time variant fuzzy time series whether it is caused by \( F(t-1),F(t-2), \ldots, \), and \( F(t-m), (m>0) \). In this forecasting method, the relation can be expressed as the fuzzy relational equation:

\[
F(t) = F(t-1) \circ R^n(t-1, t)
\]

where \( w \) is the number of years which the forecast is being affected.

2-2. Proposed Model

As mentioned earlier, this study, similar to Liu et al. [3], uses Hwang et al.’s [19] fuzzy model as a basis to develop the proposed hybrid method. The steps of their method are briefly introduced as follows:

Step 1: Compute the variations of the historical data. The variation \( V_t \) of the data between time \( t(d) \) and \( t-1 (d-1) \) is computed as \( V_t = d_t - d_{t-1} \) (\( t = 2, 3, \ldots, n \)).

Step 2: Define the universe of discourse. Find the maximum (\( D_{\text{max}} \)) and the minimum (\( D_{\text{min}} \)) among all \( V_t \). The universe of discourse \( U \) is then defined as \( U = [D_{\text{min}}, D_1, D_{\text{max}} + D_2] \) where \( D_1 \) and \( D_2 \) are two proper positive numbers.

Step 3: In this study, a fuzzy number is considered to fuzzify intervals. In this step, appropriate interval length and degree of membership values should be determined by particle swarm optimization.
Step 4: Fuzzify the variations of historical data. If the variation \( V_t \) is within the scope of \( u_i \), it belongs to fuzzy set \( A_i \). All of the variations must be classified into the corresponding fuzzy sets.

Step 5: Calculate the fuzzy time series \( F(t) \) at window base \( w \). The operation matrix \( O^w(t) \) and the criterion matrix \( C(t) \) are selected to compute the fuzzy forecasted variation \( F(t) \).

\[
C(t) = F(t-1) = [C_1 \ C_2 \ ... \ C_m]
\]

\[
O^w(t) = \begin{bmatrix}
F(t-2) \\
F(t-3) \\
\vdots \\
F(t-w)
\end{bmatrix}
\]

\[
O^w(t) = \begin{bmatrix}
O_{11} \ O_{12} \ ... \ O_{1m} \\
O_{21} \ O_{22} \ ... \ O_{2m} \\
\vdots \\
O_{(w-1)1} \ O_{(w-1)2} \ ... \ O_{(w-1)m}
\end{bmatrix}
\]

where \( C_j \) indicates the membership value at the interval \( u_j \) within fuzzy set \( A_j \). The fuzzy relation matrix \( R(t) \) is computed through performing the following fuzzy composition operation.

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

\[
R(t) = \begin{bmatrix}
0_{11} \ C_1 \ 0_{12} \ C_2 \ ... \ 0_{1m} \ C_m \\
0_{21} \ C_1 \ 0_{22} \ C_2 \ ... \ 0_{2m} \ C_m \\
\vdots \\
0_{(w-1)1} \ C_1 \ 0_{(w-1)2} \ C_2 \ ... \ 0_{(w-1)m} \ C_m
\end{bmatrix}
\]

Then, \( F(t) \) can be calculated as the maximum of every column in \( R(t) \) as follows:

\[
F(t) = \max_{k=1 \ldots m} [R_{k1}^{(t-1)} \ldots R_{km}^{(t-1)}]
\]

Step 6: Forecasted value. Suppose there are \( k \) non-zero values corresponding to intervals \( u_1, u_2, ..., u_k \), with their midpoints \( m_1, m_2, ..., m_k \), respectively. Similar to Liu et al. [21] defuzzification method, we use weighted average method to calculate the defuzzified variation \( CV_t \):

\[
CV_t = \frac{\sum_{i=1}^{k} m_i \cdot f_{i1}}{\sum_{i=1}^{k} f_{i1}}
\]

The forecasted value \( FV_t \) at time \( t \) is computed as follows:

\[
FV_t = CV_t + d_{a1}
\]

3. Particle Swarm Optimization

Particle swarm optimization (PSO) is an evolutionary and population based on the stochastic optimization technique proposed by Eberhart and Kennedy [25, 26]. PSO was first introduced to optimize various continuous nonlinear functions and requires only primitive and simple mathematical operators. In this method, each solution is like a bird in the search space, called “particle”.

