

A Novel Hybrid Fuzzy Time Series Approach with Applications to Enrollments and Car Road Accidents

Shehu Mohammed Yusuf
Ahmadu Bello University, Zaria
Department of Electrical and
Computer Engineering

M.B. Mu'azu
Ahmadu Bello University, Zaria
Department of Electrical and
Computer Engineering

O. Akinsanmi
Federal University of Oye, Ekiti
Department of Electrical and
Electronics Engineering

ABSTRACT

Fuzzy time series techniques are more suitable than traditional time series techniques in forecasting problems with linguistic values. Two shortcomings of existing fuzzy time series forecasting techniques are they lack persuasiveness in dealing with recurrent number of fuzzy relationships and assigning weights to elements of fuzzy rules in the defuzzification process. In this paper, a novel fuzzy time series technique based on fuzzy C-means clustering and particle swarm optimization is proposed to resolve these shortcomings. Fuzzy C-means clustering is adopted in the fuzzification process to objectively partition the universe of discourse. Then, particle swarm optimization is adopted to assign optimal weights to elements of fuzzy rules. Actual yearly enrollments at the University of Alabama and yearly deaths in car road accidents in Belgium are used as benchmark data. The forecasting results showed that the proposed method outperformed other existing methods.

General Terms

Fuzzy Systems, Pattern Recognition, Evolutionary Algorithms

Keywords

Forecasting, Fuzzy Time Series, Fuzzy C-means clustering, Particle Swarm Optimization.

1. INTRODUCTION

Fuzzy time series forecasting is the application of fuzzy mathematics to model and predict the future from a time series of linguistic historical observations. This technique of forecasting, compared with traditional time series forecasting, takes care of uncertainties in observations over time, does not require restrictive assumptions and too much background knowledge of the observations. Also, fuzzy time series forecasting techniques are able to work with very small set of observations [1].

The use of fuzzy time series in forecasting was first introduced in 1993 by Song and Chissom [2]. This technique comprises three processes: fuzzification, determining fuzzy relationships using matrix operation and defuzzification. Fuzzification is the process of converting crisp observations to linguistic values by identifying variations in the crisp data. In [2, 3], the fuzzification process implemented involved arbitrary defining the universe of discourse, partitioning the universe of discourse into equal lengths of intervals, defining fuzzy set sets on the universe of discourse, and fuzzification of historical observations. The contribution of [4], was the introduction of fuzzy logical relationship groups, centroid method to reduce computational overheads in the defuzzification process and, improving the performance of fuzzy time series forecasting. Huarng [5] showed that determination of length of interval is critical for improving forecasting performance in fuzzy time series. Mean and

distribution approach in determining the number of intervals was proposed. In [6], a ratio based interval length was presented and the forecasting results showed that unequal length of intervals improved fuzzy time series forecasting performance. Yu [7] pointed out the importance of assigning weights to resolve issue of recurrent fuzzy relationship and to reflect the difference in their importance. Cheng et al suggested in [8], the adoption of fuzzy C-means clustering to objectively partition the universe of discourse into unequal intervals, define memberships that best explains the unknown structure of the observations and, eliminate the need to define universe of discourse, in fuzzy time series forecasting. Models in [9 – 18] proposed optimal approaches in searching for appropriate unequal length of intervals using evolutionary algorithms. Aladag *et al* applied in [19], fuzzy C-means clustering in the fuzzification process to objectively partition the universe of discourse and define memberships in the fuzzy relational matrix. Particle swarm optimization was then used to optimize those memberships. Eleruja et al [20] adopted the concept of fuzzy set groups to establish fuzzy relationships and optimized the defuzzification process using particle swarm optimization. Uslu *et al* [1] presented a novel fuzzy time series technique that utilized genetic algorithm in the fuzzification process and assigned weights to recurrent fuzzy relationships to improve forecasting performance.

In this study, fuzzy C-means clustering was used, in the fuzzification process, to determine valid and useful unequal length of intervals for the universe of discourse, calculate observations' degree of memberships and, eliminate the need to define the universe of discourse. Then, fuzzy set group was introduced to eliminate recurrent fuzzy relations. Finally, a modified method for getting defuzzified forecasts using particle swarm optimization was proposed.

The rest of this paper is outlined as follows: A brief discussion of fuzzy C-means clustering and particle swarm optimization are described in Section 2. Section 3 discusses fuzzy time series In Section 4, the proposed forecasting model is discussed. Section 5 discusses the results obtained from the application of the proposed model to two sets of observations. Finally, Section 6 presents the conclusion.

2. THEORITICAL BACKGROUND

In this section, fuzzy C-means clustering, particle swarm optimization and fuzzy time series are briefly reviewed.

2.1 Fuzzy C-means Clustering

Fuzzy C-means clustering was first proposed by Dunn [21]. This clustering technique was later generalized by Bezdek [22]. It has been successful in partitioning and fuzzifying observations by minimizing the cluster error objective function, given in equation (1), with respect to fuzzy memberships, μ_{it} , and cluster centers, V_i :

$$J_\beta(X, U, V) = \sum_{i=1}^c \sum_{t=1}^n \mu_{it}^\beta d^2(x_i; v_i) = \sum_{i=1}^c \sum_{t=1}^n \mu_{it}^\beta \times \|x_i - v_i\| \quad (1)$$

Where $d^2(x_i; v_i)$ is the Euclidean distance measure, n is the number of observations, x_i , $\beta > 1$ is the weighting exponent used to tune out noise in the observation set and $c \geq 2$ is the number of partitions in the observation set.

