Expert Systems with Applications 40 (2013) 854–857

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks

Erol Egrioglu [a], Cagdas Hakan Aladag [b,*], Ufuk Yolcu [c]

[a] Department of Statistics, Ondokuz Mayis University, Samsun, Turkey
[b] Department of Statistics, Hacettepe University, Ankara, Turkey
[c] Department of Statistics, Giresun University, Giresun, Turkey

ARTICLE INFO

Keywords:
Artificial neural networks
Defuzzification
Forecast
Fuzzification
Fuzzy c-means
Fuzzy time series

ABSTRACT

In recent years, time series forecasting studies in which fuzzy time series approach is utilized have got more attentions. Various soft computing techniques such as fuzzy clustering, artificial neural networks and genetic algorithms have been used in fuzzy time series method to improve the method. While fuzzy clustering and genetic algorithms are being used for fuzzification, artificial neural networks method is being preferred for using in defining fuzzy relationships. In this study, a hybrid fuzzy time series approach is proposed to reach more accurate forecasts. In the proposed hybrid approach, fuzzy c-means clustering method and artificial neural networks are employed for fuzzification and defining fuzzy relationships, respectively. The enrollment data of University of Alabama is forecasted by using both the proposed method and the other fuzzy time series approaches. As a result of comparison, it is seen that the most accurate forecasts are obtained when the proposed hybrid fuzzy time series approach is used.

1. Introduction

It is obvious that forecasting activities play an important role in our daily life. Therefore, various time series forecasting approaches have been proposed in the literature. There are some real life time series with multiple values for observations. For instance, stock exchange time series have many values for same day. Therefore, when these time series are analyzed by conventional time series approach, different time series such as the high, low, open, close, etc. have to be used for forecasting. However, studies that used to take the closing price as the daily stock index are not considered to be very objective. Thus, modeling the observations found in these types of time series by using a fuzzy set (Nguyen and Wu, 2000) or a fuzzy number (Song, Leland, & Chissom, 1995) seems to be more appropriate (Huarng & Yu, 2006b).

Fuzzy time series approach based on fuzzy set theory (Zadeh, 1965) firstly proposed by Song and Chissom (1993a, 1993b, 1994). The aim of the studies on the fuzzy time series in the literature is to contribute to three steps of the approach. One of these is the fuzzification step and some studies on this step are Huarng (2001), Huarng and Yu (2006b), Chen and Chung (2006), Lee, Wang, and Chen (2007), Li, Cheng, and Lin (2008), Cheng, Cheng, and Wang (2008), Yolcu, Egrioglu, Uslu, Basaran, and Aladag (2009), Egrioglu, Aladag, Yolcu, Uslu, and Basaran (2010) and Egrioglu, Aladag, Başaran, Uslu, and Yolcu (2011). Another step is defining fuzzy relations. The studies Sullivan and Woodall (1994), Chen (1996), Huarng and Yu (2006a), Aladag, Basaran, Egrioglu, Yolcu, and Uslu (2009), Egrioglu, Aladag, Yolcu, Basaran, and Uslu (2009) and Egrioglu, Aladag, Yolcu, Uslu, and Basaran (2009) are about defining fuzzy relations. And the last step is the defuzzification and Yu (2005), Jilani and Burney (2008) and Cheng, Chen, Teoh, and Chiang (2008) are some papers about this step.

In this study, a novel hybrid fuzzy time series approach in which fuzzy c-means (FCM) method and artificial neural networks are employed for fuzzification and defuzzification, respectively is proposed. In the implementation, the proposed is applied to well known enrollment data for the University of Alabama. The time series is also forecasted by other methods available in the literature. As a result of the implementation, obtained results calculated from the proposed method are compared to those obtained from other fuzzy time series approaches.

In the second section of the paper, the definitions of fuzzy time series are given. Third section presents the FCM method. Artificial neural networks are briefly given in Section 4. The proposed hybrid fuzzy time series approach is introduced in Section 5. Section 6 includes the implementation. Finally, the last section provides the obtained results and the discussion.

