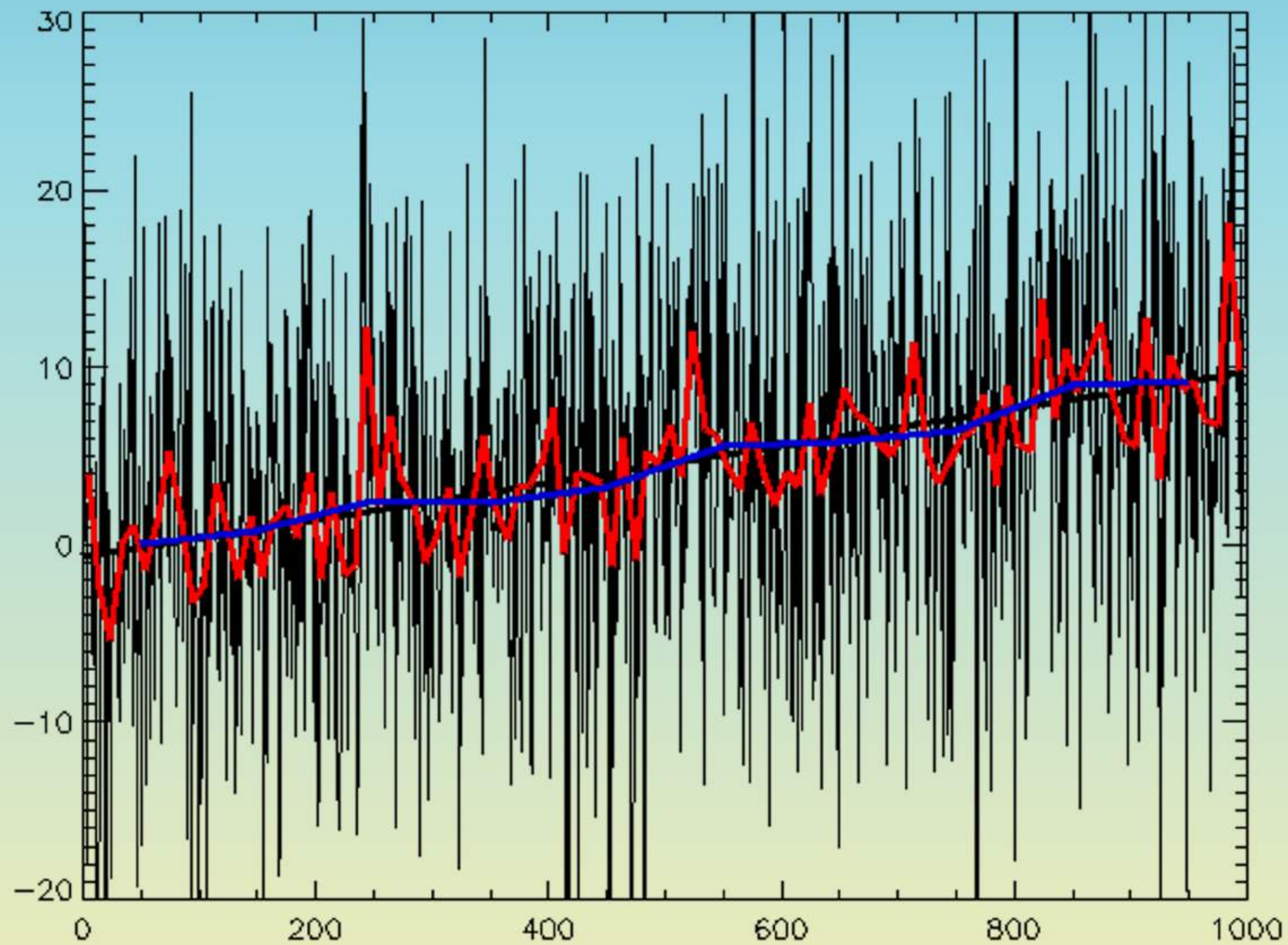
eISBN: 978-1-60805-373-5

ADVANCES IN TIME SERIES FORECASTING



Editor
Cagdas Hakan Aladag
Hacettepe University
Turkey

Co-Editor
Erol Eğrioğlu
Ondokuz Mayis University
Turkey

Bentham e Books

Advances in Time Series Forecasting

Editor

Cagdas Hakan Aladag
Department of Statistics
Hacettepe University
Turkey

Co-Editor

Erol Eğrioğlu
Department of Statistics
Ondokuz Mayis University
Turkey

CONTENTS

CHAPTERS

PART III: HYBRID METHODS

FOREWORD

Generally, advanced intelligent techniques are needed to model and solve the problems encountered in various fields and also to reach desired answers. Many institutions have used various soft computing methods to solve the problems they faced, to increase their productivity and to make strategic decisions. Hence, these methods have received more attention in recent years. In turn, practitioners and academics from various fields have been working on these approaches.

Time series forecasting is one of the most challenging contemporary tasks that are being faced in different areas. In general, different types of time series have been tried for the forecasting purpose. Unfortunately, conventional time series approaches for forecasting can be insufficient in modeling real life time series. Therefore, advanced methods such as artificial neural networks and fuzzy time series have been utilized in many applications. In this eBook, advanced forecasting approaches are described, and further explained how these approaches can be used to forecast real life time series. In particular, some new forecasting approaches are firstly introduced in this eBook. In addition, this eBook provides the background for describing new methods and improving existing advanced forecasting approaches. Dr. Cagdas Hakan Aladag and Assoc. Prof. Dr. Erol Eğrioğlu, the editors of this eBook, have made meaningful contributions to the literature regarding time series forecasting in the recent past. I believe, this eBook will be useful for both practitioners and researchers who are interested in receiving comprehensive views and insights from the variety of issues covered in this eBook.

I. Burhan Turksen
TOBB Economy and Technology University
Turkey

PREFACE

Time series analysis has got attention of many researches from different fields, such as business administration, economics, public finances, engineering, statistics, econometrics, mathematics and actuarial sciences. When many organizations are planning their future, they have to forecast the future. Time series analysis has been employed by many organizations, such as hospitals, universities, companies or government organizations in order to forecast how could be the future. Therefore, many time series forecasting methods have been proposed and improved in the literature. Firstly, linear models such as Box-Jenkins methods were used in many areas of time series forecasting. Furthermore, to overcome the restriction of the linear models and to account for certain nonlinear patterns observed in real problems, some nonlinear models have been proposed in the literature. However, since these nonlinear models were developed for specific nonlinear patterns, they are not capable of modeling other types of nonlinearity in time series. In recent years, to overcome these issues, efficient soft computing techniques such as artificial neural networks, fuzzy time series and some hybrid models have been used to forecast any kind of real life time series. Both theoretical and empirical findings in the literature show that these approaches give better forecasts than those obtained from conventional forecasting methods. In addition, conventional models require some assumptions such as linearity and normal distribution cannot be utilized efficiently for some real time series such as temperature and share prices of stockholders, since this kind of series contain some uncertainty in itself. However, when soft computing methods such as neural networks and fuzzy time series are used to forecast time series, there is no need to satisfy any assumption and the time series uncertainty can be forecasted efficiently.

This eBook contains recent applications and descriptions of these effective soft computing methods. The readers can learn how these methods work and how these approaches can be used to forecast real life time series. In addition, the hybrid forecasting model approach, which is based on combining different soft computing methods to get better forecasts, is explained and at the same time, the reader can find the applications of hybrid forecasting models. The reader of this eBook can also create a new hybrid forecasting model. Although the soft computing forecasting models have many advantages, at the same time there are still some problems with their usage. These problems are pointed out in this ebook. After researchers see those problems, they make some contributions to these forecasting methods by filling some gaps to obtain better forecast results. Furthermore, some new forecasting models are introduced in the eBook.

Cagdas Hakan Aladag

Department of Statistics
Hacettepe University
Turkey

List of Contributors

Aladag, C.H., Ph.D., Lecturer,
Department of Statistics, Hacettepe University, Turkey

Aladag, S., Ph.D., Candidate, Assistant Social Security Expert,
General Director of Service Provision, Republic of Turkey Social Security Institution, Ankara, Turkey

Bagdadioglu, N., Associate Professor,
Department of Public Finance, Hacettepe University, Ankara, Turkey

Basaran, A.A., Ph.D., Research Assistant,
Department of Public Finance, Hacettepe University, Ankara, Turkey

Degirmenci, I., Ph.D., Candidate, Research Assistant,
Department of Statistics, Hacettepe University, Ankara, Turkey

Eğrioğlu, E., Associate Professor,
Department of Statistics, Ondokuz Mayis University, Turkey

Gunay, S., Professor,
Department of Statistics, Hacettepe University, Ankara, Turkey

Kadilar, C., Professor,
Department of Statistics, Hacettepe University, Ankara, Turkey

Koc, S., M.Sc.,
Department of Statistics, Ondokuz Mayis University, Samsun, Turkey

Uslu, V.R., Associate Professor,
Department of Statistics, Ondokuz Mayis University, Samsun, Turkey

Yolcu, U., Ph.D., Assistant Professor,
Department of Statistics, Giresun University, Giresun, Turkey

CHAPTER 1

Advanced Time Series Forecasting Methods

Cagdas Hakan Aladag[1,*] and Erol Eğrioğlu[2]

[1]*Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey and* [2]*Ondokuz Mayis University, Faculty of Arts and Science, Department of Statistics, 55139, Samsun, Turkey*

Abstract: The researchers from various fields have been studying on time series forecasting for approximately one century in order to get better forecasts for the future. To achieve high forecast accuracy level, various time series forecasting approaches have been improved in the literature. During 1980s, some crucial developments happened and time series researches changed. More sophisticated algorithms could be improved since properties of computers were enhanced. Therefore, new time series forecasting approaches such as artificial neural networks and fuzzy time series could be proposed. In the applications, these approaches have proved its success in forecasting real life time series. In addition, hybrid forecasting methods which combine these new approaches have also been improved to obtain more accurate forecasts. In recent years, these advanced time series forecasting methods have been used to forecast real life time series and satisfactory results have also been obtained.

Keywords: Artificial neural networks, Fuzzy time series, Forecasting, Hybrid methods.

1. INTRODUCTION

It is needed to accurately forecast the future in order to make right decisions. One way to forecast the future is using time series whose observations depends on the time. Therefore, time series forecasting is an important issue in various implementation areas such as finance, management, health, tourism, energy, pollution, manufacturing and so on. And, researchers from different disciplines such as statistics, mathematics, economics, business administration and econometrics have been studying time series forecasting for approximately one century.

The major aim in the time series analysis is to forecast the future values, which have not been observed yet, of the time series accurately. Therefore, many forecasting methods have been suggested to increase forecasting accuracy in the literature. The proposed forecasting approaches range from extrapolation to fuzzy time series techniques.

The history of the time series forecasting was briefly given by Pino *et al.* [1] as follows:

"Before the early 1920s, forecasts were calculated by simply extrapolating time series. What might be dubbed as modern forecasting began in 1927, when Yule presented auto-regressive techniques to forecast the annual number of sun spots [2]. His model calculated forecasts as a weighted sum of previous data. If good performance was to be achieved from this linear system, an external factor called noise had to be catered for, as this noise affects the linear system. This linear system with noise was widely used for the next 50 years, when research culminated in the ARIMA methodology proposed by Box and Jenkins [3].

From this point onwards, strongly theory-based studies focused on non-stationary and/or non-linear series: bilinear, bi-spectral or threshold models are examples of this to name but a few [4-8].

During the 1980s, two crucial developments took place that changed time series research. On the one hand, ever increasing capacity and enhanced features of personal computers meant that much longer time series could be handled and more sophisticated algorithms could be used. This went hand in hand with a second aspect the development of machine learning techniques, such as artificial neural networks".

***Address correspondence to Cagdas Hakan Aladag:** Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey; E-mail: chaladag@gmail.com

Cagdas Hakan Aladag and Erol Eğrioğlu (Eds)

Artificial neural networks have been successfully used in many applications [9]. Although artificial neural networks produce very accurate forecasts for most of the time series, the real time series data such as temperature and share prices of stockholders contain some uncertainty in itself and more proper forecasting methods are needed for such time series [10]. Therefore, after artificial neural networks started to be used for time series forecasting, a new approach which is fuzzy time series method was firstly introduced by Song and Chissom [11, 12]. Then, fuzzy time series forecasting models have drawn a great amount of attention in recent years and various models have been proposed in the literature [13].

Both theoretical and empirical findings in the literature have showed that combining different methods can be an effective and efficient way to improve forecasts [14]. Therefore, in the literature, there have been various hybrid forecasting approaches that combine both the conventional and the advanced methods [15].

Advanced time series forecasting approaches such as artificial neural networks, fuzzy time series or hybrid methods have been used widely in recent years since these approaches produces more accurate forecasts than those obtained from other conventional methods and does not require satisfying any assumptions such as linearity, normal distribution and a specific observation number.

In this chapter, the information about the artificial neural networks and the fuzzy time series is presented in Section 2 and 3. The idea underlying the hybrid forecasting approaches is given in Section 4. Finally, the last section concludes the chapter.

2. ARTIFICIAL NEURAL NETWORKS

Artificial neural networks are mathematical models which are inspired from the biological neural networks. Artificial neural networks compose of tree main layers which are the input, the hidden and the output layers. And, these layers include a certain number of neurons. From a mathematical point of view, artificial neural networks method can be considered universal functional approximator [16]. Moreover, artificial neural networks are non linear by nature [17], which means that they can not only correctly estimate non-linear functions, but also extract non linear elements from the data [1].

Artificial neural networks approach is a method which has been successfully used in many areas for different purposes [18]. One of these areas is time series forecasting [19]. Since artificial neural networks can model both the linear and the nonlinear structure of time series, they have attracted more and more attention from both academic researchers and industrial practitioners in recent years [20]. Artificial neural networks have been widely used to model time series in various fields of applications [21] and used as a good alternative method for both linear and non-linear time series forecasting. Zhang *et al.* [9] presented a review of the current status in applications of neural networks for forecasting.

Although artificial neural networks have been successfully used in many implementations, there still some problems with this approach [22]. Artificial neural networks can be considered as an approach composes of three main elements such as architecture structure, learning algorithm and activation function. Determining the elements of the artificial neural networks issue that affect the forecasting performance of artificial neural networks must be considered carefully [23]. Brief information about artificial neural networks was presented by Eğrioğlu *et al.* [23] as follows:

One critical decision is to determine the appropriate architecture, that is, the number of layers, number of nodes in each layers and the number of arcs which interconnect with the nodes [24]. Feed forward artificial neural network has been used in many studies for forecasting. Therefore, our focus is on the feed forward networks. Determining architecture depend on the basic problem. Since, in the literature, there are not general rules for determining the best architecture, many architecture must be examined for the correct results. Fig. **1** depicts the broad feed forward artificial neural networks architecture that has single hidden layer and single output. Other important architectures include direct connections from input nodes to output nodes. Fig. **2** depicts these architectures.

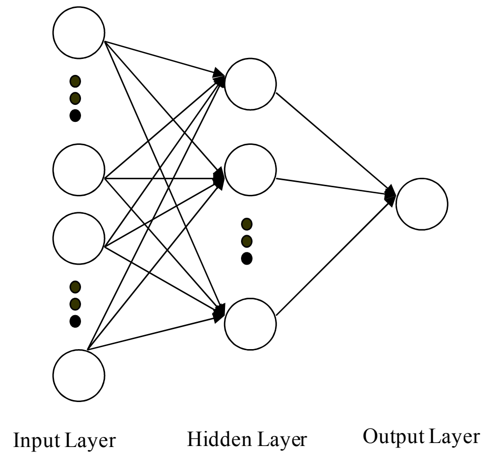


Figure 1: A broad feed forward architecture with one output neuron.

Learning of artificial neural networks for a specific task is equivalent to finding the values all of the weights such that the desired output is generated to the corresponding input. Various training algorithms have been used for the determining optimal weights values. The most popularly used training method is the back propagation algorithm [25]. In the back propagation algorithm, learning of the artificial neural network consists of adjusting all weights such as the error measure between the desired output and actual output [26].