An iterative minimization algorithm is used to minimize J_β under the following constraints [8]:

$$0 \leq \mu_{it} \leq 1, \quad \forall i, t \quad (2)$$

$$0 < \sum_{t=1}^n \mu_{it} \leq n, \quad \forall i \quad (3)$$

$$\sum_{i=1}^c \mu_{it} = 1, \quad \forall t \quad (4)$$

In each iteration, the component of v_i and x_i are updated using the following equations [19]:

$$v_i = \frac{\sum_{t=1}^n \mu_{it}^\beta x_i}{\sum_{t=1}^n \mu_{it}^\beta} \quad (5)$$

$$\mu_{it} = \frac{1}{\sum_{k=1}^c \left(\frac{d(x_i; v_i)}{d(x_i; v_k)} \right)^{\frac{2}{\beta-1}}} \quad (6)$$

The main fuzzy C-means clustering parameters to be set are; number of clusters, weighting exponents, and maximum iteration. The maximum iteration and precision are known as stopping criteria. Cheng et al [8] proposed an algorithm for fuzzy time series forecasting based on fuzzy C-means clustering. The first step of the algorithm utilizes fuzzy C-means clustering to identify unknown unequal partitions, and calculates memberships with respect to the measured partitions (cluster centers). The next step rank cluster centers in ascending order. Then, the ranks are used to define the clusters as ordered linguistic variables. When compared to Chen's model [4], each cluster center represents the mid-point of a crisp interval. Thus, cluster centers represent unequal partitions while membership values represent the degree of closeness of each observation to these mid-points.

2.2 Particle Swarm Optimization

This is a stochastic optimization technique that is robust and fast in optimizing continuous non-linear functions. It was first introduced in 1995 by Kennedy and Eberhart [24]. Particle swarm optimization is a population – based evolutionary algorithm that mimics the behavior of birds flocking or fish grouping in search for the location of food [20].

In particle swarm optimization algorithm, each solution is like a bird in the search space called 'particle' [25]. During particle swarm optimization implementation, a set of randomly generated particles, n , which are candidate solutions, are used to randomly initialize the process. Then, an iterative search process is set in motion to improve the set of current solutions. The moving particle adjusts its velocity during any iteration according to:

$$V_{i,t+1} = w \times V_{i,t} + C_1 \times Rand_1 \times (X_p - X_{i,t}) + C_2 \times Rand_2 \times (X_g - X_{i,t}) \quad (7)$$

Where $V_{i,t}$ is the velocity of particle i at t^{th} iteration and w is the inertial weight factor used to balance between local and global exploitations of the swarms. X_p , X_g and $X_{i,t}$ are particle's personal best, global best and current position, respectively. C_1 and C_2 are cognition and social learning factors, respectively, that control the influence of X_p and X_g on the search process. $Rand_1$ and $Rand_2$ are two randomly generated numbers set within the range of [0, 1].

Also, new particle position is adjusted by adding the new velocity to the current position as represented in:

$$X_{i,t+1} = X_i + V_{i,t+1} \quad (8)$$

The performance of each particle is measured according to the fitness function, J , to be solved. A fitness value is associated to each particle to be optimized, and the movement of each particle is directed by a randomly generated velocity parameter.

Fuzzy set groups (FSGs) and particle swarm optimization in the defuzzification process was adopted in [20], to obtain unique sets forecasts. Fuzzy set groups was adopted against fuzzy logical relationships to deal with recurrent fuzzy relations. Then, forecasting IF-THEN rules were deduced from the generated fuzzy set groups. In order to match forecasts with historical data they represent, optimal weights were assigned to forecasting rules using particles swarm optimization.

The main particle swarm optimization parameters applied to produce optimal solutions are; swarm size, maximum iteration, learning factors, particles initial positions, inertial weight factor, target fitness value.

In terms of ease of implementation, processing time and quality of solution in optimizing non – linear functions, particle swarm optimization (PSO) generally performs better than genetic algorithm (GA), memetic algorithm (MA) and ant colony optimization (ACO) [27].

2.3 Fuzzy Time Series

Fuzzy time series forecasting model is a mathematical forecasting model that assumes information needed to generate forecasts is contained in time series of linguistic historical data. Fuzzy time series was first developed by Song and Chissom [2], from the notion of crisp time series, to deal with incomplete and vague series of data by applying fuzzy logic. Fuzzy time series differ from traditional time series in that observations are not numeric values but linguistic values [27]. Its ability to work with a very small set of observations makes fuzzy time series forecasting outstanding [1].

Definition 1 Let $f_i(t) (i=1,2,3,...)$ be defined on $Y(t) (t=...0,1,2,...)$, the universe of discourse, which is a subset of real numbers. If $F(t)$ is a collection of $f_i(t) (i=1,2,3,...)$, then $F(t)$ is called a fuzzy time series on $Y(t) (t=...0,1,2,...)$.

