Systems Research and Behavioral Science
Syst. Res. **30**, 244–259 (2013)
Published online 1 May 2013 in Wiley Online Library
(wileyonlinelibrary.com) **DOI**: 10.1002/sres.2179

■ *Research Paper*

An ARIMA-ANN Hybrid Model for Time Series Forecasting

Li Wang[1*], Haofei Zou[2], Jia Su[3], Ling Li[4] and Sohail Chaudhry[5]

[1] *School of Economics and Management, Beihang University, Beijing, China*
[2] *China International Engineering Consulting Corporation, Beijing, China*
[3] *Telecommunication Planning Research Institute of MII, Beijing, China*
[4] *Old Dominion University, Norfolk, VA, USA*
[5] *Villanova University, Villanova, PA, USA*

Autoregressive integrated moving average (ARIMA) model has been successfully applied as a popular linear model for economic time series forecasting. In addition, during the recent years, artificial neural networks (ANNs) have been used to capture the complex economic relationships with a variety of patterns as they serve as a powerful and flexible computational tool. However, most of these studies have been characterized by mixed results in terms of the effectiveness of the ANN**s** model compared with the ARIMA model. In this paper, we propose a hybrid model, which is distinctive in integrating the advantages of ARIMA and ANNs in modeling the linear and nonlinear behaviors in the data set. The hybrid model was tested on three sets of actual data, namely, the Wolf's sunspot data, the Canadian lynx data and the IBM stock price data. Our computational experience indicates the effectiveness of the new combinatorial model in obtaining more accurate forecasting as compared to existing models. Copyright © 2013 John Wiley & Sons, Ltd.

Keywords ARIMA; Box–Jenkins methodology; artificial neural networks; time series forecasting; combined forecast; systems science

INTRODUCTION

The autoregressive integrated moving average (ARIMA) models have been successfully applied in numerous situations and prominently in economic time series forecasting. Moreover, ARIMA serves as a promising tool for modeling the empirical dependencies between successive times and failures (Walls and Bendell, 1987) with satisfactory performance (Ho and Xie, 1998). However, in the existing studies, it is generally assumed that there exists a linear correlation structure among the time series values. Whereas in complex real-world problems, this linear relationship is not factual because such an association is typically nonlinear. Hence, this nonlinear relationship needs to be ascertained because the approximation of linear models to complex real-world problems has not always provided satisfactory results (Xu 2000).

*Correspondence to: Li Wang, School of Economics and Management, Beihang University, Beijing 100191, China.
E-mail: wl2000bh@yahoo.com.cn

According to Chaudhry et al. (2000), knowledge-based decision support has been provided by evolving applied artificial intelligence systems including artificial neural networks (ANNs). During the past decades, ANN models have been used quite frequently (Duan and Xu 2012). This is mainly caused by the fact that ANNs can function in simple pattern recognition and can be applied to a wide range of application areas (Trippi and Turban, 1996). Also, it is known that ANNs mapping process can cover problems of a greater range of complexity as well as they are superior to other approaches with their powerful, flexible and easy operation (Patterson, 1996). Zhang et al. (1998) presented a review in this area. Wong and Selvi (1998) summarized studies in which ANNs were used and applied to financial applications from 1990 to 1996. One of the major advantages of neural networks lies in their flexibility of modeling nonlinear situations. Besides, the model is constructed adaptively based on the features manifested in the data. This data-driven approach is suitable for many empirical data sets where no theoretical guidance is available to suggest an appropriate data-generating process. Many empirical studies, including several large-scale forecasting experiments, have illustrated the effectiveness of ANNs in time series forecasting. Conceptually, ARIMA and ANNs are nonparametric techniques and are very similar in attempting to make appropriate internal representations of the time series data. Neural networks have been compared with traditional time series techniques in several studies.

In literature, there are studies where ANNs have been applied to time series forecasting; however, these studies have depicted mixed results on the effectiveness of the use of ANNs methodology when combined with the ARIMA models (Kohzadi et al., 1996; ElKateb et al., 1998; Ho et al., 2002). Hybrid models have been developed by combining several different models, and they have been shown to improve the forecasting accuracy as portrayed by the early work of Reid (1968) and Bates and Granger (1969). For example, Makridakis et al. (1982) showed that the forecasting performance was improved when more than one forecasting model was combined for the well-known M-competition problem. Also, Clemen (1989) provided a comprehensive review and annotated bibliography in this area. The basic idea of model combination in forecasting is to integrate each model's unique feature into an organic whole for capturing different patterns in the data. Both theoretical and empirical findings suggest that combining different methods is an effective and efficient way to improve forecasting performance (Palm and Zellner, 1992; Pelikan et al., 1992; Ginzburg and Horn, 1993; Luxhoj et al., 1996). In addition, a number of combining schemes have been proposed in forecasting studies where ANNs were used. Wedding and Cios (1996) described a combinatorial methodology based on radial basis function networks and the Box–Jenkins models. Sfetsos and Siriopoulos (2004) classified various methodologies for improving modeling capabilities into two broad categories, namely, the hybrid schemes and combinatorial synthesis of multiple models. Sfetsos and Siriopoulos (2004) examined the efficiency of the combined model in accordance with clustering algorithms and neural network in time series forecasting. Tsaih et al. (1998) presented a hybrid artificial intelligence model integrating the rule-based system technique and the neural networks technique in which they accurately predicted the direction of daily price changes in S&P 500 stock index futures. Kodogiannis and Lolis (2002) recommended the improved neural network and fuzzy model used for exchange rate prediction. Wang and Leu (1996) put forward a hybrid model to forecast the mid-term price trend of the Taiwan stock exchange weighted stock index, which was a recurrent neural network trained by features extracted from ARIMA analyses. The results achieved showed that the neural networks trained by differentiated data produced better predictions than those trained by the raw data. Tseng et al. (2002) proposed a hybrid forecasting model (SARIMABP) combining the seasonal time series ARIMA (SARIMA) and the neural network back propagation (BP) model to predict seasonal time series data. Their results demonstrated that the hybrid model produced better forecasts than either the ARIMA model or the neural network. All of the values of the measures of accuracies, such as MSE (Mean Square Error), MAE (Mean Absolute Error) and MAPE (Mean Absolute Percentage Error), were the lowest for the SARIMABP model.

Syst. Res. **30,** 244–259 (2013)
DOI: 10.1002/sres

Zhang (2003) recommended a hybrid model combining ARIMA and ANNs. These empirical results adequately illustrated the hybrid model's advantage in outperforming each component model in isolation and improving the forecasting accuracy in an effective way. The methodology for combining ANNs and other general models has been a well-studied topic in the time series forecasting area.

Autoregressive integrated moving average and ANNs possess differentiated characteristics; the former is suitable for linear prediction, whereas the latter is appropriate for nonlinear prediction. Because of the complexity associated with the historical data as well as the randomness resulting from many uncertain factors, the observed data are usually composed of linear and nonlinear components. Hence, a combined model based on ARIMA and ANNs would be a better option to improve the forecasting accuracy in which the linear part of the historical data will be handled by ARIMA whereas the nonlinear part to be processed by the ANNs model. However, the critical issue is how to combine the ARIMA and ANNs models. In this paper, a hybrid approach for time series forecasting, which combines ARIMA and ANNs models, is proposed. In general, there are four components in a time series, namely, Trend, T; Cyclical variation, C; Seasonal variation, S; and Irregular variation. Commonly, the analysis of time series can be accomplished by one of the two decomposition models. These models are the additive model: TS = T + C + S + I and the multiplicative model: TS = T × C × S × I. Furthermore, the complexity in the time series is composed of a linear component (L) and a nonlinear component (NL). Analogously, we may assume two models to analyse such a time series, an additive model (L + N) and a multiplicative model (L × N). In Zhang (2003), the additive model was used. In this paper, we propose to use the multiplicative model and explore the difference between the two models during the computational experience.

