Appl Intell
DOI 10.1007/s10489-017-0962-8

Recurrent type-1 fuzzy functions approach for time series forecasting

Nihat Tak[1] · **Atif A. Evren**[2] · **Mujgan Tez**[3] · **Erol Egrioglu**[4]

Abstract Forecasting the future values of a time series is a common research topic and is studied using probabilistic and non-probabilistic methods. For probabilistic methods, the autoregressive integrated moving average and exponential smoothing methods are commonly used, whereas for non-probabilistic methods, artificial neural networks and fuzzy inference systems (FIS) are commonly used. There are numerous FIS methods. While most of these methods are rule-based, there are a few methods that do not require rules, such as the type-1 fuzzy function (T1FF) approach. While it is possible to encounter a method such as an autoregressive (AR) model integrated with a T1FF, no method that includes T1FF and the moving average (MA) model in one algorithm has yet been proposed. The aim of this study is to improve forecasting by taking the disturbance terms into account. The input dataset is organized using the following variables. First, the lagged values of the time series are used for the AR model. Second, a fuzzy c-means clustering algorithm is used to cluster the inputs. Third, for the MA, the residuals of fuzzy functions are used. Hence, AR, MA, and the degree of memberships of the objects are included in the input dataset. Because the objective function is not derivative, particle swarm optimization is preferable for solving it. The results on several datasets show that the proposed method outperforms most of the methods in literature.

Keywords Type-1 fuzzy functions · Nonlinear time series · Forecasting · Autoregressive model · Moving average model · Particle swarm optimization

✉ Nihat Tak
nihattak@gmail.com

Atif A. Evren
aevren2006@gmail.com

Mujgan Tez
mujgantez@gmail.com

Erol Egrioglu
erole1977@yahoo.com

1 Faculty of Economics and Administrative Sciences, Kirklareli University, Kirklareli, Turkey

2 Faculty of Arts and Sciences, Yildiz Technical University, Istanbul, Turkey

3 Faculty of Arts and Sciences, Marmara University, Istanbul, Turkey

4 Faculty of Arts and Sciences, Giresun University, Giresun, Turkey

1 Introduction

A time series is defined as a variable whose observations are determined over several time intervals. This time interval can be specified as hourly, daily, weekly, monthly, seasonally, yearly, or other lengths. To forecast time series datasets, numerous methods have recently been developed. These are called time series forecasting methods. Time series forecasting methods can be divided into two categories: probabilistic models and non-probabilistic models.

Probabilistic or stochastic methods, which are also called traditional methods, make some assumptions about the time series. Stationarity is an important assumption in probabilistic time series forecasting methods. This assumption states that the time series has a constant mean, variance, and covariance function. One of the most widely used models

Published online: 17 June 2017

for probabilistic time series forecasting is the autoregressive integrated moving average (ARIMA) model, which was organized systematically to obtain the best ARIMA model parameters by Box-Jenkins [7].

However, ARIMA models assume a linear structure for the time series values. Hence, ARIMA models are not capable of dealing with a time series that has a nonlinear structure. Because most time series datasets do not satisfy the assumption of stationarity or linearity, many researchers have studied alternative methods. Of the alternative methods, artificial neural networks (ANNs) and fuzzy inference systems (FISs) are the most widely used ones. After fuzzy set theory was proposed by Zadeh [45], Zadeh introduced another paper on linguistic variables and fuzzy systems [46]. Later, several researchers combined fuzzy set theory with inference systems, called FISs. Well-known FISs include the FIS proposed by Mamdani and Assilian [34], the FIS proposed by Takagi and Sugeno [41], the adaptive neuro FIS, proposed by Jang [28], and the type-1 fuzzy function (T1FF) proposed by Turksen [42].

While the systems proposed by Mamdani and Assilian [34], and Takagi and Sugeno [41] are rule based, the system proposed by Turksen [42] is not rule based. The T1FF approach has an advantage over rule-based systems because the detection of rules is a difficult problem to tackle. Originally, FIS was not designed for time series forecasting problems but for classification problems.

Fuzzy time series forecasting methods were first proposed by Song and Chisom [37] in 1993. Song and Chisom [38, 39] defined a fuzzy time series forecasting process that consists of three stages: fuzzification, determination of the rules, and defuzzification. After Song and Chisom [38, 39] defined the fuzzy-time-series forecasting process, FIS-based time series forecasting methods were expanded by the following researchers. Chen [17] proposed a high-order fuzzy time series model for forecasting problems in 2002. ANN was used to forecast fuzzy time series by Huarng and Yu [25]. These two studies, among others, aimed to determine the fuzzy relations of fuzzy time series to improve forecasts.

Some of the studies using artificial intelligence techniques and the fuzzy set theory are listed below. Genetic algorithms were employed for time series forecasting problems by Kuo et al. [32], Chen and Chung [18], Kim and Kim [30], Egrioglu [20], and Bas et al. [4]. Multivariate fuzzy time series forecasting methods were studied by Egrioglu et al. [22], Chen and Tanuwijaya [19], Jilani et al. [29], and Huarng [24]. The particle swarm optimization (PSO) algorithm in fuzzy time series methods was employed by Chau [13], Park et al. [36], Kuo et al. [31], Aladag et al. [1], and Huang et al. [23].

