An Enhanced Neural-based Bi-Component Hybrid Model for Foreign Exchange Rate Forecasting

M. Khashei¹, S. Torbat¹, Z. H. Rahimi¹

¹Department of Industrial and Systems Engineering, Isfahan University of Technology (IUT), Isfahan, Iran

Abstract

Foreign exchange rates are among the most important economic indices in the international monetary markets. For large multinational firms, which conduct substantial currency transfers in the course of business, being able to accurately forecast movements of currency exchange rates can result in substantial improvement in the overall profitability of the firm. 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. Hybrid techniques that decompose a time series into its linear and nonlinear components are one of the most popular hybrid models categories, which have been shown to be successful for single models. However, they have yielded mixed results in some situations in comparison with components models used separately; and hence, it is not wise to apply them blindly to any type of data. In this paper, an enhanced version of hybrid neural based models is proposed, incorporating the autoregressive integrated moving average (ARIMA) and artificial neural networks (ANNs) for financial time series forecasting. In proposed model, in contrast to the traditional hybrid ARIMA/ANNs, it can be guaranteed that the performance of the proposed model will not be worse than either of the components used separately. In additional, empirical results in exchange rate forecasting indicate that the proposed model can be an effective way to improve forecasting accuracy achieved by traditional hybrid ARIMA/ANNs models. Therefore, it can be used as an appropriate alternative for exchange rate forecasting, especially when higher forecasting accuracy is needed.

© 2017 Forecast Research Laboratory. All rights reserved.

1. Introduction

Several forecasting models have been proposed and applied to exchange rate forecasting. In general, there are more than three hundred different kinds of forecasting models in modern time, which can be divided into two main categories, qualitative and quantitative models [1]. Time series models are one of the most important kinds of quantitative models in exchange rate forecasting. In these models, historical observations of the same variable are collected and analysed to develop a model that captures the underlying data generating process. Then the model is used to predict the future. This modelling approach is particularly useful when little knowledge is available on the underlying data generating process or when there is no satisfactory explanatory model that relates the prediction variable to other explanatory variables [2].

Several different time series models have been proposed and applied in foreign exchange rate forecasting, which are generally divided in two categories, linear and nonlinear models. Autoregressive integrated moving average (ARIMA) models are one of the most popular linear time series models, which are widely used in foreign exchange
rate forecasting. The popularity of the ARIMA model is due to its statistical properties as well as to the well-known Box–Jenkins methodology [3] in the model building process. Although ARIMA models are quite flexible and also have the advantages of accurate forecasting over a short period of time and ease of implementation, these models have a critical limitation that detract from their popularity for financial markets forecasting, especially exchange rate forecasting. These models assume that the future values of a time series have a linear relationship with current and past values as well as with white noise. Because of the linear limitation, the approximation of autoregressive integrated moving average models may be totally inappropriate if the underlying mechanism is nonlinear [4].

However, it is widely agreed that exchange rate movements are nonlinear and many researchers have confirmed the linear unpredictability of exchange rates [5]. Therefore, the efforts of researches have been devoted to explore the nonlinearity of the exchange rate data and develop the nonlinear models to improve exchange rate forecasting.

In the literature, several parametric nonlinear forecasting models such as autoregressive conditional heteroscedasticity (ARCH), general autoregressive conditional heteroscedasticity (GARCH), chaotic dynamics, vector autoregressive (VAR), and self-exiting threshold autoregressive models have been proposed and applied to foreign exchange rate forecasting. Some recent works by these models are summarized in Table 1. However, the literature indicates that although these models may be good for a particular situation, they perform poorly for other applications. The pre-specification of the model form restricts the usefulness of these parametric nonlinear models since there can be many other possible nonlinear patterns to be considered. One particular nonlinear specification will not be general enough to capture all the nonlinearities in the data.

A number of nonparametric forecasting models have also been proposed to forecast exchange rates. Artificial neural networks (ANNs) are one of the most popular nonlinear nonparametric 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 [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 [7]. 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 [8]. Yao and Tan [9] have reported empirical evidence that a neural network model is applicable to the prediction of foreign exchange rates including Japanese Yen, Deutsch Mark, British Pound, Swiss Franc and Australian Dollar all against the American Dollar.

[10] has investigated the predictability of spot foreign exchange rate returns rules from past buy-sell signals of the simple technical trading rules by using the nearest neighbours and the feed forward network regressions. The experimental results of this work indicate that simple technical rules provide significant forecast improvements for the current returns over the random walk model. [11] has found that neural networks are better than random walk models in predicting the Deutsche mark against the US dollar exchange rate. [12] uses 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. [13] makes 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. [14] 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 traditional linear models in exchange rate forecasting. Some other works in this field are summarized in Table 2.

