



ELSEVIER

Neurocomputing 34 (2000) 79–98

NEUROCOMPUTING

www.elsevier.com/locate/neucom

A case study on using neural networks to perform technical forecasting of forex

Jingtao Yao^{a,*}, Chew Lim Tan^b

^a*Department of Information Systems, Massey University, Palmerston North, New Zealand*

^b*School of Computing, National University of Singapore, Singapore 119260, Singapore*

Received 15 November 1997; accepted 12 April 2000

Abstract

This paper reports empirical evidence that a neural network model is applicable to the prediction of foreign exchange rates. Time series data and technical indicators, such as moving average, are fed to neural networks to capture the underlying “rules” of the movement in currency exchange rates. The exchange rates between American Dollar and five other major currencies, Japanese Yen, Deutsch Mark, British Pound, Swiss Franc and Australian Dollar are forecast by the trained neural networks. The traditional rescaled range analysis is used to test the “efficiency” of each market before using historical data to train the neural networks. The results presented here show that without the use of extensive market data or knowledge, useful prediction can be made and significant paper profits can be achieved for out-of-sample data with simple technical indicators. A further research on exchange rates between Swiss Franc and American Dollar is also conducted. However, the experiments show that with efficient market it is not easy to make profits using technical indicators or time series input neural networks. This article also discusses several issues on the frequency of sampling, choice of network architecture, forecasting periods, and measures for evaluating the model’s predictive power. After presenting the experimental results, a discussion on future research concludes the paper. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Foreign exchange rate; Neural network; Forecasting; Time series

1. Introduction

The foreign exchange market is the largest and most liquid of the financial markets, with an estimated \$1 trillion traded every day. Foreign exchange rates are amongst

* Corresponding author.

E-mail address: j.t.yao@massey.ac.nz (J. Yao).

the most important economic indices in the international monetary markets. The forecasting of them poses many theoretical and experimental challenges. Given the abandonment of the fixed exchange rates, the implementation of the floating exchange rate system in the 1970s, and the recent GATT talks towards further liberalization of trade, understanding the relationship among major currency exchange rates is of utmost importance. Foreign exchange rates are affected by many highly correlated economic, political and even psychological factors. The interaction of these factors is in a very complex fashion. Therefore, to forecast the changes of foreign exchange rates is generally very difficult. Researchers and practitioners have been striving for an explanation of the movement of exchange rates. Thus, various kinds of forecasting methods have been developed by many researchers and experts. Technical and fundamental analysis are the basic and major forecasting methodologies which are in popular use in financial forecasting. Like many other economic time series, forex has its own trend, cycle, season, and irregularity. Thus to identify, model, extrapolate and recombine these patterns and to give forex forecasting is the major challenge.

Traditionally, statistical models such as Box–Jenkins [2] models dominate the time series forecasting. White [20] suggested that the relationship between neural networks and conventional statistical approaches for time series forecasting are complementary. Refenes et al. [11] also indicated that traditional statistical techniques for forecasting have reached their limitation in applications with nonlinearities in the data set such as stock indices. Neural Network technology has seen many application areas in business especially when the problem domain involves classification, recognition and predictions. According to a survey research conducted by Wong et al. [21], more than 127 neural network business applications had been published in international journals up to September 1994. The number rose to 213 after a year [22].

This paper describes the application of neural networks in foreign exchange rates forecasting between American Dollar and five other major currencies, Japanese Yen, Deutsch Mark, British Pound, Swiss Franc and Australian Dollar. Technical indicators and time series data are fed to neural networks to capture the underlying “rules” of the movement in currency exchange rates. The rescaled range analysis is conducted to test the “efficiency” of each market before using historical data to train the neural networks. The results presented in this paper show that without the use of extensive market data or knowledge, useful prediction can be made and significant paper profit can be achieved for out-of-sample data with simple technical indicators. A further research on exchange rates between Swiss Franc and American Dollar is also conducted. However, the experiments show that with the efficient market it is not easy to profit using technical indicators or time series input neural networks. This article discusses several issues on the frequency of sampling, choice of network architecture, forecasting periods, and measures for evaluating the model’s predictive power. After presenting the experimental results, a discussion on future research concludes the paper.

