Knowledge-Based Systems 46 (2013) 12–21

Contents lists available at SciVerse ScienceDirect

Knowledge-Based Systems

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

# High-order fuzzy-neuro expert system for time series forecasting

Pritpal Singh *, Bhogeswar Borah

*Department of Computer Science and Engineering, Tezpur University, Tezpur 784 028, Assam, India*

ARTICLE INFO

*Article history:*
Received 18 July 2012
Received in revised form 6 December 2012
Accepted 25 January 2013
Available online 21 March 2013

*Keywords:*
Fuzzy time series
High-order
Temperature
Stock exchange
Interval
Fuzzy logical relation
Artificial neural network

ABSTRACT

In this article, we present a new model based on hybridization of fuzzy time series theory with artificial neural network (ANN). In fuzzy time series models, lengths of intervals always affect the results of forecasting. So, for creating the effective lengths of intervals of the historical time series data set, a new "Re-Partitioning Discretization (RPD)" approach is introduced in the proposed model. Many researchers suggest that high-order fuzzy relationships improve the forecasting accuracy of the models. Therefore, in this study, we use the high-order fuzzy relationships in order to obtain more accurate forecasting results. Most of the fuzzy time series models use the current state's fuzzified values to obtain the forecasting results. The utilization of current state's fuzzified values (right hand side fuzzy relations) for prediction degrades the predictive skill of the fuzzy time series models, because predicted values lie within the sample. Therefore, for advance forecasting of time series, previous state's fuzzified values (left hand side of fuzzy relations) are employed in the proposed model. To defuzzify these fuzzified time series values, an ANN based architecture is developed, and incorporated in the proposed model. The daily temperature data set of Taipei, China is used to evaluate the performance of the model. The proposed model is also validated by forecasting the stock exchange price in advance. The performance of the model is evaluated with various statistical parameters, which signify the efficiency of the model.



## 1. Introduction

Advance prediction of events like temperature, rainfall, stock price, population growth, and economy growth, are major scientific issues in the area of forecasting. Forecasting of all these events are tedious tasks because of their dynamic nature. Forecasting of all these events with 100% accuracy may not be possible, but the forecasting accuracy and the speed of forecasting process can be improved. So, in this article, we present a novel forecasting model, which is developed by hybridizing fuzzy time series theory with artificial neural network (ANN). The main aim of designing such a hybridized model is explained next.

For fuzzification of time series data set, determination of lengths of intervals of the historical time series data set is very important. In most of the fuzzy time series models [1–5], the lengths of the intervals were kept the same. No specific reason is mentioned for using the fix lengths of intervals. Huarng [6] shows that the lengths of intervals always affect the results of forecasting. So, for creating the effective lengths of intervals, a new "Re-Partitioning Discretization (RPD)" approach is incorporated in the proposed model.

After generating the intervals, time series data set is fuzzified based on the fuzzy time series theory. Most of the previous fuzzy time series models [1,7,2–4,8] use first-order fuzzy relationships to get the forecasting results. Many researchers show that high-order fuzzy relationships improve the forecasting accuracy of the models [9–14]. Therefore, in this work, we employ the high-order fuzzy relationships for obtaining the forecasting results.

Song and Chissom [1] adopted the following method to forecast enrollments of the University of Alabama:

$$Y(t) = Y(t-1) \circ R, \tag{1}$$

where $Y(t-1)$ is the fuzzified enrollment of year $(t-1)$, $Y(t)$ is the forecasted enrollment of year "$t$" represented by fuzzy set, "$\circ$" is the max–min composition operator, and "$R$" is the union of fuzzy relations. This method takes much time to compute the union of fuzzy relations $R$, especially when the number of fuzzy relations is more in (1) [15,16]. In 1996, Chen [3] used simplified arithmetic operations for defuzzification operation by avoiding this complicated max–min operations and their method produced better results than Song and Chissom models [1,7,2]. Most of the existing fuzzy time series models use Chen's defuzzification method [3] in order to obtain the forecasting results. However, forecasting accuracy of these models are not good enough. Also, previous fuzzy time series models use the current state's fuzzified values for forecasting. This approach, no doubt, improves the forecasting accuracy, but it degrades the predictive skill of the fuzzy time series models, because predicted values lie within the sample. So, for obtaining the forecasting results out of sample (i.e., in advance), we use the

* Corresponding author.
*E-mail addresses:* pritpal@tezu.ernet.in (P. Singh), bgb@tezu.ernet.in (B. Borah).

0950-7051/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.knosys.2013.01.030

previous state's fuzzified values (left hand side of fuzzy relations) in this model. To defuzzify these fuzzified values, an ANN based architecture is developed, and incorporated in this model. So, we have entitled this model as "High-order fuzzy-neuro time series forecasting model". The proposed model has the advantage that it can produce good forecasting results. We demonstrate the application of the proposed model using the following two real-world data set:

1. Daily average temperature data set of Taipei, China.
2. Daily stock exchange price data set of Bombay Stock Exchange (BSE), India.

The rest of this paper is organized as follows. In Section 2, we present related works for fuzzy time series models. In Section 3, we review the theory of fuzzy set with an overview of fuzzy time series. In Section 4, we give an overview of ANN along with its application in the proposed model. In Section 5, description of data set is provided. Section 6 shows the application of a new approach to find the length of intervals in the universe of discourse. The architecture of the proposed model and its training phases are presented in Sections 7 and 8 respectively. The performance of the model is assessed with various statistical parameters, which are discussed in Section 9. Empirical analysis for forecasting the daily temperature is presented in Section 10. Section 11 shows the application of the proposed model for forecasting the stock exchange price. The conclusions are discussed in Section 12.

