ARID ZONE JOURNAL OF ENGINEERING, TECHNOLOGY & ENVIRONMENT
AZOJETE, March, 2019. Vol. 15(1):40-54
Published by the Faculty of Engineering, University of Maiduguri, Maiduguri, Nigeria.
Print ISSN: 1596-2490, Electronic ISSN: 2545-5818
www.azojete.com.ng

ORIGINAL RESEARCH ARTICLE

CAT SWARM OPTIMIZATION BASED CLUSTERING ALGORITHM FOR FUZZY TIME SERIES FORECASTING

A. M. Abubakar and S. Man-Yahya

(Department of Electrical and Computer Engineering, Ahmadu Bello University Zaria, Kaduna State, Nigeria)
**Corresponding Author: muktyx5@yahoo.com*

ARTICLE INFORMATION

Submitted 31 March, 2018
Revised 11 August, 2018
Accepted 15 August, 2018

Keywords:
Cat Swarm Optimization
Clustering
fuzzy time series
forecasting

ABSTRACT
This paper presents the development of an improved Fuzzy Time Series (FTS) forecasting model with Cat Swarm Optimization based Clustering (CSO-C) algorithm. FTS forecasting is affected by some of the subjective decisions made at the fuzzification stage like size of interval length or universe of discourse, and approaches such as clustering, trend-mapping, have been adopted to address this. Traditional clustering techniques, such as fuzzy C-means (FCM) and K-means used to tackle the problems associated with the fuzzification stage encountered problems such as handling high dimensional data, sensitivity to noise and outliers and pre-mature convergence. In this paper, CSO-C algorithm was developed using MATLAB 2015a to address pre-mature convergence with a view to improving forecast accuracy in FTS forecasting. The developed model was tested on student enrolment of University of Alabama and Taiwan Future Exchange (TAIFEX). Also applied to forecast the student enrolment of University of Maiduguri (UniMaid) data. In all cases, Mean Square Error (MSE), Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) were used as the performance metrics. The developed model presents itself superior, with reduced prediction error when compared with previously existing models in the literature.

1.0 Introduction

The field of time series analysis and forecasting is useful in a variety of scientific, engineering and general applications (e.g., weather forecasting, predictive control, signal processing, population forecasting, enrolment and finance among others) (Panagiotakis *et al.*, 2016).

Linear, non-linear and clustering models are some of the model types for analysing and forecasting time series. Traditionally, time series forecasting problems are solved using linear Moving Average (MA) models, auto-regressive (AR) models and linear Auto-Regressive Integrated Moving Average (ARIMA) models (Smith and Wunsch, 2015). The traditional forecasting methods require larger observations and are unable to deal with forecasting problems in which the historical data need to be represented by linguistic values (Huang et al.,, 2011; Shah, 2012; Song and Chissom, 1993a). It is also confined to linearity assumptions only (Shah, 2012), which introduces large errors in the predicted values.

To overcome these limitations of traditional methods, many researchers used soft computing techniques such as fuzzy logic introduced by Zadeh (1965) to handle the uncertainty and vagueness available in linguistic terms. The concepts of fuzzy sets (Zadeh, 1978) have been successfully adopted to time series by Song and Chissom (1993a). Fuzzy time series methods consist of three steps. These are fuzzification, identification of fuzzified relations and defuzzification. Many studies on these three steps have been reported in literature, because they influence the forecasting accuracy of a model. Using these three steps, Song and Chissom (1993a) developed the benchmark fuzzy time series forecasting models, the time-invariant fuzzy time series model and the time-variant fuzzy time series model, to forecast the student of enrolment of the University of Alabama (Song and Chissom, 1993b, 1994). The min-max composition adopted by Song and Chissom (1993a) is computationally tasking when the fuzzy matrix is large. Chen (1996), improved on it by introducing a simple arithmetic operation to replace the complicated operations in Song and Chissom's (1993) models. Due to simplicity of Chen's (1996) model calculation, it was easy and straightforward to integrate heuristic knowledge and forecast better (Huarng, 2001).

The model has since been widely extended and adapted, with researchers (Huang *et al.,* 2011; Yu, 2005) Hu seeking to improve forecasting accuracy by adjusting the length of linguistic intervals or by changing the weighting approach or incorporating clustering techniques to tackle the interval problem. There are two major categories of FTS algorithms: FTS algorithms based on intervals of the universal set (Askari *et al.,* 2015) and FTS algorithms based on fuzzy clustering (Askari *et al*., 2015). Determination of interval length and fuzzy relations are two important factors affecting performance of fuzzy time series approaches. Many studies established the efficacy of determining an objective interval length and universe of discourse which is an issue of concern that affect forecast accuracy (Huang *et al*., 2011; Yu, 2005).