On the one hand, probabilistic or linear models can deal with time series when the linear part of the time series dominates the non-linear part. On the other hand, when the non-linear part of the time series dominates the linear part, non-linear models produces acceptable outcomes. Thus, hybrid models have been developed to deal with both the linear and non-linear parts of the time series. The seasonal ARIA (SARIMA) model and multilayer perceptron ANN (MLP-ANN) were hybridized by Tseng et al. [43]. Various hybrid methods have been introduced by Bas et al. [3], Lee and Tong [33], Chen and Wang [16], Pai and Lin [35], Zhang [47], BuHamra et al. [8], Jain and Kumar [26], and Yolcu et al. [44].

Time series forecasting methods based on fuzzy inference systems have been studied by Catalao et al. [9], Chabaa et al. [11], Chang [12], Chen and Ma [14], Chen and Zhang [15], and Egrioglu et al. [21]. The fuzzy function approach was first used in time series forecasting by Beyhan and Alici [5]. Later, Aladag et al. [2] studied the T1FF approach for time series forecasting.

The methods that have been proposed so far, in terms of the T1FF approach, have not included disturbance terms as inputs. The objective of this paper is to propose a new method that takes into account the disturbance terms to obtain better forecasts. The proposed method combines an autoregressive (AR) model, moving average (MA) model, and T1FF approach into one algorithm. Disturbance terms are determined using the T1FF residuals. To minimize the sum of squared error (SSE), PSO is used. This paper is organized as follows. The algorithm and flowchart of the proposed method are given in Section 2. In Section 3, several applications are presented to evaluate the performance of the proposed method. Finally, the conclusions are discussed in Section 4.

2 Proposed method

The T1FF approach was designed as an FIS in 2008 by Celikyilmaz and Turksen [10]. Because the classic FISs are rule based, Turksen [42] proposed the T1FF approach to address the need for a non-rule based system. T1FF has an advantage over classic FIS because an expert opinion is needed for defining rules. Originally, T1FF was proposed for classification and regression problems. Therefore, it needed to be redesigned. The T1FF approach was first adapted to time series by Beyhan and Alici [5] and later, Aladag et al. [2] adapted T1FF to time series forecasting problems. Beyhan and Alici [5] used an auto-regressive with exogenous input (ARX) model structure that was not able to search for the best model. Therefore, Aladag et al. [2]

proposed a fuzzy time series function method to search for the best model by adapting an AR model into their algorithm. Eventually, they obtained better forecasting results than Beyhan and Alici. AR and MA are the most important models in the probabilistic time series approaches. Aladag et al. [2] used an AR model in their approach. In the proposed method, the MA model is also taken into consideration. Thus, the proposed method combines the autoregressive moving average model (ARMA) with T1FF. The lagged values of the time series, lagged values of the disturbance terms, and degree of memberships obtained from the fuzzy c-means (FCM) clustering method are taken as inputs. Disturbance terms are obtained from residuals of the fuzzy functions. Because the objective function of the fuzzy functions is not derivative, the PSO algorithm is adopted.

The initial difference between a statistical approach and the proposed approach lies in the assumption of the models. A statistical approach assumes that the underlying structure of the disturbances are normally distributed and have constant covariance. However, there is no assumption about the disturbances in the proposed approach. Moreover, the most important contribution of the proposed method is that it takes the degree of membership into account.

2.1 Algorithm

A flowchart of the proposed method is shown in Fig. 1, and its steps are detailed as follows.

- Step 1. The dataset is partitioned into two datasets: a training dataset and a test dataset. To partition the dataset, a block partitioning structure is used because of the data dependency. Assume the size of the dataset is n. Then, the first k observations are chosen for the training dataset and the rest of the observations (n − k) are chosen for the test dataset.
- Step 2. The model inputs are selected as lagged variables of the dataset and disturbances. Inputs are clustered using the FCM algorithm.
- Step 3. Lagged variables, membership degrees, and the functions of the membership degrees are combined into the training dataset to obtain input matrix X. The dimension of the input matrix is $n \times p \times c \times k$, where n is the number of observations, p is the number of parameters, c is the number of clusters, and k is the number of particles.
- Step 4. Coefficients c_1 and c_2, the number of particles, and the number of iterations are specified for the PSO algorithm. The number of positions in each particle is $(p + q + 4)c$, where p is the lag for AR, q is the lag MA, and c is the number of clusters.

