Expert Systems with Applications 36 (2009) 10589–10594

Contents lists available at ScienceDirect

Expert Systems with Applications

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

A new approach based on artificial neural networks for high order multivariate fuzzy time series

Erol Egrioglu [a], Cagdas Hakan Aladag [b,*], Ufuk Yolcu [a], Vedide R. Uslu [a], Murat A. Basaran [c]

[a] Department of Statistics, Ondokuz Mayis University, Samsun 55139, Turkey
[b] Department of Statistics, Hacettepe University, Ankara 06800, Turkey
[c] Department of Mathematics, Nigde University, Nigde 51000, Turkey

ARTICLE INFO

Keywords:
Artificial neural networks
Forecasting
Fuzzy time series
Multivariate fuzzy time series approaches

ABSTRACT

Fuzzy time series methods have been recently becoming very popular in forecasting. These methods can be categorized into two subclasses that are univariate and multivariate approaches. It is a known fact that real time series data can actually be affected by many factors. In this case, the using multivariate fuzzy time series forecasting model can be more reasonable in order to get more accurate forecasts. To obtain fuzzy forecasts when multivariate fuzzy time series approach is adopted, the most applied method is using tables of fuzzy relations. However, employing this method is a computationally though task. In this study, we introduce a new method that does not require using fuzzy logic relation tables in order to determine fuzzy relationships. Instead, a feed forward artificial neural network is employed to determine fuzzy relationships. The proposed method is applied to the time series data of the total number of annual car road accidents casualties in Belgium from 1974 to 2004 and a comparison is made between our proposed method and the methods proposed by Jilani and Burney [Jilani, T. A., & Burney, S. M. A. (2008). Multivariate stochastic fuzzy forecasting models. *Expert Systems with Applications, 35*, 691–700] and Lee et al. [Lee, L.-W., Wang, L.-H., Chen, S.-M., & Leu, Y.-H. (2006). Handling forecasting problems based on two factors high order fuzzy time series. *IEEE Transactions on Fuzzy Systems, 14*, 468–477].

1. Introduction

The real time series data such as temperature and share prices of stockholders contain some uncertainty in itself. The conventional time series analyses can be inefficient to forecast such time series. Song and Chissom (1993a, 1993b, 1994) first introduced the definition of a fuzzy time series based on the concept of the fuzzy set theory proposed by Zadeh (1965). The fuzzy time series forecasting model defined by Song and Chissom is a model called as a one-factor first order model. Chen (2002) and Aladag et al. (2008) present approaches based on a one-factor high order fuzzy time series forecasting model. It is likely reasonable to think that a multivariate fuzzy time series forecasting model will provide better results in terms of prediction accuracy since the real time series data can be affected by many factors. In the literature, Yu and Huarng (2008), Lee, Wang, Chen, and Leu (2006) use two factor fuzzy time series model to forecast such time series. Then, Jilani and Burney (2008) analyze k-factor ($k \geqslant 2$) and nth order ($n \geqslant 1$) fuzzy time series forecasting models.

In this paper, a new modified method is proposed to analyze k-factor and nth order fuzzy time series forecasting model using feed forward artificial neural networks that determines fuzzy logic relations. The proposed method is applied to the total number of annual car road accidents casualties in Belgium and the results obtained from the proposed method are compared with those obtained from the methods by Jilani and Burney (2008) and Lee et al. (2006).

The sections of this paper are organized as follows. Section 2 gives the concept of the fuzzy time series. In Section 3, feed forward artificial neural networks are briefly reviewed. The proposed method is introduced in Section 4. Section 5 is the implementation of the proposed method using the data of the total number of annual car road accidents casualties in Belgium. The final section gives the comparison results and discussion.

2. Fuzzy time series

The definition of fuzzy time series was first introduced by Song and Chissom (1993a, 1993b). In contrast to conventional time series methods, various theoretical assumptions do not need to be checked in fuzzy time series approaches. The most important advantage of the fuzzy time series approaches is to be able to work with a very small set of data. The definition of fuzzy time series are given as follows:

* Corresponding author.
E-mail address: chaladag@gmail.com (C.H. Aladag).

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.02.057

Let U be the universe of discourse, where $U = \{u_1, u_2, \ldots, u_b\}$. A fuzzy set A_i of U is defined as $A_i = f_{A_i}(u_1)/u_1 + f_{A_i}(u_2)/u_2 + \cdots + f_{A_i}(u_b)/u_b$, where f_{A_i} is the membership function of the fuzzy set A_i; $f_{A_i} : U \to [0,1]$. u_a is a generic element of fuzzy set A_i; $f_{A_i}(u_a)$ is the degree of belongingness of u_a to A_i; $f_{A_i}(u_a) \in [0,1]$ and $1 \leqslant a \leqslant b$.

