Information Sciences 327 (2016) 272–287

Contents lists available at ScienceDirect

Information Sciences

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

Fuzzy time series forecasting based on fuzzy logical relationships and similarity measures

Shou-Hsiung Cheng [a,b,*], Shyi-Ming Chen [c], Wen-Shan Jian [c]

[a] Department of Information Management, Chienkuo Technology University, Changhua, Taiwan
[b] Department of Kinesiology Health Leisure Studies, Chienkuo Technology University, Changhua, Taiwan
[c] Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan

ARTICLE INFO

Article history:
Received 29 March 2015
Revised 4 August 2015
Accepted 9 August 2015
Available online 22 August 2015

Keywords:
Fuzzy logical relationships
Fuzzy time series
K-means clustering algorithm
Particle swarm optimization
Similarity measures

ABSTRACT

In this paper, we propose a new fuzzy time series forecasting method for forecasting the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) based on fuzzy time series, fuzzy logical relationships, particle swarm optimization techniques, the *K*-means clustering algorithm, and similarity measures between the subscript of the fuzzy set of the fuzzified historical testing datum on the previous trading day and the subscripts of the fuzzy sets appearing in the current states of the fuzzy logical relationships in the chosen fuzzy logical relationship group. The particle swarm optimization techniques are used to get the optimal partition of the intervals in the universe of discourse. The *K*-means clustering algorithm is used to cluster the subscripts of the fuzzy sets of the current states of the fuzzy logical relationships to get the cluster center of each cluster and to divide the constructed fuzzy logical relationships into fuzzy logical relationship groups. The experimental results show that the proposed fuzzy forecasting method gets higher forecasting accuracy rates than the existing methods. The advantages of the proposed fuzzy forecasting method is that it uses the particle swarm optimization techniques to get the optimal partition of the intervals in the universe of discourse and uses the *K*-means clustering algorithm to cluster the subscripts of the fuzzy sets of the current states of the fuzzy logical relationships to get the cluster center of each cluster and to divide the constructed fuzzy logical relationships into fuzzy logical relationship groups for increasing the forecasting accuracy rates.

1. Introduction

Song and Chissom [51–53] proposed the concepts of fuzzy time series, where the values in a fuzzy time series are represented by fuzzy sets [66]. In recent years, some fuzzy forecasting methods have been presented based on fuzzy time series, such as the Index 100 in the stocks and bonds exchange market of Istanbul (IMKB) [1] forecasting, the Taiwan Futures Exchange (TAIFEX) forecasting [1,2,18,32,35,39,42,57], enrollments forecasting [1,2,6–8,11,14,15,18–22,28–33,35,37,40,41,44,46,48–50,52–56,60,61], the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) [67] forecasting [4–6,9–13,17,19,22,23,25,26,33–36,45,58–60,62–64], temperature prediction [9,16,18,26,42,45,50,57], inventory demand forecasting [11,19,35,61], the NTD/USD exchange rates forecasting [12,17,43], the IT project cost forecasting [21], the market price of shares of the State Bank of India (SBI) at the Bombay Stock Exchange (BSE) forecasting [30], the rice production forecasting [31], the number of patents granted in Taiwan

* Corresponding author. Tel.: +886 4 7111111, x 3635; fax: +886 4 7111142.
E-mail address: shcheng@ctu.edu.tw (S.-H. Cheng).

http://dx.doi.org/10.1016/j.ins.2015.08.024
0020-0255/

forecasting [46], the Shanghai stock index forecasting [48], the daily stock exchange price of the SBI forecasting [50], the car road accident in Belgium forecasting [55], the spot gold in Turkey forecasting [55], Germany's DAX stock index (DAXSIMV) forecasting [56], etc.