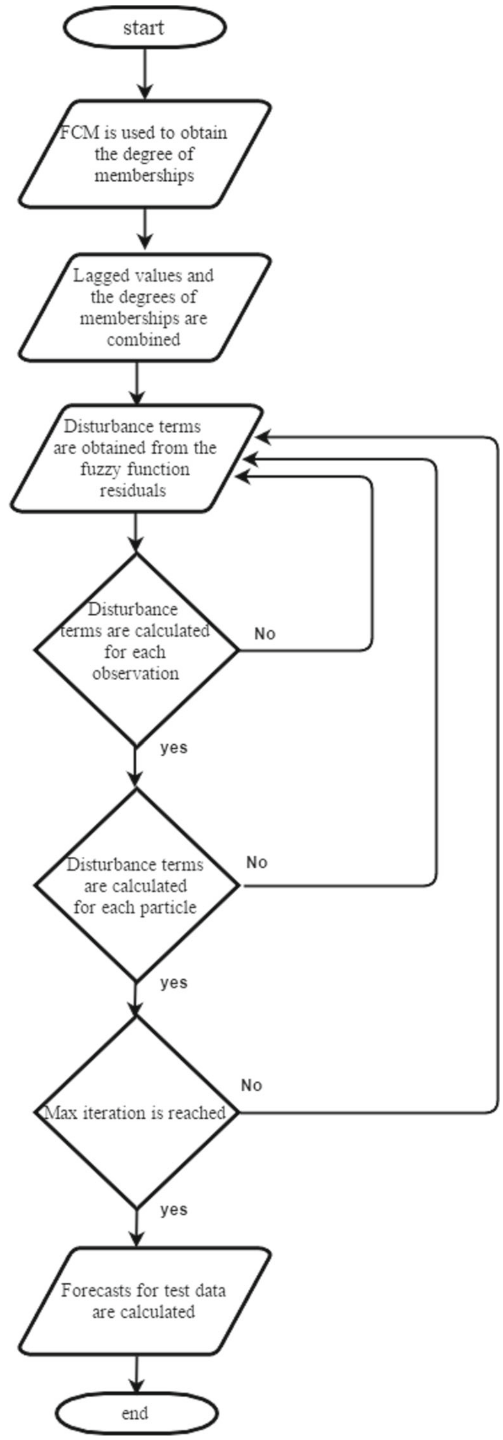


Fig. 1 Flowchart of the proposed method

- Step 5. Initial positions and corresponding velocities for each particle are generated randomly from a standard normal distribution. The initial personal best (pbest) values are assigned as initial positions. The initial global best (gbest) value is obtained using a fitness value.

- Step 6. The values for the i^{th} particle and the first observation e_{t-q} are calculated as follows.

$$\widehat{Y}_i^{(j)(k)} = X_{(i)(.)}^{(j)(k)} \beta^T{}_{(i)(.)}^{(j)(k)} \tag{1}$$

$$\widehat{Y^*}^{(j)(k)} = \widehat{Y}_{(i)}^{(j)(k)} \mu^T{}_{(i)(.)}^{(j)} \tag{2}$$

$$e_k^i = Y_i - \widehat{Y^*}_i^{(k)} \tag{3}$$

$$X_{(i+1)(p)}^{(.)(k)} = e_k^i \tag{4}$$

Equation (1) is calculated for each cluster j and the values of $\widehat{Y}_i^{(j)(k)} = [.........]$ are obtained for the number of clusters c.

In (2), $\widehat{Y^*}_i^{(k)}$ is calculated, where $\widehat{Y}_{(i)}^{(j)(k)}$ are the forecasts and $(\mu^T)_{(i)(.)}^{(j)}$ are the corresponding membership degrees.

In (3), the e_k^i values are obtained, where Y_i is the actual value and e_k^i is the same value for each cluster.

In (4), e_k^i value is assigned to input matrix X for each cluster.

- Step 7. Because a disturbance term is calculated for the first observation and the i^{th} particle in Step 6, (1)–(4) are repeated for each observation.
- Step 8. Steps 6 and 7 are repeated for each particle.
- Step 9. The pbest and gbest values are updated using the fitness value.
- Step 10. The gbest value obtained for the training dataset is used to determine e_t for the test dataset. The equations are given below.

$$\widehat{Y_{test_i}}^{(j)} = X_{test_{i,.}}^{(j)} \beta_{test\,i,.}^{T\ (j)} \tag{5}$$

$$\widehat{Y_{test_i}}^{*} = \widehat{Y_{test_i}}^{(j)} \mu_{test\,i,.}^{T\ (j)} \tag{6}$$

$$e_{test_i} = Y_{test_i} - \widehat{Y_{test_i}}^{*} \tag{7}$$

$$X_{test_{(i+1)(p)}}^{(.)} = e_{test_i} \tag{8}$$

Equation (5) is repeated for each cluster, then (6)–(8) are calculated. These steps are repeated for each observation.

- Step 11. The values for r_1 and r_2 are randomly generated from the standard normal distribution and new positions and velocities are updated as follows.

$$v_{id}^{k+1} = v_{id}^k + c_1 r_1^k (pbest_{id}^k - p_{id}^k) + c_2 r_2^k (gbest^k - p_{id}^k) \tag{9}$$

$$p_{id}^{k+1} = p_{id}^k + v_{id}^{k+1} \tag{10}$$

Table 1 Summary of the datasets and parameter selection criteria

	Series/year	Obs. number	Order of AR	Order of MA	Clusters numbers	ntest
1	ABC	147	1-10	1-2	2-10	16
2	BIST100/2009	103	1-5	1-2	2-5	7,15
3	BIST100/2010	104	1-5	1-2	2-5	7,15
4	BIST100/2011	106	1-5	1-2	2-5	7,15
5	BIST100/2012	106	1-5	1-2	2-5	7,15
6	BIST100/2013	106	1-5	1-2	2-5	7,15
7	TAIEX/1999	266	1-5	1-2	2-5	45
8	TAIEX/2000	271	1-5	1-2	2-5	47
9	TAIEX/2001	244	1-5	1-2	2-5	43
10	TAIEX/2002	248	1-5	1-2	2-5	43
11	TAIEX/2003	249	1-5	1-2	2-5	43
12	TAIEX/2004	250	1-5	1-2	2-5	45

