

Forecasting of Currency Exchange Rates using ANN: A Case Study

*Joarder Kamruzzaman
Gippsland School of Computing & IT
Monash University
Churchill, Victoria 3842
Australia*

*Ruhul A Sarker
School of Computer Science
University of NSW, ADFA Campus
Northcott Drive, Canberra 2600
Australia*

ABSTRACT

In today's global economy, accuracy in forecasting the foreign exchange rate or at least predicting the trend correctly is of crucial importance for any future investment. The use of computational intelligence based techniques for forecasting has been proved extremely successful in recent times. In this paper, we developed and investigated three Artificial Neural Network (ANN) based forecasting model using Standard Backpropagation (SBP), Scaled Conjugate Gradient (SCG) and Backpropagation with Bayesian Regularization (BPR) for Australian Foreign Exchange to predict six different currencies against Australian dollar. Five moving average technical indicators are used to build the models. These models were evaluated on five performance metrics and a comparison was made with traditional ARIMA model. All the ANN based models outperform ARIMA model. It is found that SCG based model performs best when measured on the two most commonly used metrics and shows competitive results when compared with BPR based model on other three metrics. Experimental results demonstrate that ANN based model can closely forecast the forex market.

1. INTRODUCTION

The foreign exchange market has experienced unprecedented growth over the last few decades. The exchange rates play an important role in controlling dynamics of the exchange market. As a result, the appropriate prediction of exchange rate is a crucial factor for the success of many businesses and fund managers. Although the market is well-known for its unpredictability and volatility, there exist a number of groups (like Banks, Agency and other) for predicting exchange rates using numerous techniques.

Exchange rates prediction is one of the most challenging applications of modern time series forecasting. The rates are inherently noisy, non-stationary and deterministically chaotic [3, 20]. These characteristics suggest that there is no complete information that could be obtained from the past behaviour 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 behaviour. As a result, the historical data is the major player in the prediction process. The question is how good are those predictions? The purpose of this paper is to investigate and compare two well-known prediction techniques, under different parameter settings, for several different exchange rates.

For more than two decades, Box and Jenkins' Auto-Regressive Integrated Moving Average (ARIMA) technique [1] has been widely used for time series forecasting. Because of its popularity, the ARIMA model has been used as a benchmark to evaluate some new modelling approaches [7]. However, ARIMA is a general univariate model and it is developed based on the assumption that the time series being forecasted are linear and stationary [2].

The Artificial Neural Networks, the well-known function approximators in prediction and system modelling, has recently shown its great applicability in time-series analysis and forecasting [18-21]. ANN assists multivariate analysis. Multivariate models can rely on grater information, where not only the lagged time series being forecast, but also other indicators (such as technical, fundamental, inter-marker etc. for financial market), are combined to act as predictors. In addition, ANN is more effective in describing the dynamics of non-stationary time series due to its unique non-parametric, non-assumable, noise-tolerant and adaptive properties. ANNs are universal function approximators that can map any nonlinear function without a *priori* assumptions about the data [2].

In several applications, Tang and Fishwick [15], Jhee and Lee [8], Wang and Leu [16], Hill *et al.* [6], and many other researchers have shown that ANNs perform better than ARIMA models, specifically, for more irregular series and for multiple-period-ahead forecasting. Kaastra and Boyd [9] provided a general introduction of how a neural network model should be developed to model financial and economic time series. Many useful, practical considerations were presented in their article. Zhang and Hu [21] analysed backpropagation neural networks' ability to forecast an exchange rate. Wang [17]

cautioned against the dangers of one-shot analysis since the inherent nature of data could vary. Klein and Rossin [10] proved that the quality of the data also affects the predictive accuracy of a model. More recently, Yao *et al.* [18] evaluated the capability of a backpropagation neural-network model as an option price forecasting tool. They also recognised the fact that neural-network models are context sensitive and when studies of this type are conducted, it should be as comprehensive as possible for different markets and different neural-network models.

