A novel committee machine and reviews of neural network and statistical models for currency exchange rate prediction: An experimental analysis

Trilok Nath Pandeya,⇑, Alok Kumar Jagadevb, Satchidananda Dehuric, Sung-Bae Choda

Department of Computer Science and Engineering, S’O’A (Deemed to be University), Bhubaneswar, Odisha, India
bDepartment of Computer Science and Engineering, KIIT University, Bhubaneswar, Odisha, India
cDepartment of Information and Communication Technology, Fakir Mohan University, Balasore, Odisha, India
dDepartment of Computer Science, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-749, South Korea

ABSTRACT

Prediction of currency exchange rate becomes highly desirable due to its greater role in financial and managerial decision making process. The fluctuations in exchange rate affect the economy of a country. Hence, over the years different types of neural network models along with statistical models are developed to predict the currency exchange rates of different countries with varying parameters. In this paper, we divide our effort into two parts. In first part, we have reviewed a few selected models of neural networks and statistics including fundamental and technical aspects of currency exchange rate prediction. Additionally, a thorough and careful experimental result analysis has been conducted on the models reviewed in part one. A committee machine has been proposed in part two to address the shortcomings of both neural networks and statistical models in the context of exchange rate prediction. Our study reveals that the currency exchange rates with multi-layer neural networks having Bayesian learning predictive accuracy is better than multi-layer neural networks with back-propagation learning. However, in the case of higher-order neural network multi-stage radial basis function network is predicting better than single stage radial basis function network. In the case of statistical models, it is drawn that under the umbrella of root mean square error measure, random walk is predicting better than other models of this category, whereas variance based model predicts better than rest of the models grouped under normalized mean square error measure. On the other hand, the integrated model is performing better than its counterpart like models with stand-alone mode. Moreover, our newly proposed committee machine is drawing a clear line over all the models while predicting exchange rate of GBP/USD.

Keywords:
Currency exchange rate
Neural network
Bayesian learning
Multi-layer perceptron
Radial basis function network
Functional link artificial neural network
Cascaded functional link artificial neural networks
Autoregressive integrated moving average
Committee machine

ARTICLE INFO

Article history:
Received 10 November 2017
Revised 25 January 2018
Accepted 25 February 2018
Available online xxxx

Keywords:
Currency exchange rate
Neural network
Bayesian learning
Multi-layer perceptron
Radial basis function network
Functional link artificial neural network
Cascaded functional link artificial neural networks
Autoregressive integrated moving average
Committee machine

Contents

1. Introduction ........................................................................................................ 00
2. Fundamentals of exchange rate prediction .......................................................... 00
  2.1. Fundamental approach .................................................................................. 00
  2.2. Technical approach ....................................................................................... 00
3. Neural network based exchange rate prediction .................................................. 00
  3.1. Basic neural networks based exchange rate prediction ................................. 00

⇑ Corresponding author.
E-mail addresses: trilokpandey@soa.ac.in (T.N. Pandey), alok.jagadevfcs@kiit.ac.in (A.K. Jagadev), sbcho@yonsei.ac.kr (S.-B. Cho).
Peer review under responsibility of King Saud University.

© 2018 The Authors. Production and hosting by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Exchange rate expectations play an important role in the literature on exchange rate determination. Understanding how exchange rate expectations are formed is crucial for academic analysis of exchange rate behavior, as well as for decision-making of both practitioners and policymakers. Models of exchange rate determination in open-economy macroeconomics often rely on assumptions about rationality of exchange rates expectations. It is practically impossible to test the implications of theoretical exchange rate models, without running into the problem of joint hypothesis testing in the absence of survey-based expectations. In addition to understanding exchange rate behavior, the rationality assumption can have serious implications for evaluating the effectiveness of many government policies. The availability of survey forecast allows us to evaluate the rational expectations hypothesis directly.

Although survey-based forecasts of other macroeconomic variables have been studied in the literature for at least 60 years, research on rationality and accuracy of exchange rate forecasts goes back only to the late 1980s (Engel, 1996). Limited data availability on professional exchange rate forecast is partially responsible for the short history of research on survey-based exchange rate forecast. From the early studies of Dominguez (1986) and many other researchers who, have studied the nature of exchange rate expectations using survey data, we find two most commonly examined questions in the literature on survey-based exchange rate expectations as rationality of the forecasts and their predictive accuracy.

