Expert Systems with Applications 38 (2011) 9164–9171

Contents lists available at ScienceDirect

Expert Systems with Applications

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

The adaptive fuzzy time series model with an application to Taiwan's tourism demand

Ruey-Chyn Tsaur *, Ting-Chun Kuo

Department of Management Sciences and Decision Making, Tamkang University, Taipei, Taiwan, ROC

ARTICLE INFO

Keywords:
Adaptive fuzzy time series model
Forecasting
Fuzzy logic group
Tourism demand

ABSTRACT

In this study, an adaptive fuzzy time series model for forecasting Taiwan's tourism demand is proposed to further enhance the predicted accuracy. We first transfer fuzzy time series data to the fuzzy logic group, assign weights to each period, and then use the proposed adaptive fuzzy time series model for forecasting in which an enrollment forecasting values is applied to obtain the smallest forecasting error. Finally, an illustrated example for forecasting Taiwan's tourism demand is used to verify the effectiveness of proposed model and confirmed the potential benefits of the proposed approach with a very small forecasting error MAPE and RMSE.

1. Fuzzy time series

Tourism forecasts are usually generated by either quantitative or qualitative approaches. Quantitative approaches develop and employ mathematical models and theories to fit its natural phenomena. Unlike quantitative approaches, qualitative approaches involve in-depth understanding of human behavior and the reasons behind various aspects of behavior, thus its studies in the tourism field are very limited. Since most of tourism studies have focused on regression analysis to estimate the quantitative relationship between tourism demand and its determinants, or time series models to extrapolate historic trends in tourism demand into the future without considering the underlining causes of the trends. Before the 1990s, traditional regression approaches dominated the tourism forecasting literature, but this trend changed from the mid-1990s as more researchers began to use modern econometric techniques, such as cointegration and error correction models, to model and forecast tourism demand; these studies include Song, Witt, and Li (2003), Kulendran and King (1997) and Morley (1998). However, each method has its own particular advantages/disadvantages, when the collected data are not enough to model regression model or time series model, or there exists fuzzy time series data, the statistical quantitative methods are usually failure to have smaller forecasting error. In order to improve forecasting accuracy, a fuzzy time series model is adopted in this study in order to provide a much more flexible examination for managing smaller data set or fuzzy data.

An alternative approach, fuzzy time series model (Song & Chissom, 1993a, 1994, 1993b) have been developed and applied in forecasting as if the given datum is in linguistic terms or smaller than fifty data. Song and Chissom (S&C in abbreviation) were the pioneers of studying fuzzy time series model in 1993, then fuzzy time series model had drawn much attention to the researchers. For model modifications, Chen (1996) focused on the operator used in the model and simplified the arithmetic calculations to improve the composition operations and further introduced a concept of fuzzy logical groups to improve the forecast; Huarng (2001) made a study on the effective length of intervals to improve the forecasting in fuzzy time series; Tsaur, O Yang, and Wang (2005) made an analysis of fuzzy relations in fuzzy time series on the basis of entropy of the system used it to determine the minimum value of invariant time index t to minimize errors in the forecasted values of enrollments; Cheng, Cheng, and Wang (2008) introduces a novel multiple-attribute fuzzy time series method based on fuzzy clustering in which fuzzy clustering are integrated in the processes of fuzzy time series to partition datasets objectively and enable processing of multiple attributes. For forecasting with applications, Yu (2005) proposed a weighted method to forecasting the TAIEX to tackle two issues, recurrence and weighting, in fuzzy time-series forecasting; Huarng and Yu (2006) applied a back propagation neural network to handle nonlinear forecasting problems in stock price forecasting. Chen (1996) presented high-order fuzzy time series based on multi-period adaptation model for forecasting stock markets. Further, Chen and Hwang (2000), Wang and Chen (2007), Lee, Wang, and Chen (2007) proposed methods for temperature prediction and TAIFEX forecasting based on their proposed fuzzy time series models. As we know, every prediction model is designed with the hope to obtain the characteristics of the system. The more the information that relate to the system dynamics are considered,

* Corresponding author.
E-mail address: rctsaur@yahoo.com.tw (R.-C. Tsaur).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.01.059

the better the prediction will be. In this paper, Markov chain based on statistical method is incorporated with the fuzzy time series model to further enhance the predicted accuracy. However, the above studies do not provide the user with an efficient method to significantly reduce its forecasting error. In practice, we need to known how much of a fluctuation in demand is due to randomness and how much of it is due to a shift in the base. Therefore a method is needed to help exploit the information gathered during the past periods and then improve the forecast accuracy where several adaptive forecasting methods have been developed for this purpose (e.g. Chen, Cheng, & Teoh, 2008; Cheng, Chen, Teoh, & Chiang, 2008). Therefore, in order to decrease to forecasting error in fuzzy time series model, we adopt an adaptive method to hybridize fuzzy time series model which allows a smoothing parameter to change over time and adapt to changes the characteristics of the time series data.

