Marmara Üniversitesi Öneri Dergisi • Cilt 15, Sayı 54, Temmuz 2020, ISSN 2147-5377, ss. 350-366
DOI: 10.14783/maruoneri.771818
ARAŞTIRMA MAKALESİ / RESEARCH ARTICLE

GREY WOLF OPTIMIZER BASED RECURRENT FUZZY REGRESSION FUNCTIONS FOR FINANCIAL DATASETS

FİNANSAL VERİSETLERİ İÇİN BOZKURT OPTİMİZASYON TEMELLİ GERİ BESLEMELİ BULANIK ÇIKARIM FONKSİYONLARI

Nihat TAK*

Abstract

Time series models are used extensively in many fields, such as medicine, engineering, business, economics, and finance, with the aim of making forecasts through the help of observation values from previous periods. Therefore, there are many efforts to improve time series forecasting performances in the recent literature, mainly using alternative/non-probabilistic methods. In the present study, a novel forecasting approach has been proposed by combining the type-1 fuzzy functions (T1FF) with the Autoregressive moving average (ARMA) model based on grey wolf optimizer (GWO) in order to be able to overcome the nonlinear structure in time series dataset. Considering the superiorities of GWO over other methods, such as less storage requirements and rapid convergence by striking the proper stability between the exploration and exploitation throughout the search, estimation of the coefficients of the R-T1FFs method obtained through GWO to minimize the sum of squared errors (SSE). Comparison of the proposed method and several existing forecasting methods has been performed on five real world time series datasets. The results indicate that the proposed method produces better forecasts most of the time in the terms of mean absolute percentage errors and root mean square errors along with the better running time.

Keywords: Type-1 Fuzzy Functions, Grey Wolf Optimizer, Autoregressive Moving Average, Forecasting

Öz

Zaman serisi modelleri, tıp, mühendislik, işletme, ekonomi ve finans gibi birçok alanda, önceki dönemlerden gözlem değerleri yardımıyla tahminler yapmak amacıyla yaygın olarak kullanılmaktadır. Bu nedenle, özellikle alternatif/olasılık dışı yöntemler kullanılarak, zaman serisi tahmin performanslarını geliştirmek için birçok çaba vardır. Bu çalışmada, zaman serisi veri kümesindeki doğrusal olmayan yapının üstesinden gelebilmek için, Bozkurt optimizasyon (GWO) temelli Otoregresif hareketli ortalama (ARMA) modeli ile tip-1 bulanık

* Kirklareli University, Department of Econometrics, E mail: nihattak@gmail.com

Makale Gönderim Tarihi: 16.10.2019
Yayına Kabul Tarihi: 13.03.2020

fonksiyonların (T1FFs) birleştirilmesiyle yeni bir tahmin yaklaşımı önerilmiştir. GWO'nun, arama boyunca keşif ve uygun stabiliteye hızlı ulaşması, daha az depolama gereksinimleri ve hızlı yakınsama gibi diğer yöntemler üzerindeki üstünlükleri göz önüne alındığında, kare hatalarının toplamını en aza indirgemek için geribeslemeli T1FFs yönteminin katsayılarının tahmini GWO ile elde edilmesi uygun görüşmüştür. Beş farklı gerçek veri kümesinde önerilen yöntemin ve mevcut birkaç tahmin yönteminin karşılaştırılması gerçekleştirilmiştir. Sonuçlar, önerilen yöntemin, ortalama mutlak yüzde hataları ve kök ortalama kare hataları ile birlikte daha iyi çalışma süresi açısından çoğu zaman daha iyi tahminler ürettiğini göstermektedir.

Anahtar Kelimeler: Tip-1 Bulanık Fonksiyonlar, Bozkurt Optimizasyonu, Otoregresif Hareketli Ortalamlar, Öngörü

1. Introduction

The concept of forecasting is defined as the preliminary approximation of the values that a variable may take in the future under certain assumptions. Forecasting by time series analysis is an attempt to show the extent to which predictive values can be realized under certain assumptions, using the observed values of the current and past periods of any variable. The fact that the accurate forecast that brings successful decisions and maximizes the benefits obtained in this way increases the interest in forecasting models. The autoregressive moving average (ARMA) model is one of the most used traditional time series forecasting methods in probabilistic approaches. The ARMA model assumes that there is a linear relationship between the data forming the series and has a structure that models this linear relationship. Because most of the real-world time series datasets have nonlinear structures, the majority of the time stochastic or traditional approaches fail to give satisfactory forecasting results. Therefore, numerous researchers focus on alternative approaches, such as fuzzy inference systems (FIS). Some of the well-known FISs are introduced by Mamdani and Assillian (1975), Takagi and Sugeno (1985), Jang (1993). The method of adaptive neuro FIS (ANFIS), which is introduced by Jang, is one of the most used one in terms of time series forecasting. There are numerous studies based on ANFIS in literature, some of which are introduced by Chen and Zhang (2005), Egrioglu et al. (2014), Wei (2016), Chang (2008), Sarica, Egrioglu and Asikgil (2016).

Recently, becoming more popular FIS is T1FF. Most of the FISs are rule-based systems. Because it is difficult to define the rules, Celikyilmaz and Turksen (2009) have introduced type-1 fuzzy functions (T1FF). T1FF was, first, employed in forecasting problems by Beyhan and Alici (2010). Later, Aladag et al. (2014) proposed T1FFs by including an autoregressive model into their algorithm. Tak et al. (2018) introduced another model that includes MA model in their approach to get better forecasting results. Tak (2018) proposed meta fuzzy functions to improve the forecasting ability of Tak et al. (2018). Another study was conducted by Tak (2020) that employed intuitionistic fuzzy c-means clustering algorithm in T1FFs.

Because the objective functions of alternative forecasting methods are mostly non-derivative, meta-heuristic optimization methods are frequently employed to obtain the coefficients of the models. Some of commonly used meta-heuristic optimization methods are the artificial bee colony algorithm (ABC) Karaboga and Basturk (2007), particle swarm optimization (PSO) Eberhart and Kennedy (1995). and genetic algorithm (GA) for time series forecasting methods Chau (2006), Hong (2009), Hong (2010), Tak et al. (2018), Aladag et al.(2012), Aladag, Yolcu and Egrioglu (2015), Liao and Tsao (2006), Niu, Wang and Wu (2010). However, there are some limitations of the aforementioned metaheuristic algorithms, such as premature convergence and poor local search ability Faris et al. (2018), Pradhan, Roy and Pal (2016). Therefore, in the present study, to avoid these disadvantages, a novel forecasting approach has been proposed by employing GWO with the R-T1FFs.