The remainder of this paper is organized as follows. The ARIMA and the ANNs with back-propagation models are described in Section 2. In Section 3, the details associated with the hybrid model are presented. The computational experience with four data sets using the proposed hybrid model is discussed in Section 4. In addition, the empirical results based on the four models applied to test data sets are also discussed. The conclusions are provided in Section 5.

METHODOLOGY

For the sake of completeness, ARIMA and ANNs models are explicated in this section as preconditions for constructing the hybrid model, with a focus on the basic principles and modeling processes of the two models.

Box-Jenkins ARIMA Model

In an autoregressive integrated moving average model, the future value of a variable is assumed to be a linear function of several past observations and random errors. Accordingly, a nonseasonal time series can be generally modeled as a combination of past values and errors, which can be represented as ARIMA (p,d,q) or expressed as follows:

$$Xt = \theta 0 + \phi 1Xt - 1 + \phi 2Xt - 2 + \Lambda + \phi pXt - p \\ +et - \theta 1et - 1 - \theta 2et - 2 - \Lambda - \theta qet - q \tag{1}$$

where Xt and et represent the actual value and random error at time period t, respectively; θi ($i = 1, 2, \ldots, p$) and θj ($j = 1, 2, \ldots, q$) are model parameters; and p and q are integers and are often referred to as orders of autoregressive and moving average polynomials. Random errors, et, are assumed to be independently and identically distributed with a mean zero and a constant variance, σ^2. Equation (1) entails several important special cases of the ARIMA family of models. If $q = 0$, then Equation (1) becomes an autoregressive (AR) model of order p. When $p = 0$, the model reduces to a moving average (MA) model of order q. One central task of building ARIMA model is to determine the appropriate model order (p, q). Similarly, a seasonal model can be represented as ARIMA (p, d, q)(P, D, Q).

Syst. Res. **30**, 244–259 (2013)
DOI: 10.1002/sres

Box and Jenkins (1970) developed a practical approach to building ARIMA models, which exercised a fundamental impact on time series analysis and forecasting applications. For a long period, the Box-Jenkins ARIMA linear models have remained dominant in many areas of time series forecasting.

Artificial Neural Network Approach for Time Series Modeling

The greatest advantage of an artificial neural network is its ability to model complex nonlinear relationship without priori assumptions of the nature of the relationship (Li and Li, 1999; Li and Xu 2000; Yang et al. 2001; Zhou and Xu 2001; Wang et al. 2010; Li et al., 2012; Yin et al. 2012; Zhou et al., 2012). Figure 1 shows a popular ANN model known as the feed-forward multi-layer network. This ANN consists of an input layer, a hidden layer, each with nonlinear sigmoid function, and an output layer using linear transfer function.

The ANN model performs a nonlinear functional mapping from the past observations ($Xt-1$, $Xt-2, \cdots, Xt-p$) to the future value Xt, i.e.,

$$Xt = f(Xt-1, Xt-2, \Lambda, Xt-p, w) + et \quad (2)$$

where w is a vector of all parameters and f is a function determined by the network structure and connection weights. Thus, the ANN is equivalent to a nonlinear autoregressive model.

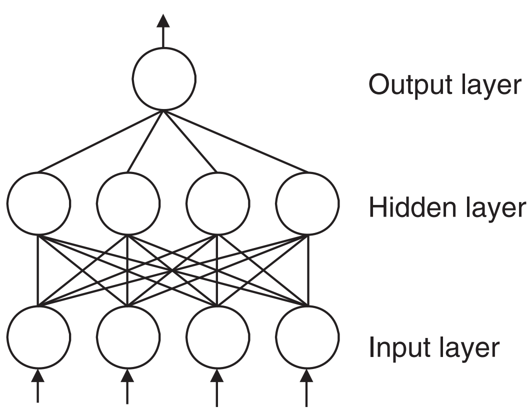

Figure 1 A typical feed-forward neural network

Training a network is an essential part for the successful operation of the neural networks. Among the several learning algorithms available, the back-propagation (BP) has been shown to be the most popular and widely implemented in all neural networks paradigms.

The important task of ANN modeling for a time series is to choose an appropriate number of hidden nodes q, as well as to select the dimension of input vector (the lagged observations), p. However, it is difficult to determine q and p in practice as there are no theoretical developments that can guide the selection process. Hence, in practice, experiments are often conducted to select the appropriate values p and q.

The Hybrid Model Based on Autoregressive Integrated Moving Average and Artificial Neural Network

As stated previously, a complex time series needs to be divided into a linear component and a nonlinear component by the decomposition techniques (e.g., Fourier decomposition and wavelet decomposition). Hence, econometric approaches can be deployed to model the linear components of the time series by the decomposition process in the form of ARIMA-based econometric linear forecasting model. As the ARIMA model fails to capture the nonlinear components of the time series, it is imperative to find a nonlinear modeling technique to account for the nonlinear components of time series. It has been shown that the ANNs are distinctively capable of modeling rather complex phenomena as they have many interactive nonlinear neurons in multiple layers, which can be utilized to address the nonlinear component of the time series data.

Artificial neural network models with hidden layers can capture nonlinear patterns in time series because they can be characterized as a class of approximate general functions, which are capable of modeling nonlinearity (Tang and Fishwick, 1993). Thus, it is prudent to combine ANNs and ARIMA in time series forecasting to deal with all of the heterogeneous components of the underlying patterns. Also, it may be judicious to consider a time series, which is composed of a linear

Syst. Res. **30**, 244–259 (2013)
DOI: 10.1002/sres

autocorrelation structure and a nonlinear component. Figure 2 shows a schematic diagram of integrating ANNs with partially known nonlinear relationships. Again, we can consider two models to analyse such time series, additive model (L + N) and multiplicative model (L × N). Figure 3 illustrates the combined models of ANNs integrating the two cases. The mathematical expressions for these two cases are given by Equations (3) and (4) below:

$$\text{Additive Model} : y_t = L_t + N_t \tag{3}$$

$$\text{Multiplicative Model} : y_t = L_t * N_t \tag{4}$$

where L_t represents the linear component and N_t the nonlinear component. These two components have to be derived from the data using these equations. In contrast, Zhang (2003) proposed a hybrid model of ARIMA and ANN for the additive model.

During the first phase of the proposed hybrid approach, an ARIMA model is applied to the linear component of time series, which is assumed to be $\{y_t, t = 1, 2, \cdots\}$, and a series of forecasts are generated, namely $\left\{\hat{L}_t\right\}$. By comparing the actual value y_t with the forecast value $\left\{\hat{L}_t\right\}$ of the linear component, we can obtain a series of nonlinear components, which are defined to be $\{e_t\}$.

