East African Scholars Journal of Engineering and Computer Sciences
Abbreviated Key Title: East African Scholars J Eng Comput Sci
ISSN 2617-4480 (Print) | ISSN 2663-0346 (Online) |
Published By East African Scholars Publisher, Kenya

Volume-2 | Issue-8 | Aug-2019 |

Review Article

A Forecasting Multi-Observations Model Based On Fuzzy Time Series and Neural Network

Nghiem Van Tinh
Thai Nguyen University of Technology – Thai Nguyen University, ThaiNguyen, Vietnam

*Corresponding Author
Nghiem Van Tinh

Abstract: The application of fuzzy time series models to forecast has been drawing a lot of attention of researchers. This study proposes a fuzzy time series model to forecast with multiple observations. The model shows how to fuzzify multiple observations into a fuzzy set. Neural networks are applied for training and forecasting the consecutive fuzzy sets. The forecasting results are defuzzified into forecasts. We use Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) as forecasting target for the period from 2000 to 2006. The TAIEX is separated into in-sample and out-of-sample observations. The in-sample observations are used for training and the out-of-sample observations are for forecasting. We consider a hit when the forecast falls between the low and high intervals for the day to forecast. To demonstrate the performance, we calculate the root mean squared errors and hit percentages for the out-of-sample forecasts.

Keywords: Fuzzy time series (FTS), fuzzy relationships, neural networks, forecasting, TAIEX.

1. INTRODUCTION

Forecasting has been considered important in various problem domains. How to model multiple observations to facilitate forecasting seems even more interesting. Meanwhile, the application of fuzzy time series models to forecast has attracting attention. The forecasting results from some of these models have shown to outperform their conventional counterparts. Therefore, this paper proposes a model to forecast a problem with multiple observations, such as multiple stock index readings in a day. First, fuzzy time series have been applied to several domain problems and have been shown to forecast better (Song & Chissom, 1993, 1994; Chen, 1996). Second, neural networks have been very popular for modeling nonlinear data (Indro, Jiang, Patuwo, & Zhang, 1999; Wasserman, 1989). They have been applied to forecasting fuzzy time series and have performed better than some other models (Huarng & Yu, 2006). In particularly, they have been applied for modeling fuzzy logical relationships between consecutive observations and have been popular for their capability in modeling nonlinear relationships (Donaldson & Kamstra, 1996; Indro, Jiang, Patuwo & Zhang, 1999). In addition, there have been different applications of neural networks, including Dow Jones forecasting (Kanas, 2001).

In this study, the proposed model first fuzzifies multiple observations into a fuzzy set. Then, neural networks are applied for training and forecasting the consecutive fuzzy sets. The forecasting results are defuzzified into forecasts. We use Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) as forecasting target for years from 2000 to 2006. The TAIEX is separated into in-sample and out-of-sample observations. The in-sample observations are used for training and the out-of-sample observations are for forecasting.

The rest of this paper is organized as follows. Section 2 reviews the definitions of fuzzy time series and neural networks. Section 3 introduces the algorithm for the proposed model. Section 4 applies the proposed model to forecast TAIEX and conclusions are presented in Section 5.

Quick Response Code

Journal homepage:
http://www.easpublisher.com/easjecs/

Article History
Received: 13.07.2019
Accepted: 22.07.2019
Published: 04.08.2019

DOI: 10.36349/easjecs.2019.v02i08.001

2. FTS Definitions and Neural Network Models

2.1 Fuzzy Time Series

Fuzzy set theory was firstly developed by Zadeh in the 1965s to deal with uncertainty using linguistic terms. (Q. Song & B.S. Chissom, 1993) successfully proposed the fuzzy forecasting model by adopting the fuzzy sets for fuzzy time series. To avoid complicated max–min composition operations, (Chen, 1996) improved the fuzzy forecasting method by using simple arithmetic operations. Let $U=\{u_1,u_2,...,u_n\}$ be an universal set; a fuzzy set A of U is defined as $A=\{f_A(u_1)/u_1+...+f_A(u_n)/u_n\}$, where f_A is a membership function of a given set A, $f_A: U\rightarrow[0,1]$, $f_A(u_i)$ indicates the grade of membership of u_i in the fuzzy set A, $f_A(u_i) \in [0, 1]$, and $1\leq i \leq n$. General definitions of fuzzy time series are given as follows

Definition 1. Fuzzy time series

Let $Y(t)$ $(t = .., 0, 1, 2..)$, a subset of R, be the universe of discourse on which fuzzy sets $f_i(t)$ (i = 1,2…) are defined and if $F(t)$ be a collection of $f_i(t)$) (i = 1, 2…). Then, $F(t)$ is called a fuzzy time series on $Y(t)$ (t . . . . . 0, 1,2 . . . .).