The social intelligence of grey wolves is the idea of developing GWO algorithm. The social intelligence refers to leadership hierarchy and hunting behavior of grey wolves in nature Mirjalili and Lewis, (2014). The superiority of the GWO has been shown to be competitive and better than most of the heuristic optimization techniques such as GA, PSO, GSA, ABC and many others Saad, Dong and Karimi (2017), Zou, Sun and Zhang (2005).

Kumaran and Ravi (2014), Mustaffa, Sulaiman and Kahar (2015), Yusof and Mustaffa (2015), Niu et al. (2016) that employ GWO in their forecasting methods are some of the studies in the literature. The outstanding forecasting abilities of these methods show that GWO can produce promising forecasts. Considering the aforementioned advantages of GWO and the successful outcomes in mentioned studies, GWO is employed in the R-T1FF method. In the proposed method, disturbance terms for the MA part are identified using the residuals. The input matrix consists of lagged variables of the time series and disturbances, and the membership grades.

The remainder of the study is as follows: The flowchart and algorithm of GWO based R-T1FFs are presented in Section 2. In section 3, several practical time series datasets are used to investigate the forecasting performance of GWO based R-T1FFs. Finally, some remarks and conclusions are argued in Section 5.

2. Proposed Method

T1FF was introduced by Celikyilmaz and Turksen (2009) as an alternative for FISs. The main advantage of the T1FF approach is that the need for defining the rules is released. The aim of Celikyılmaz and Turksen (2009) was to make contribution to regression and classification problems. However, it was later employed in forecasting methods. It was first employed in forecasting methods by Beyhan and Alici (2010). In their approach, they used autoregressive with exogenous input (ARX) model that was not capable of seeking for the best model. Later, Aladag et al. (2014) proposed a forecasting method by employing T1FFs with AR to search for the best model into their algorithm. Tak et al. (2018) has proposed another method that combines autoregressive moving average model with T1FF (R-T1FF). They used the PSO method to determine the coefficients of the model parameters.

The main disadvantage of the R-T1FFs approach was the running time. Therefore, to overcome this disadvantage, GWO is used to minimize the objective function of the proposed method. In the

GWO searches for the best three search agents; described as the alpha, beta, and delta. The rest of the agents are described as omegas. The optimization process is leaded by alpha, beta, and delta while the omegas following them.

2.1. Algorithm

The detailed steps of GWO based R-T1FFs and the flow chart (see Figure 1) is given as follows.

Step 1. The dataset is discriminated as training and test sets. The number of observations in the training and test sets are determined as $ntrain$ and $ntest$, respectively.

$$X = [X_{ij}], i = 1,2,\ldots,p; j = 1,2..,n \tag{1}$$

$$X_{train} = [X_{ij}], i = 1,2,\ldots,p; j = 1,2..,ntrain \tag{2}$$

$$X_{test} = [X_{ij}], i = 1,2,\ldots,p; j = ntrain + 1, ntrain + 2, \ldots, n \tag{3}$$

Step 2. The inputs of proposed method are the lagged variables of the time series and disturbances. After clustering the inputs by using FCM, the obtained membership grades are added as new variables into the model inputs.

Step 3. The training dataset consists of the constant terms, the degree of membership values, lagged variables of the training dataset, and lagged variables of the training disturbance terms. The training set for the first wolf looks and the first cluster like as follows.

$$X_{train}^{(1)(1)} = [C\ \mu_{train}\ log(\mu_{train})\ \mu_{train}^{2}\ Y_{t-1}\ Y_{t-2} \ldots \epsilon_{t-1}\ \epsilon_{t-2}] \tag{4}$$

Step 4. Initialize the number of wolves, a, A, C and the number of iterations. Each search agent has $(p + q + 4)c$ positions, where c is the number of clusters and p and q stand for the number of lags for AR and MA, respectively.

Step 5. The normal distribution with the expected value of 0 and the standard deviation of 1 is used to determine the initial positions of the wolves. After calculating the fitness of each wolf, we assign the best wolf as the alpha (α), the second best wolf as beta (β), the third best wolf as delta (δ), and the other wolves as omegas (ω)

Step 6. The disturbance term (e_{t-q}) for i^{th} wolf and the first observation are obtained by using Equation 5-8.

$$\bar{Y}_{i}^{(j)(\bar{k})} = X_{(i)(.)}^{(j)(k)}\, P^{T}{}_{(i)(.)}^{(j)(k)};; i = 1,\ldots,ntrain; j = 1,\ldots,c; l = 1,\ldots,k \tag{5}$$

$$\overline{Y*}^{\overline{(j)}\overline{(k)}} = \bar{Y}_{(i)}^{(j)(\bar{k})}\, \mu^{T}{}_{(i)(.)}^{(j)} \tag{6}$$

$$e_{k}^{i} = Y_{i} - \bar{Y}*_{i}^{(\bar{k})} \tag{7}$$

$$X_{(i+1)(p)}^{(.)(k)} = e_k^i \tag{8}$$

$X_{(i)(.)}^{(j)(k)}$ stands for the values of all parameters of the i^{th} observation for j^{th} cluster and k^{th} wolf. $P^{T}{}_{(i)(.)}^{(j)(k)}$ is the locations of the i^{th} observation and j^{th} cluster in k^{th} wolf. $\widehat{Y*}_{i}^{(j)(k)}$ stands for the predictions of j^{th} cluster and k^{th} wolf for $i^{t}h$ observation. $\mu^{T}{}_{(i)(.)}^{(j)}$, is the degree of memberships of the i^{th} observation in j^{th} cluster. Y_i and $\widehat{Y*}_{i}^{(k)}$ stands for the actual and forecasted value of the i^{th} observation.

$\widehat{Y}_{i}^{(j)(k)}$ is calculated for all clusters in Equation 5. e_k^i is calculated for i^{th} observation and k^{th} wolf in Equation 7. e_k^i for k^{th} wolf and i^{th} observation is assigned to the $(i+1)^{th}$ observation in the input for each cluster in Equation 8.

