Management Science Letters 2 (2012) 561–570

Contents lists available at GrowingScience

Management Science Letters

homepage: www.GrowingScience.com/msl

Exchange rate prediction with multilayer perceptron neural network using gold price as external factor

Mohammad Fathian and **Arash N. Kia***

Department of Industrial Engineering, Iran University of Science & Technology, Tehran

A R T I C L E I N F O

Article history:
Received October 10, 2011
Received in Revised form November, 14, 2011
Accepted 8 December 2011
Available online 12 December 2011

Keywords:
Forecasting
Artificial neural networks
Multilayer perceptron
Exchange rate
Gold price
Time series

A B S T R A C T

In this paper, the problem of predicting the exchange rate time series in the foreign exchange rate market is going to be solved using a time-delayed multilayer perceptron neural network with gold price as external factor. The input for the learning phase of the artificial neural network are the exchange rate data of the last five days plus the gold price in two different currencies of the exchange rate as the external factor for helping the artificial neural network improving its forecast accuracy. The five-day delay has been chosen because of the weekly cyclic behavior of the exchange rate time series with the consideration of two holidays in a week. The result of forecasts are then compared with using the multilayer peceptron neural network without gold price external factor by two most important evaluation techniques in the literature of exchange rate prediction. For the experimental analysis phase, the data of three important exchange rates of EUR/USD, GBP/USD, and USD/JPY are used.

1. Introduction

Concerning the size and importance of financial markets, foreign exchange market (FOREX) is the biggest and the most important one in the world (Baillie & McMahon, 1989). This important financial market has a 3.98$ trillion US dollars turnover per day (Bank of International Settlements, 2010). Relying upon the reasons a continuous growth is predictable for the FOREX in future (Amiri et al., 2010). Lack of a central or physical place is a distinguishing characteristic of FOREX as an over the counter (OTC) market. This is the feature by which FOREX- a 24*7 market- is differentiated from other stock markets with artificial time regulations that cause limitations in modeling and prediction methods (Berardia, & Serva, 2005). The collapse of Bretton-Woods agreement, which was an attempt to define the exchange rate in a fixed regime, resulted in a rise in the number of researches on the prediction of exchange rate time series in 1970s.

There are some items such as supply and demand of the currencies and many other macro-economic parameters as contributing factors to the definition of exchange rate and a float regime of exchange

* Corresponding author. Tel.: +98 912 1909437; Fax: +98 21 73225000.
E-mail addresses: arash.nkia@gmail.com (Arash N. Kia)

doi: 10.5267/j.msl.2011.12.008

rate emerged (McFarland et al., 1982; Friedman, & Vandersteel, 1982; Cornell et al., 1978). The termination of Bretton-Woods agreement gave rise to a series of researches aimed at making different models for exchange rate prediction (Ni, & Yin, 2009).

Different parameters such as political, economic, and psychological ones affect exchange rate fluctuations (Zhang & Wan, 2007). Empirical results reveal a close nonlinear relationship between the exchange rate and different macro-economic variables (Lee & Wong, 2007). There are many evidences, which imply significant effects of the intervention of central banks on the exchange rate (Baillie & Osterberg, 1997). The interest rate is another factor, which influences the exchange rate (Gourinchasa & Tornell, 2004). The researches indicate that some factors such as micro-traders and financial companies affect the exchange rate time series more significantly than macro-economic parameters (Devereuxa, & Engel, 2002). Financial time series such as the exchange rate time series are also affected by the news (Andersen et al., 2007). The existence of non-periodic cycles in the structurally complicated exchange time series is approved by the researches (Karuppiaha, & Los, 2005). As mentioned above there are much evidences, which indicate the complexity of predicting exchange rates as nonlinear trend. Gaining more profit and determining the possibility of predicting the market can be enumerated as two main reasons for predicting exchange rate time series (Zhang & Wan, 2007).