* Corresponding author. Tel.: +90 312 2977900.
E-mail address: aladag@hacettepe.edu.tr (C.H. Aladag).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.05.040

2. Fuzzy time series

The definition of fuzzy time series was firstly introduced by Song and Chissom (1993a, 1993b, 1994). Basic definitions of fuzzy time series are given as follows:

Let U be the universe of discourse, where $U = \{u_1, u_2, \ldots, u_b\}$. A fuzzy set A_i of U is defined as $A_i = f_{A_i}(u_1)/u_1 + f_{A_i}(u_2)/u_2 + \cdots + f_{A_i}(u_b)/u_b$, where f_{A_i} is the membership function of the fuzzy set $A_i; f_{A_i} : U \to [0,1]$. u_a is a generic element of fuzzy set $A_i; f_{A_i}(u_a)$ is the degree of belongingness of u_a to $A_i; f_{A_i}(u_a) \in [0,1]$ and $1 \leqslant a \leqslant b$.

Definition 1. Fuzzy time series. Let $Y(t)(t = \ldots, 0, 1, 2, \ldots)$, a subset of real numbers, be the universe of discourse on which fuzzy sets $f_j(t)$ are defined. If $F(t)$ is a collection of $f_1(t), f_2(t), \ldots$ then $F(t)$ is called a fuzzy time series defined on $Y(t)$.

Definition 2. Fuzzy time series relationships assume that $F(t)$ is caused only by $F(t-1)$, then the relationship can be expressed as:

$$F(t-1) \to F(t) \tag{1}$$

This model is called as first order fuzzy time series forecasting model.

3. The fuzzy c-means clustering method

FCM is a clustering method and was proposed by Bezdek (1981). This method is the most preferred fuzzy clustering method in the literature. In the FCM method, the data is partitioned into fuzzy sets by minimizing sum of square error for groups. Let u_{ij}, v_i and n represent membership value, cluster center and the number of variables, respectively. Thus, the form of the objective function tried to be minimized is as follows:

$$J_\beta(X, V, U) = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{\beta} d^2(x_j, v_i) \tag{2}$$

where β is weighting exponent ($\beta > 1$) and $d(x_j, v_i)$ is distance measure between the observation and the cluster center. J_β is tried to be minimized under the constraints given below.

$$\begin{aligned} &0 \le u_{ij} \le 1 \quad , \forall i, j \\ &0 < \sum_{j=1}^{n} u_{ij} \le n \quad , \forall i \\ &\sum_{i=1}^{c} u_{ij} = 1 \quad , \forall j \end{aligned} \tag{3}$$

Minimization process in the FCM is performed by using an iterative algorithm. In each iteration, the values of u_{ij} and v_i are updated by using the formulas given below.

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

4. Artificial neural networks

Artificial neural networks are algorithms which mimic the features of brain of human being. These features are generating new knowledge and exploring by learning. In other words, artificial neural networks are synthetic networks which imitate biological neural networks. Artificial neural networks are much more different than biological ones in terms of structure and ability, (Zurada, 1992). Artificial neural networks compose of a mathematical model (Zhang, Patuwo, & Hu, 1998). The fundamental elements of artificial neural networks can be given as follows:

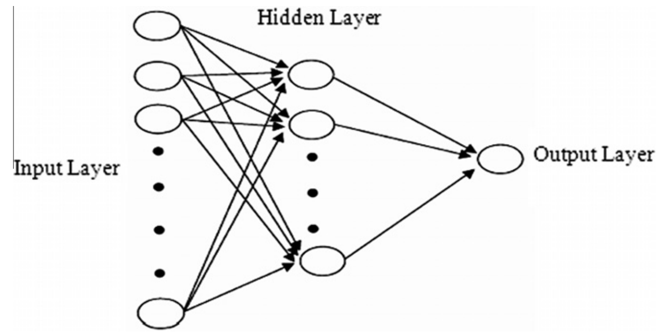

Fig. 1. Multilayer feed forward artificial neural networks with one output neuron.

4.1. Architecture structure

There are various types of artificial neural networks. Feed forward neural networks are one of these types. The structure of multilayer feed forward artificial neural networks is basically given in Fig. 1. The architecture as illustrated in the figure consists of three parts such as input, hidden, and output layers. Each layer consists of neurons. The architecture structure is determined based on deciding the number of neuron in each layer. These neurons are linked each other by weights. There is no link among the neurons in the same layer.

4.2. Learning algorithm

There have been many learning algorithms in order to determine weights. The one of the most employed algorithm is called Back Propagation learning algorithm. This learning algorithm updates the weights based on difference between real value and output value of the artificial neural networks. Learning parameter in back propagation algorithm plays an important role in reaching the desired outputs. While learning parameter is taken as a fixed value, it can be dynamically updated in the algorithm. Another efficient learning algorithm in artificial neural networks is Levenberg–Marquardt algorithm (Levenberg, 1944). This algorithm converge the result faster than Back Propagation learning algorithm.