Step 7. Equations 5-8 are calculated for each observation to determine the values of disturbance terms.

Step 8. Repeat steps 6 – 7 for all wolves.

Step 9. The positions of the wolves, $a, A, and\ C$ are updated by using the following equations.

$$\vec{A} = 2\overrightarrow{a r_1} - \vec{a} \tag{9}$$

$$\vec{C} = 2\vec{r_2} \tag{10}$$

where r_1 and r_2 are randomly generated vectors in [0,1] and $\vec{a}$ is linearly decreased from two to zero over iterations.

$$\overrightarrow{D_\alpha} = \left|\overrightarrow{C_1 P_\alpha} - \vec{P}\right|, \tag{11}$$

$$\overrightarrow{D_\beta} = \left|\overrightarrow{C_2 P_\beta} - \vec{P}\right|, \tag{12}$$

$$\overrightarrow{D_\delta} = \left|\overrightarrow{C_3 P_\delta} - \vec{P}\right| \tag{13}$$

$$\overrightarrow{P_1} = \overrightarrow{P_\alpha} - \overrightarrow{A_1}(\overrightarrow{D_\alpha}), \tag{14}$$

$$\overrightarrow{P_2} = \overrightarrow{P_\beta} - \overrightarrow{A_2}(\overrightarrow{D_\beta}), \tag{15}$$

$$\overrightarrow{P_3} = \overrightarrow{P_\delta} - \overrightarrow{A_3}(\overrightarrow{D_\delta}) \tag{16}$$

$$\vec{P}(t+1) = \frac{\overrightarrow{P_1}+\overrightarrow{P_2}+\overrightarrow{P_3}}{3} \tag{17}$$

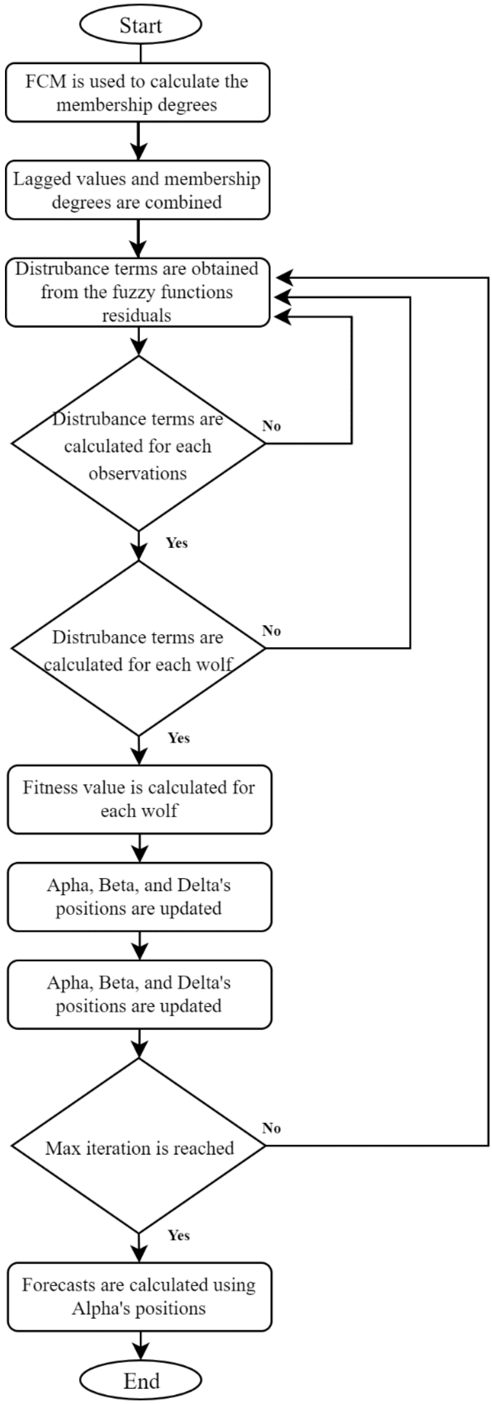

Figure 1: Flowchart of GWO based R-T1FFs

Step 10. The agent that has the best solution is selected as the alpha. The alpha's positions are used to calculate e_t for the test set by using Equations 18-21.

$$\bar{Y}_{test_t}^{(j)} = X_{test\, i_m}^{(j)} P_{test\, i_m}^{T\,(j)} \tag{18}$$

$$\widehat{Y_{test_t}}^{*} = \bar{Y}_{test_t}^{(j)} \mu_{test\, i_m}^{T\,(j)} \tag{19}$$

$$e_{test_i} = Y_{test_i} - \bar{Y}_{test_t}^{*} \tag{20}$$

$$X_{test\, t_{(i+1)(p)}}^{(.)} = e_{test_i} \tag{21}$$

Equation 18 is repeated for all clusters, then Equations 19-21 are calculated. Because the disturbances are calculated observation by observation, we repeat Step 10 for each observation in the test set.

Step 11 Repeat Steps 6-10 for each iteration.

Step 12. The final alpha's positions are used as the best candidates of the coefficients of the proposed method; thus, they are used to forecast the future values of a given time series by using the Equations 22-23.

$$\bar{Y}_{test_t}^{(j)} = X_{test\, i_m}^{(j)} P_{test\, i_m}^{T\,(j)} \tag{22}$$

$$\widehat{Y_{test_t}}^{*} = \bar{Y}_{test_t}^{(j)} \mu_{test\, i_m}^{T\,(j)} \tag{23}$$

3. Evaluation

Taiwan Stock Exchange (TAIEX) datasets TAIEX (2015) that are daily measured between 1999 and 2004, Australian Beer Consumption (ABC) dataset Janacek (2001) in which observations quarterly collected from 1956 to 1994, and Istanbul Stock Exchange (ISEX) ISEX (2015) datasets that are daily observed from 2009 to 2013 are used as the evaluation of the proposed method. These datasets are frequently used sets because stock index datasets are known as the complex time series and usually classical approaches usually fail to give desirable outputs. Thus, alternative methods usually use these datasets. TAEX, ABC, and ISEX datasets are selected in this sense for the evaluation of the proposed approach with the existing forecasting approaches. For the evaluation metrics, mean absolute percentage error (MAPE) and root mean squared errors (RMSE) that are given in Equation 19 – 20, respectively, are selected.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \widehat{y_t}}{y_t}\right| \tag{23}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \widehat{y_t})^2} \tag{24}$$

where $\widehat{y_t}$ is the forecasts and y_t is the actual values.

