A New Method For Stock Price Index Forecasting Using Fuzzy Time Series

Seyed Morteza Hosseini, Mahmoud Yahyazadeh Fard, Seyed Maryam Baboli

Department of Business management, Babol Branch, Islamic Azad University, Babol, Iran
Department of commercial management, Mazandaran University, Babolsar, Iran
Department of economic, Firozkoh Branch, Islamic Azad University, Firozkoh, Iran

Abstract: Time series forecasting is an active research area that has drawn considerable attention for applications in a variety of areas. Forecasting accuracy is one of the most important features of forecasting models. Nowadays, despite the numerous time series forecasting models which have been proposed in several past decades, it is widely recognized that financial markets are extremely difficult to forecast. Fuzzy time series have attracted attention of many researcher so that they have a good behavior with vague and incomplete data. One of the most new important methods in time series forecasting is using of genetic algorithm that have been minimum error. The most important weakness of using genetic algorithm is no have a mechanism for implementing with vague and incomplete data. This paper presents a new fuzzy time series to address this weakness. The proposed method is used for stock price forecasting.

Key words: genetic algorithms, weighted model, forecasts, fuzzy time series

INTRODUCTION

One of the important tools for development of any country is empowerment financial markets. Economic growth needs efficient capital market to allocate optimal capital resources. Forecasting future of market conditions and adopting proper methods to obtain more profits from the market is one of the main needs for industries and people. One of the tools for forecasting is time series and fuzzy time series.

A time series is a set or sequence of observed data arranged in chronological order and in an equally spaced time intervals such as daily or hourly air temperature. Time series occurs in many fields and the analysis of time series has got a wide application in areas like process control, economic forecasting, marketing, population studies, biomedical science and many more areas. Time series analysis uses systematic approaches to extract information and understand the characteristics of a physical system that creates the time series. There are a number of different approaches to deal with time series analysis including dynamic model building and performing correlations.

Fuzzy theory was originally developed to deal with problems involving linguistic terms (L.A. Zadeh, 1976; 1975; 1975). Time-series models had failed to consider the application of this theory until fuzzy time-series was defined by Song and Chissom (Q. Song, B.S. Chissom, 1993; 1994). They proposed the definitions of fuzzy time-series and methods to model fuzzy relationships among observations. The framework of Song and Chissom’s model includes six steps: (1) define and partition the universe of discourse; (2) define fuzz sets for the observations; (3) fuzzify the observations, (4) establish the fuzzy relationship, R, (5) forecast, and (6) defuzzify the forecasting results (Q. Song, B.S. Chissom, 1993; 1994).

Since (Q. Song and B. S Chissom, 1993) introduced fuzzy sets theory into time series analysis. Fuzzy time series have been applied to many areas, such as university enrollments (Q. Song and B. S Chissom, 1993; 1994), stock index forecasting (I-W. Liu, et al., 2010; I-I. Park, et al., 2010), temperature forecasting (L.-Y Hsu, et al., 2010), exchange rate (Y. Leu, et al., 2010) and tourism forecast (C-H. Wang and L.-C Hsu, 2008).

In the evolution of fuzzy time-series models, Chen, (1994) proposed another method to apply simplified arithmetic operations in forecasting algorithms rather than the complicated maximum–minimum composition operations, which are presented in Song and Chissom’s models. In subsequent research, Chen proposed several methods, such as high-order fuzzy relationships and genetic algorithms to improve his initial model (S.M. Chen, 2002; S.M. Chen, C.C. Hsu, 2004; S.M. Chen, N.Y. Chung, 2006; J.R. Hwang, et al., 1998).

The attractive of the stock market that is, every investor has the image to make millions in return from the financially buying and selling stock. Unfortunately, the reality of the nature of the market paints a less optimistic picture. Although, researchers had proposed many methods do try and do so, traditionally, the best performers have been the investors who use their professional knowledge of the markets to predict the next trend of the stock price. Therefore, a trustworthy forecasting tool is extremely desired. Because, the more accuracy the forecasting tool does, the more profitable will be made.