Another element of artificial neural networks is the activation function. It determines the relationship between inputs and outputs of a node and a network. In general, the activation function introduces a degree of the nonlinearity that is valuable for the most artificial neural networks applications. The well known activation functions are logistic, hyperbolic tangent, sine (or cosine) and the linear functions. Among them, logistic transfer function is the most popular one [9].

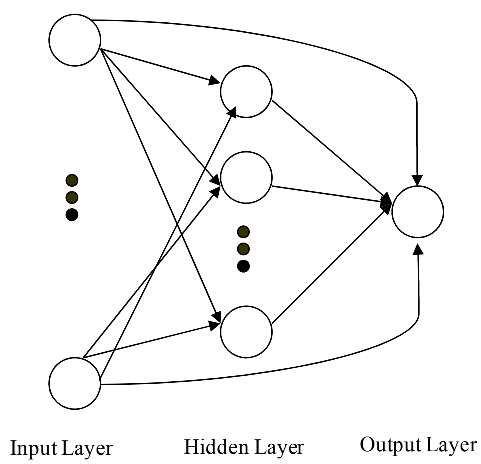


Figure 2: A direct connected feed forward architecture with one output neuron.

3. FUZZY TIME SERIES

Zadeh [27] firstly introduced fuzzy set theory. Based on this paper, fuzzy set theory has found many application areas in science. Fuzzy time series approach based on fuzzy set theory was introduced as an alternative method for conventional time series models. Recently, fuzzy time series has got much attention. Song and Chissom [11, 12] first introduced fuzzy time series. These studies have been inspired by knowledge presented in the papers [28, 29]. Song and Chissom proposed a method based on matrix

PART I: TIME SERIES FORECASTING USING ANN

CHAPTER 2

Comparison of Feed Forward and Elman Neural Networks Forecasting Ability: Case Study for IMKB

Erol Eğrioğlu[1,*], Cagdas Hakan Aladag[2] and Ufuk Yolcu[1]

[1]Ondokuz Mayis University, Faculty of Arts and Science, Department of Statistics, 55139, Samsun, Turkey and [2]Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey

Abstract: In recent years, artificial neural networks (ANN) have been widely used in real life time series forecasting. Artificial neural networks can model both linear and curvilinear structure in time series. Most of the conventional methods used in the analysis of time series are linear structure and fail to analyze non-linear time series. In conventional time series methods such as threshold autoregressive, bilinear model, which are used in non-linear time series modeling, a particular curvilinear model pattern is needed. Artificial neural network is a method based on data and does not require a model pattern. With its activation function, it provides flexible non-linear modeling. Additionally, when compared with conventional methods, successful results are obtained in forecasting time series *via* artificial neural networks in the literature. In this study, feed forward and feedback artificial neural networks which are widely used for time series forecasting were applied to Istanbul Stock Exchange Market (IMKB) time series and forecasting performances were evaluated.

Keywords: Artificial neural networks, Feed forward, Feedback, Forecasting, Time series.

1. INTRODUCTION

Box-Jenkins method has been widely used in forecasting linear time series [1]. In real life, time series generally have curvilinear structure [2]. Since curvilinear time series cannot be modeled *via* Box-Jenkins methods, curvilinear time series models such as Autoregressive Model with Changing Conditional Variance [3], Threshold Models [4] and Bilinear Models [5] are used. As curvilinear time series models are only used for particular curvilinear structures, they do not have enough flexibility [6].

Artificial neural networks have the ability to learn both curvilinear and linear structures in time series. Compared with other time series modeling methods, they provide better forecasting results. Therefore, over the last two decades, there have been studies in the literature concerning the analyses of time series using artificial neural networks [6]. It is crucial to determine the architecture structure, learning and training parameters, activation function and training algorithm which consist of components such as number of input, number of hidden layer, number of units in hidden layer and the number of output. Studies in the literature are conflicting and no common conclusion can be reached in determining these components. No artificial neural network model applicable to each problem can be identified. The simulation studies aiming to determine the components for the forecasting problem are limited and no general conclusions can be reached [7]. Therefore, all components should be reviewed for each question. Feed forward artificial neural networks and Elman feedback artificial neural networks have been widely used in the literature for time series forecasting [8-10].

In this study, forecasting performances of feed forward neural networks and Elman feedback artificial neural networks were discussed by analyzing application results of IMKB time series. Feed forward artificial neural networks and Elman feedback artificial neural networks were described in the second section. Stages of forecasting using artificial neural networks were given as algorithms in the third section. Fourth section included detailed results obtained from the analysis of IMKB time series *via* feed forward and Elman feedback artificial neural networks. Data obtained from the application were evaluated in the discussion section.

***Address correspondence to Erol Eğrioğlu:** Ondokuz Mayis University, Faculty of Arts and Science, Department of Statistics, 55139, Samsun, Turkey; E-mail: erole@omu.edu.tr

Cagdas Hakan Aladag and Erol Eğrioğlu (Eds)

2. ARTIFICIAL NEURAL NETWORKS

Artificial neural network is a data processing mechanism generated by the simulation of human nerve cells and nervous system in a computer environment. The most important feature of artificial neural network is its ability to learn from the examples. Despite having a simpler structure in comparison with the human nervous system, artificial neural networks provide successful results in solving problems such as forecasting, pattern recognition and classification.

Although there are many types of artificial neural networks in the literature, feed forward and Elman feedback artificial neural networks are frequently used for forecasting problem. Feed forward artificial neural networks consist of input layer, hidden layer(s) and output layers. An example of feed forward artificial neural network architecture is shown in Fig. **1**. Each layer consists of units called neuron and there is no connection between neurons which belong to same layer. Neurons from different layers are connected to each other with their weights. Each weight is shown with directional arrows in Fig. **1**. Bindings shown with directional arrows in feed forward artificial neural networks are forward and unidirectional. In the literature, many studies on forecasting use single neuron in output layer. An activation function is used for each neuron in hidden layer and output layer of feed forward artificial neural networks. Inputs incoming to neurons in hidden and output layers are made up multiplication and addition of neuron outputs in the previous layers with the related weights. Data from these neurons pass through the activation function and neuron output are formed. Activation function enables curvilinear match-up. Therefore, non-linear activation functions are used for hidden layer units. In addition to a non-linear activation function, linear (pure linear) activation function can be used in output layer neuron.

In feed forward artificial neural networks, learning is the determination of weights generating the closest outputs to the target values that correspond with the inputs of artificial neural network. Learning is achieved by optimizing the total errors with respect to weights. There are several types of training algorithms in the literature used for learning of feed forward artificial neural networks. One of the widely used training algorithms is Levenberg-Marquardt (LM) [12] algorithm which was also used in this study.

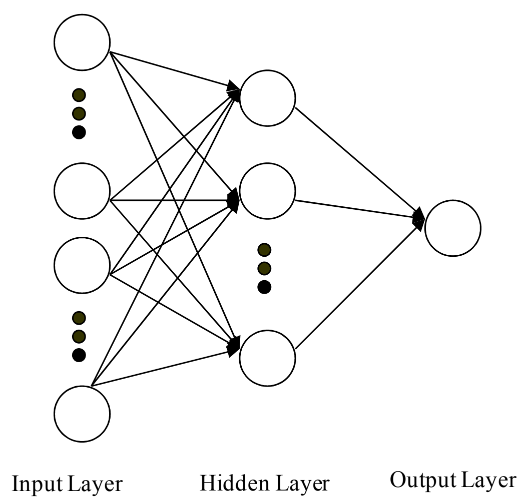


Figure 1: Multilayer feed forward artificial neural network with one output neuron.

Elman artificial neural network is one of the important artificial neural network type used in forecasting time series. Elman neural network, which has the simplest structure among feedback artificial neural networks types, was first proposed by Elman [11]. Elman feedback artificial neural networks consist of input layer, hidden layer, context layer and output layer. Context layer provides a step-delayed feedback mechanism which shows hidden layer output to network as input thus enabling artificial neural network learning with more information. An example of Elman artificial neural network architecture is shown in Fig. **2**.

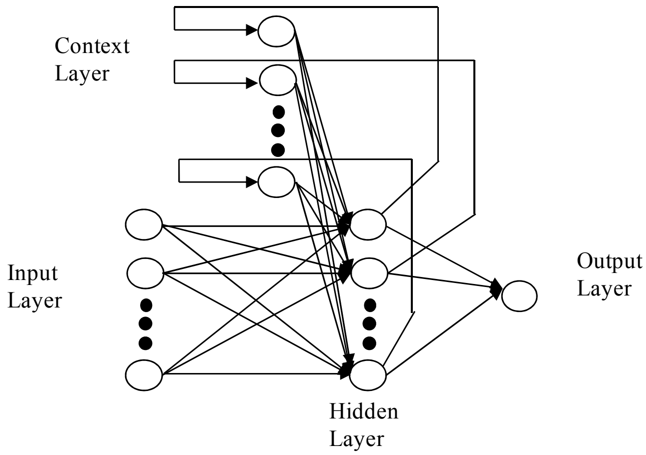


Figure 2: Elman recurrent artificial neural network.

3. FORECASTING TIME SERIES WITH ARTIFICIAL NEURAL NETWORKS

Recently, artificial neural networks have been used in most time series forecasting applications. The following factors lead the widespread usage of artificial neural networks method in time series:

- Analysis can be done without testing whether time series have curvilinear or linear structure.
- When compared with conventional time series method, artificial neural network method provides much better forecasting results.
- As conventional time series models can only be used for particular curvilinear structures, they are not flexible in general. But in the analysis using artificial neural network, no matter what the curvilinear structure of time series is.
- Artificial neural network theory is not complex as in the conventional time series forecasting methods and it is easy to understand.

The use of both feed forward artificial neural networks and feedback neural networks in forecasting time series can be summarized in seven steps.

Step 1. Selection of activation function and pre-processing of the data.

Initially, the type of activation function to be used in hidden and output layer of artificial neural network is selected. Logistic activation function was used in hidden layer units in the application of this study. Logistic activation function is given as follows:

$$f(x) = \frac{1}{1+\exp(-x)} \qquad \textbf{(1)}$$

Data are converted into the range suitable for activation function. If the logistic activation function is to be used, *xi,* to show input value can be converted in to (0, 1) range like,

$$x_i' = \frac{x_i - Min(x_i)}{Maks(x_i) - Min(x_i)} \qquad \textbf{(2)}$$

where, $Max\ (x_i)$ represents maximum input value; $Min\ (x_i)$ represents minimum input value.

CHAPTER 3

Comparison of Architecture Selection Criteria in Analyzing Long Memory Time Series

Erol Eğrioğlu[1], Cagdas Hakan Aladag[2,*] and Ufuk Yolcu[1]

[1]Ondokuz Mayis University, Faculty of Arts and Science, Department of Statistics, 55139, Samsun, Turkey and [2]Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey

Abstract: In recent years, studies including long memory time series are existed in the literature. Such time series in real life may have both linear and nonlinear structures. Linear models are inadequate for this kind of time series. An alternative method to forecast these time series is artificial neural networks which is data based and can model both linear and nonlinear structure in these time series. In order to determine the number of nodes in the layers of a network is an important decision. This decision has been made by using various architecture selection criteria. The performance of these criteria varies, depending on components of time series, such as trend and seasonality. In this study, some architecture selection criteria are compared on real time series when artificial neural networks are employed in forecasting. Some advices are given for using artificial neural networks to forecast long memory time series.

Keywords: Architecture selection criteria, Artificial neural networks, Long range dependent, Time series.

1. INTRODUCTION

In real life, many time series can have long memory structure. These time series can be called long range dependent or long memory time series. In the literature, Autoregressive Fractionally Integrated Moving Average (ARFIMA) models have been used to analyze these time series. Despite most of the long range dependent time series faced in real life can have a non linear structure, ARFIMA models are linear. Therefore, using ARFIMA models for these time series can lead to misleading results. On the other hand, using artificial neural networks would be a wiser choice since they can model both linear and non linear structures in time series.

When time series are forecasted by using artificial neural networks, some criteria such as root mean square error (RMSE), mean absolute percent error (MAPE), Akaike information criterion (AIC), Bayesian information criterion (BIC), direction accuracy criterion (DAC) are employed for architecture selection. Qi and Zhang [1] studied on using these selection criteria for time series which does not include long memory structure. In their study, they found that the mentioned selection criteria produce inconsistent results for in and out sample observations. In other words, there is no guarantee that the architecture found as the best architecture for in sample observations is the best architecture for out sample observations. Qi and Zhang [1] also pointed out that AIC and BIC criteria produce very poor results when artificial neural networks are used for forecasting.

In this study, we aim to examine whether the results reached by Qi and Zhang [1] are valid for long range dependent time series or not. In addition, it is aimed to bring out which criterion give best results for these time series. For doing these, yearly minimal water levels of the Nile River [2], VBR data [2], and Nasqad-100 index volatile [3] long range dependent time series are used. In addition to these series, the international tourism demand of Turkey time series, which has long memory structure, is used in the implementation. Besides, some time series component such as seasonality and trend can affect the performance of the selection criteria. Therefore, it should be examined that whether long range dependency has an effect on the performance of the criteria or not.

The next section introduces the long range dependent time series. In Section 3, brief information about artificial neural networks is given. Section 4 presents the architecture selection criteria such as RMSE,

***Address correspondence to Cagdas Hakan Aladag:** Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey; E-mail: chaladag@gmail.com

Cagdas Hakan Aladag and Erol Eğrioğlu (Eds)

MAPE, AIC, BIC, and DAC. While the long range dependent time series are being forecasted with artificial neural networks, performances of mentioned criteria are examined in Section 5. The obtained results are discussed in the last section.

2. THE LONG RANGE DEPENDENT TIME SERIES

Long range dependent time series can be faced in many fields such as economy, hydrology, and geophysics. Existence of long range dependent was firstly introduced by Hurst [4]. In his study, he tried to determine the minimum level of the Nile River by utilizing the concept of long range dependent [4]. Therefore, the long range dependent can be called as Hurst effect in the literature. The long range dependent time series has some characteristics given below [2]:

i) There are relatively long periods where the observations tend to stay at a high level, and on the other hand, there are long periods with low levels.

ii) When short time periods are examined, local trends and periodical behaviors are observed. However, local trends or periodical behaviors are not clearly seen when the time series data are examined.

iii) Overall, the series looks stationary.

In long range dependent process, some numerical properties can be given as follows [2]:

i) The variance of the sample mean seems to decay to zero at a slower rate than n^{-1}. In good approximation, the rate is proportional to $n^{-\alpha}$, $0 < \alpha < 1$.

ii) The sample correlations decay to zero at a rate that is in good approximation proportional to $k^{-\alpha}$,$0 < \alpha < 1$.

iii) $I(\lambda)$ Near the origin, the logarithm of the periodogram plotted against the logarithm of the frequency appears to be randomly scattered around a straight line with negative slope.

Due to these given characteristics, the definition of long range dependency can be given as follows:

Let $\alpha \in (0,1)$ is a real number and $c_\rho > 0$ is a constant. If a stationary process X_t satisfies the condition which can be given as follows:

$$\lim_{k\to\infty} \rho(k) / \left[c_\rho k^{-\alpha} \right] = 1, \qquad \textbf{(1)}$$

process X_t can be called as long range dependent process. The detailed information about long range dependent time series can be found in [5].

3. ARTIFICIAL NEURAL NETWORKS

Artificial neural networks approach has been used as a good alternative method in time series analysis. Zhang [6] summarized some time series applications in which artificial neural networks are used as analyzing method. Zhang [6] revealed basic concepts of using artificial neural networks for time series analysis. Artificial neural networks approach is a data based approach. For the examined data, therefore, using right components of the method can lead better results. Selection of the components such as architecture structure, learning algorithm and activation function has an important effect on the performance of artificial neural networks [7]. The basic elements of ANN can be given as follows [8]:

Architecture structure: Feed forward ANN has been widely used for forecasting problems because of their simple usage and success. 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. For a forecasting problem, the inputs of the network are past lagged observations.

