Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 35 (2008) 691–700
Expert Systems with Applications
www.elsevier.com/locate/eswa

Multivariate stochastic fuzzy forecasting models

Tahseen Ahmed Jilani *, Syed Muhammad Aqil Burney

Department of Computer Science, University of Karachi, Karachi 75270, Pakistan

Abstract

In this paper, we have presented two new multivariate fuzzy time series forecasting methods. These methods assume *m*-factors with one main factor of interest. Stochastic fuzzy dependence of order *k* is assumed to define general methods of multivariate fuzzy time series forecasting and control. These new methods are applied for forecasting total number of car road accidents casualties in Belgium using four secondary factors. Practically, in most of the situations, actuaries are interested in analysis of the patterns of casualties in road accidents. Such type of analysis supports in deciding approximate risk classification and forecasting for each area of a city. This directly affects the underwriting process and adjustment of insurance premium, based on risk intensity for each area. National Institute of Statistics, Belgium provides risk intensity based classification of each city. Thus, this work provides support in deciding the appropriate risk associated with an insured in a particular area.

Keywords: Average forecasting error rate (AFER); Fuzziness of fuzzy sets; Fuzzy If-then rules; Multivariate fuzzy time series; Fuzzy aggregation operators

1. Introduction

In our daily life, people often use forecasting techniques to model and predict economy, population growth, stocks, insurance/re-insurance, portfolio analysis and etc. However, in the real world, an event can be affected by many factors. Therefore, if we consider more factors for prediction, with higher complexity then we can get better forecasting results.

During last few decades, various approaches have been developed for time series forecasting. Among them ARMA models and Box–Jenkins model building approaches are highly famous. Most of the financial and economic time series are nonlinear and thus the linear models are inadequate to handle nonlinearity present in the process.

In recent years, many researchers used fuzzy time series to handle prediction problems. In this paper, we present new methods to predict total number of annual car road accidents casualties based on the *m*-factors high-order fuzzy time series. This method provides a general framework for forecasting, that can be increased by increasing the stochastic fuzzy dependence. For simplicity of computation, we have used triangular membership function. The proposed methods construct *m*-factor high-order fuzzy logical relationships based on the historical data to increase the forecasting accuracy rate.

Song and Chissom (1993) presented the concept of fuzzy time series based on the concepts of fuzzy set theory to forecast the historical enrollments of the University of Alabama. Huarng (2001b) presented the definition of two kinds of intervals in the universe of discourse to forecast the TAIFEX. Chen (2002) presented a method for forecasting based on high-order fuzzy time series. Lee, Wang, and Chen (2006) presented a method for temperature prediction based on two-factor high-order fuzzy time series. Melike and Konstsntin (2004) proposed forecasting method using first-order fuzzy time series. Lee et al. (2006) presented the handling of forecasting problems using two-factor high-order fuzzy time series for TAIFEX and daily temperature in Taipei, Taiwan.

* Corresponding author. Tel.: +92 333 3040963.
E-mail addresses: tahseenjilani@ieee.org (T.A. Jilani), Burney@computer.org (S.M.A. Burney).

0957-4174/$ - see front matter
doi:10.1016/j.eswa.2007.07.014

The rest of the paper is organized as follows. In Section 2, brief review of fuzzy time series is given. In Section 3, we present the new methods for multivariate high-order fuzzy time series modeling and in Section 4, new fuzzy forecasting models are applied on annual car road accident casualties data. We also make indirect comparison of our proposed method with existing methods. The conclusions are given in Section 5.

2. Fuzzy times series

Time series analysis plays vital role in most of the actuarial related problems. As most of the actuarial issues are born with uncertainty, therefore, each observation of a fuzzy time series is assumed to be a fuzzy variable along with associated membership function. Based on fuzzy relation and fuzzy inference rules, efficient modeling and forecasting of fuzzy time series is possible. This field of fuzzy time series analysis is not very mature due to the time and space complexities in most of the actuarial related issue.