2. Neural networks as a forecasting tool for foreign exchange rate

Foreign exchange rates were only determined by the balance of payments at the very beginning. The balance of payments was merely a way of listing receipts and payments in international transactions for a country. Payments involve a supply of the domestic currency and a demand for foreign currencies. Receipts involve a demand for the domestic currency and a supply of foreign currencies. The balance was determined mainly by the import and export of goods. Thus, the prediction of the exchange rates was not very difficult at that time. Unfortunately, interest rates and other demand–supply factors had become more relevant to each currency later on. On top of this the fixed foreign exchange rates was abandoned and a floating exchange rate system was implemented by industrialized countries in 1973. Recently, proposals towards further liberalization of trades are discussed in General Agreement on Trade and Tariffs. Increased Forex trading, and hence speculation due to liquidity and bonds, had also contributed to the difficulty of forecasting Forex.

The application of forecasting method includes two basic steps: analyzing data series and selecting the forecasting method that best fits the data series. Generally, there are three schools of thought in terms of the ability to profit from the financial market. The first school believes that no investor can achieve above average trading advantages based on the historical and present information. The major theory includes the *Random Walk Hypothesis* and *Efficient Market Hypothesis* [12]. The second school's view is that of fundamental analysis. It looks in depth at the financial condition of each country and studies the effects of supply and demand on each currency. Technical analysis belongs to the third school of thought which assumes that the exchange rates move in trends and these trends can be captured and used for forecasting. It uses such tools as charting patterns, technical indicators and specialized techniques like Gann lines, Elliot waves and Fibonacci series [14].

To maximize profits from the liquidity market, more and more “best” forecasting techniques are used by different traders. Nowadays, assisted with powerful computer technologies, traders no longer rely on a single technique to provide information about the future of the markets but rather use a variety of techniques to obtain multiple signals. Classical time series analysis, based on the theory of stationary stochastic processes [4] does not perform satisfactorily on economic time series. Economic data are not simple autoregressive-integrated-moving-average (ARIMA) processes; they are not described by simple linear structural models; they are not simple white noise or even random walks. Hence the major challenge ahead is the development of new methods, or the modification or integration of existing ones, that are capable of accurately forecasting series whose patterns or relationships change over time.

Because of the high volatility, complexity, and noise market environment neural network techniques are prime candidates for prediction purpose. Refenes et al. [10] applied a multi-layer perceptron network to predict the exchange rates between American Dollar and Deutsch Mark, and to study the convergence issues related to network architecture. Poddig [15] studied the problem of predicting the trend of the American Dollar–German Mark exchange rates, and compared results to regression

analysis. Pi [13] proposed tests for dependence among exchange rates. Other examples using neural networks in currency applications include [3,7,9,18,25,23].

Neural networks are an emerging and challenging computational technology and they offer a new avenue to explore the dynamics of a variety of financial applications. The backpropagation algorithm [16] has emerged as one of the most widely used learning procedures for multilayer networks. They have been shown to have great potential for financial forecasting [1]. Neural networks can make contributions to the maximization of returns, while reducing costs, and limiting risks. They can simulate fundamental and technical analysis methods using fundamental and technical indicators as inputs. Consumer price index, foreign reserve, GDP, export and import volume, etc. could be used as inputs. For technical methods, the delayed time series data, moving average, relative strength index, etc. could be used as inputs.

Levich et al. [6] stated that active currency risk management based on a simple application of technical trading signal (filter rule and moving average crossover rule) can lead to significant profits in the foreign exchange future market. Currencies have a reputation as “momentum trading” vehicles in which technical analysis has more validity than usual [12]. Thus technical methods are studied in this paper.