Definition 1 (*Fuzzy time series*). Let $Y(t)$ ($t = \ldots, 0, 1, 2, \ldots$), a subset of real numbers be the universe of discourse by which fuzzy sets $f_j(t)$ are defined. If $F(t)$ is a collection of $f_1(t), f_2(t), \ldots$ then $F(t)$ is called a fuzzy time series defined on $Y(t)$.

Definition 2. Fuzzy time series relationships assume that $F(t)$ is caused only by $F(t-1)$, then the relationship can be expressed as: $F(t) = F(t-1) * R(t, t-1)$, which is the fuzzy relationship between $F(t)$ and $F(t-1)$, where $*$ represents as an operator. To sum up, let $F(t-1) = A_i$ and $F(t) = A_j$. The fuzzy logical relationship between $F(t)$ and $F(t-1)$ can be denoted as $A_i \to A_j$ where A_i refers to the left-hand side and A_j refers to the right-hand side of the fuzzy logical relationship. Furthermore, these fuzzy logical relationships can be grouped to establish different fuzzy relationships.

Definition 3. Let $F(t)$ be a fuzzy time series. If $F(t)$ is caused by $F(t-1), F(t-2), \ldots$, and $F(t-n)$, then this fuzzy logical relationship is represented by

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

and it is called the nth order fuzzy time series forecasting model.

Definition 4. Let F and G be two fuzzy time series. Suppose that $F(t-1) = A_i$, $G(t-1) = B_k$ and $F(t) = A_j$. A bivariate fuzzy logical relationship is defined as $A_i, B_k \to A_j$, where A_i, B_k are referred to as the left-hand side and A_j as the right-hand side of the bivariate fuzzy logical relationship.

Therefore, first order bivariate fuzzy time series forecasting model is as follows:

$$F(t-1), G(t-1) \to F(t) \tag{2}$$

Definition 5. Let F and G be two fuzzy time series. If $F(t)$ is caused by $(F(t-1), G(t-1)), (F(t-2), G(t-2)), \ldots, (F(t-n), G(t-n))$ then this fuzzy logical relationship is represented by

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

and it is called the two factors nth order fuzzy time series forecasting model, where $F(t)$ and $G(t)$ are called the main factor fuzzy time series and the second factor fuzzy time series, respectively ($t = \ldots, 0, 1, 2, \ldots$).

Definition 6. Let F and $G_1, G_2, \ldots, G_{k-1}$ be k fuzzy time series. If $F(t)$ is caused by $(F(t-1), G_1(t-1)G_2(t-1)\cdots G_{k-1}(t-1)), \ldots, (F(t-n), G_1(t-n)G_2(t-n)\cdots G_{k-1}(t-n))$ then this fuzzy logical relationship is represented by

$$\begin{aligned}&(F(t-1), G_1(t-1)G_2(t-1)\cdots G_{k-1}(t-1)), \ldots, (F(t-n),\\ &\quad G_1(t-n)G_2(t-n)\cdots G_{k-1}(t-n)) \to F(t)\end{aligned} \tag{4}$$

and it is called the k-factors nth order fuzzy time series forecasting model, where $F(t)$ and $G_i(t)$ are called the main factor fuzzy time series and the secondary factors fuzzy time series, respectively ($t = \ldots, 0, 1, 2, \ldots$, $i = 1, 2, \ldots, k-1$).

3. Artificial neural networks

ANN consists of algorithms that mimic the features of brain of human being. These features are generating new knowledge and exploring by learning. In other words, ANN are synthetic networks that imitate biological neural networks. ANN are much more different than biological ones in terms of structure and ability (Zurada, 1992). ANN compose of a mathematical model (Zhang, Patuwo, & Hu, 1998). The fundamental elements of ANN can be given as follows:

3.1. Architecture structure

The structure of multilayer feed forward ANN is basically given in Fig. 1. Multilayer feed forward ANN as illustrated in the figure consist of three parts such as input, hidden and output layers. Each layer consists of neurons. The architecture structure is determined based on deciding the number of neuron in each layer. These neurons are linked each other by weights. There is no link among the neurons in the same layer.

