Forecasting Exchange Rates using Time Series and Neural Network Approaches

K A D S A Nanayakkara[1*], N V Chandrasekara[1] and D D M Jayasundara[1]

[1.] Department of Statistics & Computer Science
University of Kelaniya,
P.O. Box 11600,Kelaniya,
Sri Lanka
E-mail: nsajeeka@gmail.com, nvchandrasekara@kln.ac.lk, jayasund@kln.ac.lk

*Corresponding author: **K A D S A Nanayakkara**

ABSTRACT
Exchange rates play an important role in controlling dynamics of the foreign exchange market. Predicting exchange rates has become one of the most challenging applications of financial time series forecasting due to its unpredictability and volatility. This research study is to develop and compare the accuracy of two models; Generalized Auto-Regressive Conditional Heteroskedasticity (GARCH) as the time series model and Feedforward neural network with the Backpropagation algorithm as the Artificial Neural Network (ANN) model for predicting daily currency exchange rate of US Dollar against Sri Lankan Rupee (USD/LKR). For both models, past lagged observations of the data series and moving average technical indicators were employed as the explanatory variables and the predictive performance were evaluated using a number of widely used statistical metric. According to the performance of two models, it can be concluded that the ANN based model performs better when compared with the GARCH model to predict the exchange rate of USD/LKR.

Keywords: Forecasting, Exchange rate, GARCH model, Artificial Neural Network.

1. INTRODUCTION

Due to Globalization, the foreign exchange market has experienced unexpected development over the last few decades. Various economic factors such as inflation, economic growth, interest rates and monetary policies influence the value at which national currencies are traded in international markets. Therefore, the exchange rates play a vital role in controlling dynamics of the foreign exchange market.

However, exchange rates prediction has become one of the most challenging applications of modern time series forecasting as the rates are inherently noise, non-stationary and deterministically chaotic (Chandrasekara and Tilakaratne, 2009). These characteristics suggest that there is no complete information that could be obtained from the past behavior of such markets to fully capture the dependency between the future rates and that of the past. One general assumption is made in such cases is that the historical data incorporate all those behavior. As a result, the historical data is the major player in the prediction process.

Fahimifard *et. al.* (2009) states that it is well documented that many economic time series are non-linear while, a linear correlation structure is assumed among the time series values. Therefore, the ARIMA (Auto-Regressive Integrated Moving Average) model cannot capture nonlinear patterns and approximation of linear models to complex real-world problem is not always satisfactory. Much research effort has been devoted to exploring the nonlinearity of exchange rate data and to developing specific nonlinear models to improve exchange rate forecasting. Parametric nonlinear models such as the chaotic dynamic, self-exciting threshold autoregressive model, autoregressive random variance (ARV) model, autoregressive conditional heteroskedasticity (ARCH) and Generalized ARCH (GARCH) have been proposed and applied to foreign exchange rate forecasting (Huang *et al.*, 2004).

Hsieh (1989) who used daily exchange rates for five major currencies and Kugler and Lenz (1991) who used weekly data for ten currencies found GARCH model can explain large part of nonlinearities for all considered exchange rates. Furthermore, Nachne and Ray (1993) showed that accounting for conditional heteroskedasticity leads to better forecasts of monthly exchange rate as compared to random walk and other non-linear models. Also, Chappell and Chant (1998) showed that GARCH models produces better forecasts than the well-known random walk model using the South Korean Won/British Pound exchange rate to produce a set of one-step-ahead exchange rate forecasts for ten trading days.

Use of neural network based models is an alternative option available to researchers for capturing the underlying non-linearity in the exchange rate series. Kamruzzaman and Sarker (2003) stated that the Artificial Neural Networks (ANNs), the well-known function approximates in prediction and system modeling, has recently shown its great applicability in financial time series analysis and forecasting due to its ability to extract complex nonlinear and interactive effects. Concerning the application of neural networks to financial time series forecasting, Weigened *et al.* (1992) found that ANNs are better than random walk models in predicting the Deutsche mark/US dollar (DEM/USD) exchange rate. Similarly, in several applications, Wu (1995), Kamruzzaman and Sarker (2003), Fahimifard *et. al*., (2009) and many other researchers found results in favor of neural network in forecasting exchange rates. The greatest advantage of a neural network is its ability to model complex nonlinear relationship without a priori assumptions of the nature of the relationship (Karayiannis and Venetsanopoulos, 1993).

