Available online at http://scik.org
J. Math. Comput. Sci. 1 (2011), No. 1, 103-125
ISSN: 1927-5307

EXCHANGE RATE FORECASTING BETTER WITH HYBRID ARTIFICIAL NEURAL NETWORKS MODELS

MEHDI KHASHEI*, MEHDI BIJARI

Department of Industrial and Systems Engineering, Isfahan 84156-83111, Iran

Abstract: Forecasting accuracy is one of the most important features of forecasting models; hence, never has research directed at improving upon the effectiveness of time series models stopped. Nowadays, despite the numerous time series forecasting models proposed in several past decades, it is widely recognized that exchange rates are extremely difficult to forecast. Artificial neural networks (ANNs) are one of the most accurate and widely used forecasting models that have been successfully applied for exchange rate forecasting. In this paper, an improved model of the artificial neural networks is proposed using autoregressive integrated moving average models, in order to yield more general and more accurate hybrid model than artificial neural networks for time series forecasting. In our proposed model, the unique advantages of the ARIMA models in linear modeling are used in order to preprocess the under-study data for using in artificial neural networks. Empirical results in weekly Indian rupee against the United States dollar exchange rate indicate that the proposed model can be an effective way to improve forecasting accuracy achieved by artificial neural networks and traditional linear models.

Keywords: Artificial Neural Networks (ANNs); Autoregressive Integrated Moving Average (ARIMA);

---

*Corresponding author

E-mail addresses: Khashei@in.iut.ac.ir (M. Khashei), Bijari@cc.iut.ac.ir (M. Bijari）

Received November 29, 2011

Time series forecasting; Financial markets; Exchange rate.

2000 AMS Subject Classification: 92B20; 91B84; 97M30

1. Introduction

Foreign exchange markets are among the most important and the largest financial markets in the world with trading taking place twenty-four hours a day around the globe and trillions of dollars of different currencies transacted each day. Transactions in foreign exchange markets determine the rates at which currencies are exchanged, which in turn determine the costs of purchasing foreign goods and financial assets. Therefore, understanding the evolution of exchange rates will be very important for many outstanding issues in international economics and finance, such as international trade and capital flows, international portfolio management, currency options pricing, and foreign exchange risk management [1]. However, the literature shows that predicting the exchange rate movements are largely unforecastable due to their high volatility and noise and still are a problematic task for both academic researchers and business practitioners to make short term and long term forecasting efficiently [2-4].

Artificial neural networks (ANNs) are one of the most popular models, which have been proposed and examined for forecasting exchange rates. ANNs have some advantages over other forecasting models, which make it attractive in exchange rates modeling. First, neural networks have flexible nonlinear function mapping capability, which can approximate any continuous measurable function with arbitrarily desired accuracy, whereas most of the commonly used nonlinear time series models in foreign exchange markets do not have this property. Second, being nonparametric and data-driven models, neural networks impose few prior assumptions on the underlying process from which data are generated [5]. Because of this property, neural networks are less susceptible to model misspecification problem than most parametric nonlinear methods. This is an important advantage since exchange rate does not exhibit a specific nonlinear pattern. Third, neural networks are adaptive in nature. The adaptivity implies that the network's generalization capabilities remain accurate and robust in a nonstationary environment whose characteristics may change over time. Fourth, neural networks use only linearly many parameters, whereas traditional polynomial, spline, and trigonometric expansions use exponentially many parameters to achieve the same approximation rate [6].

Given the advantages of neural networks, it is not surprising that this methodology has attracted overwhelming attention in financial markets, and especially exchange rate prediction. Many researchers have investigated the artificial neural networks as models for forecasting exchange rates and have shown that neural networks can be one of the very useful tools in foreign exchange markets forecasting [5]. Weigend et al. [7] have found that neural networks are better than random walk models in predicting the Deutsche mark against the US dollar exchange rate. Kuan and Liu [8] use both feedforward and recurrent neural networks to forecast five foreign exchange rates of the British pound, the Canadian dollar, the Deutsche mark, the Japanese yen, and the Swiss franc against the US dollar. They find that neural networks are able to improve the sign predictions and its forecasts are always better than the random walk forecasts. Hann and Steurer [9] make comparisons between the neural network and the linear model in US dollar against the Deutsche mark forecasting. They report that if weekly data are used, neural network is much better than both the monetary and random walk models.

Hawley et al. [10] provide an overview of the neural networks in the field of finance. Wu [11] conducts a comparative study between neural network and ARIMA model in forecasting the Taiwan/US dollar exchange rate. His findings show that neural network produces significantly better results than the best ARIMA model. Gencay [12] compares the performance of neural networks with those of random walk and GARCH models in forecasting daily spot exchange rates for the British pound, the Deutsche mark, the French franc, the Japanese yen, and the Swiss franc. He finds that forecasts generated by neural networks are superior to those of random walk and GARCH models. Brooks [13] document some predictability of daily exchange rates using artificial neural networks.

