Human Systems Management 31 (2012) 133–143
DOI 10.3233/HSM-2012-0763
IOS Press

A hybrid model for time series forecasting

Yi Xiao[a,c], Jin Xiao[b,c] and Shouyang Wang[c,*]

[a] *Department of Information Management, Central China Normal University, Wuhan, China*
[b] *Business School, Sichuan University, Chengdu, China*
[c] *Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China*

Abstract. For time series, the problem that we often encounter is how to extract the patterns hidden in the real world data for forecasting its future values. A single linear or nonlinear model is inadequate in modeling and forecasting the time series, because most of them usually contain both linear and nonlinear patterns. This study constructs a hybrid forecasting model that combines autoregressive integrated moving average (ARIMA) with Elman artificial neural network (ANN) for short-term forecasting of time series. The proposed approach considers the linear and nonlinear patterns in the real data simultaneously so that it can mine more precise characteristics to describe the time series better. Finally, the forecasting results of the hybrid model are adjusted with the knowledge from text mining and expert system. The empirical results on the container throughput forecast of Tianjin Port show that the forecasts by the hybrid model are superior to those of ARIMA model and Elman network.

Keywords: TEI@I methodology, Elman artificial neural network, autoregressive integrated moving average, hybrid model, time series forecasting

Yi Xiao is an associate professor at Department of Information Management of Central China Normal University and a postdoctoral research fellow at Chinese Academy of Sciences. His research interest includes business intelligence, knowledge management and economic forecasting.

Jin Xiao received his Ph.D. degree from Sichuan University in 2010. He is an assistant professor at Business School of Sichuan University and a postdoctoral research fellow at Chinese Academy of Sciences. His research interest includes business intelligence data mining, customer relationship management, and knowledge management.

Shouyang Wang is currently a Bairen distinguished professor of Systems Management at Academy of Mathematics and Systems Science, Chinese Academy of Sciences. He is the editor-in-chief or a co-editor of 12 leading journals and the president of International Society of Knowledge and Systems Sciences. His research interest includes economic forecasting, knowledge management and decision support systems.

*Corresponding author: Shouyang Wang, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China. Tel.: +86 10 62651375; Fax: +86 10 62541829; E-mail: sywang@amss.ac.cn.

1. Introduction

Time series analysis and forecasting is an active research area that has drawn considerable attention for applications in variety of areas including business, economics, government, social science, environmental science, medicine, politics, finance, etc. Forecasting is an important problem because the prediction of future events is a critical input for many types of planning and decision making processes. Over the past decades, various kinds of forecasting models have been developed and have improved the accuracy of time series forecasting.

In time series forecasting, the historical data of the criterion variable can be utilized to construct a model reflecting its underlying feature. Then the established model will be used to speculate the future development of the time series. This analysis method is particularly useful when little knowledge is available on the underlying generating process of the data or when there is no satisfactory explanatory model that describes the relationship between the criterion variable and other prediction variables. In general, there are two types of methods for time series forecasting: qualitative and quantitative. Qualitative methods, e.g., Delphi method and experts meeting, forecast the future development of the object depending on the experts' experience, knowledge and analytical skills mainly. Quantitative methods usually establish mathematical forecasting models based on historical statistical data. Since the latter are more objective and precise, they have gotten more and more attention. According to the difference of the methods, they can be divided into three categories: statistical methods, causal analysis and combination forecasting. Statistical methods, such as autoregressive integrated moving average model (ARIMA), exponential smoothing, gray system method, seasonal adjustment method and Kalman filtering, construct a mathematical model only by historical data [7]. Causal analysis methods examine the correlation between the criterion variable and a series of economic indicators, and construct a forecasting model according to the relevant indicators, including regression analysis, the elasticity coefficient method and system dynamics [19]. In fact, there is no one technology or method that is able to consistently outperform all of the other methods in any case. Therefore, the combination forecasting methods get the final forecast results by integrating the results of some individual models. In this context, Wang et al. [25] proposed the TEI@I methodology integrating qualitative and quantitative analysis, which has been successfully applied to an increasingly number of areas owing to its high forecast effectiveness.

In monitoring the changes in seasonal patterns and business cycles, short-term forecasts often yield better results than long-term forecasts [12]. However, it is not easy to forecast short-term volatility due to their typical non-linearity and irregularity. The difficulty in forecasting economic time series is usually attributed to the limitation of many traditional forecasting models, which has encouraged academic researchers and business practitioners to develop more effective forecasting models. In this case, the artificial intelligence models such as artificial neural network (ANN) have been recognized as more useful than traditional statistical forecasting models. The major advantage of ANN is their flexible nonlinear modeling capability, which can approximate any continuous measurable function with arbitrarily desired accuracy [8]. As a kind of nonparametric and data-driven model, the ANN needs few prior assumptions on the underlying process from which data are generated, therefore, it is suitable for many empirical datasets where no theoretical guidance is available to understand the data generating process. For example, Lam et al. [16] developed the neural network model and forecasted 37 types of freight movements of Hong Kong Port, and it is shown that the forecasting results are more accurate compared with those of regression analysis. Although ANN has the advantages of accurate forecasting, its performance in some specific circumstances is inconsistent. Foster et al. [11] made comparisons between neural network and linear model in exchange rate forecasting. Their results show that linear regression and a simple average of exponential smoothing methods are significantly superior to ANN. Further, Taskaya and Casey [23] found that linear autoregressive model could outperform neural network in some cases. The reason may be that the data is approximately linear without much disturbance, thus, linear models are more suitable for linear relationships than ANN.