It is not a simple task to select the appropriate neural network architecture for an exchange rate forecasting due to a large number of parameters to be estimated empirically. Zhang and Hu (1998) examined the effects of the number of input and hidden nodes as well as the size of the training sample on the in-sample and out-sample performance for the British pound/US dollar (GBP/USD) exchange rate forecasting using multi-layer feedforward neural network architecture with backpropagation learning algorithm. They found that neural networks outperform linear models, particularly when the forecast horizon is short. In addition, the number of input nodes has a greater impact on performance than the number of hidden nodes, while a large number of observations do reduce forecast errors.

By revealing the literature, Generalized Auto-Regressive Conditional Hetroskedasticity (GARCH) and Artificial Neural Network (ANN) models can be identified as appropriate models to predict exchange rates. In this research study we develop and compare the accuracy of two models; GARCH as the time series model and feedforward neural network with the backpropagation algorithm as the ANN model for predicting daily currency exchange rate of US Dollar against Sri Lankan Rupee (USD/LKR).

2. THEORY AND METHODOLOGY

Nearly five years daily exchange rate of USD/LKR (from January 1, 2007 to November 18, 2011) taken from Central Bank of Sri Lanka was used in this study. This study period contains 1275 daily observations of the variable excluding the weekends. The missing values (owing to holidays and religious festivals in weekdays) were replaced by immediately preceding exchange rate.

2.1 Time Series models

ARCH is a time series model which is introduced by Engle in 1982 concerning the volatility of inflation. ARCH model is appropriate for series with variances that change over time, i.e. time varying variance processes. ARCH model can be described as follows:

$$y_t = \beta_0 + e_t \tag{1}$$

$$e_t | I_{t-1} \sim N(0, h_t) \tag{2}$$

$$h_t = \alpha_0 + \alpha_1 e_{t-1}^2 , \qquad \alpha_0 > 0,\ \ 0 \leq \alpha_1 < 1 \tag{3}$$

The Equation 1 describes the behavior of the mean of the time-series. The Equation 2 indicates that the error of the regression, e_t, are normally distributed and heteroskedastic. The variance of the current period's error depends on (are conditional on) information that is revealed in the preceding period (lagged effects). The variance of e_t is given the symbol h_t. The Equation 3 describes how the variance behaves.

The ARCH(1) model can be extended in a number of ways. One obvious extension is to allow for more lags. An ARCH(q) model that includes lags has a conditional variance function that would be:

$$h_t = \alpha_0 + \alpha_1 e_{t-1}^2 + \alpha_2 e_{t-2}^2 + \cdots + \alpha_q e_{t-q}^2 \tag{4}$$

In this case the variance or volatility in a given period depends on the magnitudes of the squared errors in the past q periods. Testing, estimating, and forecasting, are natural extensions of the case with one lag.

One of the shortcomings of an ARCH(q) model is that there are $q+1$ parameters to estimate. If q is a large number, accuracy in the estimation may be lost. The Generalized ARCH (GARCH) model is an alternative way to capture long lagged effects with fewer parameters. It is a special generalization of the ARCH model and the GARCH model would be:

$$h_t = \delta + \alpha_1 e_{t-1}^2 + \beta_1 h_{t-1} \tag{5}$$

where $\delta = (\alpha_0 - \beta_1 \alpha_1)$. This model is denoted as GARCH(1,1). In higher order GARCH(p, q) models, q refers to the number of lags of e_t and p refers to the number of lags of h_t to include in the model of the regression's variance.

2.2 Neural Network models

Neural networks are flexible and capable of nonlinear modeling. With ANN, there is no need to specify a particular model. Rather, the model is adaptively based on the features presented from the data. This data-driven approach is suitable for many empirical researches where no theoretical guidance is available to suggest an appropriate data generating process (Fahimifard *et. al.,* 2009).