According to 'multiplicative model', we have

$$e_t = y_t / \hat{L}_t \tag{5}$$

By contrast, according to 'additive model', we have

$$e_t = y_t - \hat{L}_t \tag{6}$$

Thus, a nonlinear time series is obtained. The second phase is concerned with modeling the nonlinear component of the specified time series in an ANN model.

The trained ANNs model is responsible for making a series of forecasts of nonlinear components, denoted by $\left\{\hat{N}_t\right\}$, which are based on the previously deduced nonlinear time series $\{e_t\}$ values as the inputs. That is, the ANNs time-series forecasting model is a nonlinear mapping function, as shown below:

$$e_t = f(e_{t-1}, e_{t-2}, \Lambda, e_{t-n}) + \varepsilon_t \tag{7}$$

The second phase can be seen as a process of error correction of time series prediction in the ANNs model on the basis of ARIMA model. Thus, for the multiplicative model, the combined forecast can be obtained from Equation (8) below:

$$y_t = \hat{L}_t * \hat{N}_t \tag{8}$$

However, for the 'additive model', the combined forecast is given by Equation (9) below:

$$y_t = \hat{L}_t + \hat{N}_t \tag{9}$$

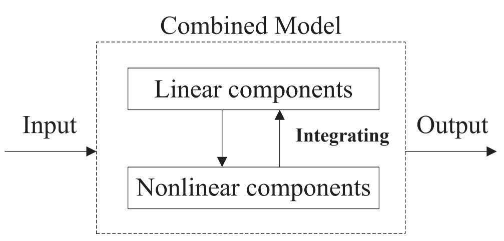

Figure 2 Schematic diagram of combined models

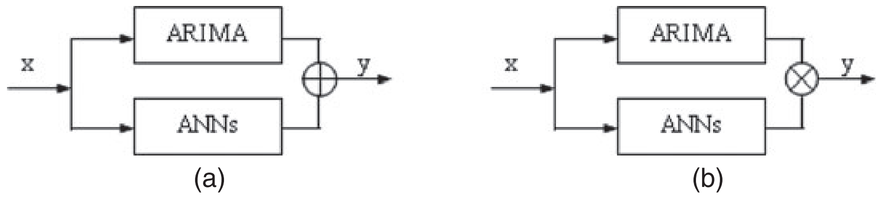

Figure 3 Combined models of artificial neural networks with two cases of integrating: (a) additive model, (b) multiplicative model

DOI: 10.1002/sres

To summarize, the proposed hybrid modeling is conducted in two phases. Initially, an ARIMA model is identified, and the corresponding parameters are evaluated, that is, an ARIMA model is constructed. As a result, the nonlinear components are computed based on this model. In the second phase, a neural network model deals with the nonlinear components.

DATA DESCRIPTION AND FORECAST EVALUATION CRITERIA

In this section, we describe the various data sets, which are used to evaluate the proposed hybrid procedure as outlined in the previous section.

Datasets

We use three well-known data sets to evaluate the efficacy of the proposed hybrid procedure. The three data sets are the Wolf's sunspot data, the Canadian lynx data and the IBM stock price data. These three time series emanate from different areas and possess different statistical characteristics. They have been widely studied in statistical research as well as the neural network literatures. Both linear and nonlinear models have been applied to these data sets.

The sunspot data contain the annual number of sunspots with 288 observations. The study of sunspot activity is of practical importance for geophysicists, environmental scientists and climatologists. The data series is regarded as nonlinear and non-Gaussian, typically useful for evaluating the effectiveness of nonlinear models. The behavior of this time series is depicted in Figure 4, which also suggests that it is characterized with a cyclical pattern with a mean cycle of about 11 years. The sunspot data have been extensively studied with a vast variety of linear and nonlinear time series models including ARIMA and ANNs.

The lynx time series contains the number of lynx trapped per year in the Mackenzie River district of Northern Canada. The lynx data are portrayed in Figure 5, showing a periodicity of approximately 10 years. The data set contains 114 observations. It has also been extensively analysed in the time series literature with a focus on the nonlinear modeling.

The third data set is for the IBM stock price. It is generally known that it is a difficult task to predict stock price or index (Zhu et al. 2008; Jiang et al. 2009; Chen and Fang 2012). Various linear and nonlinear theoretical models have been

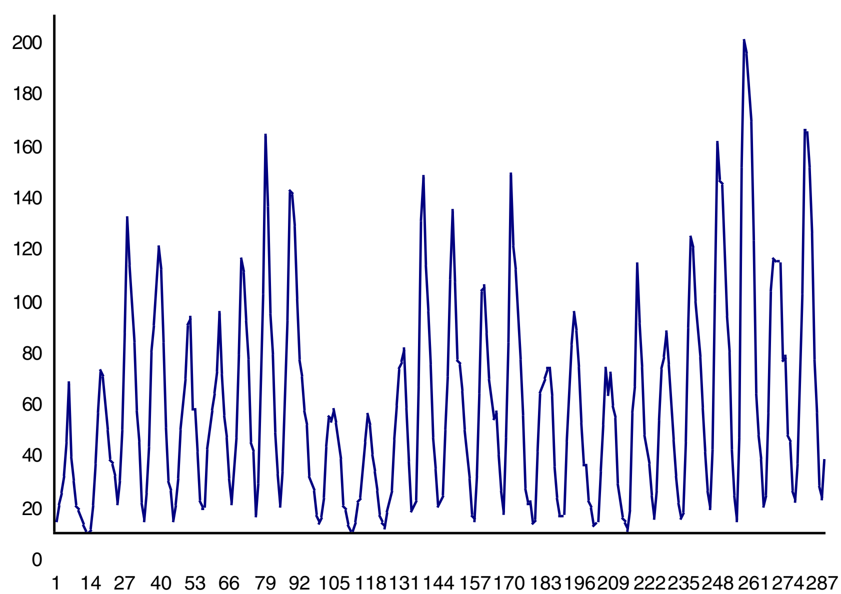

Figure 4 Sunspot series

Syst. Res. **30**, 244–259 (2013)
DOI: 10.1002/sres

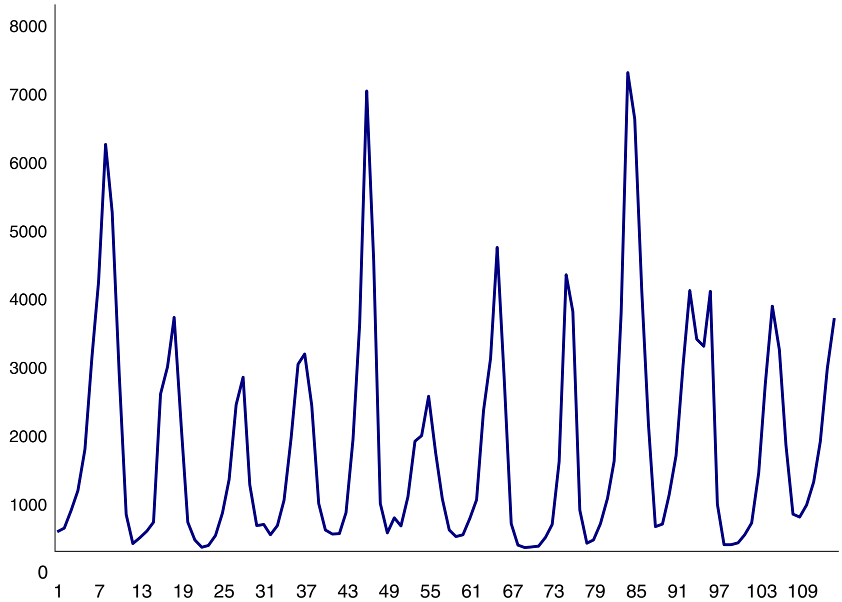

Figure 5 Canadian lynx data series

developed but few are more successful in forecasting. Recent applications of neural networks in this area have yielded mixed results. The data illustrated in this paper are obtained from the daily observations displaying 369 data points in the time series. The time series plot is shown in Figure 6, which illustrates numerous turning points in the time series.