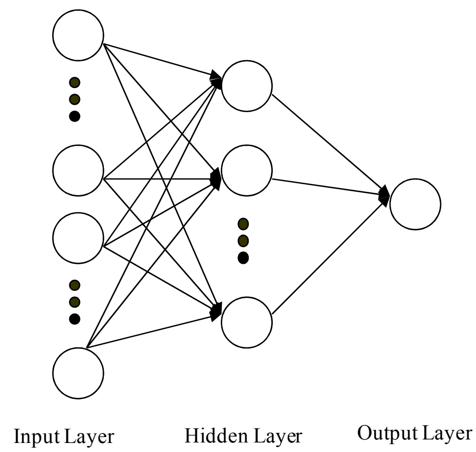


Figure 1: Multilayer feed forward ANN with one output neuron.

One critical decision is to determine the appropriate architecture, that is, the number of layers, number of nodes in each layer [9]. However, in the literature, there are not general rules for determining the best architecture [10].

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. However, back propagation networks have some disadvantages mentioned in the introduction. In light of the weakness of the conventional back propagation algorithm, a number of variations or modifications of this algorithm, such as the adaptive method, quickprop, and second-order methods *etc.*, have been proposed [11]. Among them, the second-order methods such as Levenberg Marquardt method are more efficient nonlinear optimization methods and are used in most optimization packages. Their faster convergence, robustness, and the ability to find good local minima make them attractive in ANN training [11]. Therefore, Levenberg Marquardt method [12] is used as training algorithm in the implementation.

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.

In the literature, simulation studies conducted to determine the components given above are not enough since these studies are very limited and do not include general results [13]. Lots of component combinations have to be tried to determine the best components which gives the best results. Despite this disadvantage of artificial neural networks, the method has some advantages given below.

- Time series can be analyzed by artificial neural networks without testing time series for nonlinearity.
- Artificial neural networks can produce better results those obtained from conventional methods.
- Artificial neural networks can model time series that include any type of nonlinearity. On the other hand, other non linear time series methods can only model time series that have specific non linear structures since these methods were proposed for only specific structures.

CHAPTER 4

Forecasting the Number of Outpatient Visits with Different Activation Functions

Cagdas Hakan Aladag[1,*] and Sibel Aladag[2]

[1]*Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey and* [2]*General Director of Service Provision, Republic of Turkey Social Security Institution, Ankara, Turkey*

Abstract: Forecasting the number of outpatient visits plays important role in strategic decisions for the expert of healthcare administration. In order to manage hospitals effectively, it is needed to forecast the number of outpatient visits accurately. In the literature, there have been some methods proposed to forecast these time series. One of these methods is artificial neural networks approach. Although, artificial neural networks have proved its success in forecasting, there are still some problems with using this method. Determining the elements of this method is an important issue. Activation function is a crucial element of artificial neural networks. Therefore, in this study, we examined different activation functions to obtain more accurate out sample predictions while the number of patients is being forecasted. It is found that using different activation function affects the forecasting accuracy of feed forward neural network models.

Keywords: Activation function, Artificial neural networks, Forecasting, The number of outpatient visits, Time series.

1. INTRODUCTION

Forecasting activities play important role in our daily life. To make a strategic decision for the expert of healthcare administration, forecasting the number of outpatient visits plays important role. If more accurate forecasts are obtained, a schedule for human resources and finances can be made up more reasonably and hospital material resources can be distributed more suitably. In the literature, some approaches have been employed to obtain accurate forecasts for number of patients. Loytonen [1] used Box-Jenkins method to forecast HIV seropositive population in Finland. A stepwise autoregressive method and exponential smoothing models were used to forecast the number of patients with end-stage renal disease in the United States by Xue *et al.* [2]. A stepwise linear regression analysis was performed to predict patient visits to an urgent care clinic by Batal *et al.* [3]. Guan *et al.* [4] used artificial neural network (ANN) in forecasting the incidence of hepatitis A. Seasonal autoregressive integrated moving average, time series regression, exponential smoothing, and ANN were performed to forecast the daily patient volumes in the emergency department by Jones *et al.* [5]. Cheng *et al.* [6] used fuzzy time series to forecast number of outpatient visits.

Many time series in real life may have both linear and nonlinear structures. The traditional time series method can be inefficient to forecast such time series. ANN can model both linear and nonlinear structure in time series. For the aim of obtaining accurate forecasts in real life time series, ANNs have been used successfully in many applications in the literature. Feed forward neural network models have been especially preferred for forecasting time series in many implementations because of its easy usage and providing good results.

ANNs consist of a lot of elements such as architecture structure, training algorithm, activation function and there are various types of these elements [7]. When the number of patients is forecasted by employing ANNs, small amount of selections for elements of network models have been generally examined in the literature. Although the activation function has an important effect on the forecasting accuracy, in

***Address correspondence to Cagdas Hakan Aladag:** Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey; E-mail: chaladag@gmail.com

Cagdas Hakan Aladag and Erol Eğrioğlu (Eds)

determining the best network model, different activation functions have not been tried in the analysis. Instead of employing different activation function, only one activation function has been usually utilized in the implementations. For example, Guan *et al.* [4] and Jones *et al.* [5] use only one type of activation function in their studies. Similarly, small amount of architecture is employed in the analysis. For example, Guan *et al.* [4] used 24 different architecture and Jones *et al.* [5] used only one structure including two neurons in hidden layer.

In this study, we use feed forward neural networks to forecast number of outpatient visits in health center at Hacettepe University. The monthly observations are between September 2004 and October 2008 so the time series include 50 observations. We examined different activation functions to obtain more accurate out sample predictions. Tangent sigmoid, logistic sigmoid, linear, satlin and triangular basis functions were used as activation functions in the output neuron while tangent sigmoid function is used in hidden layer neurons. We also examined 196 different architectures for each type of activation function. Also, the time series is analyzed by Box-Jenkins method, which is the well known method, for the comparison.

In the next section, brief information about ANNs is presented and some well known activation functions are given. Section 3 includes the implementation. As a result of the implementation, the obtained findings are summarized in the last section of the chapter.

2. ARTIFICIAL NEURAL NETWORKS

ANNs were originally motivated by the biological structures in the brains of humans and the animals, which are extremely powerful for such tasks as information processing, learning and adaptation. In forecasting, ANNs are mathematical models that imitate biological neural networks [8]. The function of an ANN is to produce an output pattern when presented with an input pattern. The elements of ANN are network architecture, learning algorithm and activation function.

The definition of ANN approach presented by Aladag *et al.* [9] can be given as follows:

'What is an artificial neural network?' is the first question that should be answered. Picton [10] answered this question by separating this question into two parts. The first part is why it is called an artificial neural network. It is called an artificial neural network because it is a network of interconnected elements. These elements were inspired from studies of biological nervous systems. In other words, ANN are an attempt at creating machines that work in a similar way to the human brain by building these machines using components that behave like biological neurons. The second part is what an artificial neural network does. The function of an artificial neural network is to produce an output pattern when presented with an input pattern. In forecasting, ANN are mathematical models that imitate biological neural networks. ANN consist of some elements. Determining the elements of the ANN issue that affect the forecasting performance of ANN should be considered carefully.

Network architecture: 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 [11].

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. However, back propagation networks have some disadvantages mentioned in the introduction. In light of the weakness of the conventional back propagation algorithm, a number of variations or modifications of this algorithm, such as the adaptive method, quickprop, and second-order methods *etc.*, have been proposed [12]. Among them, the second-order methods such as Levenberg Marquardt method are more efficient nonlinear optimization methods and are used in most optimization packages [13]. Their faster convergence, robustness, and the ability to find good local minima make them attractive in ANN training [12]. Therefore, Levenberg Marquardt method is used as training algorithm in the implementation.

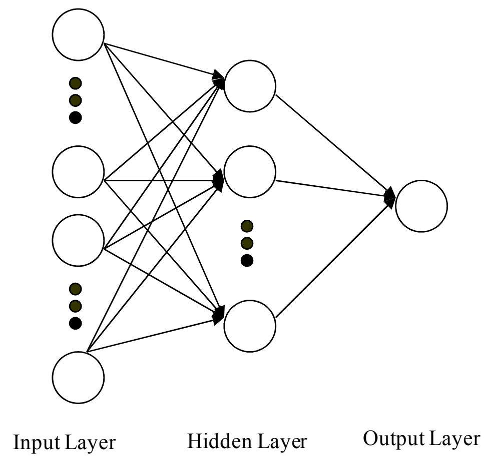
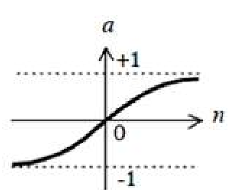


Figure 1: A broad feed forward ANN architecture.

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 [11]. Some well known activation functions are tangent sigmoid, logistic sigmoid, linear, satlin, and triangular basis functions. In Matlab computer package, these functions are represented by tansig, logsig, purelin, satlin, and tribas, respectively. These activation functions, related formulas, and their graphs can be given as follows:

Tangent sigmoid:

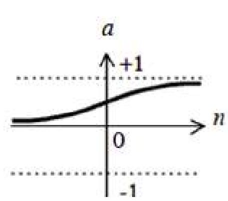
$$\boldsymbol{a = tansig(n) = \frac{2}{1+\exp(-2n)} - 1}$$



a = tansig(n)

Logistic sigmoid:

$$\boldsymbol{a = logsig(n) = \frac{1}{1+\exp(-n)}}$$



a = logsig(n)

CHAPTER 5

Adaptive Weighted Information Criterion to Determine the Best Architecture

Cagdas Hakan Aladag[1,*] and Erol Eğrioğlu[2]

[1]Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey and [2]Ondokuz Mayis University, Faculty of Arts and Science, Department of Statistics, 55139, Samsun, Turkey

Abstract: In the literature, different selection criteria are used for determining the best architecture when time series is analyzed by artificial neural networks. Criteria available in the literature measure different properties of forecasts. To obtain better forecasts, Eğrioğlu *et al.* [1] proposed a criterion which can measure all properties of forecasts. Aladag *et al.* [2] improved the criterion proposed by [1] by using optimization. In this study, both the weighted information criterion proposed by Eğrioğlu *et al.* [1] and the adaptive weighted information criterion proposed by Aladag *et al.* [2] are introduced. These criteria are used in the architecture selection to analyze time series which are the import values of Turkey and the air pollution records in Ankara. As a result of computations, obtained results are compared and discussed. As a result of the comparison, it is seen that adaptive weighted information criterion produce more consistent results.

Keywords: Artificial neural networks, Forecasting, Model selection criterion, Time series.

1. INTRODUCTION

Artificial neural networks (ANN) offer effective and alternative solutions to the problems which are the subjects of statistics. Recently, there have been numerous studies using ANN approaches for time series forecasting problem. It is well known that ANN method is effective particularly in modeling non-linear time series. Not requiring any assumption in forecasting time series makes ANN applicable for many fields. Although ANN has significant advantages, there are still problems to be solved. These problems can be expressed as follows:

- Which architecture structure should be used?
- Which activation function should be used?
- How many hidden layers should be used?
- How many units should hidden and input layers should have?
- Which algorithm should be preferred for ANN training?
- Which model selection criterion should be used for the evaluation of the forecasts?
- Which ANN type (Elman, Jordan, multi-layer) should be used?

There have been partly answers to these questions in the literature. In fact, these questions are valid not only for forecasting but also other targets such as classification. In the literature, [3] partly cover answers to these questions. It is possible to reach following conclusions from the studies available in the literature using artificial neural networks for time series forecasting [4]:

***Address correspondence to Cagdas Hakan Aladag:** Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey; E-mail: chaladag@gmail.com

- Multi-layer ANN gives successful results for forecasting problem.
- It is possible to create accurate forecasting with the use of single hidden layer.
- For satisfactory results, logistic or hyperbolic tangent activation function can be used as an activation function in all layers.
- Linear activation function can also be preferred in output layer.

Despite all these experimental results, the problem which lacks of a common solution and a systematic approach is the selection of the number of units to be used in input and hidden layers. As hidden layer units show delayed-variables in time series forecasting problem, determination of the number of units in input layer is a factor affecting forecasting performance directly. If the number of units in hidden layer is assessed accurately, modeling ability of the network will be optimized. Therefore, determination of hidden layer and input layer unit numbers is crucial. On the other hand, in the literature it is still problematic to assess the best unit number. This is called architecture style selection problem. In the literature, systematic methods not ensuring optimal solution have been suggested for architecture selection problem. Lahnajarvi *et al.* [5], Reed [6] and Siestema and Dow [7] proposed various pruning algorithms. Buhamra *et al.* [8] suggested an input layer unit number selection method based on Box-Jenkins approach. Aladag [9] and Dam and Saraf [10] proposed architecture selection methods based on tabu search algorithm and genetic algorithms, respectively. Aladag *et al.* [11] developed a new approach for architecture selection in analyzing seasonal time series with ANN methods. Zeng *et al.* [12] proposed an approach based on principal component analysis. Zhang [3] developed iterative construction algorithm. The approaches proposed in the studies [13-18] are the examples of other systematic approaches in the literature.

Eğrioğlu *et al.* [1] proposed weighted information criterion assessing forecasting in terms of growth of the error, the number of parameters and direction accuracy in the model. Roughly, weighted information criterion (WIC) is obtained by summing some weighted different selection criteria which measure the forecasting accuracy of an ANN model in different ways. Aladag *et al.* [2] improved this criterion by determining the weights used for WIC with the use of optimization. In this study, both weighted information criterion proposed by Eğrioğlu *et al.* [1] and adaptive weighted information criterion proposed by Aladag *et al.* [2] were introduced. In this application, the best ANN architecture was determined using both WIC and adaptive weighted information criterion (AWIC) in time series analysis of import values of Turkey and the air pollution records in Ankara. Finally results obtained from the application were compared and discussed.

2. WEIGHTED INFORMATION CRITERION

WIC proposed by Eğrioğlu *et al.* [1], is the weighted sum up of most preferred criteria including Akaike information criterion (AIC), Bayesian information criterion (BIC), root mean squared error (RMSE), mean absolute percentage error (MAPE), directional accuracy (DA) and modified direction accuracy. Some of the criteria are calculated by the use of following formulas.

$$AIC = \log\left(\frac{\sum_{i=1}^{T}(y_i-\hat{y}_i)^2}{T}\right)+\frac{2m}{T} \tag{1}$$

$$BIC = \log\left(\frac{\sum_{i=1}^{T}(y_i-\hat{y}_i)^2}{T}\right)+\frac{m\log(T)}{T} \tag{2}$$

$$RMSE = \left(\frac{\sum_{i=1}^{T} (y_i - \hat{y}_i)^2}{T} \right)^{1/2} \tag{3}$$

$$MAPE = \frac{1}{T} \sum_{i=1}^{T} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{4}$$

$$DA = \frac{1}{T} \sum_{i=1}^{T} a_i \quad , \quad a_i = \begin{cases} 1 \ , & (y_{i+1} - y_i)(\hat{y}_{i+1} - y_i) > 0 \\ 0 \ , & otherwise. \end{cases} \tag{5}$$

In the formulas, y_i represents real values; $\hat{y}_i$ forecasts obtained from ANN architecture, *T,* observation number in test set and *m,* weight number in ANN architecture. Another criterion, called modified direction accuracy (MDA) assessing that how well the milestones were forecasted, was proposed by Eğrioğlu *et al.* [1]. ADS criterion can be calculated as;

$$A_i = 1 \quad , \quad y_{i+1} - y_i \le 0$$
$$A_i = 0 \quad , \quad y_{i+1} - y_i > 0$$

$$F_i = 1 \quad , \quad \hat{y}_{i+1} - \hat{y}_i \le 0$$
$$F_i = 0 \quad , \quad \hat{y}_{i+1} - \hat{y}_i > 0$$

$$D_i = (A_i - F_i)^2$$

$$MDA = \frac{\sum_{i=1}^{T-1} D_i}{T-1} \tag{6}$$

Accordingly, architecture selection strategy based on WIC is presented in 5 steps in the following algorithm introduced in [1].

