Journal of Theoretical and Applied Information Technology
15th November 2018. Vol.96. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A COMPUTATIONALLY EFFICIENT METHOD FOR HIGH ORDER FUZZY TIME SERIES FORECASTING

[1]SIBARAMA PANIGRAHI, [2]H.S. BEHERA
[1]Department of Computer Science and Engineering, VSSUT, Burla, Odisha, India
[2]Associate Professor, Department of Information Technology, VSSUT, Burla, Odisha, India
E-mail: [1]panigrahi.sibarama@gmail.com, [2]hsbehera_india@yahoo.com

ABSTRACT

Despite over more than twenty years of research on fuzzy time series forecasting (TSF) and several studies indicating superior performance, an appropriate computationally efficient method have not been developed to predict various time series using fuzzy TSF method. Motivated by this, in this paper a computationally efficient method is proposed to forecast various time series by using a high order fuzzy TSF model. In this method, the fuzzy TSF parameters such as length of intervals, number of intervals and order of the model are determined deterministically. The order of the model is determined by making analysis on the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the fuzzy time series. The length of interval is determined by using single-variable constrained optimization based method and defuzzification is done by using interval average. In addition, motivated by the boost in forecasting performance due to the use of artificial neural network (ANN) for representing FLR, in this paper, a fast learning one-pass neural network called generalized regression neural network (GRNN) is used for representing the FLR. The use of GRNN model avoids the problems of traditional ANN models such as: ad hoc architecture selection and determining large number of weights and other parameters. In order to evaluate the effectiveness of the proposed model, ten univariate time series datasets are considered and three recent fuzzy time series forecasting models using ANN to represent FLR are implemented. Each model is independently executed for fifty times on each time series and extensive statistical analysis is made on the obtained results. Results revealed the robustness and statistical superiority of the proposed model considering its alternatives existing in the recent literature.

Keywords: *Time Series Forecasting, Fuzzy Time Series, Fuzzy Logical Relationship, Autocorrelation and partial Autocorrelation function, Generalized Regression Neural Network*

1. INTRODUCTION

In this internet era, a vast amount of data is collected from various domains. Therefore, various data analysis techniques have been developed to extract unknown, meaningful and potentially useful information from the collected data which assist in better planning and decision making. One of the most important data analysis techniques is time series forecasting (TSF) which predict the future outcomes of a phenomenon by analyzing its past observations.

TSF have been performed by using various stochastic and deterministic models. Traditionally, deterministic statistical models have been widely used for TSF. Although the statistical models are computationally very efficient, it is very difficult to identify the underlying model for many data generating processes (DGP). In addition, most of the statistical models presume that the time series are produced from linear processes and often fail to capture the issues like uncertainty and vagueness associated with most of the real world time series. In order to tackle this issue, methods based on fuzzy set theory called fuzzy TSF have been developed.

Fuzzy TSF model is first developed by Song and Chissom[1-2] in 1993. These models have gained tremendous attention in forecasting domain owing to its inherent capability to handle vagueness and nonlinearity observed in most of the real world time series. Therefore, since the development of these models in early nineties, a number of variants have been developed and applied in various forecasting applications like enrollment of a university, stock, electricity load, road accidents, gold price, tourism demand, temperature and many more. All variants of fuzzy TSF models undergo the following four steps:

Journal of Theoretical and Applied Information Technology
15th November 2018. Vol.96. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Step-1: Define the universe of discourse for time series, determine the effective length of interval and divide the universe of discourse.

Step-2: Define the fuzzy sets and transform the real valued time series into linguistic terms using fuzzification process.

Step-3: Identify the order of the model and based on which define the FLR.

Step-4: Using FLR forecast the linguistic terms of future points and defuzzify to obtain the true forecasts.

Several studies have been made with respect to each of the four steps either to improve the forecasting accuracy or to reduce computational complexity. Huarng[3] for the first time addressed that the determination of effective length of interval in Step 1 affects the forecasting accuracy significantly and proposed average-based and distribution-based methods for partitioning the universe of discourse into equal length intervals. Continuing his work in determining the effective length of interval, Hurang et al.[4] proposed a ratio-based method to unequally partition the universe of discourse and suggested that variable length partitioning of universe of discourse improves the forecasting accuracy. However, the ratio-based methods are computationally expensive because of calculation of relative difference and its cumulative distributions [5]. For this, Yolcu et al. [5] proposed a simplified single-variable constrained optimization based method which provided improved forecasting accuracy. Recently, various optimization techniques based on swarm and evolutionary computation like particle swarm optimization (PSO) [6-8], simulated annealing[9], genetic algorithm (GA) [10-11], ant colony optimization(ACO)[12] to determine the effective length of interval. However, the major drawback of these evolutionary and swarm based methods is that: the methods require comparatively more computational time, performance is subjected to algorithmic parameters and initial values of chromosomes or reactants or particles. Therefore in this paper, looking at a computationally efficient model, we have used Yolcu's[5] method for determining the length of intervals.

With respect to step-2, so far much work has not been done apart from using type-2 fuzzy set theory [13-15] instead of type-1 fuzzy set theory. However, the use of type-2 fuzzy logic requires higher computational overhead [15]. Therefore, in most of the studies Type-1 fuzzy logic has been chosen for fuzzy TSF.