Definition 2 If there exists a fuzzy relationship, $R(t-1, t)$, such that $F(t) = F(t-1) \circ R(t-1, t)$ where; \circ represents an operator, then $F(t)$ is said to be caused by $F(t-1)$. If $F(t-1) = A_i$ and $F(t) = A_j$, then the fuzzy logical relationship

between $F(t)$ and $F(t-1)$ can be denoted as $A_i \rightarrow A_j$ where A_i is the left – hand side which refers to the current state and, A_j is the right – hand side which refers to the next state of the fuzzy logical relationship. Fuzzy logical relationships can be grouped to establish different relationships.

Definition 3 Let $F(t)$ be a fuzzy time series. If $F(t)$ is caused by $F(t-1), F(t-2), \dots, F(t-n)$, then this fuzzy relationship is represented by $F(t-n), \dots, F(t-2), F(t-1) \rightarrow F(t)$. This is called an n – order fuzzy time series.

Definition 4 Let $F(t-2) = A_{i,t-2}$, $F(t-1) = A_{i,t-1}$ and $F(t) = A_j$. If $A_{i,t-2}, A_{i,t-1} \rightarrow A_j$, then $\{A_{i,t-2}, A_{i,t-1}\}$ is called a fuzzy set group (FSG) where $A_{i,t-2}$ and $A_{i,t-1}$ are two consecutive current states. A fuzzy set group (FSG) is originally a second order fuzzy relation. But, if a second order fuzzy set group reoccurs, the recurrent fuzzy set groups (FSGs) are extended to the third order, $\{A_{i,t-3}, A_{i,t-2}, A_{i,t-1}\}$, to generate unique set of fuzzy relations. The extension process continues until there is no recurrence of fuzzy relation.

3. PROPOSED METHOD

Recently, due to the need for improving forecasting performance, researchers have contributed to all three processes of fuzzy time series model. However, few have contributed to the defuzzification process; especially overcoming recurrent fuzzy relations and optimizing the defuzzification process. To improve on the defuzzification process, fuzzy set groups (FSGs) can be used to obtain unique set of fuzzy relations. These fuzzy relations can easily be converted to IF – THEN rules. Then, particle swarm optimization (PSO) can be utilized to assign optimal weights to elements in each forecasting rule. These weights can be used in defuzzification to obtain forecasts that match the historical observations they represent [20]. In this study, a new fuzzy time series model that uses fuzzy C-means (FCM) clustering in the fuzzification process, takes into consideration recurrent fuzzy relations and, optimize the defuzzification process using particle swarm optimization (PSO) has been proposed to improve forecasting performance.

The main advantages of the proposed model are as follows:

- There is no need to determine the universe of discourse because of utilizing fuzzy C-means clustering algorithm.
- Also, unequal interval lengths and memberships that best explain hidden structures in historical observations are objectively determined using fuzzy C-means clustering.
- Recurrent fuzzy relations are dealt with using fuzzy set groups (FSGs) to provide unique forecasts.
- Weights assigned to elements of a forecasting rule are searched using particle swarm optimization to obtain more realistic forecasts. Thus, the forecasting performance of fuzzy time series can be improved.

3.1 Algorithm

Step 1: Define the number of intervals and the parameters fuzzy C-means clustering algorithm.

After the number of intervals, c , is determined, the parameters such as the number of cluster centers, fuzzy index, precision and maximum number of iteration are decided for the fuzzy C-means clustering algorithm. The number of cluster is set to the same value as the number of intervals.

Step 2: Compute interval lengths and membership degrees using fuzzy C-means clustering algorithm.

When fuzzy C-means clustering is applied, the cluster centers v_i , are calculated iteratively. After cluster centers have been computed, the membership values, μ_{it} , of each observation, with respect to the cluster centers are computed. Then, the objective function J_{β} , is computed and compared with the set precision, ε . If the objective function is less than the set precision, the computed cluster centers are the set of unequal partitions for the historical observations otherwise the cluster centers are recomputed.

Step 3: Obtain fuzzy sets by ranking clusters in ascending order.

Cluster centers (partitions) produced by fuzzy C-means clustering algorithm are not orderly arranged like in the case of traditional fuzzy time series partitions. In the literature, fuzzy sets are only defined on ordered partitions. Thus, there is the need to rank (order) cluster centers. In this work, cluster centers are arranged in ascending order to define fuzzy sets, $A_r (r = 1, 2, 3, \dots, c)$.

Step 4: Fuzzify historical data.

Crisp historical observations are converted to their linguistic values using the defined fuzzy sets and membership degrees of each observation computed. A linguistic variable among $A_r (r = 1, 2, 3, \dots, c)$ is the linguistic value of an observation if in that fuzzy set, the observation has its highest degree.

Step 5 Establish unique FSGs.

For instance, when we observe the second order relation such as $F(t-2) = A_{i,t-2}$, $F(t-1) = A_{i,t-1}$ and $F(t) = A_j$ for any time t , this fuzzy relation is represented by $A_{i,t-2}, A_{i,t-1} \rightarrow A_j$. The relation is expressed in FSG by collecting the current states in braces as $\{A_{i,t-2}, A_{i,t-1}\}$. Also, when an FSG appears more than once, the set repeated FSGs are extended to the third order to make each group unique. The order of repeated FSGs is continuously incremented by one until every FSG is unique.