In Section 2, the basic concept of fuzzy time series model is introduced. Next, an adaptive fuzzy time series model is proposed and d an enrollment forecasting is illustrated in Section 3. Besides, the accuracy and robustness for the proposed method are evaluated and discussed with an illustration about Taiwan's tourism demand in Section 4. Finally, the summary and conclusions is drawn in Section 5.

2. Fuzzy time series model

In this section, the basic fuzzy time series model concepts are described as follows using the example illustrated for enrollment forecasting in the Alabama University. Let U be the universe of discourse, $U = \{u_1, u_2, \ldots, u_n\}$, and let A_i be a fuzzy set in the universe of discourse$U$ defined as Eq. (1) as below:

$$A_i = \mu_{A_i}(u_1)/u_1 + \mu_{A_i}(u_2)/u_2 + \cdots + \mu_{A_i}(u_n)/u_n, \tag{1}$$

where μ_{A_i} is the membership function of the fuzzy set A_i, such that $\mu_{A_i} : U \to [0,1]$, and $\mu_{A_i}(u_j)$ represents the grade of membership of u_j in A_i where $\mu_{A_i}(u_j) \in [0,1]$.

Definition 1. If $Y(t)$ $(t = 1,2,\ldots,n)$ is a subset of R^1 in which the universe of fuzzy sets $f_i(t)$ $(i = 1,2,\ldots,m)$ are defined and let $F(t)$ be collection of $f_i(t)$ $(i = 1,2,\ldots,m)$, then $F(t)$ is called a fuzzy time series on $Y(t)$ $(t = 1,2,\ldots,n)$.

Definition 2. If any two fuzzy sets $f_i(t)$ and $f_j(t-1)$ are considered as matrices $[f_i(t)]_{m\times 1}$ and $[f_j(t-1)]_{1\times m}$, respectively, then the fuzzy relation matrix between two matrices could be the max–min composite as below

$$R_{ij}(t,t-1) = \max_j \min\{f_{i1}(t), f_{1j}(t-1)\}_{\times 1}, \quad \forall\, i,j = 1,2,\ldots,m.$$

Definition 3. If for any $f_j(t) \in F(t)$ and $f_i(t-1) \in F(t-1)$ there exist a fuzzy relation $R_{ij}(t,t-1)$ such that $f_j(t) = f_i(t-1) \circ R_{ij}(t,t-1)$, where '$\circ$' is the max–min composition operator, then $F(t)$ is said to be caused only by $F(t-1)$ and is denoted by $f_i(t-1) \to f_j(t)$ or by $F(t-1) \to F(t)$.

Definition 4. If $F(t)$ is a fuzzy time series, $F(t) = F(t-1)$ for any t and $F(t)$ has only finite elements $f_i(t)$ $(i = 1,2,\ldots,m)$, then

$$R(t,t-1) = f_i(t-1) \times f_j(t) \cup f_i(t-2) \times f_j(t-1) \cup \cdots \cup f_i(t-n) \times f_j(t-n+1), \quad \forall n > 0.$$

Definition 5. Let $F(t-1) = A_i$ and $F(t) = A_j$. The relationship between two consecutive observations, $F(t-1)$ and $F(t)$, referred to as a fuzzy logical relationship can be denoted by $A_i \to A_j$, where A_i is called the left-hand side and A_j the right-hand side.

Table 1
Enrollment data of Alabama University.