In this paper, we have focused on currency exchange rate that is the rate at which currency of two countries exchange against each other. Currency exchange rate plays an important role in financial markets. Exchange rates are determined in the foreign exchange market. A stable exchange rate is helpful for financial institution for investment, where a fluctuation in exchange rate will affect interest rate, unemployment, prices, and wages of a country. Prediction of correct exchange rate will be helpful for the economical growth of any country. Many researchers have proved the effectiveness of predicting currency exchange rate by using neural network (NN) models like multi-layer perception (MLP), radial basis function neural network (RBFN), and functional link artificial neural network (FLANN) models (Albers et al., 1996; Yao and Tan, 2000; Burse et al., 2010; Bissoondeal and Binner, 2008; Ak et al., 2016). Input dimension and the time delay are two critical factors that affect the performance of neural network. In addition, many statistical models are also used for prediction of currency exchange rate such as autoregressive integrated moving average (ARIMA), generalized autoregressive conditional heteroskedasticity (GRACH), random walk (RW), stack regression, variance based, simple averaging, and simple mean square error (MSE). Researchers are taking the currency exchange rate data of different countries daily wise, month wise, and quarter wise and then predict the exchange rate and analyzed the percentage of error. The model having a less percentage of error considered as a best model for currency exchange rate prediction. From the research analysis, we found that RBFN and cascaded functional link artificial neural networks (CFLANN) models have better accuracy rate than the other models. The overall categorization of both statistical and neural network models are illustrated in Fig. 1. The categorization of neural networks are based on inputs to the neural network (i.e., linear or non-linear) and metric to predict the accuracy. However, in the case of statistical model, the categorization is based on metric towards error minimization (i.e., root mean square error (RMSE) or normalized mean square error (NMSE)).

This paper is set out as follows. Section 2 presents an overview of foreign exchange rate and the factors affecting the exchange rate.
Future works.

In the last section, we have drawn the conclusions and highlighted and analysis that we have obtained from the intensive experiments and parameters settings of different models, performance comparison, the data representation for different models, environment, and committee machine to predict the exchange rate. Section 7 presents the factors that affect the exchange rate are presented as fundamental approach and technical approach.

2. Fundamentals of exchange rate prediction

Forecasting exchange rate are necessary to determine the foreign denominated cash flow involved in international transactions, so predicting the accurate exchange rate help any country to determine the benefits and risks in the international business environment. Forecasting of exchange rate will be done using an information set selected by the forecaster (MacDonald, 2000). Factors that affect the exchange rate are as follows:

a. Interest rate – Interest rate directly affects the exchange rate. Exchange rate increases with increase in interest rate.

b. Inflation – Lower inflation increases the currency value.

c. Current account deficits – Current account is the balance of trade between a country and its trading partners. Deficit in current account shows the country is spending more on foreign trade than earning, so it decrease the currency exchange rate of that country.

d. Public debts – Public invest their money in government projects or public projects and foreign investors also invest their money if inflation is more. Government is unable to refund the money of public when the country export something and at that time other country does not give appropriate money, so exchange rate will decrease.

e. Terms of trades – If the export price of a country rises by greater than that of its import then it’s terms of trade will be improved and currency value will be increased.

f. Political stability and economic performance – Foreign investors search stable country with strong economic performance in which they can invest their capital.

Based on the information set two approaches for prediction of exchange rate are presented as fundamental approach and technical approach.

2.1. Fundamental approach

The fundamental approach is based on some fundamental economic variable on which exchange rates are predicted. Usually these variables are inflation rate, interest rate, term of trade, trade balance, etc. Fundamental model is based on structural model which is a mixture of art and science. Structural model are used by practitioners to generate equilibrium exchange rate. Projections or trading signal are generated by using these equilibrium exchange rate. Fundamental approach starts with a model which is based on purchasing power parity (PPP) theory. Based on this model the data are collected using different statistics and measure forecasting equations. Let \( S_{t-1} \) is the future value of exchange rate at time \( t \). Notation used for forecasting \( S_{t-1} \) is \( E_t[S_{t-1}] \). Where, \( E_t[\cdot] \) is expectation taken at time \( t \). Forecasting error will be determined by

\[
E_{t+1} = S_{t+1} - E_t[S_{t+1}].
\]  

(1)

Mean square error for the model is calculated as:

\[
MSE = [(E_{t+1})^2 + (E_{t+2})^2 + \ldots + (E_{t+Q})^2]/Q.
\]

(2)

where, \( Q \) is the number of data points. There are two kinds of forecasting, one is in-sample and other is out-of-sample forecasting. In-sample forecasting uses today's information to forecast the today's exchange rate, where as out-of-sample forecasting forecast the future behavior of exchange rate using today's information.

2.2. Technical approach

It is based on price information. Turning points are detected by computer and based on this trading signal is generated. Normally moving average (MA) is used for technical approach. Simple average of past price will be done in MA model. In simple moving average (SMA) model unweighted mean of the previous \( Q \) data points are used as follows:

\[
SMA = (S_t + S_{t-1} + \ldots + S_{t-(Q-1)})/Q.
\]

(3)

When, we take most recent past price then short-run MA(SRMA) will be calculated. When, longer series of past prices are taken then long-term MA (LRMA) is calculated. A double MA system uses LRMA and SRMA. In MA model when SRMA past rates cross LRMA, buy and sell signals are usually triggered. When the currency moves downward its SRMA will below its LRMA and when currency rises then it crosses LRMA by generating a buy foreign currency signal. Instead of using these direct methods we can use several supervised learning methods such as neural networks for the accurate prediction of foreign exchange rates, which are discussed in the next section.

3. Neural network based exchange rate prediction

This section is divided into two subsections. In Section 3.1, we discuss a few basic neural network models for foreign exchange.
rate prediction. Higher order neural network models for exchange rate prediction is discussed in Section 3.2.