Leung et al. [14] compare the forecasting accuracy of multilayer perceptron with the general regression neural networks (GRNNs). They show that the GRNN possessed a greater forecasting strength relative to MLFN with respect to a variety of currency exchanges. Santos et al. [15] investigate the hypothesis that the nonlinear mathematical models of multilayer perceptron and the radial basis function neural networks are able to provide a more accurate out-of-sample forecast than the traditional linear models. Their results indicate that ANNs perform better than their linear models. Panda and Narasimhan [6] compare the forecasting accuracy of neural network with those of linear autoregressive and random walk models in prediction of weekly Indian rupee against the US dollar exchange rate. They report that neural network is much better than both the linear autoregressive and random walk models in out-of-sample forecasting.

One of the major developments in neural networks over the last decade is the model combining or ensemble modeling. The basic idea of this multi-model approach is the use of each component model's unique capability to better capture different patterns in the data. Both theoretical and empirical findings have suggested that combining different models can be an effective way to improve the predictive performance of each individual model, especially when the models in the ensemble are quite different. In addition, since it is difficult to completely know the characteristics of the data in a real problem, hybrid methodology that has both linear and nonlinear modeling capabilities can be a good strategy for practical use.

In this paper, the autoregressive integrated moving average models are applied to construct a new hybrid model in order to yield more accurate model than artificial neural networks. In our proposed model, the future value of a time series is considered as nonlinear function of several past observations and random errors. Therefore, in the first phase, an autoregressive integrated moving average model is used in order to generate the necessary data from under study time series. Then, in the second phase, a neural network is used to model the generated data by ARIMA model, and predict the future values of time series.

The weekly Indian rupee against the United States dollar exchange rate is used in this paper in order to show the appropriateness and effectiveness of the proposed model for exchange rate forecasting in comparing with the artificial neural networks (ANNs), the linear autoregressive (LAR) and the random walk (RW) models. The rest of the paper is organized as follows. In the next section, the literature survey of the hybrid models for time series forecasting is briefly reviewed. The basic concepts of autoregressive integrated moving average (ARIMA) and artificial neural networks (ANNs) are presented in section 3. In Section 4, the formulation of the proposed model is introduced. In Section 5, description of used data in this paper (weekly Indian rupee against the United States dollar exchange rate) is presented. In Section 6, the proposed model is applied to weekly Indian rupee against the United States dollar exchange rate forecasting and its performance is compared with those of other models in section 7. Section 8 concludes the paper with policy implications.

2. Hybrid models for time series forecasting

In the literature, so many different combination techniques have been proposed in order to overcome the deficiencies of single models and yield results that are more accurate. The difference between these combination techniques can be described using terminology developed for the classification and neural network literature. Hybrid models can be homogeneous, such as using differently configured neural networks (all multi-layer perceptrons), or heterogeneous, such as with both linear and nonlinear models [16]. In a competitive architecture, the aim is to build appropriate modules to represent different parts of the time series, and to be able to switch control to the most appropriate. For example, a time series may exhibit nonlinear behavior generally, but this may change to linearity depending on the input conditions. Early work on threshold autoregressive models (TAR) used two different linear AR processes, each of which change control among themselves according to the input values [17]. An alternative is a mixture density model, also known as nonlinear gated expert, which comprises neural networks integrated with a feedforward gating network [16].

In a cooperative modular combination, the aim is to combine models to build a complete picture from a number of partial solutions. The assumption is that a model may not be sufficient to represent the complete behavior of a time series, for example, if a time series exhibits both linear and nonlinear patterns during the same time interval, neither linear models nor nonlinear models alone are able to model both components simultaneously. A good exemplar is models that fuse autoregressive integrated moving average with artificial neural networks. An autoregressive integrated moving average (ARIMA) process combines three different processes comprising an autoregressive (AR) function regressed on past values of the process, moving average (MA) function regressed on a purely random process, and an integrated (I) part to make the data series stationary by differencing. In such hybrids, whilst the neural network model deals with nonlinearity, the autoregressive integrated moving average model deals with the non-stationary linear component [18].

The literature on this topic has dramatically expanded since the early work of Reid [19] in 1968, and Bates and Granger [20]. Clemen [21] provided a comprehensive review and annotated bibliography in this area. Wedding and Cios [22] described a combining methodology using radial basis function networks (RBF) and the Box–Jenkins ARIMA models. Luxhoj et al. [23] presented a hybrid econometric and ANN approach for sales forecasting. Pelikan et al. [24], and Ginzburg and Horn [25] proposed to combine several feedforward neural networks in order to improve time series forecasting accuracy. Tsaih et al. [26] presented a hybrid artificial intelligence (AI) approach that integrated the rule-based systems technique and neural networks to S&P 500 stock index prediction. Voort et al. [27] introduced a hybrid method called KARIMA using a Kohonen self-organizing map and autoregressive integrated moving average method for short-term prediction. Medeiros and Veiga [28] consider a hybrid time series forecasting system with neural networks used to control the time-varying parameters of a smooth transition autoregressive model.