Although ANNs have the advantages of accurate forecasting, their performance in some specific situation is inconsistent. In the literature, several papers are devoted to comparing ANNs with traditional methods in exchange rate forecasting. Despite the numerous studies, which have shown ANNs are significantly better than the conventional linear models and their forecast considerably and consistently more accurately, some other studies have reported inconsistent results [8]. Hann and Steurer make one of the first comparisons between neural networks and linear models in exchange rate forecasting. They report that if monthly data are used, neural networks do not show much improvement over linear models [13]. [15] compares the performance of linear models with neural networks. Their results show that linear autoregressive models can outperform ANNs in some cases. [16] compares the performance of VAR with artificial neural network for exchange rate forecasting. Their results indicate that the ANN approach has superior performance of prediction capability than the VAR method. Some researchers try to answer this question that: under what conditions ANN forecasters can perform better than the linear time series forecasting methods such as Box- Jenkins models. Some researchers believe that in some specific situations where ANNs perform worse than
linear statistical models, the reason may simply be that the data is linear without much disturbance, therefore; cannot be expected that ANNs to do better than linear models for linear relationships. However, for any reason, using ANNs to model linear problems have yielded mixed results and hence; it is not wise to apply ANNs blindly to any type of data.

Table 1. Some published works of different classic time series models for exchange rate forecasting in the recent years

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Type</th>
<th>Model</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1- [17]</td>
<td>2016</td>
<td>Single</td>
<td>VAR</td>
<td>A panel VAR model is applied in order to study the behavior of U.S. listed currency hedged ETF investors towards changes in the underlying benchmark and foreign exchange rate.</td>
</tr>
<tr>
<td>2- [18]</td>
<td>2015</td>
<td>Single</td>
<td>VAR</td>
<td>A co-integrated VAR (CVAR) model is proposed in order to study the relationship between the real exchange rate (RER) and economic growth in China.</td>
</tr>
<tr>
<td>3- [19]</td>
<td>2014</td>
<td>Single</td>
<td>VAR</td>
<td>A panel vector autoregressive model (PVAR) is examined in order to study the dynamics of the overall exchange rate volatility, based on panel data for 29 economies.</td>
</tr>
<tr>
<td>4- [20]</td>
<td>2011</td>
<td>Single</td>
<td>VAR</td>
<td>This paper examines the ability of vector autoregressive (VAR) models to properly identify the transmission of monetary policy in a controlled experiment using the simulating data from a small open economy DSGE model.</td>
</tr>
<tr>
<td>5- [21]</td>
<td>2009</td>
<td>Single</td>
<td>VAR</td>
<td>This paper proposes to forecast exchange rates with a large Bayesian VAR (BVAR), using a panel of 33 exchange rates vis-a-vis the US Dollar.</td>
</tr>
<tr>
<td>6- [22]</td>
<td>2016</td>
<td>Single</td>
<td>GARCH</td>
<td>A GARCH based approach is applied to investigate the dynamic relationship between energy, stock and currency markets, using a sample of more than 10 years of daily return observations of the WTI crude oil, the Dow Jones Industrial average stock index and the trade weighted US dollar index returns.</td>
</tr>
<tr>
<td>7- [23]</td>
<td>2015</td>
<td>Single</td>
<td>GARCH</td>
<td>This paper examines the impact of exchange rate uncertainty on different components of net portfolio flows, namely net equity and net bond flows, as well as their dynamic linkages, using bilateral monthly data for the US vis-à-vis Australia, Canada, the euro area, Japan, Sweden, and the UK.</td>
</tr>
<tr>
<td>8- [24]</td>
<td>2013</td>
<td>Single</td>
<td>GARCH</td>
<td>A copula-based GARCH model is proposed in order to study and explore the dependence structure between the oil price and the US dollar exchange rate.</td>
</tr>
<tr>
<td>9- [25]</td>
<td>2015</td>
<td>Single</td>
<td>ARCH &amp; GARCH</td>
<td>This research article aimed at modeling the variations in the dollar/cedi exchange rate. It examines the applicability of a range of ARCH/GARCH specifications for modeling volatility of the series.</td>
</tr>
<tr>
<td>10- [26]</td>
<td>2014</td>
<td>Single</td>
<td>GARCH</td>
<td>This study compares in-sample forecasts from symmetric and asymmetric GARCH models with the implied volatility derived from currency options for four dollar parities.</td>
</tr>
<tr>
<td>11- [27]</td>
<td>2015</td>
<td>Single</td>
<td>STAR</td>
<td>A Smooth Transition Autoregressive (STAR) model is applied in order to forecast the real exchange rates of various OECD countries.</td>
</tr>
<tr>
<td>12- [28]</td>
<td>2009</td>
<td>Single</td>
<td>STAR</td>
<td>The main aim of this paper is to study the dynamics of the US real effective exchange rate by capturing non-linearity and long-memory features using the family of fractionally integrated STAR (FISTAR) models.</td>
</tr>
<tr>
<td>13- [29]</td>
<td>2006</td>
<td>Single</td>
<td>STAR</td>
<td>A Smooth Transition Autoregressive (STAR) model is proposed for exchange rate forecasting using quarterly data for the yen-based currencies of six major East Asian countries.</td>
</tr>
</tbody>
</table>
Both artificial neural networks and autoregressive integrated moving average models have achieved successes in their own linear or nonlinear domains. However, none of them is a universal model that is suitable for all circumstances. The approximation of ARIMA to complex nonlinear problems as well as ANNs to model linear problems may be totally inappropriate, and also, in problems that consist both linear and nonlinear correlation structures. Using hybrid models or combining several models has become a common practice in order to overcome the limitations of components models and improve the forecasting accuracy [40]. 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 modelling capabilities can be a good strategy for practical use.

