Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 34 (2008) 2732–2738
Expert Systems with Applications
www.elsevier.com/locate/eswa

Constructing and applying an improved fuzzy time series model: Taking the tourism industry for example

Chao-Hung Wang [a], Li-Chang Hsu [b,*]

[a] Department of International Business, Ling Tung University, 1 Ling Tung Road, Nantun, Taichung 40852, Taiwan, ROC
[b] Department of Finance, Ling Tung University, 1 Ling Tung Road, Nantun, Taichung 40852, Taiwan, ROC

Abstract

This study develops an improved fuzzy time series models for forecasting short-term series data. The forecasts were obtained by comparing the proposed improved fuzzy time series, Hwang's fuzzy time series, and heuristic fuzzy time series. The tourism from Taiwan to the United States was used to build the sample sets which were officially published annual data for the period of 1991–2001. The root mean square error and mean absolute percentage error are two criteria to evaluate the forecasting performance. Empirical results show that the proposed fuzzy time series and Hwang's fuzzy time series are suitable for short-term predictions.

Keywords: Fuzzy time series; Heuristic; Forecasting

1. Introduction

There is a great variety of methodologies for short-term predicting presented in the literature. Frequently used quantitative techniques include ARIMA and econometric models (Goh & Law, 2002; Lim & McAleer, 2002). However, the econometric models need large samples (minimum 50 sample data), normal distribution, and stationary data trends, limiting their application validity. Recently, numerous scholars have developed new forecasting techniques to overcome the limitations of tradition statistical methods such as neural networks (Law, 2000; Law & Au, 1999). Although neural networks still require large sample data sets for training and to establish a learning procedure, they do not require making as many assumptions as do statistical methods.

Zadeh (1965) successfully applied fuzzy theory to different research fields, including decision-making, control theory, business analysis, and forecasting. The forecasting application of fuzzy theory was first presented by Song and Chissom (hereafter SandC) (1993a, 1993b) who developed the fuzzy time series using the enrollment at the University of Alabama from 1971 to 1992 as the sample set. In addition, there are two highly effective research directions. First, Chen (1996) presented these of arithmetic operations instead of the logic max–min composition, which was used by the S&C model. Huarng (2001) incorporated the heuristic rule with Chen's model into the judging criteria of future trends, making a heuristic fuzzy time series model. Chen (2002) and Lee, Liu, and Chen (2006) further developed the model originally presented in 1996, to be high-order fuzzy time seri es. In contrast to S&C's and Chen's fuzzy models, Hwang, Chen, and Lee (1998) incorporated the variation of forecasting into the fuzzy time series model, whereby the variation value plus the actual value of the last period yields the forecast value.

Fuzzy theory methodologies for short-term predicting have well been presented in the literature, including temperature (Chen & Hwang, 2000; Lee, Wang, & Chen, 2007a, Lee, Wang, & Chen, 2007b), finance (Lee, Wang, Chen, & Leu, 2006), and disruption prediction in Tokamak

* Corresponding author. Tel.: +886 4 2389 2088x3642; fax: +886 4 2386 4342.
E-mail address: ltctht87@mail.ltu.edu.tw (L.-C. Hsu).

0957-4174/$ - see front matter
doi:10.1016/j.eswa.2007.05.042

reactors (Versaci & Morabito, 2003); however, the study of the optimal density of intervals using first-order fuzzy time series was scarce. The research reported here seeks to address that smallest interval (i.e. largest density of interval) uses sub-intervals, then judging the data of interval is upward or downward by rule. The proposed fuzzy time series model that applies and compares the rationale of these fuzzy time series and to determine which is the best forecast model based on the empirical results.

2. Methodology

When economists forecast the future trend using time series data, they must carefully examine whether the time series were stationary. If a time series were nonstationary, that implies the data had a stochastic trend and would yield an incorrect forecast. However, this paper does not examine the characteristics of the time series data, because this study uses a different methodology from that of economists. The following subsections will introduce a novel fuzzy time series model.