In all linear time series forecasting approaches, one of the most important models is ARIMA, which has been widely used to forecast social, economic, engineering, and financial problems. The popularity of ARIMA model is mainly due to its statistical properties and the well-known Box-Jenkins methodology in the modeling process. In addition, various exponential smoothing models can be implemented by ARIMA models [18]. However, a linear correlation structure is assumed among the data in ARIMA model, and therefore, it cannot capture the nonlinear patterns. The approximation of linear models to the complex problem in the real world is not always satisfactory because the real world systems are often nonlinear [30].

As we all know, it is difficult to know the characteristics of the data in the real world completely. Therefore, the hybrid model has become a common practice to improve the forecasting accuracy. The basic idea of the model combination in forecasting is to utilize each model's unique feature to capture different forms of relationship in the time series. In fact, the hybrid approach is not a new idea. As the pioneer work of combination forecasting, Reid [21], Bates and Granger [5] showed that a linear combination of forecasts would achieve a smaller error variance than any of the individual methods. Since then, the studies on the combination forecasting have expanded dramatically. Among these studies, decomposing a time series into its linear and nonlinear forms is one of the most popular hybrid models [29], which has been used to improve the forecast performance in many cases successfully [4].

This paper proposes a TEI@I based hybrid model integrating ARIMA and Elman neural network for economic time series forecasting. In order to verify the effectiveness of the proposed forecasting model, the monthly container throughput time series of Tianjin Port is used for empirical study. The experimental results show that the forecasts with the proposed model are superior to ARIMA model, as well as Elman neural network.

The remainder of this study is organized as follows: the framework of TEI@I based hybrid forecasting model we proposed is presented in Section 2; the basic concepts of ARIMA and Elman neural network are introduced in Section 3; the modeling process of the hybrid forecasting model integrating ARIMA and Elman neural network is described in Section 4; an empirical study is conducted on container throughput forecasting of Tianjin Port with monthly time series in Sections 5; Finally, some concluding remarks are drawn in Section 6.

2. Framework of TEI@I based hybrid forecasting model

TEI@I methodology, proposed by Wang [25], is a new methodology and can be applied to analyze the complicated systems which are emergent, unstable, non-linear and uncertain. In this methodology, econometrical models are used to model the linear pattern of the time series (i.e., main trends) while nonlinear

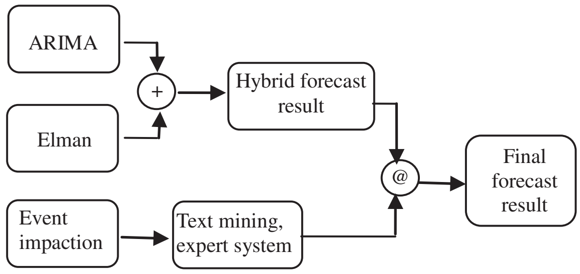

Fig. 1. Framework of TEI@I based hybrid forecasting model.

pattern of the time series (i.e., error terms) is modeled by an artificial neural network (ANN) model. In addition, the effects of irregular and infrequent events on the time series are explored by text mining technique and expert system technique. Thus, within the framework of TEI@I methodology, an integrated forecasting model with error correction and judgemental adjustment is formulated for generating more accurate forecasts. A more detailed description of TEI@I methodology can be found in reference [25].

Figure 1 shows the framework of TEI@I based hybrid forecasting model for economic time series proposed in this study. It includes three steps. In the first step, the ARIMA model is employed to explore the linear component of the economic time series. In the second step, the forecasting residuals of ARIMA model are analyzed by an Elman neural network further. The hybrid forecast results can be obtained by integrating the forecasts of ARIMA model and Elman neural network. In the third step, it utilizes the text mining technique and expert system to analyze the effects of irregular events on the system, and then adjusts the forecast results with the knowledge from the text mining and the expert system. The process of analyzing the effects of irregular events is basically the same as Wang et al. [25], and we do not give the details here. The detailed modeling processes of the first step and the second step of the proposed model are described in Section 4.

3. Time series forecasting models

There are a large number of approaches to time series modeling. The traditional statistical models such as moving average (MA), exponential smoothing, and ARIMA belong to linear forecasting, i.e., the forecasts of the future values are constrained to be linear function of past observations. In the past decades, some types of nonlinear forecasting models have been proposed to overcome the limitation of linear time series models, including bilinear model [13], threshold autoregressive (TAR) model [24], autoregressive conditional heteroscedastic (ARCH) model [10], general autoregressive conditional heteroscedastic (GARCH) [2], chaotic dynamics [3], etc. Although some improvement has been noticed with these nonlinear models, the gain of using them for general forecasting problems is limited [1]. Recently, some ANN models such as multilayer perceptron (MLP), radial basis function network (RBF), general regression neural network (GRNN) have been suggested as an alternative to time series forecasting. The main advantage of the ANN is their flexible nonlinear modeling capability. In this section, the basic principles and modeling processes of the ARIMA and ANN models are briefly introduced.