In recent years, more hybrid forecasting models have been proposed, using autoregressive integrated moving average and artificial neural networks and applied to time series forecasting with good prediction performance. Pai and Lin [29] proposed a hybrid methodology to exploit the unique strength of ARIMA models and Support Vector Machines (SVMs) for stock prices forecasting. Chen and Wang [30] constructed a combination model incorporating seasonal autoregressive integrated moving average (SARIMA) model and SVMs for seasonal time series forecasting. Zhou and Hu [31] proposed a hybrid modeling and forecasting approach based on Grey and Box–Jenkins autoregressive moving average (ARMA) models. Armano et al. [32] presented a new hybrid approach that integrated artificial neural network with genetic algorithms (GAs) to stock market forecast. Khashei and Bijri [33] based on the basic concepts of probabilistic neural classifiers, proposed a new class of hybrid models for time series forecasting.

Khashei et al. [34] presented a hybrid ARIMA and artificial intelligence approaches to financial markets prediction. Yu et al. [35] proposed a novel nonlinear ensemble forecasting model integrating generalized linear auto regression (GLAR) with artificial neural networks in order to obtain accurate prediction in foreign exchange market. Kim and Shin [36] investigated the effectiveness of a hybrid approach based on the artificial neural networks for time series properties, such as the adaptive time delay neural networks (ATNNs) and the time delay neural networks (TDNNs), with the genetic algorithms in detecting temporal patterns for stock market prediction tasks. Tseng et al. [37] proposed using a hybrid model called SARIMABP that combines the seasonal autoregressive integrated moving average (SARIMA) model and the back-propagation neural network model to predict seasonal time series data. Khashei et al. [38] based on the basic concepts of artificial neural networks, proposed a new hybrid model in order to overcome the data limitation of neural networks and yield more accurate forecasting model, especially in incomplete data situations.

3. Artificial neural networks (ANNs) and Autoregressive integrated moving average (ARIMA) models

In this section, the basic concepts and modeling approaches of the artificial neural networks (ANNs) and autoregressive integrated moving average (ARIMA) models for time series forecasting are briefly reviewed.

3.1. Artificial neural networks (ANNs) models

Recently, computational intelligence systems and among them Artificial Neural Networks (ANNs), which in fact are model free dynamics, has been used widely for approximation functions and forecasting. One of the most significant advantages of the ANN models over other classes of nonlinear models is that ANNs are universal approximators that can approximate a large class of functions with a high degree of accuracy [39]. Their power comes from the parallel processing of the information from the data. No prior assumption of the model form is required in the model building process. Instead, the network model is largely determined by the characteristics of the data. Single hidden layer feed forward network is the most widely used model form for time series modeling and forecasting. The model is characterized by a network of three layers of simple processing units connected by acyclic links (Fig. 1). The relationship between the output (y_t) and the inputs ($y_{t-1},\ldots,y_{t-p}$) has the following mathematical representation:

$$y_t = w_0 + \sum_{j=1}^{q} w_j \cdot g(w_{0,j} + \sum_{i=1}^{p} w_{i,j} \cdot y_{t-i}) + \varepsilon_t, \qquad (1)$$

where, $w_{i,j}\left(i = 0,1,2,\ldots,p,\;\; j = 1,2,\ldots,q\right)$ and $w_j\left(j = 0,1,2,\ldots,q\right)$ are model parameters often called connection weights; p is the number of input nodes; and q is the number of hidden nodes. Activation functions can take several forms. The type of activation function is indicated by the situation of the neuron within the network. In the majority of cases input layer neurons do not have an activation function, as their role is to transfer the inputs to the hidden layer. The most widely used activation function for the output layer is the linear function as non-linear activation function may introduce distortion to the predicated output. The logistic function is often used as the hidden layer transfer function that are shown in Eq. 2. Other activation functions can also be used such as linear and quadratic, each with a variety of modeling applications.

$$Sig(x) = \frac{1}{1+\exp(-x)}. \qquad (2)$$

Hence, the ANN model of (1), in fact, performs a nonlinear functional mapping from past observations to the future value y_t, i.e.,

$$y_t = f\left(y_{t-1},\ldots,y_{t-p},w\right)+\varepsilon_t, \qquad (3)$$

where, w is a vector of all parameters and *f(.)* is a function determined by the network structure and connection weights. Thus, the neural network is equivalent to a nonlinear autoregressive model. The simple network given by (1) is surprisingly powerful in that it is able to approximate the arbitrary function as the number of hidden nodes when q is sufficiently large. In practice, simple network structure that has a small number of hidden nodes often works well in out-of-sample forecasting. This may be due to the overfitting effect typically found in the neural network modeling process. An overfitted model has a good fit to the sample used for model building but has poor generalizability to data out of the sample [40].