4.3. Activation function

Activation function provides the non-linear mapping between input and output. The performance of networks depends on the proper choice of activation function. Activation function can be chosen as either linear or double polarized, or one polarized. Slope parameter should be determined when the activation is non-linear. Also, slope parameter plays a key role in reaching desired output values.

5. The proposed method

In the fuzzy time series literature, the definition and the partition of universe of discourse and the definition of discrete fuzzy set introduced by Song and Chissom (1993a, 1993b, 1994) have been used in many studies for the fuzzification process. However, unlike the algorithm of Song and Chissom (1993a), Cheng, Cheng, and Wang (2008) and Li et al. (2008) utilized FCM in the fuzzification process. When FCM method is employed, there is no need to use subjective measures for determining the length of interval, which plays important role on forecasting performance, and the membership values of discrete fuzzy sets. In the recent years, Egrioglu, Aladag, Yolcu, Uslu, and Erilli (2011) have used Gustafson Kessel clustering instead of fuzzy c-means in fuzzification stage.

In defining fuzzy relations, computations based on matrix operations or fuzzy relations group tables have been employed in the literature. On the other hand, Huarng and Yu (2006a), Aladag et al. (2009) and Egrioglu, Aladag, Yolcu, Basaran et al. (2009), Egrioglu, Aladag, Yolcu, Uslu et al. (2009) have used feed forward artificial neural networks in recent years. Using artificial neural networks for defining fuzzy relations produce more accurate forecast than those obtained from using matrix operations or fuzzy relations group tables. In addition, using artificial neural networks is easier.

In this study, an algorithm including both FCM and artificial neural networks is firstly proposed to reach high forecasting accuracy level. We were inspired by the algorithm proposed by Cheng, Chen, Teoh, and Chiang (2008), Cheng, Cheng, and Wang (2008). The proposed algorithm is introduced below in four steps.

Step 1. Apply FCM.

Firstly, FCM presented in Section 3 with the number of clusters c $(2 \leqslant c \leqslant n)$ is applied to time series includes crisp values. Values of cluster centers are calculated then, values of memberships are obtained by using these values of cluster centers.

Step 2. Fuzzify the time series.

Sorted fuzzy clusters L_r, $r = 1, 2, \ldots, c$ are obtained by sorting the cluster centers in ascending order. The fuzzy time series is obtained by mapping each observation into corresponding fuzzy set.

Step 3. Establish the fuzzy relationship with feed forward neural network.

An example will be given to explain step 3 more clearly for fuzzy time series. Because of dealing with first order fuzzy time series, one inputs are employed in neural network model, so that lagged variable F_{t-1} are obtained from fuzzy time series F_t. These series are given in Table 1. The index numbers (i) of L_i of F_{t-1} series are taken as input values whose titles are Input in Table 1 for the neural network model. Also, the index numbers of L_i of F_t series are taken as target values whose title is Target in Table 1 for the neural network model. When the third observation is taken as an example, inputs value for the learning sample $[L_6]$ is 6. Then, target value for this learning sample is 2.

Step 4. Defuzzify results.

The defuzzyfied forecasts are middle points of intervals which correspond to fuzzy forecasts obtained by neural networks in the previous stage.

6. Application

The enrollment data of University of Alabama which is shown in Table 2 is used in the implementation since this data have been used in many studies in the existence fuzzy time series literature. The proposed method is applied the yearly observed time series whose observations from 1971 to 1992 by changing the number of cluster between 5 and 15. In the establishing the fuzzy relationships process, the number of neuron in the hidden layer is changed between 1 and 8. There is one neuron in the input and the output layer. Therefore, to define fuzzy relationships well, eight different architectures are examined for each 11 clusters so 88 architectures are examined totally. In all of the neurons of the feed forward neural networks, logistic activation function is preferred. Besides, Levenberg–Marquardt algorithm (Levenberg, 1944) is used as training algorithm to reach optimal weight values of the networks. Mean square error (MSE) value is also used as performance measure and calculated MSE values for the cases are summarized in Table 3.

According to Table 3, the most accurate forecasts are obtained when the number of cluster and the number of neurons in the hidden layer are 13 and 5, respectively. MSE values of other proposed fuzzy time series methods available in the literature such as Song and Chissom (1993a,1994), Sullivan and Woodall (1994), Chen (1996), Cheng, Chen, Teoh, and Chiang (2008), Cheng, Cheng, and Wang (2008), Aladag et al. (2009), Egrioglu et al. (2010), Egrioglu, Aladag, Başaran et al. (2011), Egrioglu, Aladag, Yolcu et al. (2011) and MSE value of our proposed method are also given for

Table 1
Notations for first order fuzzy time series.