3.2. Learning algorithm

There have been many learning algorithms in order to determine weights. The one of the most employed algorithm is called back propagation learning algorithm. This learning algorithm updates the weights based on difference between real value and output value of the ANN. Learning parameter in back propagation algorithm plays an important role in reaching the desired outputs. When learning parameter is taken as a fixed value, it can be dynamically updated in the algorithm.

3.3. Activation function

Activation function provides the non-linear mapping between input and output. The performance of networks depends on the proper choice of activation function. Activation function can be chosen as either linear or double polarized, or one polarized. Slope parameter should be determined when the activation is non-linear. Also, slope parameter plays a key role in reaching desired output values.

4. The proposed method

To obtain forecasts for fuzzy time series, using multivariate fuzzy time series model, instead of using univariate one, can provide better forecasts since real time series data has a complex structure and is affected by many other factors. In the literature, Yu and Huarng (2008) presented an algorithm that analyzes a first order bivariate fuzzy time series forecasting model. Then, Yu and Huarng

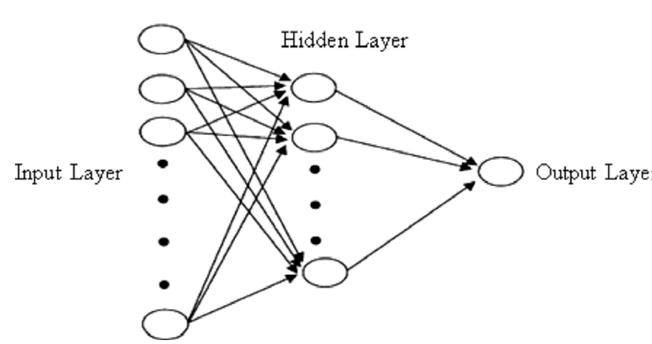

Fig. 1. Multilayer feed forward ANN with one output neuron.

(2008) proposed another algorithm that analyzes a two factor and first order fuzzy time series forecasting model using a feed forward artificial neural network. On the other hand, Jilani and Burney (2008) proposed an algorithm based on an approach that is used to forecast a multivariate high order fuzzy time series. It has been shown that the method proposed by Jilani and Burney (2008) gives better-forecast values than those generated by the method introduced by Lee et al. (2006). However, the determination of fuzzy relationships in the algorithm proposed by Jilani and Burney (2008) depends on fuzzy logic relation tables which require too many complicated calculations and so too much time.

In this paper, we propose a new algorithm that determines fuzzy relationships using a feed forward artificial neural network in order to improve the accuracy of forecast for a multi-factor high order fuzzy time series. The proposed algorithm is given as follows.

Step 1. The universe of discourses and subintervals are defined for two time series. The min and max values of the time series are denoted by $D_{\min}$ and $D_{\max}$, respectively. Then two arbitrary positive numbers D_1 and D_2 for each fuzzy time series can be chosen in order to define the universe of discourse $U = [D_{\min} - D_1, D_{\max} + D_2]$.

Step 2. Define fuzzy sets based on the universe of discourses. Based on defined U and V universe of discourses and subintervals, fuzzy sets $A_1, A_2, \ldots, A_{r_1}$ and $B_{i1}, B_{i2}, \ldots, B_{ir_2}$ are defined and given below for time series and residuals, respectively,

$$\begin{aligned}
A_1 &= a_{11}/u_1 + a_{12}/u_2 + \cdots + a_{1r_1}/u_{r_1} \\
A_2 &= a_{21}/u_1 + a_{22}/u_2 + \cdots + a_{2r_1}/u_{r_1} \\
&\vdots \\
A_{r_1} &= a_{r_11}/u_1 + a_{r_12}/u_2 + \cdots + a_{r_1r_1}/u_{r_1} \\
B_{i1} &= b_{i11}/v_{i,1} + b_{i12}/v_{i,2} + \cdots + b_{i1r_2}/v_{i,r_2} \\
B_{i2} &= b_{i21}/v_{i,1} + b_{i22}/v_{i,2} + \cdots + b_{i2r_2}/v_{i,r_2} \\
&\vdots \\
B_{ir_2} &= b_{ir_21}/v_{i,1} + b_{ir_22}/v_{i,2} + \cdots + b_{ir_2r_2}/v_{i,r_2}
\end{aligned},$$

where a_{ij} is the degree of membership values of u_i and $a_{ij} \in [0,1]$ for $1 \leqslant i, j \leqslant r_1$. r_1 is the numbers of subintervals and fuzzy sets, respectively. In similar way, where b_{ijm} is the degree of membership values of v_{ij} and $b_{ijm} \in [0,1]$ for $1 \leqslant i \leqslant k-1$ and $1 \leqslant j,\ m \leqslant r_2$. r_2 is the numbers of subintervals and fuzzy sets, respectively,