To assess the forecasting performance of different models, each data set is divided into two subsets, training set and testing set. The training set is used exclusively for model development and then, the testing set is used to evaluate the proposed models. Although there is no consensus on how to split the data for neural network applications, the general practice is to allocate more

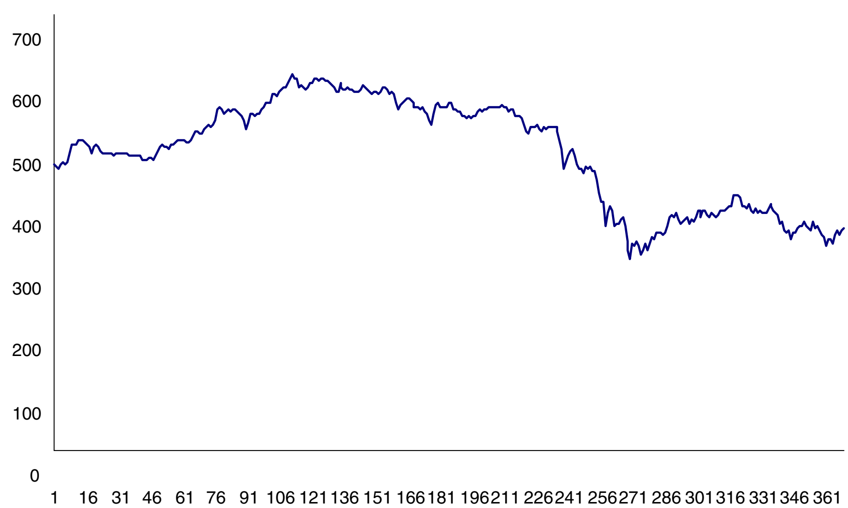

Figure 6 IBM stock price series

Syst. Res. 30, 244–259 (2013)
DOI: 10.1002/sres

data for model building and selection. Most studies in the literature use convenient ratios of splitting for in- and out-of-samples such as 70%:30%, 80%:20%, or 90%:10%. In this paper, the second approach is utilized to split the data. In addition, during the computational analysis, we make some minor adjustments to attain more accurate forecasting results. The data compositions for the three data sets are presented in Table 1.

Forecast Evaluation Criteria

In this study, we utilized two types of evaluation criteria, namely, the quantitative evaluation and turning point evaluation. For the quantitative measures, we employed the MSE, MAPE and others as the measures of accuracy.

To validate the ANN model, we selected 10-fold cross-validation and the independent test data. In the total training set, 20% of the data were chosen as the test dataset. The 10-fold cross-validation test was conducted with the remaining 80% of the trained dataset. For 10-fold cross-validation test, the training data were divided into 10 equal parts. In these 10 sets, 9 sets were used for training, and the last set was used for testing. This procedure was performed repeatedly for 10 times for all 10 sets. Based on this computational process, the average error across all 10 trials was computed.

FORECASTING PROCEDURE

In this section, we briefly describe the software package used as well as explain the forecasting procedure used to predict using the three data sets.

Table 1 Sample compositions in three data sets

Time series	Sample size	Training set (size)	Test set (size)
Sunspot	288	221	67
Lynx	114	100	14
IBM stock price	369	299	70

The Selection of the Autoregressive Integrated Moving Average Model

The time series data analyses were conducted in the following manner to determine a best possible model. First, the data were plotted, and an appropriate way to transform it was selected. Certain information was obtained from the graph of the time series, such as trend, cyclical variation, seasonal variation, outliers and other variations. Second, the autocorrelation function (ACF) and the partial autocorrelation function (PACF) were calculated based on the original data. Then, the model values of p, q and d can be deduced.

The Wolf's Sunspot Data Set

For the Wolf's sunspot data set, ACF followed an attenuating sine wave pattern that reflects the random periodicity of the data and possible indication of nonseasonal and/or seasonal AR terms in the model. The performance of the PACF also signified the need for some type of AR model. In addition to possessing significant values at lags 1 and 2, the PACF also has rather large values at lags 6–9. Thus, AR(2) and AR(9) are considered to model the sunspot data. The AIC was employed to verify which model was better. The AIC for the AR(2) and AR(9) was 1332.03 and 1288.63. Therefore, AR(9) was selected. The plot of the PACF showed the residuals were uncorrelated, and all estimated values of the ACF and PACF fell within the 5% significance interval. Thus, ARIMA(9,0,0) was considered to be a fairly suitable model for sunspot data.

The Canadian Lynx Data

The lynx time series does not depict a stationary pattern based on the graph as depicted in Figure 5. This behavior is confirmed and discussed in other studies (Campbell and Walker, 1977; Rao and Sabr, 1984), and the natural logarithms were taken to transform the lynx data. The ACF of the logarithmic lynx data followed an attenuating sine wave pattern that reflected the random periodicity of the data and the possible indication for the need of nonseasonal and/or seasonal AR terms in the model. The performance of the PACF also signified

Syst. Res. 30, 244–259 (2013)
DOI: 10.1002/sres

the need for some type of AR model. In addition to possessing significant values at lags 1, the PACF also had rather large values at lags 2 and 11, and other PACF fell within the 5% significance interval. In line with R package, AR(11) was selected. The plot of the PACF showed the residuals were uncorrelated. All of the estimated values of the ACF and PACF fell within the 5% significance interval. ARIMA(11,0,0) was a fairly suitable model for lynx data.

The IBM Stock Price Data

The ACF of IBM stock price demonstrated that the data were composed of seasonal patterns caused by the pronounced peaks of the ACF at lags. The ACF attenuates very slowly, which indicated the need for seasonal and/or nonseasonal differencing. It was found that the first-difference series became stationary and the seasonal differencing removed the seasonal wave pattern in the ACF. In addition to possessing significant values at lags 1, the ACF and PACF both had large values at lags 6. ARIMA(1,1,2) was selected, and the plot of the PACF showed the residuals were uncorrelated. All of the estimated values of the ACF and PACF fell within the 5% significance interval. ARIMA(1,1,2) was a fairly suitable model for IBM stock price data.