So many researchers in the last three decades have used various techniques in order to address problems of subjective decisions associated with the fuzzification stage. Such techniques include; Genetic Algorithm (GA) (Chen and Chung, 2006), Particle Swarm Optimization (PSO) (Huang *et al*., 2011; Kuo *et al.,* 2009), Artificial Bee Colony (ABC), Ant Colony Optimization (ACO) (Cai et al., 2015), Fuzzy C-Means (FCM) (Egrioglu *et al.,* 2013), Gustafson-kessel clustering algorithm (Egrioglu et al., 2011) and hierarchical clustering (Chen and Tanuwijaya, 2011).

Even though some of these approaches try avoiding subjective judgements in determining intervals, membership values were still chosen arbitrarily. Fuzzy clustering approaches have been developed to overcome this problem. FCM clustering technique was used by Cheng *et al.* (2008) and Li *et al.* (2008), Gustafson Kessel fuzzy clustering technique was used by Egrioglu *et al.* (2011) in order to eliminate the problem of subjective determination of membership values.

In most of these studies, fuzzy sets having the highest membership values are considered and other values are ignored in the determination of fuzzy relations. Yu and Huarng (2010) suggested an approach in which all membership values are considered in the determination of fuzzy relations with feed forward artificial neural network (FFANN) but again the membership values are specified subjectively. Aladag *et al.,* (2012) and Yolcu *et al.,* (2013) considered all the membership values which were determined objectively using FCM technique for first order FTS forecasting models. However, most of the time series encountered in real life problems involve more complicated relations, which makes first order FTS forecasting models insufficient to forecast these kind of time series (Cagcag *et al*., 2016).

Corresponding author's e-mail address: muktyx5@yahoo.com

A method considering all the membership values at the defuzzification step was introduced by Yolcu *et al.* (2013), which Bas *et al.* (2015) and Cagcag *et al.* (2016) adopted and developed a high order model. However, the FCM used has been associated with sensitivity to noise, difficulty in handling outlier points and high dimensional data and pre-mature convergence. This, consequently, affects and reduces the forecasting accuracy. Premature convergence is typically caused by loss of diversity, a measure of the amount variety, i.e. number of different solutions in the population, and how different they are (distance between alternative solutions).

This loss of diversity could be as a result of too strong selective pressure towards best solution i.e. too much exploitation of existing building blocks from current population. As such, Cat Swarm Optimization based Clustering (CSO-C) algorithm was developed to overcome the problem of pre-mature convergence, in order to improve forecast accuracy.

2.0 Materials and Methods

2.1 Data Collection

In this paper, the benchmark fuzzy time series data of student enrolment of University of Alabama and Taiwan Future Exchange (TAIFEX) data were used for testing the model and UniMaid enrolment data was used for the model application.

The time series data of Student enrolment of University of Alabama between 1971 and 1992 was obtained from the work of Song and Chissom (1993a) as can be found in Table 1.

Table 1: Student Enrolment Data of University of Alabama

Years	Actual	Years	Actual
1971	13055	1982	15433
1972	13563	1983	15497
1973	13867	1984	15145
1974	14696	1985	15163
1975	15460	1986	15984
1976	15311	1987	16859
1977	15603	1988	18150
1978	15861	1989	18970
1979	16807	1990	19328
1980	16919	1991	19337
1981	16388	1992	18872

Table 2 shows the time series data of TAIFEX obtained from the work of Bas *et al.,* (2015). The time series has 47 daily observations between 03.08.1998 and 30.09.1998. The last 16 observations between 10.09.1998 and 30.09.1998 were used as the test set to evaluate the model's performance.