Step 6: Extract IF – THEN rules from FSGs.

These rules are generated on the basis of the content of the fuzzy set groups. The IF-THEN rule is of the form:

$$if \left(\begin{array}{l} F(t-1) = A_{r,t-1} \wedge F(t-2) = A_{r,t-2} \wedge \dots \wedge F(t-n+1) = A_{r,t-n+1} \\ \wedge F(t-n) = A_{r,t-n} \end{array} \right)$$

$$then \quad w_{t-1} = ? \wedge w_{t-2} = ? \wedge \dots \wedge w_{t-n+1} = ? \wedge w_{t-n} = ? \quad (9)$$

w_{t-n} is the weight of the previous historical observation at time $(t-n)$, t is the time (period) the future observation whose forecast is required, $n=t$ is the time (period) of the previous historical data point matched in a forecasting rule.

The weight w_{t-n} , represents the strength of fuzzy logical relationship between the previous historical data at n and future forecast at t .

Step 7: Train IF – THEN rules using PSO algorithm.

After deciding on the parameters of PSO algorithm, weights w_{t-n} , are initially assigned to represent the contribution of individual elements in a forecasting rule such that the more recent current state in the forecasting rule are assigned much larger weights. Then, optimal weights are generated for the content of fuzzy set groups (FSGs), by training the IF-THEN rules using particle swarm optimization. Each weight represents the strength of relationship between a past value and a future forecast. Generating optimal weights for forecasting rules will reduce mismatch between the forecasts and actual historical observations. When particle swarm optimization algorithm is applied on an IF-THEN rule, initial particle positions, $X_{n,i}$ and velocities, $V_{i,t}$, are set. Next, the initial squared error, J_n , is computed as a fitness value for each particle using equation (10) and the minimum among the computed fitness values is set as an intermediate squared error, J_{int} .

$$J = (Y(t) - A(t))^2 \quad (10)$$

Here, $Y(t)$ is the forecast and $A(t)$ is the actual observation.

The forecast $Y(t)$ is obtained as follows:

$$Y(t) = \sum_{i=1}^n a_{t-i} \times X_i \quad (11)$$

Here, a_{t-i} is the cluster center (defuzzified) value of a linguistic historical observation in the forecasting rule. X_i is the corresponding weight assigned to the linguistic historical observation in the forecasting rule. Next, a loop counter k , is set and the current velocities, $V_{i,t+1}$, is updated using equation (7). Then, the current particles positions, $X_{i,t+1}$, are updated using equation (8). Next, the current squared error, J_{nc} , for each particle is computed.

If the minimum of the computed current squared error is less than or equal to the intermediate squared error and target fitness value, the global best particle's position is obtained and taken as the optimal weights assigned to the forecasting rule. The loop counter is incremented as much as the maximum iteration number set and the process of updating velocities and position are repeated, if the current minimum squared error computed is greater than the target fitness value.

Step 8: Derive forecasts.

After particle swarm optimization is applied in the defuzzification process, completely generated IF-THEN forecasting rules are obtained. By utilizing optimal weights obtained and equation (11) is used to compute forecasts.

4. APPLICATION

In order to verify the performance of the proposed hybrid forecasting technique, this is applied to two different time series observations: enrollments at the University of Alabama and yearly deaths in car road accidents in Belgium [28]. The obtained results are compared with the results obtained from other fuzzy time series techniques in the literature using Root

Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) criteria as shown in:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (Y(t) - A(t))^2}{n}} \quad (12)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|Y(t) - A(t)|}{A(t)} \times 100 \quad (13)$$

For each time series observations, the decided FCM clustering parameters are as follows:

- Maximum number of iterations as 100 [23].
- Precision $\varepsilon = 0.00001$ [23].
- $\beta = 2$ [6].
- c is experimented as from 7 to 20 [1].

A suitable value of c is selected by evaluating the RMSE value between historical observations and their defuzzified (cluster center) values. After experimentation, the C value with the smallest RMSE value was taken as the best solution among these solutions.

Also, for each time series observation, the PSO parameters are defined as follows [20]:

- Swarm size, $n = 5$.
- Maximum number of iterations, $k_{max} = 500$.
- Target fitness value as minimum squared error of 1.
- Minimum and maximum particles' position is limited to $[0, 1]$.
- Minimum and maximum velocity should be in the range $[-0.01, 0.01]$.
- Learning factors, C_1 and C_2 , is both equal to 2.
- Inertial coefficient, w as 1.4.

For all possible case, FCM clustering and PSO are executed in C#.

4.1 Forecasting enrollments at University of Alabama

The proposed model is first implemented on the time series observations of students at the University of Alabama. The best result was obtained in the case where $c = 13$.

Table 1 shows the cluster centers and their corresponding linguistic values obtained by ranking in ascending order.

Table 2 presents the generated fuzzy set groups and optimal weights assigned to the content of the forecasting rules for the enrollment observation set.

A comparative presentation of enrollments' forecasts and the RMSE and MAPE values for the proposed method and some other methods is given in Table 3.