3.1. Basic neural networks based exchange rate prediction

Under this section, we have discussed several basic neural network models to predict the exchange rates of different countries.

3.1.1. Multi-layer perceptron neural network

Multi-layer perceptron neural network is an artificial neural network used for classification, pattern recognition, and prediction. MLP consists of input layer, hidden layers, and output layer, the number of hidden layer changes depending on complexity of data (Dehuri et al., 2012). Input layer receives input and then multiply with weight and forward to hidden layer (Guresen et al., 2011; Emmerson, 1993). In hidden layer nonlinear activation function is used which convert nonlinear form of a problem to its linear form so that it is easily separable, different activation functions are acquired for different networks for better performance. Most common activation function for MLP are sigmoid and hyperbolic tangent. All nodes in hidden layer uses same activation (Bissoondeal and Binner, 2008; Gill et al., 2010; Galeshchuk, 2016). MLP uses a supervised learning technique where desired output was known by the network. A typical architecture of MLP is given in Fig. 2. Output of the MLP network ‘y’ is calculated as

\[ y = f_j \left( \sum_{j=1}^{N} w_{ij} h_j - T \right) \]  

Where, N is the number of neurons in the hidden layer, w_{ij} is the weight between hidden layer to the output neuron, h_j is the output of neuron j, T is the threshold of the output neuron and f_j is the sigmoid activation function of the output neuron.

3.1.2. MLP with back-propagation learning

Back-propagation learning allows supervised learning procedure where desired output is known by the network before training process. Learning occurs by changing connection weights based on the amount of error in the output. Error produced in the output propagate backward direction of the network for minimizing the overall error produced by the network. It uses gradient descent method for minimizing the error in weight space because weight is the solution of the learning problem. If the state of system moving in the opposite direction to the largest local scope then weight are updated in downward direction. An activation function is used in the hidden layer and output layer which compute the gradient of error.

Back-propagation learning has two phases one is propagation and other is weight updation (Bilski, 2000; Rocio Cogollo and David Velasquez, 2014). In propagation phase forward propagation of input is done by forwarding the input data through the network by the help of activation function and backward propagation error produced by the network is calculated to produce delta. In weight updating phase, each weight(w_{ij}) multiply with its output delta and input activation to get the gradient of the weight and subtract a ratio (percentage) of the gradient from the weight. In Fig. 3 input is propagate towards forward direction, output of the system is calculated as \( h_j = g(h_j) \), where, g is the activation function and h_j is the output of neuron j.

MLP network can also be trained with Bayesian learning technique as discussed in the next subsection.

3.1.3. MLP with Bayesian learning

Bayesian network is a directed acyclic graph, used for regression, classification, and inverse problem. In Bayesian learning entire distribution of model prediction estimation is done instead of the mean prediction estimation of the model (Madsen et al., 2017; Villanueva and Maciel, 2014; Gasse et al., 2014; Zhao, 1997; Majhi et al., 2009). This estimation take noise in the data and variance of the model. For all unknown data Bayesian approach produce posterior probability distributions. In Bayesian learning for MLP neural network the natural end variables are the predictions of the model for new input. For input x_k and training data D = \((x_1, y_1), (x_2, y_2), \ldots, (x_k, y_k)\), for the output y_k posterior predictive distribution is calculated by integrating the predictions of the model with respect to the posterior distribution of the model (Huang et al., 2008; Wang et al., 2012; Zen et al., 2017; Zhijun, 2013).

\[ p(y_{k+1}|x_{k+1}, D) = \int p(y_{k+1}|x_{k+1}, \theta)p(\theta|D)d\theta \]  

Where, \( \theta \) denotes all the model parameters and hyper parameters of the model structures. The MLP neural network uses many layers and many hidden neurons in each layer, which potentially increases the complexity of the overall network. So, we can also use the simpler higher-order neural network models to predict the exchange rate of any country, which is discussed in the next Section.

3.2. Higher order neural network based Exchange rate prediction

In this section, we have analysed some of the higher order neural network based models to predict the exchange rate of different countries.

3.2.1. Exchange rate prediction using radial basis function network

RBFN is a simple neural network having powerful problem solving ability. Name of the network is RBFN because it uses radial basis activation function. It has three layers, that are input, hidden, and output layer as shown in Fig. 4. In the input layer data are given to the network, then this input multiply with the weight and given to the hidden layer. c_k is the cluster vector in this hidden layer and there must be cluster elements c_k (i = 1 to N, k = 1 to L), where N and L are number of input and hidden nodes, respectively (Chen...
et al., 2016; Majhi et al., 2009). Each node in the hidden layer calculates the Euclidean distance between cluster head and the input and basically Gaussian type radial bias function is used as activation function (Cheng et al., 2006; Dash et al., 2016; Mahanta et al., 2016). The outputs of the hidden layer multiply with weights are given to the output layer.

For better accuracy we can use combination of more then one RBFN networks, which is discussed in the next Section.