The review of the related literature shows that the modeling and prediction of exchange rate time series have been investigated from three different aspects of economic, statistical, and artificial intelligence or machine learning. By employing the economic theories and macro-economic parameters some people have tried to explain the fluctuations of exchange rate time series from the economic point of view (Gourinchasa, & Tornell, 2004; Devereuxa, & Engel, 2002; Kohli, 2003; Kim, & Sheen, 2006). The other types of researches were conducted by using conventional statistical methods with the aim of explaining the exchange rate time series attributes and prediction of these time series (Berardia & Serva, 2005; Chen et al., 2008; Baviera et al., 2002; Yu et al., 2005). The last group of researches (Ni & Yin, 2009; Zhang, & Wan, 2007; Lee & Wong, 2007; Yu et al., 2005; Yu et al., 2009; Yao, & Tan, 2000; Anastasakis & Mort, 2009) where the model and predict exchange rate time series with the help of machine learning and artificial intelligence is the focal point in their studies. There are many instances of the researches to use neural networks to model and predict financial time series (Bildirici & Ersin, 2009). These researches have employed machine-learning techniques such as artificial neural networks in modeling the time series with the belief that complex and nonlinear exchange rate time series cannot be modeled and predicted by conventional statistical methods.

Since financial time series are complex, the researchers consider the processes, which change the time series data as a black box and they just study the fluctuations of the series (Anastasakis & Mort, 2009). Contrary to the fundamental analysis approach in which the prediction of financial markets depends on the processes, which affects the prices, this is a technical analysis approach towards prediction. The chartists who try to predict financial markets such as foreign exchange rate market by the technical analysis methods use technical indicators and statistical methods. Among the statistical methods for prediction of time series, ARIMA is considered as one of the most famous and most powerful ones. In order to predict nonlinear time series such as exchange rate time series, it is better to employ neural networks that are nonlinear in nature.

According to the researches, artificial neural networks are better predictors of exchange rates than the ARIMA models for two main reasons: First, ARIMA models are linear. Second, time series should be stationary in mean and variance if modeling with ARIMA (Huang, & Wu, 2008). Multilayer Perceptron (MLP) neural networks, radial basis function (RBF) neural networks, and self organizing map (RBF) networks are three techniques used mostly and to a great degree in artificial neural networks (Haykin, 1998). Sometimes a hybrid model of different types of artificial neural networks and statistical models are used to achieve a better prediction (Ni, & Yin, 2009; Zhang, & Wan, 2007; Lee, & Wong, 2007; Yu et al., 2005).

The use of an external factor has not been widely discussed in the literature of exchange rate prediction. Many economic and macro-economic variables are dependent to each other and fluctuation in one of them can make fluctuation in others. In this research, the researchers use gold price to see whether it could help predicting the exchange rate time series with MLP neural networks or not. The results of both using and not using gold price factor are tested to see which approach makes a better and more accurate prediction. The universal gold price comes from London market but the gold price in each countries market will be different due to the transportation costs and many other reasons like different central bank policies or different New Year holidays, which affect the price and many more. Because of this, the price of gold in two different countries is used together for achieving better results.

After discussing the theories behind MLP neural networks in brief, the next section of the paper goes directly through the methodology phase and discusses the proposed model used for prediction. Section 3 discusses the data preparation and compilation process used as the input of neural network model. Section 4 talks about the evaluation methodology used in the literature of exchange rate prediction and also in this research. In section 5, the results of both models of using and not using gold price in neural networks are shown and the next section analyzes and discusses the results. The last two sections are for conclusion and some further research suggestions.

1.1 MLP Neural Networks

Unidentified relation between the input and output data can be modeled by neural networks (Bildirici et al., 2010). A type of neural networks, which is most commonly used in the literature of artificial intelligence is multilayer perceptions. In neural networks, one or more than one input nodes are connected to one or more than one output nodes by the means of one or more than one hidden layers. The input will be changed under the effect of the hidden layers, which behave in nonlinear way and the results of the changes multiply the weights will be transferred to the output nodes. An approximation model of any nonlinear complex function can be made via the learning and training phase of this process (Kay & Titterington, 2000). A multilayer perceptron neural network with two hidden layers is depicted in Fig. 1.