Step-3 is one of the most crucial steps affecting the forecasting performance of fuzzy TSF models in which the order of the model is identified and FLR are constructed. Generally, rule based methods have been predominantly used to represent the FLR. However, to improve the forecast accuracy and achieve better computational complexity, authors have used artificial neural network (ANN) like multilayer feed forward neural network [16-20, 22-24] multiplicative neuron model [21, 25], fuzzy inference system [26], k-means clustering [27], PSO [28], rough set theory [29] as an alternative. Out of all these alternatives the use of ANN to define FLR has gained tremendous attention [25]. However, the choice of ANN architecture (number of inputs, number of hidden neurons) [40] can affect the performance of the model significantly. In addition, the computational overhead and architecture selection problem increases with the increase in number of layers of ANN. Motivated by this, in our model, we have used GRNN for representing FLR because GRNN is a one-pass fast learning neural network with a single design parameter [30]. Pertaining to step-4, most of the authors have used mid-point of intervals, interval average, centroid based methods in the defuzzification step. However, defuzzification by using interval average provided superior forecasting accuracy than other defuzzification alternatives [15]. Motivated by this, in the proposed model, interval average is used for defuzzification.

The organization of remainder of this paper is as follows. Section 2 provides the basic definitions relating to fuzzy TSF. Section 3 provides brief introduction to GRNN. The proposed methodology is illustrated in Section 4. In Section 5, the experimental setup is presented and simulation results are analyzed. Finally, conclusions are drawn and presented in Section 6.

2. FUZZY TIME SERIES FORECASTING

Definition 1:

A fuzzy time series $F(t)$ is a set of linguistic terms $f_i(t)(i = 1,2, \dots)$ representing a real valued time series $\mathrm{Y(t)(t} = 1,2, \dots)$ with the universe of discourse is a subset of real numbers R.

Definition 2:

The fuzzy relation $R(t, t-1)$ between the fuzzy sets $F(t-1)$ and $F(t)$ can be denoted as $F(t) = F(t-1) \circ R(t, t-1)$ or $F(t-1) \rightarrow F(t)$ which means $F(t)$is caused by $F(t-1)$, with $\circ$ representing an operator. The consequential model is called the first-order fuzzy time series model.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Definition 3:

The FLR between two successive observations $F(t-1)$ and $F(t)$ is represented by $f_i \rightarrow f_j$ with $F(t-1) = f_i$ and $F(t) = f_j$. Here f_i is the current state or left hand side of FLR and f_j is the next state or right hand side of FLR.

Definition 4:

If a fuzzy time series $F(t)$is caused by k-past values of the series, then $F(t)$ is a k-order fuzzy time series denoted by $F(t-1), F(t-2), F(t-3) \ldots F(t-k) \rightarrow F(t)$.

3. GENERALIZED REGRESSION NEURAL NETWORK (GRNN)

GRNN [30] (as in Figure 1) is a class of radial basis function network (RBFN). It has three layers: an input layer, a hidden layer having radial basis neurons and an output layer with linear activation function. GRNN with enough number of neurons can estimate any nonlinear function to any preferred level of accuracy [31]. In GRNN, the number of hidden neurons is usually set to the number of training patterns. However, to reduce the number of neurons in hidden layer for problems having larger training patterns, Husain et al.[32] suggested to first cluster the training patterns and then based on the number of clusters the neurons are set. The distance between the training pattern and input vector is taken as input by the radial basis neurons to produce RBF transformation of the input by scaling with spread factor. Usually, multivariate Gaussian function (with A_i being the center and σ being the radius) is used as the RBF function which is defined as follows:

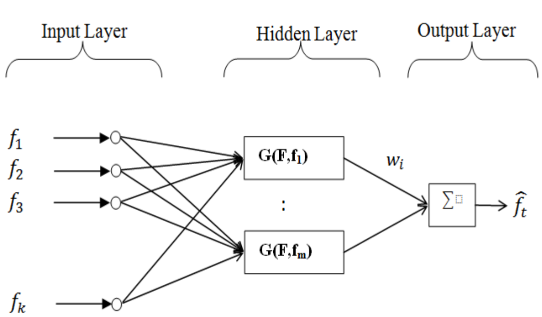

Figure 1: Architecture of GRNN

$$G(F, f_i) = \exp(-\frac{1}{2\sigma^2} ||F - f_i||^2) \quad (1)$$

Considering m-training patterns and the basic GRNN model, the output $\hat{y}_t$ for t-test point can be computed as:

$$\hat{f}_t = \sum_{i=1}^{m} w_i f_i \quad (2)$$

Where

$$w_i = \frac{\exp(-\frac{||F - f_i||^2}{2\sigma^2})}{\sum_{k=1}^{m} \exp(-\frac{||F - f_i||^2}{2\sigma^2})} \quad (3)$$

4. PROPOSED MODEL

The main goal in this paper is to develop a higher order fuzzy time series forecasting model which can be applied on various time series with least human intervention and overall good forecasting accuracy. In order to obtain a better fuzzy forecasting model, all the four steps need to be optimized. Therefore, in this paper, computationally efficient Yolcu's[5] method is used to find out the length of interval, type-1 fuzzy set theory is used, autocorrelation function is used to identify the suitable order of the model[33], the fuzzy time series is transformed to a pattern set based on the order of the model, computationally efficient GRNN is used to define FLR, interval average[15] is used to defuzzify the time series.

Our proposed model is presented in Algorithm 1. It takes a univariate time series $y = [y_1, y_2, \ldots, y_n]^T$, percentage of data for training, validation and testing as input and produces the forecasts and forecast accuracy on test set as output. The proposed model works in four steps.

In the first step, the universe of discourse U for the time series is determined by making a 20% practical increment and decrement on the maximum and minimum value of the series. In addition, to partition the universe of discourse $U = \{u_1, u_2, u_3 \ldots u_p\}$ into a set of p intervals Yolcu's[5] method is used because of its better computational efficiency and forecasting performance than other alternatives.

In step 2, the real valued time series is fuzzified by using the interval index (to which the data point belongs) in place of each data point.