Table 1. Cluster Centers and Defined Fuzzy Sets for Enrollments

Cluster	Center	Fuzzy Set	Cluster	Center	Fuzzy Set
13	13079	A ₁	2	16388	A ₈
9	13750	A ₂	5	16862	A ₉
4	14695	A ₃	1	18150	A ₁₀
7	15155	A ₄	3	18876	A ₁₁
12	15320	A ₅	8	18970	A ₁₂
6	15493	A ₆	10	19332	A ₁₃
11	15921	A ₇			

As seen in Table 3, it is obvious that the proposed technique produces a superior forecasting performance. Figure 1 shows the plot the forecasts and real observations against time.

As seen in Figure 1, there is a good matching between the actual and forecasted enrollment observations.

4.2 Forecasting Car Road Accident in Belgium

The proposed technique is also implemented on the time series observations of number of those killed in car road accident in Belgium. The best result was obtained in the case where $c = 14$. Table 4 presents the cluster centers and their corresponding linguistic values obtained by ranking in ascending order. The generated fuzzy set groups and optimal weights assigned to the content of the forecasting rules, for the observation set is shown in Table 5. Additionally, a comparative presentation of enrollments' forecasts and the RMSE and MAPE values for the proposed method and some other methods is given in Table 6. As seen in Table 6, the proposed technique produces a superior result with respect to the forecasting performance.

Table 2. Generated Fuzzy Set Groups and Optimal Weights for Enrollments

Year	Enrollment	Linguistic	FSG	Optimal Weights	Year	Enrollment	Linguistic	FSG	Optimal Weights
1971	13055	A ₁			1982	15433	A ₆	{A ₉ , A ₈ }	{0.341, 0.591}
1972	13563	A ₂			1983	15497	A ₆	{A ₈ , A ₆ }	{0.363, 0.617}
1973	13876	A ₂	{A ₁ , A ₂ }	{0.390, 0.638}	1984	15145	A ₄	{A ₆ , A ₆ }	{0.364, 0.613}
1974	14696	A ₃	{A ₂ , A ₂ }	{0.411, 0.658}	1985	15163	A ₄	{A ₆ , A ₄ }	{0.371, 0.622}
1975	15460	A ₆	{A ₂ , A ₃ }	{0.413, 0.666}	1986	15984	A ₇	{A ₄ , A ₄ }	{0.398, 0.657}
1976	15311	A ₅	{A ₃ , A ₆ }	{0.382, 0.625}	1987	16859	A ₉	{A ₄ , A ₇ }	{0.415, 0.664}
1977	15603	A ₆	{A ₆ , A ₅ }	{0.383, 0.631}	1988	18150	A ₁₀	{A ₄ , A ₇ , A ₉ }	{0.084, 0.382, 0.640}
1978	15861	A ₇	{A ₅ , A ₆ }	{0.386, 0.642}	1989	18970	A ₁₂	{A ₉ , A ₁₀ }	{0.413, 0.662}
1979	16807	A ₉	{A ₆ , A ₇ }	{0.407, 0.660}	1990	19328	A ₁₃	{A ₁₀ , A ₁₂ }	{0.393, 0.643}
1980	16919	A ₉	{A ₆ , A ₇ , A ₉ }	{0.061, 0.358, 0.609}	1991	19337	A ₁₃	{A ₁₂ , A ₁₃ }	{0.382, 0.626}
1981	16388	A ₈	{A ₉ , A ₉ }	{0.362, 0.610}	1992	18876	A ₁₁	{A ₁₃ , A ₁₃ }	{0.363, 0.613}

Table 3. A Comparative Presentation of Enrollments' Forecasts

Year	Actual enrollment	[2]	[4]	[5]	[6]	[1]	Proposed Model
1971	13055						
1972	13563	14000	14000	14000	14242	13650	
1973	13867	14000	14000	14000	14242	13650	13873.3
1974	14696	14000	14000	14000	14242	14836	14685
1975	15460	15500	15500	15500	15474.3	15332	15465.6
1976	15311	16000	16000	15500	15474.3	15447	15312.1
1977	15603	16000	16000	16000	15474.3	15447	15600.7
1978	15861	16000	16000	16000	15474.3	15447	15860
1979	16807	16000	16000	16000	16146.5	16746	16813.5
1980	16919	16813	16833	17500	16988.3	17075	16913.7
1981	16388	16813	16833	16000	16988.3	16380	16389.9
1982	15433	16789	16833	16000	16146.5	15457	15435.3
1983	15497	16000	16000	16000	15474.3	15447	15508
1984	15145	16000	16000	15500	15474.3	15447	15136.7
1985	15163	16000	16000	16000	15474.3	15332	15174.3
1986	15984	16000	16000	16000	15474.3	16027	15988.5
1987	16859	16000	16000	16000	16146.5	16746	16860.9
1988	18150	16813	16833	17500	16988.3	18211	18146.5
1989	18970	19000	19000	19000	19144	19059	18979.3
1990	19328	19000	19000	19000	19144	19059	19330.7
1991	19337	19000	19000	19500	19144	19059	19348.4
1992	18876		19000	19000	19144	19059	18887.4
	RMSE	650	638	476	478	178	7.02
	MAPE	3.22%	3.11%	2.45%	2.20%	0.90%	0.04%