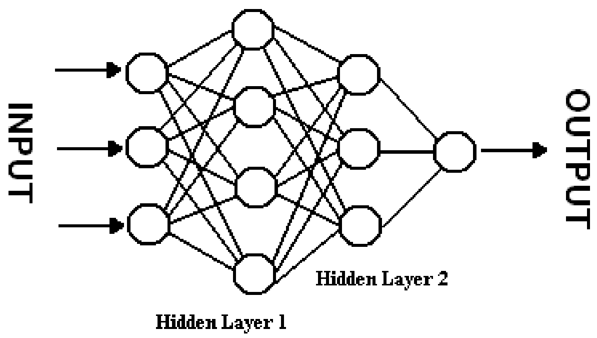

Fig 1. A simple schema of a two layer perceptron neural network

For representing a network structure like the one shown in Fig. 1 there is a way such as (3-4-3-1) which is used in this paper and some other researches (O'Connor & Madden, 2006). This way of representation says that the neural network has 3 inputs and two hidden layers with 4 and 3 nodes and 1 output node. Here a simple single layer perceptron is illustrated in the form of equation in Eq. (1).

$$o = f\left(\sum_{i=0}^{n} w_i \, \emptyset(x)\right). \tag{1}$$

In Eq. (1), the output result of the MLP could be a number or a vector of zero numbers. The variable, x, is the input vector or number, and $\emptyset$ is called the activation function. The variable, w, is the weight of any edge in the structure of the neural network and f is the output function. Activation function, which is the core of any node in the hidden layers of MLP changes the input of the node to the output. Log-sigmoid and hyperbolic tangent are the most famous activation functions, which can be found in the neural networks literature. The activation function used in this research is log-sigmoid and Log-sigmoid function is presented in Eq. (2).

$$f(a) = \frac{1}{1+e^{-a}}. \tag{2}$$

As it was referred to earlier, the MLP neural network can have more than one hidden layer and each hidden layer can transfer its output to the next hidden layer. Therefore, a multilayer perceptron with two hidden layers can be illustrated as Eq. (3).

$$o_k = f(\sum_j w_{kj}^{(2)} g(\sum_i w_{ij}^{(1)} x)). \tag{3}$$

In Eq. (3), x is the input vector and w is the weight of any given edge in the neural network structure, $w_{ij}^{(1)}$ is considered as the weights of the first hidden layer, and $w_{kj}^{(2)}$ is representation of the weights of the second hidden layer. The function g is the activation function, and the function f is the output function. Finally, o is the output produced by the multilayer perceptron neural network.

2 Methodology

The use of external factors for improving the prediction accuracy has been used in many financial time series but there is a lack of studying external factors in foreign exchange rate time series. In a research by O'Connor and Madden in 2006, the oil price was used to improve the forecast of Dow Jones Index. The relationship between some macro-economic variables like interest rate and exchange rate (Gourinchass, & Tornell, 2004), has been studied in some researches but the focus of these studies were not on building prediction model. It was just on the economical concepts of these relationships.

Many factors can be considered in a research for forecast modeling of exchange rate time series. These factors should have some characteristics: they should have different daily data, which are sensitive to market fluctuations just like the exchange rate data. Some data like interest rate time series are constant for a period and cannot follow the fluctuations of market. Therefore, they are not helpful in improving the forecast of exchange rate time series. The oil price, the stock market indices, and the gold price were those considered in this research. The oil price was used in the study of O'Connor and Madden in 2006 for prediction of stock market index. The problem of stock market indices as an external factor in prediction of exchange rate time series was that there were many indices tested to see which one would improve the prediction. The gold price had none of these problems and therefore it was used to see whether it would improve the neural network exchange rate prediction model or not.

The difference between the proposed neural network model and other neural network models in the literature is not in the structure of it but in the input. If considering the exchange rate as the first currency divided to the second currency, the time series of the gold price in the first currency market and the second currency market are both used as external factor and input of our prediction MLP neural network. Therefore, the exchange rate series themselves and two gold price time series will be fed into the model.

The schema of the model is presented in Fig. 2, which is obvious that a MLP neural network is used and the input data are fed into the model in an n-day time delay format. In this research n is 5 as it was discussed before that 5 day delay is for the weekly cyclic behavior of the input series minus two days off for weekly holidays (7 – 2 = 5).

The reason of presenting the model as a triangle was to emphasize on the pyramid structure of the MLP models used in learning phase of the neural network. The pyramid structure is faster and easier to be learnt and therefore many different structures could be tested in the trial and error of building the best model for prediction to find the best one. In the pyramid structured neural networks, the number of neurons of the hidden layers in each layer is less than the previous layer.