Fitting Neural Network Models to the Data

Criteria of Selection of the Models

Data normalization often precedes the training process. In this paper, the following formula (linear transformation to [0,1]) was used:

$$x^*_t = (x_t - x\min)/(x\max - x\min) \quad (10)$$

As far as time series forecasting is concerned, the number of input nodes corresponds to the number of lagged observations, which were used to discover the underlying pattern in a time series and forecast values. The most common way of determining the number of hidden nodes is by performing experiments or conducting trial-and-error approach. Therefore, the experimental design determines the number of input and hidden nodes by selecting the 'best' models according to delineated criteria, which include the root mean squared error (RMSE) and MAE. In addition, we use two types of information criteria, namely, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The following are the general formulas for AIC and BIC:

$$AIC = \log(ASSE) + \frac{2p}{T} \quad (11)$$

$$BIC = \log(ASSE) + \frac{p\log(T)}{T} \quad (12)$$

where $ASSE = \sum_{t=1}^{T}(yt - \hat{y}t)^2/T$. Also, p represents the number of parameters in the model, and T represents the number of observations. The first item in Equations (11) and (12) measures the applicability of the model to specified data, whereas the second item sets a penalty in the case of overfitting. The total number of parameters (p) in a neural network with m inputs and n hidden nodes is $p = m(n+2)+1$. Too few or too many input nodes can affect either the learning or predictive nature of the network (Zhang, 1998). It is the same as the selection of the number of hidden nodes. Therefore, we set an upper limit of lag period 6, and the number of hidden nodes varied from 1 to 6. A total of 36 different models for the data set were taken into consideration in building the ANN model.

Neural Network Design and Architecture Selection

In our study, neural network models were built with three-layered back-propagation neural networks, with a particular focus on ANNs with a single hidden layer (which is the usual case). Thus, the choice of architecture was primarily concerned with the number of neurons in the hidden layers. The transfer functions of the hidden layer and of the output layer were tan-Sigmoid and linear, respectively. The ANN model was built with Neural Network Toolbox of MatLab. All ANN models were trained with the modifications of conventional back propagation algorithm operating on the training set as well as the test set.

Syst. Res. 30, 244–259 (2013)
DOI: 10.1002/sres

The convergence criterion used for training was MSE of less than or equal to 0.001 or a maximum of 1000 iterations. The averages of MAE, MSE, MAPE, AIC and BIC were computed based on 10 replications of the experiment.

In terms of the criteria for selecting the best model mentioned in the previous section, the neural network structures were delimited as $4 \times 4 \times 1$, $7 \times 5 \times 1$ and $6 \times 5 \times 1$, which were selected to model the sunspot data, lynx data, and IBM stock price data, respectively.

EXPERIMENTAL RESULTS

In Table 2, we present the forecasting results for the sunspot data. A subset of the autoregressive model of order 9 was found to be the most parsimonious among all ARIMA models that were evaluated based on the residual analysis (Rao and Sabr, 1984; Hipel and McLeod, 1994). The ANN-based model was a $4 \times 4 \times 1$ network used to forecast for 35 and 67 periods. For the forecasts obtained, the results showed that the measures of accuracy associated with the ANNs model were lower than those which were related to the ARIMA model. In fact, according to the results displayed in Table 2, the multiplicative model performed considerably better than the other three models in improving the forecasts based on the various measures of accuracy. It should be noted that the multiplicative model did not provide improved forecasts for every data point; however, in general, the model provided significantly more accurate forecasts than the other models on this data set. The improvement in forecasting ability over the first 35 periods associated with the multiplicative model based on MAD over the ARIMA and ANN, and the additive model was 55.43%, 44.72% and 50.02%, respectively. Likewise, the multiplicative model provided better predictions over the other models based on MSE and MAPE. For the longer term forecasts, that is, 67 periods, the multiplicative method generated superior results as indicated by the improvements in MSE by 50.47%, 46.68% and 46.76% and in MAD by 40.28%, 36.13% and 35.60% over the ARIMA, ANN and additive model, respectively. Also, it should be noted that

Table 2 Forecasts for sunspot data

	Measures of fit (training set)*					Forecast accuracy (testing set)					
						35 points ahead			67 points ahead		
	MD[a]	SD[b]	MAD	MSE	MAPE (%)	MAD	MSE	MAPE (%)	MAD	MSE	MAPE (%)
ARIMA	39.4915	47.7292	11.3387	225.0567	53.4703	12.7421	257.8309	50.6676	14.4167	359.9119	43.2316
ANN	38.1501	45.1802	9.4436	147.5602	38.6672	10.2724	157.654	44.1603	13.3899	336.4765	32.5668
Additive model	39.876	47.8622	9.6785	164.9467	39.4593	11.3616	214.2265	41.8799	13.4117	333.7192	37.1394
Multiplicative model	35.8559	40.5158	6.4401	106.8691	22.6874	5.6789	95.5013	12.3727	7.1399	214.922	12.9885

*The values are based on averaging 10 runs of 10-fold cross-validation.
[a]Mean deviation.
[b]Standard deviation.

the value of MAPE associated with the multiplicative model was the lowest among the four models. The comparison between the actual values and the forecasted values for the 67 periods are given in Figure 7.

In a similar fashion, the analysis of the Canadian lynx data was undertaken by a subset of AR model of order 11. Table 3 shows the forecast results for the last 14 years. The neural network structure of $7 \times 5 \times 1$ made slightly better forecasts than the

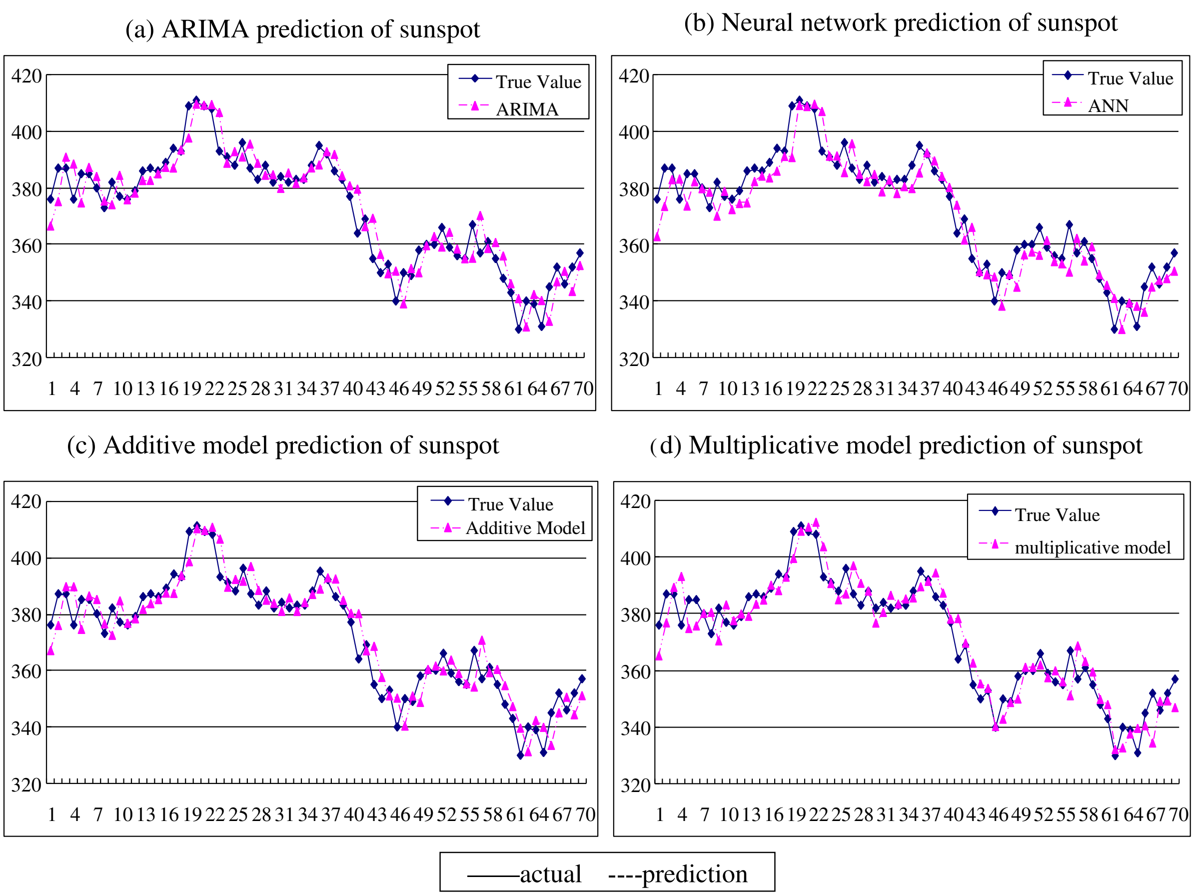

Figure 7 (a) ARIMA prediction of sunspot, (b) neural network prediction of sunspot, (c) additive model prediction of sunspot and (d) multiplicative model prediction of sunspot

