International Journal of Industrial Engineering & Production Research

December 2012, Volume 23, Number 4
pp. 269-276

http://IJIEPR.iust.ac.ir/

ISSN: 2008-4889

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

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

***Mehdi Mahnam**, Department of Industrial Engineering, Amirkabir University of Technology, Tehran, Iran*
***Seyyed Mohammad Taghi Fatemi Ghomi**, Professor of Industrial Engineering, Amirkabir University of Technology, Tehran, Iran*

KEYWORDS

Forecasting, Fuzzy Time Series, Time Variant,
Particle Swarm Optimization

ABSTRACT

Fuzzy time series have been developed during the last decade to improve the forecast accuracy. Many algorithms have been applied in this approach of forecasting such as high order time invariant fuzzy time series. In this paper, we present a hybrid algorithm to deal with the forecasting problem based on time variant fuzzy time series and particle swarm optimization algorithm, as a highly efficient and a new evolutionary computation technique inspired by birds' flight and communication behaviors. The proposed algorithm determines the length of each interval in the universe of discourse and degree of membership values, simultaneously. Two numerical data sets are selected to illustrate the proposed method and compare the forecasting accuracy with four fuzzy time series methods. The results indicate that the proposed algorithm satisfactorily competes well with similar approaches.

1. Introduction

Forecasting has always been a crucial challenge for managers and scientists to arrive 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 variant and time-invariant. If the relations are only between time *t* and its prior time *t*-1, 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

* **Corresponding author S. Mohammad Taghi Fatemi Ghomi**
Email: *fatemi@aut.ac.ir*
Paper first received July. 05, 2012, and in revised form Oct. 9, 2012.

(TAIFEX) forecasting, based on the two-factor high order fuzzy 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.

Huarng 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 $U=\{u_1,u_2,...,u_m\}$and let A be a fuzzy set in the universe of discourse U defined as follows:

$$A= f_A(u_1)/u_1 + f_A(u_2)/u_2 +\ldots+ f_A(u_m)/u_m \quad (1)$$

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

Let X (t) (t = 0, 1, 2, …) a subset of a real number, be the universe of discourse on which fuzzy sets f_j(t), j=1,2,...,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,...).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 ○ is an Max–Min composition operator. Considering a fuzzy logical relationship $A_i\rightarrow A_j$, where $A_i=F(t-1)$ and $A_j= F(t)$, A_i 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 $F(t-1)$,$F(t-2)$, …, and $F(t-m)$, (m>0). In this forecasting method, the relation can be expressed as the fuzzy relational equation:

$$F(t)= F(t-1) \circ R^w(t-1, t) \quad (2)$$

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 $t(d_t)$ and t-1 (d_t-1) is computed as $V_t=d_t-d_{t-1}$ (t = 2, 3, ..., n).

Step 2: Define the universe of discourse. Find the maximum (D_{max}) and the minimum (D_{min}) among all V_t. The universe of discourse U is then defined as $U=[D_{\min}-D_1 \quad D_{\max}+D_2]$ 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_j, it belongs to fuzzy set A_j. 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 $F(t)$.

$$C(t)= F(t-1) = [C_1 C_2 \ldots C_m] \quad (3)$$

$$O^w(t) = \begin{bmatrix} F(t-2) \\ F(t-3) \\ \vdots \\ F(t-w) \end{bmatrix} = \begin{bmatrix} O_{11} O_{12} & \cdots & O_{1m} \\ O_{21} O_{22} & \cdots & O_{2m} \\ \vdots \quad \vdots & \vdots & \vdots \\ O_{(w-1)1} O_{(w-1)2} & \cdots & O_{(w-1)m} \end{bmatrix} \quad (4)$$