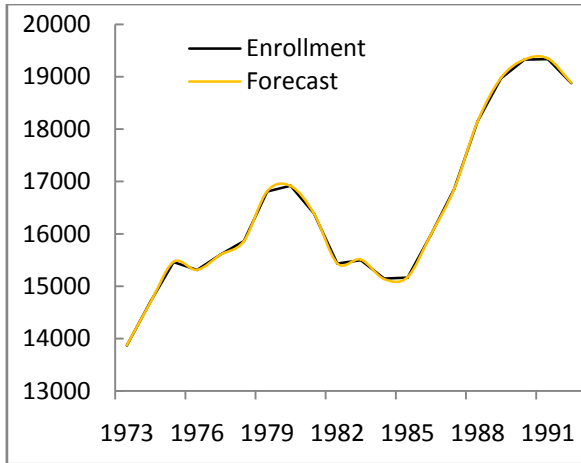


Figure 1: Actual and Forecasted Enrollment vs. Time

Table 4. Cluster Centers and Defined Fuzzy Sets for Yearly Deaths from Road Accidents in Belgium

Cluster	Center	Fuzzy Set	Cluster	Center	Fuzzy Set
2	953	A ₁	5	1306	A ₈
1	1035	A ₂	7	1376	A ₉
13	1122	A ₃	6	1422	A ₁₀
12	1148	A ₄	11	1469	A ₁₁
8	1173	A ₅	9	1573	A ₁₂
3	1226	A ₆	10	1629	A ₁₃
14	1258	A ₇			

Figure 2 shows the plot the forecasts and real observations against time. The graph also shows a good match between the forecasts and actual observations

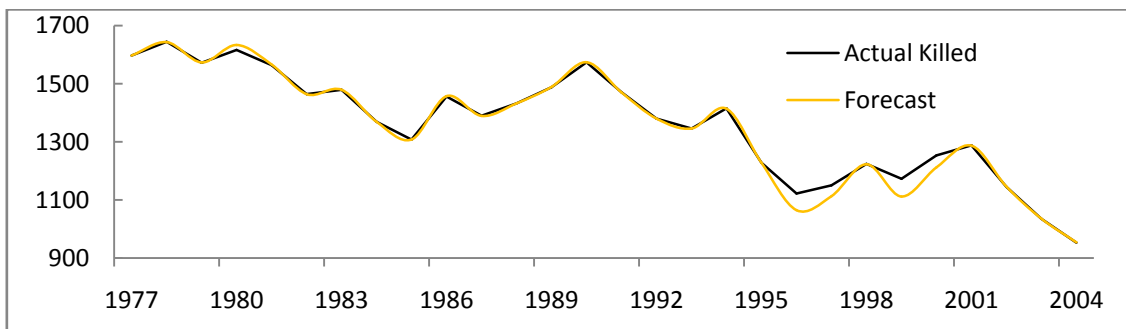


Figure 2: Graph of Forecasts of the Proposed Method and Actual Observations of YearlyDeaths from Accidents in Belgium

Table 5. Generated Fuzzy Set Groups and Optimal Weights for Yearly Deaths from Accidents in Belgium

Year	Actual	Ling.	FSG	Optimal Weights	Year	Actual	Ling.	FSG	Optimal Weights
1974	1574	A ₁₃			1990	1574	A ₁₃	{A ₁₀ , A ₁₁ }	{0.419, 0.666}
1975	1460	A ₁₁			1991	1471	A ₁₁	{A ₁₁ , A ₁₃ }	{0.353, 0.605}
1976	1536	A ₁₂			1992	1380	A ₉	{A ₁₁ , A ₁₃ , A ₁₁ }	{0.024, 0.321, 0.572}
1977	1597	A ₁₃	{A ₁₁ , A ₁₂ }	{0.401, 0.654}	1993	1346	A ₉	{A ₁₃ , A ₁₁ , A ₉ }	{0.035, 0.333, 0.583}
1978	1644	A ₁₄	{A ₁₂ , A ₁₃ }	{0.402, 0.652}	1994	1415	A ₁₀	{A ₉ , A ₉ }	{0.392, 0.636}
1979	1572	A ₁₃	{A ₁₂ , A ₁₃ , A ₁₄ }	{0.047, 0.342, 0.591}	1995	1228	A ₆	{A ₉ , A ₉ , A ₁₀ }	{0.000, 0.320, 0.554}
1980	1616	A ₁₄	{A ₁₂ , A ₁₃ , A ₁₄ , A ₁₃ }	{0.011, 0.051, 0.353, 0.611}	1996	1122	A ₃	{A ₁₀ , A ₆ }	{0.309, 0.557}
1981	1564	A ₁₃	{A ₁₄ , A ₁₃ , A ₁₄ }	{0.041, 0.341, 0.591}	1997	1150	A ₄	{A ₆ , A ₃ }	{0.371, 0.620}
1982	1464	A ₁₁	{A ₁₄ , A ₁₃ , A ₁₄ , A ₁₃ }	{0.001, 0.026, 0.321, 0.571}	1998	1224	A ₆	{A ₃ , A ₄ }	{0.414, 0.661}
1983	1479	A ₁₁	{A ₁₄ , A ₁₃ , A ₁₁ }	{0.041, 0.341, 0.596}	1999	1173	A ₅	{A ₄ , A ₆ }	{0.358, 0.622}
1984	1369	A ₉	{A ₁₁ , A ₁₁ }	{0.339, 0.593}	2000	1253	A ₇	{A ₆ , A ₅ }	{0.399, 0.653}
1985	1308	A ₈	{A ₁₁ , A ₁₁ , A ₉ }	{0.028, 0.323, 0.576}	2001	1288	A ₈	{A ₅ , A ₇ }	{0.402, 0.648}
1986	1456	A ₁₁	{A ₉ , A ₈ }	{0.419, 0.674}	2002	1145	A ₄	{A ₇ , A ₈ }	{0.320, 0.569}
1987	1390	A ₉	{A ₈ , A ₁₁ }	{0.371, 0.616}	2003	1035	A ₂	{A ₈ , A ₄ }	{0.307, 0.553}
1988	1432	A ₁₀	{A ₈ , A ₁₁ , A ₉ }	{0.053, 0.356, 0.610}	2004	935	A ₁	{A ₄ , A ₂ }	{0.319, 0.568}
1989	1488	A ₁₁	{A ₁₁ , A ₉ , A ₁₀ }	{0.067, 0.372, 0.618}					