Table 2: TAIFEX Data

Date	TAIFEX	Date	TAIFEX	Date	TAIFEX
03.08.1998	7552.0	24.08.1998	695.5.00	11.09.1998	6726.50
04.08.1998	7560.00	25.08.1998	6949.00	12.09.1998	6774.55
05.08.1998	7487.00	26.08.1998	6790.00	15.09.1998	6762.00
06.08.1998	7462.00	27.08.1998	6835.00	16.09.1998	6952.75
07.08.1998	7515.00	28.08.1998	6695.00	17.09.1998	6906.00
10.08.1998	7365.00	29.08.1998	6728.00	18.09.1998	6842.00
11.08.1998	7360.00	31.08.1998	6566.00	19.09.1998	7039.00
12.08.1998	7330.00	01.09.1998	6409.00	21.09.1998	6861.00
13.08.1998	7291.00	02.09.1998	6430.00	22.09.1998	6926.00
14.08.1998	7320.00	03.09.1998	6200.00	23.09.1998	6852.00
15.08.1998	7320.00	04.09.1998	6403.20	24.09.1998	6890.00
17.08.1998	7219.00	05.09.1998	6697.50	25.09.1998	6871.00
18.08.1998	7220.00	07.09.1998	6722.30	28.09.1998	6840.00
19.08.1998	7285.00	08.09.1998	6859.40	29.09.1998	6806.00
20.08.1998	7274.00	09.09.1998	6769.60	30.09.1998	6787.00
21.08.1998	7225.00	10.09.1998	6709.75		

The time series enrolment data of University of Maiduguri (UniMaid) shown in Table 3 was obtained from the Academic planning Unit of the University. The time series has 18 yearly enrolments between 1976 and 1993.

Table 3: UniMaid Enrolment Data

Year	Enrolment	Year	Enrolment
1976	743	1985	5800
1977	1128	1986	6168
1978	1882	1987	6711
1979	2500	1988	7238
1980	2925	1989	7687
1981	3251	1990	7960
1982	4561	1991	8302
1983	5329	1992	9884
1984	5719	1993	11410

2.2 Fuzzy Time Series and its concept

The definition of fuzzy time series was firstly introduced by Song and Chissom (1993a). Basic definitions of fuzzy time series not including constraints such as linear model and observation number can be given as follows (Song and Chissom, 1993a).

Corresponding author's e-mail address: muktyx5@yahoo.com

Definition 1 Fuzzy time series

Let $Y(t)(t=...,0,1,2,...)$, a subset of real numbers, be the universe of discourse by which fuzzy sets $f_j(t)$ are defined. If $F(t)$ is a collection of $f_1(t), f_1(t),...$, then $F(t)$ is called a fuzzy time series defined on $Y(t)$.

Definition 2 Suppose $F(t)$ is caused by $F(t-1)$ only, that is, $F(t-m) \rightarrow F(t)$. Then this relation can be expressed as $F(t) = F(t-1)^{\circ} R(t, \mathrm{t}-1)$ where $R(t,t-1)$ is the fuzzy relationship between $F(t-1)$ and $F(t)$, and $F(t) = F(t-1)^{\circ} R(t, \mathrm{t}-1)$ is called the first order model of $F(t)$.

Definition 3 Suppose $R(t,t-1)$ is a first order model of $F(t)$. If for any $t, R(t,t-1)$ is independent of t, that is, for any $t, R(t,t-1) = \mathrm{R(t-1)(t-2)}$, then, $F(t)$ is called a time invariant fuzzy time series otherwise, it is called a time variant fuzzy time series.

Definition 4 High order fuzzy time series forecasting model.

Let $F(t)$ be a fuzzy time series. If $F(t)$ is caused by $F(t-1), \mathrm{F}(t-2),...,F(t-n)$, then this fuzzy logical relationship is represented by

$F(t-n),...,F(t-2),F(t-1) \rightarrow F(t)$ and it's called the nth order fuzzy time series forecasting model.

2.3 Cat swarm optimization clustering

In contrast to most of the existing clustering techniques, the CSO-C algorithm requires no prior knowledge of the data to be classified. Rather, it determines the optimal number of partitions of the Data"on the run" (Santosa and Ningrum, 2009).

The CSO Clustering consists of two main parts, clustering data and searching for the best cluster center using the CSO algorithm (Chu and Tsai, 2007). The inputs for clustering CSO will be the population of data that are going to be clustered, number of cluster k, and number of copy (will be used in seeking mode). Steps of clustering CSO are described below (Santosa and Ningrum, 2009).

Step 1: Defining the initial cluster center

Randomly choose k point from the data to be the initial cluster centre.

Step 2: Grouping data into clusters

Put the data into cluster using Euclidean distance as follows (Santosa and Ningrum, 2009).

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{1}$$

Step 3: Calculating the Sum of Squared-Error (SSE)

SSE is the fitness function in this algorithm. The value of SSE are calculated, subject to the following constraint, using (Santosa and Ningrum, 2009).

$$SSE = \min \sum_{i=1}^{k} \sum_{x \in Di} (u_i, \|x - m_i\|)^2 \tag{2}$$

$$0 \le u_i \le 1 \tag{3}$$

In solving the minimization problem, the values of m_i are updated using the following formula:

$$u_{ij} = \frac{1}{\sum_{k=1}^{k}\left(\frac{d\left(x_j, v_i\right)}{d\left(x_j, v_k\right)}\right)^{2/(k-1)}} \tag{4}$$

Where; x is data, mi the cluster centeri, k the number of cluster and ui the membership i.