2.2.1 Architecture of Neural Networks

An ANN is defined as a data processing system consisting of a large number of simple highly interconnected processing elements in architecture inspired by the way biological nervous systems of the brain. There are many factors which should determine under the architecture of a neural network, namely, number of input and output nodes, number of hidden layers, number of nodes in each hidden layer, biases, and weights, etc. and how they are connected.

The most common type of artificial neural network consists of three groups or layers of units: a layer of "input" units is connected to a layer of "hidden" units, which is connected to a layer of "output" units. Number of nodes in each layer is depends on the nature of the study.

Networks in which signals are propagated only in the forward direction are known as Feedforward networks. In this network, the information moves from the input nodes, through the hidden nodes (if any) and to the output nodes. There are no cycles or loops in the network. In other words, there is no feedback from the output to input. Figure 1 illustrates the architecture of feedforward neural network.

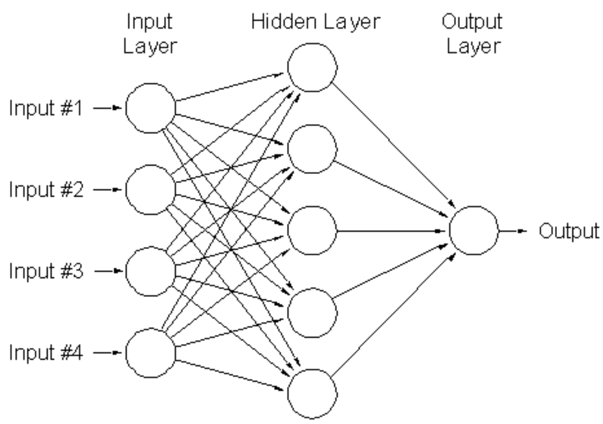

Figure 1 - Architecture of Feedforward Neural Network

2.2.2 Supervised Training

When the network is trained for a specific target, the training process is called Supervised Training. As the inputs are applied to the network, the network outputs are compared to the targets. Then the training is carried out to adjust the weights and biases of the network in order to move the network outputs closer to the targets.

An important issue concerning supervised learning is the problem of error convergence, which is also known as the minimization of error between the desired and computed unit values. The aim is to determine a set of weights which minimizes the error.

2.2.3 Backpropagation Algorithm

ANNs apply many learning rules, of which backpropagation is one of the most frequently used algorithms in time series forecasting, because it is capable of resolving a wide variety of problems. The backpropagation algorithm is commonly applied to feedforward networks with supervised learning rules.

Backpropagation algorithm uses steepest gradient descent technique to minimize the sum of squared error (E) over all training data. During training, each desired output is compared with actual output and E is calculated at the output layer. The weights are updated using the delta rule that the adjustments to the weights are made in order to minimize the error. The weight ω_j is updated in the n^{th} training cycle according to the following equation where the parameters η and α are the learning rate and the momentum factor, respectively.

$$\Delta\omega_j(n) = -\eta \frac{\partial E}{\partial \omega_j} + \alpha\, \Delta\omega_j(n-1) \qquad (6)$$

The value of the learning rate significantly affects the convergence of the neural network learning algorithm. For a large-scale problem Backpropagation learns very slowly and its convergence largely depends on choosing suitable values of learning rate and momentum factor by the user.

2.3 Performance Criteria

Evaluating performance is a crucial factor to select the best model and to draw conclusions. Some of the widely used performance metrics in time series predictions are namely, Normalized Mean Square Error (NMSE), Mean Absolute Error (MAE), Directional Symmetry (DS), Correct Up trend (CU) and Correct Down trend (CD). NMSE and MAE measure the deviation between actual and forecasted value. DS measures correctness in predicted directions. CU and CD measure the correctness of predicted up and down trend, respectively. These criteria are defined as follows:

$$NMSE = \frac{\sum_k (x_k - \hat{x}_k)^2}{\sum_k (x_k - \bar{x}_k)^2} = \frac{1}{\sigma^2 N}\sum_k (x_k - \hat{x}_k)^2 \qquad (7)$$

$$MAE = \frac{1}{N}\sum_k |x_k - \hat{x}_k| \qquad (8)$$

$$DS = \frac{100}{N}\sum_k d_k\,; \quad d_k = \begin{cases} 1, & if(x_k - x_{k-1})(\hat{x}_k - \hat{x}_{k-1}) \geq 0 \\ 0, & otherwise \end{cases} \qquad (9)$$