3.1. ARIMA model

Suppose that z_t is the observation of a time series, and its mean is μ. A mixed autoregressive moving average (ARMA) model is expressed as

$$\phi(B)\bar{z}_t = \theta(B)a_t, \tag{1}$$

where $\bar{z}_t = z_t - \mu$, a_t and B are the random error of the time series at time t and the backward shift operator respectively. $\phi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i$ and $\theta(B) = 1 - \sum_{j=1}^{q} \theta_j B^j$ are polynomials of degree p and q respectively, $\phi_i (i = 1, 2, \ldots, p)$ and $\theta_j (j = 1, 2, \ldots, q)$ are model parameters, p and q are integers and often refer to the orders of the model. Further, the random errors a_t are assumed to be independently and identically distributed with a mean of zero and a constant variance of σ^2.

The ARMA process is stationary if the roots of $\phi(B) = 0$ lie outside the unit circle and it exhibits explosive non-stationary behaviour if they lie inside the unit circle. If $\phi(B)$ is a stationary autoregressive operator, then the autoregressive integrated moving average model (ARIMA) process is derived as

$$\phi(B)(1 - B)^d \bar{z}_t = \theta(B)a_t, \tag{2}$$

where d is an integer and often referred to the order of difference. Introducing the backward difference operators $\nabla = 1 - B$ and $\nabla^d \bar{z}_t = \nabla^d z_t$, the above equation becomes

$$\phi(B)\nabla^d z_t = \theta(B)a_t, \tag{3}$$

where considering the various values of p, d and q, the ARIMA model is written as ARIMA(p, d, q). In ARIMA, the future value of a variable is assumed to be a linear function of several past observations and random errors.

Box and Jenkins [6] developed a practical approach to build ARIMA models, which has a fundamental impact on the applications of time series analysis and forecast. Their methodology involves the following three iterative steps: model identification, parameter estimation and diagnostic checking. The autocorrelation function (ACF) and the partial ACF (PACF) of the sample data are utilized as the basic tools to identify the order of the best ARIMA model. The data transformation is often needed to make the time series stationary. A stationary time series is characterized by statistical characteristics such as the mean and the autocorrelation structure being constant over time. When the observed time series presents trend and heteroscedasticity, difference and power transformation are applied to the data to remove the trend and to stabilize the variance before an ARIMA model can be fitted.

After the function of the ARIMA model has been specified, the parameters of the function need to be estimated further. The parameters are estimated so that an overall measure of errors is minimized. This usually involves the use of a least squares estimation process. The last step of model building is the diagnostic checking of model adequacy, which mainly checks if the model assumptions about the errors a_t are satisfied. Several diagnostic statistics and plots of the residuals can be used to examine the goodness of fit of the tentatively entertained model to the historical data. If the model is not adequate, a new tentative model should be identified, which is again followed by the steps of parameter estimation and model verification. The diagnostic information can help us to find some alternative models. This three-step model building process is typically repeated several times until a satisfactory model is selected finally. The final selected model can then be used for forecasting.

3.2. Artificial neural network

Artificial neural network (ANN), first proposed in 1943 [17], is derived through neuropsychology. The basic idea of ANN is to emulate the biological system of the human brain to learn and identify patterns. Different ANN models have been proposed, in which feedforward neural network (FNN) is the most widely used one. Usually, the FNN learning mechanism consists of two phases: training and testing. During the training phase, the input and output layers are set to represent a training pair (x, y) where x is the independent variable vector and y is the dependent variable vector. The goal is to construct a following equation:

$$y = f(x) = f_a(x) = f(x; a), \quad a \in A, \tag{4}$$

where $f = f_a$ is defined by specifying parameters $a \in A$ from an explicitly parameterized family A of the models. Usually the training algorithm of FNN repeatedly adjusts the connection-weight matrices by minimizing the error for each training pair. When the average squared error over all training pairs is acceptable, the training procedure stops and the final connection-weight matrices are stored as knowledge for use in the testing phase.

During the testing phase, the input layer of the FNN is activated by the new input vectors. This activation of FNN spreads from the input layer to the output layer through the connection-weight matrices stored in the training phase. That is, $f = f_a$, determined by the training phase, is applied to new feature vectors x' unused in training phase and the forecasting output $y = f(x')$ of FNN is produced. We usually measure the effectiveness of the model in a test set, which is distinctly different from the training set.