Table 1 represents the parameter specification and the features of the datasets of GWO based R-T1FFs.

Table 1: Parameter selections and summary of the sets

Series/Year	Number of Observations	Number of Lag (MA)	Number of Lag (AR)	Number of Cluster	ntest
ABC	147	1-2	1-10	2-10	16
ISEX-09	103	1-2	1-7	2-7	7,15
ISEX-10	104	1-2	1-7	2-7	7,15
ISEX-11	106	1-2	1-7	2-7	7,15
ISEX-12	106	1-2	1-7	2-7	7,15
ISEX-13	106	1-2	1-7	2-7	7,15
TAIEX-99	266	1-2	1-7	2-7	45
TAIEX-00	271	1-2	1-7	2-7	47
TAIEX-01	244	1-2	1-7	2-7	43
TAIEX-02	248	1-2	1-7	2-7	43
TAIEX-03	249	1-2	1-7	2-7	43
TAIEX-04	250	1-2	1-7	2-7	45

The complexity with running time of R-T1FFs and the proposed method for each data set is represented in Tables 2 and 3. The calculations of the proposed method are computed on a computer equipped with 512 GB SSD HDD, 8 GB RAM, and I7 CPU.

Table 2: Calculation times of ISEX time series

	ntest=7		ntest=15	
Data Set	ARMA-T1FF	Proposed Method	ARMA-T1FF	Proposed Method
ISEX-09	2.64 sec	0.43 sec	4.95 sec	0.5 sec
ISEX-10	6.45 sec	0.28 sec	9.02 sec	0.53 sec
ISEX-11	7.31 sec	0.71 sec	5.72 sec	0.5 sec
ISEX-12	6.80 sec	0.28 sec	3.57 sec	0.42 sec
ISEX-13	8.72 sec	0.36 sec	3.51 sec	0.42 sec

Table 3: Calculation times of TAIEX and ABC

Data Set	ntest	ARMA-T1FF	Proposed Method
TAIEX-99	45	8.04 sec	3.01 sec
TAIEX-00	47	11.18 sec	3.21 sec
TAIEX-01	43	10.46 sec	1.16 sec
TAIEX-02	43	5.73 sec	1.14 sec
TAIEX-03	43	7.65 sec	0.53 sec
TAIEX-04	45	12.54 sec	0.67 sec
ABC	16	10.02 sec	1.33 sec

3.1. ABC dataset

ABC dataset contains 148 observations that were quarterly measured between 1956 and 1994. To evaluate the performance of GWO based R-T1FFs, the following methods are used: SARIMA, modified ANFIS (MANFIS), adaptive neuro fuzzy inference system (ANFIS), and R-T1FFs. The outcomes of the existing methods are obtained from Tak et al. (2018).

Table 4: Outcomes of the existing methods and proposed method

Test Data	SARIMA	FANN	ANFIS	MANFIS	ARMA-T1FF	Proposed
430.5	452.72	453.88	446.71	445.23	442.82	437.1018
600	578.29	557.81	553.73	575.63	554.44	554.7321
464.5	487.71	497.52	482.07	494.07	477.33	461.0611
423.6	446.28	437.39	434.19	434.56	443.47	448.4949
437	456.77	449.01	438.55	444.69	422.46	419.6111
574	583.51	569.01	559.01	575.42	571.21	579.5689
443	492.13	471.08	472.52	481.28	463.85	443.1361
410	450.36	424.33	427.57	414.44	410.47	418.853
420	461.01	448.87	445.01	430.31	420.02	428.5535
532	588.96	560.04	562.94	565.18	551.34	555.7687
432	496.77	447.01	459.14	452.05	436.37	424.7154
420	454.64	408.64	416.16	392.14	390.61	400.0747
411	465.46	428.11	431.71	419.33	398.69	409.5583
512	594.71	537.69	544.98	536.88	517.62	511.5785
449	501.67	438.43	444.31	446.32	430.76	421.3092
382	459.17	420.58	426.01	406.64	408.27	415.5937
RMSE	47.04	24.11	25.05	21.37	19.21	19.56064
MAPE	0.0949	0.0476	0.0467	0.0401	0.0333	0.03197752

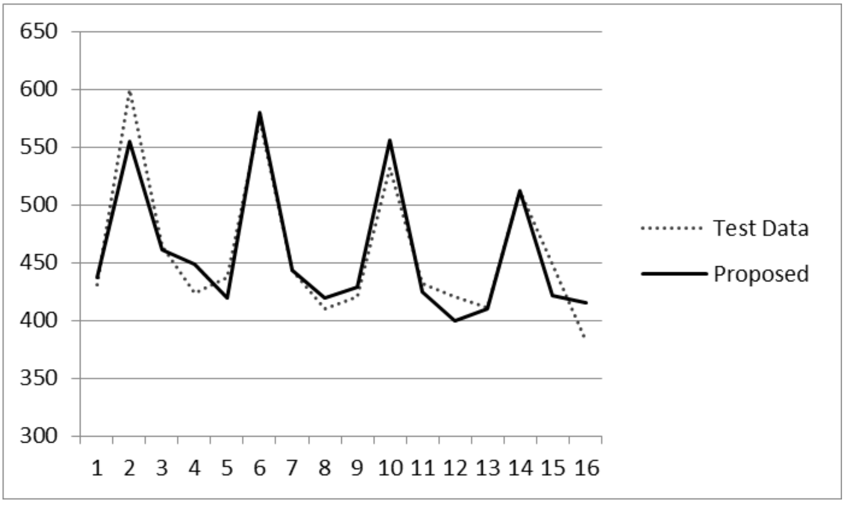

Figure 2: The line plot of the actual observations and the forecasts that are obtained from GWO based R-T1FFs

16 observations are left out for the test purpose in ABC dataset. The optimum cluster numbers are searched iteratively from 2 to ten, and the optimum number of lags for MA and AR are searched between 1 to 2 and 1 to 10, respectively. The number of iterations and wolves are set to 12 and 25, respectively, for GWO. The minimum RMSE value is obtained for the lag length of 8 for AR, the lag length of 1 for MA and 6 clusters, and MAPE values are calculated accordingly for GWO based R-T1FFs. Inspecting Table 4, the best performance is obtained by R-T1FFs in terms of RMSE. The second-best forecasts, however, are obtained from the GWO based R-T1FFs. Looking at the MAPE values, it is clear that the GWO based R-T1FFs outperform the other forecasting methods. Figure 2 represents the obtained forecasts from the GWO based R-T1FFs and the actual values.