The rest of this paper is organized as follows: section 2 first presents two forecasting methods including weighted Fuzzy Time Series and Time series prediction based on genetic algorithms then it presents proposed
method. Section 3 includes implication of proposed method for a dataset. Section 4 contains results, discussion, conclusion and future research guidelines.

MATERIAL AND METHODS

2.1. weighted Fuzzy Time Series:
According to (Hui-Kuang Yu, 2005), one of the disadvantages of conventional fuzzy time series is ignoring number of Fuzzy Logic Relations (FLR), that this subject can increase uncertainty in the predictions. Yu (L.A. Zadeh, 1976) presented a model for reducing this uncertainty as following.
At first the universal set is defined as \( U = [\text{Start, End}] \) and then it converts into some partitions. After determination of linguistic variables \( (A_1) \), fuzzy relations make between them. In the next step of establishing FLRs, the appropriate weights are allocated to them based on expert’s knowledge or order of FLR. Then based on the assigned weights, Final \( (t) \) is calculated. Finally, mean square error (MSE) is computed. Less error, better forecasting. The main advantage of this model to previous fuzzy time series is better considering the uncertainty, and also lower than error.

2.2. Time Series Prediction Based On Genetic Algorithms:
Using genetic algorithms for fuzzy time series were presented by Chen (Shyi-Ming Chen, Nien-Yi Chung, 2006). The advantage of this method is determination of variable lengths for the universal set (Shyi-Ming Chen, Nien-Yi Chung, 2006).
At first the universal set \( U \) is defined as \( U = [D_{\text{min}}-D_1, D_{\text{max}}+D_2] \) based on \( D_{\text{min}} \) and \( D_{\text{max}} \) then it to partition in equal intervals. In the next step, after the production of chromosomes, for each of them a certain number of randomly genes are generated and linguistic values is defined as \( A_1 \). After fuzzification, appropriate fuzzy relations are established between them and in the next step, forecasting are carried out. Then mutation and crossover is implemented and finally, the mean square error (MSE) is calculated for each chromosome. In this step, the numbers of chromosomes that have lower MSE are selected as next population, then the new series of data is produced and after re-implementation mutation and crossover MSE is calculated. This process continues until the error rate is closed to zero. Finally, the chromosome with minimum MSE is selected as best forecasting.
The most advantage of this method is the dynamic length of the range. This method will be increased the output accuracy. But the disadvantage of this method is no having policies for dealing with uncertainty. For addressing this shortcoming, presented paper combines this method with the fuzzy weighted time series recommended by Yu.

2.3. Proposed Method:
One of the best methods that have the lower error than other methods are fuzzy time series is combined with genetic algorithm. But this method doesn’t consider repeated FLRs for improving forecasting accuracy that this approach increase uncertainly and decrease forecasting error. This paper combines defined weight in Yu model with Chen model to increase uncertainly and decrease forecasting error.

Algorithms that deal with uncertainty must be implemented in the next step of Chen algorithm. This algorithm is described as follows. The overall flowchart of the proposed method is shown in Figure 1.
1- If there is only one fuzzy relation, there isn’t uncertainty and Eq. (1) is considered:

\[
\text{Group} 1: A_1 \rightarrow A_2
\]  

(1)

If there are some of fuzzy relations as Eq. (2), there is a need to have a mechanism to deal with uncertainty

\[
\text{Group} 2: A_1 \rightarrow A_2, A_3
\]  

(2)