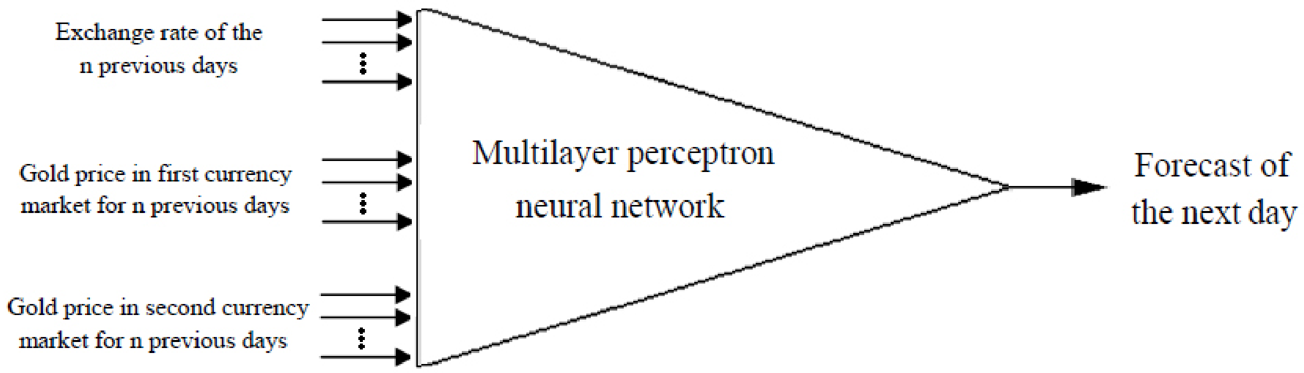

Fig. 2. Schema of the n-day delay model using gold price as external factor

3 Data preparation and compilation

The time series of GBP/USD, EUR/USD, and USD/JPY are used as the input of the model. Also one ounce gold price in the London, New York, Tokyo, and European market (gold price in countries which use euro as their currency), are used as the external factors. These external factors are also fed into the model as input. Exchange rate data come from the Federal Reserve Bank of St. Luis economic research center, which is accessible to all the researchers from their website[1]. The gold price of all the local markets comes from the yahoo finance website, which gathers many historical data of financial time series[2].

Exchange rate and gold price data are from first of April 2001 to 31 July 2011. This includes 3440 data in each time series. The exchange rate data are the mean of buy rate and sell rate from the Federal Reserve Bank of New York at 10 AM each day. The Saturday and Sunday holidays were missing data of the time series. We used the mean of 3 days before to fill the missing data of the weekly holidays. In technical analysis it is called MA3[3]. When a 5 day delay MLP model is used the first output of the model for the learning phase will be the 6th of April 2001 data. The 5 day delay model works better because of the 5 day data of the week, representing the weekly nature of the time series. When learning a neural network model, the data should be partitioned into three sets called training, validation, and test sets. The training and validation sets are used in the model building

[1] Federal Reserve Bank of St. Luis Economic Research Center, http://research.stlouisfed.org [online]
[2] Financial Site of Yahoo, http://finance.yahoo.com [online]
[3] Moving average of 3 days before.

phase. Training set is used for learning the model and validation set is used for finding the best structure for the model. The testing set will be used for the evaluation with the help of evaluation parameters that are described in the next section. According to Yao and Tan (2000), the best ratio for partitioning the data into three sets is 7:2:1 and this ratio are also used for this research. Table 1 shows the date range of training, validation, and testing set in this study.

Table 1
Data partitioning for training, validation, and test sets

7:2:1 Ratio	Start Date	Finish Date	Total Data
Training Set	2001/4/1	2007/10/22	2395
Validation Set	2007/10/23	2009/9/10	690
Testing Set	2009/9/11	2010/7/31	355

4. Forecast evaluation in financial time series

There are three important evaluation criteria for the prediction of financial time series (Yao, & Tan, 2000). The first type of criterion is the class of mean square error (MSE) evaluation techniques, like root mean square error (RMSE) or normalized mean square error (NMSE) or mean square error (MSE). All these evaluative values try to find the difference between the original time series and the predicted time series in terms of level. This means that they just emphasize on how far or near the prediction is from the real series in terms of value. The RMSE is presented in Eq. (4), T is the number of the items in the series and x_t is the t'th element, in the original time series and x_t is the t'th element, of the predicted time series.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}(x_t - \widehat{x_t})^2}{T}}. \tag{4}$$