The hybrid techniques that decompose a time series into its linear and nonlinear forms are one of the most popular hybrid models categories, which have been shown to be successful for single models. [41] presented a hybrid ARIMA and ANN approaches for time series forecasting using mentioned technique. In Zhang’s hybrid model, the linear ARIMA and the nonlinear multilayer perceptron models are jointly used in order to capture different forms of relationship in the time series data. His motivation to build the model comes from following perspectives. First, it is difficult to determine whether a time series under study is generated from a linear or nonlinear underlying process: thus, the problem of model selection can be eased by combining linear ARIMA and nonlinear ANN models. Second, real-world time series are rarely pure linear or nonlinear and often contain both linear and nonlinear patterns, which neither ARIMA nor ANN models alone can be adequate for modelling in such cases; hence the problem of modelling the combined linear and nonlinear autocorrelation structures in time series can be solved by combining linear ARIMA and nonlinear ANN models. Third, it is almost universally agreed in the forecasting literature that no single model is the best in every situation, due to the fact that a real-world problem is often complex in nature and any single model may not be able to capture different patterns equally well. Therefore, the chance in order to capture different patterns in the data can be increased by combining different models.

These hybrid models, despite all their advantages, have two assumptions [15] that will degenerate their performance if the opposite situation occurs; therefore, they may be inadequate in some specific situations. [42]
indicates that these architectures do not always lead to better estimates when compared to single models and combined forecasts do not necessarily outperform the results captured by a linear model. Hibon and Evgeniou present experiments, using the 3003 series of the M3-competition that challenge this belief that combining forecasts improves accuracy relative to individual forecasts [43]. Their results indicate that the advantage of combining forecasts is not that the best possible combinations perform better than the best possible individual forecasts, but that it is less risky in practice to combine forecasts than to select an individual forecasting method. Taskaya and Casey try to answer the question that whether the performance of hybrid models shows consistent improvement over single models. Their results show that hybrid models are not always better, and hence, the process of model selection, still remains an important step despite the popularity of hybrid models. They believe that, despite the popularity of hybrid models, which rely upon the success of their components, single models themselves can be sufficient [15].

In this paper, ARIMA models are applied to construct a new hybrid model in order to overcome the above-mentioned limitation of artificial neural networks and to yield more general and more accurate model than traditional hybrid ARIMA and artificial neural networks models. In our proposed model, a time series is considered as function of a linear and a nonlinear component, so, in the first phase, an autoregressive integrated moving average model is first used in order to identify and magnify the existing linear structures in data. In the second phase, a multilayer perceptron is used as a nonlinear neural network in order to model the pre-processed data, which the existing linear structures are identified and magnified by ARIMA, and to predict the future value of time series. The rest of the paper is organized as follows. In the next section, the literature survey of the hybrid models for financial time series forecasting, especially exchange rate forecasting, is briefly reviewed. In Section 3, the formulation of the proposed model is introduced. In Section 4, the proposed model is applied to forecast the weekly Indian rupee, the weekly British pound, and the daily Euro all against the United States dollar exchange rates and its performance is compared with those of other models. Section 5 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 [44]. The literature on this topic has dramatically expanded since the early work of Reid, and Bates and Granger. [45] presented a hybrid artificial intelligence (AI) approach that integrated the rule-based systems technique and neural networks to S&P 500 stock index prediction. [46] considers 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.