$$CU = 100\frac{\sum_k d_k}{\sum_k t_k}\,; \quad d_k = \begin{cases} 1, & if(\hat{x}_k - \hat{x}_{k-1}) > 0, (x_k - x_{k-1})(\hat{x}_k - \hat{x}_{k-1}) \geq 0 \\ 0, & otherwise \end{cases} \quad t_k = \begin{cases} 1, & if(x_k - x_{k-1}) > 0 \\ 0, & otherwise \end{cases} \qquad (10)$$

$$CD = 100\frac{\sum_k d_k}{\sum_k t_k}\,; \quad d_k = \begin{cases} 1, & if(\hat{x}_k - \hat{x}_{k-1}) < 0, (x_k - x_{k-1})(\hat{x}_k - \hat{x}_{k-1}) \geq 0 \\ 0, & otherwise \end{cases} \quad t_k = \begin{cases} 1, & if(x_k - x_{k-1}) < 0 \\ 0, & otherwise \end{cases} \qquad (11)$$

3. RESULTS AND DISCUSSION

Technical and fundamental analyses are the two major financial forecasting methodologies. Literature reveals that in recent times, technical analysis has drawn particular academic interest due to the increasing evidence that markets are less efficient than was originally thought. In this study, several inputs / explanatory variables such as MA5 (moving average on 5 days), MA10, MA20, MA30, MA40, MA60, MA120 and past lagged exchange rates were considered to identify most influential variables. By using graphical techniques (autocorrelation and moving average), it is possible to identify the potential input variables among them.

Lag 1, lag 2, MA5, MA10, MA20 and MA30 of daily exchange rates were identified as the potential influence variables to predict exchange rates of USD/LKR from the preliminary analysis.

3.1 Time Series Model

The considered exchange rate series is segregated for testing, 10% of latest time period data and the remaining is for training. i.e. 1151 observations for training (in-sample data) and remaining 124 observations are kept for testing (out-of-sample data).

The Lagrange Multiplier test was used to identifying the ARCH effects. Since the ARCH effect is present, GARCH (1,1) model is more appropriate as it takes the time-varying variance into account.

The results of GARCH(1,1) model proved that all considered explanatory variables are significantly effect on exchange rate USD/LKR. Further, R-squared value and F-statistic suggest that GARCH(1,1) model is significant. Moreover, Autocorrelation and Partial Autocorrelation plots, Ljung-Box test and Lagrange Multiplier test for residuals implied that the selected model is adequate with fewer parameters.

Therefore, GARCH(1,1) model is more appropriate for predicting exchange rate of USD/LKR in time series approach and the model is as follows:

Conditional Mean Equation:

$$\begin{aligned} LKR = {} & 0.4874 + 0.2226*(LAG1) - 0.5969*(LAG2) + 1.8278*(MA5) \\ & - 0.3390*(MA10) - 0.1975*(MA20) + 0.1406*(MA30) \end{aligned} \tag{12}$$

Conditional Variance Equation:

$$GARCH = 0.00196 + 1.4238* RESID(-1)^{\wedge}2 + 0.194401*GARCH(-1) \tag{13}$$

3.2 Neural Network Model

As outlined in Section 2, USD/LKR exchange rates is segregated to train, validate and test the neural network using the proportions, 75%, 15% and 10% respectively. i.e. 927 records for training, 185 records for validation and remaining 124 records are kept for testing (out-of-sample data).

The performance of a neural network depends on a number of factors. As mentioned earlier, appropriate input combination, number of hidden layers and neurons in each layer and parameters of learning algorithm (learning rate and momentum) in a neural network are selected using trial and error method. Training process was carried out a large number of times, using the training data set, iteratively by modifying above mentioned factors to find the best model. The model which indicates the highest performance was selected as the best performed model to predict the USD/LKR exchange rate.

In order to achieve better forecasting performance, input combination is more important factor. Most applicable input combination for neural network model is lag 1, lag 2, MA5, MA10, MA20 and MA30.