The forecasting of the $t+1$ period is compared to that of the t period either upward or downward. Therefore, a novel fuzzy time series model states that the forecast should use a logical relationship to judge the upward or downward movement of the forecast curve, and then yield the forecast value. The main difference between the proposed model and other fuzzy models is that the proposed model forecasts the trend of forecast curve by mean of changing length in each interval of the universes of discourse and using the differences of variations. The proposed fuzzy model was established by the following steps:

Step 1: Define the universe of discourse. The universe U is defined as, $U=[D_{\min}-D_1, D_{\max}+D_2]$, where $D_{\min}$ and $D_{\max}$ denote the minimum and maximum number of units among the historical data, respectively; and D_1 and D_2, which divide the U into intervals of equal length, are two proper positive numbers.

Step 2: Calculate the density of intervals. The number of intervals depends on the amount of data. This paper divides the largest density of interval into three sub-intervals of equal length (namely, the density of interval is 3). Furthermore, we divide the second largest density of interval into two sub-intervals of equal length (namely, the density of interval is 2). Finally, we can find the smallest interval. If the data were not distributed in the interval, we can delete the interval.

Step 3: Define the fuzzy set A_i and fuzzify the historical data using the intervals mentioned in step 2, $A_i=f_{Ai}(u_1)/u_1+f_{Ai}(u_2)/u_2+\cdots+f_{Ai}(u_t)/u_t$; where $f_{Ai}(u_i)$ denotes the grade of membership of u_i in A_i and $f_{Ai}(u_i)\in[0,1]$, the symbol " / " separates the membership degrees for each element degrees in the universe of discourse U, and the symbol " + " means "union" rather than the commonly used algebraic symbol of summation.

Step 4: Establish the fuzzy logical relationships given the fuzzifying the data. Fuzzy logical relationships can be found: $A_i\to A_q, A_i\to A_r,\ldots,$ where $A_i=F(t-1)$ and $A_j=F(t)$, and so $F(t)$ is said to be caused by $F(t-1)$.

Step 5: Forecast the future value. According to step 2, we divide the redistributed interval into four equal lengths; meanwhile, whether the forecasting curve, will be upward or downward, depends on the one-fourth point and three-fourth point within the interval.

The changing trend of forecasting is conducted using the following rules. Assuming the value of t period to be P_t, then the variation of value is denoted by $\Delta P_t = P_t - P_{t-1}$, and the difference of the variation is denoted by $\Delta\widetilde{P}_t = \Delta P_t - \Delta P_{t-1}$. Let $Q_t = |\Delta\widetilde{P}_t|\times 2 + t - 1$ period of value; $\tilde{Q}_t = |\Delta\widetilde{P}_t|/2 + t - 1$ period of value; $\widehat{Q_t} = |\Delta\widetilde{P}_t|/2$.

We propose the rules to decide the changing trend of the $t+1$ period as follows:

Rule 1:

(1) If Q_t belongs to A_j having the membership degree 1, then we judge the changing direction of the $t+1$ period to be upward.
(2) If $\widetilde{Q}_t$ belongs to A_j having the membership degree 1, then we judge the changing direction of the $t+1$ period to be downward.
(3) Excepting rule (1) and rule (2), the forecasts for the $t+1$ period will be the midpoint of A_j having the membership degree 1.

Rule 2: If we can exactly know the variation rather than $\Delta\widetilde{P}_t$, then for

(1) $\widehat{Q} > |\widehat{A_j}|/2$, we judge the change of $t+1$ period will be upward.
(2) $\widehat{Q} = |\widehat{A_j}|/2$, we judge the change of $t+1$ period will be constant.
(3) $\widehat{Q} > |\widehat{A_j}|/2$, we judge the change of $t+1$ period will be downward.

Meanwhile, the interval length of $\widehat{A_j}$ belongs to the fuzzy set A_j with the membership degree 1.

3. Empirical study for tourism demand forecasting

In order to explore the application of the proposed novel fuzzy time series model, the present paper uses the sample of Taiwan tourists to the USA as published by the Taiwan Tourism Bureau. Empirical analysis is conducted to compare the forecasting result of the proposed fuzzy time series with those of the other models.

3.1. The empirical of novel fuzzy time series

The forecasted method is presented as follows:

(1) Table 1 lists the actual visitors from historical data. The minimum value $D_{\min}$ and maximum value $D_{\max}$