In [5], Cai et al. presented a fuzzy time series forecasting model combined with ant colony optimization and auto-regression techniques for forecasting the TAIEX. In [9], Chen and Chang presented a multi-variable fuzzy time series forecasting method for forecasting the temperature and the TAIEX based on fuzzy clustering techniques and fuzzy interpolation techniques. In [10], Chen and Chen presented a method to forecast the TAIEX based on fuzzy time series and fuzzy variation groups, where the TAIEX is called the main factor and the Dow Jones, the NASDAQ, the M1B, or their combination are called the secondary factors, respectively. In [13], Chen et al. presented a method for forecasting the TAIEX based on fuzzy time series and automatically generated weights of multiple factors. In [17], Chen et al. presented a method for forecasting the TAIEX and the exchange rates based on two-factors second-order fuzzy-trend logical relationship groups and particle swarm optimization techniques. In [36], Huarng et al. presented a multivariate heuristic model for fuzzy time-series forecasting for forecasting the TAIEX. In [62], Yu presented weighted fuzzy time series models for forecasting the TAIEX. In [63], Yu and Huarng presented a bivariate fuzzy time series model for forecasting the TAIEX. However, because the forecasting accuracy rates of the methods presented in [5,9,10,12,13,17,36,62,63] for forecasting the TAIEX are not good enough, we need to develop a new fuzzy forecasting method based on fuzzy time series, fuzzy logical relationships, particle swarm optimization (PSO) techniques [27,38], the *K*-means clustering algorithm [47] proposed by MacQueen, and similarity measures to get higher forecasting accuracy rates for forecasting the TAIEX and to overcome the drawbacks of the existing methods [5,9,10,12,13,17,36,62,63]. The PSO techniques are used to get the optimal partition of the interval in the universe of discourse of the TAIEX and the *K*-means clustering algorithm [47] proposed by MacQueen is used to cluster the subscripts of the fuzzy sets of the current states of the fuzzy logical relationships to get the cluster center of each cluster and to divide the constructed fuzzy logical relationships into fuzzy logical relationship groups.

In this paper, we propose a new fuzzy time series forecasting method for forecasting the TAIEX based on fuzzy time series, fuzzy logical relationships, particle swarm optimization (PSO) techniques [27,38], the *K*-means clustering algorithm [47] proposed by MacQueen, and similarity measures between the subscript of the fuzzy set of the fuzzified testing datum of the TAIEX on the previous trading day and the subscripts of the fuzzy sets of the current states of the fuzzy logical relationships in the chosen fuzzy logical relationship group. First, the proposed method uses the PSO techniques to get the optimal partition of the intervals in the universe of discourse based on the historical training data of the TAIEX. Then, it fuzzifies the historical training data of the TAIEX to construct fuzzy logical relationships. Then, it applies the *K*-means clustering algorithm to cluster the subscripts of the fuzzy sets of the current states of the fuzzy logical relationships to get the cluster center of each cluster and to divide the constructed fuzzy logical relationships into fuzzy logical relationship groups. Then, it calculates the distance between the subscript of the fuzzy set of the fuzzified historical testing datum of the TAIEX on the previous trading day and the cluster center of each obtained cluster, respectively. Then, it chooses the fuzzy logical relationship group which has the minimum distance between the subscript of the fuzzy set of the fuzzified historical testing datum of the TAIEX on the previous trading day and the cluster center of the cluster containing the subscripts of the fuzzy sets of the current states of the fuzzy logical relationships. Finally, it forecasts the historical testing datum of the TAIEX on the next trading day based on the subscript of the fuzzy set of the fuzzified historical testing datum of the TAIEX on the previous trading day and the subscripts of the fuzzy sets of the current states of the fuzzy logical relationships in the chosen fuzzy logical relationship group. The experimental results show that the proposed fuzzy forecasting method gets higher forecasting accuracy rates than the methods presented in [5,9,10,12,13,17,36,62,63]. The advantages of the proposed fuzzy forecasting method are that it uses the particle swarm optimization techniques to get the optimal partition of the intervals in the universe of discourse and uses the *K*-means clustering algorithm [47] proposed by MacQueen to cluster the subscripts of the fuzzy sets of the current states of the fuzzy logical relationships to get the cluster center of each cluster and to divide the constructed fuzzy logical relationships into fuzzy logical relationship groups for increasing the forecasting accuracy rates.

The rest of this paper is organized as follows. In Section 2, we briefly review basic concepts of fuzzy time series [51–53], fuzzy logical relationships [7] and the *K*-means clustering algorithm [47] proposed by MacQueen. In Section 3, we propose a new fuzzy time series forecasting method for forecasting the TAIEX based on fuzzy time series, fuzzy logical relationships, the *K*-means clustering algorithm [47] proposed by MacQueen, and similarity measures between the subscript of the fuzzy set of the fuzzified testing datum of the TAIEX on the previous trading day and the subscripts of the fuzzy sets appearing in the current states of the fuzzy logical relationships in the chosen fuzzy logical relationship group. In Section 4, we make a comparison of the experimental results of the proposed method with the ones of the existing methods [5,9,10,12,13,17,36,62,63] for forecasting the TAIEX. The conclusions are discussed in Section 5.

2. Preliminaries

2.1. Fuzzy time series

In the following, we briefly review some basic concepts of fuzzy time series [51,52], where the values in a fuzzy time series are represented by fuzzy sets [66]. A fuzzy set A in the universe of discourse U, $U = \{u_1, u_2, \ldots, u_n\}$, can be represented by

$$A = f_A(u_1)/u_1 + f_A(u_2)/u_2 + \cdots + f_A(u_n)/u_n,$$