Whenever number of hidden layers in a neural network increases, it becomes more complex and slower to converge (Chandrasekara and Tilakaratne, 2009). Therefore many researchers suggest that use minimum number of layers to design neural networks. Hence number of hidden layers which were concerned in this study, were varied from one to three while keeping other factors consistently. Neural network with two hidden layers is performed well than using one or three hidden layers.

In general, networks with too few hidden nodes may not have enough power to model and learn the data. But networks with too many hidden nodes may cause overfitting problems, leading to poor forecasting ability (Zhang and Hu, 1998). The number of hidden neurons was varied between 2 to 6 and different networks were trained with different learning parameters. The appropriate numbers of neurons in first and second hidden layers are four and three respectively, which indicate highest performance.

When identifying parameters, learning rate and momentum which associated with backpropagation learning algorithm are considered. To have better performance, either low learning rate or high momentum is expected to accelerate the convergence of the training process. In this case, large number of different neural networks with different values of learning rate and momentum were trained to come up with the most suitable values. Learning rate of 0.005 together with momentum of 0.8 gives the best performance when compared to other parameter combinations.

3.3 Model Comparison

The forecasting results of exchange rate of USD against LKR measured in terms of the performance metrics of best performed models selected from each approach over 124 days are showed in Table 1.

Table 1- Performance of best GARCH and Neural Network models

Performance Criteria	NMSE	MAE	DS	CU	CD
GARCH(1,1) Model	0.0422	0.0627	69.3548	47.619	42.8571
Neural Network Model	0.0708	0.0600	82.2581	72.2222	65.6250

Although the NMSE of GARCH model is lesser than in neural network model, MAE of GARCH model is higher than in neural network model. Furthermore, accuracy of upward trend predictions in GARCH model which is 48% and in neural network model it is 72%. Also downward trend prediction accuracy of GARCH model is 43% while the neural network model shows 66% accuracy in downward trend prediction. The neural network model exhibits 82% overall prediction accuracy where as the GARCH model exhibits 69%.

The results indicate that neural network based forecasting model produce better performance than time series model; GARCH for predicting exchange rate of USD/LKR. Moreover, time series plots of actual/predicted exchange rates for testing sample of both approaches were drawn in order to clarify the accuracy of the predictions of each approaches and illustrated in Figure 2 and 3.