[47] proposed a hybrid methodology to exploit the unique strength of ARIMA models and Support Vector Machines (SVMs) for stock prices forecasting and obtained very promising results. [48] constructed a combination model incorporating seasonal autoregressive integrated moving average (SARIMA) model and SVMs for seasonal time series forecasting. Their experimental results showed that the hybrid model was superior to the individual models (SARIMA and SVM) for the test cases of the production value of the Taiwanese machinery industry. [49] proposed a hybrid modelling and forecasting approach based on Grey and Box–Jenkins autoregressive moving average (ARMA) models. They concluded that the hybrid method had a higher forecasting precision to the complex problems than the single methods. [50] presented a new hybrid approach that integrated artificial neural network with genetic algorithms (GAs) to stock market forecast.

[51] presented a hybrid ARIMA and artificial intelligence approaches to financial markets prediction. Their results indicated that the integrated model could be used as an appropriate alternative forecasting tool for financial markets forecasting. [52] 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. Obtained results of this work revealed that the prediction using the proposed nonlinear ensemble model was generally better than those obtained using the other models. [53] 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. [54] 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. Their results showed that SARIMABP was superior to the SARIMA model, the BP with deseasonalized data, and the BP with differenced data for the test cases of machinery production time series and soft drink time series. [55] 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. Other works that use the hybrid models for exchange rate forecasting are summarized in Table 3.

### Table 3. Some published neural network based hybrid models for exchange rate forecasting in the recent years

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Type</th>
<th>Model</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-</td>
<td>[56]</td>
<td>2015</td>
<td>Hybrid</td>
<td>The motivation of this paper is to introduce a hybrid Rolling Genetic Algorithm-Support Vector Regression (RG-SVRs) model for optimal parameter selection and feature subset combination. The algorithm is applied to the task of forecasting and trading the EUR/USD, EUR/GBP and EUR/JPY exchange rates.</td>
</tr>
<tr>
<td>2-</td>
<td>[57]</td>
<td>2009</td>
<td>Hybrid</td>
<td>This paper describes a hybrid model formed by a mixture of various regressive neural network models, such as temporal self-organizing maps and support vector regressions, for modeling and prediction of foreign exchange rate time series. A genetic algorithm is applied to fuse all the information from the mixture regression models and the economic indicators.</td>
</tr>
<tr>
<td>3-</td>
<td>[58]</td>
<td>2006</td>
<td>Hybrid</td>
<td>This paper propose a two stage forecasting model which incorporates parametric exchange rate models such as autoregressive integrated moving average (ARIMA), vector autoregressive (VAR) and co-integration techniques, and nonparametric techniques such as support vector regression (SVR) and multilayer perceptrons (MLP).</td>
</tr>
<tr>
<td>4-</td>
<td>[59]</td>
<td>2014</td>
<td>Hybrid</td>
<td>In this paper, the forecasting results obtained by conventional time-series models and by the Inter-active Artificial Bee Colony (IABC), which is a young artificial intelligent method, are compared with each other with 4 years historical data.</td>
</tr>
<tr>
<td>5-</td>
<td>[60]</td>
<td>2015</td>
<td>Hybrid</td>
<td>The main goal of this study is to enhance the ability of GARCH-type family models in forecasting the Euro/Dollar exchange rate volatility. For this purpose, a new neural-network-based hybrid model is developed in which a predefined number of simulated data series generated by the calibrated GARCH-type model along with other explanatory variables is used as input variables.</td>
</tr>
<tr>
<td>6-</td>
<td>[61]</td>
<td>2014</td>
<td>Hybrid</td>
<td>This study focuses upon the applications of currently available intelligence techniques to forecast exchange rates in short and long horizons. The predictability of exchange rate is investigated through the use of a novel co-integration-based neuro-fuzzy system (FIS), which is a combination of a Fuzzy Inference System; and Multilayer perceptron.</td>
</tr>
<tr>
<td>7-</td>
<td>[62]</td>
<td>2009</td>
<td>Hybrid</td>
<td>In this paper, a combined approach is proposed where the parametric (MLPs) and nonparametric self-organizing methods (SOMs) are combined sequentially, exploiting the advantages of the individual methods with the aim of improving their performance.</td>
</tr>
<tr>
<td>8-</td>
<td>[63]</td>
<td>2007</td>
<td>Hybrid</td>
<td>A neuro-fuzzy decision-making technology is designed and implemented to obtain the optimal daily currency trading rule. It is found that a non-linear, multilayer perceptron exchange rate microstructure (hybrid) model combined with a fuzzy logic controller generates a set of trading strategies that earn a higher rate of return compared to the simple buy-and-hold strategy.</td>
</tr>
</tbody>
</table>