Once the time series data is fuzzified, GRNN is used represent the FLR in Step 3. However, motivated by the work of Panigrahi et al. [34-35], to improve the forecasting performance, the data is first normalized using min-max [34-35] normalization technique. Then the fuzzy time series is transformed into a pattern set based on the order of the model k which is determined by using the ACF and PACF function of the time series. Then by taking a k-input GRNN model, train and validation patterns the number of RBF neurons and spread factor is determined by using neural network toolbox of MATLAB. Then, forecasts on test

Journal of Theoretical and Applied Information Technology
15th November 2018. Vol.96. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

patterns are obtained which is denormalized to obtain the forecasted interval indices.

Algorithm 1: Proposed Model

Input: a time series $y = [y_1, y_2, \ldots\ldots, y_n]^T$ and percentage of data for training, validation and test.
Output: Future values and forecast accuracy
Step 1: Identification of universe of discourse and its partitioning
1.1: Define the universe of discourse U considering a 20% decrement in the minimum value and 20% increment in the maximum value of time series. $U = [U_l \;\; U_h]$ with $U_l = min_y - D_1$, $U_h = max_y + D_2$, $D_1 = 0.2 \times min_y$ and $D_2 = 0.2 \times max_y$.
1.2: Partition the universe of discourse $U = \{u_1, u_2, u_3 \ldots u_p\}$ into p-intervals using single-variable constrained optimization based method [28], number the intervals and determine the mid-point of each interval.
Step 2: Fuzzification
Fuzzify the real valued time series $y = [y_1, y_2, \ldots\ldots, y_n]^T$ to obtain the fuzzy time series $f = [f_1, f_2, \ldots\ldots, f_n]^T$ by replacing each data point y_j of time series with the index number i of the interval u_i to which the data point belongs.
Step 3: Establishing Fuzzy logical relationships using GRNN
3.1: Normalize the fuzzy time series using min-max normalization.
3.2: Identify the order *k* of the fuzzy TSF model by making analysis on the autocorrelation and partial autocorrelation function of the train set of fuzzified series $f = [f_1, f_2, \ldots\ldots, f_n]^T$.
3.3: Transform the normalized fuzzy time series into a pattern set of *n-k+1* samples.
3.4: The number of inputs of GRNN is set to *k*, the number of hidden neurons and spread factor is determined using neural network toolbox of MATLAB. Once the parameters are determined, forecasts are obtained on test set.
3.5: The forecasts are denormalized to obtain the index of intervals to which the forecasts belong.
Step 4: Defuzzification and Forecasting
4.1: Defuzzification is carried out by using the interval average of the intervals i.e. the predicted index of intervals obtained in Step 3.5 are replaced by the respective interval average.
4.2: Forecasts are obtained and accuracy is measured.

In Step 4, the forecasted index intervals are defuzzified by using interval average [15] and forecast accuracy is measured.

5. EXPERIMENTAL SETUP AND SIMULATION RESULTS

In order to assess the effectiveness of the proposed model, we have considered three fuzzy forecasting models namely: Aladag et al.[21], Egrioglu et al. [22] and Yolcu et al. [24]. The chosen Aladag et al.[21] model uses a multiplicative neuron model whereas Egrioglu et al. [22] and Yolcu et al. [24] use feed forward MLP model. In addition, Aladag et al.[21] and Egrioglu et al. [22] do not consider membership value in their methodology while Yolcu et al. [24] considers membership values. Therefore, to consider all types of ANN models used to represent FLR found in the literature, these three models are chosen as comparative model. For Aladag et al.[21] and Egrioglu et al. [22] models, the order of the model i.e. the number of inputs to ANN is determined by using the ACF and PACF function of the time series. The other simulation parameters of all methods are kept same to the respective papers.

Table 1: Time Series Datasets

Time Series	Description of Time Series Dataset
TAIEX	Daily Taiwan Stock Exchange Capitalization Weighted Stock Index during 01/2004–09/2004.
Accidental	Monthly number of deaths in accident in USA during 1973 – 1978.
Acres Burned	Yearly number of Acres of forest land burned in fires in Canada during 1918 – 1988.
Gas	Monthly residential gas usage (in cubic feet) in Iowa during 1/1971 – 10/1979
Car	Monthly number of cars sold in Quebec during 1960 – 1968.
Milk	Monthly milk production per cow(in pounds) during 1962-1975.
Internet	Yearly internet traffic data (in bits) in united kingdom academic network during 19/11/2004–26/01/2005.
Pollution	Monthly number of pollution equipment shipped during 01/1966 – 10/1976
Passenger	Number of airline passengers (in thousands) travelling monthly during 01/1949 – 12/1960
Rainfall	Yearly rainfall(in inches) at London during 1813–1912

Table 1 presents the ten time series datasets [36] considered for evaluating the efficiency of the proposed model. The statistical properties of each time series presented in Table 2 indicate the diversity of datasets considered. Each time series is divided into train set (earliest 60%), validation set (next 20%) and test set (latest 20%). Table 3 presents the dataset division and order of the model with the order being calculated by analyzing the ACF and PACF of normalized fuzzy time series. For e.g. Figure 2 shows the ACF and PACF plot for Passenger time series. The ACF plot indicates that the series is non stationary and there is a gradual tailing off of correlation coefficients for successive lags. Therefore, the PACF plot is analyzed to determine the significant lag. It can be observed from PACF plot that the correlation coefficient cuts off after lag 12. Therefore, 12 is used as the significant lag.

