2011 International Conference on Innovation, Management and Service
IPEDR vol.14(2011) © (2011) IACSIT Press, Singapore

Application of Intelligent Systems and Econometric Models for Exchange Rate Prediction

Abu Hassan Shaari Md Nor[1], Behrooz Gharleghi[1+]

[1] University Kebangsaan Malaysia

[1+] University Kebangsaan Malaysia

Abstract. In most of the studies done by researchers for exchange rate prediction, linear models such as econometric models and non-linear models such as neural networks have been applied. The lack of studies on the application of dynamic networks is the most important motivation of this study. In this paper NNARX as a dynamic non-linear neural network, artificial neural network (ANN) as a static neural network, GARCH as a non-linear econometric model and ARIMA as a linear econometric model are applied to forecast the Singaporean dollar/US dollar (SGD/USD) exchange rate in three time horizons. Comparison of the performance of different models is measured by different criteria. Results reveal that among all models, NNARX outperformed other models and among non-linear models, NNARX outperformed ANN and both outperformed the GARCH model.

Keywords: NNARX, ANN, GARCH, ARIMA, exchange rate

1. Introduction

Recently, numerous studies have been done for time series forecasting using ARIMA model. This model follows a linear structure and since most of time series represent a non-linear pattern, application of nonlinear models becomes more important since nonlinear models can predict the time series variables with higher accuracy than linear models. Artificial intelligence are widely accepted as an nonlinear models for time series forecasting since they can learn the past behavior of variables and recognize the complexity and nonlinearity in the pattern of data set.[6]

As an example of intelligent systems, a neural network is a processor that has the ability for storing experiential knowledge and making it available for use at a latter stage. ANNs can model the complex nonlinear relationship among the data set without any prior assumption. [8]

Neural networks are categorized into two types; dynamic networks and static networks. Static networks as in feed-forward network have no feedback element and contain no delay in the network. The output of network is calculated directly from inputs through the feed forward connections. In dynamic networks, output not only depends on inputs but also depends on previous inputs, outputs and the state of network. Dynamic networks can learn the sequential or time-varying patterns. [7]

Neural networks have been used for time series forecasting by numerous researchers. For instance, Wu [1] compare the performance of neural networks and ARIMA model for Taiwan/USD exchange rate. His result reveals that the performance of neural networks is better than ARIMA for one step ahead as well as six steps ahead predictions. Zhang and Hu [5] find the result of their paper in favor of neural networks compared to other econometric models. Gencay [11] compare the performance of neural networks with GARCH model in daily spot exchange rate for GBP, DM, and JPY. His finding shows the higher accuracy of neural networks compared to GARCH model.

There are very few literature on application of dynamic neural networks to time series forecasting compared to static neural networks, so this motivate us to perform this study of comparison between these

[1 +] Corresponding author. Tel.: +60173461920; fax: +60389215789
E-mail address: gharleghi.bn@gmail.com

two types of networks. In this paper, we compare the neural network autoregressive with exogenous input (NNARX) as a dynamic neural network, feed forward ANN as a nonlinear static neural network, GARCH as a nonlinear model when we consider the volatility and ARIMA as a linear model for prediction. In order to compare the performance of these models, the common performance measures such as Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Theil's U statistics (U) are used. We compare the performance of the above mentioned models for three time horizons (3, 6 and 12) steps ahead for SGD/USD monthly exchange rate.

2. Methodology

2.1. Autoregressive Integrated Moving Average

Box-Jenkins [10] is one of the most popular approaches for time series prediction that is known as ARIMA method. The assumption underlying the ARIMA model is that the future value of a variable is a linear function of past observations and random errors. In this model it is possible to find an adequate description of data set. This method consists of four steps: (i) model identification, (ii) parameter estimation, (iii) diagnostic checking and (iv) forecasting. In the identification step, it can be seen that if a model generated from an ARIMA process it may contain some autocorrelation properties, so there will be some potential models that can fit the data set but the best fitted model is selected according to AIC information criteria. Stationarity is a necessary condition in building an ARIMA model used for forecasting, so data transformation is often required to make the time series to be stationary. In this paper, the Augmented Dickey Fuller unit root test [3], Phillips Perron unit root test [2] and Zivot-Andrews unit root test [4] are used to test the stationarity of the series. Based on the result obtained, the data set is stationary at first difference even with the introduction of structural break.