Observation No	F_{t-1}	F_t	Input	Target
1	–	L_6	–	–
2	L_6	L_2	6	2
3	L_2	L_3	2	3
4	L_3	L_7	3	7
5	L_7	L_4	7	4
6	L_4	L_2	4	2

Table 2
Enrollment data.

Years	Actual	Years	Actual
1971	13055	1982	15433
1972	13563	1983	15497
1973	13867	1984	15145
1974	14696	1985	15163
1975	15460	1986	15984
1976	15311	1987	16859
1977	15603	1988	18150
1978	15861	1989	18970
1979	16807	1990	19328
1980	16919	1991	19337
1981	16388	1992	18876

Table 3
MSE values obtained from the proposed method.

Number of cluster	Number of Hidden Layer Neuron			
	1	2	3	4
5	368492	368492	1949823	368492
6	476898	403050	403050	403050
7	369754	338626	306948	338626
8	348910	340688	309225	309225
9	390597	321385	313171	284341
10	527050	349865	349865	320078
11	482139	287954	258238	278299
12	364970	2857618	266284	202998
13	471361	471361	228329	228329
14	398909	296575	254856	355590
15	465229	422687	237637	227426
	5	6	7	8
5	368492	368492	368492	368492
6	330408	330408	330408	330408
7	306948	306948	306948	306948
8	307516	278294	276585	307516
9	284341	219073	219073	217351
10	201972	3136799	287378	201972
11	262794	262794	1639037	177776
12	137863	328921	137863	137863
13	32849*	219527	34354	34348
14	157029	157029	162438	171794
15	227426	241050	109132	230830

* Minimum MSE value.

Table 4
The comparison of the results.

Method	MSE
Song and Chissom (1993b)	412499
Song and Chissom (1994)	775687
Sullivan and Woodall (1994)	386055
Chen (1996)	407507
Cheng, Cheng, and Wang (2008)	228918
Aladag et al. (2009)	78073
Egrioglu et al. (2010)	60714
Egrioglu, Aladag, Başaran et al. (2011)	66661
Egrioglu, Aladag, Yolcu et al. (2011)	60140
The proposed method	32849

comparison purpose in Table 4. The result of our proposed method has the smallest MSE value when compared with the other methods so it can be said that the new proposed method produces better forecasts.

7. Conclusion and discussion

In fuzzy time series approach, some soft computing methods such as fuzzy clustering, artificial neural networks and genetic algorithms have been employed in order to get more accurate forecasts. In this study, we proposed a novel hybrid fuzzy time series approach in which FCM and artificial neural networks methods are utilized for the fuzzification and defining fuzzy relationships processes, respectively. The proposed method is the first fuzzy time series method that uses both FCM and artificial neural networks at the same time. The proposed hybrid approach is applied to well known enrollment data for the University of Alabama. Obtained MSE value calculated from the proposed method and MSE values obtained from other fuzzy time series approaches proposed in the literature are compared. As result of the comparison, it is seen that the proposed method produce the smallest MSE value. In other words, the most accurate forecasts are obtained when the proposed hybrid fuzzy time series approach is used.

Besides the obtained good result in the implementation, the proposed method has some advantages. Since FCM is used in the fuzzification step, some problems caused by partition of discourse of universe are removed. Also, there is no need to use difficult matrix operations or complex fuzzy group relationships tables since fuzzy relationships are defined by artificial neural networks. Therefore, using the hybrid fuzzy time series approach proposed in this study for forecasting will be a good choice.

References

Aladag, C. H., Basaran, M. A., Egrioglu, E., Yolcu, U., & Uslu, V. R. (2009). Forecasting in high order fuzzy time series by using neural networks to define fuzzy relations. *Expert Systems with Applications, 36*, 4228–4231.

Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*. NY: Plenum Press.

Chen, S. M. (1996). Forecasting enrollments based on fuzzy time-series. *Fuzzy Sets and Systems, 81*, 311–319.

Chen, S. M., & Chung, N. Y. (2006). Forecasting enrolments using high order fuzzy time series and genetic algorithms. *International Journal of Intelligent Systems, 21*, 485–501.

Cheng, C.-H., Chen, T.-L., Teoh, H. J., & Chiang, C.-H. (2008). Fuzzy time-series based on adaptive expectation model for TAIEX forecasting. *Expert Systems with Applications, 34*, 1126–1132.