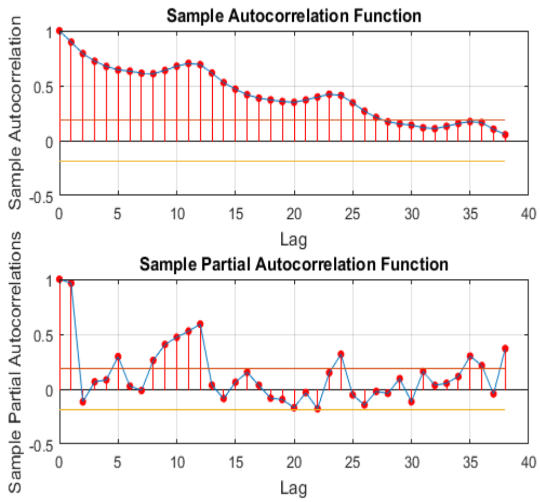

Figure 2:ACF and PACF plot for Passenger

In order to evaluate the forecasting accuracy, Hyndman et al. [37] demonstrated several forecasting measures. However, in the NN3 and M3 forecasting competitions, symmetric mean absolute percentage error (SMAPE) (eq. 4) measure is considered to evaluate the models. Therefore, in this study, SMAPE measure is used to evaluate the models. Additionally, a scale dependent measure namely root mean square error (RMSE) (eq. 5) is also considered.

$$SMAPE = \frac{1}{n}\sum_{j=1}^{n}\frac{|y_j - \hat{y}_j|}{(|y_j| + |\hat{y}_j|)/2} \tag{4}$$

$$RMSE = \sqrt{\frac{1}{NOP}\sum_{i=1}^{NOP}(y_i - \hat{y}_i)^2} \tag{5}$$

Where y_j is the jth actual value of the time series, $\hat{y}_j$ is the predicted value for the jth time series observation and n is the number of observations in the time series.

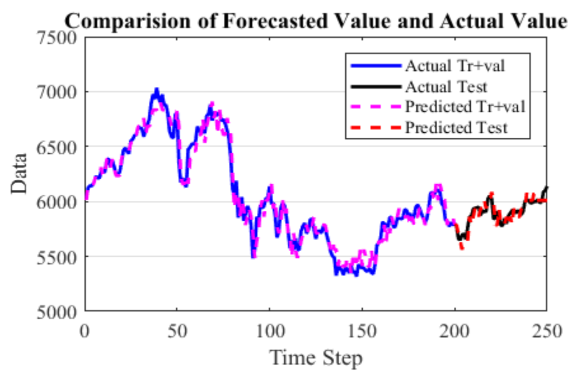

Figure:3 TAIEX Time Series Actual and Forecasted value

Table 2: Statistical properties of Time Series Datasets

Time Series	Min	Max	Mean	Std. Dev.
TAIEX	5316.87	7034.1	6033.77	418.08
Acres Burned	264000	10856000	2318563.38	2053601.11
Accidental	6892	11317	8787.7	958.34
Gas Usage	30	302	124.71	84.15
Car Sells	5568	26099	14595.11	4525.20
Milk	553	969	754.70	1.02
Internet	459055.57	2002629	2002629	368088.17
Pollution	120.89	5566.10	1439.30	1261.20
Passenger	104	622	280.29	119.96
Rainfall	16.93	38.10	24.82	4.21

Table 3: Length of Train, Validation and Test set and Order of the model

Time Series	Train	Validation	Test	Order of the model
TAIEX	150	50	50	30
Acres Burned	43	14	14	19
Accidental	43	14	15	12
Gas	64	21	21	12
Car	66	21	21	12
Milk	93	31	32	12
Internet	42	14	13	6
Pollution	78	26	26	10
Passenger	86	29	29	12
Rainfall	60	20	20	7

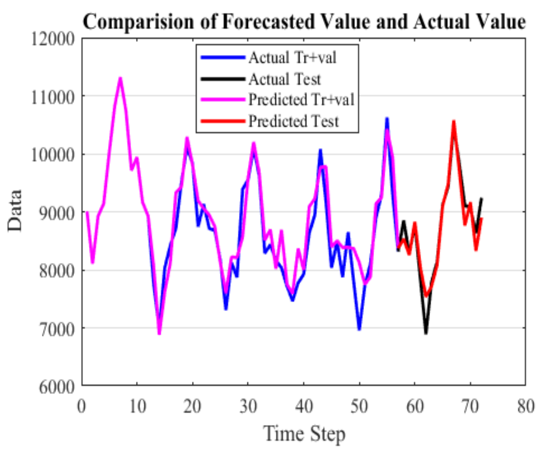

Figure:4 Accidental Time Series Actual and Forecasted value

In this study, we have made one-step-ahead prediction and based on which models are evaluated. By using the abovementioned experimental setup, we have carried out 50 independent simulations for each model on each dataset and measured the forecast accuracy.
The mean RMSE and SMAPE over 50 independent simulations are obtained and presented in Table 4 and Table 5 respectively. One can observe from Table 4 that the proposed model provided the best mean RMSE on six time series. In addition, the proposed model povided better mean RMSE than Aladag[21], Egrioglu [22] and Yolcu [24] in eight, nine and eight time series datasets respectively. One can observe from Table 5 that the proposed model has the best mean SMAPE on five time series. In addition, the proposed model povided better mean SMAPE than Aladag[21], Egrioglu [22] and Yolcu [24] in eight, nine and six time series datasets respectively.