$$A_j = a_{j1}/u_1 + a_{j2}/u_2 + \cdots + a_{jr_1}/u_{r_1},$$

$$a_{jm} = \begin{cases} 1, & m = j \\ 0.5, & m = j-1 \\ 0.5, & m = j+1 \\ 0, & o.w. \end{cases}, \quad j, m = 1, \ldots, r_1$$

and

$$B_{ij} = b_{ij1}/v_{i,1} + b_{ij2}/v_{i,2} + \cdots + b_{ij,r_2}/v_{i,r_2}, \quad i = 1, 2, \ldots, k-1$$

$$b_{ijm} = \begin{cases} 1, & m = j \\ 0.5, & m = j-1 \\ 0.5, & m = j+1 \\ 0, & o.w. \end{cases}, \quad j, m = 1, \ldots, r_2$$

Step 3. Fuzzify the observations. Each crisp value is mapped into a fuzzy set where its membership degree has maximum value. Fuzzy main factor time series is demonstrated by $F(t)$, fuzzy secondary factors time series are demonstrated by $G_i(t)$, $i = 1, 2, \ldots, k-1$.

Step 4. Establish fuzzy relationship. In order to establish fuzzy relationships, ANN is employed. The lagged variables $F(t-1), F(t-2), \ldots, F(t-n)$ of the main factor fuzzy time series are used for training, and the lagged variables $G_1(t-1), G_1(t-2), \ldots, G_1(t-n), \ldots, G_{k-1}(t-1)$, $G_{k-1}(t-2)$, …, $G_{k-1}(t-n)$ of secondary factors are taken as the inputs of network. The main factor fuzzy time series $F(t)$ is used for the output of network. Feed forward neural network is trained in terms of these inputs and output. The number of neurons in the input layer is taken as the product of k times n while the number of neurons in the hidden layer (NNHL) can be decided by trial and error method. It is obvious that the number of neurons in the output layer should be one.

A small example will give an idea of what is being done in Step 4 clearly for three variables second order fuzzy time series. Because of dealing with a second order fuzzy time series, four inputs are employed in neural network model, so that lagged variables $F(t-1)$, $F(t-2)$, $G_1(t-1)$, $G_1(t-2)$, $G_2(t-1)$, $G_2(t-2)$ are obtained from three fuzzy time series $F(t)$, $G_1(t)$ and $G_2(t)$. These series are given in Table 1. The index numbers (j) of A_j of $F(t-1)$ and $F(t-2)$ series are taken as input values whose titles are Input 1 and Input 2, The index numbers (j) of $B_{i,j}$ of $G_i(t-1)$, $G_i(t-2)$, $i = 1, 2$ series are taken as input values whose titles are Input 3, Input 4, Input 5, Input 6 in Table 1 for the neural network model. Also, the index numbers of A_j of $F(t)$ series are taken as target values whose title is Target in Table 1 for the neural network model. When the third observation is taken as an example, inputs values for the learning sample $[A_5, A_6, B_{15}, B_{12}, B_{23}, B_{21}]$ are 5, 6, 5, 2, 3 and 1. Then, target value for this learning sample is 2.

Step 5. Forecast. Prepare data for forecasting: $F(t+l-1), \ldots, F(t+l-n+1), F(t+l-n)$ and $G_1(t+l-1), \ldots, G_1(t+l-n+1)$, $G_1(t+l-n), \ldots, G_{k-1}(t+l-1), \ldots, G_{k-1}(t+l-n+1), G_{k-1}(t+l-n)$ are taken as the inputs for the trained feed forward neural network and the output from the model is the fuzzy forecast for $F(t+l)$. For example, while the input values for the learning sample are $[A_5, A_6, B_{15}, B_{12}, B_{23}, B_{21}]$ if the output of ANN is 4 then the fuzzy forecast will be A_4.

Step 6. Defuzzify each fuzzy forecast $F(t+l)$. Apply "Centroid" method to get the results. This procedure (also called center of area, center of gravity) is the most often adopted method of defuzzyfication. Suppose that the

Table 1
Symbolic representation of a second order three variables fuzzy time series.