Once a tentative model is obtained, estimation of the model parameters is applicable. The parameters are estimated such that an overall measure of errors is minimized. The third step is diagnostic checking for model adequacy. Autocorrelation and also serial correlation of the residuals are used to test the goodness of fit of the tentatively obtained model to the original data. When the final model is approved then it will be used for prediction of future value of exchange rate. The ARIMA model can be written as follows when the data set is stationary:

$$Z_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Z_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \qquad 1.$$

The above equation implies that the forecasted value is depended on the past value and on previous shocks.

2.2. Generalized Autoregressive Conditional Heteroscedasticity

Volatility is one of the features in exchange rate data set and can be measured through the GARCH model. In this model, the conditional variance of a time series depends on the past variance and squared residuals of the process, and it has the advantage of incorporating heteroscedasticity into the estimation procedure of the conditional variance. GARCH model is the reduced form of a more complicated dynamic structure for the time varying conditional second order moments [12]. The GARCH model can be presented by the following form:

$$y_t = \mu + \varepsilon_t \qquad 2.$$

$$\varepsilon_t / \Omega_{t-1} \sim N(0, \sigma^2_t) \qquad 3.$$

$$\sigma_t^2 = \dot{\omega} + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 \qquad 4.$$

where;

yt is equal to log (et / et-1), et is the exchange rate, μ is the mean yt conditioned on past information (Ωt-1) and the following inequality restrictions ώ >0, βj >0, αj >0 are imposed to ensure that the conditional variance (σ2t) is positive. The size and significance of αj indicate the magnitude of the effect imposed by the lagged error term (εt-j) on the conditional variance (σ2t). In other forms of interpretation, the size and significance of αj indicate the ARCH process in the residuals (volatility clustering in the data).

2.3. Artificial Neural Network

Neural Network is a modeling method based on human brain that can learn the rules on the foreign exchange rate through the past data, save these rules and forecast the exchange rate in future. For ANN, there is no need to specify any particular model because ANN can be adapted based on the features presented in data set.

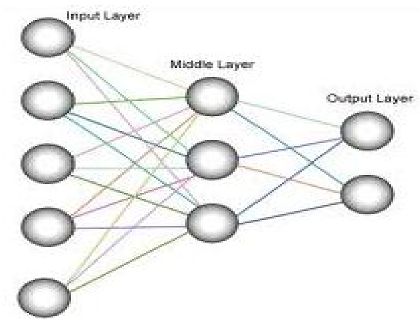
Feed Forward Neural Network is the most widely used network in which all layers except input layer receive weights from their previous layer. This network is consisted of three layers; input layer which includes explanatory variables (inputs) in the model. Hidden layer; lies between the input and output layers. There can be many hidden layers, which allow the network to learn, adjust, and generalize from the previously learned facts (data sets) to the new input. The number of hidden layers and nodes in the network are determined by experimentation, and this paper follows this technique. Output layer is including the output of network.

Single hidden layer feed forward network is represented as follow for time series modeling and forecasting, it has three layers of simple processing units connected by acyclic links:

$$y_t = w_0 + \sum_{j=1}^{q} w_j . g(w_{0j} + \sum_{i=1}^{p} w_{i,j} . y_{t-i}) + \varepsilon_t \qquad 5.$$

where, w_{ij} ($i = 0,1,2,\ldots,p, j=1,2,\ldots,q$) and w_j ($j=0,1,2,\ldots,q$) are model parameters called connection weights; p is the number of input nodes; and q is the number of hidden nodes. Figure1 represents the simple structure of feed forward neural network:

Fig.1: Three layer neural network



Activation function can take several forms; the type of this function is specified by the situation of the neuron within the network. The logistic and tangent hyperbolic activation functions are mostly used as the hidden layer transfer function that represents in Eqs. (6) and (7), respectively:

$$Sig(x) = 1/(1+e^{-x}) \qquad 6.$$

$$Tanh(x) = (1-e^{-2x})/(1+e^{-2x}) \qquad 7.$$

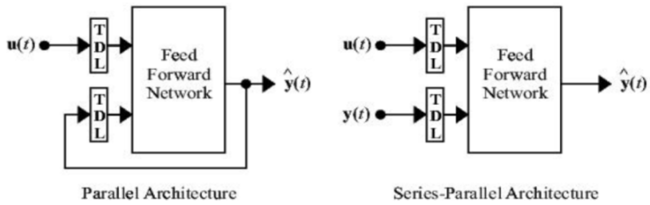
2.4. Neural Network Autoregressive Model with Exogenous Inputs