The work discussed in this paper would represent a violation of the efficient market hypothesis postulated by the first school. The inclusion of fundamental factors proposed by the second school is under study as part of our continuing work.

3. Rescaled range analysis and market efficiency testing

The most well known, widely tested but little believed hypotheses to support the first school of thought mentioned in Section 2 are the Random Walk Hypothesis and Efficient Market Hypothesis. The Random Walk hypothesis states that the market prices wander in a purely random and unpredictable way. The Efficient Market Hypothesis states that the markets fully reflect all of the available information and prices are adjusted fully and immediately once new information become available. In the actual market, some people do react to information immediately after they have received the information while other people wait for the confirmation of information. The waiting people do not react until a trend is clearly established.

The Hurst exponent [5] is a measure of the bias in fractional Brownian motion.¹ The method could be used in economic and financial market time series to see whether these series are also biased random walks which indicates the possibility of forecasting.

The rescaled range analysis (or R/S analysis hereafter) [5,12] is able to distinguish a random series from a fractal series, irrespective of the distribution of the underlying series (Gaussian or non-Gaussian). It is a robust statistics for measuring the amount of noise in a system. It can be used to determine the persistence of trends and the average

¹ The erratic path followed by a particle suspended in a fluid. The distance of a random particle covered increases with the square root of time used.

length of non-periodic cycles. In this paper, it is used to detect the long-memory effect in the foreign exchange rate time series over a time period. R captures the maximum and minimum cumulative deviations of the observations x_t of the time series from its mean (μ), and it is a function of time (the number of observations is N):

$$R_N = \max_{1 \leq t \leq N} [x_{t,N}] - \min_{1 \leq t \leq N} [x_{t,N}], \quad (1)$$

where $x_{t,N}$ is the cumulative deviation over N periods and defined as follows:

$$x_{t,N} = \sum_{u=1}^t (x_u - \mu_N), \quad (2)$$

where $\mu_N = \text{Average of } x_u \text{ over } N \text{ periods}$.

The R/S ratio of R and the standard deviation S of the original time series can be estimated by the following empirical law: $R/S = N^H$ when observed for various N values. For some value of N , the H exponent can be calculated by

$$H = \log(R/S)/\log(N), \quad 0 < H < 1 \quad (3)$$

and the estimate of H can be found by calculating the slope of the log/log graph of R/S against N using regression. The R/S used in the regression between $\log(R/S)$ and $\log(N)$ for various N , is the average R/S value for each N . By regressing over the range of different values of N , the highest value obtained indicates the average length of the non-periodic cycle (mean orbital period) of the observed system.

The Hurst exponent H describes the probability that two consecutive events are likely to occur. There are three distinct classifications for the Hurst exponent: (1) $H = 0.50$, (2) $0 \leq H < 0.50$, and (3) $0.50 < H < 1.00$. The type of series described by $H = 0.5$ is random, consisting of uncorrelated events. However, value $H = 0.50$ cannot be used as an evidence of a Gaussian random walk, it only proves that there is no evidence of a long memory effect. A value of H different from 0.50 denotes that the observations are not independent. When $0 \leq H < 0.5$, the system is an antipersistent or ergodic series with frequent reversals and high volatility. Despite the prevalence of the mean reversal concept in economic and financial literature, few antipersistent series have yet been found. For the third case ($0.5 < H \leq 1.0$), H describes a persistent or trend-reinforcing series which is characterized by long-memory effects. The strength of the bias depends on how far H is above 0.50. However, even in the case that the Hurst process describes a biased random walk, the bias can change abruptly either in direction or magnitude. Therefore, only the average cycle length of observed data can be estimated. In this range, the lower the value of H , the more noise there is in the system and the more random-like the series is. The higher the H value, the more efficient the market is since it has less noise than the others.