Step 4: Clustering optimization with CSO

In this algorithm, cluster centre represents the cat and the solution set will be the new cluster centre which is expected to come up with the smaller SSE value than before.

Step 4.1: Seeking mode

Seeking mode is intended to look for points in an area around the cluster centre which have possibilities in resulting in more than an optimal fitness value.

Step 4.2: Updating SSE and cluster center

The value of SSE obtained from seeking mode will be compared with the previous value of SSE, if seeking SSE < earlier SSE then it becomes the new cluster centre. Otherwise, if seeking SSE ≥ earlier SSE, the previous cluster centre will be used.

Step 4.3: Tracing Mode

Tracing mode is intended to shift the point so it will be concentrated to a better position with a more optimal fitness value.

Step 4.4: Repeat step 4.2 for tracing SSE and cluster center

The value of SSE obtained from tracing mode will be compared with the previous value of SSE, if tracing SSE < earlier SSE then it becomes the new cluster center. If tracing SSE ≥ earlier SSE, the previous cluster center will be used.

Step 5: Repeat step 4 until it reach the stopping criteria.

2.4 Artificial Neural Network

Artificial neural networks is a mathematical algorithm inspired by biological neural networks, which can learn from examples and generalize what is learnt (Bas et al., 2015). Architecture structure, learning algorithm and activation function are the three basic operational component of the artificial neural networks (Bas et al., 2015).

2.5 Developed Model Procedure

The flowchart of the developed model presented in figure 1 shows the steps carried out in the model implementation aimed at improving forecast accuracy.

2.5.1 Training/testing time series data

The data's of tables 1, 2 and 3 were used as the input data for the developed model.

Corresponding author's e-mail address: muktyx5@yahoo.com

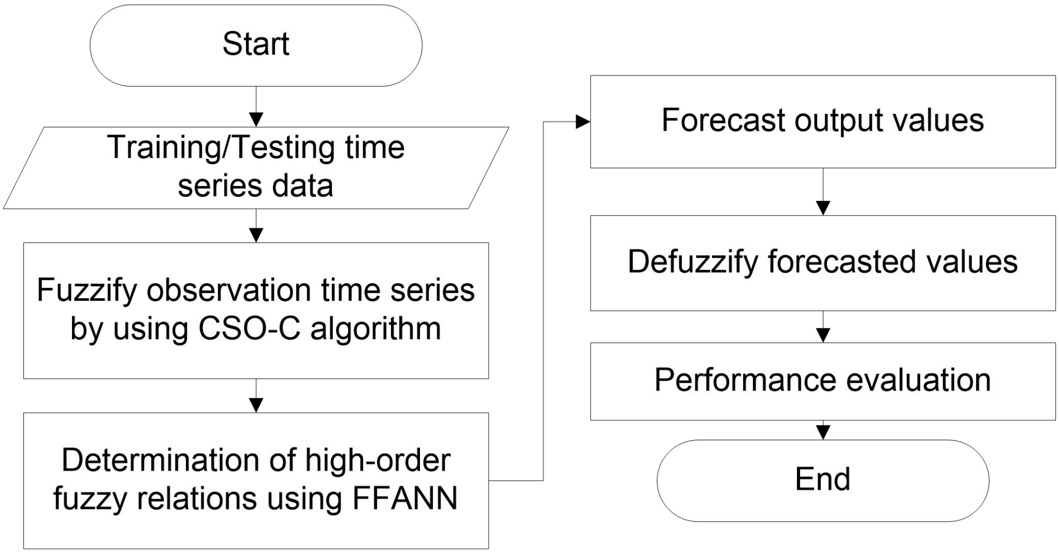

Figure 1: Flowchart of the developed model

2.5.2 Fuzzify time series observation by using CSO-C algorithm

At this stage, the CSO-C algorithm was used in fuzzifying the time series data specified in section 2.5.1 as a means of determining an optimized cluster center and membership values for fuzzy time series forecasting. The population of data to be clustered used comprised of data shown in Table’s 1, 2 and 3. The number of clusters chosen was seven as used in the work of Bas et al., (2015) as specified in Table 4

Table 4: Input Parameter to the CSO-C Algorithm

Parameters	Specifications
D	Data
K	7

The initial cluster center has been chosen randomly from the data set, the data were then grouped into clusters by their closeness using Euclidean distance formula of equation 1 to calculate the distance between the data and the cluster center. The sum of squared error between the data and the cluster center was then calculated using equation 2.