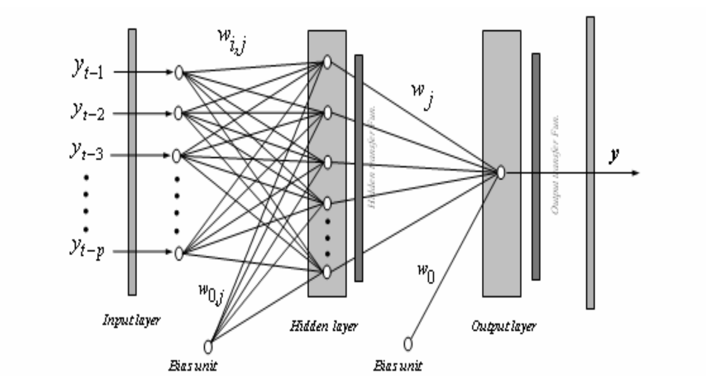

Fig. 1: Neural network structure (N (p-q-1).).

The choice of q is data-dependent and there is no systematic rule in deciding this parameter. In addition to choosing an appropriate number of hidden nodes, another important task of ANN modeling of a time series is the selection of the number of lagged observations, p, and the dimension of the input vector. This is perhaps the most important parameter to be estimated in an ANN model because it plays a major role in determining the (nonlinear) autocorrelation structure of the time series.

3.2. The autoregressive integrated moving average models

For more than half a century, autoregressive integrated moving average (ARIMA) models have dominated many areas of time series forecasting. In an ARIMA (p,d,q) model, the future value of a variable is assumed to be a linear function of several past observations and random errors. That is, the underlying process that generates the time series with the mean μ has the form:

$$\phi(B)\nabla^{d}(y_t-\mu)=\theta(B)a_t \tag{4}$$

where, y_t and a_t are the actual value and random error at time period t, respectively; $\phi(B)=1-\sum_{i=1}^{p}\varphi_i B^i$, $\theta(B)=1-\sum_{j=1}^{q}\theta_j B^j$ are polynomials in B of degree p and q, $\phi_i \quad (i=1,2,...,p)$ and $\theta_j \quad (j=1,2,...,q)$ are model parameters, $\nabla=(1-B)$, B is the backward shift operator, p and q are integers and often referred to as orders of the model, and d is an integer and often referred to as order of differencing. 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 Box–Jenkins [41] methodology includes three iterative steps of model identification, parameter estimation, and diagnostic checking. The basic idea of model identification is that if a time series is generated from an ARIMA process, it should have some theoretical autocorrelation properties. By matching the empirical autocorrelation patterns with the theoretical ones, it is often possible to identify one or several potential models for the given time series. Box and Jenkins [41] proposed to use the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the sample data as the basic tools to identify the order of the ARIMA model. Some other order selection methods have been proposed based on validity criteria, the information-theoretic approaches such as the Akaike's information criterion (AIC) [42] and the minimum description length (MDL) [43- 45]. In addition, in recent years different approaches based on intelligent paradigms, such as neural networks [46], genetic algorithms [47, 48] or fuzzy system [49] have been proposed to improve the accuracy of order selection of ARIMA models.

In the identification step, data transformation is often required to make the time series stationary. Stationarity is a necessary condition in building an ARIMA model used for forecasting. 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, differencing and power transformation are applied to the data to remove the trend and to stabilize the variance before an ARIMA model can be fitted. Once a tentative model is identified, estimation of the model parameters is straightforward. The parameters are estimated such that an overall measure of errors is minimized. This can be accomplished using a nonlinear optimization procedure. The last step in model building is the diagnostic checking of model adequacy. This is basically to check 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 will again be followed by the steps of parameter estimation and model verification. Diagnostic information may help suggest alternative model(s). This three-step model building process is typically repeated several times until a satisfactory model is finally selected. The final selected model can then be used for prediction purposes.

4. Formulation of the proposed model

Despite the numerous time series models available, the accuracy of time series forecasting currently is fundamental to many decision processes, and hence, never research into ways of improving the effectiveness of forecasting models been given up. Many researches in time series forecasting have argued that predictive performance improves in combined models. In hybrid models, the aim is to reduce the risk of using an inappropriate model by combining several models to reduce the risk of failure and obtain results that are more accurate. Typically, this is done because the underlying process cannot easily be determined. The motivation for combining models comes from the assumption that either one cannot identify the true data generating process or that a single model may not be sufficient to identify all the characteristics of the time series.

In this paper, a novel hybrid model of artificial neural networks is proposed in order to yield more accurate results using the autoregressive integrated moving average models. In our proposed model, based on the Box–Jenkins [56] methodology in linear modeling, a time series is considered as nonlinear function of several past observations and random errors as follows:

$$y_t = f[(z_{t-1}, z_{t-2}, \ldots, z_{t-m}), (e_{t-1}, e_{t-2}, \ldots, e_{t-n})] \tag{5}$$

where f is a nonlinear function determined by the network, $z_t = (1-B)^d (y_t - \mu)$, e_t is the residual at time t and m and n are integers. So, in the first stage, an autoregressive integrated moving average model is used in order to generate the residuals (e_t).