The FNN is widely used for time series forecasting [15, 22, 31] because its nonlinear modeling capability can capture the nonlinear characteristics of time series well. When applying FNN to time series forecasting, the final output can be represented as

$$y_t = a_0 + \sum_{j=1}^{n} w_j f(a_j + \sum_{i=1}^{m} w_{ij} y_{t-i}) + \xi_t, \tag{5}$$

where $a_j (j = 0, 1, 2 \ldots, n)$ is a bias on the j-th unit, and $w_{ij} (i = 1, 2, \ldots, m; j = 1, 2 \ldots, n)$ is the connection weight between layers of the model, $f(\cdot)$ is the transfer function of the hidden layer, and m is the number of input nodes and n is the number of hidden nodes. Actually, the FNN model in Equation (4) performs a nonlinear function mapping from the past observed values $(y_{t-1}, y_{t-2}, \ldots, y_{t-m})$ to the future value y_t, i.e.,

$$y_t = \varphi(y_{t-1}, y_{t-2}, \ldots, y_{t-m}, v) + \xi_t, \tag{6}$$

where v is the parameter vector and φ is a function determined by the network structure and connection weights. Thus, in some senses, the FNN model is equivalent to a nonlinear autoregressive model [28].

A major advantage of FNN is its ability to provide flexible mapping between the inputs and outputs. Furthermore, Hornik et al. [14] and White [26] have proved theoretically that a three-layer feedforward neural network (FNN) with an identity transfer function in the output nodes and logistic functions in the middle-layer nodes can approximate any continuous function to any desire accuracy well given a sufficient number of middle-layer nodes. Therefore, a three-layer FNN model is used as a basic learning paradigm in this study.

3.3. Elman neural network

Elman neural network belongs to the class of recurrent neural networks architecture. It is a three layers feedforward neural network with the addition of a recurrent connection from the output of the hidden layer to its input. The network is augmented at the input level by additional units, called context units. The number of context units is equal to the number of hidden units. The augmented input units, including both the input units and the context units activate the hidden units [9]. The inputs to the context units are the outputs of the hidden neurons forming the second or hidden layer. The outputs of the context units and the external input neurons are fed to the hidden neurons. The context unit values at time step $t+1$ are exactly the same as the hidden unit values at time step t. Therefore, the current units which transfer the previous state of the hidden units to the input layer are recognized as a one-step time delay. Context units are also known as memory units as they store the previous output of the hidden neurons.

During Elman neural network operation, the activation values of the input units are set to a desired input pattern. The activation value of every hidden unit is computed by multiplying each input and context activation value by the value of the weight from the unit to the hidden unit. These values are then summed, the bias of the hidden unit is added, and the sum is passed through a squashing function f. The result value is then regarded as the output value of the hidden unit. In the Elman neural network, the squashing function used is the logistic f. Then, the activations of output units are calculated based on the hidden units in an analogous manner. This represents one time step. Next, the activation of each hidden unit is copied into a corresponding context unit on a one-for-one basis with fixed weights of 1, and then the next time step is performed. This is equivalent to a recurrent connection from every hidden unit to itself and is more restrictive than the arbitrary recurrent connections allowed by Minsky's claim [20].

Suppose that n, l, m are the number of the input units, hidden units and output units respectively. The primary input is x_i $(i = 1, 2, \ldots, n)$, and the network output is y_k $(k = 1, 2, \ldots, m)$. w_{ji} $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, l)$, w_{jr} $(r = 1, 2, \ldots, l; j = 1, 2, \ldots, l)$, w_{kj} $(j = 1, 2, \ldots, l; k = 1,$

2, ..., m) are the weights of the connections between the input and hidden units, the recurrent and the hidden units, and the hidden and the output units, respectively. b_j and b_k are biases of hidden units and output units, and $f(\cdot)$ and $g(\cdot)$ are hidden and output functions, respectively. The architecture of Elman neural network can be written mathematically as follows.

The output of the hidden unit:

$$y'_j(t) = f\left(\sum_{i=1}^{n} w_{ji}x_i + \sum_{r=1}^{l} w_{jr}y'_r(t-1) + b_j\right). \quad (7)$$

The output of the output unit:

$$y_k(t) = g\left(\sum_{j=1}^{l} w_{kj}y'_j(t) + b_k\right). \quad (8)$$

The conventional neural network is not suitable for the patterns that vary over time. However, the historical economic series data as the inputs of the neural network are time-varying. Therefore, a network with temporal processing ability should be considered. In this study, an Elman recurrent neural network is trained to forecast the residual series of economic time series.

4. The hybrid method integrating ARIMA with Elman neural network

In the real world, the time series forecasting is far from simple due to high volatility, complexity, irregularity and noise. Moreover, practical time series are rarely pure linear or nonlinear. They often contain both linear and nonlinear patterns. Although both ARIMA and Elman neural network models have succeeded in their own linear or nonlinear domains, neither ARIMA nor Elman neural network can adequately model and forecast time series since the linear models cannot deal with nonlinear relationships well while the Elman neural network is not able to handle both linear and nonlinear patterns equally well. Therefore, there is no universal model that is suitable for all kinds of time series. On the other hand, as previously mentioned, for time series forecasting the relationship between ARIMA and ANN is complementary. ARIMA is one of the linear models that can capture the linear characteristics of a time series, while Elman neural network trained by back-propagation with one hidden layer is one of the general function approximators with strong capability of modeling nonlinearity and they can capture nonlinear patterns in time series. It is not wise to apply only ARIMA or Elman neural network blindly to any time series because it is difficult to know the characteristics of the time series completely in the real problems. Therefore, in this study we propose a hybrid model integrating ARIMA and Elman neural network in an adaptive manner for economic time series forecasting. By combining the two models, different aspects of the underlying patterns may be captured. The purpose of the hybrid method is to integrate linear and nonlinear modeling capabilities and improve the prediction power in practical applications further.