The second evaluation method in financial time series prediction is a more practical method. The question of why the traders try to forecast a financial series comes from the need to gain more profit by understanding the future of market prices. Therefore, a good prediction is the one, which makes the trader more profit. By defining a trading strategy and assuming to have a certain amount of money at first of the trade, the trader can find out if trading with the predicted series, how much profit or loss he/she would make at the end. The model, which gives the trader more profit will be a better model for prediction. Using this method for evaluation depends on trading strategy and needs a simulation of the trade so that it could be found out how much profit or loss will be gained at the end, using the model.

The third evaluation method is more applicable than the second one and does not rely on the trading strategy or the amount of the money the traders start with. This evaluation method is widely used in the literature of exchange rate prediction (Yu et al., 2005; Yu et al., 2009; Yao, & Tan, 2000; O'Connor, & Madden, 2006). This method is called directional success or directional status. The value of directional status shows the correctness of the prediction in terms of direction, which means if the original series goes up or down in the value the predicted series behaves the same. The importance of this evaluation method is due to the way the trader trades in the real market. If the price of a commodity is expected to rise the traders will buy that commodity or won't sell it and if the price falls the traders will sell it or won't buy it. Due to this strategy, having a good directional success in the prediction is the most important evaluative criteria in the field of financial time series modeling and forecasting. The directional success value can be calculated as presented in Eq. (5) where the description of the parameters are like those in Eq. (5), and Dstat is the directional status evaluative.

$$D_{stat} = {}^{1}/_{T} \sum_{t=1}^{T} a_k \qquad (5)$$

$$if\ (x_{t+1} - x_t)(\hat{x}_{t+1} - x_t) \geq 0$$

$$then\ a_k = 1$$

$$else\ a_k = 0$$

5. Experimental results and analysis

The parameters used for the learning phase are presented in Table 2. In Table 3, the structures of the MLP networks built for the three time series of GBP/USD, EUR/USD, and USD/JPY are shown. Both the MLP models using the gold price as external factors and simple MLP model structures are presented in Table 3.

Table 2
Parameters of the learning phase for MLP neural networks

	Alpha	Initial Eta	Low Eta	High Eta	Eta Decay
MLP	0.9	0.3	0.01	0.1	30
RBF	0.9	0.4	-	-	-

Table 3
MLP Structures after the learning phase

	MLP without using external factor	MLP with gold price as external factor
EUR/USD	5-14-9-1	15-9-8-1
GBP/USD	5-8-5-1	15-7-6-1
USD/JPY	5-7-11-1	15-18-11-1

After the learning phase and building the model, the graph of the original and predicted time series for the last 50 data of the test set are shown in the Fig. 2.

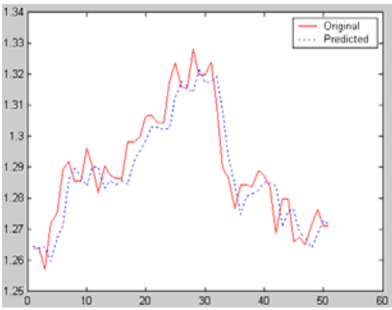

Fig 3.a. EUR/USD

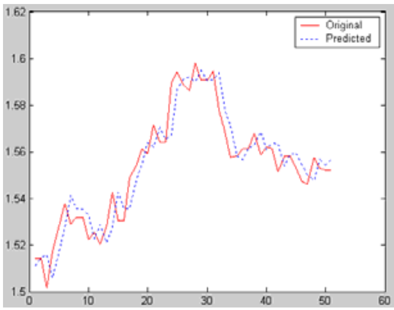

Fig 3.b. GBP/USD

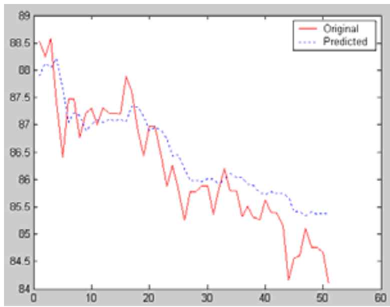

Fig 3.c. USD/JPY

Currency prediction with MLP using gold price as external factor

The evaluation results of both MLP models with and without using the external factors are presented in Table 4 and Table 5. The evaluative results are from RMSE and Dstat parameters for both modeling approaches. The RMSE comparison of two modeling approaches shows that the error level of the models with external factor is lower than the models that do not use the external factor. The Dstat parameter, that according to the section 4 of this study is a more important evaluative parameter in the business world, shows that models with external factor achieve a higher percentage of success in direction prediction.