All particles have fitness values which are evaluated by fitness functions. Also, each particle has a velocity which determines its flight direction. Generally, particles fly in search space following particles with best solution. Initially, PSO consists of a randomly produced population and velocity. Then, the velocity is dynamically adjusted at each step according to the experience by itself and its colleagues as given by Eq. (9). The new particle position is found by adding the new velocity to the current position (Eq. (10)).

\[
V_{i,t+1} = w \cdot V_{i,t} + R_1 \cdot C_1 \cdot (P_{i,t} - X_{i,t}) + R_2 \cdot C_2 \cdot (P_{g,t} - X_{i,t})
\]

\[
X_{i,t+1} = X_{i,t} + V_{i,t+1}
\]

Where \( i \) is the \( i \)th particle; \( X_{i,t} \) is the position of particle \( i \) in iteration \( t \); \( V_{i,t} \) is the velocity of particle \( i \) in iteration \( t \); \( P_i \) is the best previous position of particle \( i \) so far (pbest) and \( P_g \) is the best previous position among all the particles (gbest). \( w \) is inertia weight and its function is to balance global and local exploitations of the swarm. The most applicable way of using inertia weight is linear decreasing [27], which is determined as follows:

\[
w = w_{max} - \left( \frac{w_{max} - w_{min}}{iter_{max}} \right) \cdot iter
\]

Where \( w_{max} \) is the initial value of weighting coefficient; \( w_{min} \), the final value of weighting coefficient; \( iter_{max} \), maximum number of iterations; and, \( iter \) is the current iteration. \( C_1 \) and \( C_2 \) are two learning factors which control the influence of pbest and gbest on the search process and are usually set 2 to cover the whole region of pbest and gbest. \( R_1 \) and \( R_2 \) are two random numbers within the range of \([0, 1]\).

The process of PSO algorithm is as follows:

Step 1: Initialize a population of particles with random positions and velocities in the \( D \)-dimensional problem space.

Step 2: Evaluate the objective values of all particles, set pbest of each particle equal to its current position, and set gbest equal to the position of the best initial particle.

Step 3: Update the velocity and position of particles according to Eqs. (9) and (10).

Step 4: Evaluate the objective values of all particles.

Step 5: For each particle, compare its current objective value with its pbest value. If the current value
is better, update pbest with the current position and objective value.

**Step 6:** Determine the best particle of the current whole population with the best objective value. If the objective value is better than that of gbest, update gbest with the current best particle.

**Step 7:** If a stopping criterion is met, output gbest and its objective value; otherwise, go back to Step 3.

### 4. Proposed Algorithm

This paper develops a novel integrated approach to determine the length of intervals and membership value, simultaneously. In this algorithm, to obtain the best forecasted value, different possible combinations of window base (w) and number of intervals (m) are considered then for each combination of (w, m), PSO finds a near-optimal solution for the length of intervals and degree of membership values by updating the particles based on MAD objective function and their pbest and gbest in each iteration.

#### 4-1. Particle Representation

In this approach, each particle consists of the two parts, $X_1$ and $X_2$, the former defining the length of intervals (consists of $m\times1$ elements) and the latter indicating the value of memberships (consisting of two elements).

Let the number of the intervals be $m$, a particle is a vector consisting of $m\times1$ elements, such that $0<q_i<1$ for $i=1,...,m\times1$ and $\sum_{i=1}^{mn} q_i = 1$. In order to determine the intervals, the elements of the corresponding particle should be multiplied by the universe of discourse, i.e. $b_i = q_i \times (D_{\max} - D_{\min})$. Therefore, the $m$ intervals are respectively as follows:

$$
\begin{align*}
I_1 &= [D_{\min}, D_{\max} + b_1], \\
I_2 &= [D_{\min}, D_{\max} + b_1 + b_2], \\
&\vdots \\
I_{m-1} &= [D_{\min} + \sum_{i=1}^{m-2} b_i, D_{\max} + \sum_{i=1}^{m-2} b_i], \\
I_m &= [D_{\min} + \sum_{i=1}^{m-1} b_i, D_{\max}]
\end{align*}
$$

By updating the position of particles in step 3, the elements of the corresponding new vector need to be divided by the summation of elements to ensure that each element $b_i$ is in domain (0, 1). Moreover, the value of memberships should be in domain (0, 1).

Moreover, $X_2$ indicates the membership values of fuzzy intervals. To reduce the solution space of $X_2$, only three intervals can be considered as non-zero membership value. These three intervals are entitled lower, normal, and upper intervals, respectively, in this method. Also the membership values of lower, normal, and upper intervals are bounded between 0.3 and 0.7, which improved the results.

The important feature of this representation is that all off-springs formed by the algorithm are feasible solutions and its application is faster than other representations in the literature. Fig. 1 shows a sample of particle representation.