- Step 12. Steps 6–11 are repeated for each iteration up to the maximum number of iterations.
- Step 13. The final gbest value is used to forecast the future values of the time series using the equations below.

$$\widehat{Y_{test_i}}^{(j)} = X_{test_{i,.}}^{(j)} \beta_{test\,i,.}^{T\ (j)} \tag{11}$$

$$\widehat{Y_{test_i}}^{*} = \widehat{Y_{test_i}}^{(j)} \mu_{test\,i,.}^{T\ (j)} \tag{12}$$

3 Evaluation

The proposed method's performance was evaluated as follows. First, some initial possible solutions are generated from the standard normal distribution to determine the coefficients of the models. The disturbance terms are calculated recursively, observation by observation, for each model. Third, the models are evaluated in terms of SSE and the best solution (gbest) of the possible initial solutions is chosen. In short, the evaluation of the models in the proposed method is made sample by sample.

Table 2 Computation times of the BIST100 datasets

Dataset	ntest=7	ntest=15
BIST100/2009	2.64 sn	4.95 sn
BIST100/2010	6.45 sn	9.02 sn
BIST100/2011	7.31 sn	5.72 sn
BIST100/2012	6.80 sn	3.57 sn
BIST100/2013	8.72 sn	3.51 sn

Table 3 Computation times of the ABC and TAIEX datasets

Dataset	ntest	Elapsed
TAIEX1999	45	8.04 sn
TAIEX2000	47	11.18 sn
TAIEX2001	43	10.46 sn
TAIEX2002	43	5.73 sn
TAIEX2003	43	7.65 sn
TAIEX2004	45	12.54 sn
ABC	16	10.02 sn

To evaluate the performance of the proposed method, 12 real world time series datasets were analyzed using R, a statistical programming language. The first dataset is the Australian Beer Consumption (ABC) dataset [27]. The observations of this dataset were observed for each quarter from 1956 to 1994. The next five datasets are from the Istanbul Stock Exchange (BIST100) datasets [6]. The data of the BIST100 datasets were observed daily for the first half of the years between 2009 and 2013. The data of the Taiwan Stock Exchange (TAIEX) dataset [40] were observed daily from 1999 to 2004. These datasets were chosen so that the performance of the proposed method could be compared with other methods that previously used the same datasets. The methods are evaluated using the root mean squared error (RMSE) and mean absolute percentage error (MAPE), calculated as follows.

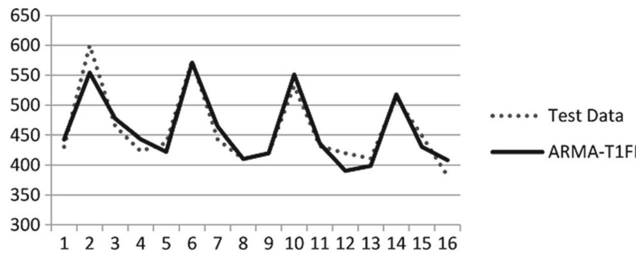


Fig. 2 Forecasts and test data for the ABC dataset (ntest =16)

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(x_t - \widehat{x}_t)^2} \tag{13}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{x_t - \widehat{x}_t}{x_t}\right| \tag{14}$$

The number of observations of the original datasets, number of observations of the test datasets, order of the AR model, order of the MA model, and number of clusters are listed in Table 1. ARMA -T1FF refers to the proposed method.

The complexity of the proposed method is difficult to compute. However, the calculation time for each dataset is given Tables 2 and 3. The calculations were computed on a PC equipped with an I7 CPU, 8 GB RAM, and 512 GB SSD HDD.

Table 4 Results obtained for the ABC test dataset when ntest=16

Test data	WMES	SARIMA	MLP-ANN	ANFIS	MANFIS	ARMA-T1FF*
430.50	453.91	452.72	453.88	446.71	445.23	442.82
600.00	575.22	578.29	557.81	553.73	575.63	554.44
464.50	502.32	487.71	497.52	482.07	494.07	477.33
423.60	444.73	446.28	437.39	434.19	434.56	443.47
437.00	459.66	456.77	449.01	438.55	444.69	422.46
574.00	582.48	583.51	569.01	559.01	575.42	571.21
443.00	508.64	492.13	471.08	472.52	481.28	463.85
410.00	450.31	450.36	424.33	427.57	414.44	410.47
420.00	465.40	461.01	448.87	445.01	430.31	420.02
532.00	589.74	588.96	560.04	562.94	565.18	551.34
432.00	514.96	496.77	447.01	459.14	452.05	436.37
420.00	455.89	454.64	408.64	416.16	392.14	390.61
411.00	471.15	465.46	428.11	431.71	419.33	398.69
512.00	597.00	594.71	537.69	544.98	536.88	517.62
449.00	521.28	501.67	438.43	444.31	446.32	430.76
382.00	461.46	459.17	420.58	426.01	406.64	408.27
RMSE	53.33	47.04	24.11	25.05	21.37	19.21*
MAPE	0.1072	0.0949	0.0476	0.0467	0.0401	0.0333*

3.1 ABC dataset

In the first evaluation, the ABC dataset was used. This dataset consists of 148 observations that were observed quarterly from 1956 to 1994. To compare the results of the proposed method, the performances of the SARIMA model,