Table 6. A Comparative Presentation of Yearly Deaths from Car Accidents' Forecasts

Year	Actual	[11]	[28]	[29]	[1]	Proposed Model
1974	1574		1497			
1975	1460		1497		1506	
1976	1536		1497		1453	
1977	1597	1500	1497	1497	1598	1597
1978	1644	1500	1497	1497	1584	1643
1979	1572	1500	1497	1497	1584	1573
1980	1616	1500	1497	1598	1506	1633
1981	1564	1500	1497	1598	1584	1566
1982	1464	1500	1497	1498	1506	1464
1983	1479	1500	1497	1498	1453	1479
1984	1369	1500	1497	1398	1375	1369
1985	1308	1400	1396	1298	1383	1308
1986	1456	1300	1296	1498	1454	1457
1987	1390	1500	1497	1398	1453	1389
1988	1432	1400	1396	1398	1383	1432
1989	1488	1400	1396	1498	1509	1489
1990	1574	1500	1497	1598	1598	1574
1991	1471	1500	1497	1498	1506	1470
1992	1380	1500	1497	1398	1375	1380
1993	1346	1400	1396	1298	1383	1346
1994	1415	1300	1296	1398	1383	1414
1995	1228	1400	1396	1198	1231	1228
1996	1122	1100	1095	1098	1135	1065
1997	1150	1200	1196	1198	1180	1113
1998	1224	1200	1196	1198	1245	1223
1999	1173	1200	1196	1198	1135	1112
2000	1253	1300	1296	1298	1245	1212
2001	1288	1300	1296	1298	1284	1287
2002	1145	1100	1095	1098	1143	1146
2003	1035	1000	995	997	970	1036
2004	953	1000	995	997	970	954
	RMSE	85.35	83.12	46.78	41.61	19.2
	MAPE	5.25%	5.06%	2.70%	2.29%	0.67%

5. CONCLUSION

It has been observed that recurrence of fuzzy relation and assigning weights to elements of a fuzzy rule affects forecasting performance of fuzzy time series.

In this study, a novel fuzzy time series technique which considers the number of recurrences in a second order fuzzy logic relationship table and assigns optimal weights to elements of a fuzzy rule has been proposed. Fuzzy C-means clustering is integrated in the fuzzification process to partition observation sets objectively while fuzzy set group is adopted in the defuzzification process to generate unique set of fuzzy relations. Also, particle swarm optimization is integrated in the defuzzification process to assign optimal weights to elements of a fuzzy rule. Experimental results on the enrollments and 'killed in car road accident' observations demonstrate that the proposed forecasting technique provides more accurate forecasts.

In the future, there is the need for more applications of the proposed model to validate the general applicability. And, also different techniques can be introduced to deal with recurrent number of fuzzy relations. Other clustering techniques and evolutionary algorithms may be used in the fuzzification process.

6. REFERENCES

- [1] Uslu V. R, Bas E., Yolcu U., Egrioglu E. 2014 A fuzzy time series approach based on weights determined by the recurrences of fuzzy relations. Swarm and Evolutionary Computation. 19-26.
- [2] Song Q., Chissom B. S. 1993 Forecasting enrollments with fuzzy time series—Part I. Fuzzy Sets Syst. 54 pp. 1–9.
- [3] Song Q., Chissom B. S. 1993 Fuzzy time series and its models. Fuzzy Sets Syst. 54 pp. 269–277.
- [4] S.M. Chen, Forecasting enrollments based on fuzzy time-series. Fuzzy Sets Syst. 81 pp. 311–319.
- [5] Huarng K. 2001 Effective length of intervals to improve forecasting in fuzzy time- series. Fuzzy Sets Syst. 123 pp. 387–394.
- [6] Huarng K., Yu T. H. K. 2006 Ratio-based lengths of intervals to improve fuzzy time series forecasting. IEEE Trans. Syst. ManCybern. Part B: Cybern. 36 pp. 328–340.
- [7] Yu H. K. 2005 Weighted fuzzy time series models for TAIEX forecasting, Physica A 349 pp. 609–624.
- [8] Cheng C. H., Cheng G. W, Wang J. W. 2008 Multi-attribute fuzzy time series method based on fuzzy clustering. Expert Syst. Appl. 34 pp. 1235–1242.
- [9] Kuo I. H., Horng S. J., Kao T. W, Lin T. L., Lee C. L., Pan Y. 2009 An improved method of forecasting enrolments based on fuzzy time series and particle swarm optimization. Expert Syst. Appl. 36 pp. 6108–6117.