No.	$F(t)$	$G_1(t)$	$G_2(t)$	Input 1 $F(t-1)$	Input 2 $F(t-2)$	Input 3 $G_1(t-1)$	Input 4 $G_1(t-2)$	Input 5 $G_2(t-1)$	Input 6 $G_2(t-2)$	Target $F(t)$
1	A_6	B_{12}	B_{21}	–	–	–	–	–	–	–
2	A_5	B_{15}	B_{23}	–	–	–	–	–	–	–
3	A_2	B_{13}	B_{21}	5	6	5	2	3	1	2
4	A_4	B_{13}	B_{22}	2	5	3	5	1	3	4
5	A_8	B_{14}	B_{28}	4	2	3	3	2	1	8
6	A_7	B_{16}	B_{25}	8	4	4	3	8	2	7

fuzzy forecast of $F(t+l)$ is A_i. The defuzzified forecast is equal to the midpoint of the interval (with a biggest membership value) that corresponds to A_i.

5. Application

We have applied this new technique to the time series data of the total number of annual car road accidents casualties in Belgium from 1974 to 2004 that is given in Table 2.

In this application, the main factor fuzzy time series is the number of killed persons, secondary factor time series are mortally wounded, died within 30 days, severely wounded, light casualties, respectively. The application is given step by step below.

Step1. Universe of discourses and subintervals for five time series are defined as
$D_{\min} = 953$, $D_{\max} = 1644$, $D_1 = 103$, $D_2 = 6$, $U = [850, 1650]$ for killed,
$D_{\min} = 90$, $D_{\max} = 819$, $D_1 = 40$, $D_2 = 31$, $V_1 = [50, 850]$ for mortally wounded,
$D_{\min} = 1094$, $D_{\max} = 2393$, $D_1 = 44$, $D_2 = 57$, $V_2 = [1050, 2450]$ for died within 30 days,
$D_{\min} = 5949$, $D_{\max} = 16{,}645$, $D_1 = 149$, $D_2 = 255$, $V_3 = [5800, 16{,}900]$ for severely wounded,
$D_{\min} = 38{,}390$, $D_{\max} = 46{,}818$, $D_1 = 40$, $D_2 = 282$, $V_4 = [38{,}350, 47{,}100]$ for light casualties.
The intervals for five time series are defined as given below:

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \\ u_7 \\ u_8 \end{bmatrix} = \begin{bmatrix} 850 & 950 \\ 950 & 1050 \\ 1050 & 1150 \\ 1150 & 1250 \\ 1250 & 1350 \\ 1350 & 1450 \\ 1450 & 1550 \\ 1550 & 1650 \end{bmatrix}, \quad \begin{bmatrix} v_{1,1} \\ v_{1,2} \\ v_{1,3} \\ v_{1,4} \\ v_{1,5} \\ v_{1,6} \\ v_{1,7} \\ v_{1,8} \end{bmatrix} = \begin{bmatrix} 50 & 150 \\ 150 & 250 \\ 250 & 350 \\ 350 & 450 \\ 450 & 550 \\ 550 & 650 \\ 650 & 750 \\ 750 & 850 \end{bmatrix},$$

$$\begin{bmatrix} v_{2,1} \\ v_{2,2} \\ v_{2,3} \\ v_{2,4} \\ v_{2,5} \\ v_{2,6} \\ v_{2,7} \\ v_{2,8} \end{bmatrix} = \begin{bmatrix} 1050 & 1225 \\ 1225 & 1400 \\ 1400 & 1575 \\ 1575 & 1750 \\ 1750 & 1925 \\ 1925 & 2100 \\ 2100 & 2275 \\ 2275 & 2450 \end{bmatrix}$$

$$\begin{bmatrix} v_{3,1} \\ v_{3,2} \\ v_{3,3} \\ v_{3,4} \\ v_{3,5} \\ v_{3,6} \\ v_{3,7} \\ v_{3,8} \end{bmatrix} = \begin{bmatrix} 5800 & 7100 \\ 7100 & 8500 \\ 8500 & 9900 \\ 9900 & 11,300 \\ 11,300 & 12,700 \\ 12,700 & 14,100 \\ 14,100 & 15,500 \\ 15,500 & 16,900 \end{bmatrix}, \quad \begin{bmatrix} v_{4,1} \\ v_{4,2} \\ v_{4,3} \\ v_{4,4} \\ v_{4,5} \\ v_{4,6} \\ v_{4,7} \\ v_{4,8} \end{bmatrix} = \begin{bmatrix} 38,350 & 39,400 \\ 39,400 & 40,500 \\ 40,500 & 41,600 \\ 41,600 & 42,700 \\ 42,700 & 43,800 \\ 43,800 & 44,900 \\ 44,900 & 46,000 \\ 46,000 & 47,100 \end{bmatrix}$$