In second stage, a neural network is used in order to model the nonlinear and linear relationships existing in residuals and original data. Thus,

$$z_t = w_0 + \sum_{j=1}^{Q} w_j \cdot g(w_{0,j} + \sum_{i=1}^{p} w_{i,j} \cdot z_{t-i} + \sum_{i=p+1}^{p+q} w_{i,j} \cdot e_{t+p-i}) + \varepsilon_t , \tag{6}$$

where, $w_{i,j}(i = 0,1,2,\ldots, p+q,\ \ j = 1,2,\ldots,Q)$ and $w_j(j = 0,1,2,\ldots,Q)$ are connection weights; p, q, Q are integers, which are determined in design process of final neural network.

It must be noted that any set of above–mentioned variables $\{\ e_i (i = t-1,\ldots,t-n)\}$ or $\{\ z_i (i = t-1,\ldots,t-m)\}$ may be deleted in design process of final neural network. This maybe related to the underlying data generating process and the existing linear and nonlinear structures in data. For example, if data only consist of pure nonlinear structure, then the residuals will only contain the nonlinear relationship. Because the

ARIMA is a linear model and does not able to model nonlinear relationship. Therefore, the set of residuals $\{ e_i (i = t-1,...,t-n)\}$ variables maybe deleted against other of those variables.

As previously mentioned, in building autoregressive integrated moving average as well as artificial neural networks models, subjective judgment of the model order as well as the model adequacy is often needed. It is possible that suboptimal models will be used in the hybrid model. For example, the current practice of Box–Jenkins methodology focuses on the low order autocorrelation. A model is considered adequate if low order autocorrelations are not significant even though significant autocorrelations of higher order still exist. This suboptimality may not affect the usefulness of the hybrid model. Granger [50] has pointed out that for a hybrid model to produce superior forecasts, the component model should be suboptimal. In general, it has been observed that it is more effective to combine individual forecasts that are based on different information sets [50].

5. Description of data

In this paper, the weekly spot rates of Indian rupee against the United States dollar exchange rate [6] is used in order to show the appropriateness and effectiveness of the proposed model for exchange rate forecasting. The data set is from FX database for the period of January 6, 1994–July 10, 2003, for a total 497 observations. The weekly returns are calculated as the log differences of the levels. Let p_t be the exchange rate price for the period t. Then the exchange rate return at time t is calculated as follows:

$$y_t = (\log(p_t) - \log(p_{t-1})) \times 100 \tag{7}$$

The log difference is multiplied by one hundred in order to reduce round-off errors. The summary statistics of the data are presented in Table 1. Both the skewness and kurtosis are substantially high. The kurtosis coefficient, i.e. 42.5818, is larger than that of the standard normal distribution (which is equal to 3), which in turn indicates the leptokurtosis of INR/USD return series. Jarque-Bera statistic also rejects the normality of exchange rate, which is common in high frequency financial time series data. The first 20 autocorrelations are calculated but only five autocorrelations i.e. ρ1, ρ5, ρ10, ρ15, ρ20 are reported in the table. The series shows evidence of autocorrelation. The Ljung-Box (LB) statistic for the first 20 lags is 42.34, which rejects the hypothesis of identical and independent observations. The value for Augmented Dicky-Fuller (ADF) test statistic confirms the stationarity of weekly exchange rate return series.

Table 1

Summary statistics for the weekly exchange rates: log first difference January 6, 1994–July 10, 2003

Description	*INR/USD*
Sample size	496
Mean	0.0455
Median	0.0052
S.D.	0.2806
Skewness	2.4834
Kurtosis	42.5818
Maximum	3.1904
Minimum	−1.7316
Jarque-Bera	32888 (0.000)
ρ1	0.131
ρ5	−0.002
ρ10	−0.008
ρ15	0.080
ρ20	−0.059
LB statistic (20)	42.34 (0.000)
ADF	−9.348

Note: The ρ-value for Jarque-Bera and LB statistic is given in parentheses. The

MacKinnon critical values for ADF test are −3.98, −3.42, and −3.13 at 1%, 5%, and 10% significance level, respectively.

6. Application of the proposed model to exchange rate forecasting

In this section, the procedure of the proposed model for Indian rupee against the United States dollar exchange rate is illustrated. Only the one-step-ahead forecasting is considered. To assess the forecasting performance, the Indian rupee against the United States dollar exchange rate data set is divided into two samples of training and testing. The training data set, 350 observations (i.e. in-sample data), is exclusively used in order to formulate the models and then the test sample, the last 146 observations (i.e. out-of-sample data), is used in order to evaluate the performance of the established models.

Stage I : Using the Eviews package software, the best-fitted model is a autoregressive model of order two, AR (2).