3.2. TAIEX datasets

TAIEX data consists of six datasets yearly from 1999 to 2004. The objects of the TAIEX sets are observed daily. The following forecasting methods are selected for the evaluation of GWO based R-T1FFs: Chen, Chu and Sheu (2012), Chen and Chen (2011), Chen Chang (2010), Chen (1996), and R-T1FFs. The outcomes of the selected methods in Table 5 are cited from Bas et al. (2015) and Tak et al. (2018).

Table 5: Outcomes of the existing methods and proposed method

Methods	99	00	01	02	03	04	Mean
Chen and Chen (2011)	112.47	123.62	115.33	71.01	58.06	57.73	89.7
Chen and Chang (2010)	101.97	129.42	113.33	66.82	53.51	60.48	87.58
Chen, Chu and Sheu (2012)	99.87	119.98	114.47	67.17	52.49	52.27	84.37
Chen (1996)	120	176.32	147.84	101.18	74.46	84.28	117.34
MANFIS (2014)	101.94	124.92	112.47	62.57	52.33	53.66	84.64
Tak et al. (2018)	98.33	128.18	106.48	65.14	52.38	53.78	84.05
Proposed Method	96.82	126.37	106.64	65.01	52.16	53.7	83.45*

The forecasts of the proposed approach and the existing methods are evaluated based on RMSE. The best forecasts were obtained from the proposed method for 1999 and 2003. For 2000 and 2004, the method proposed by Chen et al. (2012) outperformed the others. MANFIS gave the best outcomes for 2002. R-T1FFs gave the best forecasting results for 2001. However, GWO based R-T1FFs have better forecasting accuracy than others by looking at the mean of all years. Figure 3 gives a comparison graph of the outcomes of GWO based R-T1FFs and the selected forecasting methods.

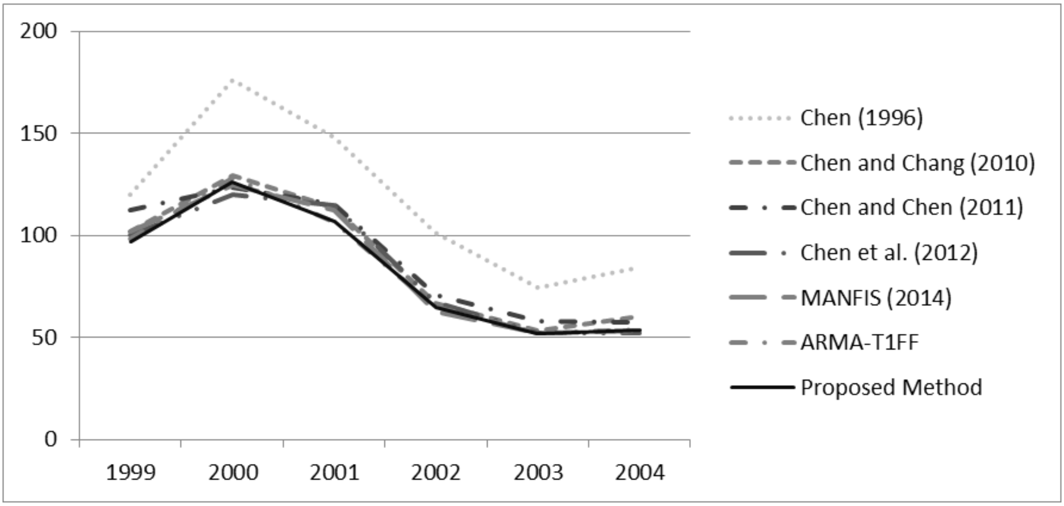

Figure 3: RMSE values of GWO based R-T1FFs and the selected methods

3.3. ISEX Datasets

There are five data sets for ISEX data. Observations are daily observed for the first half of the year from 2009 to 2013. To evaluate the forecasting accuracy of GWO based R-T1FFs, the outcomes of exponential smoothing (ES), ARIMA, T1FFs, fuzzy time series network (FTS-N), MLP-ANN, R-T1FFs, and the proposed method are compared. The outcomes of the selected methods for the comparison are taken from Bas et al. (2015) and Tak et al. (2018).

Holt and Winter's method performed best for ES procedure. The best model for the MLP-ANN method is searched iteratively with setting the number of inputs and the hidden layer neurons between 1 to 5. The best model for T1FFs is searched when the cluster numbers and the lag length for AR are varied from 2 to 10 and from 1 to 5, respectively. To obtain the best outcomes from FTS-N, the cluster numbers is varied from 5 to 15, the lag length (p) from 1 to 5. R-T1FFs searched for the best outcomes with setting the lags length for AR and MA models between 1 and 5 and 1 and 2, respectively, with the cluster numbers between 2 and 5. In addition, 100 iterations and 25 particles are initialized for particle swarm optimization in R-T1FFs. The best outcomes of the proposed method is searched iteratively when the lag length for AR was varied from 1 to 5, the lag length for MA was varied from 1 to 2, the number of clusters was varied from 2 to 5, and the number of wolves and iterations were set to 12 and 25, respectively.

The best forecasting outcomes of GWO based R-T1FFs were determined for each dataset by using the parameters in Table 6.