Table 3 Forecasts for Lynx data

	Measures of fit (training set)*					Forecast accuracy (testing set)		
	MD[a]	SD[b]	MAD	MSE	MAPE (%)	MAD	MSE	MAPE (%)
ARIMA	0.322252	0.3803021	0.1159084	0.020183	3.702328	0.112559	0.016941	3.813114
ANN	0.3526945	0.4093142	0.1198109	0.020466	3.604238	0.111033	0.017857	3.598819
Additive model	0.309976	0.3679077	0.103972	0.016233	3.443875	0.109997	0.017521	3.643365
Multiplicative model	0.2998202	0.3532316	0.110035	0.019075	3.54008	0.103524	0.018625	3.589712

*The values are based on averaging 10 runs of 10-fold cross-validation;
[a]mean deviation;
[b]standard deviation.

Syst. Res. 30, 244–259 (2013)
DOI: 10.1002/sres

ARIMA model. The multiplicative model generated more accurate forecast results based on the measures of accuracy of MAD and MAPE. In relation to MAD, the multiplicative model improved the forecasting ability by 8.03%, 6.81% and 5.88% from the ARIMA, ANN and additive model, respectively. It can also be noticed that the values of MAPE associated with the multiplicative model was the lowest. However, the same cannot be stated for the MSE measure for the multiplicative model when compared with the other models. Perhaps, this is most likely caused by the fact that there is relatively small number of periods to predict. Figure 8 shows the actual and forecast values produced by the four different models.

For the IBM stock price data set, a neural network of $6 \times 5 \times 1$ was selected to model the nonlinear patterns. Table 4 presents the measures of accuracy results during the training and testing phases based on the four models. During the testing phase for the 3 and 10 weeks, it can be seen that the hybrid model performed better in accuracy than the other models. Furthermore, the multiplicative model made more accurate predictions than the additive model for long-term forecasting (10 weeks), whereas the opposite was true for the short-term situation. With regard to MAD in long-term forecasting, the multiplicative model was 6.93%, 11.5% and 5.51% better as compared with ARIMA, ANN and additive model, respectively. As for MAPE, the multiplicative model improved the forecasting accuracy by 7.07%, 11.11%, and 5.77% for the three models, respectively. Also, there is a

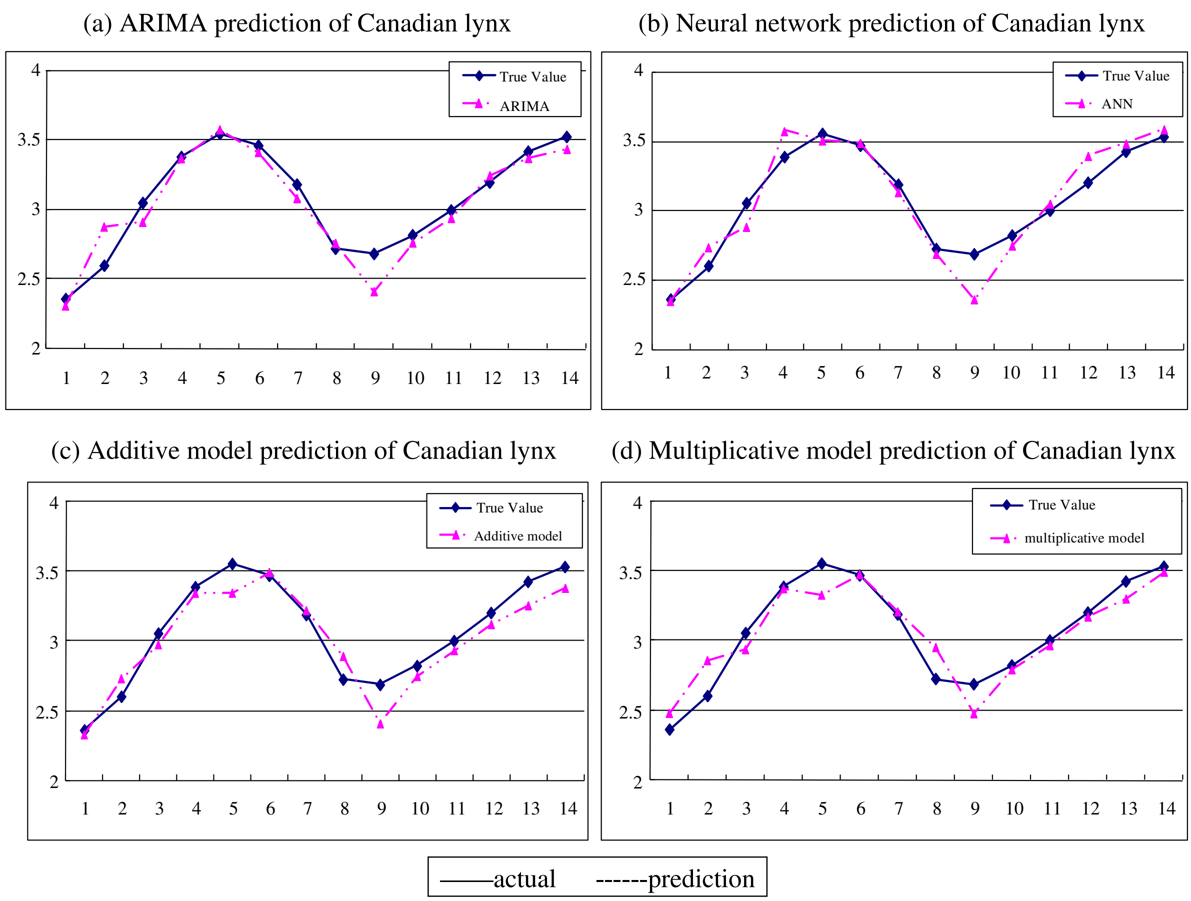

Figure 8 (a) ARIMA prediction of Canadian lynx, (b) neural network prediction of Canadian lynx, (c) additive model prediction of Canadian lynx and (d) multiplicative model prediction of Canadian lynx

Syst. Res. **30**, 244–259 (2013)
DOI: 10.1002/sres

Table 4 Forecasts for IBM stock price data

	Measures of fit (training set)					Forecast accuracy (testing set)					
						3 weeks			10 weeks		
	MD[a]	SD[b]	MAD	MSE	MAPE (%)	MAD	MSE	MAPE (%)	MAD	MSE	MAPE (%)
ARIMA	17.1265	19.8081	5.1031	53.1132	1.3089	4.9286	40.9368	1.2782	5.5086	47.0342	1.4971
ANN	17.1238	19.755	5.0355	51.7971	1.2807	6.2423	63.3021	1.6135	5.7436	52.0047	1.5651
Additive model	17.1193	19.796	4.8305	37.1861	1.1907	4.9872	42.1258	1.2953	5.4355	46.1861	1.4764
Multiplicative model	17.4039	20.1444	5.0352	40.9916	1.2077	5.6809	53.3387	1.4783	5.1516	45.606	1.3913

*The values are based on averaging 10 runs of 10-fold cross-validation;
[a]mean deviation;
[b]standard deviation.