In clustering with the CSO-C algorithm, the cluster center represents the cat and the solution set will be the new cluster center which is expected to come up with the smaller SSE value than what was obtained. Here, all the cats go through the process of tracing and seeking. The CSO-C parameters initialization were specified in table 5, which include count of dimension to change (CDC), seeking memory pool (SMP), seeking range of the selected dimension (SRD) and self-position consideration (SPC). The CDC in seeking mode were assumed 100% value, so every dimension of cat’s copy changes, SMP represents how many copy a cluster center has, SRD declares a mutation ratio with a value between [0,1] and SPC is a Boolean random value{0,1}.

Table 5: CSO Parameters (Santosa and Ningrum, 2009)

Parameters	Specifications
SMP	5
CDC	100%
SRD	0.2
Const1	2
r1	[0,1]
Velmax	0.9

The simulation parameters of table 5 were selected based on the optimal performance achieved as reported in the work of Santosa and Ningrum (2009).

The membership values were computed and updated using equations 2 and 4. The next step is activating tracing mode which looks for an area around the cluster center with possibilities of generating a more optimal fitness value. The distance between the data and the new cluster center where been calculated using equation 1, the new cluster center calculated using 2 and the membership values computed with equation 4.

The SSE of the tracing mode computed were then compared with the previous values of SSE and if the tracing SSE is less than the previous SSE, then the membership value and cluster center resulting from the tracing becomes the new membership value and cluster center. Conversely, if the value of the tracing SSE was greater than or equal the earlier SSE, then the previous membership values and cluster centers would be used. The process continues until the number of iterations is met, after which, the number of cluster K, the membership values and the best cluster center were then obtained.

2.5.3 Determination of high order fuzzy relation using FFANN

The feed forward artificial neural nework (FFANN) was used in establishing fuzzy relationship. That is, the produced degree are employed as the inputs and the target values of FFNNs. As such, this prevents arbitrarily degrees of membership determination, hard matrix operations or complex group tables need not be used. The input to the FFNN consists of membership values obtained from section 2.5.2 and the activation function used in all layers of the FFNN of whose architecture is given as follows:

$$f(x) = \left(1 + \exp(-x)\right)^{-1} \tag{5}$$

Optimal weights were obtained by training using Lavenberg-Marquardt learning algorithm.

2.5.4 Defuzzify forecasted values

In defuzzifying the forecasted values, weighted averages were used instead of the central method, and so, the fuzzy sets with a less degree of membership were considered. This helps in avoiding possible knowledge loss. Using the following expression, the degrees of membership of the fuzzy forecast obtained from the previous section were transformed into weights.

$$w_{it} = \frac{\hat{u}_{it}}{\hat{u}_{1t} + \hat{u}_{2t} + \ldots + \hat{u}_{kt}} \tag{6}$$

Corresponding author's e-mail address: muktyx5@yahoo.com

The defuzzified forecasts of the observation at time t are calculated using the following expression

$$\hat{X}_t = \sum_{i=1}^{k} w_{it} m_i \tag{7}$$

Where, $\hat{u}_{it}, i = 1,2,...,k,$ denotes the membership degrees for the fuzzy forecasting value of the time series data at t obtained from the output of the FFANN. $w_{it}, t = 1,2,...,k,$ denotes the weights used for defuzzifying the forecasts. $m_i, i = 1,2,...,k,$ denotes the cluster center values.

2.5.5 Performance evaluation

In evaluating the performance of the developed model, mean absolute performance error (MAPE) and root mean square error (RMSE) were used as metrics as presented in the following expressions respectively.

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left(\left|\mathrm{actual}_i - \mathrm{forecast}_i\right|\right)}{\mathrm{actual}_i}\times 100 \tag{8}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{actual}_i - \mathrm{forecast}_i\right)^2} \tag{9}$$

The percentage improvement of the developed model over other models reported in the literature was computed using the expression as follows:

$$Percentage\ Improvement = \frac{Initial - Final}{Initial}\times 100\% \tag{10}$$

Where initial stands for RMSE or MAPE values computed from the approaches of the Benchmark FTS or FCM based, and final for the values obtained from the developed CSO-C based model.