#### 4-2. Initial Swarm

The population size selected is problem-dependent with the sizes of 20-50 as the most common sizes [27], and initial particles are generated randomly.

#### 4-3. Neighborhood Structure

The neighborhood of a particle is the social environment a particle encounters. As generalized form of $gbest$, $nbest$ can be used as the best position of $n$ neighbor particles achieved so far. Proposed algorithm uses social neighborhood which is a list of particles regardless of their positions. Since it may occasionally happen that the algorithm is easily trapped into the local optima, a neighborhood structure is used to prevent the particle from pre-convergence. Therefore, if after $g_1$ iterations, $gbest$ shows no improvement, global neighborhood will replace the local one. This decision provides the possibility for searching boundaries between neighborhoods. Finally, if the algorithm reaches the local optima in the second stage – i.e., if the result is not improved after $g_2$ iterations – the algorithm will stop.

#### 5. Experimental Results and Analyses

To analyze the efficiency of the proposed forecasting method, experiments were conducted on two numerical examples: (1) the enrollments of the University of Alabama, and (2) the production value of the machinery industry in Taiwan.

The algorithm was executed on the Microsoft Visual C++ software on a Pentium IV 2 GHz and 512MB RAM computer. Experimental results for PSO model are compared with those of existing methods. The PSO algorithm is executed 5 runs for each scenario, and the best result is taken to be the final result. In our pilot experiment, we tested different values of algorithm parameters, and it was proved that the following values were more effective:

$N$, swarm size: 30;
$L$, neighborhood size: 10;
$g_1$, iteration with no improvement in neighborhood: 300;
$g_2$, iteration with no improvement: 400
$w_{\max}$, maximum inertial weight: 1.4;
$w_{\min}$, minimum inertial weight: 0.4;
$C_1$, acceleration coefficient toward $pbest$: 1;
$C_2$, acceleration coefficient toward $gbest$: 2.

Also, in order to evaluate the efficiency of the forecasting model, a well-known criterion employed in the literature, Mean Absolute Deviation (MAD), used and calculated as follows:

$$\text{MAD} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$

Fig. 1. A sample of the particle representation

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>0.7</td>
<td>0.8</td>
</tr>
<tr>
<td>…</td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td></td>
</tr>
</tbody>
</table>

International Journal of Industrial Engineering & Production Research, December 2012, Vol. 23, No. 4
where \( n \) is the number of historical data.

5.1. Enrollments of the University of Alabama
The first example is the yearly number of students enrolled at the University of Alabama. The historical data of enrollments from 1971 to 1991 are shown in Table 1. The goal is to predict the student enrollment in 1992.

<table>
<thead>
<tr>
<th>No.</th>
<th>Year</th>
<th>Enrollment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1971</td>
<td>13055</td>
</tr>
<tr>
<td>2</td>
<td>1972</td>
<td>13563</td>
</tr>
<tr>
<td>3</td>
<td>1973</td>
<td>13867</td>
</tr>
<tr>
<td>4</td>
<td>1974</td>
<td>14696</td>
</tr>
<tr>
<td>5</td>
<td>1975</td>
<td>15460</td>
</tr>
<tr>
<td>6</td>
<td>1976</td>
<td>15311</td>
</tr>
<tr>
<td>7</td>
<td>1977</td>
<td>15603</td>
</tr>
<tr>
<td>8</td>
<td>1978</td>
<td>15861</td>
</tr>
<tr>
<td>9</td>
<td>1979</td>
<td>16007</td>
</tr>
<tr>
<td>10</td>
<td>1980</td>
<td>16919</td>
</tr>
<tr>
<td>11</td>
<td>1981</td>
<td>16388</td>
</tr>
<tr>
<td>12</td>
<td>1982</td>
<td>15433</td>
</tr>
<tr>
<td>13</td>
<td>1983</td>
<td>15497</td>
</tr>
<tr>
<td>14</td>
<td>1984</td>
<td>15145</td>
</tr>
<tr>
<td>15</td>
<td>1985</td>
<td>15163</td>
</tr>
<tr>
<td>16</td>
<td>1986</td>
<td>15984</td>
</tr>
<tr>
<td>17</td>
<td>1987</td>
<td>16859</td>
</tr>
<tr>
<td>18</td>
<td>1988</td>
<td>18150</td>
</tr>
<tr>
<td>19</td>
<td>1989</td>
<td>18970</td>
</tr>
<tr>
<td>20</td>
<td>1990</td>
<td>19328</td>
</tr>
<tr>
<td>21</td>
<td>1991</td>
<td>19337</td>
</tr>
</tbody>
</table>