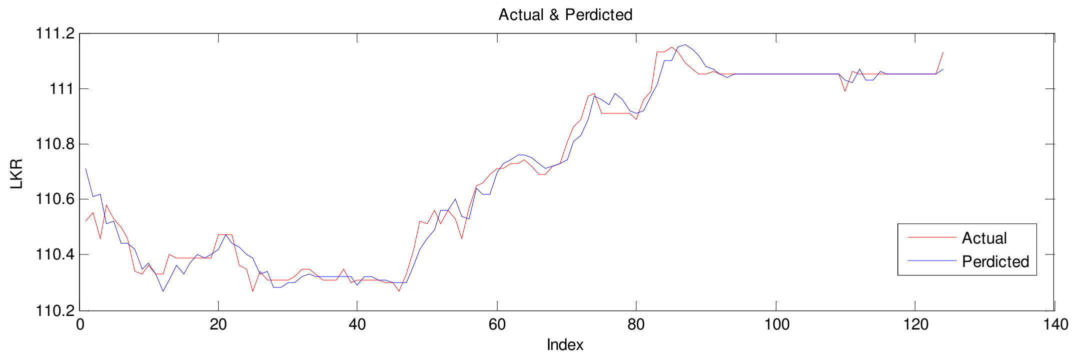

Figure 2- Time series plot of actual/predicted of GARCH(1,1) model

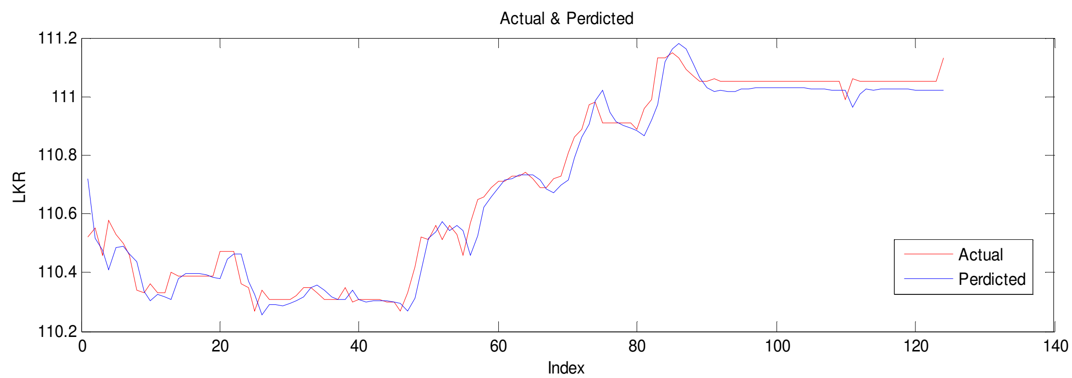

Figure 3- Time series plot of actual/predicted of neural network model

Figure 2 and 3 show that the deviation of predicted values from actual values is approximately equal in both approaches. However, the directions of predictions are more parallel in neural network model than GARCH model.

Considering all performance, neural network can be suggested as better approach to predict the USD/LKR than time series approach.

4. CONCLUSION

The potential influence variables to exchange rates of USD/LKR were lag 1, lag 2, MA5, MA10, MA20 and MA30. The GARCH(1,1) model was found as the best model in time series approach with 69% prediction accuracy. The optimal architecture of neural network was two hidden layers with four neurons in first hidden layer and three neurons in second hidden layer together with learning rate of 0.005 and momentum of 0.8 which are the parameters of learning algorithm. Prediction accuracy of the best feedforward neural network model with backpropagation algorithm is 82%. Neural network based forecasting model is more appropriate than the time series model (GARCH) to predict exchange rate of US Dollar against Sri Lankan Rupee.

5. REFERENCES

Chandrasekara, N.V. and Tilakaratne, C.D., (2009). Forecasting Exchange rates using Artificial Neural Networks. *Sri Lankan Journal of Applied Statistics*, 10, 187-201.

Chappell, D. and Chant, L., (1998). A non-linear time series model for the South Korean Won/British pound exchange rate: 1.1.97-9.30.98. *Global Economic Review (impact factor: 0.2)*, 27(3), 65-75.

Hsieh, D.A., (1989). Testing for non-linear dependence in daily foreign exchange rates. *Journal of Business*, 62, 339–368.

Huang, W., Lai, K. K., Nakamori, Y., Wang, S., (2004). Forecasting foreign exchange rates with Artificial Neural Networks: a review. *International Journal of Information Technology & Decision Making,* 3(1), 145-165.

Fahimifard, S.M., Homayounifar, M., Sabouhi, M., and Moghaddamnia, A.R., (2009). Comparison of ANFIS, ANN, GARCH and ARIMA Techniques to Exchange Rate Forecasting. *Journal of Applied Sciences,* 9(20), 3641-3651.

Kamruzzaman, J. and Sarker, R.A., (2003). Forecasting of currency exchange rate: a case study.to appear in *Proc. IEEE International Conference on Neural Networks & Signal Processing (ICNNSP'03), Nanjing*.

Karayiannis, N.B. and Venetsanopoulos, A.N., (1993). Artificial Neural Networks: Learning Algorithms. Performance Evaluation and Applications. *Kluwer Academic Publishers, Boston, USA*., ISBN-13: 978-0792392972.

Kugler, P. and Lenz, C., (1991). Chaos ARCH and Foreign Exchange Markets.*Conference on Nonlinear Dynamics and Econometrics*, *Los Angeles*, April 5-6.

Nachne D.M. and Ray D. (1993). Modelling exchange rate dynamics: new perspectives from the frequency domain.*Journal of Forecasting,* 12(5), 379–394.

Weigend, A. S., Huberman, B. A. and Rumelhart, D. E., (1992). Predicting sunspots and exchange rates with connectionist networks.*Nonlinear modeling and forecasting*, 395-432.

Wu, B., (1995). Model-free forecasting for non-linear time series (with application to exchange rates).*Computational Statistics & Data Analysis*, 19, 433-459.

Zhang, G. and Hu, M.Y., (1998). Neural network forecasting of the British pound/USdollar exchange rate. *OMEGA: Int. Journal of Management Science*, 26, 495-506.