Step 2. Based on defined U, V_1, V_2, V_3, V_4 universe of discourses and subintervals, $r_1 = r_2 = 8$ and fuzzy sets A_j, $j = 1,2,\ldots,8$ and B_{ij}, $i = 1,2,3,4$, $j = 1,2,\ldots,8$, are defined and given below for main factor and secondary factor time series, respectively. The membership values are

$$a_{jm} = \begin{cases} 1, & m = j \\ 0.5, & m = j-1 \\ 0.5, & m = j+1 \\ 0, & o.w. \end{cases}, \quad j, m = 1, \ldots, 8,$$

$$b_{ijm} = \begin{cases} 1, & m = j \\ 0.5, & m = j-1 \\ 0.5, & m = j+1 \\ 0, & o.w. \end{cases}, \quad j, m = 1, \ldots, 8,\ i = 1,2,3,4$$

Table 2
Yearly car accident causalities in Belgium from 1974 to 2004.

Year	Killed	Mortally wounded	Died within 30 days	Severely wounded	Light casualties
1974	1574	819	2393	16,506	44,640
1975	1460	701	2161	15,794	42,423
1976	1536	728	2264	16,057	44,227
1977	1597	701	2298	15,830	44,995
1978	1644	728	2372	16,645	44,797
1979	1572	544	2116	15,750	42,346
1980	1616	557	2173	15,915	42,670
1981	1564	454	2018	15,091	41,915
1982	1464	406	1870	14,601	40,936
1983	1479	412	1891	14,864	42,023
1984	1369	363	1732	14,471	42,456
1985	1308	352	1660	13,287	39,879
1986	1456	330	1786	13,764	42,965
1987	1390	380	1770	13,809	44,090
1988	1432	339	1771	14,029	44,956
1989	1488	312	1800	14,515	46,667
1990	1574	190	1764	13,864	46,818
1991	1471	209	1680	12,965	43,578
1992	1380	173	1553	12,113	41,772
1993	1346	171	1517	11,680	41,736
1994	1415	149	1564	11,160	40,294
1995	1228	109	1337	10,267	39140
1996	1122	115	1237	9123	38,390
1997	1150	105	1255	9229	39,594
1998	1224	121	1345	8784	41,038
1999	1173	126	1299	8461	41,841
2000	1253	103	1356	7990	39,719
2001	1288	90	1378	7319	38,747
2002	1145	118	1263	6834	39,522
2003	1035	101	1136	6898	42,445
2004	953	141	1094	5949	41,627

Table 3
Fuzzified observations for the killed time series.

Year	Killed	Fuzzified killed	Year	Killed	Fuzzified killed
1974	1574	A_8	1990	1574	A_8
1975	1460	A_7	1991	1471	A_7
1976	1536	A_7	1992	1380	A_6
1977	1597	A_8	1993	1346	A_5
1978	1644	A_8	1994	1415	A_6
1979	1572	A_8	1995	1228	A_4
1980	1616	A_8	1996	1122	A_3
1981	1564	A_8	1997	1150	A_4
1982	1464	A_7	1998	1224	A_4
1983	1479	A_7	1999	1173	A_4
1984	1369	A_6	2000	1253	A_5
1985	1308	A_5	2001	1288	A_5
1986	1456	A_7	2002	1145	A_3
1987	1390	A_6	2003	1035	A_2
1988	1432	A_6	2004	953	A_2
1989	1488	A_7			

Table 4
MSE and AFER values obtained from changing the number of neurons in the hidden layer and the model order.

Number of hidden layer neurons	1		2		3		4	
	MSE	AFER	MSE	AFER	MSE	AFER	MSE	AFER
Model order								
1	5782.8670	4.9773	3554.6900	3.7957	2013.9290	2.745486	1651.1480	2.4979
2	3922.8670	3.7930	989.1724	2.1774	978.2143	2.1715	1014.1110	2.2449
3	1169.5330	2.2978	989.1724	2.1774	978.2143	2.1715	1014.1110	2.2449
4	2296.2000	2.7527	86,409.8600	19.4696	978.2143	2.1715	1014.1110	2.2449
5	1009.5330	2.1961	989.1724	2.1774	978.2143	2.1715	1014.1110	2.2449