3.0 Results and Discussion

In this section, the performance of the developed fuzzy time series forecasting is evaluated and analyzed using two different data sets that were previously studied with other methods in the literature and a new data set that has not been previously used

3.1 Obtained Results for Student Enrolment Data of University of Alabama

The optimized cluster center values of fuzzy sets and the membership values obtained from the developed model on the training data set of Table 1 as earlier explained in section 2.5 were presented in Table 6 and Table 7 respectively. The defuzzified forecast obtained were given in Table 8.

Table 6: Obtained Cluster Centre Values for University of Alabama Data

m1	m2	m3	m4	m5	m6	m7
13490.51	15117.90	15621.22	16790.81	18145.80	18921.80	19331.84

Table 7: Obtained Membership Values for University of Alabama Data

Year	k1	k2	k3	k4	k5	k6	k7
1971	0.9053	0.0403	0.0261	0.0123	0.0066	0.0050	0.0044
1972	0.9955	0.0022	0.0012	0.0005	0.0002	0.0002	0.0002
1973	0.8538	0.0773	0.0393	0.0142	0.0066	0.0047	0.0041
1974	0.0872	0.7121	0.1481	0.0289	0.0107	0.0071	0.0059
1975	0.0054	0.1776	0.7993	0.0117	0.0029	0.0017	0.0014
1976	0.0079	0.7014	0.2719	0.0120	0.0033	0.0020	0.0016
1977	0.0000	0.0014	0.9981	0.0002	0.0000	0.0000	0.0000
1978	0.0085	0.0866	0.8314	0.0553	0.0092	0.0051	0.0040
1979	0.0000	0.0000	0.0002	0.9994	0.0002	0.0000	0.0000
1980	0.0014	0.0049	0.0094	0.9671	0.0106	0.0040	0.0027
1981	0.0129	0.0674	0.1849	0.6701	0.0352	0.0169	0.0125
1982	0.0067	0.2557	0.7166	0.0138	0.0035	0.0021	0.0018
1983	0.0034	0.0955	0.8888	0.0082	0.0020	0.0012	0.0009
1984	0.0003	0.9960	0.0033	0.0003	0.0000	0.0000	0.0000
1985	0.0007	0.9884	0.0096	0.0008	0.0002	0.0001	0.0001
1986	0.0146	0.1207	0.6877	0.1391	0.0194	0.0105	0.0081
1987	0.0004	0.0015	0.0030	0.9904	0.0028	0.0011	0.0008
1988	0.0000	0.0000	0.0000	0.0000	0.9999	0.0000	0.0000
1989	0.0000	0.0002	0.0002	0.0005	0.0033	0.9784	0.0174
1990	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.9999
1991	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.9998
1992	0.0000	0.0001	0.0002	0.0005	0.0039	0.9853	0.0100

Table 8 shows the results of forecast obtained from the developed model, where the performance criteria explained in section 2.5 were used in computing the model performance. The obtained MSE, RMSE and MAPE were 785.17, 88.46 and 0.0046 respectively. From the results comparison of Table 9, it can be seen that the developed model has the lowest prediction error in terms of MSE, this is an indication of better forecast exhibited by the developed model. It shows that the developed model recorded 98.10% over Song and Chissom (1993a), 86.98%, over Egrioglu *et al.*, (2011), 83.14% over Aladag *et al.*, (2012), 70.02% over Bas *et al.*, (2015) and 28.77% over Cagcag *et al.,* (2016) computed with equation 10.

Corresponding author's e-mail address: muktyx5@yahoo.com

Table 8: Defuzzified Forecast for University of Alabama Data

Year	Data	Forecast
1973	13867	13945
1974	14696	14702
1975	15460	15397
1976	15311	15386
1977	15603	15546
1978	15861	15934
1979	16807	16834
1980	16919	16828
1981	16388	16301
1982	15433	15389
1983	15497	15416
1984	15145	15295
1985	15163	15316
1986	15984	15950
1987	16859	16834
1988	18150	18161
1989	18970	19056
1990	19328	19186
1991	19337	19200
1992	18876	18990
	MSE	7825.17
	RMSE	88.46
	MAPE	0.0046

Table 9: The Results Comparison

Methods	MSE
Song and Chissom (1993a)	412499
Egrioglu et al., (2011)	60140
Aladag et al., (2012)	46422
Bas et al., (2015)	26108
Cagcag et al., (2016)	10987
The Developed	7825.17

3.2 Obtained Results for TAIFEX Data

Table 10 shows results of forecast obtained from TAIFEX data of table 2 following methods discussed in section 2.5.

The results further shows the forecasts obtained from the model of Bas *et al.*, (2015) and Cagcag *et al.*, (2016) on the same data set. After close examination, it can be seen that the developed model recorded the least prediction error with a significant improvement in terms of MSE, RMSE and MAPE.