2.1. Multivariate fuzzy inferencing

Using fuzzy relational calculus (Zimmerman, 2001; Klir & Yuan, 2005), we can extend the concept for many antecedents and single consequent. For example, in designing two-factor kth-order fuzzy time series model with X be the primary and Y be second factor. We assume that there are k antecedent $((X_1,Y_1),(X_2,Y_2),\dots,(X_k,Y_k))$ and one consequent X_{k+1}

$$\text{If } (X_1=x_1,Y_1=y_1),(X_2=x_2,Y_2=y_2),\dots,\\ (X_k=x_k,Y_k=y_k)\to(X_{k+1}=x_{k+1}) \tag{1}$$

In the similar way, we can define m-factor $(i=1,2,\dots,m)$ and kth order $(k=1,2,\dots,k)$ fuzzy time series as

$$\begin{aligned}\textbf{If } &(X_{11}=x_{11},X_{12}=x_{12},\dots,X_{1k}=x_{1k}),\\ &(X_{21}=x_{21},X_{22}=x_{22},\dots,X_{2k}=x_{2k}),\dots,\\ &(X_{m1}=x_{m1},X_{m2}=x_{m2},\dots,X_{mk}=x_{mk})\ \textbf{then}\\ &(X_{m+1,k+1}=x_{m+1k+1})\quad \text{for } i=1,2,\dots,m,\ \ j=1,2,\dots,k\end{aligned} \tag{2}$$

2.2. Fuzzy time series

Let $Y(t)$, $(t=\dots,0,1,2,\dots)$ be the universe of discourse and $Y(t)\subseteq R$. Assume that $f_i(t)$, $i=1,2,\dots$ is defined in the universe of discourse $Y(t)$ and $F(t)$ is a collection of $f(t_i),(i=\dots,0,1,2,\dots)$, then $F(t)$ is called a fuzzy time series of $Y(t)$, $i=1,2,\dots$. Using fuzzy relation, we define $F(t)=F(t-1)\circ R(t,t-1)$ where $R(t,t-1)$ is a fuzzy relation and "$\circ$" is the max–min composition operator, then $F(t)$ is caused by $F(t-1)$ where $F(t)$ and $F(t-1)$ are fuzzy sets.

3. New forecasting method based multivariate high-order fuzzy time series

In Table 1, we have presented the data taken from National Institute of Statistics, Belgium for the period of 1974–2004. The main factor of interest is the yearly road accident casualties and secondary factors are mortally wounded, died within one month, severely wounded and light casualties.

For forecasting purpose, we can define relationship among present and future state of a time series with the help of fuzzy sets. Assume the fuzzified data of the ith and $(i+1)$th day are A_j and A_k, respectively, where $A_j,A_k\in U$, then $A_j\to A_k$ represented the fuzzy logical relationship between A_j and A_k.

Let $F(t)$ be a fuzzy time series. If $F(t)$ is caused by $F(t-1),F(t-2),\dots,F(t-n)$, then the fuzzy logical relationship is represented by

$$F(t-n),\dots,F(t-2),F(t-1)\to F(t) \tag{3}$$

is called the one-factor nth order fuzzy time series forecasting model.

Let $F(t)$ be a fuzzy time series. If $F(t)$ is caused by $F_1(t-1)$, $F_2(t-(1)),(F_1(t-2),F_2(t-2)),\dots,(F_1(t-n),F_2(t-n))$, then this fuzzy logical relationship is represented by

$$(F_1(t-n),F_2(t-n)),\dots,(F_1(t-2),F_2(t-2)),\\ (F_1(t-1),F_2(t-1))\to F(t) \tag{4}$$

is called the two-factors nth-order fuzzy time series forecasting model, where $F_1(t)$ and $F_2(t)$ are called the main factor and the secondary factor fuzzy time series respectively. In the similar way, we can define m-factor nth-order fuzzy logical relationship as

$$(F_1(t-n),F_2(t-n),\dots,F_m(t-n)),\dots,(F_1(t-2),\\ F_2(t-2),\dots,F_m(t-2)),\quad (F_1(t-1),F_2(t-1),\dots,\\ F_m(t-1))\to F(t) \tag{5}$$

Here $F_1(t)$ is called the main factor and $F_2(t)$, $F_3(t),\dots,F_m(t)$ are called secondary factor fuzzy time series. Here we can implement any of the fuzzy membership function to define the fuzzy time series in above equations. Comparative study by using different membership functions is also possible. We have used triangular membership function due to low computational cost.

Using fuzzy composition rules, we establish a fuzzy inference system for fuzzy time series forecasting with higher accuracy. The accuracy of forecast can be improved by considering higher number of factors and higher dependence on history. Now we present an extended method for handling forecasting problems based on m-factors high-order fuzzy time series. The proposed method is now presented as follows.

Step 1. Define the universe of discourse U of the main factor $U=[D_{\min}-D_1,D_{\max}-D_2]$, where $D_{\min}$ and $D_{\max}$ are the minimum and the maximum values of the main factor of the known historical