Let \( Y(t) \) be the historical data on year \( t \). According to step 1, the variations of the enrollments \( (V_t) \) can be easily calculated, as shown in Table 2. Then, the universe of discourse on \( V_t \) is considered as domain \([-1000, 1300]\) where \( D_{nu}=1291, D_{nu}=955, D_t=45, \) and \( D_n=8.\)

Assume that the number of intervals \( (m) \) determined by PSO be seven and the lengths of intervals are as follows:

\[
\begin{align*}
I_1 &= (-1000, -973) \\
I_2 &= (-973, -744) \\
I_3 &= (-744, -384) \\
I_4 &= (-384, 117) \\
I_5 &= (117, 821) \\
I_6 &= (821, 848) \\
I_7 &= (848, 1300)
\end{align*}
\]

Also, the degree of membership values of lower, normal, and upper intervals are 0.69 and 0.7, respectively. So, in this example, the definitions of fuzzy sets and their linguistic variables are described based on Eq. (1) would be defined as follows:

\[
\begin{align*}
A_1 &= \frac{1}{I_1} + 0.7/ I_2 + 0/ I_3 + 0/ I_4 + 0/ I_5 + 0/ I_6 + 0/ I_7 \\
&\quad \text{“not many”} \\
A_2 &= 0.69/ I_1 + 0.7/ I_2 + 0/ I_3 + 0/ I_4 + 0/ I_5 + 0/ I_6 + 0/ I_7 \\
&\quad \text{“not too many”} \\
A_3 &= 0/ I_1 + 0.69/ I_2 + 0.7/ I_3 + 0/ I_4 + 0/ I_5 + 0/ I_6 + 0/ I_7 \\
&\quad \text{“many”} \\
A_4 &= 0/ I_1 + 0/ I_2 + 0.69/ I_3 + 1/ I_4 + 0.7/ I_5 + 0/ I_6 + 0/ I_7 \\
&\quad \text{“very many”} \\
A_5 &= 0/ I_1 + 0/ I_2 + 0/ I_3 + 0.69/ I_4 + 1/ I_5 + 0.7/ I_6 + 0/ I_7 \\
&\quad \text{“too many”} \\
A_6 &= 0/ I_1 + 0/ I_2 + 0/ I_3 + 0/ I_4 + 0.69/ I_5 + 1/ I_6 + 0.7/ I_7 \\
&\quad \text{“too many many”} \\
A_7 &= 0/ I_1 + 0/ I_2 + 0/ I_3 + 0/ I_4 + 0.69/ I_5 + 1/ I_6 + 0.7/ I_7 \\
&\quad \text{“very many many”}
\end{align*}
\]

In order to fuzzify the historical data, each should be assigned to the fuzzy set (linguistic value) which its interval contains the enrollment value. For example, the enrollment deviation on year 1975 is 764, and it belongs to interval \( I_3 \), therefore, fuzzy set \( A_5 \) is assigned to this value. More details of fuzzification for this example are shown in Table 2.