Since the real economic time series usually contain many complex patterns, it is not enough to only model the linear component for a time series. Therefore, the nonlinear relationship of the time series should be considered in order to reflect actual circumstances better. It may be reasonable to suppose that a time series is composed of a linear autocorrelation structure and a nonlinear component:

$$y_t = L_t + N_t, \quad (9)$$

where y_t is actual value, L_t and N_t denote linear component and the nonlinear component respectively.

According to Equation (9), the two components have to be estimated from the data to obtain the final forecasting result. In the first phase, an ARIMA model is used to extract the linear component of time series. By comparing the actual value y_t of the time series and the forecast value $\hat{L}_t$ of linear component, we can obtain a series of residuals, which is defined as e_t:

$$e_t = y_t - \hat{L}_t. \quad (10)$$

Residuals are important in diagnosis of the sufficiency of linear models. A linear model is not sufficient if there are still linear correlation structures left in the residuals. In this study, we assume that the ARIMA model constructed in the first phase is sufficient. Therefore, the residuals from the linear model will not contain the linear relationship any more, and the linear model is not able to deal with it. Any significant nonlinear pattern in the residuals will indicate the limitation of the ARIMA. The second phase is how to fit the nonlinear component of series. In this phase, an Elman neural network model is used to model the above nonlinear time series. The nonlinear time series generated previously is regarded as the inputs of Elman model, and then the trained Elman model is used to generate a series of forecasts of nonlinear components of time series. With m

input nodes, the Elman neural network model for the residuals will be described as

$$e_t = \varphi(e_{t-1}, e_{t-2}, \dots, e_{t-m}, v) + \xi_t, \tag{11}$$

where v is the parameter vector, φ is a function determined by the Elman network structure and connection weights, and ξ_t is the random error.

In order to obtain the synergetic forecasting results, we only need to integrate the forecasts of linear and nonlinear components of time series. Thus, the final forecast $\hat{y}_t$ can be calculated as follows:

$$\hat{y}_t = \hat{L}_t + \hat{N}_t, \tag{12}$$

where $\hat{L}_t$ is the forecasting result of linear component and $\hat{N}_t$ is the forecasting value of nonlinear component.

In summary, the proposed hybrid method consists of three phases. In the first phase, an ARIMA model should be constructed to analyze the linear pattern of the time series. The procedure of building the ARIMA model can be summarized as follows.

1. Identify the order of the ARIMA model with graphs, statistics, autocorrelation function (ACF), partial autocorrelation function (PACF), Akaike's information criterion (AIC), the minimum description length (MDL), as well as intelligent approaches such as neural networks, genetic algorithms or fuzzy system;
2. Estimate the parameters of the identified model by the methods of least squares, maximum likelihood methods, etc.;
3. The diagnostic checking of model adequacy.

In the second phase, an Elman neural network model is introduced to model the nonlinear pattern. Since ARIMA model cannot capture the nonlinear component of the economic time series, the residuals of linear model will contain information about the nonlinear pattern. Therefore, the results from the Elman neural network can be used as predictions of the error term for the ARIMA model.

In the third phase, the combined forecasting results are obtained from Equation (12). The hybrid model can explore the unique feature and strength of ARIMA model as well as ANN model in determining different patterns. Theoretically speaking, the hybrid model presented in this study is still a class of single forecasting model. However, it is able to model linear and nonlinear patterns separately by using different models, and can improve the forecasting performance further.

5. Empirical study

5.1. Data preparation and preprocessing

The monthly container throughput time series of Tianjin Port in the period from January 2001 to December 2011 has been used in our experiments. The data are downloaded from the CEIC database (see Fig. 2). The monthly data from January 2001 to December 2010 are used as the training set for modeling, including the ARIMA and Elman neural network models, and the remaining data from January 2011 to December 2011 as the test set for model verification and comparison.

The data need to be normalized before training the models, which can be described as the following formula:

$$x_n(i) = \frac{x(i) - \min(X)}{\max(X) - \min(X)}, \tag{13}$$

where $x_n(i) \in [0, 1]$ is the normalized data, $x(i)$ is the original data, $\min(X)$ and $\max(X)$ represent the minimum and maximum of the original time series respectively.

5.2. Performance measures

In order to evaluate the performance of hybrid forecasting model, three different forecast performance measures are used.

The first is the mean absolute error (abbreviated as MAE), which is described as

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |T_i - T_i'|. \tag{14}$$

The second is the mean absolute percentage error (MAPE) that can be written as

$$MAPE = \frac{1}{N}\sum_{i=1}^{N} \left|\frac{T_i - T_i'}{T_i}\right|. \tag{15}$$

The third is the root mean squared error (RMSE), which can be presented as

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (T_i - T_i')^2}, \tag{16}$$

where N is the total number of observed data. T_i and $T_i{}'$ represent the actual and forecasting values respectively.