The datasets of five currencies studied in this paper comprise 2910 daily rates for a sampling period between 18 May 1984 and 7 July 1995. In a real situation, there is no closing price for Forex. Forex trading takes place 24 h a day. The 24 h data should be used in order to capture the underlying rules of the movement in Forex rates. The more data one uses the more rules one may get. However, in view of the difficulties in

Table 1

Weekly foreign exchange rates statistics: number of observations: 510; period: May 1984–July 1995

Exchange	Mean	Standard deviation	Variance	Max	Min
AUD	0.7424	0.0605	0.003642	0.9035	0.5959
CHF	1.6203	0.3828	0.1466	2.9245	1.1145
DEM	1.9450	0.4770	0.2263	3.4520	1.3529
GBP	1.6025	0.1978	0.03938	2.0040	1.0520
JPY	152.240	42.2177	1782.4335	262.80	80.83

Table 2

Hurst exponent and correlation for the experimented five currencies

Exchange	Hurst exponent	Correlation
AUD	0.532681	0.046347
CHF	0.553941	0.077645
DEM	0.554672	0.078737
GBP	0.544408	0.063497
JPY	0.540706	0.058053

managing the huge data, fixed sampled time series are often used to model the behavior of exchange rate changes. Daily data are easily available compared with tick-by-tick data. In addition, trading too frequently will cost much on transaction. Weekly forecasts over 7 daily data were adopted to reduce the amount of transaction. In this research, the weekly data are used assuming that they have enough information to capture the “rules”. The weekly closing prices are used as the prediction target of our experiment. They refer to Friday’s closing prices in the Singapore forex market which ranks the fourth largest in the world. In the event of Friday being a holiday, the most recently available closing price for the currency was used. The dataset for each currency in this study consists of 510 weekly data.

The statistics summary of the weekly data used in this study are shown in Table 1. As shown in Table 2, the value of Hurst exponent for the logarithmic returns of daily exchange rates data is higher than 0.5 for all the observed time series. This shows that all the markets are fractal, and not random walk. The highest value is 0.554 for the exchange rate of DEM/USD. In other words, it is the most efficient market amongst the five studied markets. Peters [12] found that US stock market is the most efficient market, followed by Germany, UK and Japan. The foreign exchange rates markets efficiency happens to be in the same sequence, i.e., Mark, Pound and Yen, with Peters’ finding in stock markets. The results obtained in this study imply that all the five markets are not random walk and are not highly efficient thus forecasting of the exchange rates is possible.

4. Building a neural network forecasting model

To build a neural network forecasting model, historical data are divided into three portions: training, validation and testing sets. The training set contains 70% of the collected data, while the validation and the testing sets contain 20% and 10%, respectively. The division is based on a rule of thumb derived from the experience of the authors. A model is considered good if the error of out-of-sample testing is the lowest compared with the other models. If the trained model is the best one for validation and also the best one for testing, one can assume that it is a good model for future forecasting. The data are chosen and segregated in time order. In other words, the data of the earlier period are used for training, the data of the later period are used for validation, and the data of the latest time period are used for testing. The objective is to discover the underlying “structure” of the mechanism generating the data, i.e., to discover the relationship between present, past and future observations.

4.1. Measurement of neural networks

A usual measure to evaluate and compare the predictive power of the model is the normalized mean squared error (NMSE) which was used to evaluate entries into the Santa Fe Time Series Competition [19]. Given a set P comprising pairs of the actual values (or targets, x_k) and predicted values (\hat{x}_k), the NMSE which normalizes the MSE by dividing it through the variance of respective series can be defined as

$$NMSE = \frac{\sum_{k \in P} (x_k - \hat{x}_k)^2}{\sum_{k \in P} (x_k - \bar{x}_k)^2} = \frac{1}{\sigma_P^2} \frac{1}{N} \sum_{k \in P} (x_k - \hat{x}_k)^2, \tag{4}$$