### 3. Formulation of the Proposed Method

In this section, a novel hybridization of artificial neural networks and ARIMA model is proposed in order to overcome the above-mentioned limitations of traditional hybrid models and also limitations of linear and nonlinear models by using the unique advantages of ARIMA and ANNs models in linear and nonlinear modelling, respectively. In our proposed model, there are no above-mentioned assumptions of the traditional hybrid ARIMA-ANNs models. In addition, in our proposed model in contrast to the traditional hybrid ARIMA-ANNs models, we can guarantee that the performance of the proposed model will not be worse than both ARIMA and artificial neural networks.

Based on the previous works in the linear and nonlinear hybrid models literature, a time series can be considered to be composed of a linear autocorrelation structure and a nonlinear component. In this paper, a time series is considered as function of a linear and a nonlinear component. Thus,

$$ y_t = f_c(L_t, N_t) $$

(1)

where $L_t$ denotes the linear component and $N_t$ denotes the nonlinear component. These two components have to be estimated from the data. In the first stage, the main aim is linear modeling; therefore, an autoregressive integrated moving average (ARIMA) model is used to model the linear component. The residuals from the first stage will contain the nonlinear relationships that linear model does not able to model them, and maybe the linear relationships that linear model does not able to completely model them [64]. Thus,
Currently, no general diagnostic statistics for nonlinear autocorrelation – $N\tilde{z}_t$, and $e_t$, all three aforementioned components are often consist both linear and nonlinear structures. Therefore, each data set may be ignored against the other components. For example, if data only consist of pure linear structure, then we can logically conclude that a component may be ignored against the other components. For example, if data only consist of pure linear structure, then the residual analysis is able to detect any nonlinear relationships in the data. In fact, there is currently no general diagnostic statistics for nonlinear autocorrelation relationships. Therefore, even if a model has passed diagnostic checking, the model may still not be adequate in that nonlinear relationships have not been appropriately modeled. Any significant nonlinear pattern in the residuals will indicate the limitation of the ARIMA model. The forecast values and residuals of linear modeling are the results of first stage that are used in next stage. In addition, the linear patterns are magnified by ARIMA model in order to apply in second stage.

In second stage, the main aim is nonlinear modeling; therefore, a multilayer perceptron is used in order to model the nonlinear and probable linear relationships existing in residuals of linear modeling and original data. Thus,

$$L_t = \left[ \sum_{j=1}^{p} \varphi_j z_{t-j} - \sum_{j=1}^{q} \theta_j e_{t-j} \right] + e_t = \hat{L}_t + e_t,$$

where $\hat{L}_t$ is the forecast value for time $t$, estimated by ARIMA model. $z_t = (1-B)^d (\gamma_t - \mu)$, and $e_t = e_t$ is the residual at time $t$ from the linear model.

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. However, residual analysis is not able to detect any nonlinear patterns in the data. In fact, there is currently no general diagnostic statistics for nonlinear autocorrelation relationships. Therefore, even if a model has passed diagnostic checking, the model may still not be adequate in that nonlinear relationships have not been appropriately modeled. Any significant nonlinear pattern in the residuals will indicate the limitation of the ARIMA model. The forecast values and residuals of linear modeling are the results of first stage that are used in next stage. In addition, the linear patterns are magnified by ARIMA model in order to apply in second stage.

In second stage, the main aim is nonlinear modeling; therefore, a multilayer perceptron is used in order to model the nonlinear and probable linear relationships existing in residuals of linear modeling and original data. Thus,

$$N^1_t = f^1(e_{t-1},...,e_{t-n})$$

(3)

$$N^2_t = f^2(z_{t-1},...,z_{t-m})$$

(4)

$$N_t = f^N(N^1_t, N^2_t)$$

(5)

where $f^1$, $f^2$, $f^N$ are the nonlinear functions determined by the neural network; $n$, $m$ are integers, which are often determined by trial-and-error, and often referred to as orders of the model. Thus, the combined forecast will be as follows:

$$y_t = f^N\left(N^1_t, \hat{L}_t, N^2_t \right)$$

(6)

where $f$ are the nonlinear functions determined by the neural network. $n_j \leq n$ and $m_j \leq m_j$ are integers that are also determined by trial-and-error in design process of final neural network. When we don’t have so many input variables, we can use the second form of the equation (6); otherwise, we have to use the general form (first form) of this equation in order to overcome the limitation of input variables number in the artificial neural networks.