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

$$R(t) = O^w(t) \otimes C(t)$$
$$= \begin{bmatrix} O_{11}, C_1 O_{12}, C_2 \ldots O_{1m}, C_m \\ O_{21}, C_1 O_{22}, C_2 \quad \cdots \quad O_{2m}, C_m \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ O_{(w-1)1}, C_1 \, O_{(w-1)2}, C_2 \quad \cdots \quad O_{(w-1)m}, C_m \end{bmatrix} \quad (5)$$

$$= \begin{bmatrix} R_{11} R_{12} & \cdots & R_{1m} \\ R_{21} R_{22} & \cdots & R_{2m} \\ \vdots \quad \vdots & \vdots & \vdots \\ R_{(w-1)1}, R_{(w-1)2} & \cdots & R_{(w-1)m} \end{bmatrix}$$

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

$$F(t)=[max_{k=1,\ldots,w-1}\{R_{k1}\} \ \ldots \ max_{k=1,\ldots,w-1}\{R_{km}\}] \quad (6)$$

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

$$Cv_t = \frac{\sum_{i=1}^{m} f_i m_i}{\sum_{i=1}^{m} f_i} \quad (7)$$

The forecasted value Fv_t at time t is computed as follows:

$$Fv_t = Cv_t + d_{t-1} \quad (8)$$

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)).

$$V_{i,t+1} = w.V_{i,t} + R_1.C_1(P_i - X_{i,t}) + R_2.C_2(P_g - X_{i,t}) \quad (9)$$

$$X_{i,t+1} = X_{i,t} + V_{i,t+1} \quad (10)$$

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 inertial 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:

$$w = w_{\max}-((w_{\max}-w_{\min})/ \, iter_{\max}) \, iter \quad (11)$$

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

$I_1=(D_{min}, D_{min}+b_1]$,

$I_2=(D_{min}+b_1, D_{min}+b_1+b_2]$,

…

$I_{n-1}=(D_{min}+\sum_{i=1}^{m-2} b_i, D_{min}+\sum_{i=1}^{m-1} b_i]$

$I_n=(D_{min}+\sum_{i=1}^{m-1} b_i, D_{max}]$

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 bi 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.

X					Y	
0.4	0.7	…	0.3		0.6	0.8

Fig. 1. A sample of the 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:

$$MAD = \frac{\sum_{t=1}^{n} |Fv_t - d_t|}{n} \quad (12)$$

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.

Tab. 1. The historical data of enrollments at the University of Alabama

No.	Year	Enrollment
1	1971	13055
2	1972	13563
3	1973	13867
4	1974	14696
5	1975	15460
6	1976	15311
7	1977	15603
8	1978	15861
9	1979	16807
10	1980	16919
11	1981	16388
12	1982	15433
13	1983	15497
14	1984	15145
15	1985	15163
16	1986	15984
17	1987	16859
18	1988	18150
19	1989	18970
20	1990	19328
21	1991	19337

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_{max}=1291, D_{min}=-955, D_1=45, and D_2=8.

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

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]

Also, the degree of membership values of lower, normal, and upper intervals are 0.69and 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:

A_1= 1/ I_1 + 0.7/ I_2 + 0/ I_3 + 0/ I_4 + 0/ I_5 + 0/ I_6 + 0/ I_7 "not many"
A_2= 0.69/I_1+ 1/I_2+ 0.7/I_3 +0/I_4 +0/I_5+ 0/ I_6+ 0/ I_7 "not too many"
A_3= 0/ I_1 + 0.69/ I_2 + 1/ I_3 + 0.7/ I_4 + 0/ I_5 + 0/ I_6 + 0/ I_7 "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 "many 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 "very 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 "too many"
A_7= 0/I_1+ 0/I_2+0/I_3 +0/I_4 + 0/I_5 +0.69/I_6 +1/I_7 "too many many"

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_5, therefore, fuzzy set A_5 is assigned to this value. More details of fuzzification for this example are shown in Table 2.