Table 5 Results obtained for TAIEX

Methods	1999	2000	2001	2002	2003	2004	Mean
Chen (1996)	120	176.32	147.84	101.18	74.46	84.28	117.34
Chen et al. (2010)	101.97	129.42	113.33	66.82	53.51	60.48	87.58
Chen et al. (2011)	112.47	123.62	115.33	71.01	58.06	57.73	89.7
Chen et al. (2012)	99.87	119.98*	114.47	67.17	52.49	52.27*	84.37
ANFIS (1993)	101.16	137.02	114.72	65.99	57.04	61.36	89.54
MANFIS (2014)	101.94	124.92	112.47	62.57*	52.33*	53.66	84.64
ARMA-T1FF*	98.33*	128.18	106.48*	65.14	52.38	53.78	84.05*

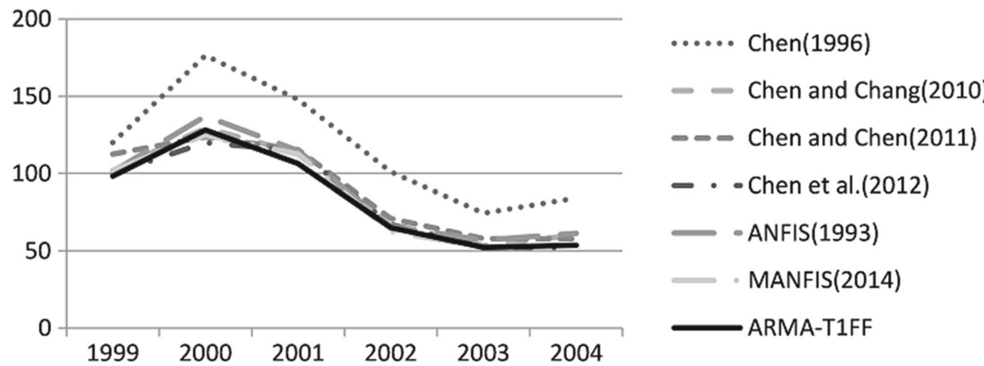


Fig. 3 RMSE values of the proposed method and the other methods

Table 6 Parameter specifications for the best model

Year	Number of clusters	Order of MA	Order of AR	ntest
2009	2	2	2	7
2009	3	2	2	15
2010	4	2	2	7
2010	5	1	1	15
2011	5	1	3	7
2011	3	1	2	15
2012	4	2	3	7
2012	3	1	2	15
2013	5	1	2	7
2013	2	1	1	15

Table 8 MAPE values for BIST100 when ntest=7

Year	ARIMA	ES	MLP-ANN	T1FF	FTS-N	ARMA-T1FF*
2009	0.0087	0.0087	0.0083	0.0101	0.0058	0.00541
2010	0.0183	0.0185	0.0143	0.0179	0.0159	0.0157
2011	0.0144	0.0144	0.0128	0.0153	0.0105	0.0079
2012	0.0084	0.0084	0.0111	0.0162	0.0084	0.0085
2013	0.0116	0.0116	0.0109	0.0131	0.0065	0.0058
Mean	0.01228	0.01232	0.01148	0.01452	0.00942	0.008662

Table 7 RMSE values for BIST100 when ntest=7

Year	ARIMA	ES	MLP-ANN	T1FF	FTS-N	ARMA-T1FF*
2009	344.91	344.93	325.1011	445.5147	266.6011	235.96
2010	1221	1208.1	1077.4	1179.9	1049.5	1057.097
2011	1057.6	1057	919.9204	1083.2	765.07	714.1724
2012	650.56	650.7387	774.6103	1034.2	590.3545	547.13
2013	1361.6	1361.6	1314.9	1511.6	786.13	783.9803
Mean	927.134	924.4737	882.3864	1050.883	691.5311	667.6679

Table 9 RMSE values for BIST100 when ntest=15

Year	ARIMA	ES	MLP-ANN	T1FF	FTS-N	ARMA-T1FF*
2009	540.2	540.2087	525.7264	534.1345	514.5627	478.1365
2010	1611.5	1611.5	1603	1852	1357.4	1332.159
2011	1129.6	1129.7	1095.7	1145.6	916.5411	1017.41
2012	620.7892	620.829	783.3547	1037.6	581.71	529.69
2013	1268.7	1268.7	1232.5	1278.6	1207.9	1159.598
Mean	1034.16	1034.188	1048.056	1169.587	915.6228	903.3987

Table 10 MAPE values for BIST100 when ntest=15

Year	ARIMA	ES	MLP-ANN	T1FF	FTS-N	ARMA-T1FF*
2009	0.012	0.012	0.0114	0.0438	0.0112	0.0093
2010	0.0220	0.022	0.0220	0.0264	0.0202	0.019
2011	0.015	0.015	0.0146	0.0156	0.0121	0.0134
2012	0.0088	0.0088	0.0117	0.0161	0.0087	0.0076
2013	0.0109	0.0109	0.0107	0.0108	0.0106	0.0106
Mean	0.013737	0.01374	0.014076	0.02254	0.01256	0.01198

Winter's Exponential Smoothing (WMES) method, MLP-ANN, adaptive neuro fuzzy inference system (ANFIS), and modified ANFIS (MANFIS) from the study of Egrioglu et al. [21] are used.