where σ^2 is the estimated variance of the data, \bar{x}_k being the mean, and N is the size of a set P . If the estimated mean of the data is used as predictor, $NMSE = 1.0$ is obtained. The normalized mean squared error is related to R^2 which measures the (linear) dependence between pairs of desired values and predictions by $NMSE = 1 - R^2$. Additional evaluation measures include the calculation of correct matching number of the actual and predicted values, x_t and \hat{x}_t , respectively, in the testing set with respect to the sign and directional change (expressed in percentages). Sign statistics can be expressed as

$$S_{\text{stat}} = \frac{1}{N} \sum_{k \in P} a_k, \tag{5}$$

where $a_k = 1$ if $x_t \hat{x}_t > 0$ or $x_t = \hat{x}_t = 0$, and $a_k = 0$ otherwise. Similarly, directional change statistics can be expressed as

$$D_{\text{stat}} = \frac{1}{N} \sum_{k \in P} b_k, \tag{6}$$

where $b_k = 1$ if $(x_{t+1} - x_t)(\hat{x}_{t+1} - \hat{x}_t) \geq 0$, and $b_k = 0$ otherwise.

These statistics are desirable because the normalized mean square errors measure prediction only in terms of levels. Hence, the quality of the forecast can be measured by R^2 , the correctness of gradient predictions (D_{stat}) and the sign changes.

Table 3

The out of sample forecasting results for neural network models (delay method) for weekly foreign exchange data^a

Exchange	Model	Test. NMSE (R^2)	Gradient (%)	US Ret 1 (%)
AUD	5-3-1	0.0543 (0.9456)	55.00	1.09
CHF	5-3-1	0.1100 (0.8900)	56.00	8.40
DEM	6-3-1	0.3153 (0.6847)	51.00	4.36
GBP	6-3-1	0.1555 (0.8445)	54.74	2.30
JPY	5-3-1	0.1146 (0.8853)	53.40	3.00

^aTraining 1984–1993, testing Nov 1993–July 1995 US Return1 refers to the return achieved by USD for trading Strategy 1.

Table 4

The testing results for neural network models using indicators for weekly foreign exchange data^a

Exchange	Model	Test. NME	Gradient (%)	$Ret1_{US}$ (%)	$Ret1$ (%)	$Ret2_{US}$ (%)	$Ret2$ (%)
AUD	5-3-1	0.035105	73.86	8.82	12.19	12.43	15.90
AUD	6-4-1	0.032362	76.14	8.97	12.34	12.67	16.16
CHF	5-3-1	0.068819	65.91	28.49	9.99	22.49	4.85
CHF	6-4-1	0.065962	64.77	32.36	13.31	21.64	4.15
DEM	5-3-1	0.063462	61.36	22.86	8.86	15.20	2.07
DEM	6-4-1	0.061730	64.77	27.84	13.27	18.00	4.55
GBP	5-3-1	0.061370	73.86	7.22	2.87	14.78	10.13
GBP	6-4-1	0.053650	72.73	10.62	6.13	16.48	11.76
JPY	5-4-1	1.966195	46.59	19.71	3.47	0.00	– 13.57
JPY	6-4-1	1.242099	46.59	23.42	6.67	0.00	– 13.57
SGD	5-3-1	0.268410	39.77	7.70	2.42	5.95	0.75
SGD	6-4-1	0.376993	38.64	7.00	2.26	5.28	0.62

^aTraining 18 May 1984–12 July 1991, Validation 19 Nov 1991–29 Oct 1993, Testing 5 Nov 1993–7 July 1995. $Ret1$: Return using the difference between the predictions this and next period as a criterion for trading; $Ret2$: Return using difference between the prediction next period and the actual level this period; $Ret1_{US}$ denotes the starting currency which is USD.

The real aim of forecasting is the trading profits or financial gains based on prediction results. It does not matter whether the forecasts are accurate or not in terms of NMSE or gradient. After the forecast results were obtained from the neural network models, a program simulating the real trading was developed to test the possible monetary gains. Since this is not the real trading, we name it as paper profits.