## 2. Related works

Forecasting using fuzzy time series is applied in several areas including forecasting university enrollments, sales, road accidents and financial forecasting. In a conventional time series models, the recorded values of a special dynamic process are represented by crisp numerical values. But, in a fuzzy time series model, the recorded values of a special dynamic process are represented by linguistic values. Based on fuzzy time series theory, first forecasting model was introduced by Song and Chissom [1,7,2]. They presented the fuzzy time series model by fuzzy relational equations involving max–min composition operation and applied the model to forecast the enrollments in the University of Alabama. In 1996, Chen [3] used simplified arithmetic operations avoiding the complicated max–min operations and their method produced better results. Later, many studies provided some improvements to the fuzzy time series methods in determining the lengths of intervals, fuzzification process and defuzzification techniques. Hwang et al. [4] used the differences of the available historical data as fuzzy time series rather than direct usage of raw numeric values. Sah and Degtiarev also used a similar approach in [17]. Huarng tried to improve the forecasting accuracy based on determination of the length of intervals [6] and heuristic approaches [5]. Lee and Chou [18] forecasted the university enrollments with the average error rate less than Chen's method [3] by defining the supports of the fuzzy numbers that represent the linguistic values of the linguistic variables more appropriately.

Yu [19] proposed weighted fuzzy time series model to resolve issues of recurrence and weighting in fuzzy time series forecasting. Cheng et al. [8] used entropy minimization to create the intervals. They also used trapezoidal membership functions in the fuzzification process. Chang et al. [20] presented cardinality-based fuzzy time series forecasting model, which builds weighted fuzzy rules by calculating the cardinality of fuzzy relations. To enhance the performance of fuzzy time series models, Chen et al. [21] incorporates the concept of the Fibonacci sequence in the existing models as proposed by Song and Chissom [1,2] and Yu [19]. To obtain less number of intervals, Cheng et al. [22] proposed a model using fuzzy clustering technique to partition the data effectively. The K-means clustering algorithm has been applied to partition the universe of discourse in [23]. Chou et al. [24] forecasted the tourism demand based on hybridization of rough set with fuzzy time series. Singh and Borah [25] forecasted the university enrollments with the help of new proposed algorithm by dividing the universe of discourse of the historical time series data into different length of intervals. Recent advancement in fuzzy time series forecasting models can be found in [26–29].

Recently, many researchers have proposed various hybridization based models to solve complex problems in forecasting. For example, Hadavandi et al. [30] presented a new approach based on genetic fuzzy systems and ANNs for building a stock price forecasting expert system to improve the forecasting accuracy. Cheng et al. [31] proposed a new stock price forecasting model based on hybridization of genetic algorithm with rough set theory. Kuo et al. [32] hybridized the particle swarm optimization with fuzzy time series to adjust the lengths of intervals in the universe of discourse. Aladag et al. [33] introduced a new approach to define fuzzy relation in high order fuzzy time series using feed forward neural networks. Teoh et al. [34] proposed a fuzzy-rough hybrid forecasting model, where rules (fuzzy logical relationships) are generated by rough set algorithm. Pal and Mitra [35] proposed a rough-fuzzy hybridization scheme for case generation. They used the fuzzy set theory for linguistic representation of patterns and then obtained the dependency rules by using the rough set theory. For advance prediction of dwelling fire occurrence in Derbyshire (United Kingdom), Yang et al. [36] employed three approaches: logistic regression, ANN and Genetic Algorithm. Keles et al. [37] proposed a model for forecasting the domestic dept by Adaptive Neuro-Fuzzy Inference System. Chang et al. [38] developed a hybrid model by integrating K-mean cluster and fuzzy neural network to forecast the future sales of a printed circuit board factory. Huarng and Yu [39], and Yu and Huarng [40] presented a new hybrid model based on neural network and fuzzy time series to forecast TAIEX. Kuo et al. [41] and Huang et al. [42] introduced a new enrollments forecasting model based on hybridization of fuzzy time series and particle swarm optimization.

## 3. Fuzzy sets and fuzzy time series

In 1965, Zadeh [43] introduced fuzzy sets theory involving continuous set membership for processing data in presence of uncertainty. He also presented fuzzy arithmetic theory and its application [44–46]. In this section, we will briefly review fuzzy sets theory from [43] and fuzzy time series concepts from [1,7,2].

**Definition 3.1** (*Fuzzy Set [43]*). A fuzzy set is a class with varying degrees of membership in the set. Let $U$ be the universe of discourse, which is discrete and finite, then fuzzy set $A$ can be defined as follows:

$$A = \{\mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + \cdots\} = \Sigma_i \mu_A(x_i)/x_i, \tag{2}$$

where $\mu_A$ is the membership function of $A$, $\mu_A$: $U \to [0,1]$, and $\mu_A(x_i)$ is the degree of membership of the element $x_i$ in the fuzzy set $A$. Here, the symbol "+" indicates the operation of union and the symbol "/" indicates the separator rather than the commonly used summation and division in algebra respectively.

When $U$ is continuous and infinite, then the fuzzy set $A$ of $U$ can be defined as:

$$A = \left\{ \int \mu_A(x_i)/x_i \right\}, \quad \forall x_i \in U, \tag{3}$$