Table 4
The RMSE factor for both modeling approaches

RMSE	MLP without using external factor	MLP with gold price as external factor
EUR/USD	0.0085	0.0083
GBP/USD	0.0098	0.0099
USD/JPY	0.6715	0.5955

Table 5
The Dstat parameter for both modeling approaches.

Dstat	MLP without using external factor	MLP with gold price as external factor
EUR/USD	%60	%62.3
GBP/USD	%61.4	%62.6
USD/JPY	%62	%64.3

6. Discussion on USD/JPY prediction

By comparing the prediction results of EUR/USD, GBP/USD, and USD/JPY, we will see that there is an obvious weakness in forecast of USD/JPY. This is not something that has happened just for this research and the weakness in forecasting USD/JPY exchange rate compared with other exchange rates has already been specified by other studies. Fig. 4 shows the prediction of USD/JPY in the research of Yao and Tan in 2000. The gap between the original series and the predicted series is discussed there and the reason is said to be the market efficiency of USD/JPY, which is higher than other exchange rates, which makes the prediction harder for the models. It is worth noting that the data used in the Yao and Tan research was from November 1993 to July 1995. This study confirms the results of other studies like Yao and Tan in 2000 about the efficiency of USD/JPY exchange rate market. The efficiency of the market makes the movements in the market more randomly; hence the prediction will be harder. Our model shows better results than the simple MLP model presented by Yao and Tan in 2000.

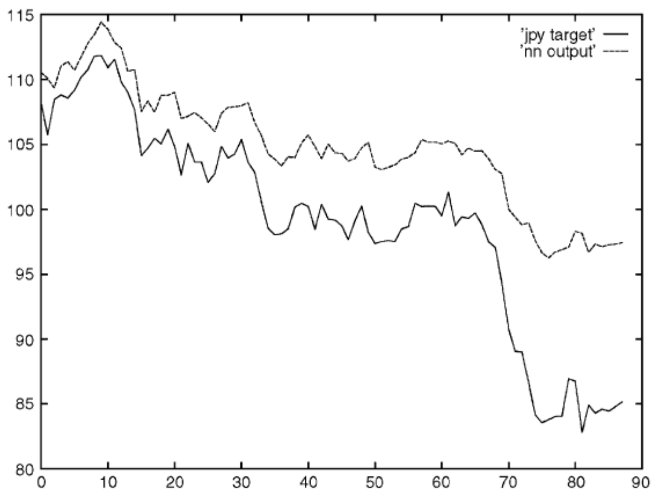

Fig 4. Predicting USD/JPY by simple neural network, (Yao & Tan, 2000)

7. Conclusions and further researches

The research shows that using the gold price as the input of the MLP neural networks improves the forecasting ability of the models. This is proved by both RMSE and Dstat evaluative parameters in the financial forecasting literature. The results for all the three series of EUR/USD, GBP/USD, and USD/JPY show that using gold price as the external factor improves the forecast. The improvement is less for the USD/JPY and this shows the efficiency of the USD/JPY market in comparison to the EUR/USD, and GBP/USD market as discussed in the previous section.

This study can be done by using RBF neural network instead of MLP neural networks or it can be done by using hybrid neural network or other hybrid modeling techniques that use nonlinear machine learning methodologies. Other exchange rates can be tested with this model to see whether the results of this study are confirmable or not. For example, fixed exchange rates like RMB of China can be fed to the model to see how much improvement will be achieved using the gold price external factor for fixed exchange rate regimes.

References

Bildirici, M., & Ersin, O.O. (2009). Improving forecasts of GARCH family models with the artificial neural networks: An application to the daily returns in Istanbul Stock Exchange. *Expert Systems with Applications*, 7355-7362.