In dynamic networks, output not only depends on inputs but also depends on previous inputs, outputs and the state of network. Dynamic networks can learn the sequential or time-varying patterns. [6] NNARX is a recurrent dynamic network, with feedback connections enclosing several layers of the network. The NNARX model is based on the linear ARX model, which is commonly used in time-series modeling and forecasting. NNARX model can be represented as follows:

$$y(t) = f(y_{t-1}, y_{t-2}, \ldots, y_{t-n_y}, u_{t-1}, u_{t-2}, \ldots, u_{t-n_u}) \qquad 8.$$

Where the next value of the dependent output signal $y(t)$ is regressed on previous values of the output signal and previous values of an independent (exogenous) input signal. The output is feedback to the input of the feed-forward neural network as part of the standard NNARX architecture as shown in Figure2 left side. Since the true output is available during the training, we could create a series parallel architecture [9] in which the true output is used instead of feeding back the estimated output as shown in Figure 2 right side. This has two advantages, first is that the input to the feed-forward network is more accurate, second is that the resulting network has purely feed-forward architecture and static back propagation can be used for training. [9]

Fig.2: NNARX network



3. Performance Comparison

In order to compare the performance of different models, the following criteria as shown in table1. are used:

Table1. performance comparison criterion

Root Mean Square Error	$RMSE = \sqrt{\sum_{t=1}^{n} (F_t - X_t)^2 / n}$	Mean Absolute Percentage Error	$MAPE = [\sum_{t=1}^{n} \left\lvert \frac{F_t - X_t}{X_t} \right\rvert] / n \times 100$
Mean Absolute Error	$MAE = [\sum_{t=1}^{n} \lvert F_t - X_t \rvert] / n$	Theil's U statistics	$U = \sqrt{\frac{\sum_{t=1}^{n} (F_t - X_t)^2}{n}} / [\sqrt{\frac{\sum_{t=1}^{n} (X_t)^2}{n}} + \sqrt{\frac{\sum_{t=1}^{n} (F_t)^2}{n}}]$

For all the above formula, F denotes the forecasted value and X is the actual value, RMSE and MAE criteria depend on the scale of the dependent variable. These should be used as relative measures to compare the forecast value for the same series across different models; the smaller value means the better the forecasting performance of that model. MAPE and Theil's U are scale invariant. The Theil's U lies between zero and one where zero indicates a perfect fit.

4. Empirical Findings

Applying the best fitted models for exchange rate prediction gives the following values for both in sample and out of sample forecasting. Table 2 represents the results for in sample forecasting while Table 3 represents the results for out of sample forecasting respectively.

Table2. In sample forecasting

Model	ARIMA (0,1,0)	GARCH(1,1)	ANN(1,15.20.15.1)	NNARX (1,25,1)
RMSE	0.016429	0.016443	0.008482	0.00019189
MAE	0.010785	0.010815	0.005927	0.00006200
MAPE	126.08	144.23	0.944168	0.016960
Theil'S U	0.9074	0.8755	0.006617	0.00016658

Based on the results obtained in table 2, it is clear that when we consider the volatility of exchange rate through the GARCH model, the performance of forecasting increases (it can be shown through the comparison of different criteria). When ANN is used, the performance of forecasting by this model increases because this model used the pattern of past behavior of the exchange rate. Considering the dynamic network (NNARX) for prediction will improve the results dramatically.