In this paper, we apply ARIMA and ANNs for predicting currency exchange rates of Australian Dollar with six other currencies such as US Dollar (USD), Great British Pound (GBP), Japanese Yen (JPY), Singapore Dollar (SGD), New Zealand Dollar (NZD) and Swiss Franc (CHF). A total 500 weeks (closing rate of the week) data are used to build the model and 65 weeks data to evaluate the models. Under ANNs, three models using standard backpropagation, scaled conjugate gradient and Bayesian regression were developed. The outcomes of all these models were compared with ARIMA based on five different error indicators. The results show that ANN models perform much better than ARIMA models. Scaled conjugate gradient and Bayesian regression models show competitive results and these models forecasts more accurately than standard Backpropagation which has been studied considerably in other studies.

In section 2, ANN forecasting model and performance metrics are defined. Section 3 and section 4 describe experimental results and conclusion, respectively.

2. NEURAL NETWORK FORECASTING MODEL

Recently neural networks have been used for modelling nonlinear economic relationship because of its ability to extract complex nonlinear and interactive effects. Neural networks are a class of nonlinear model that can approximate any nonlinear function to an arbitrary degree of accuracy and have the potential to be used as forecasting tools in many different areas. There are many different neural net learning algorithms found in the literature. No study has been reported to analytically determine the generalization performance of each algorithm. In this study we experimented with three different neural network learning algorithms, namely standard Backpropagation (BP), Scaled Conjugate Gradient Algorithm (SCG) and Backpropagation with regularization (BPR) in order to evaluate which algorithm predicts the exchange rate of Australian dollar most accurately. In the following we describe the three algorithms briefly.

2.1 Learning Algorithms

Standard BP: BP [14] uses steepest gradient descent technique to minimize the sum-of-squared error E over all

training data. During training, each desired output d_j is compared with actual output y_j and E is calculated as sum of squared error at the output layer.

The weight ω_j is updated in the n -th training cycle according to the following equation.

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

The parameters η and α are the learning rate and the momentum factor, respectively. The learning rate parameter controls the step size in each iteration. For a large-scale problem Backpropagation learns very slowly and its convergence largely depends on choosing suitable values of η and α by the user.

SCGA: In conjugate gradient methods, a search is performed along conjugate directions, which produces generally faster convergence than steepest descent directions [5]. In steepest descent search, a new direction is perpendicular to the old direction. This approach to the minimum is a zigzag path and one step can be mostly undone by the next. In CG method, a new search direction spoils as little as possible the minimization achieved by the previous one and the step size is adjusted in each iteration. The general procedure to determine the new search direction is to combine the new steepest descent direction with the previous search direction so that the current and previous search directions are conjugate as governed by the following equations.

$$\omega_{k+1} = \omega_k + \alpha_k p_k,$$

$$p_k = -E'(\omega) + \alpha_k p_{k+1}$$

where p_k and p_{k+1} are the conjugate directions in successive iterations. α_k and β_k are calculated in each iteration. An important drawback of CG algorithm is the requirement of a line search in each iteration which is computationally expensive. Moller introduced the SCG to avoid the time-consuming line search procedure of conventional CG. SCG needs to calculate Hessian matrix which is approximated by

$$E''(\omega_k) p_k = \frac{E'(\omega_k + \sigma_k p_k) - E'(\omega_k)}{\sigma_k} + \lambda_k p_k$$

where E' and E'' are the first and second derivative of E . p_k , σ_k and λ_k are the search direction, parameter controlling the second derivation approximation and parameter regulating indefiniteness of the Hessian matrix. Considering the machine precision, the value of σ should be as small as possible ($\leq 10^{-4}$). A detailed description of the algorithm can be found in [13].