- [10] Chen S. M., Chung N. Y. 2006 Forecasting enrolments using high order fuzzy time series and genetic algorithms. *Int. J. Intell. Syst.* 21 pp. 485–501.
- [11] Egrioglu E., Aladag C. H., Yolcu U., Uslu V. R., Basaran M. A. 2010 Finding an optimal interval length in high order fuzzy time series. *Expert Syst. Appl.* 37 pp. 5052–5055.
- [12] Egrioglu E., Aladag C. H., Basaran M. A., Uslu V. R., Yolcu U. 2011 A new approach based on the optimization of the length of intervals in fuzzy time series. *J. Intell. Fuzzy Syst.* 22 pp.15–19.
- [13] Lee L. W., Wang L. H., Chen S. M., Leu Y. H. 2006 Handling forecasting problems based on two factor high-order fuzzy time series. *IEEE Trans. Fuzzy Syst.* 14 (3) pp. 468–477.
- [14] Kuo I. H., Horng S. J., Chen Y. H., Run R. S., Kao T. W., Chen R. J., Lai J. L., Lin T. L. 2010 Forecasting TAIEX based on fuzzy time series and particle swarm optimization. *Expert Syst. Appl.* 37 pp. 1494–1502.
- [15] Kuo I. H., Horng S.J., Kao T. W., Lin T. L., Lee C. L., Pan Y. 2009 An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization. *Expert Syst. Appl.* 36 pp. 6108–6117.
- [16] Davari S., Zarandi M. H. F., Turksen I.B. 2009 An Improved fuzzy time series forecasting model based on particle swarm intervalization. The 28th North American Fuzzy Information Processing Society Annual Conferences, NAFIPS 2009, Cincinnati, Ohio, USA.
- [17] Park J. I., Lee D. J., Song C. K., Chun M. G. 2010 TAIEX and KOSPI200 forecasting based on two factors high order fuzzy time series and particle swarm optimization. *Expert Syst. Appl.* 37 pp. 959–967.
- [18] Hsu L. Y., Horng S. J., Kao T. W., Chen Y. H., Run R. S., Chen R. J., Lai J. L., Kuo I. H. 2010 Temperature prediction and TAIEX forecasting based on fuzzy relationships and MTPSO techniques, *Expert Syst. Appl.* 37 pp. 2756–2770.
- [19] Aladag C. H., Yolcu U., Egrioglu E., Dalar A. Z. 2012 A new time invariant fuzzy time series forecasting method based on particle swarm optimization. *Appl. Soft Comput.* 12 pp. 3291–3299.
- [20] Eleruja S. A., Mu'azu M. B., Dajab D. D. 2012 Application of trapezoidal fuzzification approach (TFA) and particle swarm optimization (PSO) in fuzzy time series (FTS) forecasting. *Proceedings of ICAI*, 1 pp. 80–89.
- [21] Dunn J. C. 1974 A fuzzy relative of ISODATA process and its use in detecting compact well – separated clusters. *Cybernetics.* 3 pp. 32 – 57.
- [22] Bezdek J. C. 1981 *Pattern recognition with fuzzy objective function algorithms.* New York. Plenum.
- [23] Jafar O. A. M., Sivakumar R. 2013 A comparative study of hard and fuzzy data clustering algorithms with cluster validity indices. *Proceedings of International Conference on Emerging Research in Computing, Information, Communication and Applications*, 775 – 782.
- [24] Kennedy J., Eberhart R. 1995 Particle swarm optimization. *Proceedings of IEEE International Conference on Neural Network*, 1942 – 1948.
- [25] Mahnam M., Ghomi S. M. T. F. 2012 A particle swarm optimization algorithm for forecasting based on time variant fuzzy time series. *IJIEPR*, ISSN: 2008-4889, 23 pp.269-276.
- [26] Elbeltagi E., Hegazy T., Grierson D. 2005 Comparison among evolutionary – based optimization algorithms. *Advance Engineering Informatics*, 19 pp. 43 – 53.
- [27] Li S. T., Cheng Y. C., Lin S. Y. 2008 A FCM – based deterministic forecasting model for fuzzy time series. *Computer and Mathematics with Applications*, 56 pp. 3052 – 3063.
- [28] Jilani T. A., Burney S. M. A., Ardil C. 2007 Multivariate high order fuzzy time series forecasting for car road accident. *World Acad. Sci. Eng. Technol.* 25 pp. 288–293.