MAE, MAPE and RMSE are the metrics which are used to estimate the forecasting error of the model.

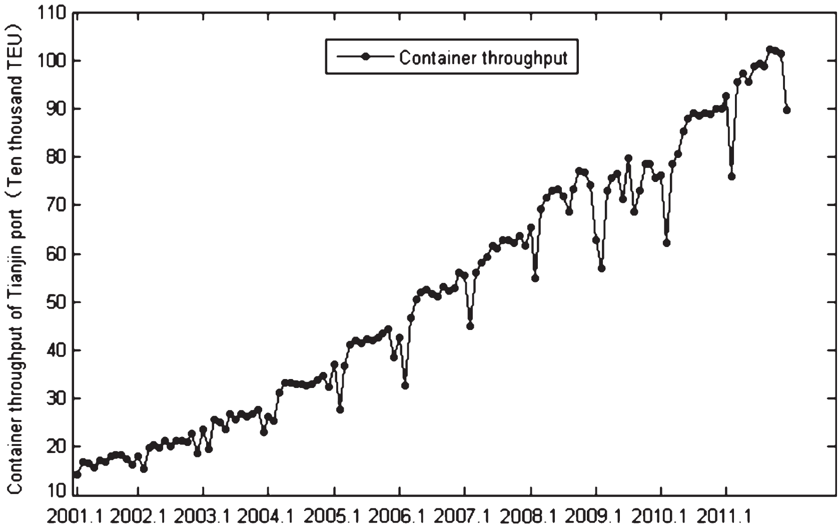

Fig. 2. Monthly container throughput data of Tianjin Port from January 2001 to December 2011.

Smaller values of these metrics indicate higher accuracy in forecasting.

5.3. ARIMA modeling

ARIMA model is used to model the linear component of container throughput time series. To construct ARIMA model with the available container throughput time series data, a three-stage procedure of model identification, estimation of model parameters and diagnostic checking of the estimated parameters was employed. We found that the best fitted model was a autoregressive model of order 2, i.e., AR(2) by the Eviews software. Further, the trained ARIMA model was utilized to forecast the monthly container throughput data from January 2011 to December 2011. As it is shown in Fig. 3 that the forecasting results of AR(2) have some biases compared with the actual values. These biases are mainly due to the limitations of the linear modeling of ARIMA model. Thus, we can conclude that ARIMA model is generally not suitable to identify and explore the nonlinear pattern of container throughput time series.

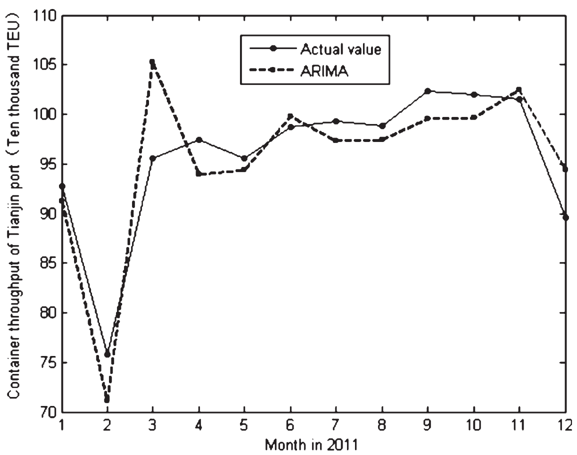

Fig. 3. The forecasting results of container throughput based on ARIMA model.

5.4. Elman neural network modeling

Elman neural network is used to model the nonlinear component of container throughput time series by gradient descent with momentum and optimized back-propagation training algorithm. Elman network is a feedforward network with the addition of a recurrent connection from the output of the hidden layer to its input. The delay in this connection stores the values from the previous time step, which can be used in the current time step. Thus, the Elman neural network can learn to recognize and generate temporal patterns, as well as spatial patterns with the feedback path. In this section, an Elman network was trained in MATLAB software according to the residuals of linear model. Finally, the optimal Elman network architecture composed of three inputs, four hidden and one output neurons, called E(3-4-1), was obtained by pruning technology. The forecasting results of the residuals for test set are shown in Fig. 4. The results indicate that

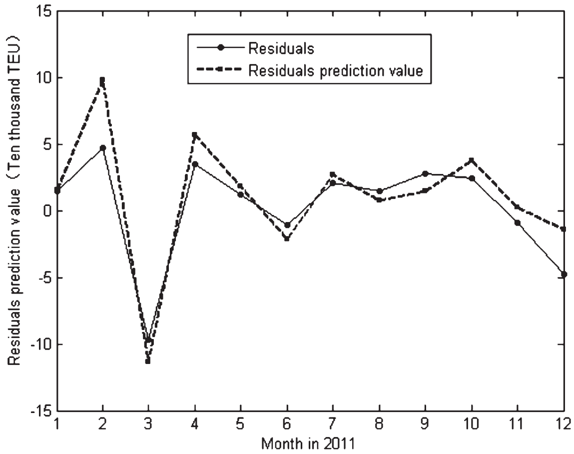

Fig. 4. The forecasting results of the residuals based on Elman network.