Stage II : In order to obtain the optimum network architecture, based on the concepts of artificial neural networks design and using pruning algorithms in MATLAB 7 package software, different network architectures are evaluated to compare the ANNs performance. The best fitted network which is selected, and therefore, the architecture which presented the best forecasting accuracy with the test data, is composed of eight inputs, four hidden and one output neurons (in abbreviated form, N(8-4-1)). The structure of the best-fitted network is shown in Fig. 2. The performance measures of the proposed model for train and test data sets are given in Table 2.

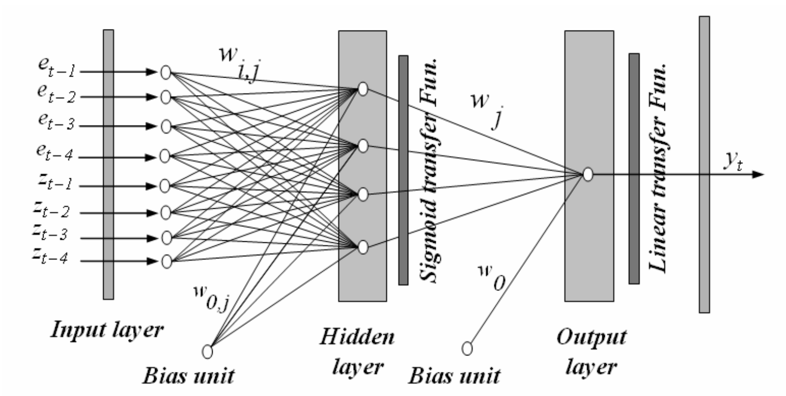

Fig. 2: Structure of the best fitted network, N(8-4-1).

Table 2

Performance measures of the proposed model.

Train					Test				
RMSE	*MAE*	*CORR*	*DA*	*SIGN*	*RMSE*	*MAE*	*CORR*	*DA*	*SIGN*
0.1590	0.0894	0.8219	0.8485	0.6844	0.0929	0.0632	0.2418	0.8276	0.6232

7. Comparison with other models

In this section after selecting the appropriate models for the proposed model, the predictive capabilities of the proposed model for in-simple and out-of-simple data are compared with the artificial neural networks, the linear autoregressive, and the random walk models in Indian rupee against the United States dollar exchange rate forecasting. Five performance criteria such as root mean square (RMSE), mean absolute error (MAE), Pearson correlation coefficient (CORR), direction accuracy (DA)1 and sign predictions (SIGN)2 are employed in order to evaluate the predictive power of the proposed model in comparison with those other forecasting models.

---

1 $$DA = \frac{1}{N}\sum_{i=1}^{N} a_i \quad where \quad a_i = \begin{cases} 1 & if \quad (y_{i+1} - y_i)(\hat{y}_{i+1} - y_i) > 0 \\ 0 & if \quad otherwise \end{cases}$$

2 $$SING = \frac{1}{N}\sum_{i=1}^{N} b_i \quad where \quad b_i = \begin{cases} 1 & if \quad y_{i+1} \cdot \hat{y}_{i+1} > 0 \\ 0 & if \quad otherwise \end{cases}$$

7.1. In-sample forecasts

In-sample performance of the proposed model, the neural network, the linear autoregressive, and the random walk models is presented in Table 3. The results show that the proposed model outperforms the neural network, the linear autoregressive, and the random walk models by all the evaluation criteria. The RMSE of the proposed model i.e. 0.1590 is significantly lower than the RMSEs of the neural network, the linear autoregressive, and the random walk models which are equal to 0.2048, 0.2518, and 0.3325, respectively. The proposed model has also got smaller values for MAE as compared to the values of the neural network, the linear autoregressive and the random walk models. The proposed model fitted values have higher correlation, which is equal to 0.8219, with the actual series as compared to the values of the neural network (0.6687), the linear autoregressive (0.4062), and the random walk (0.2720) models. The most important evaluation criteria, from investor's point of view, are direction accuracy and sign prediction. Because an investor is more enthusiastic to know about the directional and sign change in the exchange rate return for tomorrow rather than the exact magnitude of it. In the Table 3, it can be seen that the proposed model gives better sign predictions than the neural network, the linear autoregressive, and the random walk models. It can also be seen that direction accuracy of the proposed model, which is equal to 0.8485, is higher than the corresponding values of the neural network (0.6848), the linear autoregressive (0.6424). Here it should be noted that the direction accuracy is zero for random walk. This is because, by definition, the random walk has absolutely no ability to predict whether the exchange rate return will go up or down since, it simply takes the exchange rate return of current period as the forecast of the next period and predicts no change.

Table 3

In-sample performance of the proposed, neural network, linear autoregressive and random walk models on exchange rate data.

	proposed	artificial neural	linear	random walk

	model	network	autoregressive	
RMSE	0.1590	0.2048	0.2518	0.3325
MAE	0.0894	0.1057	0.1341	0.1577
CORR	0.8219	0.6687	0.4062	0.2720
DA	0.8485	0.6848	0.6424	0.0000
SIGN	0.6844	0.5181	0.5030	0.5030

7.2. Out-of-sample forecasts

The results for out-of-sample performance are presented in Table 4. The out-of-sample forecasts of the proposed model are also more accurate than the neural network, the linear autoregressive, and the random walk forecasts by all criteria, except for the Pearson correlation coefficient of the random walk model. The correlation coefficient for the proposed model is 0.2418, as compared with the corresponding figure of random walk, which is equal to 0.2535, is slightly lower. However, the superiority of the proposed model over random walk in out-of-sample forecasts provides evidence against the efficient market hypothesis.