BPR: A desired neural network model should produce small error on out of sample data, not only on sample data alone. To produce a network with better generalization ability, MacKay [12] proposed a method to constrain the

size of network parameters by regularization. Regularization technique forces the network to settle to a set of weights and biases having smaller values. This causes the network response to be smoother and less likely to overfit [5] and capture noise. In regularization technique, the cost function F is defined as

$$F = \gamma E + \frac{1-\gamma}{n} \sum_{j=1}^n \omega_j^2$$

where E is the sum-squared error and γ (<1.0) is the performance ratio parameter, the magnitude of which dictates the emphasis of the training. A large γ will drive the error E small whereas a small γ will emphasize parameter size reduction at the expense of error and yield smoother network response. Optimum value of γ can be determined using Bayesian regularization in combination with Levenberg-Marquardt algorithm [4]

2.2 Forecasting Model

Technical and fundamental analyses are the two major financial forecasting methodologies. In recent times, technical analysis has drawn particular academic interest due to the increasing evidence that markets are less efficient than was originally thought [11]. Like many other economic time series model, exchange rate exhibits its own trend, cycle, season and irregularity. In this study, we used time delay moving average as technical data. The advantage of moving average is its tendency to smooth out some of the irregularity that exists between market days [19]. In our model, we used moving average values of past weeks to feed to the neural network to predict the following week's rate. The indicators are MA5, MA10, MA20, MA60, MA120 and X_i , namely, moving average of one week, two weeks, one month, one quarter, half year and last week's closing rate, respectively. The predicted value is X_{i+1} . So the neural network model has 6 inputs for six indicators, one hidden layer and one output unit to predict exchange rate. Historical data are used to train the model. Once trained the model is used for forecasting.

2.3 Performance Criteria

The forecasting performance of the above model is evaluated against a number of widely used statistical metric, namely, Normalized Mean Square Error (NMSE), Mean Absolute Error (MAE), Directional Symmetry (DS), Correct Up trend (CU) and Correct Down trend (CD). These criteria are defined in Table 1. NMSE and MAE measure the deviation between actual and forecasted value. Smaller values of these metrics indicate higher accuracy in forecasting. DS measures correctness in predicted directions. CU and CD measure the correctness of predicted up and down trend, respectively.

3. EXPERIMENTAL RESULTS

The data used in this study is the foreign exchange rate of six different currencies against Australian dollar from January 1991 to July 2002 made available by the Reserve Bank of Australia. We considered exchange rate of US dollar, British Pound, Japanese Yen, Singapore dollar, New Zealand dollar and Swiss Franc. As outlined in Section 2.2, 565 weekly data was considered of which first 500 weekly data was used is training and the remaining 65 weekly data for evaluating the model.

Table 1: Performance metrics used in the experiment.

$NMSE = \frac{\sum_k (x_k - \hat{x}_k)^2}{\sum_k (x_k - \bar{x}_k)^2} = \frac{1}{\sigma^2 N} \sum_k (x_k - \hat{x}_k)^2$
$MAE = \frac{1}{N} x_k - \hat{x}_k $
$DS = \frac{100}{N} \sum_k d_k,$ $d_k = \begin{cases} 1 & \text{if } (x_k - x_{k-1})(\hat{x}_k - \hat{x}_{k-1}) \geq 0 \\ 0 & \text{otherwise} \end{cases}$
$CU = 100 \frac{\sum_k d_k}{\sum_k t_k},$ $d_k = \begin{cases} 1 & \text{if } (\hat{x}_k - \hat{x}_{k-1}) > 0, (x_k - x_{k-1})(\hat{x}_k - \hat{x}_{k-1}) \geq 0 \\ 0 & \text{otherwise} \end{cases}$ $t_k = \begin{cases} 1 & \text{if } (x_k - x_{k-1}) > 0 \\ 0 & \text{otherwise} \end{cases}$
$CU = 100 \frac{\sum_k d_k}{\sum_k t_k}$ $d_k = \begin{cases} 1 & \text{if } (\hat{x}_k - \hat{x}_{k-1}) < 0, (x_k - x_{k-1})(\hat{x}_k - \hat{x}_{k-1}) \geq 0 \\ 0 & \text{otherwise} \end{cases}$ $t_k = \begin{cases} 1 & \text{if } (x_k - x_{k-1}) < 0 \\ 0 & \text{otherwise} \end{cases}$

The performance of a neural network depends on a number of factors, e.g., initial weights chosen, different learning parameters used during training (described in section 2.1) and the number of hidden units. For each algorithm, we trained 30 different networks with different initial weights and learning parameters. The number of hidden units was varied between 3~7 and the training was terminated at iteration number between 5000 to 10000.