<table>
<thead>
<tr>
<th>No.</th>
<th>Year</th>
<th>Actual Data</th>
<th>( V_t )</th>
<th>Fuzzy sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1971</td>
<td>13055</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>1972</td>
<td>13563</td>
<td>508</td>
<td>A5</td>
</tr>
<tr>
<td>3</td>
<td>1973</td>
<td>13867</td>
<td>304</td>
<td>A5</td>
</tr>
<tr>
<td>4</td>
<td>1974</td>
<td>14696</td>
<td>829</td>
<td>A6</td>
</tr>
<tr>
<td>5</td>
<td>1975</td>
<td>15460</td>
<td>764</td>
<td>A5</td>
</tr>
<tr>
<td>6</td>
<td>1976</td>
<td>15311</td>
<td>-149</td>
<td>A4</td>
</tr>
<tr>
<td>7</td>
<td>1977</td>
<td>15603</td>
<td>292</td>
<td>A5</td>
</tr>
<tr>
<td>8</td>
<td>1978</td>
<td>15861</td>
<td>258</td>
<td>A5</td>
</tr>
<tr>
<td>9</td>
<td>1979</td>
<td>16807</td>
<td>946</td>
<td>A7</td>
</tr>
<tr>
<td>10</td>
<td>1980</td>
<td>16919</td>
<td>112</td>
<td>A4</td>
</tr>
<tr>
<td>11</td>
<td>1981</td>
<td>16388</td>
<td>-531</td>
<td>A3</td>
</tr>
<tr>
<td>12</td>
<td>1982</td>
<td>15433</td>
<td>-955</td>
<td>A2</td>
</tr>
<tr>
<td>13</td>
<td>1983</td>
<td>15497</td>
<td>64</td>
<td>A4</td>
</tr>
<tr>
<td>14</td>
<td>1984</td>
<td>15145</td>
<td>-352</td>
<td>A4</td>
</tr>
<tr>
<td>15</td>
<td>1985</td>
<td>15163</td>
<td>18</td>
<td>A4</td>
</tr>
<tr>
<td>16</td>
<td>1986</td>
<td>15984</td>
<td>821</td>
<td>A6</td>
</tr>
<tr>
<td>17</td>
<td>1987</td>
<td>16859</td>
<td>875</td>
<td>A7</td>
</tr>
<tr>
<td>18</td>
<td>1988</td>
<td>18150</td>
<td>1291</td>
<td>A7</td>
</tr>
<tr>
<td>19</td>
<td>1989</td>
<td>18970</td>
<td>820</td>
<td>A5</td>
</tr>
<tr>
<td>20</td>
<td>1990</td>
<td>19328</td>
<td>358</td>
<td>A5</td>
</tr>
<tr>
<td>21</td>
<td>1991</td>
<td>19337</td>
<td>9</td>
<td>A4</td>
</tr>
</tbody>
</table>

In order to calculate the forecasted value in 1980 in our example, by considering window base 3, the operation matrix \( O^3(1980) \) and the criterion matrix \( C(1980) \) are as follows:

\[
C(1980) = F(1979) = [0 0 0 0 0 0 0.69 1]
\]

\[
O^3(1980) = [F(1978)] = [0 0 0 0.69 1 0.7 0 0]
\]

So, \( R(1980) \) is computed as follows:

\[
R(1980) = \begin{bmatrix}
0 & 0 & 0 & 0 & 0.488 & 0 & 0 & 0
\end{bmatrix}
\]
Based on Eq. (8), $F(1980)$ is

$$F(1980) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.488 & 0 \end{bmatrix}$$

$Cv_{1980}$ and $F_{1980}$ are computed based on Eqs. (8) and (9) and the forecasted value in 1980 is obtained as follows:

$$F_{1980} = Cv_{1980} + d_{1980} = 407.50 + 16078 = 17214.5$$

It is necessary to analyze the objective value of different values of $m$ and $w$ and select the best results for comparing with other methods. Table 3 shows the results obtained from the proposed algorithm for different values of intervals ($m$) and window base ($w$).

**Tab. 3. Residual values of $m$ and $w$ for enrollments of the University of Alabama**

<table>
<thead>
<tr>
<th>$m$</th>
<th>$w$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>419.3</td>
<td>390.2</td>
<td>390.2</td>
<td>396.91</td>
<td>417.16</td>
<td>422.61</td>
<td>357.42</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>402.7</td>
<td>383.9</td>
<td>379.03</td>
<td>394.66</td>
<td>414.81</td>
<td>421.94</td>
<td>353.01</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>406.7</td>
<td>378.9</td>
<td>385.06</td>
<td>387.76</td>
<td>410.29</td>
<td>424.08</td>
<td>337.29</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>407.5</td>
<td>367.5</td>
<td>375.87</td>
<td>381.77</td>
<td>406.13</td>
<td>412.34</td>
<td>352.67</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>403.7</td>
<td>383.9</td>
<td>395.01</td>
<td>376.41</td>
<td>403.77</td>
<td>418.53</td>
<td>343.55</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>410.3</td>
<td>373.8</td>
<td>356.42</td>
<td>374.46</td>
<td>405.23</td>
<td>412.50</td>
<td>346.94</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>400.9</td>
<td>374.1</td>
<td>375.12</td>
<td>370.74</td>
<td>405.45</td>
<td>417.11</td>
<td>343.84</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>401.1</td>
<td>377.3</td>
<td>378.70</td>
<td>371.52</td>
<td>404.86</td>
<td>415.63</td>
<td>350.84</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>411.5</td>
<td>369.9</td>
<td>373.49</td>
<td>396.69</td>
<td>412.78</td>
<td>344.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>394.4</td>
<td>373.2</td>
<td>387.56</td>
<td>361.07</td>
<td>403.67</td>
<td>413.10</td>
<td>323.93</td>
<td></td>
</tr>
</tbody>
</table>

Also, the trends of combinations are shown in Fig. 2. Although, firstly, an increasing trend is observed in the objective values by increasing the window base, in the last few series a decreasing trend is shown. Moreover, no special trend is observed in the number of intervals ($m$). So, the best result is the combination of $m=14$ and $w=8$.