Table 6: The best parameter specifications for ISEX series

Year	# of lags for MA	# of lags for AR	# of Clusters	ntest
2009	2	2	2	7
2009	2	2	2	15
2010	2	2	2	7
2010	1	2	4	15
2011	2	1	5	7
2011	2	2	3	15
2012	2	1	2	7
2012	1	2	3	15
2013	2	1	2	7
2013	2	1	2	15

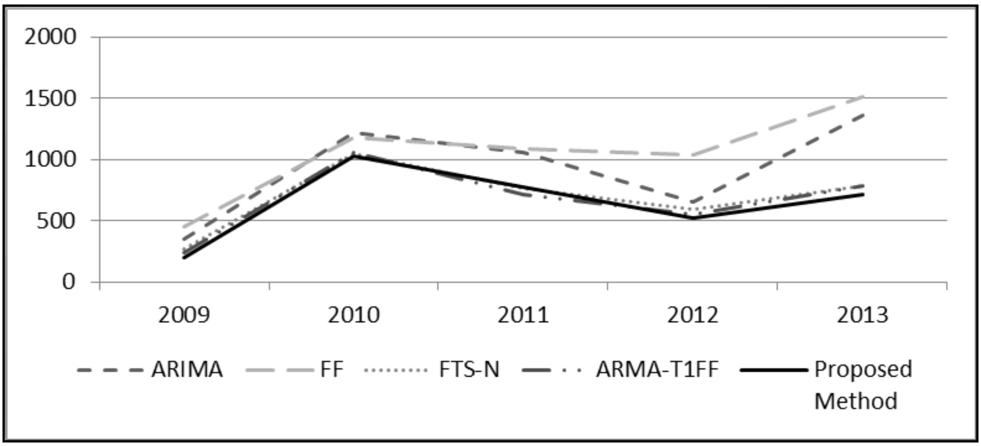

Figure 4: RMSE values of GWO based R-T1FFs and the selected methods when ntest=7

RMSE and MAPE values of GWO based R-T1FFs and the selected methods for comparison purpose are listed in Tables 7 and 8, respectively for ISEX datasets when the length of the test data is seven. Figure 4 and 5 compare the RMSE and MAPE values of GWO based R-T1FFs with the selected existing methods. Inspecting tables and figures it can be seen that GWO based R-T1FFs give very competitive and most of the time more accurate outcomes compared to the others.

Table 7: RMSE values for ISEX datasets for selected methods and GWO based R-T1FFs for the length of test set is 7

	ARIMA	ES	FF	FTS-N	ARMA-T1FF	Proposed Method
2009	344.91	344.93	445.5147	266.6011	235.96	200.91
2010	1221	1208.1	1179.9	1049.5	1057.097	1020.97
2011	1057.6	1057	1083.2	765.07	714.1724	772.3
2012	650.56	650.7387	1034.2	590.3545	547.13	524.19
2013	1361.6	1361.6	1511.6	786.13	783.9803	712.16
Mean	927.134	924.4737	1050.883	691.5311	667.6679	646.107*

Table 8: MAPE values for ISEX datasets for selected methods and GWO based R-T1FFs for the length of test set is 7

	ARIMA	ES	FF	FTS-N	ARMA-T1FF	Proposed Method
2009	0.0087	0.0087	0.0101	0.0058	0.0054	0.0049
2010	0.0183	0.0185	0.0179	0.0159	0.0157	0.0147
2011	0.0144	0.0144	0.0153	0.0105	0.0079	0.009
2012	0.0084	0.0084	0.0162	0.0084	0.0085	0.0072
2013	0.0116	0.0116	0.0131	0.0065	0.0058	0.0061
Mean	0.01228	0.01232	0.01452	0.00942	0.00866	0.00838*

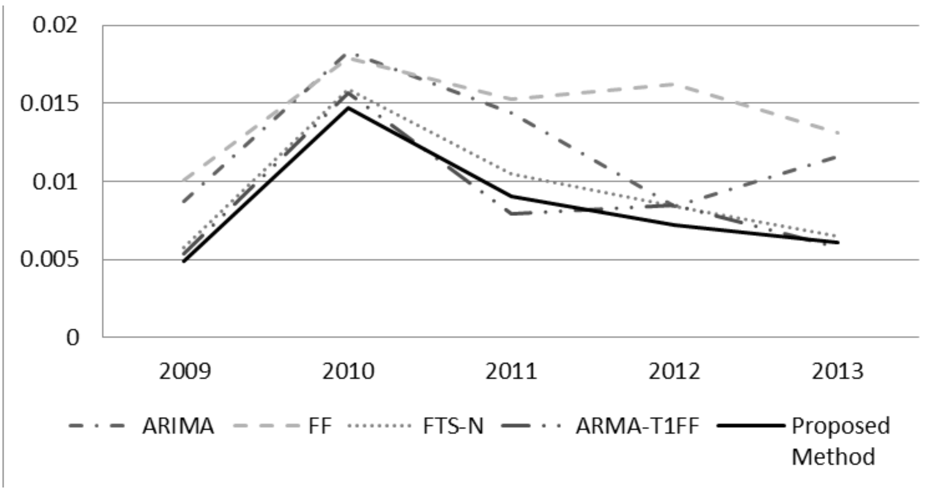

Figure 5: RMSE values of GWO based R-T1FFs and the selected methods when ntest=15

Tables 9 and 10 represent the outcomes of the selected methods and GWO based R-T1FFs for the length of the test dataset is 15 in terms of RMSE and MAPE, respectively. GWO based R-T1FFs overall gives better forecasting accuracy most of the time than the others. To visualize the forecasting accuracy of the proposed method, the line plot of RMSE and MAPE values of the proposed method and the existing methods are given in Figure 6 and 7.

Table 9: RMSE values for ISEX datasets for selected methods and GWO based R-T1FFs for the length of test set is 15

Years	ARIMA	ES	FF	FTS-N	ARMA-T1FF	Proposed
2009	540.21	540.2087	534.1345	514.5627	478.1365	439.1954
2010	1611.5	1611.5	1852	1357.4	1332.159	1315.429
2011	1129.6	1129.7	1145.6	916.5411	1017.41	954.6677
2012	620.7892	620.829	1037.6	581.71	529.69	510.8357
2013	1268.7	1268.7	1278.6	1207.9	1159.598	1056.979
Mean	1034.15984	1034.18754	1169.5869	915.62276	903.3987	855.42136*