Tab. 2. An instance of fuzzification for enrollment data

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

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]

$$O^3(1980) = \begin{bmatrix} F(1978) \\ F(1977) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0.69 & 1 & 0.7 & 0 & 0 \\ 0 & 0 & 0 & 0.69 & 1 & 0.7 & 0 & 0 \end{bmatrix}$$

So, R(1980) is computed as follows:

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

Based on Eq. (8), $F(1980)$ is
$F(1980) = [0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0.488\ \ 0\ \ 0]$
Cv_{1980} and Fv_{1980} are computed based on Eqs. (8) and (9) and the forecasted value in 1980 is obtained as follows:

$$Fv_{1980} = Cv_{1980} + d_{1979} = 407.50 + 16807 = 17214.5$$

It is necessary to analyze the objective value of different values of *m* and *w* and select the best results for comparing with the 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

m	*w* 2	3	4	5	6	7	8
5	419.3	390.2	390.72	396.91	417.16	422.61	357.42
6	402.2	383.9	379.03	394.66	414.81	421.94	353.01
7	406.7	378.9	385.06	387.76	410.29	424.08	**337.29**
8	407.5	369.7	375.87	381.77	406.13	412.34	352.67
9	403.7	383.9	395.01	376.41	403.77	418.53	**343.55**
10	410.3	373.8	356.42	374.46	405.23	412.50	**346.94**
11	400.9	374.1	375.12	370.74	405.45	417.11	**343.84**
12	401.1	377.3	378.70	371.52	404.86	415.63	**350.84**
13	411.5	369.9	360.19	373.49	396.69	412.78	**344.25**
14	394.4	373.2	387.56	361.07	403.67	413.10	**323.93**

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.

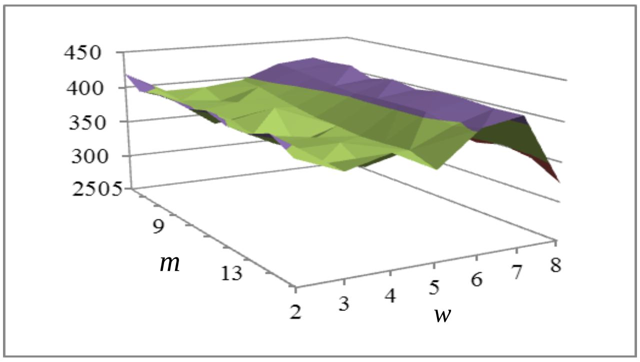

Fig. 2. Residual trends of *m* and *w* for enrollments of the University of Alabama

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.

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.