Step 1. Possible architectures are defined. For example, when both input layer unit number and hidden layer unit number vary 1 and 12, it is possible to create a total of 144 architectures.

Step 2. The best values of ANN are determined using training data and AIC, BIC, RMSE, MAPE, DA and MDA criteria were calculated for the test data.

Step 3. AIC, BIC, RMSE, MAPE, DA and MDA criteria calculated for all possible architectures are standardized. For example, AIC criterion for 144 architectures is standardized as follows;

$$AIC_i = \frac{AIC_i - \min(AIC)}{\max(AIC) - \min(AIC)}, \quad i = 1, \ldots, 144$$

Step 4. Weighted information criterion (WIC) is calculated with the following formula for each architecture:

$$WIC = 0{,}1(AIC + BIC) + 0{,}2(RMSE + MAPE) + 0{,}2\big((1 - DA) + MDA\big) \tag{7}$$

Step 5. Architecture with minimum WIC value is selected.

CHAPTER 6

Public Expenditure Forecast by Using Feed Forward Neural Networks

Alparslan A. Basaran[1], Cagdas Hakan Aladag[2,*], Necmiddin Bagdadioglu[1] and Suleyman Gunay[2]

[1]Hacettepe University, Faculty of Economics and Administrative Sciences, Department of Public Finance, 06800, Ankara, Turkey and [2]Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey

Abstract: The accurate forecast of public expenditure is crucial for the success of the new public financial management approach developed in Turkey since the financial crisis of 2001. The public institutions are now obliged to align their expenditure with the framework shaped by the Public Financial Management and Control Law (No: 5018), the Middle-Term Programme of 2010-2012, and recently the Fiscal Rule envisaged to apply in the next budgetary period. This necessitates a better forecasting method than the traditional way of budget forecasting, which is typically based on the expenditures of previous years adjusted by inflation. Particularly focusing on the expenditure side of the budget, this chapter applies various artificial neural networks models to the expenditures of 1973-2008 of two Turkish public institutions, namely, the State Planning Organization and the Court of Accounts to achieve accurate forecast levels. The artificial neural networks approach is rarely applied for the forecasting of public expenditures, and as far as we know this is the first of such attempts involving Turkish data. The artificial neural networks application provided very accurate public expenditure forecasts for these public institutions, suggesting that the artificial neural networks is a very useful method for the public expenditure forecasting, as well.

Keywords: Artificial neural networks, Budget forecasting, Public expenditure, Time series.

1. INTRODUCTION

In recent years, the artificial neural networks approach has been applied to many areas; one of them is the time series forecasting [1]. Since the artificial neural networks can model both nonlinear and linear structure of time series, using the artificial neural networks in forecasting can give more accurate results than the other methods [2]. The accurate forecast of public expenditure has crucial importance for Turkey considering her chronic budget imbalances derived from excess expenditures over collected revenues. A recent study by Bagdigen [3] suggests that the severe errors in budget forecasting in Turkey are caused by the under-forecasting of expenditures and the over-forecasting of revenues.

Turkey has shown her willingness for not tolerating such budget imbalances anymore by initiating an economic recovery programme, including a new public financial management approach since the financial crisis of 2001 [4]. As part of the economic recovery programme, presently, the public institutions are obliged to align their expenditure with the framework shaped by the Public Financial Management and Control Law (No: 5018), the Middle-Term Programme of 2010-2012, and recently the Fiscal Rule envisaged to apply in the next budgetary period. The success of this initiative necessitates a better forecasting method than the traditional method of budget forecasting in Turkey, which is typically based on the adjustment of expenditures and revenues of previous years by inflation.

The aim of this chapter is to show how useful the artificial neural networks can be in getting more accurate public budget expenditure forecasts in Turkey. The artificial neural networks approach is rarely applied for the forecast of public expenditures [5-7], and as far as we know this is the first of such attempts involving Turkish data. The chapter is organized under three main sections. The following section

***Address correspondence to Cagdas Hakan Aladag:** Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey; E-mail: chaladag@gmail.com

introduces the artificial neural networks. The third section explains the basics of the new budgetary approach in Turkey, and the involvement of the State Planning Organization (SPO) and the Court of Accounts (CoA) in the budgetary process. The fourth section applies the artificial neural networks on the expenditures of the SPO and the CoA. The chapter ends with an assessment of the findings.

2. THE ARTIFICIAL NEURAL NETWORKS

Aladag *et al.* [8] gave brief information about the artificial neural networks as follows:'What is an artificial neural network?' is the first question that should be answered. Picton [9] answered this question by separating this question into two parts. The first part is why it is called as artificial neural network. It is called as artificial neural network because it is a network of interconnected elements. These elements were inspired from studies of biological nervous systems. In other words, the artificial neural networks are an attempt at creating machines that work in a similar way to the human brain by building these machines using components that behave like biological neurons.

The second question is what an artificial neural network does? The function of an artificial neural network is to produce an output pattern when presented with an input pattern. In forecasting, the artificial neural networks are mathematical models that imitate the biological neural networks. The artificial neural networks consist of some elements. Determining the elements of the artificial neural networks issue that affect the forecasting performance of artificial neural networks should be considered carefully. Elements of the artificial neural networks are generally given as network architecture, learning algorithm and activation function [10].

One critical decision is to determine the appropriate architecture, that is, the number of layers, the number of nodes in each layers and the number of arcs which interconnects with the nodes [11]. However, in the literature, there are not general rules for determining the best architecture. Therefore, several architectures should be tried for the correct results. There are various types of artificial neural networks. One of them is called as feed forward neural networks. The feed forward neural networks have been used successfully in many studies [10]. In the feed forward neural networks, there are no feedback connections. The broad feed forward neural network architecture that has single hidden layer and single output is given as an illustration in Fig. **1**.

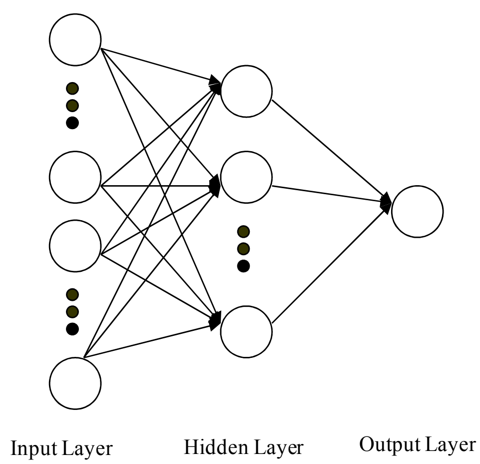


Figure 1: A broad feed forward neural network architecture.

Learning of an artificial neural network for a specific task is equivalent to finding the values of all weights such that the desired output is generated by the corresponding input. Various training algorithms have been used for the determination of the optimal weights values. The most popularly used training method is the back propagation algorithm presented by Smith [12]. In the back propagation algorithm, learning of the artificial neural networks consists of adjusting all weights considering the error measure between the desired output and actual output [13].

Another element of the artificial neural networks is the activation function. It determines the relationship between inputs and outputs of a network. In general, the activation function introduces a degree of the non-linearity that is valuable in most of the artificial neural networks applications. The well known activation functions are logistic, hyperbolic tangent, sine (or cosine) and the linear functions. Among them, logistic activation function is the most popular one [14].

The artificial neural networks approach, as indicated above, has been rarely used for public expenditure forecasting particularly in Turkey. Before the application of the artificial neural networks on public expenditures forecasting, the next section briefs about the new budgetary approach recently become the norm by the publications of the Public Financial Management and Control Law (No: 5018) in Turkey.

3. THE NEW BUDGETARY APPROACH IN TURKEY

By means of the Public Financial Management and Control Law (No: 5018) published in 2004 and the switch on the Middle-Term Performance Based Budgetary approach, Turkey has taken important steps to attain the budgetary and financial discipline in the public sector. Undoubtedly the most important of these steps is the initiation of performance based budgetary process designed according to the strategic plan of each public institution. The strategic plan is prepared for five years in the light of the public institution's mission, vision, medium and long term goals. The strategic plan includes performance programme, performance indicators and monitoring systems which will be used to assess the achievements of the related public institution. The performance programme is the yearly reflection of the strategic plan, assessing whether the public resources are used in effective, economic and efficient way during the year. The strategic plan is prepared following a participatory process of shareholders and regarded as the main document establishing the link between the future perspective, performance and budget allowances of the public institution.

In the performance based budgeting approach, the central government guides the public institutions by two main documents used for the determination of macroeconomic goals and the budget allowances limitations. These are the Middle-Term Programme and the Middle-Term Plan prepared by the SPO and the Ministry of Finance (MoF), respectively. The Middle-Term Programme focuses on the macroeconomic indicators, while the Middle-Term Plan defines the limits of allowances that the public institutions ought to take into consideration.

Under the performance based budgeting approach there are three main documents determining the boundary of expenditures of the public institutions. These are the middle-term programme, the middle-term financial plan, and the strategic plan and performance programme. According to the Public Financial Management and Control Law (No: 5018), the middle-term programme prepared by the SPO determines the basic economic thresholds by taking into consideration of general economic conditions, strategic and development plans. Similarly, the middle-term plan prepared in line with the middle-term programme by the MoF defines the budget revenues and expenditures, the likely amounts of deficit/surplus, and the budget allowances ceilings for the next three years.

Ultimately, the public institutions are expected to forecast their expenditures as performance based and in line with the five year strategic plans and the yearly performance programmes. Meanwhile, these expenditure forecasts are expected to be compatible with the middle-term programme and the middle-term financial plan, as well. Unlike the classical budgetary approach, the performance based budgetary approach is pursued in a participatory involvement of shareholders in the budget process.

The accuracy of budget forecasts is also important in terms of two budgeting principles, namely, accuracy and sincerity, concerning the main stages of budget preparation, implementation and control. The principle of accuracy requires that the budget is prepared accordingly to the current macroeconomic situation. That is the budget ought to include objective forecasts of both revenues and expenditures. The principle of sincerity on the other hand requires that those who prepare the budget should not determine the revenues and the expenditures more or less than needed.

PART II: FUZZY TIME SERIES

CHAPTER 7

A New Method for Forecasting Fuzzy Time Series with Triangular Fuzzy Number Observations

Erol Eğrioğlu[1,*], Cagdas Hakan Aladag[2] and Ufuk Yolcu[1]

[1]Ondokuz Mayis University, Faculty of Arts and Science, Department of Statistics, 55139, Samsun, Turkey and [2]Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey

Abstract: Most of the time series faced in real life are fuzzy time series and these time series have to be forecasted by fuzzy time series forecasting methods. Therefore, there have been many studies in the literature in which various fuzzy time series approaches are proposed. The fuzzy time series methods introduced in the literature have been generally proposed to analyze fuzzy time series whose observations are fuzzy sets. On the other hand, Song *et al.* firstly improved a fuzzy time series model to analyze fuzzy time series whose observations are triangular fuzzy numbers [1]. Their method requires complex arithmetic operations for triangular fuzzy numbers. We propose a novel fuzzy time series forecasting approach based on simulation and feed forward neural networks to forecast fuzzy time series including triangular fuzzy numbers. The proposed method is applied to gold prices in Turkey series to show the applicability of the method.

Keywords: Artificial neural networks, Fuzzy time series, Forecasting, Gold prices in Turkey, Triangular fuzzy number.

1. INTRODUCTION

Fuzzy time series are time series whose observations are fuzzy sets. In the literature, various approaches have been proposed and have been used to analyze fuzzy time series. Observations of fuzzy time series can be fuzzy sets and can also be triangular fuzzy numbers which are special type of fuzzy sets. A definition of fuzzy time series based on triangular fuzzy numbers was firstly introduced by Song *et al.* [1]. Then, Hong [2] defined the arithmetic operations related to the triangular fuzzy numbers given in [1] by using weakest t-norm and extended the definitions presented in [1]. Although there have been many studies in fuzzy time series literature, there have been only these two studies which are based on triangular fuzzy numbers. Furthermore, in these studies, any implementation was not included and it was not explained how the proposed methods work.

The most of the studies in fuzzy time series literature have been focused on the methods that proposed to forecast fuzzy time series whose observations are fuzzy sets consist of intervals. Song and Chissom [3] firstly proposed an approach to forecast such fuzzy time series. Then, Chen proposed new approaches in [4, 5] and these methods are well known ones in the literature. In addition, Tseng [6, 7] proposed new models which are like traditional time series models. In these models, however, the parameters are fuzzy numbers.

In this study, a novel fuzzy time series forecasting model based on simulation and feed forward neural networks is proposed to forecast fuzzy time series whose observations are triangular fuzzy numbers. Using the proposed method easier than using the methods presented in [1, 2] since the proposed method does not require complex arithmetic operations for triangular fuzzy number. Also the proposed approach is applied to gold prices in Turkey series in the implementation.

This chapter gives in Section 2 brief information about artificial neural networks. Section 3 includes the basic definitions for the proposed method and the algorithm of the method is introduced in this section. The implementation is given in Section 4. Finally, Section 5 provides the concluding remarks.

***Address correspondence to Erol Eğrioğlu:** Ondokuz Mayis University, Faculty of Arts and Science, Department of Statistics, 55139, Samsun, Turkey; E-mail: erole@omu.edu.tr

Cagdas Hakan Aladag and Erol Eğrioğlu (Eds)

2. ARTIFICIAL NEURAL NETWORKS

Artificial neural networks are mathematical algorithms that simply mimic the biological neural networks. Although their structure is not as complex as the structure of biological neural networks, many real life problems can be solved by using artificial neural networks. There are different types of artificial neural networks [8]. Especially feed forward neural networks have been preferred in time series forecasting problem since this kind of neural network has been proved its success in many forecasting studies [9].

Feed forward neural networks compose of layers such as input, hidden and output layers. And each layer consists of elements which are called neuron. Layers are connected with each other by weights. However, there is no interconnection with neurons in the same layer. The number of hidden layer can be more than one. In time series forecasting, the input values are the lagged variables of the time series. Target value is the value belongs to the next period and the output of the network will be the prediction value for the next period. A broad multi layer feed forward neural network architecture that contains one hidden layer and one neuron in the output layer is illustrated in Fig. **1**.

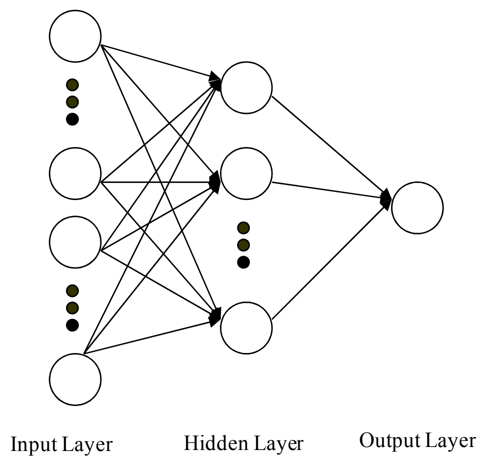


Figure 1: Multilayer feed forward artificial neural network with one output neuron.

After the inputs for neurons in the hidden and the output layers are multiplied by corresponding weights, they are summed and the value of the activation function is calculated for this sum. By employing activation function, artificial neural networks can model the non linear part of time series. To produce accurate forecasts, artificial neural networks can be trained. In the training process, the best values of the weights are tried to be found. The closer the difference between the prediction and the target value is, the better the weight values are. To train artificial neural networks, various optimization algorithms are used and these algorithms are called as training algorithms [10].

Feed forward neural networks have produced very good results in time series forecasting applications. On the other hand, to reach accurate forecasts, the elements of artificial neural networks such as architecture structure, activation function and training algorithm should be carefully determined since these main components have important effect on forecasting performance of the method [11].