Table 4

Out-of-sample performance of the proposed, artificial neural network, linear autoregressive and random walk models on weekly exchange rate data.

	proposed model	artificial neural network	linear autoregressive	random walk
RMSE	0.0929	0.1087	0.1096	0.1342
MAE	0.0632	0.0676	0.0740	0.0854
CORR	0.2418	0.1851	0.2201	0.2535
DA	0.8276	0.7063	0.5793	0.0000
SIGN	0.6232	0.5714	0.6031	0.5634

8. Summary and policy implications

Applying quantitative methods for forecasting and assisting investment decision making has become more indispensable in business practices than ever before. In this paper, a new hybrid artificial neural network and autoregressive integrated moving average model is proposed as an alternative forecasting technique to nonlinear neural networks and linear autoregressive and random walk models in the one-step-ahead prediction of weekly Indian rupee/US dollar exchange rate. Empirical results indicate that our proposed model has superior in- and out-of simple forecast than the neural network, the linear autoregressive, and the random walk models in all evaluation criteria, except for Pearson correlation coefficient of the random walk model in out-of simple.

The findings of this paper have following policy implications. First, the superiority of the proposed model over random walk model suggests that the foreign exchange market is not efficient. There always exists a possibility of forecasting exchange rate. The better forecasting or understanding of the movements of exchange rate by using hybrid linear and nonlinear forecasting models may help the policy makers to conduct a suitable monetary policy, which will in turn achieve its desired objectives of price stability and higher economic activity. Second, the better forecast of exchange rate by our proposed model over linear autoregressive model recommends that the linear unpredictability of exchange rate can be improved. Third, the superiority of the proposed model over both linear autoregressive and nonlinear neural network models show that the information that are hidden in exchange rate can be better extracted by using such hybrid models. This finding also indicates that the exchange market participants' expectations are better modeled by our proposed model. Therefore, the use of the proposed model may help policy makers in extracting useful information about the economic and financial conditions. Finally, firms or investors in the foreign exchange market can usefully apply our proposed model in predicting exchange rate for the future and thus can take a profitable trading strategy and a proper decision on asset allocation.

REFERENCES

[1] Y. Hong, H. Li, F. Zhao, Can the random walk model be beaten in out-of-sample density forecasts? Evidence from intraday foreign exchange rates, Journal of Econometrics 141, (2007), 736-776.

[2] A. Trapletti, A. Geyer, F. Leisch, Forecasting exchange rates using cointegration models and intra-day data, Journal of Forecasting 21, (2002), 151– 166.

[3] L. Kilian, M.P. Taylor, Why is it so difficult to beat random walk forecast of exchange rates?, Journal of International Economics 60, (2003), 85– 107.

[4] A. Preminger, R. Franck, Forecasting exchange rates: A robust regression approach, International Journal of Forecasting 23, (2007), 71– 84.

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

[6] C. Panda, V. Narasimhan, Forecasting exchange rate better with artificial neural network, Journal of Policy Modeling 29, (2007), 227–236.

[7] A. S. Weigend, D. E. Rumelhart, B. A. Huberman, Generalization by weight-elimination with application to forecasting, Advances in Neural Information Processing Systems 3, (1991), 875-882.

[8] C. M. Kuan, T. Liu, Forecasting exchange rates using feedforward and recurrent neural networks, Journal of Applied Econometrics 10, (1995), 347-364.

[9] T. H. Hann, E. Steurer, Much ado about nothing? Exchange rate forecasting: neural networks vs. linear models using monthly and weekly data, Neurocomputing 10, (1996), 323-339.

[10] D. Hawley, J. Johnson, D. Raina, Artificial neural systems: a new tool for financial decision-making, Financial Analysts Journal, (1990), 63-72.

[11] B. Wu, Model-free forecasting for nonlinear time series (with application to exchange rates), Computational Statistics and Data Analysis19, (1995), 433-459.

[12] R. Gencay, Linear, non-linear, and essential foreign exchange rate prediction with simple technical trading rules, Journal of International Economics 47, (1999), 91–107.

[13] C. Brooks, Testing for non-linearity in daily sterling exchange rates, Applied Finance Economics 6, (1996), 307–317.

[14] M.T. Leung, A. S. Chen, H. Daouk, Forecasting exchange rates using general regression neural networks, Computers &Operations Research27, (2000), 1093–110.

[15] A. Santos, N. da Costa Jr, L. Coelho, Computational intelligence approaches and linear models in

case studies of forecasting exchange rates, Expert Systems with Applications 33, (2007), 816–823.