then

$$\begin{matrix} A_1 = a_{11}/u_1 + a_{12}/u_2 + \cdots + a_{1,8}/u_8 \\ A_2 = a_{21}/u_1 + a_{12}/u_2 + \cdots + a_{2,8}/u_8 \\ \vdots \\ A_8 = a_{8,1}/u_1 + a_{8,2}/u_2 + \cdots + a_{8,8}/u_8 \end{matrix},$$

$$\begin{matrix} B_{i,1} = b_{i,1,1}/v_{i,1} + b_{i,1,2}/v_{i,2} + \cdots + b_{i,1,8}/v_{i,8} \\ B_{i,2} = b_{i,2,1}/v_{i,1} + b_{i,2,2}/v_{i,2} + \cdots + b_{i,2,8}/v_{i,8} \\ \vdots \\ B_{i,8} = b_{i,8,1}/v_{i,1} + b_{i,8,2}/v_{i,2} + \cdots + b_{i,8,8}/v_{i,8} \end{matrix}, \quad i = 1,2,3,4$$

Step 3. The observations are fuzzified. The fuzzified observations for the killed time series are demonstrated in Table 3.

Step 4. Feed forward ANN is used to determine fuzzy relationships. The inputs of the ANN are being changed according to the order of the model since it is changed between 1 and 4 in the application. For example, the inputs of ANN for the second order five variables (one main factor and four secondary factor time series) model are given as follows:

$$F(t-1), F(t-2), G_1(t-1), G_1(t-2), G_2(t-1), G_2(t-2), G_3(t-1), G_3(t-2), G_4(t-1), G_4(t-2)$$

The target value of ANN is $F(t)$. The outputs of ANN will be the predicted values of the fuzzy time series denoted by $\widehat{F}(t)$. The number of neurons in the hidden layer for the ANN has been determined by changing the number of neurons between 1 and 5. The algorithm is applied to each possible case that is 20 in total. In every stage of ANN, the logistic activation function has been used. Also, the backward learning algorithm is employed in the training stage of the network.

Step 5. The outputs of ANN will be the predicted values of the fuzzy time series denoted by $\widehat{F}(t)$.

Step 6. The centroid method is used to defuzzify the fuzzy forecasts. For example, if $\widehat{F}(t) = A_4$ is obtained then the defuzzified forecast will be the midpoint of the interval u_4 which has got the biggest membership value as defining this linguistic variable A_4. For each of the 20 different cases, mean squared errors (MSE) and the average forecasted error rate (AFER) are calculated based on the model order and the number of neurons in the hidden layer. The formulas of mean squared error (MSE) and the average forecasted error rate (AFER) are given below. Table 4 shows different MSE and AFER values that are obtained from changing the model order and the number of neurons in the hidden layer:

$$MSE = \sum_{j=1}^{n} \frac{(\text{Forecasted value of day } j - \text{Actual value of day } j)}{n} \quad (5)$$

$$AFER = \sum_{j=1}^{n} \frac{|(\text{Forecasted value of year } j - \text{Actual value of year } j)/\text{Actual value of year } j|}{n} \times 100\% \quad (6)$$

Table 5
The comparison of the methods proposed by Jilani and Burney (1–2) (2008), Lee et al. and the method proposed in this paper for yearly car accident casualties in Belgium from 1974 to 2004.

Years	Actual killed	Jilani and Burney (2008) Method 1	Jilani and Burney (2008) Method 2	Lee et al. (2006)	Proposed method
1974	1574	–	–	–	–
1975	1460	–	–	–	–
1976	1536	–	–	–	–
1977	1597	1497	1497	1500	1600
1978	1644	1497	1497	1500	1600
1979	1572	1497	1497	1500	1600
1980	1616	1598	1497	1500	1600
1981	1564	1598	1497	1500	1600
1982	1464	1498	1497	1500	1500
1983	1479	1498	1497	1500	1500
1984	1369	1398	1497	1500	1400
1985	1308	1298	1396	1400	1300
1986	1456	1498	1296	1300	1500
1987	1390	1398	1497	1500	1400
1988	1432	1398	1396	1400	1400
1989	1488	1498	1396	1400	1500
1990	1574	1598	1497	1500	1600
1991	1471	1498	1497	1500	1500
1992	1380	1398	1497	1500	1400
1993	1346	1298	1396	1400	1300
1994	1415	1398	1296	1300	1400
1995	1228	1198	1396	1400	1200
1996	1122	1098	1095	1100	1100
1997	1150	1198	1196	1200	1200
1998	1224	1198	1196	1200	1200
1999	1173	1198	1196	1200	1200
2000	1253	1298	1296	1300	1300
2001	1288	1298	1296	1300	1300
2002	1145	1098	1095	1100	1100
2003	1035	997	995	1000	1000
2004	953	997	995	1000	1000
AFER		2.6951	5.2444	5.2483	2.1715