It must be noted that anyone of above-mentioned variables $e_t(i=t-L,...,t-n)$, $z_j(i=t-L,...,t-m)$, and $\hat{L}_t$ or set of them $\{ e_t(i=t-L,...,t-n) \}$ or $\{ z_j(i=t-L,...,t-m) \}$ may be deleted in design process of final network architecture. This may be related to the underlying data generating process and the existing linear and nonlinear structures in data. It is clear that before designing the final network, we cannot sure which component may be ignored or used. However, in some specific situations that we have information about the data generation process (pure linear or pure nonlinear), we can logically conclude that a component may be ignored against the other components. For example, if data only consist of pure linear structure, then $\{ e_t(i=t-L,...,t-n) \}$ variables will be probably ignored against other of those variables. In contrast, if data only consist of pure nonlinear structures, then $\hat{L}_t$ variable will be probably ignored against other of those variables. However, in real world situations, data sets rarely are pure linear or pure nonlinear and often consist both linear and nonlinear structures. Therefore, all three aforementioned components are often applied in order to model real data sets.

Although the proposed model, such as traditional hybrid artificial neural networks and autoregressive integrated moving average models, exploits the unique feature and strength of autoregressive integrated moving average model as well as artificial neural network in determining different linear and nonlinear patterns, it has no assumptions that limit the process of linear and nonlinear patterns modeling separately by using different models and also combining process. Thus, the proposed model can capture the linear and nonlinear autocorrelation structures in data more and
better than traditional hybrid models, so can be an effective way in order to yield more general and more accurate model than of those hybrid models and either of the components models.

4. Application of the Proposed Model to Exchange Rate Forecasting

In this section, the procedure of the proposed model for the weekly Indian rupee, the weekly British pound, and daily Euro all against the United States dollar exchange rates is illustrated. Only the one-step-ahead forecasting is considered.

4.1. The exchange rate (Indian rupee/US dollar) forecasts

The first data set that is considered in this section in order to show the appropriateness and effectiveness of the proposed model for exchange rate forecasting is the weekly spot rates of Indian rupee against United States dollar exchange rate [65]. The data set is from FX database for the period of January 6, 1994–July 10, 2003, for a total 497 observations. 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. In this paper, following the previous works [65], the log differences of the levels are used as follows:

\[ y_t = (\log(p_t) - \log(p_{t-1})) \times 100 \]  

(7)

where \( p_t \) is the exchange rate price for the period \( t \). The log difference is also multiplied by one hundred in order to reduce round-off errors.

**Stage I.** According to the autocorrelation function (ACF) and the partial autocorrelation function (PACF) and using the *Eviews* package software, the best-fitted model for modeling linear structures is an 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 nine inputs, seven hidden and one output neurons (in abbreviated form, \( N^{9-7-1} \)). The structure of the network is shown in Figure 1. The performance measures of the proposed model for train and test data sets are given in Table 4.

![Figure 1. Structure of the best fitted network (INR/USD exchange rate), N^{9-7-1}](image)

**Table 4. Performance measures of the proposed model (INR/USD exchange rate)**

<table>
<thead>
<tr>
<th></th>
<th>Train</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>MAE</td>
</tr>
<tr>
<td></td>
<td>0.1306</td>
<td>0.0797</td>
</tr>
</tbody>
</table>


4.2. The exchange rate (British pound/US dollar) forecasts

The second data set that is considered in this investigation is the exchange rate between British pound and United States dollar. This data contain the weekly observations from 1980 to 1993, giving 731 data points in the time series. This data set is divided into two samples of training and testing. The training data set, 679 observations, is exclusively used in order to formulate the model and then the test sample, the last 52 observations, is used in order to evaluate the performance of the established models. In this paper, the natural logarithmic transformed data is used in the modeling and forecasting analysis, according to [66] and [41].

**Stage I.** In a similar fashion, based on the autocorrelation function (ACF) and the partial autocorrelation function (PACF) and using the Eviews package software, the best-fitted ARIMA model is a random walk model, which has been used by Zhang [41].

**Stage II.** Similar to the previous sections, using pruning algorithms in MATLAB 7 package software, the best fitted network which is selected, is composed of twelve inputs, four hidden and one output neurons (N(12-4-1)). The structure of the best-fitted network is shown in Figure 2. The performance measures of the proposed model for exchange rate data are given in Table 5.