Year	Historical data $Y(t)$	u_1	u_2	u_3	u_4	u_5	u_6	u_7
1971	13,055	1	0.5	0	0	0	0	0
1972	13,563	1	0.8	0.1	0	0	0	0
1973	13,867	1	0.9	0.2	0	0	0	0
1974	14,696	0.8	1	0.8	0.1	0	0	0
1975	15,460	0.2	0.8	1	0.2	0	0	0
1976	15,311	0.2	0.8	1	0.2	0	0	0
1977	15,603	0	0.6	1	0.6	0.1	0	0
1978	15,861	0	0.5	1	0.7	0.2	0	0
1979	16,807	0	0.1	0.5	1	0.9	0.2	0
1980	16,919	0	0.1	0.5	1	0.9	0.2	0
1981	16,388	0	0.2	0.8	1	0.5	0	0
1982	15,433	0.2	0.8	1	0.2	0	0	0
1983	15,497	0.2	0.8	1	0.2	0	0	0
1984	15,145	0.2	0.8	1	0.2	0	0	0
1985	15,163	0.2	0.8	1	0.2	0	0	0
1986	15,984	0	0.2	1	0.7	0.2	0	0
1987	16,859	0	0.1	0.5	1	0.8	0.1	0
1988	18,150	0	0	0.1	0.5	0.8	1	0.7
1989	18,970	0	0	0	0.25	0.55	1	0.8
1990	19,328	0	0	0	0.3	0.5	0.8	1

Then, the stepwise procedure proposed by S& C's model is described as below.

Step 1. Define the universe of discourse U for the historical data. First, we find the minimum data $D_{\min}$ and the maximum data $D_{\max}$ individually in the historical time series data, then we define the universal discourse U as $[D_{min} - D_1, D_{max} + D_2]$, where D_1 and D_2 are two proper positive numbers.

Step 2. Partition the universe of discourse into equal length of intervals: $u_1, u_2, \ldots, u_n$. The number of intervals will be in accordance with the number of fuzzy sets $A_1, A_2, \ldots, A_n$ to be considered.

Step 3. Define the fuzzy sets A_i on universe of discourse U in Step 2. If there are fuzzy sets$A_1, A_2, \ldots, A_n$, then the fuzzy sets A_i, $\forall i = 1,2,\ldots,n$ can describe as $A_i = \mu_{A_i}(u_1)/u_1 + \mu_{A_i}(u_2)/u_2 + \cdots + \mu_{A_i}(u_n)/u_n$.

For example, the linguistic variable 'enrollment' can be described the fuzzy sets as A_1 = (not many), A_2 = (not too many), A_3 = (many), A_4 = (many many), A_5 = (very many), A_6 = (too many), A_7 = (too many many). Thus, all the fuzzy sets are expressed as follows:

$$A_1 = \{1/u_1, 0.5/u_2, 0/u_3, 0/u_4, 0/u_5, 0/u_6, 0/u_7\},$$
$$A_2 = \{0.5/u_1, 1/u_2, 0.5/u_3, 0/u_4, 0/u_5, 0/u_6, 0/u_7\},$$
$$A_3 = \{0/u_1, 0.5/u_2, 1/u_3, 0.5/u_4, 0/u_5, 0/u_6, 0/u_7\},$$
$$A_4 = \{0/u_1, 0.5/u_2, 0.5/u_3, 1/u_4, 0.5/u_5, 0/u_6, 0/u_7\},$$
$$A_5 = \{0/u_1, 0/u_2, 0/u_3, 0.5/u_4, 1/u_5, 0.5/u_6, 0/u_7\},$$
$$A_6 = \{0/u_1, 0/u_2, 0/u_3, 0/u_4, 0.5/u_5, 1/u_6, 0.5/u_7\},$$
$$A_7 = \{0/u_1, 0/u_2, 0/u_3, 0/u_4, 0/u_5, 0.5/u_6, 1/u_7\}.$$

Step 4. Fuzzify the historical data. Find the fuzzy sets A_i $(i = 1,2,\ldots,n)$ which each historical data belonged, and show the fuzzified enrollment data of the Alabama University is listed as Table 1.

Step 5. Determine fuzzy relation matrix R.
We have the fuzzy logical relationships from Table 1 as follows:

$$A_1 \to A_1, \quad A_1 \to A_2, \quad A_2 \to A_3, \quad A_3 \to A_3,$$
$$A_3 \to A_4, \quad A_4 \to A_4, \quad A_4 \to A_3, \quad A_4 \to A_6, \quad A_6 \to A_6,$$
$$A_6 \to A_7, \quad A_7 \to A_7, \quad A_7 \to A_6.$$