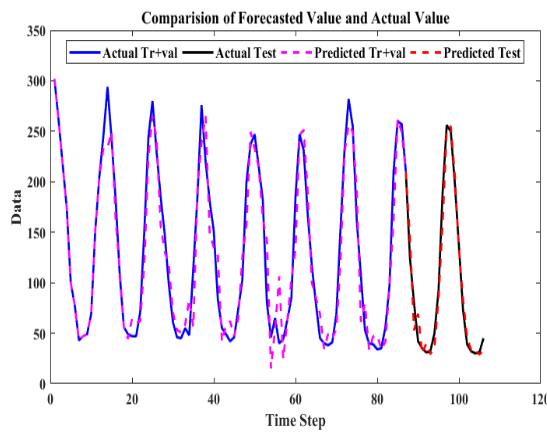

Figure:5 Gas Time Series Actual and Forecasted value

Furthermore, in order to make a statistical significance test of obtained results, we have conducted a Post Hoc analysis using Duncan's multiple range test. The significance level was set to 95% and the obtained results considering RMSE and SMAPE measure are presented in Table 6 and Table 7 respectively in which the fields of the table indicate the better model (+), worse model (−) and equivalent model (≈) to that of the proposed model. One can observe from Table 6 that the proposed model provided the statistically best RMSE on four time series. In addition, the proposed model provided statistically better RMSE than Aladag[21], Egrioglu [22] and Yolcu [24] in four, nine and six time series datasets respectively. The proposed model provided statistically inferior SMAPE than Aladag[21], Egrioglu [22] and Yolcu

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[24] in one time series than each model and in all other cases the proposed model provided statistically equivalent RMSE. One can observe from Table 7 that the proposed model provided the statistically best SMAPE on four time series. Additionally, the proposed model has shown statistically superior SMAPE than Aladag[21], Egrioglu [22] and Yolcu [24] in six, nine and six time series datasets respectively. The proposed model provided statistically inferior SMAPE than Aladag[21], Egrioglu [22] and Yolcu [24] only in one time series than each model and in all other cases the proposed model provided statistically equivalent RMSE.

Table 4: Mean RMSE over 50 Simulations (bold face denote best results)

Time Series	Proposed	Aladag[21]	Egrioglu [22]	Yolcu [24]
TAIEX	**75.32**	76.56	78.43	75.43
Acres Burned	3012932.56	3125909.76	3245804.32	**2862400.00**
Accidental	**471.87**	504.73	576.48	587.36
Gas	**20.15**	20.44	24.86	46.57
Car	**2174.71**	2201.88	2438.72	3478.71
Milk	**18.65**	23.24	28.72	61.43
Internet	180461.57	165131.53	**164021.53**	172400.00
Pollution	**879.79**	961.54	1026.78	1248.43
Passenger	49.10	**42.98**	52.48	88.48
Rainfall	**4.21**	4.79	5.04	4.97

Table 5: Mean SMAPE over 50 Simulations (bold face denote best results)

Time Series	Proposed	Aladag[21]	Egrioglu [22]	Yolcu [24]
TAIEX	1.02	1.03	1.05	**1.01**
Acres Burned	61.09	61.96	62.48	**57.42**
Accidental	**4.61**	5.07	5.34	5.38
Gas	**19.55**	23.81	32.42	40.26
Car	**10.60**	10.82	11.84	15.47
Milk	**1.80**	2.19	3.02	5.73
Internet	12.16	11.22	**10.88**	11.84
Pollution	**20.91**	25.94	28.26	32.07
Passenger	9.14	**8.39**	10.08	15.20
Rainfall	12.49	13.11	14.48	**11.65**

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 6: Individual Hypothesis Test Results considering RMSE

Time Series	Aladag[21]	Egrioglu [22]	Yolcu [24]
TAIEX	≈	–	≈
Acres Burned	≈	–	+
Accidental	–	–	–
Gas	–	–	–
Car	≈	–	–
Milk	–	–	–
Internet	+	+	≈
Pollution	–	–	–
Passenger	+	–	–
Rainfall	≈	–	≈

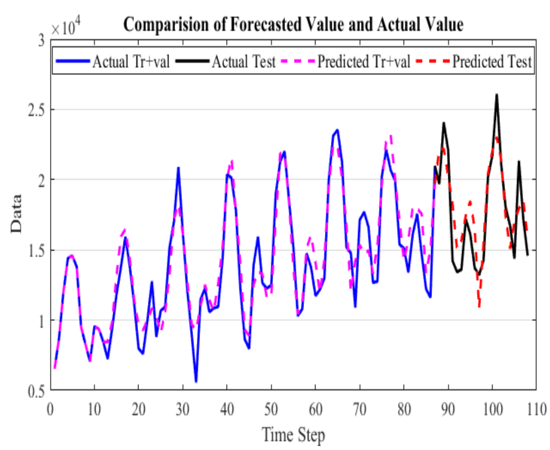

Figure:6 Car Time Series Actual and Forecasted value

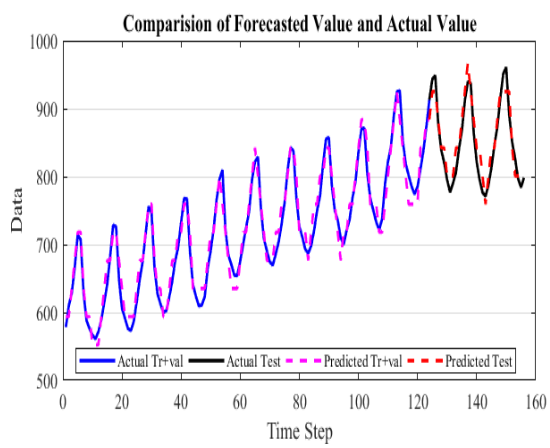

Figure:7 Milk Time Series Actual and Forecasted value

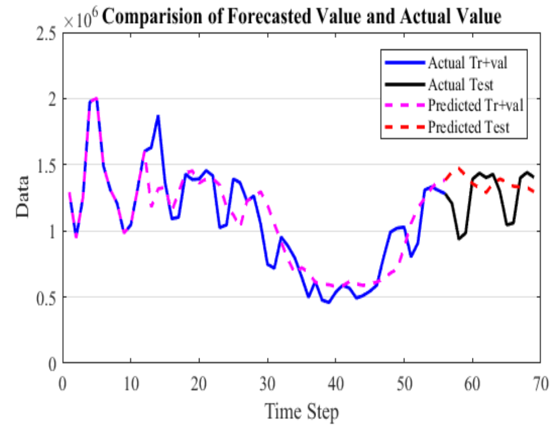

Figure:8 Internet Time Series Actual and Forecasted value

Table 7: Individual Hypothesis Test Results using SMAPE

Time Series	Aladag[21]	Egrioglu [22]	Yolcu [24]
TAIEX	–	–	–
Acres Burned	≈	–	+
Accidental	–	–	≈
Gas	–	–	–
Car	≈	–	–
Milk	–	–	–
Internet	≈	+	≈
Pollution	–	–	–
Passenger	+	–	–
Rainfall	–	–	≈