Table3. out of sample forecasting

Time Periods	Criterion	ARIMA (0,1,0)	GARCH(1,1)	ANN(1,20,25,20,1)	NNARX(1,25,1)
3 steps ahead	RMSE	0.010596	0.010158	0.01586	0.007932
	MAE	0.007823	0.007598	0.01294	0.007130
	MAPE	132.40	174.20	2.9993	0.07964
	Theil'S U	0.7953	0.7256	0.01780	0.003967
6 steps ahead	RMSE	0.11376	0.010918	0.01381	0.007690
	MAE	0.009069	0.008844	0.01143	0.007683
	MAPE	147.57	181.72	2.6466	0.15535
	Theil'S U	0.8046	0.7370	0.01554	0.003871
12 steps ahead	RMSE	0.011890	0.011629	0.010399	0.007358
	MAE	0.009351	0.009239	0.008566	0.007328
	MAPE	139.81	163.65	1.9348	0.30053
	Theil'S U	0.8356	0.7803	0.011650	0.003751

According to the results in table 3, generally, in all time horizons, the performance of the dynamic model (NNARX) is better than the other models (linear and nonlinear). More specific, when we consider the volatility inside the model through the GARCH model, performance of prediction becomes better than ARIMA model in all time horizons. As the number of steps increases, the performance of ARIMA and GARCH model decreases. When it comes to the comparison between ANN and both previous models, the ANN performance increases as the number of steps increases. Finally our dynamic model (NNARX) outperformed all the previous models because in all time horizons, its performance is better. In addition, as the number of steps for prediction increases, its performance increases as well. Thus it can be concluded that for the longer forecasting period, the performance of artificial intelligence becomes better compared to the econometric models. It can be concluded that we have a better forecasting performance in a shorter period for econometric models, while artificial intelligence is better in the longer period.

5. Conclusion

Recently, the application of different models for predicting the most important variables in the economy such as exchange rate, stock market and interest rate for decision making like foreign direct investment, international trade and investment becomes more important in business and economics. In this paper four different models are applied to predict the SGD/USD exchange rate for three time horizons, i.e. 3, 6 and 12 steps ahead. Since the exchange rate exhibits a nonlinear pattern and exhibits volatility in its own behavior (we did not present the evidence due to the limited space), non-linear models predict better than linear models. The results reveal that the performance of econometric models get worse when we consider longer period of time but the performance of artificial intelligence is better in the longer period of time. Among the non-linear models, NNARX outperformed the ANN and both of them outperformed the GARCH model.

6. Acknowledgment

The authors would like to gratefully acknowledge the financial support from the fundamental research grant scheme (MOHE, Malaysia).

7. References

[1]. B. Wu. Model-free forecasting for non-linear time series (with application to exchange rates). Computational Statistics and Data Analysis. 1995, 19: 433–459.

[2]. C.B. Phillips, p. Perron. testing for unit roots in time series regression. Biometrica. 1988, 75: 335-346.

[3]. D. Dickey, W. Fuller. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association. 1979, 74: 427-431

[4]. E. Zivot, D. W. K. Andrews. Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit Root Hypothesis. Journal of Business & Economic Statistics. 1992, 10 (3): 251-270.

[5]. G. Zhang, M.Y. Hu, Neural Network Forecasting of the British Pound/US Dollar Exchange Rate. International Journal of Management Science. 1998, 26: 495–506.

[6]. I. Kamwa, R. Grondin, V. K. Sood, C. Gagnon, V. T. Nguyen, J. Mereb. Recurrent neural networks for phasor detection and adaptive identification in power system control and protection. IEEE Transactions on Instrumentation and Measurement. 1996, 45, 657–664.

[7]. J. S. Racine. On The Nonlinear Predictability of Stock Returns Using Financial and Economic Variables. forthcoming. Journal of Business and Economic Statistics. 2001. 19: 80-382.

[8]. N.B. Karayiannis, A.N. Venetsanopoulos. Artificial Neural Networks: Learning Algorithms. Performance Evaluation and Applications. Kluwer Academic Publishers. Boston. USA. 1993.

[9]. O. De Jesús, J.M. Horn, M.T. Hagan. Analysis of Recurrent Network Training and Suggestions for Improvements. Proceedings of the International Joint Conference on Neural Networks. 2001, 2632–2637.

[10]. P. Box, G.M. Jenkins. Time series analysis: Forecasting and control. San Francisco, CA: Holden-day Inc. 1976

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

[12]. T. Bollerslev. Generalized Autoregressive Conditional Heteroscedasticity. Journal of Econometrics. 1986, 31: 307-327