Cheng, C.-H., Cheng, G.-W., & Wang, J.-W. (2008). Multi-attribute fuzzy time series method based on fuzzy clustering. *Expert Systems with Applications, 34*, 1235–1242.

Egrioglu, E., Aladag, C. H., Yolcu, U., Basaran, M. A., & Uslu, V. R. (2009). A new hybrid approach based on SARIMA and partial high order bivariate fuzzy time series forecasting model. *Expert Systems with Applications, 36*, 7424–7434.

Egrioglu, E., Aladag, C. H., Yolcu, U., Uslu, V. R., & Basaran, M. A. (2009). A new approach based on artificial neural networks for high order multivariate fuzzy time series. *Expert Systems with Applications, 36*, 10589–10594.

Egrioglu, E., Aladag, C. H., Yolcu, U., Uslu, V. R., & Basaran, M. A. (2010). Finding an optimal interval length in high order fuzzy time series. *Expert Systems with Applications, 37*, 5052–5055.

Egrioglu, E., Aladag, C. H., Başaran, M. A., Uslu, V. R., & Yolcu, U. (2011). A new approach based on the optimization of the length of intervals in fuzzy time series. *Journal of Intelligent and Fuzzy Systems, 22*, 15–19.

Egrioglu, E., Aladag, C. H., Yolcu, U., Uslu, V. R., & Erilli, N. A. (2011). Fuzzy time series forecasting method based on Gustafson–Kessel fuzzy clustering. *Expert Systems with Applications, 38*, 10355–10357.

Huarng, K. (2001). Heuristic models of fuzzy time series for forecasting. *Fuzzy Sets and Systems, 123*(3), 369–386.

Huarng, K., & Yu, T. H.-K. (2006a). The application of neural networks to forecast fuzzy time series. *Physica A, 363*, 481–491.

Huarng, K., & Yu, T. H.-K. (2006b). Ratio-based lengths of intervals to improve fuzyy time series forecasting. *IEEE Transactions on Systems, Man., and Cybernetics-Part B: Cybernetics, 36*, 328–340.

Jilani, T. A., & Burney, S. M. A. (2008). A refined fuzzy time series model for stock market forecasting. *Physica A, 387*, 2857–2862.

Lee, L. W., Wang, L. H., & Chen, S. M. (2007). Temperature prediction and TAIFEX forecasting based on fuzzy logical relationships and genetic algorithms. *Expert Systems with Applications, 33*, 539–550.

Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *The Quarterly of Applied Mathematics, 2*, 164–168.

Li, S.-T., Cheng, Y.-C., & Lin, S.-Y. (2008). A FCM-based deterministic forecasting model for fuzzy time series. *Computers and Mathematics with Applications, 56*, 3052–3063.

Nguyen, T. H., & Wu, B. (2006). Fundamentals of statistics with fuzzy data. In *Studies in Fuzziness and Soft Computing* (Vol. 198). Berlin: Springer. (ISSN: 1434–9922).

Song, Q., & Chissom, B. S. (1993a). Fuzzy time series and its models. *Fuzzy Sets and Systems, 54*, 269–277.

Song, Q., & Chissom, B. S. (1993b). Forecasting enrollments with fuzzy time series – Part I. *Fuzzy Sets and Systems, 54*, 1–10.

Song, Q., & Chissom, B. S. (1994). Forecasting enrollments with fuzzy time series – Part II. *Fuzzy Sets and Systems, 62*(l), 1–8.

Song, Q., Leland, R. P., & Chissom, B. S. (1995). A new fuzzy time-series model of fuzzy number observations. *Fuzzy Sets and Systems, 73*, 341–348.

Sullivan, J., & Woodall, W. H. (1994). A comparison of fuzzy forecasting and Markov modeling. *Fuzzy Sets and Systems, 64*(3), 279–293.

Yolcu, U., Egrioglu, E., Uslu, V. R., Basaran, M. A., & Aladag, C. H. (2009). A new approach for determining the length of intervals for fuzzy time series. *Applied Soft Computing, 9*, 647–651.

Yu, T. H.-K. (2005). Weighted fuzzy time series models for TAIEX forecasting. *Physica A, 349*, 609–624.

Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control, 8*, 338–353.

Zhang, G., Patuwo, B. E., & Hu, Y. M. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting, 14*, 35–62.

Zurada, J. M. (1992). *Introduction of artificial neural systems*. St. Paul: West Publishing.