3. THE PROPOSED METHOD

In recent years, various approaches have been proposed to forecast fuzzy time series. The observations of fuzzy time series can be fuzzy sets or fuzzy numbers, which can be considered as special type of fuzzy sets. Most of fuzzy time series forecasting approaches available in the literature were improved to forecast fuzzy time series whose observations consist of elements that are subsets of universe of discourse. On the other hand, in the studies [1, 2], two forecasting approaches were proposed to forecast fuzzy time series consist of observations that are fuzzy numbers. These two approaches are based on arithmetic operations for fuzzy numbers. In this study, we propose a novel forecasting model based on feed forward neural networks and simulation in order to forecast fuzzy time series whose observations are triangular fuzzy numbers. The proposed method can be used for time series such as share prices of stockholders and gold prices which have lowest and highest values for a specific time interval. Besides, it is possible to apply the proposed method to other time series which can be considered as fuzzy time series. The fuzzy time series forecasting approach proposed in this study provides some advantages. And, these advantages can be summarized as follows:

- Using the proposed method is easier than using the methods introduced in [1, 2] since the proposed method does not require complex arithmetic operations related to fuzzy numbers.

- The proposed method does not require satisfying any assumptions such as linearity, normal distribution and a specific observation number.

- The proposed method has the ability of flexible modeling which is also included by artificial neural networks.

- The proposed approach produces more accurate forecasts than those obtained from most of the methods available in the literature.

The algorithm of the proposed forecasting method is given below:

Algorithm of the Proposed Method

Step 1. Fuzzification

The fuzzification process can be changed depend on the observed time series. If the time series has the lowest and the highest values in time intervals in which the time series observations are observed, in the fuzzification step, the lowest and the highest values are taken as the left and the right side values of the triangular fuzzy number for each observations, respectively. The center value of the triangular fuzzy number is equal to mean value of the corresponding left and right side values. In this way, each observation is turned into a triangular fuzzy number. If there is only a crisp value for the time interval in which the corresponding observation of the time series are observed, this crisp value is taken as the center value of the triangular fuzzy number. Then, a spread value is determined and the left and the right side values are calculated by using this spread value. Therefore, the obtained triangular fuzzy number will be symmetrical one.

Step 2. Determining Fuzzy Relations

To establish fuzzy relationships, feed forward neural networks are used. By using simulation method, the fuzzy time series whose observations are triangular fuzzy numbers is turned into a time series composes of crisp observations in order to apply feed forward neural networks. A fuzzy time series $F(t)$ can be given as follows:

$$F(t) = (l(t), m(t), r(t)),\ t = 1,2,\ldots,n \tag{1}$$

where $l(t)$,$m(t)$ and $r(t)$ are the left side, the center and the right side value, respectively, for the i th observation of the fuzzy time series.

Step 2.1. Calculating the Crisp Time Series

First of all, it is determined how many crisp time series will be analyzed and this number can be represented by itr. When each crisp time series is generated, the following way is used.

- It is assumed that α has a uniform distribution (α~Unifrom(0,1)) so an α value can be generated by using this distribution.

- The obtained α cut values for every observation can be calculated by using the equation given in (2) depend on the value of that α.

CHAPTER 8

New Criteria to Compare Interval Estimates in Fuzzy Time Series Methods

Erol Eğrioğlu*, V. Rezan Uslu and Senem Koc

Ondokuz Mayis University, Faculty of Arts and Science, Department of Statistics, 55139, Samsun, Turkey

Abstract: The idea of exploring fuzzy set theory to time series forecasting issues has been enormously attracted researcher's attention in recent years. Several new approaches on fuzzy time series have been put forward. These approaches have got some advantages related to classical methods and are complementary of them. Two of these kinds of procedures are FARIMA and FSARIMA. FARIMA and FSARIMA do not require a restriction of at least 50 observations and linearity assumption. The methods of FARIMA and FSARIMA provide interval estimates of a time series. ARIMA and SARIMA also provide interval estimation but it has been put forward that estimated intervals are large, therefore not informative. The width of estimated intervals obtained from FARIMA and SARIMA may generally tend to be less than ones from ARIMA and SARIMA. In the literature, there has been no study which provides a criterion for the comparisons of time series with respect to interval estimates. In this study, two criteria for such comparisons are presented.

Keywords: ARIMA, Fuzzy ARIMA, Fuzzy SARIMA, Interval estimates, SARIMA.

1. INTRODUCTION

In time series analysis the linear and nonlinear procedures are generally classified into two groups; univariate and bivariate time series. In time series analysis, the Box–Jenkins methodology, applies autoregressive moving average (ARMA) or autoregressive integrated moving average (ARIMA) models to find the best fit of a time series to past values of this time series, in order to make forecasts. In order to apply the Box-Jenkins methodology, however, two assumptions are generally required; one of them is that the applied model must be linear and the other is that there must be at least 50 observations. Most often researches have been faced some difficulties in providing these assumptions dealing with real life time series data. In that point fuzzy time series approaches can be cure when they face such problems. Since these approaches are data- based approaches there might not need some kind of limitations. Fuzzy logic concept was firstly introduced by [1]. The first implementation of fuzzy logic on time series was explored in the study of [2], by introducing a fuzzy linear regression model. This implementation was used by [3] who puts forward fuzzy ARIMA.

Fuzzy ARIMA and fuzzy SARIMA are aimed to get interval estimates. Traditional ARIMA and SARIMA can provide the point estimation whereas they also provide interval estimates when the normality assumption is maintained. Fuzzy approaches do not require the normality assumption, moreover since the fuzzy interval estimates are generally narrower than the traditional ones they can interpret more easily. In fuzzy literature the comparisons are being made graphically. So far there are no criteria for the comparisons of interval estimates. In this study two new criteria have been proposed to do this. This is shown empirically on the data of the amount of carbon dioxide measured monthly in Ankara. SARIMA and Fuzzy SARIMA have been applied to this data set. For this analysis some subroutines are programmed in MATLAB.

In the second section, fuzzy ARIMA method is summarized. Fuzzy SARIMA method is briefly represented in the third section. In the fourth section, the proposed criteria are introduced. Application of real life time series is given section 5. The obtained results of application are discussed and the conclusions are given in the last section.

***Address correspondence to Erol Eğrioğlu:** Ondokuz Mayis University, Faculty of Arts and Science, Department of Statistics, 55139, Samsun, Turkey; E-mail: erole@omu.edu.tr

Cagdas Hakan Aladag and Erol Eğrioğlu (Eds)

2. FUZZY ARIMA

ARIMA is a most powerful analyzing method for time series as long as its assumptions, such as linearity and the number of observations, are satisfied. Most real life time series data cannot, however, provide these assumptions. In ARIMA model, the value of observation at time t of a time series is a linear combination of the value of the observation at time t-1, t-2 … and error term [4]. In practice it is very difficult to construction such a relationship. [5] proposed Fuzzy SARIMA by exploring Tanaka's possibility approach to Seasonal ARIMA. [3] provides a experimental study of analyzing of a real life time series data *via* Fuzzy ARIMA. The method in this experimental study is summarized as follows:

In Fuzzy ARIMA approach the model parameters are assumed fuzzy as Tanaka's fuzzy regression model. The parameters $\tilde{\boldsymbol{\varphi}} = \left(\tilde{\varphi}_1, ..., \tilde{\varphi}_p\right)$ and $\tilde{\boldsymbol{\theta}} = \left(\tilde{\theta}_1, ..., \tilde{\theta}_q\right)$ are used instead of $\boldsymbol{\varphi} = \left(\varphi_1, ..., \varphi_p\right)$ and $\boldsymbol{\theta} = \left(\theta_1, ..., \theta_q\right)$ in fuzzy ARIMA model, respectively. Then the model is written as

$$\tilde{Z}_t = \tilde{\varphi}_1 Z_{t-1} + \cdots + \tilde{\varphi}_p Z_{t-p} + a_t - \tilde{\theta}_1 a_{t-1} - \cdots - \tilde{\theta}_q a_{t-q} \tag{1}$$

For the simpler notation the model is rewritten as

$$\tilde{Z}_t = \tilde{\beta}_1 Z_{t-1} + \cdots + \tilde{\beta}_p Z_{t-p} + a_t - \tilde{\beta}_{p+1} a_{t-1} - \cdots - \tilde{\beta}_{p+q} a_{t-q} \tag{2}$$

The membership functions for the fuzzy parameters, which are expressed in the form of triangular fuzzy number, are as below.

$$\mu_{(\tilde{\beta}_i)}(\beta_i) = \begin{cases} 1 - \dfrac{|\beta_i - \alpha_i|}{c_i} & , \ \alpha_i - c_i \le \beta_i \le \alpha_i + c_i \\ 0 & , \ o.w \end{cases} \tag{3}$$

With respect to the extension principal, the membership function of $\tilde{Z}_i$ is

$$\mu_Z(Z_t) = \begin{cases} 1 - \dfrac{\left|Z_t - \sum_{i=1}^{p} \alpha_i Z_{t-i} - a_t + \sum_{i=p+1}^{p+q} \alpha_i a_{t+p-i}\right|}{\sum_1^p c_i \left|Z_{t-i}\right| + \sum_{i=p+1}^{p+q} c_i \left|a_{t+p-i}\right|} & , \quad Z_t \ne 0, \quad a_t \ne 0, \\ 0 & , \quad o.w. \end{cases} \tag{4}$$

Similar to fuzzy regression, the membership degree of each observation Z_i should be greater than a predetermined value h ($h \in [0,1]$). The choice of h is influenced on the extension of scale parameters. The case is expressed by

$$\boldsymbol{\mu_Z}(Z_t) \ge h, \quad t = 1, 2, \ldots, k \tag{5}$$

The total uncertainty in fuzzy ARIMA is given by

$$S = \sum_{i=1}^{p} \sum_{t=1}^{k} c_i \left|\phi_{ii}\right| \left|Z_{t-i}\right| + \sum_{i=p+1}^{p+q} \sum_{t=1}^{k} c_i \left|\rho_{i-p}\right| \left|a_{t+p-i}\right| \tag{6}$$

where φ_{ii} is the ith partial autocorrelation coefficient and ρ_i is the ith autocorrelation coefficient. The estimates of fuzzy ARIMA model parameters are obtained by solving the following linear programming model.

$$Min \quad S=\sum_{i=1}^{p}\sum_{t=1}^{k}c_i\left|\varphi_{ii}\right|\left|Z_{t-i}\right|+\sum_{i=p+1}^{p+q}\sum_{t=1}^{k}c_i\left|\rho_{i-p}\right|\left|a_{t+p-i}\right|$$

$YŞ$;

$$\sum_{i=1}^{p}\alpha_i Z_{t-i}+a_t-\sum_{i=p+1}^{p+q}\alpha_i a_{t+p-i}+(1-h)\left(\sum_{i=1}^{p}c_i\left|Z_{t-i}\right|+\sum_{i=p+1}^{p+q}c_i\left|a_{t+p-i}\right|\right)\geq Z_t \quad ,t=1,2,\ldots,k \tag{7}$$

$$\sum_{i=1}^{p}\alpha_i Z_{t-i}+a_t-\sum_{i=p+1}^{p+q}\alpha_i a_{t+p-i}-(1-h)\left(\sum_{i=1}^{p}c_i\left|Z_{t-i}\right|+\sum_{i=p+1}^{p+q}c_i\left|a_{t+p-i}\right|\right)\leq Z_t \quad ,t=1,2,\ldots,k$$

$$c_i\geq 0$$

[3] used the three stage procedure, which is given in below, for the application of fuzzy ARIMA.

Stage 1. When the time series is not fuzzy time series ARIMA(p,d,q) model is analyzed by Box-Jenkins procedure. The obtained estimates are taken as the centered value of the fuzzy parameters.

$$\boldsymbol{\alpha}^*=\left(\alpha_1^*,...,\alpha_{p+q}^*\right)=\left(\hat{\varphi}_1,...,\hat{\varphi}_p,\hat{\theta}_1,...,\hat{\theta}_q\right)$$

Stage 2. The centered values $\boldsymbol{\alpha}^*=\left(\alpha_1^*,...,\alpha_{p+q}^*\right)$ and the residuals (a_t) from Stage 1 are used as the inputs of the problem of minimization. Then it is solved and found the estimates of scale parameters. The fuzzy ARIMA model

$$\tilde{Z}_t=<\alpha_1,c_1>Z_{t-1}+\cdots+<\alpha_p,c_p>Z_{t-p}+a_t-\cdots-<\alpha_{p+q},c_{p+q}>a_{t-q} \tag{8}$$

Stage 3. The possibility fuzzy regression method is sensitive to outliers then the observations, which are close to the estimates of lower and upper limits of fuzzy ARIMA, are deleted and the fuzzy ARIMA is repeated.

3. FUZZY SARIMA

There are some limitations as well as fuzzy ARIMA. At least 50 observations are at hand. In fuzzy SARIMA approach it is also assumed that the parameters are fuzzy. $\tilde{\boldsymbol{\varphi}}=\left(\tilde{\varphi}_1,...,\tilde{\varphi}_p\right)$, $\tilde{\Phi}=(\tilde{\Phi}_1,...,\tilde{\Phi}_p)$, $\tilde{\Theta}=(\tilde{\Theta}_1,...,\tilde{\Theta}_p)$ and $\tilde{\boldsymbol{\theta}}=\left(\tilde{\theta}_1,...,\tilde{\theta}_q\right)$ $\boldsymbol{\varphi}=\left(\varphi_1,...,\varphi_p\right)$ are also formed as triangular fuzzy number. The parameter of a_t is $\tilde{\gamma}_t$. Then the fuzzy SARIMA model with fuzzy parameters $(p,d,q)(P,D,Q)$ is

$$\tilde{\phi}(B)\tilde{\Phi}(B^s)W_t=\tilde{\beta}_0+\tilde{\theta}(B)\tilde{\Theta}(B^s)a_t \quad W_t=(1-B)^d(1-B^s)^D Z_t$$

$$\tilde{W}_t=\tilde{\beta}_0+\sum_{i=1}^{p}\tilde{\phi}_i W_{t-i}+\sum_{i=1}^{p}\tilde{\Phi}_i W_{t-is}-\sum_{i=1}^{p}\phi_i\Phi_1 W_{t-s-i}-\sum_{i=1}^{p}\phi_i\Phi_2 W_{t-2s-i}-...$$

$$-\sum_{i=1}^{p}\phi_i\Phi_p W_{t-Ps-i}+\tilde{\gamma}a_t-\sum_{i=1}^{q}\tilde{\theta}_i a_{t-i}-\sum_{i=1}^{Q}\Theta_i a_{t-is+}$$

where Z_t's are the observations, The membership degree of $\tilde{\gamma}_t$ is 1. The equation above is rewritten as

$$\tilde{W}_t=\tilde{\beta}_0+\sum_{i=1}^{p}\tilde{\beta}_i W_{t-i}+\sum_{i=1}^{P}\tilde{\beta}_{p+i}W_{t-is}-\sum_{j=1}^{P}\sum_{i=1}^{p}\tilde{\beta}_i\tilde{\beta}_{p+j}W_{t-js-i}+\tilde{\beta}_{p+P+1}a_t-\sum_{i=1}^{q}\tilde{\beta}_{p+P+1+i}a_{t-i}$$

$$\sum_{i=1}^{Q}\tilde{\beta}_{p+P+q+1-i}a_{t-is}+\sum_{j=1}^{Q}\sum_{i=1}^{q}\tilde{\beta}_{p+P+1+i}\tilde{\beta}_{p+P+q+1+j}a_{t-js-i}.$$

CHAPTER 9

The Effect of the Length of Interval in Fuzzy Time Series Models on Forecasting

Erol Eğrioğlu[1] and Cagdas Hakan Aladag[2,*]

[1]Ondokuz Mayis University, Faculty of Arts and Science, Department of Statistics, 55139, Samsun, Turkey and [2]Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey

Abstract: Due to the vagueness that they contain in their observations, fuzzy time series models worked in two main categories such as first order and high order models, has an ever expending field of study. Fuzzy time series analysis method is highly effective in uncovering the relations of this type of time series structure. In the implementation of fuzzy time series methods, it is crucial to determine the model order in terms of forecasting performance. Besides, regardless of the model order, the length of interval determined in the partition phase of the universe of discourse, greatly affects forecasting performance. Therefore, there have been numerous studies focusing on determining the length of interval in the literature. This study aims to introduce the significance of interval length determination in fuzzy time series analysis method on forecasting performance. For this purpose, related methods are introduced, implementation of two real time series is shown and some comparisons between methods are made and finally obtained results are discussed.