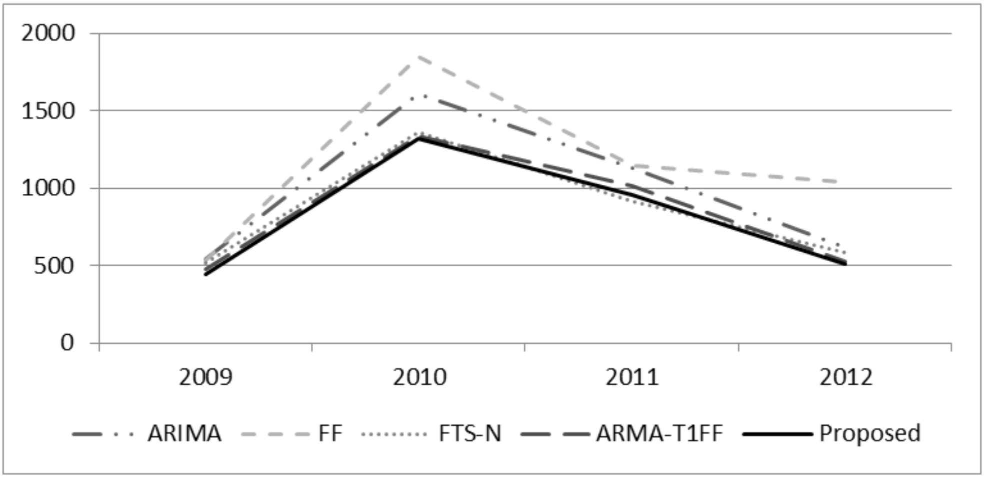

Figure 5: RMSE values of GWO based R-T1FFs and the selected methods when ntest=15

Table 10: MAPE values for ISEX datasets for selected methods and GWO based R-T1FFs for the length of test set is 15 }

	ARIMA	ES	FF	FTS-N	ARMA-T1FF	Proposed Method
2009	0.012	0.012	0.0438	0.0112	0.0093	0.0099
2010	0.022	0.022	0.0264	0.0202	0.019	0.021
2011	0.015	0.015	0.0156	0.0121	0.0134	0.011
2012	0.0088	0.0088	0.0161	0.0087	0.0076	0.0076
2013	0.0109	0.0109	0.0108	0.0106	0.0106	0.0086
Mean	0.01374	0.01374	0.02254	0.01256	0.01198	0.01162*

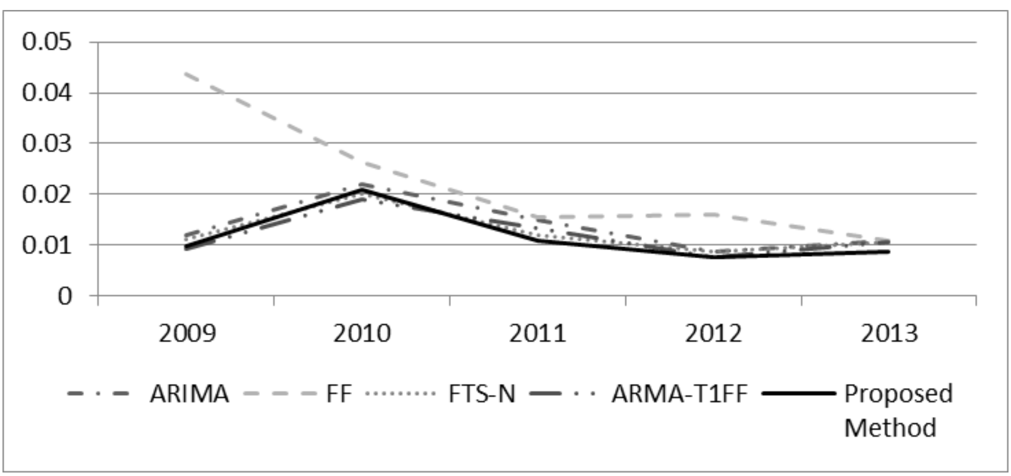

Figure 6: MAPE values of GWO based R-T1FFs and the selected methods when ntest=15

4. Conclusions

The proposed method contributes the R-T1FFs approach in terms of improving calculation time as well as the forecasting results. GWO, which is relatively a new approach for optimization problems, is adapted to R-T1FFs in this study. The proposed method has the following advantages and contributions.

- No assumption is required for the proposed method as the classical time series approaches.
- Because the objective function that is to be minimized in the proposed method is not derivative, relatively a newer approach, GWO is adapted to the proposed method. The main advantages of the GWO is that it is less likely to stuck in a local optimum, there are only two parameters to be adjusted, it needs less storage requirement, and it is easier to implement. Thus, the proposed method presents better outcomes with much less calculation time.
- The results showed that the proposed method was able to provide very competitive results compared with the other methods in literature.

The results obtained for the datasets show that GWO based R-T1FFs give very competitive results compared the other methods. The outcomes of the ABC dataset emphasize that GWO based R-T1FFs give the most accurate and the second most accurate forecasts in terms of MAPE and RMSE respectively. GWO based R-T1FFs have better forecasting results on average in terms of RMSE and MAPE for the ISEX dataset from 2009 to 2013. The similar results are obtained for TAEX datasets. Investigating the mean of RMSE values all years, it is obvious that the best forecasting accuracy is obtained from the proposed method. In summary, considering the ABC, ISEX, and TAIEX datasets, GWO based R-T1FFs obtain very competitive forecasting results.

References

Aladag, C.H., Turksen, I.B., Dalar, A.Z., Egrioglu, E. & Yolcu, U. (2014). Application of type-1 fuzzy functions approach for time series forecasting, *Turkish Journal of Fuzzy Systems*, 5(1), 1-9.

Aladag, C.H., Yolcu, U., Egrioglu. E. & Dalar, A.Z. (2012). A new time invariant fuzzy time series method based on particle swarm optimization, *Applied Soft Computing*, 12(10), 3291-3299.

Aladag, C.H., Yolcu, U. & Egrioglu, E. (2015). A new multiplicative seasonal neural network model based on particle swarm optimization, *Neural Processing Letters*, 37(3), 251-262.

Bas, E., Egrioglu, E., Yolcu, U. & Aladag, C.H. (2015). Fuzzy time series network used to forecast linear and nonlinear time series, *Applied Intelligence*, 43(2), 343-355.

Beyhan, S. & Alci, M. (2010). Fuzzy functions based arx model and new fuzzy basis function models for nonlinear system identification, *Applied Soft Computing*, 10(2), 439-444.

ISEX. (2015). Istanbul stock exchange index dataset. http://www.borsaistanbul.com/veriler/gecmise-donuk-veri-satisi. (Accessed 5 November 2015).