Definition 2. Fuzzy logical relationship

If there exists a fuzzy relationship R(t-1,t), such that F(t) = F(t-1)∗R(t-1,t), where ∗ is an Max – min operator, then F(t) is said to be caused by F(t-1). (t) and F(t-1) can be denoted by F(t-1)⟶F(t) (Chen, 1996). Various operations have been applied to compute the fuzzy relationship between F(t) and F(t-1) (Hwang, Chen & Lee, 1998; Song & Chissom, 1994; Sullivan & Woodall, 1994). Chen (1996) suggested that when the maximum degree of membership of F (t) belongs to A_i, F(t) is considered to be A_i. So, if let A_i = F(t) and A_j = F(t-1), the relationship between F(t) and F(t -1) is denoted by fuzzy logical relationship $A_i \rightarrow A_j$ where A_i and A_j refer to the current state or the left hand side and the next state or the right-hand side of fuzzy time series. The advantage of this proposal is that it simplifies the complicated calculations, but it falls into a hurdle when the out-of-sample observations do not appear in the in-sample observations. And this problem can be solved by the proposed model in this study.

Definition 3. Two – variable fuzzy logical relationship

Let F_1 and F_2 be two fuzzy time series. Suppose $F_1(t-1) = A_i$, $F_2(t-1) = B_k$ and F(t) = A_j. Two - variable FLR is defined as $A_i, B_k \rightarrow A_j$, where A_i, B_k are referred to as the left-hand side and A_j, as the right-hand side of the Two - variable FLR.

2.2 Neural Network Models

Neural networks have been successfully applied to the forecasting of different applications as credit ratings (Kumar & Bhattacharya, 2006), Dow Jones forecasting (Kanas, 2001), customer satisfaction analysis (Gronholdt & Martensen, 2005), stock ranking (Refenes, Azema-Barac & Zapranis, 1993), and tourism demand (Law, 2000; Law & Au, 1999; Martin & Witt, 1989; Palmer, Montaño & Sesé, 2006). The nonlinear structures of neural networks have been very useful in forecasting and they have been shown to discover nonlinear relationships among the observations (Donaldson & Kamstra, 1996; Indro, Jiang, Patuwo & Zhang, 1999). Hence, this study chooses the neural network to establish the fuzzy relationships in a multi-variable fuzzy time series, which is also nonlinear.

A simple neural network is listed in Figure. 1. Neural networks consist of an input layer, an output layer, and one or more hidden layers. Each of the layers contains nodes, and these nodes of two consecutive layers are connected with each other. The leftmost layer is the input layer, consisting of input nodes. Each input node is for an input variable. Hence, the number of input variables is equal to the number of input nodes. The rightmost layer is the output layer, consisting of output nodes. Similarly, each output node is for an output variable, with the number of output variables being equal to the number of output nodes.

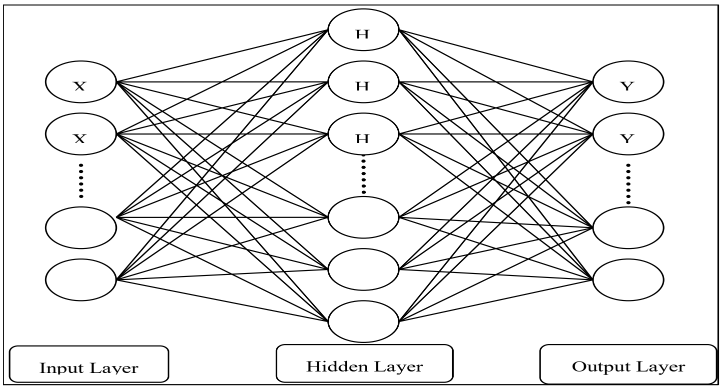

Figure1. A neural network structure

3. A Forecasting Model Based on FTS and Neural Networks

The main steps for the FTS forecasting model based on neural networks are shown as follows:

Step 1: Partition the universe of discourse into intervals and define the fuzzy sets for FTS.

Step 2: Determine a fuzzy observation (*l*, *m*, *h*) to represent all the observations at time *t*.
Step 3: Calculate the fuzzy logical relationship between each fuzzy observation.

Step4: Fuzzy Relationship=Max ∗ min (fuzzy observation, each fuzzy set of the fuzzy time series), where those degrees of memberships below 0.0 are considered as 0.0.