Keywords: Fuzzy time series, Forecasting, Length of interval, Optimization.

1. INTRODUCTION

The implementation of fuzzy time series methods in most common time series with vagueness gives successful results. As well as its effective forecasting performance and not containing constraints found in conventional approaches make fuzzy time series approach applicable and attractive.

Fuzzy time series, which are widely used nowadays, was first proposed by Song and Chissom [1-3]. This followed by simpler methods such as Sullivan and Woodall's models [4] based on Markov model and Chen's model [5] without needing operational matrix processes and using fuzzy logic relation and group relation tables in the relation determination. Additionally, Huarng [6] proposed average and distribution based approaches and he showed that lengths of interval determined in the partition stage of universe of discourse have effect on forecasting performance. Eğrioğlu *et al.* [7] used univariate restricted optimization which minimizes the forecasting error in the determination of the length of interval.

It is a well known fact that model order affects the forecasting performance in fuzzy time methods. Therefore, researchers have used various high order models in the analysis of fuzzy time series. Chen [8], Aladag *et al.* [9] and Eğrioğlu *et al.* [7] aimed to improve forecasting performance using high order models with proposed fuzzy time series analyses approaches. In their studies, while Chen [8] and Eğrioğlu *et al.* [7] determined the fuzzy relations with fuzzy logic relations and group relations tables, Aladag *et al.* [9] used artificial neural networks.

All fuzzy time series methods, whether first order or high order, are sensitive to length of interval determined in the partition stage of universe of discourse in terms of forecasting accuracy. Therefore, interval lengths should be selected in a way that provides the best forecasting accuracy. In the literature, several studies were put forward concerning determination of interval lengths. While the length of interval has been determined subjectively in some approaches, others have tried to improve forecasting performance by bringing more systematic approaches. It should be noted that selection of larger intervals will destroy fluctuation in the series and prevent the emergence of existing relations. Also selection of smaller intervals will eliminate the difference between conventional analysis and fuzzy time series.

***Address correspondence to Cagdas Hakan Aladag:** Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey; E-mail: chaladag@gmail.com

In the second section of the study, fuzzy time series and its definitions are introduced. Third section introduces first order and high order approaches proposed by Chen [8] and approaches proposed by Huarng [6] and Eğrioğlu *et al.* [7]. Fourth section deals with the application and the application results of the approaches concerning determination of interval length. Finally, the last section provides a conclusion.

2. FUZZY TIME SERIES

The definition of fuzzy time series was firstly introduced by Song and Chissom [1, 2]. In fuzzy time series approaches, the validation of theoretical assumptions does not need to be checked just as in conventional time series procedures. The most important advantage of fuzzy time series approaches is to be able to work with a very small set of data and not to require the linearity assumption. Some basic definitions of fuzzy time series can be given as follows [7]:

Let U be the universe of discourse, where $U = \{u_1, u_2, ..., u_b\}$. 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 $1 \le a \le b$.

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 $f_1(t), f_2(t), ...$ 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 relationship.

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

$$F(t-m), ..., F(t-2), F(t-1) \to F(t),$$

and it is called the m^{th} order fuzzy time series forecasting model.

3. THE RELATED METHODS AVAILABLE IN THE LITERATURE

In this section, the methods which were used in the analysis are introduced to emphasize the effect of interval length determination on forecasting performance in fuzzy time series analyses.

3.1. Chen's First Order Fuzzy Time Series Method

The method proposed by Chen [5] is based on fuzzy logical relationships and group relation tables. In this method, the lengths of intervals used in partitioning universe of discourse are determined subjectively and the effect of these intervals on forecasting is evident. The following algorithm can be given for the first order fuzzy time series forecasting model proposed by Chen [5]:

Step 1. Define the discourse of universe and subintervals. Based on min and max values in the data set, $D_{\min}$ and $D_{\max}$ variables are defined. Then choose two arbitrary positive numbers which are D_1 and D_2 in order to divide the interval evenly.

$$U = [D_{\min} - D_1, D_{\max} + D_2]. \quad (1)$$

Step 2. Define fuzzy sets based on the universe of discourse and fuzzify the historical data.

Step 3. Fuzzify observed rules.

Step 4. Establish fuzzy logical relationships and group them based on the current states of the data of the fuzzy logical relationships.

For example, $A_1 \to A_2, A_1 \to A_1, A_1 \to A_3$, can be grouped as: $A_1 \to A_2, A_3, A_1$.

Step 5. Forecast.

$$\text{Let } F(t-1) = A_i$$

Case 1: There is only one fuzzy logical relationship in the fuzzy logical relationship sequence. If $A_i \to A_j$, then $F(t)$, forecast value, is equal to A_j .

Case 2 : If $A_i \to A_i, A_j, \ldots, A_k$, then $F(t)$, forecast value, is equal to $A_i, A_j, \ldots, A_k$.

Case 3: If $A_i \to empty$, then $F(t)$, forecast value, is equal to A_i

Step 6. Defuzzify.

Apply "Centroid" method to get the results. This procedure (also called center of area, center of gravity) is the most often adopted method of defuzzification. Suppose that the fuzzy forecast of F(t) is A_k. The defuzzyfied forecast is equal to the midpoint of the interval which corresponds to A_k .

3.2. Chen's High Order Fuzzy Time Series Method

Chen proposed a method based on high order fuzzy time series which enable to obtain forecasts [8]. The method proposed by Chen [8] produces more accurate forecasts than the first order fuzzy time series methods. In this method, the lengths of intervals are determined subjectively. This affects forecasting accuracy as in the other methods. The model given in definition 3 can be analyzed by the high order fuzzy time series approach. The steps of the algorithm of the method proposed by Chen can be given as follows [8]:

Step 1. Define the discourse of universe and subintervals. Based on min and max values in the data set, $D_{\min}$ and $D_{\max}$ variables are defined. Then choose two arbitrary positive numbers which are D_1 and D_2 in order to divide the interval evenly.

$$U = \left[D_{\min} - D_1, D_{\max} + D_2\right]. \qquad \textbf{(2)}$$

Step 2. Define fuzzy sets based on the universe of discourse and fuzzify the historical data.

Step 3. Fuzzify observed rules.

Step 4. Establish fuzzy logical relationships and group them based on the current states of the data of the fuzzy logical relationships. Based on the linguistically defined variables, k^{th} order fuzzy logical relationship $A_{ik}, A_{i(k-1)}, \ldots, A_{il} \to A_j$ can be established. For example, the values of the year $i-1$ and i corresponds to fuzzy values A_a and A_b. Also, the values of the year $i+1$ corresponds to fuzzy value A_j. Therefore, 2^{th} order fuzzy logical relationship can be written as $A_a, A_b \to A_j$. In a similar manner, the more high order fuzzy logical relationships and fuzzy logical groups for 3th, 4th and other high orders are constructed.

Step 5. Forecast and Defuzzify.

In here, fuzzy values are defuzzyfied and forecasts are done.

CHAPTER 10

Determining Interval Length in Fuzzy Time Series by Using an Entropy Based Approach

Cagdas Hakan Aladag*, Irem Degirmenci and Suleyman Gunay

Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey

Abstract: Various theoretical assumptions in conventional time series methods do not need to be checked in fuzzy time series approach. Therefore fuzzy time series are preferred in many applications. The identification of the length of intervals is an important issue and affects the forecasting performance. But in many studies in the literature, the length of intervals is determined randomly. Starting from this point, Huarng [1] has proposed two novel approaches which are based on the distribution and the average to choose a more effective length. Huarng and Yu [2] used a dynamic approach for adjusting lengths of interval. Huarng [3] suggested a different method which is called ratio based lengths of intervals. Cheng *et al.* [4] have proposed a new approach by using entropy. Eğrioğlu *et al.* [5] and Yolcu *et al.* [6] have determined the lengths of intervals by using optimization. At the first stage of the method proposed by Cheng *et al.* [4], a specific method has not been used and classes have been assigned intuitively while classes to which data belong were generating. In this study, the approach proposed by Degirmenci *et al.* [7] is applied to the enrollment data at the University of Alabama and the yearly data of the quantities of clean water used in Istanbul. Then obtained forecasts are compared with those obtained from other methods available in the literature.

Keywords: Entropy, Forecasting, Length of interval, Fuzzy c-means clustering, Fuzzy time series.

1. INTRODUCTION

Fuzzy set theory was firstly introduced by Zadeh [8] and it has found many application areas since that time. In fuzzy time series, one of the most important application areas is forecasting. Fuzzy time series approach has been introduced as an alternative method for conventional time series models. In contrast to conventional time series methods, various theoretical assumptions do not need to be checked in fuzzy time series approach. The traditional time series approaches require having the linearity assumption and at least 50 observations. The most important advantages of fuzzy time series approach are to be able to work with a few observations and not to require the linearity assumption. In addition to this, the traditional time series methods may be insufficient to forecast the time series whose observations include uncertainties (temperature, stock, *etc.*), even in some cases, traditional approximations can not produce any solution [9]. Fuzzy time series can also be applied to this kind of time series.

Fuzzy time series was firstly introduced by Song and Chissom [10-12]. Afterwards, Sullivan and Woodall [13] proposed an approach based on Markov model. Then, Chen [9] proposed a new method based on Song and Chissom's approach. This method is easier than the method proposed by Song and Chissom and consists of considerably simple calculations. Wang and Lee [14] have revealed a new method which used a different method of fuzzification. In later years, Chen [15] has proposed a new algorithm for high order fuzzy time series forecasting model.

Although many studies have been made about forecasting with fuzzy time series, there are some problems in these studies. The identification of the length of intervals is one of these problems. In fuzzy time series approaches, at the forecasting process, the length of intervals affects the forecasting performance. Hence, it is important to choose an effective length of intervals for improving forecasting accuracy in fuzzy time series approaches. Starting from this point, Huarng [1] has proposed two novel approaches which are based

***Address correspondence to Cagdas Hakan Aladag:** Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey; E-mail: chaladag@gmail.com

Cagdas Hakan Aladag and Erol Eğrioğlu (Eds)

on the distribution and the average to choose a more effective length. Huarng and Yu [2] used a dynamic approach for adjusting lengths of interval. Huarng [3] suggested a different method which is called ratio based lengths of intervals. Cheng *et al.* [4] have proposed a new approach by using entropy. Eğrioğlu *et al.* [5] and Yolcu *et al.* [6] have determined the lengths of intervals by using optimization.

At the first stage of the method proposed by Cheng *et al.* [4], a specific method has not been used and classes have been assigned intuitively while classes to which data belong were generating. In the approach proposed by Degirmenci *et al.* [7], fuzzy c-means data clustering technique was used to classify data. Afterwards, lengths of intervals were determined by entropy approach.

In this study, the approach proposed by Degirmenci *et al.* [7] is applied to the enrollment data at The University of Alabama which have been used in the most of the other studies in the literature. The yearly data of the quantities of clean water used in Istanbul are also used to reveal that the performance of the proposed approach. In section 2, the fundamental definitions about fuzzy time series are presented. In the third section, Chen's method is given briefly. In subsequent section, the approach proposed by Degirmenci *et al.* [7] and its applications results are presented and compared with other methods in the literature. In final section, the results are discussed.

2. FUZZY TIME SERIES

Some basic definitions of fuzzy time series are given as follows [16]:

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 and $f_{A_i}: U \rightarrow [0,1]$. 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 $1 \leq a \leq b$.

Definition 1. Let $Y(t)$ $(t = .,0,1,2, \ldots)$ 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. 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 $F(t)$ and $F(t-1)$, where * represents as an operator. To sum up, let $F(t-1) = A_i$ and $(t) = A_j$. The fuzzy logical relationship between $F(t)$ and $F(t-1)$ can be denoted as $A_i \rightarrow A_j$ 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 relationship.

3. CHEN'S METHOD [9]

Chen [9] has improved the approximation given by Song and Chissom [10, 11] and he proposes a new method which uses a simpler operation instead of complex matrix operations. The algorithm of Chen's method can be given as follows:

Step 1. Define the universe of discourse and intervals.

Based on the domain issue, The universe of discourse can be defined as: $U = [starting, ending]$. As the length of interval is determined, U can be partitioned into several equally length intervals.

Step 2. Define fuzzy sets based on the universe of discourse and fuzzify the historical data.

Step 3. Fuzzify observed rules.

Step 4. Establish fuzzy logical relationships and group them based on the current states of the data of the fuzzy logical relationships.

For example, $A_1 \rightarrow A_2, A_1 \rightarrow A_1, A_1 \rightarrow A_3$, can be grouped as: $A_1 \rightarrow A_2, A_3, A_1$.

Step 5. Forecast.

Let $F(t-1) = A_i$.

Case 1: There is only one fuzzy logical relationship in the fuzzy logical relationship sequence.

If $A_i \rightarrow A_j$, then $F(t)$, forecast value, is equal to A_j.

Case 2: If $A_i \rightarrow A_i, A\ \ , \ldots, A_k$, then $F(t)$, forecast value, is equal to $A_i, A_j, \ldots, A_k$.

Case 3: If $A_i \rightarrow empty$, then $F(t)$, forecast value, is equal to A_i.

Step 6. Defuzzify.

Apply "Centroid" method to get the results. This procedure is the most often adopted method of defuzzification.

4. THE METHOD PROPOSED BY DEGIRMENCI *et al.* [7]

One of the most important problems in the fuzzy time series forecasting is determining the lengths of intervals. The length of intervals affects the performance of forecasting significantly. One of the methods for choosing the length of intervals is the entropy approach which was proposed by Cheng *et al.* [4].

The entropy of a probability distribution is a measure of the uncertainty of the distribution [4]. Let X be a random variable. $(p_1, p_2, \ldots, p_n)$ to the probability distribution of this variable, entropy H is,

$$\boldsymbol{H = -\sum_{i=1}^{n} p_i ln p_i}$$

In this study, Minimize Entropy Principle Approach is used. This approach is preferred method to use when the decision maker does not have any prior knowledge [4].

For partitioning the universal set to the subintervals, determining the threshold between classes of data is needed. In the method proposed by Cheng *et al.* [4], every threshold is obtained by minimizing the entropy. After obtaining the first threshold, the segmentation process is started. The primary threshold firstly divides the data into two classes. At the end of this repeated partitioning with threshold value calculations, the data set is partitioned into any number of fuzzy sets.

Assume that a threshold value is sought for a sample in the range between x_1 and x_2. An entropy equation with each value of x is written for the regions $[x_1, x]$ and $[x, x_2]$ and the first region is denoted by p, the second region is denoted by q. An entropy with each value of x in the region between x_1 and x_2 is expressed as

$$\boldsymbol{S(x) = p(x)S_p(x) + q(x)S_q(x)}$$

where

$$\boldsymbol{S_p(x) = -[p_1(x) ln p_1(x) + p_2(x) ln p_2(x)]}$$

$$\boldsymbol{S_q(x) = -[q_1(x) ln q_1(x) + q_2(x) ln q_2(x)]}$$

PART III: HYBRID METHODS

CHAPTER 11

An Architecture Selection Method Based on Tabu Search

Cagdas Hakan Aladag*

Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey

Abstract: In recent years, the most preferred forecasting method in time series forecasting has been artificial neural networks. In many applications, artificial neural networks have been successfully employed to obtain accurate forecasts in the literature. This approach has been preferred to conventional time series forecasting models because of its easy usage and providing accurate results. On the other hand, there are still some problems with using this method. Fining a good artificial neural network architecture which gives the most accurate forecasts is an important issue when the method is used for forecasting. Although, there are some systematical methods proposed to determine the best architecture, the most preferred method is trial and error method [1]. To solve the architecture selection problem, Aladag [2] also proposed an approach based on tabu search algorithm. In this study, the air pollution in Ankara time series is forecasted by utilizing artificial neural networks and the architecture selection algorithm proposed by Aladag [2] is used to determine the best architecture. The obtained results show that high accuracy level is reached when Aladag's [2] algorithm is employed.