The length of the test dataset (ntest) is taken as 16. The algorithm searched for the best model when the number of clusters was varied from two to ten, the AR lag was varied one to ten, and the MA lag was varied from one to two. The number of particles and iterations were set to 35 and 100, respectively. Under these conditions, the minimum RMSE and MAPE values were obtained for five clusters, seven AR lags, and one MA lag. Looking at the RMSE and MAPE values in Table 4, it is obvious that the minimum RMSE and MAPE values are obtained by the proposed method. The forecasts obtained by the proposed method and the original observations are given in Fig. 2.

3.2 TAIEX dataset

To evaluate the performance of the proposed method for TAIEX dataset, the following methods were chosen: Chen (1996), Chen and Chang (2010), Chen and Chen (2011), and Chen et al. (2012). The results of the methods in Table 5 are taken from the paper of Bas et al. [3]. For the TAIEX datasets from 1999 to 2004, the results obtained by the proposed method are listed in Table 5.

The forecasting results from the proposed and other methods are compared in terms of the RMSE. For 1999 and 2001, the best forecasting results are obtained from the proposed method. For 2002 and 2003, MANFIS outperforms the other methods, and for 2004, Chen et al.'s method gives the best results. However, looking at the means of the years, it is obvious that the proposed method has better forecasting results than others. A comparison of the results of the proposed method with those of other methods by year is given in Fig. 3.

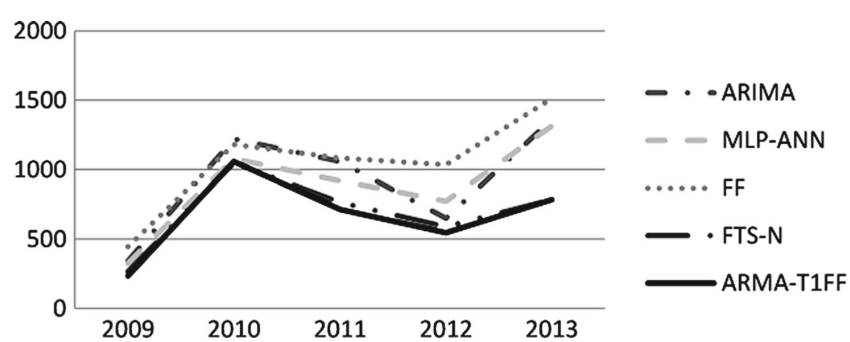
3.3 BIST100 dataset

For the BIST100 dataset, the outcomes of ARIMA, exponential smoothing (ES), MLP-ANN, T1FF, fuzzy time series network (FTS-N), and recurrent T1FF (ARMA-T1FF) are compared. The outcomes for ARIMA, ES, MLP-ANN, T1FF, and FTS-N are taken from Bas et al. [3].

The best results are obtained using the Box-Jenkins procedure for the ARIMA procedure. For the ES procedure, the best results are obtained from Holt and Winter's method. For the MLP-ANN method, the hidden layer neurons and number of inputs were varied from 1 to 5 and the best model of these was chosen. While performing T1FF, the number of clusters and the model order was varied from 5 to 15 and from 1 to 5, respectively. The best outcomes were selected. For the results of FTS-N, the model order p was varied from 1 to 5 and the number of clusters from 5 to 15. To obtain the best results for the proposed method (ARMA-T1FF), the number of clusters was varied from 2 to 5, the AR lag was varied from 1 to 5, and the MA lag was varied from 1 to 2. In addition, the number of iterations was set to 100 and the number of particles was 25. The best forecasting results, year by year, were determined using the parameters given in Table 6. The results are listed in Tables 7, 8, 9 and 10.

When the length of the test dataset is seven, the best forecasting results of the proposed method for the applications

Fig. 4 RMSE values of the proposed and other methods when ntest=7



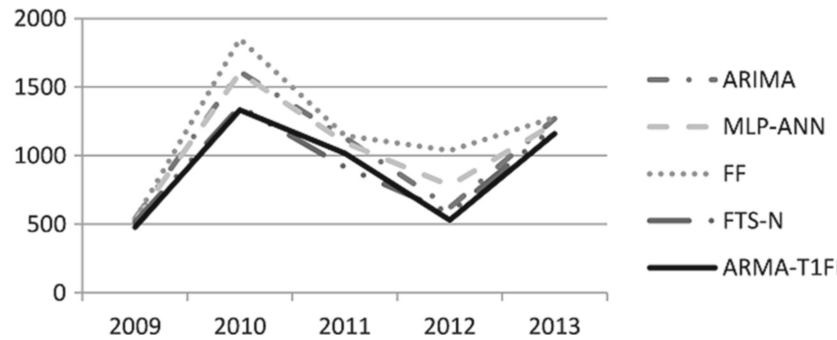


Fig. 5 RMSE values of the proposed and other methods when ntest=15

of the BIST100 are listed in Tables 7 and 8 in terms of RMSE and MAPE values, respectively. The proposed method has better forecasting performance than the other methods. Figure 4 compares the RMSE values of the proposed method (ARMA-T1FF) with those obtained by the other methods.

Similarly, when the length of the test dataset is 15 for BIST100, the best forecasting results in terms of RMSE and MAPE values are given in Tables 9 and 10, respectively. Again, the proposed method outperforms the others. Figure 5 compares the RMSE values of the proposed method with those obtained by the other methods.