Dealing with the uncertainty is carried out in the fourth stage that is the weight assigned stage.
2- defuzzification
For defuzzification, if \( f(t) \) be \( A_{j_1}, A_{j_2}, ..., A_{j_k} \), non-fuzzy matrix is a matrix with midpoints of \( A_{j_1}, A_{j_2}, ..., A_{j_k} \) as \( M(t) = [m_{j_1}, m_{j_2}, ..., m_{j_k}] \) that \( M(t) \) is non-fuzzy prediction of \( F(t) \)
3- Weight allocation
To determine the weights we have: if \( f(t) = A_{j_1}, A_{j_2}, ..., A_{j_k} \) the determinant weight for \( A_{j_1}, A_{j_2}, ..., A_{j_k} \) are as \( W_1, ..., W_k \). The likelihood of matrix is defined as:
\[ W(t) = [w_1, w_2, \ldots, w_k] = \left[ \frac{w_1}{\sum_{h=1}^{k} h}, \frac{w_2}{\sum_{h=1}^{k} h}, \ldots, \frac{w_k}{\sum_{h=1}^{k} h} \right] \]

4- Forecasting
Forecasting is obtained by non-fuzzy matrix and weighted matrix as follows:
\[ nal(t) = M(t) \times W(t)^T \]

\[ \text{Divide amount of changing by the number of range} \]

\[ \text{By using genetic algorithm determine best length for each range} \]

\[ \text{By using weighted times series determine forecasting of stock price} \]

Fig. 1: proposed method

RESULT AND DISCUSSION

To demonstrate the effectiveness of the proposed model, large amounts of data are needed. For this reason we use the daily Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) closing prices covering the period from 1990 to 1999. TAIEX is the Technical Assistance and Information Exchange instrument managed by the Directorate-General Enlargement of the European Commission. TAIEX supports partner countries with regard to the approximation, application and enforcement of EU legislation. It is largely demand driven and facilitates the delivery of appropriate tailor-made expertise to address issues at short notice. An empirical analysis is conducted and the respective performances of the methods are compared with Chen and Yu methods from the literature. The TAIEX closing prices dataset covering the period from 1990 to 1999 are used in our empirical analysis; the dataset contains 2824 objects with only one attribute, closing price. The previous ten months of each year, from January to October, is used for training, and the last two months of each year, November to December, is for testing (L.A. Zadeh, 1975).

The proposed method uses the linguistic value (seven) for forecasting. The forecasting performances are compared with Chen’s and Yu’s methods, and produce a performance comparison table and figure as shown in Table 1 and Figure 2 to Figure 4.

Chen (1996) presented a new method to forecast university enrollments based on fuzzy time series. The data of historical enrollments of the University of Alabama shown in Song and Chissom (1993a, 1994) were adopted to illustrate the forecasting process of the proposed method. The robustness of the proposed method was also tested. The proposed method not only could make good forecasts of the university enrollments, but also could make robust forecasts when the historical data were not accurate. The proposed method was more efficient than the one presented in Song and Chissom (1993a) due to the fact that the proposed method used simplified arithmetic operations rather than the complicated max-min composition operations presented in Song and Chissom.

Yu (2005) proposed weighted fuzzy time series models to resolve two issues in fuzzy time series forecasting: recurrence and weighting. The TAIEX from 1990 to 1999 was used as the forecasting target. Both average- and distribution-based lengths of intervals were chosen as the lengths of intervals. The forecasting results showed that the weighted models outperform Chen’s model. The weighted models could thus be applied to improve fuzzy time series forecasting.

In the proposed method, From empirical results indicated that the proposed method using seven linguistic values wins 9 out of 10 on the comparison. It was obvious that the proposed method outperforms Yu’s and Chen’s methods.

<table>
<thead>
<tr>
<th>Table 1: Performance comparison table</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chen's method (1996)</td>
</tr>
<tr>
<td>Yu's method (2004)</td>
</tr>
<tr>
<td>Proposed methods</td>
</tr>
</tbody>
</table>
Fig. 2: stock price index forecasting using weighted times series, MSE=0.176

Fig. 3: stock price index forecasting using time's series combined with genetic algorithm, MSE=0.168

Fig. 4: stock price index forecasting using proposed method, MSE=0.127
Conclusion:
In this research, we combine Chen model that is one of the most accurate models in fuzzy time series models with Yu model to improve the accuracy of forecasting. The results of experiment on test data from the currency market shows that this combined approach have to be more careful from Yu and Chen model. The practical results obtained show that the proposed model have a good application.

REFERENCES