Step5: Neural network training for the consecutive fuzzy relationships.
($\mu^1_{t-1,t}$, $\mu^2_{t-1,t}$,…, $\mu^k_{t-1,t}$, …); $1 \le k \le n$; n is number of observations

Step6: Defuzzify the forecasting value for fuzzy logical relationship Forecast (t) $= \frac{\sum_{k=1} \mu^k_{t-1,t} \times m^k}{\sum_{k=1} \mu^k_{t-1,t}}$

Where forecast (t) is the forecast for *t*; $\mu^k_{t-1,t}$ is the forecasted degrees of membership and m^k is the corresponding midpoints of the interval, $\mu^k_{t-1,t}$.

Step7: Evaluate the performance of forecasting model
Two indicators are used to show the performance of the proposed model. Root mean squared errors (RMSEs) have been used to measure the performance of fuzzy time series (Huarng & Yu, 2006).

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(F_i - R_i)^2} \qquad (1)$$

Where, R_i notes actual data (observation) on date i, F_i forecasted value on date i, n is number of the forecasted data
In addition, the percentage of the defuzzified forecast falling into the range of its forecasting fuzzy observation.

4. APPLY THE PROPOSED MODEL FOR FORECASTING TAIEX

This study uses The Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) as forecasting target to demonstrate how the proposed model can be used to model the time series with multiple observations.

Step1. Partition the universe of discourse into intervals and define the fuzzy sets for FTS.

We set the universe of discourse as [0, 12000]. The length of interval is set as 1000. Hence, we have the following intervals: u_1= [0, 1000], u_2= [1000, 2000], …u_n= [1000*(n-1), 1000*n]. We define a fuzzy time series as follows (Chen, 1996).

$$A_1 = \frac{1}{u_1} + \frac{0.5}{u_2} + \frac{0}{u_3} + \cdots + \frac{0}{u_n}$$
$$A_2 = \frac{0.5}{u_1} + \frac{1}{u_2} + \frac{0.5}{u_3} + \cdots + \frac{0}{u_n}$$

$$A_n = \frac{0}{u_1} + \frac{0}{u_2} + \cdots + \frac{1}{u_{n-1}} + \frac{0.5}{u_n}$$

For simplicity, the membership values of fuzzy set A_i either are 0, 0.5 or 1. The value 0, 0.5 and 1 indicate the grade of membership of u_j ($1 \le j \le n$) in the fuzzy set A_i ($1 \le i \le n$).

Where, the symbol '+' denotes fuzzy set union, the symbol '/' denotes the membership of u_j which belongs to A_i.

Step2. Determine a fuzzy observation
For each day, there are multiple observations for stock index. We choose low, high, and close to determine a fuzzy observation for that day. For example, on January 4, 2001, the low is 5026.36, the close is 5137.21, and the high is 5171.28. Hence, the fuzzy observation for January 4, 2001 is (5026.36, 5026.36, and 5171.28). Similarly, we can obtain the fuzzy observations for the indices of other days.

Step3. Calculate the fuzzy logical relationships
Following the equation, the fuzzy relationship for January 4, 2001, is calculated as
(0.0, 0.0, 0.0, 0.0, 0.87, 1.0, 0.17, 0.0, 0.0, 0.0, 0.0, 0.0)
Similarly, other fuzzy relationships can be calculated.

Step4. Neural network training
Based on the constructing of neural network in Figure 1. The fuzzy relationships of *t*-1 and *t* form the input layer with X_i representing an input node, and output layer with Y_i representing an output node, of the neural network, respectively. In Figure 1, H_i represents the hidden node of the hidden layer. For example, the fuzzy relationship of January 4, 2001, is (0.0, 0.0, 0.0, 0.0, 0.87, 1.0, 0.17, 0.0, 0.0, 0.0, 0.0, 0.0) and that of January 5 is (0.0, 0.0, 0.0, 0.0, 0.76, 1.0, 0.32, 0.0, 0.0, 0.0, 0.0, 0.0). In this case, the former can be the input and the latter can be the output of the neural network.

The observations are sampled randomly. Each year, we sample 20% from all the observations as out-of-sample, the rest are in-sample. The fuzzy relationships of the stock index of the in-sample observations of year 2001 can then be established via the neural network training.