Moreover, to evaluate the performance of the proposed method, it is compared with three previous fuzzy time series models developed for this application in Table 4 [10, 16, 21]. It is revealed that the objective value of proposed algorithm is better than the previous methods for all $w$ values. It is because of this fact that the proposed algorithm considered more flexibility on its parameters by taking into account variability on them.

**Tab. 4. The comparison of proposed method to previous algorithms for enrollments of the University of Alabama**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1971</td>
<td>13655</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>1972</td>
<td>13563</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>1973</td>
<td>13867</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>1974</td>
<td>14696</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>1975</td>
<td>15460</td>
<td>15296</td>
<td>15296</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>1976</td>
<td>15311</td>
<td>16260</td>
<td>16260</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>1977</td>
<td>15603</td>
<td>15711</td>
<td>15511</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>1978</td>
<td>15861</td>
<td>16003</td>
<td>16003</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>1979</td>
<td>16807</td>
<td>16261</td>
<td>16261</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>1980</td>
<td>16919</td>
<td>17407</td>
<td>17407</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>1981</td>
<td>16388</td>
<td>17319</td>
<td>17119</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>1982</td>
<td>15433</td>
<td>16188</td>
<td>16188</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>1983</td>
<td>15497</td>
<td>14833</td>
<td>14833</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>1984</td>
<td>15145</td>
<td>15097</td>
<td>15497</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>1985</td>
<td>15163</td>
<td>14945</td>
<td>14745</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>1986</td>
<td>15984</td>
<td>14963</td>
<td>15163</td>
<td>15163</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>1987</td>
<td>16859</td>
<td>16384</td>
<td>16384</td>
<td>16384</td>
<td>16384</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>1988</td>
<td>18150</td>
<td>17659</td>
<td>17659</td>
<td>17659</td>
<td>17659</td>
<td>17659</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>1989</td>
<td>18970</td>
<td>19150</td>
<td>19150</td>
<td>19150</td>
<td>19150</td>
<td>18369</td>
<td>19030</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>1990</td>
<td>19228</td>
<td>19970</td>
<td>19770</td>
<td>19770</td>
<td>19770</td>
<td>19770</td>
<td>19454</td>
<td>19454</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>1991</td>
<td>19337</td>
<td>19928</td>
<td>19928</td>
<td>19928</td>
<td>19928</td>
<td>19728</td>
<td>19454</td>
<td>19454</td>
<td>19454</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>1992</td>
<td>18876</td>
<td>19537</td>
<td>19537</td>
<td>19537</td>
<td>19537</td>
<td>19454</td>
<td>19454</td>
<td>19454</td>
<td>19454</td>
<td>19417</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

**Fig. 2. Residual trends of $m$ and $w$ for enrollments of the University of Alabama**

**5.2. Sales Volume of Products for a Manufacturing Company**

The second example considered in the literature is the sales volume of products for one of the largest polypropylene manufacturing companies in Taiwan. The sales volume of the products is from January to December 2002 which is shown in Table 5.
Liu et al. [3] have shown that the datum in June is an outlier and it should not be considered. So, we do not consider this datum too and remove it from the historical data. All steps of the proposed method are followed similar to the first example. Table 6 shows the results of the proposed algorithm for different values of intervals (m) and window base (w).

![Fig. 3. Residual trends of m and w for sales volume of products](image)

Furthermore, the forecasted results of the four methods are shown in Table 7. Similar to the first example, the MAD criterion is better than the previous methods for all window base values. In addition, the best result is obtained at w=4 and m=11.

### 6. Conclusions

This paper presented a novel hybrid algorithm for the forecasting problem based on time variant fuzzy time series and particle swarm optimization. The proposed algorithm determines the length of each interval in the universe of discourse and degree of membership values, simultaneously. In order to analyze the number of intervals and window base parameters and evaluate the efficiency of the proposed algorithm, two numerical data sets are selected, and compared with previous fuzzy time series methods in the literature based on the forecasting accuracy. The results indicate that the proposed algorithm satisfactorily competes well with similar approaches.
Future research may focus on the development of other heuristic and metaheuristic methods such as genetic algorithm and ant colony optimization trying to improve the quality of the forecasts.

References