The performance improvement computed with equation 10 was used to obtain the performance improvement of the developed model over some methods reported in literature. The model developed recorded an RMSE improvement of 65.57% over that of Bas *et al.*, (2015) and 64.22% over that of Cagcag *et al.*, (2016).

Table 10: Forecast for TAIFEX data

Date	TAIFEX	Bas et al., (2015)	Cagcag et al., (2016)	Developed
10.09.1998	6709.75	6804.31	6719.92	6721.21
11.09.1998	6726.50	6804.31	6719.92	6721.92
14.09.1998	6774.55	6761.17	6843.75	6773.72
15.09.1998	6762.00	6758.51	6719.92	6774.30
16.09.1998	6952.75	6856.13	6843.75	6946.40
17.09.1998	6906.00	6856.13	6843.75	6901.86
18.09.1998	6842.00	6856.13	6843.75	6841.70
19.09.1998	7039.00	6856.13	6843.75	7038.93
21.09.1998	6861.00	6856.13	6843.75	6861.63
22.09.1998	6926.00	6856.13	6843.75	6919.85
23.09.1998	6852.00	6856.13	6843.75	6853.61
24.09.1998	6890.00	6856.13	6843.75	6893.58
25.09.1998	6871.00	6807.48	6843.75	6863.90
28.09.1998	6840.00	6793.07	6843.75	6840.41
29.09.1998	6806.00	6793.07	6843.75	6819.25
30.09.1998	6787.00	6856.13	6778.28	6879.66
	MSE	4869.25	4507.78	576.96
	RMSE	69.78	67.14	24.02
	MAPE	0.00761	0.0066	0.00152

3.3 Obtained Results for UniMaid Enrolment Data

The UniMaid enrolment data of table 3 was not previously applied for any fuzzy time series forecasting research. In view of that, the developed model was applied to obtain the forecast as shown in table 11. It can be seen that the forecast produced has an MSE of 1395543.21, RMSE 0f 373.55 and MAPE 0.0667.

Corresponding author's e-mail address: muktyx5@yahoo.com

Table 11: Forecast for UniMaid Data

Year	Enrolment	Forecast
1978	1882	2702.84
1979	2500	2680.11
1980	2925	2737.43
1981	3251	2837.88
1982	4561	4497.60
1983	5329	5550.53
1984	5719	5849.53
1985	5800	5965.73
1986	6168	6457.61
1987	6711	6992.32
1988	7238	7372.57
1989	7687	7591.29
1990	7960	7807.23
1991	8302	8120.24
1992	9884	10031.38
1993	11410	10432.95
	MAPE	0.0667
	MSE	139543.21
	RMSE	373.55

4.0 Conclusion

In minimizing the subjective decisions made at the fuzzification stage, traditional clustering techniques such as fuzzy C-means (FCM) and K-means, have been previously used, but problems such as handling high dimensional data, sensitivity to noise, outliers and pre-mature convergence still persist. In this research, CSO-C algorithm was developed to address this issue of pre-mature convergence with a view to improving forecast accuracy in fuzzy time series forecasting. In the implementation, the developed model was tested and compared with other models forecast reported in literature for the student enrolment of University of Alabama and TAIFEX data.

The forecast results of enrolment of University of Alabama data produced an MSE improvement of 98.10% over Song and Chissom (1993a), 86.98%, over Egrioglu et al., (2011), 83.14% over Aladag et al., (2012), 70.02% over Bas et al., (2015) and 28.77% over Cagcag *et al.,* (2016). Also, the forecast results of TAIFEX data recorded an RMSE improvement of 65.57% over that of Bas *et al.,* (2015) and 64.22% over that of Cagcag et al., (2016). This clearly shows that the model developed has better forecast with reduced prediction error.

The developed model was further to UniMaid enrolment data, which was not previously reported to have been used in literature with respect the research area of fuzzy time series forecasting. The forecast produced has an MSE of 1395543.21, RMSE 0f 373.55 and MAPE 0.0667. It is expected that future researchers will forecast with the UniMaid enrolment data and see how there model suit into the data.

Reference

Aladag, CH., Yolcu, U., Egrioglu, E., and Dalar, AZ. 2012. A new time invariant fuzzy time series forecasting method based on particle swarm optimization. Applied Soft Computing, 12(10), 3291-3299.

Askari, S., Montazerin, N. and Zarandi, MF. 2015. A clustering based forecasting algorithm for multivariable fuzzy time series using linear combinations of independent variables. Applied Soft Computing, 35, 151-160.