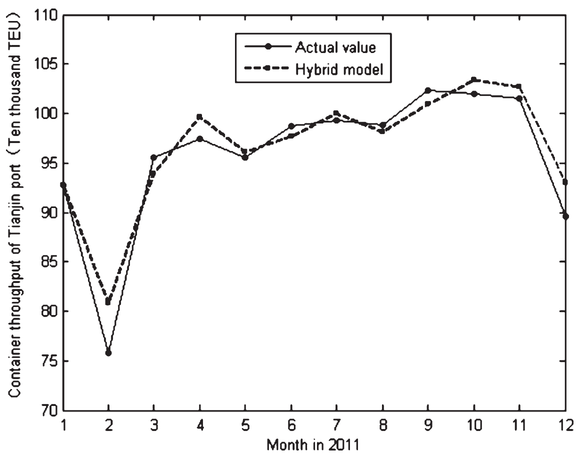

Fig. 5. The forecasting results of the hybrid model for container throughput.

Elman network is able to capture the nonlinear pattern of the container throughput time series to provide good prediction performances of the monthly fluctuation of container throughput.

5.5. *Integrating forecast based on ARIMA model and Elman neural network*

After getting the forecasting results of linear component of container throughput time series by AR(2) and the results of nonlinear component by Elman neural network, we only need to integrate the two forecast time series to get the final forecasting results. The forecast values of hybrid model in test set are shown in Fig. 5. The verification results suggest that the prediction results from the proposed hybrid model can match the observations more reasonably.

5.6. *Comparison of model performance*

Table 1 shows the simulation results of the ARIMA model, Elman network and hybrid model in test set, where each value is the average performance of 10 experiments.

It can be seen from Table 1 that: (a) according to all indices, the forecasting performance of ARIMA model is better than that of Elman network. The reason may be that in the container throughput time series we considered, the dominant factor is linear components; (b) the prediction performance of hybrid model is better than that of ARIMA model and Elman network, which demonstrates that the container throughput time series include both nonlinear components and linear components. Therefore, it is not suitable for the container throughput forecasting to utilize only ARIMA or Elman network; (c) the hybrid model is superior to the model proposed by Xiao et al. in another study [27]. Finally, we can conclude that the hybrid model proposed in this study is more favorable to model container throughput time series.

As can be seen in Fig. 2, because of the global financial crisis originated from United States, the container throughput growth of Tianjin Port slowed down from 2009 to 2010. But with the rapid recovery of Chinese economy from the financial crisis and the rapid development of Tianjin New Coastal Area, the container throughput of Tianjin Port began to grow quickly from 2010. As a causal factor, the financial crisis is not reflected in models based on the historical data, so the

Table 1
Comparison of average performance of different models over 10 experiments

Model	MAE	Order	MAPE	Order	RMSE	Order
Elman network	6.6482	3	0.0743	3	8.0681	3
ARIMA model	2.7227	2	0.0286	2	3.8018	2
TEI@I based hybrid model	2.2362	1	0.0234	1	3.618	1

Table 2
Forecasts of container throughput of Tianjin Port by TEI@I based hybrid model (Ten thousands TEU)

Month	2012.1	2012.2	2012.3	2012.4	2012.5	2012.6
Forecasting values	106.05	101.53	104.74	113.82	108.67	109.62

forecasting errors of ARIMA model or Elman network are relatively large. However, the prediction performance of TEI@I based hybrid model is significantly better than that of the other two models by adjusting with the knowledge from text mining and expert system. Therefore, we utilize the TEI@I based hybrid model to predict the container throughput of Tianjin Port in the next 6 months. The forecasting results are listed in Table 2 and ready for future validation.

6. Conclusions

In time series forecasting, the problem that we often encounter is how to increase the prediction accuracy as much as possible with the non-stationary and noise data. In this study we hope to design a TEI@I based hybrid model that can effectively improve the performance of container throughput forecasting. In the proposed hybrid model, the unique capability of ARIMA model in linear modeling is used in order to extract the existing linear pattern in data, and then an Elman network is adopted to capture nonlinear pattern from the residuals of linear model. We experiment ARIMA model, Elman network and the proposed hybrid model in monthly container throughput time series of Tianjin Port, compare the forecasting results based on the mean absolute error, mean absolute percent error and root mean squared error, and find that the forecasting performance of the proposed hybrid model is better than that of ARIMA model and Elman network in general. The proposed hybrid model may be used as an alternative forecasting tool and helpful for the short-term time series forecasting.

Acknowledgments

This research was supported by the Humanities and Social Sciences Youth Foundation of the Ministry of Education of PR China under Grant No. 11YJC870028, the Special Fund for Basic Scientific Research of Central Colleges under Grant No. CCNU10A01031, the Natural Science Foundation of China under Grant No. 71101100, New Teachers' Fund for Doctor Stations, Ministry of Education under Grant No. 20110181120047, China Postdoctoral Science Foundation under Grant No. 2011M500418, Center for Forecasting Science of the Chinese Academy of Sciences and the National Center for Mathematics and Interdisciplinary Sciences. This paper was completed during the corresponding author's visit to Department of Management Sciences, City University of Hong Kong. He is grateful to the department and the university for a financial support to his visit.