3.2.2. Multistage radial basis function network

Combination of RBFN models produce less amount of error and gives better result than single RBFN model. So a multistage RBFN model was develop where more number of single RBFN are combine together and consider as a single RBFN model (Hoori and Motai, 2017; Zhao, 1997; Majhi et al., 2009; Wang et al., 2012; Zen et al., 2017). A multistage RBFN have three stage as shown in Fig. 5 and are presented as:

a. Producing multiple single RBFN predictors
b. Choosing appropriate ensemble members
c. Combining the selected members

After selecting the appropriate model the out put of all the model are combine and final output is produced $y = f(x)$,

$$f(x) = \sum_{i=1}^{m} w_i f_i(x).$$

where, $w_i$ is the weight of $f_i(x)$. Four method are used for determining weight in the network. That are simple averaging, simple MSE, stacked regression method and variance-based weighting method.

We can also use another simple technique such as FLANN to predict the exchange rate, which is discussed in the next Section.

3.3. Functional link artificial neural network for exchange rate prediction

FLANN is a single layer artificial neural network. It is capable of performing complex decisions i.e it can work with nonlinear data with out hidden layer. Hidden layer is removed to reduce its computational cost (Majhi et al., 2012; Naik et al., 2016; Patra and Kot, 2002). Inputs are functionally expanded using some trigonometry expansion. Let’s consider a two dimensional input $x = [x_1x_2]^T$. The enhanced pattern is obtained by using trigonometric function which, is used by network for equalization purpose. Block diagram of FLANN is given in Fig. 6. FLANN consists of 3 process.

a. Functional expansion process: In this input element is non-linearly expanded to create more number of inputs. The number of element after expansion becomes more than original input.

b. Estimation process: It compute the output of adaptive model and generate the error signal.

c. Adaptive process: It adjust the weight by weight update learning rule.

The learning process of FLANN is given in the next section.

3.3.1. FLANN with back-propagation learning

FLANN use back-propagation learning algorithm for training the network. Error produced by it is calculated by substituting the estimated output from desired output and propagate backward for weight updating using weight update rule as shown in Fig. 7.

Where, $K$ is the number of sine and cosine expansion. After expansion $S$ matrix is generated. Each row of $S$ is denoted as
We can also use CFLANN to predict the exchange rate as discussed with the new value and the testing process will be conducted.

When the training process is completed weight are computed correction weight vector (Zhao and Zhang, 2010; Dai et al., 2014). The model can be represented as:

$$Y(m) = X^T(m)W(m) + \hat{w}_0(m).$$ (7)

$$\hat{d}_1(m) = f(X, m) = \frac{1 - e^{-\gamma(m)}}{1 + e^{-\gamma(m)}}.$$ (8)

where, the system output is denoted as \(\hat{d}_1(m)\), and \(d_1(m)\) is the desired output. Error is calculated as the difference between the desired output and system output \(e(m) = \hat{d}_1(m) - d_1(m)\). Then error and input vector are applied to weight updated algorithm to compute correction weight vector (Zhao and Zhang, 2010; Dai et al., 2014). When the training process is completed weight are fixed with the new value and the testing process will be conducted. We can also use CFLANN to predict the exchange rate as discussed in the next subsection.

3.3.2. Cascaded FLANN

In this model two single FLANN are connected in series as shown in Fig. 8. Each FLANN are passed through back-propagation algorithm and output of first FLANN given to the second FLANN as its input. This value also expanded using trigonometric function (Bebarta et al., 2012; Pandey et al., 2013). The block diagram of CFLANN is presented in Fig. 8.

Where, FL1 is the first FLANN and \(\hat{d}_1(m)\) is the output of first FLANN, which is calculated as follows:

$$\hat{d}_1(m) = f(X, m) = \frac{1 - e^{-\gamma(m)}}{1 + e^{-\gamma(m)}}.$$ (9)

and the output \(Y(m)\) is calculated as in FLANN.

The operational principle is same as single FLANN till we get a minimum MSE. In the next Section we have analysed some of the statistical models integrated with neural networks.

4. Statistical models for exchange rate prediction

In this section we have discussed some of the statistical models to predict the exchange rates of different countries.

4.1. Random walk model

A random walk (naïve) model assumes that the most recent observation is the best predictor of the future observation (Trucos and Hotta, 2016). The model can be represented as:

$$y_t = y_{t-1} + \epsilon_t.$$ (10)

where, \(y_t\) is the predicted value and \(\epsilon_t\) is a white noise process.

4.2. GARCH model

GARCH model became a staple tool in the field of finance. It is useful not only for analyzing but also forecasting volatility. The GARCH model (Bollerslev, 1986) is a generalisation of the Autoregressive Conditional Heteroskedasticity (ARCH) model. Instead of being conditionally fixed over time, the variance of the ARCH models is modelled as being dependent on lags of past squared residuals. Bollerslev (1986) extended ARCH models to produce GARCH models in which the variance also depends on lags of past variances. The GARCH model can be considered as an infinite order ARCH model (Bollerslev, 1986). A basic GARCH(p,q) model consists of a mean equation and a variance equation.

The simplest and most commonly used variance equation is the GARCH(1,1) and therefore will be used in this paper. The mean equation is specified as an AR(k) process, where k is chosen such that the residuals are white noise. Thus, \(k = 3\) for the Australian dollar/US dollar rate and \(k = 1\) for the British pound/US dollar rate.