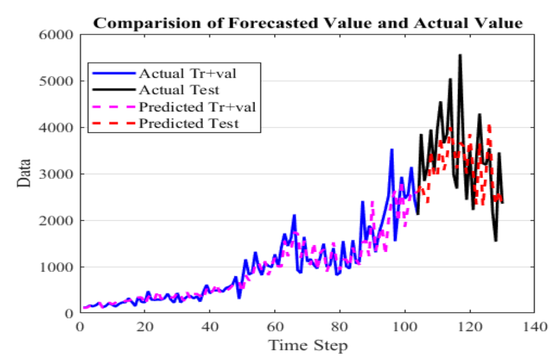

Figure:9 Pollution Time Series Actual and Forecasted value

In order to rank the models considering all time series, we have conducted a Friedman and Nemenyi hypothesis test [38] on the results obtained with respect to the scale free SMAPE measure. Friedman's test is being for ranking the models and upon rejection of null hypothesis a Nemenyi hypothesis test is conducted [39]. Table 8 presents the test results. It is seen that the proposed model

Journal of Theoretical and Applied Information Technology
15th November 2018. Vol.96. No 21
© 2005 – ongoing JATIT & LLS

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

has the lowest Rank SMAPE and gained the first rank among all the models which indicate statistical superiority of proposed model. For a clear understanding of closeness of forecasted values to that of actual values, we have plotted comparison graphs (Figure 3 - Figure 11) for some of the time series data considered in this study.

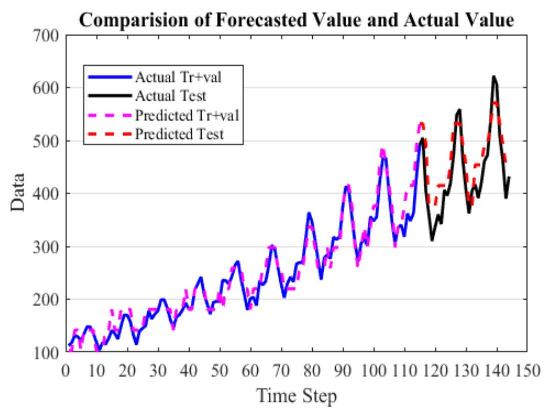

Figure:10 Passenger Time Series Actual and Forecasted value

Table 8: Ranking of all models considering SMAPE

Model	Rank SMAPE	Mean Rank
Proposed	110.06	1
Aladag[21]	117.36	2
Egrioglu [22]	129.07	3
Yolcu [24]	140.68	4

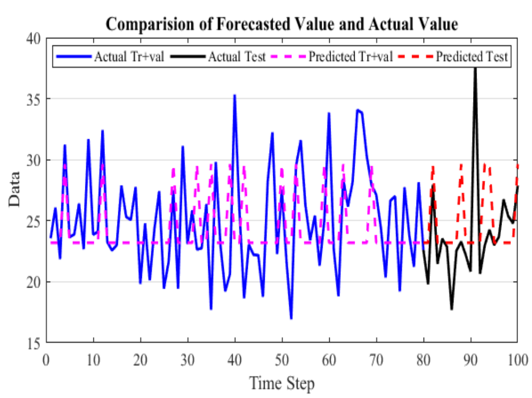

Figure:11 Rainfall Time Series Actual and Forecasted value

6. CONCLUSION

Fuzzy time series forecasting is an interesting and challenging task since it provides better forecasting accuracy by implicitly handling uncertainty. In this study, we made an attempt to develop a high order fuzzy time series forecasting model by considering the issues found in all the four steps of fuzzy TSF. Since the order of the model significantly affects the performance, the order of the model for a time series is determined by using ACF and PACF function of the fuzzified series. In order to determine the length of interval, computationally efficient single-variable constrained optimization based method is employed. In addition, a one-pass fast learning GRNN model is used to represent the fuzzy logical relationships. Ten time series data sets are considered for evaluating the efficiency of the proposed model. In addition to the proposed model, three recent fuzzy forecasting models Aladag et al.[21], Egrioglu et al. [22] and Yolcu et al. [24] are implemented. Each model is applied on every time series dataset for fifty different times and forecast accuracy is measured. The proposed model statistically outperformed all other models in four time series datasets considering both RMSE and SMAPE measure. The proposed model achieved statistically significantly better RMSE than Aladag[21] in four time series datasets, Egrioglu [22] in nine time series datasets and Yolcu [24] in six time series datasets. The proposed model achieved statistically significantly better SMAPE than Aladag[21] in six time series datasets, Egrioglu [22] in nine time series datasets and Yolcu [24] in six time series datasets. Moreover, the proposed model only provided statistically poor performance than each of the considered model in one dataset. Furthermore, to identify the best model considering all time series datasets, a Friedman and Nemenyi hypothesis test is being conducted on the obtained SMAPE measures. The test results revealed that the proposed model has the lowest rank SMAPE and acquired the best rank. From experimental results it can be concluded that the proposed model is statistically superior in terms of forecasting accuracy when compared with other alternatives considered in this study. Additionally, it is computationally very efficient. This can be understood from the fact that, in addition to the use of one-pass fast learning GRNN model for representing FLR, we have used the computationally efficient method in every step found in the literature so far.