References

[1] C.H. Aladag, E. Egrioglu and C. Kadilar, Forecasting nonlinear time series with a hybrid methodology, *Applied Mathematics Letters* **22** (2009), 1467–1470.

[2] G. Ali, ARCH, GARCH, and ARMAX models for forecasting pathogen indicators and advisories at marine recreational sites, *Marine Resource Economics* **26** (2011), 211–224.

[3] M. Ardalani-Farsa and S. Zolfaghari, Residual analysis and combination of embedding theorem and artificial intelligence in chaotic time series forecasting, *Applied Artificial Intelligence* **25** (2011), 45–73.

[4] P. Areekul, T. Senjyu, H. Toyama and A. Yona, A hybrid ARIMA and neural network model for short-term price forecasting in deregulated market, *IEEE Transactions on Power Systems* **25** (2010), 524–530.

[5] J.M. Bates and C.W.J. Granger, The combination of forecasts, *Operations Research Quarterly* **20** (1969), 451–468.

[6] G.E.P. Box and G.M. Jenkins, Time series analysis, forecasting and control, holden-day, 1970.

[7] S.H. Chen and J.N. Chen, Forecasting container throughputs at ports using genetic programming, *Expert Systems with Applications* **37** (2010), 2054–2058.

[8] G. Cybenko, Approximations by superpositions of a sigmoidal function, *Mathematics of Control, Signals and Systems* **2** (1989), 303–314.

[9] J.L. Elman, Finding structure in time, *Cognitive Science* **14** (1990), 179–211.

[10] R.F. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. *Econometrica* **50** (1982), 987–1008.

[11] W.R. Foster, F. Collopy and L.H. Ungar, Neural network forecasting of short, noisy time series, *Computers and Chemical Engineering* **16** (1992), 293–297.

[12] P.H. Franses and D. Van Dijk, The forecasting performance of various models for seasonality and nonlinearity for quarterly industrial production, *International Journal of Forecasting* **21** (2005), 87–102.

[13] C.W.J. Granger and A.P. Anderson, *An Introduction to Bilinear Time Series Models*, Vandenhoeck-Ruprecht, 1978, pp. 1–8.

[14] K. Hornik, M. Stinchocombe and H. White, Multilayer feedforward networks are universal approximators, *Neural Networks* **2** (1989), 359–366.

[15] H.B. Hwang, Insights into neural-network forecasting time series corresponding to ARMA (p, q) structures, *Omega* **29** (2001), 273–289.

[16] W.H.K. Lam, P.L.P. Ng, W. Seabrooke and E.C.M. Hui, Forecasts and reliability analysis of port cargo throughput in Hong Kong, *Journal of Urban Planning and Development* **130** (2004), 133–144.

[17] W.S. McCulloch and W.A. Pitts, Logical calculus of the ideas imminent in nervous activity, *Bulletin and Mathematical Biophysics* **5** (1943), 115–133.

[18] E.D. McKenzie, General exponential smoothing and the equivalent ARMA process, *Journal of Forecasting* **3** (1984), 333–344.

[19] W.Y. Peng and C.W. Chu, A comparison of univariate methods for forecasting container throughput volumes, *Mathematical and Computer Modeling* **50** (2009), 1045–1057.

[20] D.T. Pham and X. Liu, Dynamic system identification using partially recurrent neural networks, *Journal of Systems Engineering* **2** (1992), 90–97.

[21] D.J. Reid, Combining three estimates of gross domestic product, *Economica* **35** (1968), 431–444.

[22] Z. Tang, C. Almeida and P.A. Fishwick, Time series forecasting using neural networks vs. Box–Jenkins methodology, *Simulation* **57** (1991), 303–310.

[23] T. Taskaya and M.C. Casey, A comparative study of autoregressive neural network hybrids, *Neural Networks* **18** (2005), 781–789.

[24] H. Tong, *Threshold Models in Non-Linear Time Series Analysis*, Springer-Verlag, 1983, pp. 3–9.

[25] S.Y. Wang, L. Yu and K.K. Lai, Crude oil price forecasting with TEI@I methodology, *International Journal of Systems Science and Complexity* **18** (2005), 145–166.

[26] H. White, Connectionist nonparametric regression: Multilayer feedforward networks can learn arbitrary mappings, *Neural Networks* **3** (1990), 535–549.

[27] Y. Xiao, J. Xiao and S.Y. Wang, A hybrid forecasting model for non-stationary time series: An application to container throughput prediction, *International Journal of Knowledge and Systems Sciences* **3** (2012), in press.

[28] L. Yu, S.Y. Wang and K.K. Lai, A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates, *Computers and Operations Research* **32** (2005), 2523–2541.

[29] G.P. Zhang, Time series forecasting using a hybrid ARIMA and neural network model, *Neurocomputing* **50** (2003), 159–175.

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

[31] G.P. Zhang, E.B. Patuwo and M.Y. Hu, A simulation study of artificial neural networks for nonlinear time-series forecasting, *Computers and Operations Research* **28** (2001), 381–396.