4.3. Stack regression model

Stacking regressions is a method for forming linear combinations of different predictors to give improved prediction accuracy. The idea is to use cross-validation data and least squares under non negativity constraints to determine the coefficients in the combination. Its effectiveness is demonstrated in stacking regression trees of different sizes and in a simulation stacking linear subset and ridge regressions. Reasons why this method works are explored.

4.4. ARIMA model

ARIMA models are generally combination of autoregressive (AR) and moving average model (MA), a non seasonal ARIMA model is denoted ARIMA(p, d, q) where parameters p, d, and q are non-negative integers is the order of the Autoregressive model, d is the degree of differencing, and q is the order of the moving-average and a seasonal ARIMA model is denoted ARIMA(p, d, q)\(_s\), where m is the number of period (Li and Chiang, 2013; Yunus et al., 2016).

We can also use integrated ARIMA model with neural networks to produce better results as discussed in the next Section.

5. An integrated model using ARIMA, MLP, and RBFN

An integrated model is produce by using ARIMA, MLP and RBFN model (Kia et al., 2012; Khashei and Bijari, 2011; Chen et al., 2010). In this model real exchange rate is called relax and predicted exchange rate is called arimaout. First the time series data is given to the ARIMA model and the produced output and error is called error1 as shown in Fig. 9. Error1 is the error produced by ARIMA model. This error1 is given to MLP and result is mlperror as shown in Fig. 10. Relax = arimaout + mlperror. RBFN is used to further reduced the error. Output of ARIMA is added to output of MLP.

Fig. 8. Block diagram of CFLANN.

Fig. 9. Error in ARIMA Model.

Fig. 10. Error in MLP Model.
and the remainder of this result and the real exchange rate will be called error2 and given to the RBFN as shown in Fig. 11.

\[ \text{error2} = \text{realex} - (\text{arimaout} + \text{mlpererror}) \]

The final result of the proposed integrated model comes from the summation of the error modeled by RBFN and the error modeled by MLP with the time series that was modeled by ARIMA. Final model is given in Fig. 12.

6. Committee machine to predict exchange rate

From the above experimental studies we have found that it is a standard practice with neural networks to train many different candidate networks, and then to select and keep the best while discarding the rest (Chen and Lin, 2006). In this process all the effort involved in training the discarded network is wasted and the validation set has a random component, so the network that has the best performance on it will not necessarily have the best performance on the test set. These drawbacks can be overcome by combining the networks together to form a committee machine. With this we can improve the performance of the testing data with a little extra computational effort. In fact, the committee can better than the best single constituent network in isolation. In this work, we have considered a mixture of experts, where the constituent outputs are non-linearly combined by some form of gating system as shown in the Fig. 13.

With mixture of experts, the principle of divide-and-conquer distributes the given learning task between a set of expert networks, and combines the constituent outputs to produce an overall output that is superior to that of any single network acting on its own. In the mixture of experts approach the outputs are gated according to the inputs as follows:

\[ y(p) = \sum_{i=1}^{k} \frac{1}{K_i} y_i(p). \quad (11) \]

where, \( y(p) \) is the predicted output, \( y_i(p) \) is the input to the expert system and \( k \) is the number of inputs to the system. In our proposed committee machine we have selected the best three models form each category as expert system and the weights are selected by probability method.

7. Experimental study

In this work, we have analysed different models that takes different exchange rates from different countries with varying time duration. The number of training and testing data are also vary with different types model and are described in this section.

7.1. Dataset preparation

In this section, we have explained the various type of data used in different models for their analysis.

a. Neural network with Random Walk(RW), GRACH and ARIMA Models

For this model data has been collected as follows:

In order to substantiate the results in the paper, two currencies are analysed, namely the Australian dollar/US dollar and the UK pound/US dollar. The daily data were obtained from Data stream for the period of 2nd-January-1986 to 13th-September-2007, consisting of 5661 observations. The stationarity properties of the data were investigated using the Augmented Dickey-Fuller (ADF) test (Dickey and Fuller, 1979). To avoid the complications arising in building forecasting models from the issue of possible non-stationarity, the log returns of the series are considered. Such an approach has also been used by, for example, Brooks (1997). Let \( p_i, i = 1, 2, \ldots, T \) be the daily exchange rate series, then the log-returns are calculated as \( x_i = \log(p_i) - \log(p_{i-1}) \).

b. MLP with back-propagation learning

For this model data has been collected as follows:

1. Daily wise

Daily exchange rate of EUR/USD, GBP/USD, USD/JPY are taken from the site.

http://www.global-view.com/forex-trading-tools/forex-history/index.html. Data are collected for the period from 1st-Jan-2014 to 25th-Apr-2014. From 78 training vectors 60 vectors are used for for the training and 18 for the testing.

2. Monthly wise

The monthly exchange rate of EUR/USD,GBP/USD,USD/JPY are taken from http://www.oanda.com/currency/historical-rates/. Data are collected for the period from May-2009 to May-2014. From 55 training vectors 40 vectors are used for for the training and 15 for the testing.