Celikyilmaz, A. & Turksen, B. (2009). *Modeling Uncertainty with Fuzzy Logic: With Recent Theory and Applications*. Berlin: Springer.

Chang, B.R. (2008). Resolving the forecasting problems of overshoot and volatility clustering using ANFIS coupling nonlinear heteroscedasticity with quantum tuning, *Fuzzy Sets and Systems*, 159(23), 3183-3200.

Chau, K.W. (2006). Particle swarm optimization-training algorithm for ANNs in stage prediction of shing mun river, *Journal of Hydrology,* 329(3, 4), 363-367.

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

Chen, S. M., & Chang, Y. C. (2010). Multi-variable fuzzy forecasting based on fuzzy clustering and fuzzy rule interpolation techniques. Information sciences, 180(24), 4772-4783.

Chen, S. M., & Chen, C. D. (2011). TAIEX forecasting based on fuzzy time series and fuzzy variation groups. IEEE Transactions on Fuzzy Systems, 19(1), 1-12.

Chen, S. M., Chu, H. P., & Sheu, T. W. (2012). TAIEX forecasting using fuzzy time series and automatically generated weights of multiple factors. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 42(6), 1485-1495.

Chen, D. & Zhang, J. (2005). Time series prediction based on ensemble ANFIS, *Proceedings of the Fourth International Conference on Machine Learning and Cybernetics*, 3552-3556.

Eberhart, R. & Kennedy, J. (1995). A new optimizer using particle swarm theory, *Micro Machine and Human Science Proceedings of the Sixth International Symposium on. IEEE (MHS'95).*

Egrioglu, E., Aladag, C.H., Yolcu, U. & Bas, E. (2014). A new adaptive network based fuzzy inference system for time series forecasting, *Aloy Journal of Soft Computing and Applications*, 2(1), 25-32.

Faris, H., Aljarah, I., Al-Betar, M.A. & Mirjalili, S. (2018). Grey wolf optimizer: A review of recent variants and applications, *Neural Computing and Applications*, 30(2), 413-435.

Hong, W.C. (2009). Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model, *Energy Conversion and Management*, 50(1), 105-117.

Hong, W.C. (2010). Application of chaotic ant swarm optimization in electric load forecasting, *Energy Policy*, 38(10), 5830-5839.

Huang, C.M., Huang, C.J. & Wang, M.L. (2005). A particle swarm optimization to identifying the ARMAX model for short-term load forecasting, *IEEE Transactions on Power Systems*, 20(2), 1126-1133.

Janacek, G. & Janacek, G. J. (2001). *Practical Time Series*. London: Arnold.

Jang, J.S.R. (1993). ANFIS: Adaptive-network-based fuzzy inference system, *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3), 665-685.

Karaboga, D. & Basturk, B. (2007). A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm, *Journal of Global Optimization*, 39(3), 459-471.

Kumaran, J. & Ravi, G. (2014). Long-term forecasting of electrical energy using ANN and HSA. *International Review on Modelling and Simulations (IREMOS)*, 7(3), 489-496.

Liao, G. & Tsao, T. (2006). Application of a fuzzy neural network combined with a chaos genetic algorithm and simulated annealing to short-term load forecasting, *IEEE Transactions on Evolutionary Computation*, 10(3), 330-340.

Mirjalili, S.M.S. & Lewis, A. (2014) Grey wolf optimizer, *Advances in Engineering Software*, 69,496-517.

Mustaffa, Z., Sulaiman, M.H. & Kahar, M.N.M. (2015, August). LS-SVM hyper-parameters optimization based on GWO algorithm for time series forecasting. *In Software Engineering and Computer Systems (ICSECS), 2015 4th International Conference on* (pp. 183-188). IEEE.

Niu, D., Wang, Y. & Wu, D.D. (2010). Power load forecasting using support vector machine and ant colony optimization, *Expert Systems with Applications,* 37(3), 2531-2539.

Niu, M., Wang, Y., Sun, S. & Li, Y. (2016). A novel hybrid decomposition-and-ensemble model based on CEEMD and GWO for short-term PM 2.5 concentration forecasting, *Atmospheric Environment*, 134,168-180.

Pradhan, M., Roy, P.K. & Pal, T. (2016). Grey wolf optimization applied to economic load dispatch problems, *International Journal of Electrical Power and Energy Systems*, 83, 325-334.

Saad, A.E.H., Dong, Z. & Karimi, M. (2017). A comparative study on recently-introduced nature-based global optimization methods in complex mechanical system design, *Algorithms*, 10(4), 120.

Wei, L. (2016). A hybrid anfis model based on empirical model decomposition for stock time series forecasting, *Applied Soft Computing,* 42, 368-376.

Mamdani, E.H. & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller, *International Journal of Man-Machine Studies,* 7(1), 1-13.

TAIEX. (2015). Taiwan stock exchange index dataset. http://www.taiwanindex.com.tw/index/history/t00. Accessed 17 October 2015.

Tak , N., Evren, A.A., Tez, M. & Egrioglu, E. (2018). Recurrent type-1 fuzzy functions approach for time series forecasting, *Applied Intelligence,* 48(1), 68-77, doi: 10.1007/s10489.017.0962-8.

Tak , N. (2018). Meta fuzzy functions: Application of recurrent type-1 fuzzy functions, *Applied Soft Computing*, 73, 1-13.

Tak, N. (2020). Type-1 recurrent intuitionistic fuzzy functions for forecasting, *Expert Systems with Applications*, 140, 112913.

Takagi, T. & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control, *IEEE Transactions on Systems, Man, and Cybernetics,* 15(1), 116-132.

Sarica, B., Egrioglu, E. & Asikgil, B. (2016). A new hybrid method for time series forecasting, *Neural Computing and Applications*, 29(3), 749-760.

Yusof, Y. & Mustaffa, Z. (2015, March). Time series forecasting of energy commodity using grey wolf optimizer, *In Proceedings of the International Multi Conference of Engineers and Computer Scientists (IMECS'15*) (Vol. 1).

Zou, Z.D., Sun, Y.M. & Zhang, Z.S. (2005). Short-term load forecasting based on recurrent neural network using ant colony optimization algorithm, *Power System Technology*, 3, 59-63.