The best results obtained by each algorithm are presented below.

Table 2 shows the forecasting results measured in terms of the performance metrics over 35 weeks and 65 weeks for three neural network models in case of US dollar and presents a comparison with the traditional ARIMA model.

Table 2: Forecasting results of neural network model for US Dollar.

Prediction Period	Criteria	Neural Network Model			ARIMA (1,0,1)
		SBP	SCG	BPR	
35 Week	NMSE	0.5041	0.2624	0.2787	1.0322
	MAE	0.0047	0.0035	0.0036	0.0069
	DS	71.4286	80.00	82.8571	52.9411
	CP	76.4706	82.3529	82.3529	0
	CD	70.5882	82.3529	88.2353	105.882
65 Week	NMSE	0.0937	0.0418	0.0441	1.7187
	MAE	0.0043	0.0029	0.0030	0.0171
	DS	75.3846	81.5385	83.0769	42.1875
	CP	81.5789	78.9474	78.9474	0
	CD	69.2308	88.4615	92.3077	130.8462

The results show that neural network models produce better performance than linear ARIMA model indicating its suitability for financial modelling. Both SCG and BPR forecasts are better than SBP in terms of all metrics. In our experiment this is consistently observed in all other currencies also. In terms of the most commonly used criteria, i.e., NMSE and MAE, SCG perform better than BPR in all currencies except Japanese Yen. In terms of other metrics, SCG yields slightly better performance in case of Swiss Franc, BR slightly better in US Dollar and British Pound, both perform equally in case of Japanese Yen, Singapore and New Zealand Dollar. Table 3 shows the performance metrics for other currencies. Fig 1(a)-(f) shows the actual and forecasted time series of six currency rates.

4. CONCLUSION

In this study, we investigated three ANN based forecasting models to predict six foreign currencies against Australian dollar using historical data and moving average technical indicators, and a comparison was made with traditional ARIMA model. All the ANN based models outperformed ARIMA model measured on five performance metrics. Results demonstrate that ANN based model can forecast the forex rates closely. Among the three ANN based models, SCG based model yields best results measured on two popular metrics and shows results comparable to BPBR based models when measured on three other metrics.

5. REFERENCES

[1] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, CA.

- [2] L. Cao and F. Tay, "Financial Forecasting Using Support Vector Machines," *Neural Comput & Applic*, vol. 10, pp.184-192, 2001.
- [3] G. Deboeck, *Trading on the Edge: Neural, Genetic and Fuzzy Systems for Chaotic Financial Markets*, New York Wiley, 1994.
- [4] F. D. Foresee and M.T. Hagan, "Gauss-Newton approximation to Bayesian regularization," *Proc. IJCNN 1997*, pp. 1930-1935.
- [5] M.T. Hagan, H.B. Demuth and M.H. Beale, *Neural Network Design*, PWS Publishing, Boston, MA, 1996.
- [6] T. Hill, M. O'Connor and W. Remus, "Neural Network Models for Time Series Forecasts," *Management Science*, vol. 42, pp 1082-1092, 1996.
- [7] H. B. Hwang and H. T. Ang, "A Simple Neural Network for ARMA(p,q) Time Series," *OMEGA: Int. Journal of Management Science*, vol. 29, pp 319-333, 2002.
- [8] W. C. Jhee and J. K. Lee, "Performance of Neural Networks in Managerial Forecasting," *Intelligent Systems in Accounting, Finance and Management*, vol. 2, pp 55-71, 1993.
- [9] I. Kaastra and M. Boyd, "Designing a Neural Network for Forecasting Financial and Economic Time-Series," *Neurocomputing*, vol. 10, pp215-236, 1996.
- [10] B. D. Klein and D. F. Rossin, "Data Quality in Neural Network Models: Effect of Error Rate and Magnitude of Error on Predictive Accuracy," *OMEGA: Int. Journal of Management Science*, vol. 27, pp 569-582, 1999.
- [11] B. LeBaron, "Technical trading rule profitability and foreign exchange intervention," *Journal of Int. Economics*, vol. 49, pp. 124-143, 1999.
- [12] D.J.C. Mackay, "Bayesian interpolation," *Neural Computation*, vol. 4, pp. 415-447, 1992.
- [13] A.F. Moller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks*, vol. 6, pp.525-533, 1993.
- [14] D.E. Rumelhart, J.L. McClelland and the PDP research group, *Parallel Distributed Processing*, vol. 1, MIT Press, 1986.
- [15] Z. Tang and P. A. Fishwick, "Back-Propagation Neural Nets as Models for Time Series Forecasting," *ORSA Journal on Computing*, vol. 5(4), pp 374-385, 1993.
- [16] J. H. Wang and J. Y. Leu, "Stock Market Trend Prediction Using ARIMA-based Neural Networks," *Proc. of IEEE Int. Conf. on Neural Networks*, vol. 4, pp. 2160-2165, 1996.
- [17] S. Wang (1998) An Insight into the Standard Back-Propagation Neural Network Model for Regression Analysis, *OMEGA: Int. Journal of Management Science*, 26, pp133-140.
- [18] J. Yao, Y. Li and C. L. Tan, "Option Price Forecasting Using Neural Networks," *OMEGA: Int. Journal of Management Science*, vol. 28, pp 455-466, 2000.
- [19] J. Yao and C.L. Tan, "A case study on using neural networks to perform technical forecasting of forex," *Neurocomputing*, vol. 34, pp. 79-98, 2000.
- [20] S. Yaser and A. Atiya, "Introduction to Financial Forecasting," *Applied Intelligence*, vol. 6, pp 205-213, 1996.
- [21] G. Zhang and M. Y. Hu, "Neural Network Forecasting of the British Pound/US Dollar Exchange Rate," *OMEGA: Int. Journal of Management Science*, 26, pp 495-506, 1998.