3. Quarterly wise

The quarterly wise data of EUR/USD,GBP/USD,USD/JPY are taken from site: http://www.oanda.com/currency/historical-rates/. Data are collected for the period from May-1999 to May-2014. From 54 training vectors 42 vectors are used for for the training and 12 for the testing.

c. MLP with Bayesian learning

For this model data has been collected as: U.S. dollar against the British Pound (GBP/USD) and Japanese Yen (JPY/USD) from Jan-1990 to Dec-2002 are taken daily wise, which was given by Professor Werner Antweiler, University of British Columbia, Canada. 60 patterns are used for testing and rest of the data are taken for training the network.

d. Integrated model

For this model data has been collected as follows EUR/USD (Euro to US Dollar) exchange rate is used for the model

Out-of-sample forecasts of British Pound.

<table>
<thead>
<tr>
<th>Models</th>
<th>RMSE (RMSE/RMSE)**</th>
<th>MAE (MAE/MAE)**</th>
<th>MAPE (MAPE/MAPE)**</th>
</tr>
</thead>
<tbody>
<tr>
<td>RW</td>
<td>0.003521 (1.000)</td>
<td>0.002743 (1.000)</td>
<td>0.3962% (1.000)</td>
</tr>
<tr>
<td>ARMA</td>
<td>0.003527 (1.002)</td>
<td>0.002754 (1.004)</td>
<td>0.3977% (1.004)</td>
</tr>
<tr>
<td>GARCH</td>
<td>0.003525 (1.001)</td>
<td>0.002746 (1.001)</td>
<td>0.3966% (1.001)</td>
</tr>
<tr>
<td>NN</td>
<td>0.003506 (0.996)</td>
<td>0.002733 (0.996)</td>
<td>0.3947% (0.996)</td>
</tr>
</tbody>
</table>

The RMSE shows the error in terms of the level and not the direction. Mean absolute error (MAE) is a measure of difference between two continuous variables. Assume X and Y are variables of paired observations that express the same phenomenon. MAE is the average vertical distance between each point and the Y = X line, which is also known as the One-to-One line. Mathematically MAE can be presented as:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |Y_i - \hat{Y}_i|$$

(13)

MAPE is a measure of prediction accuracy of a forecasting method in statistics. It usually expresses accuracy as a percentage, and is defined by the formula:

$$\text{MAPE} = \frac{1}{n} \left( \sum_{i=1}^{n} \frac{|Y_i - \hat{Y}_i|}{Y_i} \right)$$

(14)

7.2. Environment and parameters setting

In this work, most of the models have used neural network tool box provided by Matlab software package. The author has taken one input variable, six hidden nodes and 6000 iterations to compare NN models with RW, GARCH, and ARIMA model. For MLP neural networks the authors have used (7-6-5-1), (7-4-2-1) and (5-10-1) network architectures and for RBF neural networks the authors have used (7-11-1) and (7-9-1) network structures. The learning parameters for learning algorithm of MLP neural network were eta = 0.3 and alpha = 0.9, and for the RBF neural network, the learning parameters were eta = 0.4 and alpha = 0.9. In ARIMA model ARIMA (1, 1, 4) architecture was used. Each set of data are normalized by dividing each value by the maximum value of each set such that each normalized value is less than or equal to unity.

7.3. Error function

In this section, we have discussed various error functions as follows:

We can determine the error of a system as Root-mean-square error and can be calculated as follows

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (\hat{Y}_t - Y_t)^2}.$$  

(12)

The ability to predict movement direction or turning points can be measured by a statistic. The NMSE (Normalised Mean Square Error) is an estimator of the overall deviations between predicted and measured values. It is defined as:

$$\text{NMSE} = \frac{1}{N} \sum_{t=1}^{N} \frac{\hat{Y}_t - Y_t}{\hat{Y}_t}^2$$

(15)

$$\hat{Y} = \frac{1}{N} \sum_{t=1}^{N} \hat{Y}_t$$

(16)

$$\bar{Y} = \frac{1}{N} \sum_{t=1}^{N} Y_t$$

(17)

where, \(\hat{Y}_t\) is a vector of n predictions, and Y is the vector of observed values corresponding to the inputs to the function which generated the predictions. Directional change statistics (Dsat) can be expressed as Dsat.

$$D_{sat} = \frac{1}{T} \sum_{t=1}^{T} a_t$$

(18)

if \((x_{t+1} - x_t)/(x_{t+1} - x_t) >= 0\) then \(a_t=1\), else \(a_t=0\) and T is the number of testing samples.

7.4. Results and analysis

In this section, we have analysed the performance of different models and determined which model produces less amount of error and more accurate output based on the analysis. The details of the analysis are given below. At last, we have provided a comparison table for the above analysis.