4 Conclusions

This study proposed a new method that uses T1FF as the AR and MA model terms. The proposed method is the first recurrent fuzzy function approach regarding times series forecasting. To estimate the coefficients of the model, the PSO method is preferred. The proposed approach has the following advantages and contributions.

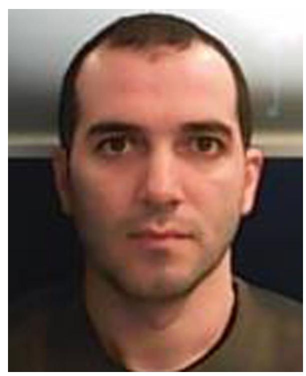
- Unlike most FISs, rules do not need to be defined in the proposed method.
- The assumptions of classical time series forecasting methods are not needed for the proposed method. In other words, there is need for any assumptions on the time series for ARMA-T1FF.
- Because the function that is to be optimized is not derivative, the PSO algorithm is preferred for estimating the coefficients of the model. The advantage of PSO is that it is less likely become stuck in a local optimum.
- The ARMA-T1FF approach is the first method that uses a recurrent learning structure.
- The number of inputs is fewer than for other methods because of the contribution of the MA model.
- The proposed method obtains better forecasting results than many methods in the literature.

The results obtained for the datasets show that ARMA-T1FF obtains overall better forecasting results than other methods. The results of the ABC dataset clearly show that the proposed method outperforms the other methods in terms of RMSE and MAPE. For the BIST100 dataset from 2009 to 2013, the ARMA-T1FF approach has better forecasting results on average in terms of RMSE and MAPE. For the TAIEX dataset, the mean RMSE for all years is used as the performance metric of the models, and for this metric, the proposed method gives the best results. In summary, considering the ABC, BIST100, and TAIEX datasets, the ARMA-T1FF approach obtains better forecasting results.

References

1. Aladag CH, Yolcu U, Egrioglu E, Dalar AZ (2012) A new time invariant fuzzy time series method based on particle swarm optimization. Appl Soft Comput 12:3291–3299
2. Aladag CH, Turksen IB, Dalar AZ, Egrioglu E, Yolcu U (2014) Application of type-1 fuzzy functions approach for time series forecasting. Turk J Fuzzy Syst 5(1):1–9
3. Bas E, Egrioglu E, Yolcu U, Aladag CH (2015) Fuzzy time series network used to forecast linear and nonlinear time series. Appl Intell 43(2):343–355
4. Bas E, Uslu VR, Yolcu U, Egrioglu E (2014) A modified genetic algorithm for fuzzy time series to find the optimal interval length. Appl Intell 42(2):453–463
5. Beyhan S, Alci M (2010) Fuzzy functions based arx model and new fuzzy basis function models for nonlinear system identification. Appl Soft Comput 10:439–444
6. BIST100 (2015) Istanbul Stock Exchange Index Dataset. http://www.borsaistanbul.com/veriler/gecmise-donuk-veri-satisi. Accessed 5 November 2015
7. Box GEP, Jenkins GM (1976) Time series analysis: forecasting and control. Holdan-Day, San Francisco
8. BuHamra S, Smaoui N, Gabr M (2003) The box-Jenkins analysis and neural networks: prediction and time series modeling. Apple Math Model 27(10):805–815
9. Catalao JPS, Pousinho HMI, Mendes VMF (2011) Hybrid wavelet PSO ANFIS approach for short term wind power forecasting in Portugal. IEEE Trans Sustain Energy 2(1):50–59
10. Celikyilmaz A, Turksen B (2009) Modeling uncertainty with fuzzy logic: with recent theory and applications. Springer, Berlin
11. Chabaa S, Zeroual A, Antari J (2009), ANFIS method for forecasting internet traffic time series. In: IEEE microwave symposium
12. Chang B (2008) Resolving the forecasting problems of overshoot and volatility clustering using ANFIS coupling nonlinear heteroscedasticity with quantum tuning. Fuzzy Set Syst 159(23):3183–3200