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

No.	Year	Actual Enrol.	Hwang *et al.* [19] *w*=2	Hwang *et al.* [19] *w*=3	Hwang *et al.* [19] *w*=4	Hwang *et al.* [19] *w*=5	Hwang *et al.* [19] *w*=6	Lee and Chou [10]	Liu *et al.* [16]	Prop. method *w*=2	Prop. method *w*=3	Prop. method *w*=4	Prop. method *w*=5	Prop. method *w*=6	Prop. method *w*=7	Prop. method *w*=8
1	1971	13055	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	1972	13563	-	-	-	-	-	14025	-	-	-	-	-	-	-	-
3	1973	13867	-	-	-	-	-	14568	-	-	-	-	-	-	-	-
4	1974	14696	14267	-	-	-	-	14568	-	14160	-	-	-	-	-	-
5	1975	15460	15296	15296	-	-	-	15654	15068	14696	15290	-	-	-	-	-
6	1976	15311	16260	16260	16260	-	-	15654	15807	15460	15859	15549	-	-	-	-
7	1977	15603	15711	15711	15511	15511	-	15654	15311	15480	15538	15400	15381	-	-	-
8	1978	15861	15803	16003	16003	16003	16003	15654	15739	15772	15680	15692	15820	15872	-	-
9	1979	16807	16261	16261	16261	16261	16261	16197	16018	16154	15894	15950	16078	16130	16178	-
10	1980	16919	17409	17407	17407	17607	17607	17283	17015	16807	17119	16807	16807	16807	17472	17534
11	1981	16388	17319	17119	17119	16919	16919	17283	16999	16919	16952	17008	16805	16884	16864	16668
12	1982	15433	16188	16188	16188	16188	16188	16197	16348	16388	16161	16234	16198	16175	16087	16137
13	1983	15497	14833	14833	14833	14833	14833	15654	15153	15433	15433	15433	15433	15433	15433	15433
14	1984	15145	15097	15297	15497	15497	15497	15654	15457	15497	15270	15544	15378	15429	15394	15246
15	1985	15163	14945	14745	14745	14745	14745	15654	15089	15145	15178	15192	15026	15051	15042	14894
16	1986	15984	14963	15163	15163	15163	15163	15654	15458	15163	15196	15252	15044	15069	15026	14912
17	1987	16859	16384	16384	16384	16384	16384	16197	15984	15984	16296	16386	15984	16168	16227	16711
18	1988	18150	17659	17659	17659	17659	17659	17283	17277	17509	17545	17522	17590	17283	17524	17586
19	1989	18970	19150	19150	19150	19150	19150	18369	19030	18941	18836	18911	18881	18855	18976	18877
20	1990	19328	19970	19770	19770	19770	19770	19454	19681	19471	19564	19372	19275	19424	19238	19410
21	1991	19337	19928	19928	19928	19928	19728,00	19454	19432	19328	19555	19417	19328	19560	19645	19328
22	1992	18876	19537	19537	19537	19537	19537	19454	19417	19506	19414	19426	19407	19421	19282	19086
m									8	14	8	10	14	13	8	14
MAD			496	488	517	491	504	436	431	394	370	356	361	397	412	324

Tab. 5. The historical data of sales volume of product

No.	Month	Products
1	January	26658
2	February	29216
3	March	30035
4	April	30846
5	May	29171
6	June	21068
7	July	28416
8	August	25651
9	September	22122
10	October	26423
11	November	22467
12	December	23274

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. Table6 shows the results of the proposed algorithm for different values of intervals (m) and window base (w).

Tab. 6. Residual values of *m* and *w* for sales volume of products

m	*w*				
	2	3	4	5	6
5	1938.26	2080.22	1792.17	1805.6	2946.5
6	1873.09	1638.29	2290.8	1989.6	2934.23
7	1865.92	1836.38	2008.83	1537.6	2621.58
8	1885.66	1482.57	2082.72	1965.6	2946.5
9	1559.04	2007.33	2008.83	1965.6	2584.92
10	1787.79	1908.29	2082.72	1983.6	2927.75
11	1559.04	1745.90	1352.17	2151.6	2621.58
12	1802.38	1846.86	1896.06	1727.6	2488.25
13	1717.38	1754.00	1871.61	1983.6	2094.92
14	1559.04	1545.90	1903.83	1405.6	2488.25

Also, the trends of different values of w are shown in Fig 3. Window base 2 and 3 have the best trends and window base 6 is the worst. Therefore, this example is less stable, and thus less window bases are more efficient. On the other words, the objective value (residuals) would increase by increasing the window base.

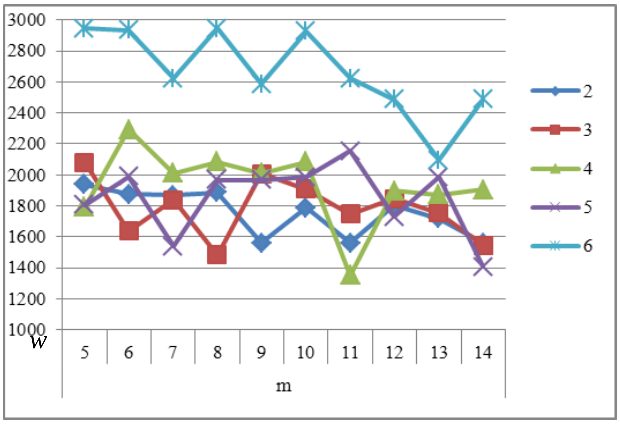

Fig. 3. Residual trends of *m* and *w* for sales volume of products

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.