7.4.1. Neural network with Random Walk (RW), GARCH, and ARIMA Models

Here, we have investigated whether NN models can offer improvements in terms of forecasting accuracy over RW models and extensively used ARIMA and GARCH models. Given that most rate data contain nonlinear structures, one would expect NN models to be able to exploit the nonlinear structures to provide better forecasts as NN models can approximate any continuous function to a good degree of accuracy without the imposition of the assumptions regarding the form of nonlinearity. The results obtained in this paper, indicate that NN models can provide better forecasts than RW models and traditional
models, such as, ARIMA and GARCH models. The better performance of NN models is likely to have stemmed from the fact that they are nonlinear models, that can exploit the nonlinearity in the exchange rate data without the imposition of assumptions about the form of nonlinearity. Table 1 shows the daily wise exchange rate predictions of British Pound using linear and non-linear models. It has been observed from the experimental analysis that neural network (NN) models have less error rates recomparisons to other linear models.

7.4.2. MLP with back-propagation learning

In this section we have analysed that the short-term prediction method provides good accuracy of the prediction and can be used in practical systems to predict the exchange rate for one step ahead. Table 2 shows the daily wise, monthly wise and quarterly wise error rate of different exchange rates using MLP with back propagation learning. The prediction results with daily, monthly and quarterly step are depicted in Fig. 14, Fig. 15, and Fig. 16 for EUR/USD, GBP/USD and USD/JPY. These results show that the
short-term prediction method using MLP provides good accuracy of the prediction.

7.4.3. MLP with Bayesian learning

In this section we have analysed that the RMSE of BPNN is larger than the MLP neural networks by Bayesian learning. It indicates that Bayesian learning is better at avoiding over-fitting than the traditional parameter optimization method. Because Bayesian learning does not attempt to find a single best parameter vector; but rather attempts to infer the posterior distribution of the parameters from the given the data. Table 3 shows the error rate of different exchange rates using MLP with bayesian learning. The experimental results shows that predicting exchange rates using MLP with Bayesian learning is producing better results than BPNN model.

7.4.4. Performance of Integrated model

In this section we have compared the proposed integrated model to simple ARIMA and simple MLP and RBF in terms of error level with RMSE. Table 4 shows the error rate of different exchange rates using Integrated model. The results from the Table 4 shows that the integrated model works better than the other single models both in terms of error level and directional status.

Table 3
Performance of Bayesian learning.

<table>
<thead>
<tr>
<th>Input nodes</th>
<th>BPNN</th>
<th>Bayesian learning</th>
<th>MLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Error in Neural network Models(GBP/USD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.005471</td>
<td>0.005468</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.004491</td>
<td>0.004482</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.004496</td>
<td>0.004494</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.005467</td>
<td>0.005465</td>
<td></td>
</tr>
<tr>
<td>(b) Error in the Neural network Models(JPY/USD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.007438</td>
<td>0.007121</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.006348</td>
<td>0.006302</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.006358</td>
<td>0.006285</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.007293</td>
<td>0.006994</td>
<td></td>
</tr>
</tbody>
</table>

Table 4
Performance of Integrated model.

<table>
<thead>
<tr>
<th>EUR/USD</th>
<th>Integrated Model</th>
<th>Single ARIMA</th>
<th>Single MLP</th>
<th>Single RBF</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.0068</td>
<td>0.0081</td>
<td>0.0076</td>
<td>0.0085</td>
</tr>
<tr>
<td>D-sat(%)</td>
<td>71</td>
<td>65</td>
<td>67</td>
<td>59</td>
</tr>
</tbody>
</table>

Table 5
Performance of multistage RBFN model.

<table>
<thead>
<tr>
<th>Models used for Prediction of GBP/USD Data used for training and testing</th>
<th>Error Rate (RMSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN</td>
<td>0.003521</td>
</tr>
<tr>
<td>MLP-BP</td>
<td>0.005467</td>
</tr>
<tr>
<td>GARCH</td>
<td>0.003525</td>
</tr>
<tr>
<td>Committee Machine (Using Neural networks as Expert system)</td>
<td>0.001495</td>
</tr>
</tbody>
</table>

Table 6
Accuracy in FLANN and CFLANN.

<table>
<thead>
<tr>
<th>Month ahead</th>
<th>Actual</th>
<th>LMS</th>
<th>FLANN</th>
<th>CFLANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rupees conversion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>45.63</td>
<td>44.20</td>
<td>44.682</td>
<td>45.9620</td>
</tr>
<tr>
<td>3</td>
<td>44.20</td>
<td>44.78</td>
<td>44.7953</td>
<td>44.0019</td>
</tr>
<tr>
<td>6</td>
<td>44.82</td>
<td>44.78</td>
<td>44.9954</td>
<td>44.2958</td>
</tr>
<tr>
<td>12</td>
<td>45.97</td>
<td>42.83</td>
<td>45.2923</td>
<td>46.1454</td>
</tr>
<tr>
<td>Pound conversion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.5764</td>
<td>0.5524</td>
<td>0.5725</td>
<td>0.5776</td>
</tr>
<tr>
<td>3</td>
<td>0.5654</td>
<td>0.5181</td>
<td>0.5763</td>
<td>0.5651</td>
</tr>
<tr>
<td>6</td>
<td>0.5656</td>
<td>0.4926</td>
<td>0.5802</td>
<td>0.5698</td>
</tr>
<tr>
<td>12</td>
<td>0.5181</td>
<td>0.4716</td>
<td>0.5900</td>
<td>0.5283</td>
</tr>
<tr>
<td>Yen conversion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>118.45</td>
<td>114.78</td>
<td>115.3831</td>
<td>116.6212</td>
</tr>
<tr>
<td>3</td>
<td>115.48</td>
<td>113.68</td>
<td>115.4728</td>
<td>115.7532</td>
</tr>
<tr>
<td>6</td>
<td>117.07</td>
<td>113.89</td>
<td>116.4297</td>
<td>116.7695</td>
</tr>
<tr>
<td>12</td>
<td>111.00</td>
<td>108.00</td>
<td>121.6428</td>
<td>113.2557</td>
</tr>
</tbody>
</table>