Keywords: Architecture selection, Artificial neural networks, Forecasting, Tabu search, Time series.

1. INTRODUCTION

Forecasting in time series is an important issue on which many researchers from different disciplines have still working. In the literature, various approaches have been proposed to obtain more accurate forecasts. In recent years, artificial neural networks approach has the most preferred method for forecasting since the method has proved its success in many forecasting applications [3]. However, there are some drawbacks in using artificial neural networks to forecast time series. When this method is utilized for forecasting, selection of the components of the method is a vital issue to reach good forecasting results [4]. Determining the best artificial neural network model can be defined as selection of the components such as architecture structure, learning algorithm and activation function [5]. Selection of the best model, especially determining the best architecture and weights, remains a problem in artificial neural networks applications [6].

When the best model is being searched, an important decision is the selection of architecture structure consists of determining the numbers of neurons in the layers of a network. Various techniques have been proposed to determine the best artificial neural network architecture [2]. Some of them are constructive and pruning algorithm [7], polynomial time algorithm [8], network information criterion [9], iterative construction algorithm [10], a method based on Box-Jenkins analysis [11], a method based on information entropy [12], genetic algorithms [13], the principle component analysis [14], weighted information criterion [5], a deletion/substitution/addition algorithm [15], an architecture selection strategy for autoregressive seasonal time series [16], and design of experiments [17]. In spite of these proposed method, the most preferred method is trial and error [1, 2]. Trial and error approach, however, is not rigorous and offers no guarantee of arriving at a truly optimal structure [6].

Both theoretical and empirical findings in the literature show that combining different methods can be an efficient way to improve forecasts [4]. Therefore, in this study, a hybrid forecasting method combines artificial neural networks and tabu search algorithm is used to forecast the time series of air pollution of Ankara which is the capitol city of Turkey. The time series is forecasted by using artificial neural networks and tabu search algorithm proposed by Aladag [2] is used to determine the best architecture. For

***Address correspondence to Cagdas Hakan Aladag:** Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey; E-mail: chaladag@gmail.com

Cagdas Hakan Aladag and Erol Eğrioğlu (Eds)

comparison, the time series is also forecasted using trial and error method which is the most preferred technique in the literature. Then, the obtained results are examined and discussed.

The next section briefly gives artificial neural networks. Section 3 introduces the hybrid approach in which an architecture selection method based tabu search algorithm is employed. Section 4 presents the implementation. Finally, the last section concludes the chapter.

2. ARTIFICIAL NEURAL NETWORKS

Artificial neural networks have attracted more and more attention from both academic researcher and industrial practitioners in the recent years [19]. This approach has been widely used to model time series in various fields of applications [20]. Due to ability of modeling both linear and non-linear structures, artificial neural networks have been used as a good alternative method for both linear and non-linear time series forecasting. Zhang *et al.* [21] presented a review of the current status in applications of neural networks for forecasting.

In time series forecasting, artificial neural networks are mathematical models that imitate biological neural networks. Artificial neural networks consist of some elements. Determining the elements of the artificial neural networks issue that affect the forecasting performance of artificial neural networks should be considered carefully [16]. Elements of the artificial neural networks can be given as network architecture, learning algorithm and activation function [22].

For forecasting problems, feed forward neural networks have been widely preferred as architecture structure because of their simple usage and success. The structure of multilayer feed forward neural network is basically given in Fig. **1**. A multilayer feed forward neural network as illustrated in the figure consists 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 [6].

When artificial neural networks are used for forecasting, one critical decision is to determine the appropriate architecture, that is, the number of layers, number of nodes in each layer [2]. However, in the literature, there are not general rules for determining the best architecture [6]. Although there are some systematic approaches in the literature, the most preferred method to find the best architecture is trial and error method.

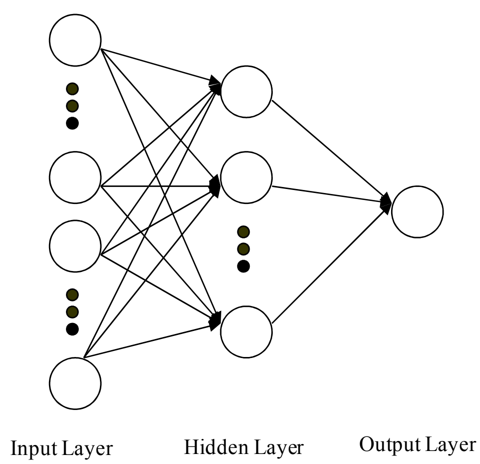


Figure 1: Multilayer feed forward artificial neural network with one output neuron.

Learning of an artificial neural network for a specific task is equivalent to finding the values of all weights such that the desired output is generated by the corresponding input. Various training algorithms have been used for the determination of the optimal weights values. One of the training algorithms is Levenberg Marquardt algorithm [23]. This algorithm is used as training algorithm in most optimization packages since it is an efficient nonlinear optimization method.

Another element of the artificial neural networks is the activation function. It determines the relationship between inputs and outputs of a network. In general, the activation function introduces a degree of the non-linearity that is valuable in most of ANN applications. The well known activation functions are logistic, hyperbolic tangent, sine (or cosine) and the linear functions. Among them, logistic activation function is the most popular one [21].

3. THE ARCHITECTURE SELECTION METHOD BASED ON TABU SEARCH

When artificial neural networks are used for forecasting, determining the best architecture which gives the best forecasting accuracy is an important issue. Each of possible architectures and their performance criterion values can be considered as solutions and objective function values, respectively. Thus, the problem of finding the best architecture can be interpreted as an optimization problem [6]. Aladag [2] solved this optimization problem by using tabu search algorithm. He improved a tabu search algorithm to determine the best architecture which gives the most accurate forecasts. In his algorithm, he focused on the feed forward neural networks and claimed that his algorithm can also be extended to other type of neural networks such as feed-back neural networks.

Aladag's [2] architecture selection method was proposed determine the best feed forward neural network architecture that includes one hidden layer and one neuron in the output layer which is shown in Fig. **1**. The tabu search algorithm proposed by him starts with an initial solution. This solution is generated using a starting pool strategy [2]. Why the starting pool strategy is used in Aladag's [2] algorithm can be explained as follows [6]:

The tabu search algorithm is a meta heuristic method and the starting conditions such as starting solution and initial state of the memories are important for reaching better global solutions. In order to make good choices in the beginning, a strategy called "starting pool" proposed by Aladag [2] is utilized. In this strategy, a specified number of solutions are examined and the information gathered from this examination is used to generate starting solution and the initial state of the memories. Therefore, the number of examined solutions in the starting pool is a parameter for the proposed algorithm. Detailed information about the starting pool strategy can be found in [2]. After the starting conditions are determined by employing the starting pool strategy, the proposed tabu search algorithm can be used to find a good architecture which produces accurate forecasts.

As mentioned before, the tabu search algorithm is used to determine a feed forward neural network architecture including one hidden layer and one output neuron given in Fig. **1**. Therefore, this algorithm will find the number of inputs and the number of neurons in the hidden layer. The elements of tabu search algorithm proposed by Aladag [2] are briefly given in below.

Let vector x represents a solution whose elements are the numbers of neurons in the input and the hidden layers. The number of output neuron is always equal to one so there is no need to include this number in the solution vector. Thus, each solution x has two elements. The solution space X is a set of solutions. The solution space includes all of the possible ANN architectures. The objective function is root mean square error (RMSE) which measures the difference between the actual and the forecasted values obtained from a solution x. RMSE value is computed over the test set. The objective function can be defined by

$$f(x) = \left(\frac{\sum_{i=1}^{T} (y_i - \hat{y}_i)^2}{T} \right)^{1/2} \quad \textbf{(1)}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value forecasted by using the solution x, and T represents the size of test data. Since the objective function is RMSE, the tabu search algorithm will try to minimize RMSE value. In other words, the algorithm will try to find an architecture that produces the most accurate forecasts. The detailed information about the all elements of this tabu search algorithm can be found in Aladag [2].

CHAPTER 12

A Hybrid Forecasting Approach Combines SARIMA and Fuzzy Time Series

Erol Eğrioğlu[1], Cagdas Hakan Aladag[2,*] and Ufuk Yolcu[1]

[1]Ondokuz Mayis University, Faculty of Arts and Science, Department of Statistics, 55139, Samsun, Turkey and [2]Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey

Abstract: Fuzzy time series, subjected to many scientific studies, have been used in forecasting in recent years. Due to their uncertainty, time series encountered in daily life should be perceived as fuzzy time series and analyzed by fuzzy time series methods. Instead of representing time series, which may have different values during the time they measured, by instantaneous value of each observation, representing a fuzzy set which may contain several values provides more information and thus more realistic analyses. In such a situation, forecasting problem of time series whose observations are fuzzy sets emerges. In the literature, there are several methods and algorithms proposed for forecasting these types of fuzzy time series. However, one can say that most of the observed fuzzy time series contain seasonal structures. From this stand point, using seasonal fuzzy time series forecasting methods in analyzing fuzzy time series containing seasonal relations would be effective in terms of both forecasting performance and explanation of the relation of the data contained in. This study aims to introduce a partial high order bivariate fuzzy time series forecasting method hybridized with Box-Jenkins method seasonal autoregressive integrated moving average model (SARIMA), one of the conventional time series analysis methods used in forecasting seasonal time series, and its advantages. For this purpose, two real data are analyzed using this seasonal fuzzy time series forecasting method and results are evaluated with certain fuzzy and conventional seasonal time series methods.

Keywords: Bivariate fuzzy time series, Forecasting, High order, Seasonal fuzzy time series.

1. INTRODUCTION

Fuzzy time series methods are effective methods in forecasting time series which contain uncertainty in their observations and which can be frequently encountered in real life. Not having constraints which are seen in other methods, such as number of observation and model assumption, have attracted attention on fuzzy time series methods.

Fuzzy set theory was first introduced by Zadeh [1]. Then, Song and Chissom proposed the fuzzy time series [2, 3] based on fuzzy set theory. Fuzzy time series analysis method consists of three steps as fuzzifying the observations, identifying fuzzified relations and defuzzifying. While Song and Chissom used matrix operations and compound processes [2-4], Chen proposed an easy method based on fuzzy logical relation tables [5]. Moreover, Huarng and Yu [6] proposed a model in which fuzzy relations are determined by artificial neural networks. In the literature, all the methods proposed by Sullivan and Woodal [7], Hwang, Chen and Lee [8], Chen and Hwang [9], Huarng and Yu [10] and Yu [11, 12] include first order fuzzy time series forecasting models. However, analyzing fuzzy time series with first order models, gives insufficient results as they contain high order relations. Therefore, Chen [13] proposed a new model to analyze fuzzy time series that contains high order relations. All lagged variables belong to fuzzy time series are present in the high order method proposed by Chen [13]. For instance, when a fourth order fuzzy time series forecasting model is created, the model includes all the lagged variables in the first, second, third and fourth orders. For seasonal time series, model order should be equal to the period. However, this adds pure lagged variables to the model and unnecessarily increases the number of input in the model. Additionally, Song [14] proposed a new model where lagged variable $F(t-m)$ is the input and $F(t)$ is the output to analyze a seasonal time series whose period is m. It is a fact that this method will fail to analyze seasonal time series containing more complex relations.

***Address correspondence to Cagdas Hakan Aladag:** Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey; E-mail: chaladag@gmail.com

Cagdas Hakan Aladag and Erol Eğrioğlu (Eds)

In order to eliminate this problem, Eğrioğlu *et al.* [15] proposed a model including seasonal autoregressive and seasonal moving average (MA) terms in analyzing fuzzy seasonal time series. Since both time series (X_t) and residuals series (a_t) are included in the model, the method proposed by Eğrioğlu *et al.* [15] has bivariate structure. Bivariate fuzzy time series method was first proposed by Huarng [16], Hsu, Tse, and Wu [17] and Yu and Huarng [18]. In his method, Huarng [16] used only first order lagged variables. On the other hand, Yu and Huarng used feed forward artificial neural networks in determining fuzzy relations.

In the method proposed by Eğrioğlu *et al.* [15], time series is primarily determined by Box-Jenkins method with SARIMA model. Then, partial high order, bivariate fuzzy time series using the inputs of determined SARIMA model is created. Determination of fuzzy relations is done by using artificial neural networks. The advantages of this model are as follows:

- It establishes a high order model in forecasting seasonal fuzzy time series.
- Model order is determined systematically.
- It is the first method in the literature including the term of MA.
- It improves the forecasting performance.

The method proposed by Eğrioğlu *et al.* [15] and emphasized in this study is applied to the time series of "the average of sulfur dioxide (SO_2) in Samsun province between January 2004 and December 2008 (SAMSO)" and "the number of foreign tourist arriving in Turkey between January 1998 and December 2008 (TOURISM DATA)" and obtained results are compared with those in the other methods. In the second section, SARIMA models and artificial neural networks are introduced and basic definitions of fuzzy time series are given. The method and its algorithm proposed by Eğrioğlu *et al.* [15] are introduced in the third section. Fourth section deals with the application. In the last section, results obtained from the study are discussed considering the results of other methods in the literature.

2. RELATED METHODS

In this section, SARIMA models and artificial neural networks, the basis of method intended to be introduced, and basic definitions of fuzz time series are given.

2.1. SARIMA Models

Let Z_t be a time series with μ mean. Then, the model is

$$\varphi(B)\Phi(B^s)(1-B)^d(1-B^s)^D(Z_t-\mu)=\theta(B)\Theta(B^s)a_t \tag{1}$$

where,

$$\varphi(B)=(1-\varphi_1 B-\cdots-\varphi_p B^p) \tag{2}$$

$$\theta(B)=(1+\theta_1 B+\cdots+\theta_q B^q) \tag{3}$$

$$\Phi(B)=(1-\Phi_1 B^s-\cdots-\Phi_P B^{sP}) \tag{4}$$

$$\Theta(B)=(1+\Theta_1 B^s+\cdots+\Theta_Q B^{sQ}) \tag{5}$$

Box-Jenkins first proposed a method for SARIMA, expressed by $\text{SARIMA(p,d,q)(P,D,Q)}_s$ [19]. This method is known as Box-Jenkins method and consists of determination, estimation, diagnostic control and forecasting stages. Detailed information for SARIMA models and Box-Jenkins method can be obtained from [19].

2.2. Artificial Neural Networks

Artificial neural networks are synthetic networks that imitate biological neural networks. There are great differences between artificial neural networks and biological neural networks in terms of their architectures and abilities [20]. Artificial neural networks constitute mathematical models and known as convergence of function [21].

More recently, artificial neural networks, as an alternative to other methods, began to be used in some stages such as determination of relation in the analysis of fuzzy time series. In the studies [6] and [18], feed forward artificial neural networks were used for the determination of fuzzy relations in fuzzy time series. The determination of fuzzy relations with artificial neural networks removes the complexity in defining fuzzy logical relation and group relation tables and improves forecasting accuracy.

There are three components directing function of artificial neural networks. These are architecture structure, learning algorithm and activation function.

Architecture structure: Multi layer feed forward artificial neural network architecture used in artificial neural networks literature, consists of three layers. These are, input layer, hidden layer (or layers) and output layer. Layers consist of units called neuron (node). Accurate determination of architecture is achieved by deciding the number of neuron in layers. In artificial neural networks, neurons are interconnected *via* weights. In feed forward networks, these connections are unidirectional. And, there is no connection between the units in the same layer.