Table 7. The comparison of proposed method to previous algorithms for sales of products of a manufacturing company

No.	Month	Actual data	Hwang *et al.* [19] *w*=2	Hwang *et al.* [19] *w*=3	Hwang *et al.* [19] *w*=4	Hwang *et al.* [19] *w*=5	Hwang *et al.* [19] *w*=6	Lee and Chou [10]	Liu *et al.* [16]	Prop. method *w*=2	Prop. method *w*=3	Prop. method *w*=4	Prop. method *w*=5	Prop. method *w*=6
1	Jan.	26658												
2	Feb.	29216						25959						
3	Mar.	30035						27221						
4	Apr.	30846	32235					27221	30561	30035				
5	May	29171	31746	31746				29745	31201	28739,3	27866			
6	Jun.	21068	28771	28771	28771			27221						
7	Jul.	28416	16768	16768	16768	16768		27851	29171	28331	26191	27211		
8	Aug.	25651	28416	28416	31916	31916	33216	25959	27528	26309,3	25436	25616	25496	
9	Sep.	22122	25651	21351	23951	23951	23951	25959	24967	23544,3	22671	22731	22731	23544,3
10	Oct.	26423	19122	19122	17822	17822	17822	25012	20226	22122	22122	22122	22122	22122
11	Nov.	22467	26423	27323	29923	29923	29923	25959	26423	26423	23443	23623	23623	24316,3
12	Dec.	23274	22467	18167	18167	18167	18167	25012	22467	22467	22467	22467	22467	22467
m										14	8	11	14	13
MAD			4630,3	5340,8	6944,1	6817,7	6111,6	2524,9	2344,0	1559,0	1482,6	1352,2	1405,6	2094,9

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

[1] Gooijer, J., Hyndman, R., "*Twenty Five Years of Time-Series Forecasting*," International Journal of Forecasting, Vol. 22, 2006, pp. 443–473.

[2] Behnamian, J., Fatemi Ghomi, S.M.T., "*Development of a PSO–SA Hybrid Metaheuristic for a New Comprehensive Regression Model to Time-Series Forecasting*," Expert Systems with Applications, Vol. 37, 2010, pp. 974–984.

[3] Liu, H.T., Wei, N.C., Yang, C.G., "*Improved Time-Variant Fuzzy Time Series Forecast,"* Fuzzy Optimization Decision Making, Vol. 8, 2009, pp. 45–65.

[4] Song, Q., Chissom, B.S., *"Fuzzy Time Series and its Models,*" Fuzzy Sets and Systems, Vol. 54, 1993, pp. 269–277.

[5] Chen, S.M., "*Forecasting Enrollments Based on Fuzzy Time Series*," Fuzzy Sets and Systems, Vol. 81, No. 3, 1996, pp. 311–319.

[6] Chen, S.M., "*Forecasting Enrollments Based on High-Order Fuzzy Time Series*," Cybernetics and Systems, Vol. 33, No. 1, 2002, pp. 1–16.

[7] Chen, S.M., Chung, N.Y., "*Forecasting Enrollments using High-Order Fuzzy Time Series and Genetic Algorithms*," International Journal of Intelligent Systems, Vol. 21, No. 5, 2006, pp. 485–501.

[8] Huarng, K.H., "*Heuristic Models of Fuzzy Time Series for Forecasting*," Fuzzy Sets and Systems, Vol. 123, 2001, pp. 369–386.

[9] Huarng, K.H., Yu, T.H.K., "*A Type 2 Fuzzy Time Series Model for Stock Index Forecasting*," Physica A: Statistical Mechanics and its Applications, Vol. 353, No. 1–4, 2005, pp. 445–462.

[10] Lee, H.S., Chou, M.T., *"Fuzzy Forecasting Based on Fuzzy Time Series,"* International Journal of Computer Mathematics, Vol. 81, No. 7, 2004, pp. 781–789.