Table 7
Statistical models vs committee machine exchange rate prediction error with respect to GBP/USD.

<table>
<thead>
<tr>
<th>Models used for Prediction of GBP/USD Data used for training and testing</th>
<th>Error Rate (RMSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RW</td>
<td>0.003521</td>
</tr>
<tr>
<td>ARIMA</td>
<td>0.005467</td>
</tr>
<tr>
<td>GARCH</td>
<td>0.003525</td>
</tr>
<tr>
<td>Committee Machine (Using Statistical Models as Expert System)</td>
<td>0.001495</td>
</tr>
</tbody>
</table>

Table 8
Higher order neural network models vs committee machine exchange rate prediction error with respect to GBP/USD.

<table>
<thead>
<tr>
<th>Models used for Prediction of GBP/USD Data used for training and testing</th>
<th>Error Rate (RMSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single RBF model</td>
<td>0.0614</td>
</tr>
<tr>
<td>Multistage RBF</td>
<td>0.0388</td>
</tr>
<tr>
<td>Committee Machine (Using Higher order neural networks as Expert system)</td>
<td>0.0065</td>
</tr>
</tbody>
</table>

7.4.5. Performance of multistage RBFN model

In this section, four main currency exchange rates are used to test the proposed RBF neural network ensemble forecasting model. Table 5 shows the error rate of different exchange rates using Multi stage RBFN model. The low NMSE does not necessarily mean that there is a high hit ratio for foreign exchange movement direction prediction. Thus, the Dstat comparison is necessary for business practitioners. Focusing on Dstat of Table 5, it is not hard to find that the proposed RBF neural network ensemble forecasting model outperforms the other ensemble models and the single RBF model (Yu et al., 2008; Yu et al., 2006; Zhang et al., 2013).

7.4.6. Performance of FLANN model

In this section the authors have proposed two efficient low complexity neural network based forecasting models for exchange rate prediction. The first model is very simple with one layer and single neuron but with nonlinearly mapped input features. The second model consists of two stage FLANNs. The output of the first stage undergoes nonlinear expansion and then fed to the second FLANN for predicting the exchange rate. Computer simulation study of both the models reveal that each of them offer better prediction performance compared to the LMS model. However, out of the FLANN and CFLANN models, the later offers superior prediction performance in all cases. Table 6 shows the error rate of different exchange rates using FLANN and CFLANN model.

7.4.7. Performance of Committee machine model using a common data set

In this section we have compared our proposed committee machine model with the other categories of prediction models and we have found that for the common data set our proposed committee machine model is producing better results in comparison to the other models. The experimental results are shown in Table 7–10.

7.5. Result analysis w.r.t different models based on a common dataset

In this section we have compared different models having a common dataset and the experimental results are depicted in Tables 11 and 12 based on RMSE and NMSE measures. It has been
observed from the experimental analysis presented in Tables 11 and 12 that for the same dataset the RBF ensemble model and MLP with Bayesian learning model have less error rates and hence are producing better predictions. It has also been observed that the committee machines for different categories are producing better results then the individual and ensemble models.

Finally in Table 13 we have summarised the whole experimental study on exchange rate prediction based on different types of statistically, bio-inspired algorithms, and committee machine.

8. Conclusions and future work

From the review and experimental results analysis it is concluded that neural network models are better than the various linear models and in MLP if the data are taken daily wise to train the network then error rate will be less in comparison to other models. We have also observed that reduced complexity of neural network, we can combine more number of RBF to produce more accurate result than single RBF network. It is observed that combination of RBF, MLP, ARIMA resulting an integrated model is producing better performance than individual model. To avoid complexity and for better accuracy FLANN can be used to predict the exchange rates of different countries. It is observed that combination of FLANN which is called CFLANN gives more accurate result than individual FLANN. From this intensive experimental study we conclude that the integrated models are producing better results than the basic models and also the exchange rate prediction depends on the economic of the developed countries. With this inspiration, we have developed a committee machine and tested on GBP/USD exchange rate prediction. It is observed that the committee machine is performing better than all other models reported in this work. In future, we will include exchange rates of different developed and under developed countries for rigorous analysis. Our future work also extend to develop robust hybrid models by combining the best attributes of neural networks and meta-heuristic algorithms for prediction of exchange rate more accurately.

Conflict of interest

We do not have any conflict of interest with other authors.

References