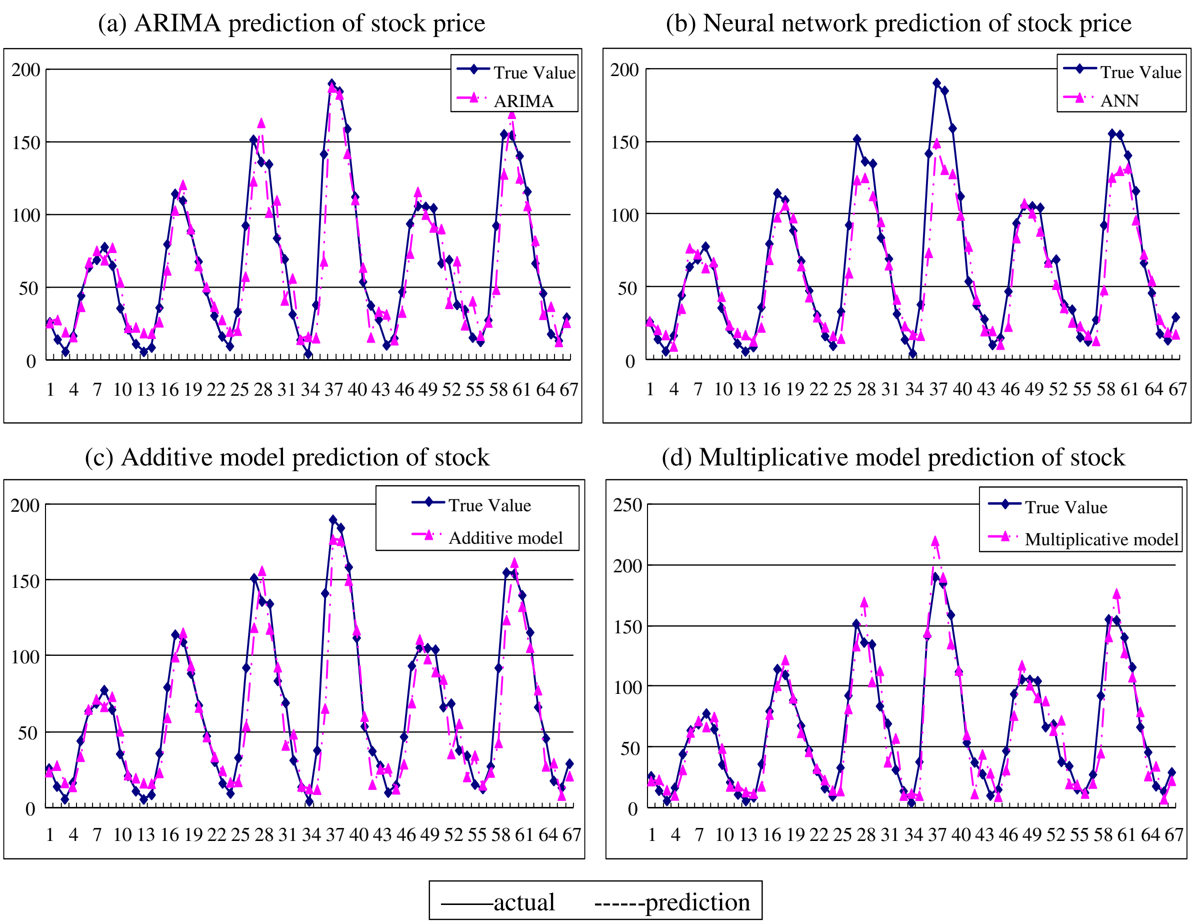

Figure 9 (a) ARIMA prediction of stock price, (b) neural network prediction of stock price, (c) additive model prediction of stock price and (d) multiplicative model prediction of stock price

Syst. Res. **30**, 244–259 (2013)
DOI: 10.1002/sres

considerable improvement of accuracy related to MSE from utilizing the multiplicative model. By contrast, for short-term forecasting (3 weeks), ARIMA and the additive model provided more accurate predictions than ANN and the multiplicative model. This suggests that for longer time spans, based on these data set, the multiplicative model outperforms the other three models, whereas for short-term forecasting, the ARIMA model worked better on the data set to capture the pattern of the time series. However, as the computational experience indicates that the additive and the multiplicative models increased the forecasting accuracy to a greater extent. The graphs between actual and predicted values are illustrated in Figure 9.

CONCLUSION

Both ARIMA and ANNs have been extensively utilized in time series analysis and forecasting because of their flexibility and effectiveness in modeling a variety of problems over the past few decades. However, many researchers have found that neither of them is universally suitable as a model for making forecasts regardless of the uniqueness of the situation. To improve the accuracy and performance of time series forecasting, a number of recent studies have focused on building hybrid models.

In this paper, a hybrid model (the multiplicative model) was proposed, which combined the time series ARIMA model and the neural network model to make predictions with nonseasonal time series data. The linear ARIMA model and the nonlinear ANNs model were employed jointly, with the aim to capture the different patterns in the time series data. The hybrid model integrated the versatilities of ARIMA and ANNs in linear and nonlinear modeling. The combinatorial approach was devised as an effective way to improve forecasting performance because of the complexity in linear and nonlinear structures (Shi et al. 1996, 1999; Feng and Xu 1999). Our results showed that multiplicative model was superior to the ARIMA model, the ANN model and the additive model, which were based on the computational experience with three different data sets. The measures of accuracy of MSE, MAD and MAPE were used as the evaluation criteria. Based on our computational experience, the lowest MSE, MAD and MAPE were achieved for the multiplicative model in almost all the situations with the exception of some short-term forecasts. It should be noted that for the sunspots time series forecasting, multiplicative model significantly increased the accuracy levels. Also, the empirical results with three data sets clearly showed that the multiplicative model outperformed each individual model.

ACKNOWLEDGEMENT

The authors acknowledge the support of the National Natural Science Foundation of China (Grant Nos. 71271013, 91224007, 70971005, 70971004 and 71132008), National Social Science Foundation of China (Grant No. 11AZD096), and Quality Inspection Project (Grant Nos. 201010268,200910088 and 200910248).

REFERENCES

Bates J, Granger C. 1969. The combination of forecasts. *Operational Research Quarterly* **20**(4): 451–468.

Box G, Jenkins G. 1970. *Time Series Analysis, Forecasting and Control*. Holden-Day: San Francisco, CA.

Campbell M, Walker A. 1977. A survey of statistical work on the MacKenzie River series of annual Canadian lynx trappings for the years 1821–1934, and a new analysis. *Journal of the Royal Statistical Society Series A* **140**(4): 411–431.