**Table 5.** Performance measures of the proposed model (BP/USD exchange rate)

<table>
<thead>
<tr>
<th></th>
<th>Train</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE  MAE</td>
<td>RMSE  MAE</td>
</tr>
<tr>
<td></td>
<td>0.0056  0.0043</td>
<td>0.0060  0.0049</td>
</tr>
</tbody>
</table>

![Image of Figure 2](image)

**Figure 2.** Structure of the best fitted network (BP/USD exchange rate), N(12-4-1)

4.3. The exchange rate (Euro/US dollar) forecasts

In order to also evaluate the performance of the proposed model in shorter horizon exchange rate series, which may contain more distinct nonlinear feature, the last data set that is considered in this investigation is the daily exchange rate between the Euro and United States dollar. This data contain the daily observations from 21 March 2005 to 20 March 2006, giving 365 data points in the time series. According to the previous sections, this data set is divided into two samples of training and testing. The training data set, 335 observations, is exclusively used in order to formulate the model and then the test sample, the last 30 observations, is used in order to evaluate the performance of the established models.

**Stage I.** In a similar fashion, according to the autocorrelation function (ACF) and the partial autocorrelation function (PACF) and using the Eviews package software, the best-fitted ARIMA model for this data set is a autoregressive model of order two, AR (2).

**Stage II.** Similar to the previous sections, using pruning algorithms in MATLAB 7 package software, the best fitted network which is selected, is composed of six inputs, four hidden and one output neurons (N(6-4-1)). The structure of the best-fitted network is shown in Figure 3. The performance measures of the proposed model for this exchange rate data are given in Table 6.
4.4. Comparison with other models

In this section after selecting the appropriate models for the proposed model, its predictive capabilities for in-sample and out-of-sample data is compared with Zhang’s hybrid model [41], artificial neural networks, linear autoregressive, and the random walk models in the weekly Indian rupee, the weekly British pound, and daily Euro against the United States dollar exchange rate forecasting. Two performance criteria including root mean square (RMSE) and mean absolute error (MAE) are employed in order to evaluate the predictive power of the proposed model in comparison with those other forecasting models.

4.4.1. In-sample forecasts

In-sample performance of the proposed model, neural network, linear autoregressive, and the random walk models for the weekly Indian rupee against the United States dollar (INR/USD), the weekly British pound against the United States dollar (BP/USD), and daily Euro against the United States dollar (Euro/USD) exchange rates are respectively presented in Tables 7-9. The obtained results in the INR/USD, BP/USD, and Euro/USD exchange rate forecasting cases show that the proposed model outperforms the neural network, the linear autoregressive, and the random walk models in both evaluation criteria. For example in INR/USD exchange rate case, the RMSE of the proposed model i.e. 0.1306 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.

Table 7. Comparison of the in-sample performance of the proposed model with those of other forecasting models (INR/USD)

<table>
<thead>
<tr>
<th></th>
<th>Proposed Model</th>
<th>Neural Network</th>
<th>Linear Autoregressive</th>
<th>Random Walk</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.1306</td>
<td>0.2048</td>
<td>0.2518</td>
<td>0.3325</td>
</tr>
<tr>
<td>MAE</td>
<td>0.0797</td>
<td>0.1057</td>
<td>0.1341</td>
<td>0.1577</td>
</tr>
</tbody>
</table>

Table 8. Comparison of the in-sample performance of the proposed model with those of other forecasting models (BP/USD)

<table>
<thead>
<tr>
<th></th>
<th>Proposed Model</th>
<th>Neural Network</th>
<th>Linear Autoregressive</th>
<th>Random Walk</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.0056</td>
<td>0.0060</td>
<td>0.0061</td>
<td>0.0061</td>
</tr>
<tr>
<td>MAE</td>
<td>0.0043</td>
<td>0.0046</td>
<td>0.0046</td>
<td>0.0046</td>
</tr>
</tbody>
</table>

Table 9. Comparison of the in-sample performance of the proposed model with those of other forecasting models (Euro/USD)

<table>
<thead>
<tr>
<th></th>
<th>Proposed Model</th>
<th>Neural Network</th>
<th>Linear Autoregressive</th>
<th>Random Walk</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.0004</td>
<td>0.0019</td>
<td>0.0049</td>
<td>0.0054</td>
</tr>
<tr>
<td>MAE</td>
<td>0.0017</td>
<td>0.0023</td>
<td>0.0032</td>
<td>0.0036</td>
</tr>
</tbody>
</table>
4.4.2. Out-of-sample forecasts