13. Chau KW (2006) Particle swarm optimization training algorithm for ANNs in stage prediction of Shing Mun River. J Hydrol 329:363–367
14. Chen B, Ma Z (2009) Short-term traffic flow prediction based on ANFIS. In: IEEE communication software and networks, pp 791–793
15. Chen D, Zhang J (2005) Time series prediction based on ensemble ANFIS. In: Proceedings of the fourth international conference on machine learning and cybermetics, pp 3552–3556
16. Chen KY, Wang CH (2007) A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan. Expert Syst Appl 32(1):254–264
17. Chen SM (1996) Forecasting enrollments with fuzzy time series. Fuzzy Set Syst 81:311–319
18. Chen SM, Chung NY (2006) Forecasting enrollments using high-order fuzzy time series and genetic algorithms. Int J Intell Syst 21:485–501
19. Chen SM, Tanuwijaya K (2011) Multivariate fuzzy forecasting based on fuzzy time series and automatic clustering techniquesh. Expert Syst Appl 38:10594–10605
20. Egrioglu E (2012) A new time invariant fuzzy time series forecasting method based on genetic algorithm. Adv Fuzzy Syst 2012:Article ID 785709
21. Egrioglu E, Aladag CH, Yolcu U, Bas E (2014) A new adaptive network based fuzzy inference system for time series forecasting. Aloy J Soft Comput Appl 2:25–32
22. Egrioglu E, Aladag CH, Yolcu U, Uslu VR, Basaran MA (2009) A new approach based on artificial neural networks for high order multivariate fuzzy time series. Expert Syst Appl 36: 10589–10594
23. Huang CM, Huang CJ, Wang ML (2005) A particle swarm optimization to identifying the ARMAX model for short-term load forecasting. IEEE T Power Syst 20(2):1126–1133
24. Huarng KH (2007) A multivariate heuristic model for fuzzy time-series forecasting. IEEE Trans Syst Man Cybern 37(4):836–846
25. Huarng K, Yu HK (2006) The application of neural networks to forecast fuzzy time series. Phys A 363:481–491
26. Jain A, Kumar AM (2007) Hybrid neural network models for hydrological time series forecasting. Appl Soft Comput 7:585–592
27. Janacek GJ (2001) Practical time series. Arnold
28. Jang JSR (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23(3):665–685
29. Jilani TA, Burney SMA, Ardil C (2008) Multivariate high order fuzzy time series forecasting for car road accidents. Int J Comput Int Syst 4:15–20
30. Kim D, Kim C (1997) Forecasting time series with genetic fuzzy predictor ensemble. IEEE Trans Fuzzy Syst 5(4):523–535
31. Kuo IH, Horng SJ, Kao TW, Lin TL, Lee CL, Pan Y (2009) An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization. Expert Syst Appl 36:6108–6117
32. Kuo RJ, Chen CH, Hwang YC (2001) An intelligent trading decision support system through integration of genetic algorithm based fuzzy neural network and artifical neural network. Fuzzy Set Syst 118:21–45
33. Lee YS, Tong LI (2011) Forecasting time series using a methodology based on autoregressive integrated moving average and genetic algorithm. Knowl-Based Syst 24(1):66–72
34. Mamdani EH, Assilian S (1975) An experiment in linguistic synthesis with a fuzzy logic controller. Int J Man Mach Stud 7(1):1–13
35. Pai PF, Lin CS (2005) A hybrid ARIMA and support vector machines model in stock price forecasting. Omega 33(7):497–505
36. Park FI, Lee DJ, Song CK, Chun (2010) TAIEX and KOSPI forecasting based on two-factor high-order fuzzy time series and particle swarm optimization. Expert Syst Appl 37:959–967
37. Song Q, Chissom BS (1993) Fuzzy time series and its models. Fuzzy Set Syst 54:269–277
38. Song Q, Chissom BS (1993) Forecasting enrollments with fuzzy time series – part I. Fuzzy Set Syst 54:1–10
39. Song Q, Chissom BS (1994) Forecasting enrollments with fuzzy time series – part II. Fuzzy Set Syst 62(1):1–8
40. TAIEX (2015) Taiwan Stock Exchange Index Dataset. http://www.taiwanindex.com.tw/index/history/t00. Accessed 17 October 2015
41. Takagi T, Sugeno M (1985) Fuzzy identification of systems and its applications to modeling and control. IEEE Trans Syst Man Cybern 15(1):116–132
42. Turksen IB (2008) Fuzzy functions with LSE. Appl Soft Comput 8(3):1178–1188
43. Tseng FM, Yu HC, Tseng GH (2002) Combining neural network model with seasonal time series ARIMA model. Technol Forecast Soc 69:71–87
44. Yolcu U, Aladag CH, Egrioglu E (2013) A new linear and nonlinear artificial neural network model for time series forecasting. Decis Support Syst 54(3):1340–1347
45. Zadeh LA (1965) Fuzzy sets. Inf Control 8:338–353
46. Zadeh LA (1973) Outline of a new approach to the analysis of complex systems and decision process. IEEE Trans Syst Man Cybern 3(1):28–44
47. Zhang G (2003) Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 50:159–175



Nihat Tak is a graduate assistant at Kırklareli University, Turkey. He received the B.Sc. degree in statistics and computer sciences from Karadeniz Technical University in 2008. He received the M.Sc. degree in statistics from the University of Arkansas in 2012 and Ph.D. degree from Yıldız Technical University in 2016. His research interest include: fuzzy approaches in statistics, computational optimization techniques, time series analysis, and Bayesian approaches.

Atif A. Evren is associate professor at the department of Statistics at Yıldız Technical University. He was graduated from the Department of Industrial Engineering of Bosphorus University in 1988. He received Ph.D. degree on nonlinear regression models in Statistics program at Istanbul University in 1996. His research interests include statistical modeling, regression models, and applications of entropy in statistics.

Mujgan Tez is a professor at the Marmara University, Turkey. She received the M.Sc. degree in statistics at the Faculty of Science of the Ankara University, Turkey, in 1982 and the Ph.D. degree in statistics at the Institute of Natural and Applied Sciences, of the Cukurova University, Turkey, in 1985. Her research interests include Nonlinear Regression, Geometry of Statistical Models, Compartment Models.

Erol Egrioglu is a professor at the University of Giresun, Turkey. He received the M.Sc. degree M.Sc. thesis title: Bayesian analysis of ARMA model and an application) in statistics at the Faculty of Science and Arts of the University of Ondokuz Mayis, Turkey, in 2002 and the Ph.D. degree in statistics at the Faculty of Science, of the University of Hacettepe, Turkey, in 2006. His research interests include time series analysis, Bayesian approaches, fuzzy time series and artificial neural networks, heuristic and evolutionary algorithms.