Chaudhry S, Varano M, Xu L. 2000. Systems research, genetic algorithms and information systems. *Systems Research and Behavioral Science* **17**: 149–162.

Chen X, Fang Y. 2012. Enterprise systems in financial sector- an application in precious metal trading forecasting. *Enterprise Information Systems* DOI:10.1080/17517575.2012.698022

Clemen R. 1989. Combining forecasts: a review and annotated bibliography. *International Journal of Forecasting* **5**: 559–608.

Duan L, Xu L. 2012. Business Intelligence for Enterprise Systems: a Survey. *IEEE Transactions on Industrial Informatics* **8**(3): 679–687.

Elkateb M, Solaiman K, Al-Turki Y. 1998. A comparative study of medium-weather-dependent load forecasting using enhanced artificial/fuzzy neural network and statistical techniques. *Neurocomputing* **23**(1–3): 3–13.

Feng S, Xu L. 1999. Hybrid Artificial Intelligence Approach to Urban Planning. *Expert Systems* **16**(4): 248–261.
Ginzburg I, Horn D. 1993. *Combined neural networks for time series analysis*. *Proceedings of Advances in Neural Information Processing Systems 6*, 7th NIPS Conference, Denver, Colorado, USA, 1993, pp. 224–231.
Hipel K, McLeod A. 1994. *Time Series Modeling of Water Resources and Environmental Systems*. Amsterdam: Elsevier.
Ho S, Xie M. 1998. The use of ARIMA models for reliability forecasting and analysis. *Computers and Industrial Engineering* **35**(1–2): 213–216.
Ho S, Xie M, Goh T. 2002. A comparative study of neural network and Box-Jenkins ARIMA modeling in time series prediction. *Computers and Industrial Engineering* **42**(2–4): 371–375.
Jiang Y, Xu L, Wang H, Wang H. 2009. Influencing Factors for Predicting Financial Performance based on Genetic Algorithms. *Systems Research and Behavioral Science* **26**(6): 661–673.
Kodogiannis V, Lolis A. 2002. Forecasting Financial Time Series using Neural Network and Fuzzy System-based Techniques. *Neural Computing & Applications* **11**(2): 90–102.
Kohzadi N, Boyad M, Kermanshahi B, Kaastra I. 1996. A comparison of artificial neural network and time series models for forecasting commodity price. *Neurocomputing* **10**(2): 169–181.
Li H, Li L. 1999. Representing diverse mathematical problems using neural networks in hybrid intelligent systems. *Expert Systems* **16**(4): 262–272.
Li H, Xu L. 2000. A Neural Network Representation of Linear Programming. *European Journal of Operational Research* **124**: 224–234.
Li L, Ge R, Zhou S, Valerdi R. 2012. Guest editorial Integrated healthcare information systems. *IEEE Transactions on Information Technology in Biomedicine* **16**(4): 515–517.
Luxhoj J, Riis J, Stensballe B. 1996. A hybrid econometric-neural network modeling approach for sales forecasting. *International Journal of Production Economics* **43**(2–3): 175–192.
Makridakis S, Anderson A, Carbone R, et al. 1982. The accuracy of extrapolation (time series) methods: results of a forecasting competition. *Journal of Forecasting* **1**(2): 111–153.
Palm F, Zellner A. 1992. To combine or not to combine? Issues of combining forecasts. *Journal of Forecasting* **11**(8): 687–701.
Patterson D. 1996. *Artificial Neural Networks*. Prentice-Hall: Englewood Cliffs, NJ.
Pelikan E, de Groot C, Wurtz D. 1992. Power consumption in West-Bohemia: improved forecasts with decorrelating connectionist networks. *Neural Network World* **2**: 701–712.
Rao T, Sabr M. 1984. *An Introduction to Bispectral Analysis and Bilinear Time Series Models, Lecture Notes in Statistics, Vol. 24*. Springer-Verlag: New York.
Reid D. 1968. Combining three estimates of gross domestic product. *Economica* **35**(140): 431–444.
Sfetsos A, Siriopoulos C. 2004. Combinatorial time series forecasting based on clustering algorithms and neural networks. *Neural Computing & Applications* **13**(1): 56–64.
Shi S, Xu L, Liu B. 1996. Applications of Artificial Neural Networks to the Nonlinear Combination of Forecasts. *Expert Systems* **13**: 195–201.
Shi S, Xu L, Liu B. 1999. Improving the Accuracy of Nonlinear Combined Forecasting using Neural Networks. *Expert Systems with Applications* **16**: 49–54.
Tang Z, Fishwick P. 1993. Feedforward neural nets as models for time series forecasting. *ORSA Journal on Computing* **5**: 374–385.
Trippi R, Turban E. 1996. *Neural Networks in Financial and Investing*. Irwin: Chicago.
Tsaih R, Hsu Y, Lai C. 1998. Forecasting S&P 500 stock index futures with a hybrid AI system. *Decision Support Systems* **23**(2): 161–174.
Tseng F, Yu H, Tzeng G. 2002. Combining neural network model with seasonal time series ARIMA model, *Technological Forecasting and Social Change* **69**(1): 71–87.
Walls L, Bendell A. 1987. Time series methods in reliability. *Reliability Engineering* **18**(4): 239–265.
Wang J, Leu J. 1996. Stock market trend prediction using ARIMA-based neural networks. *Proceedings of IEEE International Conference on Neural Networks 1996*, pp. 2160–2165.
Wang P, Xu L, Zhou S, Fan Z, Li Y, Feng S. 2010. A Novel Bayesian Learning Method for Information Aggregation in Modular Neural Networks. *Expert Systems with Applications* **37**(2): 1071–1074.
Wedding D, Cios K 1996. Time series forecasting by combining RBF networks, certainty factors, and the Box–Jenkins model. *Neurocomputing* **10**(2): 149–168.
Wong B, Selvi Y. 1998. Neural network application in finance: a review and analysis of literature (1990–1996). *Information Management* **34**(13): 129–139.
Xu L. 2000. Systems Science, Systems Approach and Information Systems Research. *Systems Research and Behavioral Science* **17**: 103–104.
Yang B, Li L, Ji H, Xu J. 2001. An early warning system for loan risk assessment using artificial neural networks. *Knowledge-based Systems* **14**(5): 303–306.
Yin H, Fan Y, Xu L. 2012. EMG and EPP-integrated human-machine interface between the paralyzed and rehabilitation exoskeleton. *IEEE Transactions on Information Technology in Biomedicine* **16**(4): 542–549.
Zhang G. 2003. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* **50**: 159–175.

Syst. Res. **30,** 244–259 (2013)
DOI: 10.1002/sres

Zhang G, Patuwo B, Hu M. 1998. Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting* **14**(1): 35–62.

Zhou S, Xu L. 2001. A New Type of Recurrent Fuzzy Neural Network for Modeling Dynamic Systems. *Knowledge-Based Systems* **14**: 243–251.

Zhou Z, Valerdi R, Zhou S. 2012. Guest editorial Special section on enterprise systems. *IEEE Transactions on Industrial Informatics* **8**(3): 630.

Zhu X, Wang H, Xu L, Li H. 2008. Predicting Stock Index Increments by Neural Networks: The Role of Trading Volume under Different Horizons. *Expert Systems with Applications* **34**(4): 3043–3054.

Syst. Res. **30,** 244–259 (2013)
DOI: 10.1002/sres