Recent studies [24-25] claimed that the use of membership value while representing FLR using ANN improves the forecasting accuracy. In this paper, we have implemented Yolcu et al. [24] method which provided the best SMAPE for three time series datasets (TAIEX, Acres Burned, Rainfall) while provided the worst SMAPE as well as RMSE in six time series datasets(Car, Accidental, Gas, Pollution, milk and Passenger). It can be observed from the plots of these six time series that all these six time series have seasonality and trend component. Therefore, it can be concluded from simulation results that, one should avoid considering membership value while modeling FLR using ANN when the time series has both trend and seasonality component.

Although the proposed method provided the best forecasting performance considering all time series, it has a limitation. It uses computationally efficient ACF and PACF to identify the significant number of inputs to GRNN. However, ACF and PACF only capture linear relationship existing in time series [42]. Therefore, in future one may employ the computationally expensive genetic programming technique [42] which has the ability to handle complex nonlinear relationship existing in time series so that better forecasting accuracy may be achieved. One may also use deep learning methods instead of GRNN model to boost the performance of the proposed model without hampering the computational efficiency.

REFRENCES:

[1] Qiang Song and Brad S. Chissom, "Forecasting Enrollments with fuzzy time series – Part I", Fuzzy Sets and Systems, Vol. 54, No. 1, 1993, pp.1-9.

[2] Qiang Song and Brad S. Chissom, "Fuzzy time series and its model", Fuzzy Sets and Systems, Vol. 54, No. 3, 1993, pp. 269-277.

[3] Kunhuang Huarng, "Effective lengths of intervals to improve forecasting in fuzzy time series", Fuzzy Sets and Systems, Vol. 123, No. 3, 2001, pp. 387–394.

[4] Kunhuang Huarng and Tiffany Hui-Kuang Yu, "Ratio-Based lengths of intervals to improve Fuzzy Time Series Forecasting", IEEE Transactions On Systems, Man, And Cybernetics—Part B, Cybernetics, Vol. 36, No. 2, 2006, pp. 328-340.

[5] Ufuk Yolcu, Erol Egriogl, Vedide R.Uslu,Murat A.Basaran and Cagdas H.Aladag, "A new approach for determining the length of intervals for fuzzy time series", Applied Soft Computing, Vol. 9, No. 2, 2009, pp. 647–651.

[6] I-Hong Kuo, Shi-Jinn Horng, Tzong-Wann Kao, Tsung-Lieh Lin, Cheng-Ling Lee, and Yi Pan, "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.

[7] I-Hong Kuo, Shi-Jinn Horng, Yuan-Hsin Chen, Ray-Shine Run, Tzong-Wann Kao, Rong-Jian Chen, Jui-Lin Lai, and Tsung-Lieh Lin, "Forecasting TAIFEX based on fuzzy time series and particle swarm optimization", *Expert Systems with Applications,* Vol. 37, No. 2, 2010,pp. 1494-1502.

[8] Shyi-Ming Chen and Pei-Yuan Kao, "TAIEX forecasting based on fuzzy time series, particle swarm optimization techniques and support vector machines", *information Sciences,* Vol. 247, 2013, pp. 62-71.

[9] Li-Wei Lee, Li-Hui Wang and Shyi-Ming Chen, "Temperature prediction and TAIFEX forecasting based on high-order fuzzy logical relationships and genetic simulated annealing techniques" *Expert Systems with Applications,* Vol. 34, No. 1, 2008, pp. 328-336.

[10] Li-Wei Lee, Li-Hui Wang and Shyi-Ming Chen, "Temperature prediction and TAIFEX forecasting based on fuzzy logical relationships and genetic algorithms", *Expert Systems with Applications,* Vol. 33, No. 3, 2007, pp. 539-550.

[11] Vedide Rezan Uslu, Eren Bas, Ufuk Yolcu and Erol Egrioglu, "A fuzzy time series approach based on weights determined by the number of recurrences of fuzzy relations", *Swarm and Evolutionary Computation,* Vol. 15, 2014, pp. 19-26.

[12] Qisen Cai, Defu Zhang, Wei Zheng and Stephen CH Leung, "A new fuzzy time series forecasting model combined with ant colony optimization and auto-regression", *Knowledge-Based Systems,* Vol. 74, 2015, pp.61-68.

[13] Narges Shafaei Bajestani and Asef Zare, "Application of optimized type 2 fuzzy time series to forecast Taiwan stock index", In *Computer, Control and Communication, 2009. IC4 2009. 2nd International Conference on computer,control and communication*, IEEE, 2009, pp. 1-6.

[14] Narges Shafaei Bajestani and Assef Zare, "Forecasting TAIEX using improved type 2 fuzzy time series", *Expert Systems with Applications,* Vol. 38, No. 5, 2011, pp. 5816-5821.

[15] Wenyu Zhang, Shixiong Zhang, Shuai Zhang, Dejian Yu and NingNing Huang, "A multi-factor and high-order stock forecast model based on Type-2 FTS using cuckoo search and self-adaptive harmony search", *Neurocomputing,* Vol. 240, 2017,pp. 13-24.

[16] Mehdi Khashei, Seyed Reza Hejazi and Mehdi Bijari, "A new hybrid artificial neural networks and fuzzy regression model for time series

forecasting", *Fuzzy sets and systems,* Vol. 159, No. 7, 2008, pp. 769-786.