Bas, E., Yolcu, U., Egrioglu, E. and Aladag, CH. 2015. A Fuzzy Time Series Forecasting Method Based on Operation of Union and Feed Forward Artificial Neural Network. American Journal of Intelligent Systems, 5(3), 81-91.

Cagcag, O., Yolcu, U., Egrioglu, E. and Aladag, CH. 2016. A High Order Fuzzy Time Series Forecasting Method Based on Operation of Intersection. Applied Mathematical Modelling.

Cai, Q., Zhang, D., Zheng, W. and Leung, SC. 2015. A new fuzzy time series forecasting model combined with ant colony optimization and auto-regression. Knowledge-Based Systems, 74, 61-68.

Chen, SM. 1996. Forecasting enrollments based on fuzzy time series. Fuzzy sets and systems, 81(3), 311-319.

Chen, SM. and Chung, NY. 2006. Forecasting enrollments of students by using fuzzy time series and genetic algorithms. International journal of information and management sciences, 17(3), 1-17.

Chen, SM., and Tanuwijaya, K. 2011. Multivariate fuzzy forecasting based on fuzzy time series and automatic clustering techniques. Expert Systems with Applications, 38(8), 10594-10605.

Cheng, CH., Cheng, GW. and Wang, JW. 2008. Multi-attribute fuzzy time series method based on fuzzy clustering. Expert Systems with Applications, 34(2), 1235-1242.

Chu, SC. and Tsai, PW. 2007. Computational intelligence based on the behavior of cats. International Journal of Innovative Computing, Information and Control, 3(1), 163-173.

Egrioglu, E., Aladag, C., Yolcu, U., Uslu, VR. and Erilli, NA. 2011. Fuzzy time series forecasting method based on Gustafson–Kessel fuzzy clustering. Expert Systems with Applications, 38(8), 10355-10357.

Egrioglu, E., Aladag, CH. and Yolcu, U. 2013. Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks. Expert Systems with Applications, 40(3), 854-857.

Huang, YL., Horng, SJ., He, M., Fan, P., Kao, TW., Khan, MK., Lai, JL. and Kuo, IH. 2011. A hybrid forecasting model for enrollments based on aggregated fuzzy time series and particle swarm optimization. Expert Systems with Applications, 38(7), 8014-8023.

Huarng, K. 2001. Heuristic models of fuzzy time series for forecasting. Fuzzy sets and systems, 123(3), 369-386.

Kuo, IH., Horng, SJ., Kao, TW., Lin, TL., Lee, CL. and Pan, Y. 2009. An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization. Expert Systems with Applications, 36(3), 6108-6117.

Li, ST., Cheng, YC., and Lin, SY. 2008. FCM-based deterministic forecasting model for fuzzy time series . Computer and Mathematics with Applications, 56, 3052-3063.

Corresponding author's e-mail address: muktyx5@yahoo.com

Panagiotakis, C., Ramasso, E., Fragopoulou, P. and Aloise, D. 2016. Theory and Applications of Data Clustering.

Santosa, B. and Ningrum, MK. 2009. Cat swarm optimization for clustering. Paper presented at the Soft Computing and Pattern Recognition, 2009. SOCPAR'09. International Conference of.

Shah, M. 2012. Fuzzy based trend mapping and forecasting for time series data. Expert Systems with Applications, 39(7), 6351-6358.

Smith, C. and Wunsch, D. 2015. Time series prediction via two-step clustering. Paper presented at the International Joint Conference on Neural Networks (IJCNN), 2015

Song, Q. and Chissom, BS. 1993a. Forecasting enrollments with fuzzy time series—part I. Fuzzy sets and systems, 54(1), 1-9.

Song, Q. and Chissom, BS. 1993b. Fuzzy time series and its models. Fuzzy sets and systems, 54(3), 269-277.

Song, Q. and Chissom, BS. 1994. Forecasting enrollments with fuzzy time series—part II. Fuzzy sets and systems, 62(1), 1-8.

Yolcu, U., Aladag, CH., Egrioglu, E., and Uslu, VR. 2013. Time-series forecasting with a novel fuzzy time-series approach: an example for Istanbul stock market. Journal of Statistical Computation and Simulation, 83(4), 599-612.

Yu, HK. 2005. A refined fuzzy time-series model for forecasting. Physica A: Statistical Mechanics and its Applications, 346(3), 657-681.

Zadeh, LA. 1965. Fuzzy sets. Information and control, 8(3), 338-353.

Zadeh, LA. 1978 Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1, 3-28.