Table 3: Prediction performance for other currencies. The first figure in each cell shows the metric on 35 weeks prediction while the second figure on 65 weeks prediction.

Currency	SCG NN Model					BPR NN Model				
	NMSE	MAE	DS	CU	CD	NMSE	MAE	DS	CU	CD
B. Pound	0.1578 0.0729	0.0030 0.0023	77.14 84.61	81.25 87.87	73.68 83.87	0.1724 0.0790	0.0031 0.0024	82.85 87.69	93.75 93.93	73.68 83.87
J. Yen	0.1264 0.0411	0.6243 0.5188	80.00 81.53	81.81 83.78	76.92 78.57	0.1091 0.0367	0.5806 0.5043	80.00 81.53	81.81 83.78	76.92 78.57
S. Dollar	0.2321 0.0760	0.0076 0.0060	82.85 86.15	82.35 88.23	83.33 83.87	0.2495 0.0827	0.0080 0.0063	82.85 86.15	82.35 88.23	83.33 83.87
NZDollar	0.0878 0.0217	0.0038 0.0033	85.71 84.61	87.50 82.14	84.21 88.88	0.0898 0.0221	0.0039 0.0033	85.71 84.61	87.50 82.14	84.21 88.88
S. Franc	0.0485 0.0389	0.0059 0.0052	82.85 84.61	80.00 84.61	86.66 86.66	0.0496 0.0413	0.0057 0.0051	80.00 81.53	75.00 77.14	86.66 86.66