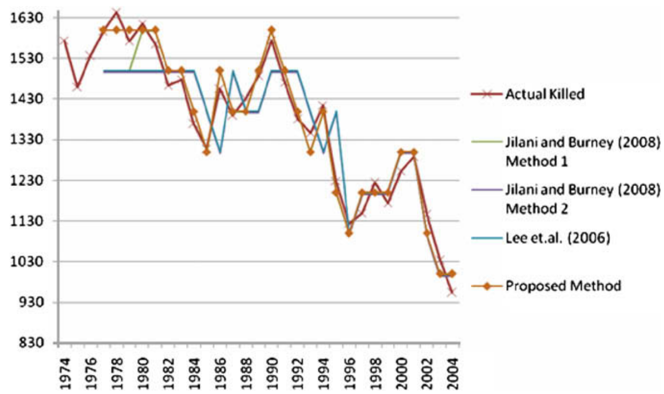

Fig. 2. A comparison of the proposed method with other methods.

As seen from Table 4, the smallest values for MSE and AFER obtained from the proposed method are 978.2143 and 2.1715, respectively. In our proposed method, the best result has been found for the third order model. When the model order is 3, for all cases in which the number of neurons in the hidden layer is 2, 3, 4 and 5, the same MSE and AFER values have been found. Consequently the forecasted values obtained from our new techniques for the case with third order model and two neurons of the hidden layer are listed in Table 5. This table also presents the forecasts from the method of Jilani and Burney (2008) and Lee et al. (2006).It has also seen that the smallest AFER value was obtained from our proposed method which is 2.1715. The real values and the forecasts obtained from all the methods in the literature and the forecasts obtained from our proposed method are shown in Fig. 2. When Fig. 2 is examined, it is easily seen that our proposed method has performed better than existing methods.

6. Conclusion

To analyze multivariate high order fuzzy time series forecasting models, the method proposed by Jilani and Burney (2008) has to deal with too many computations to get fuzzy relation table. In this study, we propose a new method which is based on feed forward artificial neural networks to analyze multivariate high order fuzzy time series forecasting models. Using artificial neural networks for determining fuzzy relationships avoids intense computations and therefore saves time. For the implementation purpose, our proposed method is applied to time series that are yearly car accident causalities in Belgium from 1974 to 2004. The proposed method provides forecasts with a smaller AFER value than ones obtained from the methods presented by Jilani and Burney (2008) and Lee et al. (2006). The advantage of the proposed method can be summarized with two words that are more accurate forecasts and easy implementation. As a future works, other architectures available in ANN should be searched to reach more accurate forecasts.

References

Aladag, C. H., Basaran, M. A., Egrioglu, E., Yolcu, U., & Uslu, V. R. (2008). Forecasting in high order fuzzy time series by using neural networks to define fuzzy relations. *Expert Systems with Applications, 36*, 4228–4231.

Chen, S. M. (2002). Forecasting enrollments based on high-order fuzzy time series. *Cybernetics and Systems an International Journal, 33*, 1–16.

Jilani, T. A., & Burney, S. M. A. (2008). Multivariate stochastic fuzzy forecasting models. *Expert Systems with Applications, 35*, 691–700.

Lee, L.-W., Wang, L.-H., Chen, S.-M., & Leu, Y.-H. (2006). Handling forecasting problems based on two factors high order fuzzy time series. *IEEE Transactions on Fuzzy Systems, 14*, 468–477.

Song, Q., & Chissom, B. S. (1993a). Fuzzy time series and its models. *Fuzzy Sets and Systems, 54*, 269–277.

Song, Q., & Chissom, B. S. (1993b). Forecasting enrollments with fuzzy time series. Part I. *Fuzzy Sets and Systems, 54*, 1–10.

Song, Q., & Chissom, B. S. (1994). Forecasting enrollments with fuzzy time series. Part II. *Fuzzy Sets and Systems, 62*, 1–8.

Yu, T. K., & Huarng, K. (2008). A bivariate fuzzy time series model to forecast the TAIEX. *Expert Systems with Applications, 34*(4), 2945–2952.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*, 338–353.

Zhang, G. P., Patuwo, B. E., & Hu, Y. M. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting, 14*, 35–62.

Zurada, J. M. (1992). *Introduction of artificial neural systems*. St. Paul: West Publishing.