[17] Cagdas H. Aladag, Murat A. Basaran, Erol Egrioglu, Ufuk Yolcu and Vedide R. Uslu, "Forecasting in high order fuzzy times series by using neural networks to define fuzzy relations", *Expert Systems with Applications,* Vol.36, No. 3,2009, pp. 4228-4231.

[18] Erol Egrioglu, Cagdas Hakan Aladag, Ufuk Yolcu, Vedide R. Uslu and Murat A. Basaran, "A new approach based on artificial neural networks for high order multivariate fuzzy time series", *Expert Systems with Applications,* Vol. 36, No. 7, 2009, pp. 10589-10594.

[19] Cagdas Hakan Aladag, Ufuk Yolcu and Erol Egrioglu, "A high order fuzzy time series forecasting model based on adaptive expectation and artificial neural networks", *Mathematics and Computers in Simulation,* Vol.81, No. 4,2010, pp. 875-882.

[20] Tiffany Hui-Kuang Yu and Kun-Huang Huarng, "A neural network-based fuzzy time series model to improve forecasting", *Expert Systems with Applications,* Vol. 37, No. 4, 2010, pp. 3366-3372.

[21] Cagdas Hakan Aladag, "Using multiplicative neuron model to establish fuzzy logic relationships", *Expert Systems with Applications,* Vol. 40, No. 3,2013, pp. 850-853.

[22] Erol Egrioglu, Cagdas Hakan Aladag and Ufuk Yolcu, "Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks", *Expert Systems with Applications,* Vol. 40, No. 3, 2013, pp. 854-857.

[23] Mu-Yen Chen, "A high-order fuzzy time series forecasting model for internet stock trading", *Future Generation Computer Systems,* Vol. 37, 2014, pp. 461-467.

[24] Ozge Cagcag Yolcu, Ufuk Yolcu, Erol Egrioglu and C. Hakan Aladag, "High order fuzzy time series forecasting method based on an intersection operation", *Applied Mathematical Modelling,* Vol. 40, No. 19-20, 2016, pp. 8750-8765.

[25] Ozge Cagcag Yolcu and Hak-Keung Lam, "A combined robust fuzzy time series method for prediction of time series", *Neurocomputing,* Vol. 247, 2017, pp. 87-101.

[26] Milad Avazbeigi, Seyed Hossein Hashemi Doulabi, and Behrooz Karimi, "Choosing the appropriate order in fuzzy time series: A new N-factor fuzzy time series for prediction of the auto industry production", *Expert Systems with Applications, Vol.* 37, No. 8, 2010, pp. 5630-5639.

[27] Shou-Hsiung Cheng, Shyi-Ming Chen and Wen-Shan Jian, "Fuzzy time series forecasting based on fuzzy logical relationships and similarity measures", *Information Sciences, Vol.* 327, 2016, pp. 272-287.

[28] Cagdas Hakan Aladag, Ufuk Yolcu, Erol Egrioglu and Ali Z. Dalar, "A new time invariant fuzzy time series forecasting method based on particle swarm optimization", *Applied Soft Computing,* Vol. 12, No. 10, 2012, pp. 3291-3299.

[29] Hia Jong Teoh, Tai-Liang Chen, Ching-Hsue Cheng and Hsing-Hui Chu, "A hybrid multi-order fuzzy time series for forecasting stock markets", *Expert Systems with Applications,* Vol. 36, No. 4,2009, pp. 7888-7897.

[30] Simon Haykin, "*Neural networks: a comprehensive foundation*", Prentice Hall PTR, 1994.

[31] Donald F Specht, "A general regression neural network", *IEEE transactions on neural networks,* Vol. 2, No. 6, 1991, pp. 568-576.

[32] Hafizah Husain, Marzuki Khalid and Rubiyah Yusof, "Automatic clustering of generalized regression neural network by similarity index based fuzzy c-means clustering", In *TENCON 2004. 2004 IEEE Region 10 Conference*, Vol. 2, Nov. 2004, pp. 302-305.

[33] Qiang Song, "A note on fuzzy time series model selection with sample autocorrelation functions", *Cybernetics &Systems, Vol.* 34, No. 2,2003, pp. 93-107.

[34] Sibarama Panigrahi and H. S. Behera, "Effect of Normalization Techniques on Univariate Time Series Forecasting using Evolutionary Higher Order Neural Network", *Int. J. Eng. Adv. Technol,* Vol. 3, No. 2, 2013, pp. 280-285.

[35] Sibarama Panigrahi, Y. Karali and H. S. Behera, "Normalize time series and forecast using evolutionary neural network" *Int. J. Eng. Res. Technol,* Vol. 2, No. 9, 2013, pp. 2518-2522.

[36] Rob J Hyndman, "Time Series Data Library", 2010. URL ⟨http://data.is/TSDLdemo⟩.

[37] Rob J. Hyndman and Anne B. Koehler, "Another look at measures of forecast accuracy", *International journal of forecasting,* Vol. 22, No. 4, 2006, pp. 679-688.

[38] Myles Hollander, Douglas A. Wolfe, Eric Chicken, "Nonparametric statistical methods,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Nonparametric Statistical Methods", John Wiley& Sons, Hoboken, NJ, (1999).

[39] Demsar, Janez, "Statistical comparisons of classifiers over multiple data sets", *Journal of Machine learning research,* Vol. 7, 2006, pp. 1-30.

[40] Fang-Mei Tseng, Gwo-Hshiung Tzeng, Hsiao-Cheng Yu, and Benjamin JC Yuan, "Fuzzy ARIMA model for forecasting the foreign exchange market", *Fuzzy sets and systems,* Vol. 118, No. 1, 2001, pp. 9-19.

[41] H.D. Tran, N. Muttil, B.J.C. Perera, "Selection of significant input variables for time series forecasting", *Environmental Modelling & Software*, Vol. 64, 2015, pp. 156-163.