While multi layer artificial neural network architectures have single input and output layer, they may contain more than one hidden layers. The determination of architecture means determination of neuron number in the layers. A multi layer feed forward artificial neural network architecture is presented in Fig. **1**.

In Fig. **1**, w_{ij} represents the weight between neuron i in input layer and neuron j in hidden layer and w_{jk}, represents the weight between neuron j in hidden layer and neuron k in hidden layer. As there is only one neuron in the figure, k index shows this single neuron. In another architecture, a direct connection can be established between neuron in input layer and neuron(s) in output layer.

Learning Algorithm: There are several learning algorithms used in the determination of weights in artificial neural networks. Two of the most widely used algorithms are Back Propagation and Levenberg Marquardt learning algorithms. Detailed information on Back Propagation and Levenberg Marquardt learning algorithm can be obtained from [22-28] respectively.

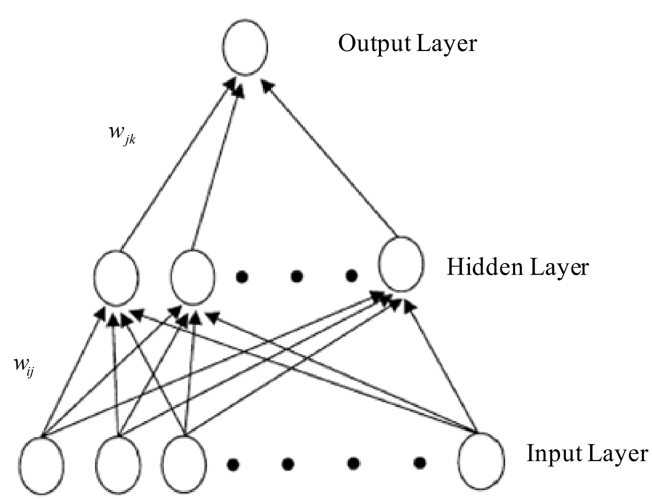


Figure 1: Multilayer feed forward artificial neural network with one output neuron.

CHAPTER 13

Forecasting Gold Prices Series in Turkey by the Forecast Combination

Cagdas Hakan Aladag[1,*], Erol Eğrioğlu[2] and Cem Kadilar[1]

[1]Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey and [2]Ondokuz Mayis University, Faculty of Arts and Science, Department of Statistics, 55139, Samsun, Turkey

Abstract: Forecast combination is a method used for obtaining more accurate forecasts. Forecast combination consists of the combination of forecasts obtained from different models with various methods. There are several types of forecast combination in the forecasting literature. In this study, various fuzzy time series approaches are applied to Turkey's daily highest gold prices series and forecasts obtained from these approaches are combined with variance covariance method (VCM), mean square forecast error method (MSFE) and artificial neural networks (ANN) approach. Results obtained from all of these methods are analyzed and the optimal forecast technique for Turkey's daily highest gold prices series is determined.

Keywords: Artificial neural networks, Forecasting, Forecast combination, Fuzzy time series.

1. INTRODUCTION

Numerous studies have been done and various approaches have been proposed in the literature to get better forecasts in time series. Among these methods, the most widely used two approaches in recent years are ANN and fuzzy time series methods which give very promising results for the forecasting problem [1, 2]. Another effective method used for the time series forecasting problem is the combination of forecasts of various methods. This method was first proposed in [3]. Then, [4-6] proposed forecast combination for more than two models. [7] compared three different forecast combinations for the combination of four time series models.

When forecast combination method is used, the forecasting value is obtained by a linear combination in which various models are combined with weights. The contribution of each model in forecasts of combination approach differs. The weights in forecast combination can be determined by both the particular assumptions and various optimization methods. The key point in forecast combination is the determination of optimal weights and combination function that give the best forecasts. Forecasts of combination function may be linear or curvilinear. [8] proposed a new forecast combination model in which forecasts obtained from fuzzy time series forecasting models are used as feed forward artificial neural network input and output of artificial neural network is obtained as combined forecasts. When this artificial neural network is optimized, optimal weights that provide the best curvilinear match up are determined. Therefore, when artificial neural networks method is used in forecast combination, both combination function and optimal weights can be determined without any problem.

In this study, various fuzzy time series approaches are applied to Turkey's daily highest gold prices series and forecasts obtained from the approaches are combined with variance covariance method (VCM), mean squared forecast error method (MSFE), and ANN approach. Results obtained from all of these methods are analyzed and the optimal forecasting technique for Turkey's daily highest gold prices series is developed.

In the second section of the study, forecast combination techniques existing in the literature are introduced. Third section deals with the general definition of fuzzy time series and a brief summary of fuzzy time series methods. In the fourth section, forecast combination techniques are applied to Turkey's daily highest gold prices series. Obtained results are discussed in the last section.

***Address correspondence to Cagdas Hakan Aladag:** Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey; E-mail: chaladag@gmail.com

2. FORECAST COMBINATION TECHNIQUES

In the literature, various forecast combinations have been developed. Three of the best known are; simple forecast combination, variance covariance combination and mean squared forecast error method. [8] also proposed a forecast combination based on artificial neural networks.

2.1. Simple Forecast Combination (SFC) Method

In the simple forecast combination method, combined forecasts are obtained by multiplying and adding forecasts obtained from two or more models with weights. Combine forecast is calculated by using the formula given below.

$$\boldsymbol{f_c = \sum_{i=1}^{n} w_i f_i} \tag{1}$$

where, f_i and n represent the forecast value obtained from the forecasting model i, and number of forecasting models, respectively. In this method, weights are not necessarily equal. If intervals are equal,

$$\boldsymbol{w_i = 1/n} \tag{2}$$

then the method is called simple mean combination. As a result, the combined forecast, no matter what the weights are, is a linear combination of model forecasts.

2.2. Variance Covariance Method (VCM)

Variance covariance method is a linear forecast combination approach in which the weights are determined based on performance of model forecasts. In this method, weights are calculated by using the formula given below.

$$\boldsymbol{w' = u'\Sigma^{-1}/u'\Sigma^{-1}\mathbf{u}} \tag{3}$$

where Σ represents covariance matrix, $u = (1,1,.,1)'$ and $\sum_{i=1}^{n} w_i = 1$. In some cases, weights may be negative and in this case, the use of variance covariance method is not suitable. When the weights are obtained according to equation (3), combined forecasts are calculated according to equation (1).

2.3. Mean Squared Forecast Error Method (MSFE)

Unlike variance-covariance method, in the MSFE method, combination weights are determined by increasing contribution of the last forecast. In this method, weights are calculated by using the formula as

$$\boldsymbol{w_i = \frac{1/\sum_{t=1}^{T}\beta^{T-t+1}e_{it}^2}{1/\sum_{i=1}^{n}\sum_{t=1}^{T}\beta^{T-t+1}e_{it}^2}} \tag{4}$$

where $0 < \beta < 1$ is the arbitrary discount factor and e_{it}^2 is the forecast error obtained from model i for observation t. When the weights are obtained according to equation (4), combined forecasts are calculated according to equation (1).

2.4. Forecast Combination Approach Based On Artificial Neural Network (ANNFC)

Although the first three combinations use different techniques in determining weights, combination functions of three methods are linear. Selection of linear combination function makes easy to obtain optimal weight values. However, curvilinear function is preferred in obtaining optimal combination. In [8], a forecast combination, in which forecasts obtained from various models are taken as input of artificial neural network and output of artificial neural network is obtained as combined forecast, was proposed. In ANN techniques proposed by [8], weights of combination function was a curvilinear function.

In ANNFC technique, the number of input units in artificial neural network used for the combination is equal to the number of forecasting models (n). Hidden layer unit number is taken as 1 to ensure the simple combination function structure. Output number is also taken as 1. Artificial neural network used for combination is given in Fig. **1**. Artificial neural network whose architecture, given in Fig. **1**, can be combined with forecasts obtained from three models.

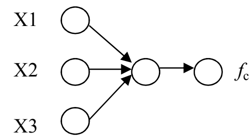


Figure 1: Feed forward ANN model for forecast combination.

Feed forward ANN model is used for forecast combination, logistic activation function, given in (5), is used for hidden layer units and a linear activation function, given in (6), is used for output layer.

$$\boldsymbol{f(net)} = \frac{\mathbf{1}}{\mathbf{1+exp\,(-\boldsymbol{net})}} \qquad \textbf{(5)}$$

$$\boldsymbol{f(net) = net} \qquad \textbf{(6)}$$

Mathematical representation of feed forward artificial neural network can be given by

$$\boldsymbol{f_c} = \left(\left(\sum_{\boldsymbol{i=1}}^{\mathbf{3}} \boldsymbol{w}(\mathbf{1},\boldsymbol{i})\frac{\mathbf{1}}{\mathbf{1+exp}(-\boldsymbol{X_i})}\right) + \boldsymbol{w}(\mathbf{3},\mathbf{1})\right)\boldsymbol{w}(\mathbf{2},\mathbf{1}) + \boldsymbol{w}(\mathbf{3},\mathbf{2}) \qquad \textbf{(7)}$$

where x_i is the forecast obtained from model i, $w(1,i)$ is connection weights between input layer and hidden layer, where i=1,2,3, $w(3,1)$ is a connection weight of input layer side to weight hidden layer, $w(2,1)$ is a connection weight between hidden layer and output layer, and $w(3,2)$ is a weight between hidden layer side and output layer. It is clear that function, given in (7), is in a curvilinear form. The weights obtained from this function are optimal weights for forecast combination and output of artificial neural network is combined forecasts.

3. FUZZY TIME SERIES METHODS

Fuzzy time series approaches depend on basic definitions defined as follows: Let $U = \{u_1, ..., u_b\}$ be the universe of discourse, where u_1, …, u_b are intervals. These intervals are obtained from partition of the universe of discourse including all values of time series. A fuzzy set A_i of U is defined as

$$A_i = f_{A_i}(u_1)/u_1 + \cdots + f_{A_i}(u_b)/u_b$$

where f_{A_i} is the membership function of the fuzzy set A_i, $f_{A_i} : U \to [0,1]$. $f_{A_i}(u_b)$ is the degree of belongingness of u_b to A_i.

Definition 1. Let $Y(t)$, $t = ..., 0,1,2,...$ be a subset of real numbers. After defining the universe of discourse, the new time series $F(t)$ consisting of A_i is called a fuzzy time series.

Definition 2. Fuzzy time series relationships assume that $F(t)$ is caused only by $F(t-1)$. Then, the fuzzy relationship is called the first order fuzzy time series. The first order fuzzy time series forecasting model can be written as follows:

$$F(t-1) \to F(t) \qquad \textbf{(8)}$$

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

CHAPTER 14

A Hybrid Forecasting Model Based on Multivariate Fuzzy Time Series and Artificial Neural Networks

Cagdas Hakan Aladag[1,*] and Erol Eğrioğlu[2]

[1]Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey and [2]Ondokuz Mayis University, Faculty of Arts and Science, Department of Statistics, 55139, Samsun, Turkey

Abstract: Fuzzy time series approaches have been recently used for forecasting in many studies [1]. These approaches can be categorized into two subclasses that are univariate and multivariate approaches. It is a fact that many factors can actually affect real time series data. Therefore, using a multivariate fuzzy time series forecasting model can be more reasonable in order to get more accurate forecasts. The most preferred method is using tables of fuzzy relations for determining fuzzy relations in multivariate fuzzy time series approaches in the literature. However, employing this method is a computationally though task. In this study, we propose a new method based on utilizing artificial neural networks in determining fuzzy logic relations and using the formula defined by Jilani and Burney [2] in calculating defuzzyfied forecasts. Hence, it is aimed to produce more accurate forecasts and avoid intense computations. 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 [2] and Lee *et al.* [3].

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

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 produce unsatisfactory forecasts for such time series. Song and Chissom [4-6] first introduced the definition of a fuzzy time series based on the concept of the fuzzy set theory proposed by Zadeh [7]. The Song and Chissom's model [4] is a fuzzy time series-forecasting model called as a one factor first order model. Chen [8] and Aladag *et al.* [1] present approaches based on a one factor high order fuzzy time series forecasting model. Using a multivariate fuzzy time series forecasting model can give more accurate forecasts since the real time series data can be affected by many factors. In order to forecast such time series, two-factor fuzzy time series model was used by Yu and Huarng [9], and Lee *et al.* [3] in the literature. Then, Jilani and Burney [2] analyze k-factor ($k \geq 1$) and n^{th} order fuzzy time series forecasting models.

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 [9] presented an algorithm that analyzes a first order bivariate fuzzy time series-forecasting model. Then, Yu and Huarng [9] 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 [2] 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 [2] gives better-forecast values than those generated by the method introduced by Lee *et al.* [3]. However, the determination of fuzzy relationships in the algorithm proposed by Jilani and Burney [2] depends on fuzzy logic relation tables which require too many complicated calculations and so too much time.

***Address correspondence to Cagdas Hakan Aladag:** Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey; E-mail: chaladag@gmail.com

In this chapter, we proposed a new method to analyze k-factor and n^{th} order fuzzy time series-forecasting model. In the proposed method, feed forward artificial neural networks (ANN) are utilized to determine fuzzy logic relations. In addition, we use the formula defined by Jilani and Burney [2] in the defuzzyfication step. 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 [2] and Lee *et al.* [3]. We would like to note that the method proposed in this chapter is different from the method proposed by Eğrioğlu *et al.* [12] although some parts of this chapter were taken from the study of Eğrioğlu *et al.* [12].

Section 2 gives the fundamental fuzzy time series definition. The brief information related to ANN is given in Section 3. 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. Section 6 is the last part that gives comparison results and discussion.

2. FUZZY TIME SERIES

The definition of fuzzy time series was first introduced by Song and Chissom [4,5]. The definitions related to fuzzy time series are given as follows:

Let U be the universe of discourse, where $U = \{u_1, u_2, ..., u_b\}$. 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 $1 \le a \le b$.

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 $f_1(t), f_2(t), ...$ 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 $F(t-1), F(t-2)$,…, and $F(t-n)$, then this fuzzy logical relationship is represented by

$$F(t-n), ..., F(t-2), F(t-1) \to F(t) \quad \textbf{(1)}$$

and it is called the n^{th} 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:

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

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

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

and it is called the two-factors n^{th} 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=…0,1,2,…).

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

$(F(t-1), G_1(t-1)G_2(t-1)\ldots G_{k-1}(t-1)), \ldots, \ (F(t-n), G_1(t-n)G_2(t-n)\ldots G_{k-1}(t-n))$ then this fuzzy logical relationship is represented by

$$(F(t-1), G_1(t-1)G_2(t-1)\ldots G_{k-1}(t-1)), \ldots,$$

$$(F(t-n), G_1(t-n)G_2(t-n)\ldots G_{k-1}(t-n)) \rightarrow F(t) \tag{4}$$

and it is called the k-factors n^{th} 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 [10]. ANN compose of a mathematical model [11]. The fundamental elements of ANN can be given as follows:

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.

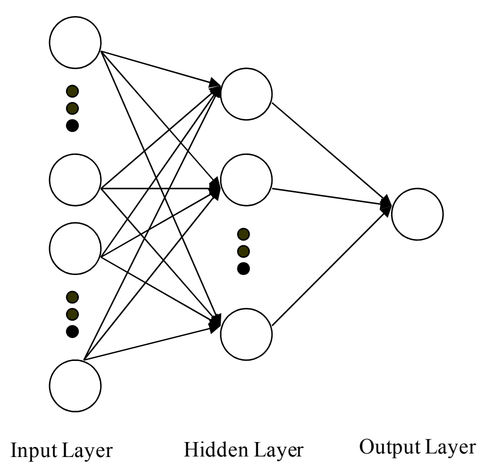


Figure 1: Multilayer feed forward ANN with one output neuron.

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

Subject Index

Cagdas Hakan Aladag and Erol Eğrioğlu (Eds)

Author Index