Step 5. Neural network forecasting

After neural network training, we can proceed to forecast the out-of-sample observations. For example, the fuzzy observation for January 10, 2001 was (5352.32, 5367.24, 5534.23) whose fuzzy relationship is (0.0, 0.0, 0.0, 0.0, 0.62, 1.0, 0.42, 0.0, 0.0, 0.0, 0.0, 0.0). Hence, if we use that fuzzy relationship as input to the trained neural network, the output of the neural network is the forecast for the fuzzy relationship on January 11, 2001:

$(\mu^1_{t-1,t}, \mu^2_{t-1,t}, \ldots)$ = (0.0, 0.0, 0.000082, 0.002487, 0.599091, 0.998332, 0.463215, 0.000182, 0.0, 0.0, 0.0, 0.0)

Step 6. Defuzzify forecasting output value

Weighted average by the fuzzy relationship and the midpoints of the interval is used to defuzzify the fuzzy relationship:
(0.000082×2500+0.002487×3500+0.599091×4500+0.998332×5500+0.463215×6500+0.000182×7500)/6=5436.65

Step 7. Evaluate the performance of forecasting model

Based on formula (1), the forecasting RMSE for year 2000 by the proposed model is 160.56; the hit percentage is 47%. The RMSEs and hit percentages for other years from 2000 to 2006 are shown in Table 1.

Table 1. The performance of forecasting model

Year	2000	2001	2002	2003	2004	2005	2006
RMSE	160.56	99.2	92.06	74.15	72.98	46.48	70.72
Hit percentage	47.2%	48.4%	66.3%	55.4%	62.4%	61.3%	55.2%

5. CONCLUSION

This paper presents a fuzzy time series model based on Neural networks to model multiple observations in forecasting . Neural networks are applies to establish the fuzzy relationships between consecutive observations. Stock index (TAIEX) forecasting is taken as the forecasting example to demonstrate the forecasting process. The observations are separated into in-sample and out-of-sample observations where the in-sample observations are used for neural network training and the out-of-sample for forecasting. The empirical results show that the average hit percentage is 56%.

REFERENCES

1. Chen S.-M. (1996). Forecasting enrollments based on fuzzy time series. *Fuzzy Sets and Systems*, vol. 81, no. 3, pp. 311-319,.
2. Donaldson R.G. & Kamstra M. (1996). Forecast combining with neural networks. Journal of Forecasting, vol. 15, no. 1, pp. 49-61.
3. Gronholdt L. & Martensen A. (2005). Analysing customer satisfaction data: A comparison of regression and artificial neural networks. International Journal of Market Research, vol. 47, no. 2, pp, 121-130.
4. Huarng K. & Yu H.-K. (2006). The application of neural networks to forecast fuzzy time series. Physica A, vol. 363, no. 2, pp. 481-491.
5. Hwang J.R., Chen S.-M., & Lee C.-H. (1998). Handling forecasting problems using fuzzy time series. Fuzzy Sets and Systems, vol. 100, no. 1-3, pp. 217-228.
6. Indro D. C., Jiang C. X., Patuwo B.E., & Zhang G. P. (1999). Predicting mutual fund performance using artificial neural networks. Omega, vol. 27, pp. 373-380.
7. Kanas A. (2001). Neural network linear forecasts for stock returns. International Journal of Finance and Economics, vol. 6, no. 3, pp. 245-54.
8. Kumar K. & Bhattacharya S. (2006). Artificial neural network vs. linear discriminant analysis in credit ratings forecast: A comparative study of prediction performances. Review of Accounting & Finance, vol. 5, no. 3, p. 216.
9. Law R. (2000). Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. Tourism Management, vol. 21, pp. 331-340.
10. Law R. & Au N. (1999). A neural network model to forecast Japanese demand for travel to Hong Kong. Tourism Management, vol. 20, pp. 89-97.
11. Martin C. A. & Witt S. F. (1989). Accuracy of econometric forecasts of tourism. Annals of Tourism Research, vol. 16, pp. 407-428.
12. Palmer A., Montaño J. J., & Sesé A. (2006). Designing an artificial neural network for forecasting tourism time series. Tourism Management, vol. 27, pp. 781-790.
13. Refenes A.N., Azema-Barac M., & Zapranis A.D. (1993). Stock ranking: Neural networks vs. multiple linear regression. IEEE International Conference on Neural Networks, pp. 1419-1426.
14. Song.Q & Chissom B.S. (1993). "Forecasting Enrollments with Fuzzy Time Series – Part I," Fuzzy set and system, vol. 54, pp. 1-9
15. Song Q. & Chissom B.S. (1994). Forecasting enrollments with fuzzy time series - part 2. Fuzzy

Sets and Systems, vol. 62, no. 1, pp. 1-8.

16. Sullivan J. & Woodall W.H. (1994). A comparison of fuzzy forecasting and Markov modeling. Fuzzy Sets and Systems, vol. 64, no. 3, pp. 279-293.
17. Wasserman, P. D. (1989). Neural computing: theory and practice. New York: Van Nostrand Reinhold.