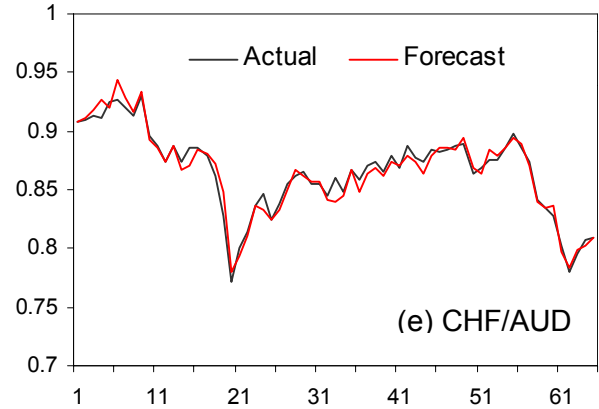
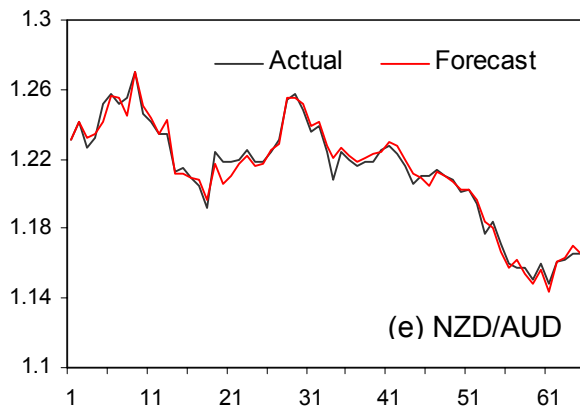
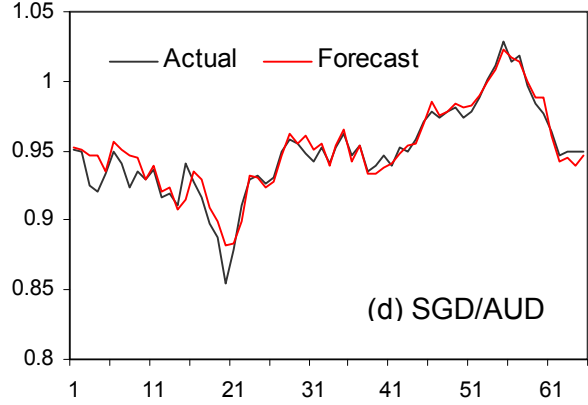
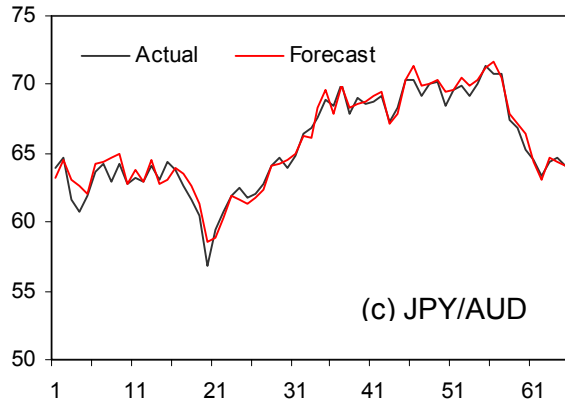
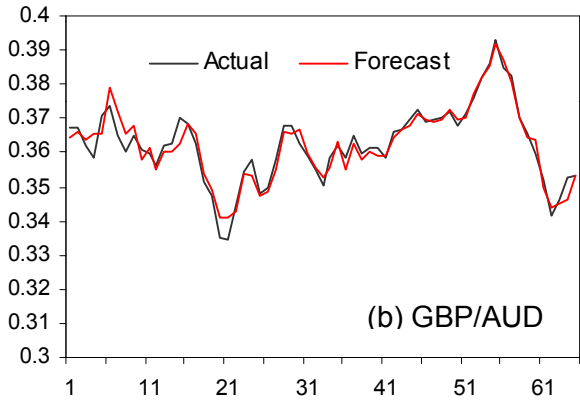
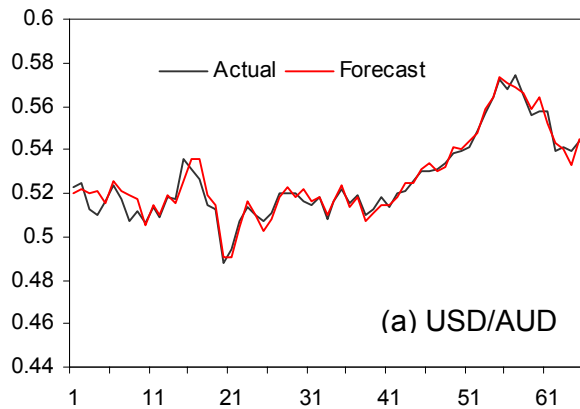


Fig. 1. Forecasting of different currencies by SCG based neural network model over 65 weeks.