The results for out-of-sample performance of the proposed model, Zhang’s hybrid ARIMA/ANNs, neural network, linear autoregressive, and the random walk models for above-mentioned exchange rates are respectively presented in Table 10, 11, and 12. The out-of-sample forecasts of the proposed model are also more accurate than Zhang’s hybrid model, neural network, linear autoregressive, and the random walk forecasts by both criteria in all INR/USD, BP/USD, and Euro/USD exchange rates cases. For example in the case of BP/USD exchange rate, the MAE of the proposed model i.e. 0.0049 is significantly lower than the MAEs of the above-mentioned models which are equal to 0.0051, 0.0053, 0.0054, and 0.0054, respectively.

Table 10. Comparison of the out-of-sample performance of the proposed model with those of other forecasting models (INR/USD)

<table>
<thead>
<tr>
<th></th>
<th>Proposed Model</th>
<th>Zhang’s Model</th>
<th>Neural Network</th>
<th>Autoregressive</th>
<th>Random Walk</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.0898</td>
<td>0.0978</td>
<td>0.1087</td>
<td>0.1096</td>
<td>0.1342</td>
</tr>
<tr>
<td>MAE</td>
<td>0.0620</td>
<td>0.0666</td>
<td>0.0676</td>
<td>0.0740</td>
<td>0.0854</td>
</tr>
</tbody>
</table>

Table 11. Comparison of the out-of-sample performance of the proposed model with those of other forecasting models (BP/USD)

<table>
<thead>
<tr>
<th></th>
<th>Proposed Model</th>
<th>Zhang’s Model</th>
<th>Neural Network</th>
<th>Autoregressive</th>
<th>Random Walk</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.0060</td>
<td>0.0066</td>
<td>0.0067</td>
<td>0.0067</td>
<td>0.0067</td>
</tr>
<tr>
<td>MAE</td>
<td>0.0049</td>
<td>0.0051</td>
<td>0.0053</td>
<td>0.0054</td>
<td>0.0054</td>
</tr>
</tbody>
</table>

Table 12. Comparison of the out-of-sample performance of the proposed model with those of other forecasting models (Euro/USD)

<table>
<thead>
<tr>
<th></th>
<th>Proposed Model</th>
<th>Zhang’s Model</th>
<th>Neural Network</th>
<th>Autoregressive</th>
<th>Random Walk</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.0013</td>
<td>0.0039</td>
<td>0.0047</td>
<td>0.0053</td>
<td>0.0062</td>
</tr>
<tr>
<td>MAE</td>
<td>0.0020</td>
<td>0.0032</td>
<td>0.0034</td>
<td>0.0038</td>
<td>0.0041</td>
</tr>
</tbody>
</table>

5. Conclusions

Theoretical as well as empirical evidences in the literature suggest that combining different methods together can be an effective and efficient way to improve forecasts. The traditional hybrid techniques that decompose a time series into its linear and nonlinear form are one of the most popular hybrid models categories, which have been shown to be successful for single models. However, these hybrid models, despite of all advantages cited for them, have some assumptions that may degenerate their performance if the opposite situation occurs. In these models, the linear and nonlinear patterns in a time series are assumed to be separately modeled by components, the residuals from the linear model are contain only the nonlinear relationship, and the relationship between the linear and nonlinear components is additive. Therefore, the relationship between the components may be underestimated in these models; and hence their performance may be degraded, for example, if the linear and nonlinear patterns cannot be separately modeled or the residuals of the linear component do not comprise valid nonlinear patterns or there is not be any additive association between the linear and nonlinear elements and the relationship is different (for example multiplicative).

In this paper, a new hybrid artificial neural network and autoregressive integrated moving average model is proposed as an alternative forecasting technique to the traditional hybrid ARIMA/ANNs models for time series forecasting. In our proposed model, similar to the traditional hybrid ARIMA/ANNs models, the unique strength of ARIMA and ANN in linear and nonlinear modeling are jointly used, aiming to capture different forms of relationship in the data; especially, in complex problems that have both linear and nonlinear correlation structures. However, there are no aforementioned assumptions in modeling process of the proposed model. Therefore, in proposed model, in contrast to the traditional hybrid ARIMA/ANNs, it can be generally guaranteed that the performance of the proposed model will not be worse than either of their components used separately. In addition, empirical results in both weekly and daily exchange rate forecasting indicate that the proposed model can be an effective way to improve forecasting accuracy achieved by traditional hybrid ARIMA/ANNs models.

References


https://books.google.com.tr/books?id=1WVHAAAAAMAAJ