[11] Lee, L.W., Wang, L.H., Chen, S.M., Leu, Y.H., *"Handling Forecasting Problems Based on Two-Factors Hgh-Order Fuzzy Time Series,"* IEEE Transactions on Fuzzy Systems, Vol. 14, No. 3, 2006, pp. 468–477.

[12] Liu, H.T., "*An Improved fuzzy Time Series Forecasting Method using Trapezoidal Fuzzy Numbers*," Fuzzy Optimization Decision Making, Vol. 6, 2007, pp. 63–80.

[13] Tsaur, R.C., Yang, J.C.O., Wang, H.F., *"Fuzzy Relation Analysis in Fuzzy Time Series Model*," Computer and Mathematics with Applications, Vol. 49, 2005, pp. 539–548.

[14] Yu, H.K., *"A Refined Fuzzy Time Series Model for Forecasting,"* Physica A, Vol. 346, 2005, PP. 657–681.

[15] Yu, H.K., "Weighted fuzzy time series model for TAIEX forecasting," Physica A, Vol. 349, 2005, PP. 609–624.

[16] Liu, H.T., *"An Integrated Fuzzy Time Series Forecasting System,"* Expert Systems with Applications, Vol. 36, 2009, pp. 10045-10053.

[17] Hsu, L.Y., Horng, S.H., Kao, T.W., Chen, Y.H., Run, R.S., Chen, R.J., Lai, J.L., & Kuo, I.H., "*Temperature Prediction and TAIFEX Forecasting Based on Fuzzy Relationships and MTPSO Techniques*," Expert Systems with Applications, Vol. 37, 2010, pp. 2756–2770.

[18] Song, Q., Chissom, B.S. "*Forecasting Enrollments with Fuzzy Time Series – Part II,"* Fuzzy Sets and Systems, Vol. 64, 1994, pp. 1–8.

[19] Hwang, J.R., Chen, S.M., & Lee, C.H., "*Handling Forecasting Problems using Fuzzy Time Series*," Fuzzy Sets and Systems, Vol. 100**,** 1998, pp. 217–228.

[20] Chen, S.M., Hwang, J.R., "*Temperature Prediction using Fuzzy Time Series*," IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics, Vol. 30**,** 2000, pp. 263–275.

[21] Lee, L.W., Wang, L.H., Chen, S.M., *"Temperature Prediction and TIFEX Forecasting Based on High-Order Fuzzy Logical Relationships and Genetic Simulated Annealing Techniques,"* Expert Systems with Applications, Vol. 34, 2008, pp. 328– 336.

[22] Kuo, I.H., Horng, S.J., Kao, T.W., Lin, T.L., Lee, C.L., Pan, Y., *"An Improved Method for Forecasting Enrollments Based on Fuzzy Time Series and Particle Swarm Optimization*," Expert Systems with Applications, Vol. 36, No. 3, 2009, pp. 6108-6117.

[23] Huarng, K.H., Yu, T.H.K., "*Ratio-Based Lengths of Intervals to Improve Fuzzy Time Series Forecasting*," IEEE Transactions on SMC – Part B: Cybernetics, Vol. 36, 2006, pp. 328–340.

[24] Liu H.T., Wei, N.C., *"An Improved Fuzzy Forecasting Method for Seasonal Time Series*." Expert Systems with Applications, Vol. 37, No. 9, 2010, pp. 6310-6318.

[25] Eberhart, R., Kennedy, J., "*A New Optimizer using Particle Swarm Theory*," In proceedings of the 6th international symposium on micro machine and human science, 1995, pp. 39–43.

[26] Eberhart, R.C., Kennedy, J.F., Swarm Intelligence, Morgan Kaufmann, 2002.

[27] Hu, X., Shi, Y. & Eberhart, R., "*Recent Advances in Particle Swarm*," In Proceedings of the IEEE Congress on Evolutionary Computation, 2004, pp. 90–97.