Bildirici, M., Alp, E.A., & Ersin, O.O. (2010). TAR-Cointegration neural network model: An empirical analysis of exchange rates and stock returns. *Expert Systems with Applications*, 2-11.

Box, G.E.P., Jenkins, G.M., & Reinsel, G.C. (1994). *Time Series Analysis: Forecasting and Control*. Englewood Cliffs, NJ: Prentice-Hall.

Chen, Ch.I., Chen, H.L, & Chen, Sh.P. (2008). Forecasting of foreign exchange rates of Taiwan's major trading partners by novel nonlinear Grey Bernoulli model NGBM(1,1). *Communications in Nonlinear Science and Numerical Simulation*, 13, 1194–1204.

Cornell, W.B, Dietrich, Kimbell, J. (1978). The efficiency of the market for foreign exchange rates. Review of Economic Studies. *Journal of Finance*, 60, 111– 120.

Devereuxa, M.B., & Engel, Ch. (2002). Exchange rate pass-through, exchange rate volatility, and exchange rate disconnect. *Journal of Monetary Economics*, 49, 913–940.

Friedman, D., & Vandersteel, S. (1982). Short-run fluctuations in foreign exchange rates evidence from data, 1973–79. *Journal of International Economics*, 13, 171– 186.

Gourinchasa, P.O., & Tornell, A. (2004). Exchange rate puzzles and distorted beliefs. *Journal of International Economics*, 64, 303– 333.

Haykin, S. (1998, July 16). *Neural Networks—A Comprehensive Foundation* (2nd ed.). Englewood Cliffs: Prentice-Hall.

Huang, Sh.Ch., & Wu, T.K. (2008). Integrating GA-based time-scale feature extractions with SVMs for stock index forecasting. *Expert Systems with Applications*, 35, 2080–2088.

Karuppiaha, J., & Los, C.A. (2005). Wavelet multiresolution analysis of high-frequency Asian FX rates, Summer 1997. *International Review of Financial Analysis*, 14, 211– 246.

Kay, J.W., & Titterington, D.M. (2000). *Statistics and neural networks: advances at the interface*. USA: Oxford University Press.

Kim, S.J, & Sheen, J. (2006). Interventions in the Yen–dollar spot market: A story of price, volatility and volume. *Journal of Banking & Finance*, 30, 3191–3214.

Kohli, R. (2003). Real exchange rate stabilisation and managed floating: exchange rate policy in India, 1993–2001. *Journal of Asian Economics*, 14, 369–387.

Lee, V.C.S., & Wong, H.T. (2007). A multivariate neuro-fuzzy system for foreign currency risk, management decision making. *Neurocomputing*, 70, 942–951.

McFarland, J.W., Pettit, Richardson, R., & Ksung, S. (1982). The distribution of foreign exchange price changes: Trading day effects and risk management. *Journal of Finance*, 37, 693– 715.

Ni, H., & Yin, H. (2009). A self-organising mixture autoregressive network for FX time series modelling and prediction. *Neurocomputing*, 72, 3529–3537.

O'Connor, N., & Madden, M.G. (2006). A neural network approach to predicting stock exchange movements using external factors. *Knowledge-Based Systems*, 19, 371-378.

White, H. (1989). Learning in artificial neural networks: a statistical perspective. *Neural Computing*, 1, 425-464.

Yao, J., & Tan, C.J. (2000). A case study on using neural networks to perform technical forecasting of FOREX. *Neurocomputing*, 34, 79-98.

Yu, L., Lai, K., & Wang, S. (2008). Multistage RBF neural network ensemble learning for exchange rate forecasting. *Neurocomputing*, 71, 3295-3302.

Yu, L., Wang, S., & Lai, K. (2009). A neural-network-based nonlinear metamodeling approach to financial time series forecasting. *Applied Soft Computing*, 9, 563-574.

Yu, L., Wang, S., & Lai, K. (2005). A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates. *Computers & Operations Research*, 32, 2523-2541.

Zhang, Y.Q., & Wan, X. (2007). Statistical fuzzy interval neural networks for currency exchange rate time series prediction. *Applied Soft Computing*, 70, 1149–1156.