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

[17] H. Tong, K.S. Lim, Threshold autoregressive, limit cycles and cyclical data, Journal of the Royal Statistical Society Series B 42 (1980), 245–292.

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

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

[20] J.M. Bates, W.J. Granger, The combination of forecasts, Operation Research 20, (1969), 451–468.

[21] R. Clemen, Combining forecasts: a review and annotated bibliography with discussion , International Journal of Forecasting 5, (1989), 559–608.

[22] D.K. Wedding, K.J. Cios, Time series forecasting by combining RBF networks, certainty factors, and the Box–Jenkins model, Neurocomputing 10, (1996), 149–168.

[23] J.T. Luxhoj, J.O. Riis, B. Stensballe, A hybrid econometric-neural network modeling approach for sales forecasting, Int. J. Prod. Econ. 43, (1996), 175–192.

[24] E. Pelikan, C. de Groot, D. Wurtz, Power consumption in West-Bohemia: improved forecasts with decorrelating connectionist networks, Neural Network World 2, (1992), 701–712.

[25] I. Ginzburg, D. Horn, Combined neural networks for time series analysis, Adv. Neural Inf. Process.Systems 6, (1994), 224–231.

[26] R. Tsaih, Y. Hsu, C.C. Lai, Forecasting S&P 500 stock index futures with a hybrid AI system, Decision Support Systems 23, (1998), 161–174.

[27] M.V.D. Voort, M. Dougherty, S. Watson, Combining Kohonen maps with ARIMA time series models to forecast traffic flow, Transportation Research Part C: Emerging Technologies 4, (1996), 307–318.

[28] M.C. Medeiros, A. Veiga, A hybrid linear-neural model for time series forecasting, IEEE Transaction on Neural Networks 11, (2000), 1402–1412.

[29] P.F. Pai, C.S. Lin, A hybrid ARIMA and support vector machines model in stock price forecasting, Omega 33, (2005), 497 – 505.

[30] K.Y. Chen, C.H. Wang, A hybrid SARIMA and support vector machines in forecasting the

production values of the machinery industry in Taiwan, Expert Systems with Applications 32, (2007), 254–264.

[31] Z.J. Zhou, C.H. Hu, An effective hybrid approach based on grey and ARMA for forecasting gyro drift, Chaos, Solitons and Fractals 35, (2008), 525–529.

[32] G. Armano, M. Marchesi, A. Murru, A hybrid genetic-neural architecture for stock indexes forecasting, Information Sciences 170, (2005), 3- 33.

[33] M. Khashei, M. Bijari, A new class of hybrid models for time series forecasting, Expert Systems with Applications 39, (2012), 4344–4357.

[34] M. Khashei, M. Bijari, GH A. Raissi, Improvement of Auto-Regressive Integrated Moving Average Models Using Fuzzy Logic and Artificial Neural Networks (ANNs) , Neurocomputing 72, (2009), 956- 967.

[35] L. Yu, S. ang, 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.

[36] H. Kim, K. Shin, A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets, Applied Soft Computing 7, (2007), 569–576.

[37] F.M. Tseng, H.C. Yu, G.H. Tzeng, Combining neural network model with seasonal time series ARIMA model, Technological Forecasting & Social Change 69, (2002), 71–87.

[38] M. Khashei, S. R. Hejazi, M. Bijari, A new hybrid artificial neural networks and fuzzy regression model for time series forecasting, Fuzzy Sets and Systems 159, (2008), 769 – 786.

[39] G.P. Zhang, G. Min Qi, Neural network forecasting for seasonal and trend time series, European Journal of Operational Research 160, (2005), 501–514.

[40] H. Demuth, B. Beale, Neural Network Toolbox User Guide, The Math works Inc, Natick, 2004.

[41] P. Box, G.M. Jenkins, Time Series Analysis: Forecasting and Control, Holden-day Inc, San Francisco, CA, 1976.

[42] R. Shibata, Selection of the order of an autoregressive model by Akaike's information criterion, Biometrika AC-63, (1976), 117–126.

[43] R.H. Jones, Fitting autoregressions, J. Amer. Statist. Assoc. 70, (1975), 590–592.

[44] C. M. Hurvich, C. L. Tsai, Regression and time series model selection in small samples, Biometrica 76, (1989), 297–307.

[45] L. Ljung, System Identification Theory for the User, Prentice-Hall, Englewood Cliffs, NJ, 1987.

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

[47] T. Minerva, I. Poli, Building ARMA models with genetic algorithms, in: Lecture Notes in Computer Science, (2001), 335–342.

[48] C. S. Ong, J. J. Huang, G. H. Tzeng, Model identification of ARIMA family using genetic algorithms, Appl. Math. Comput. 164, (2005), 885–912.

[49] M. Haseyama, H. Kitajima, An ARMA order selection method with fuzzy reasoning, Signal Process 81, (2001), 1331–1335.

[50] C.W.J. Granger, Combining forecasts—Twenty years later, J. Forecasting 8, (1989), 167–173.