The paper profits are calculated by the return one can expect if one starts with either the USD or the currency under consideration. Assume that a certain amount of seed money is used in this program. The seed money is used to buy a certain amount of another currency when the prediction shows a rise in that currency. At the end of the testing period, the currency should be converted to the original currency of the seed money using the exact direct or cross rate of that day. The results obtained are

shown in Tables 3 and 4. The paper profit is calculated as follows:

$$\text{Return} = \left(\frac{\text{MoneyObatined}}{\text{SeedMoney}} \right)^{52/w} - 1, \quad (7)$$

where *MoneyObatined* is the amount of the money obtained on the last day of testing; *SeedMoney* is amount of money used for trading on the the first day of testing; and *w* the number of weeks in testing period.

4.2. Trading strategies

Trading is an art. As there is no perfect forecasting technique, trading profit is ensured only by a good trading strategy taking “full” advantage of a good forecasting method. There are two kinds of trading strategies used in this study. One uses the difference between predictions, and another uses the difference between the predicted and the actual levels to trade.

Strategy 1:

$$\text{if } (\hat{x}_{t+1} - \hat{x}_t) > 0 \quad \text{then buy else sell.} \quad (8)$$

Strategy 2:

$$\text{if } (\hat{x}_{t+1} - x_t) > 0 \quad \text{then buy else sell.} \quad (9)$$

Here x_t is the actual level at time t , where \hat{x}_t is the prediction of the neural networks.

The reason for the different results of the different strategies lies in the noise. Assume that the noise of the prediction at time t is δ_t , then the following equation holds:

$$x_t = \hat{x}_t + \delta_t. \quad (10)$$

Which strategy is more accurate depends on $|\delta_{t+1} - \delta_t|$ and $|\delta_{t+1}|$. If δ_{t+1} and δ_t have different signs, then Strategy 1 is better than Strategy 2. Otherwise Strategy 2 is better than Strategy 1.

Transaction cost will be considered in real trading. In this paper, 1% of transaction cost was included in the calculation. The transaction cost of a big fund trading, and thus affecting the market rates was not taken into consideration. To be more realistic, a specific amount of transaction cost has to be included in the calculation.

Using one currency as an example, say JPY, if the forecast shows that JPY is going up and the currency at hand is Dollar, then the strategy is to buy Yen. If the forecasts show that Yen is going down and the currency at hand is Yen, then the strategy is to sell Yen. Otherwise, no action is taken.

What are the criteria for determining whether the currency is going up or down? If the output of neural network is given in percentage of changes, we can use positive or negative output to show that the currency is going up or down. The returns of the two different strategies based on the best results so far are as follows: The GBP trades that depend on the difference between the prediction of this period and the prediction of the next period (Strategy 1) achieved 14.78% in annual return. The GBP trades that depend on the difference between the prediction next period of the neural network and the actual level of this time period (Strategy 2) received 7.22% in annual return. In

actual trading, practitioners may choose one of the strategies. A conservative trading strategy would require a trader to act only when both strategies recommend the same actions.

4.3. Neural network forecasting model

Basically, three experiments are conducted in this research. First, a purely time-delayed model is built to capture the relationship between next week's exchange rate and the historical exchange rates. The purely time-delayed forecast method is one of the simplest technical analysis methods. The normalized exchange rates of the previous periods are fed to a neural network to forecast the next period exchange rate in this model. For example, the inputs to a 5-x-1 neural network are FX_{i-4} , FX_{i-3} , FX_{i-2} , FX_{i-1} and FX_i while the output of the neural network is FX_{i+1} , the next week's exchange rate where FX_i stands for the current week's exchange rate. The architecture of the neural network is denoted by $i-h-o$. The $i-h-o$ stands for a neural network with i neurons in input layer, h neurons in hidden layer, and o neurons in output layer. In our experiments, up to eight weeks of time-delayed data are used. We find that there is no significant improvement in terms of NMSE for the networks with more than six inputs. Sometimes the results are even worse. This shows that increasing the number of inputs does not necessarily increase the accuracy of predictions. One reason may be that the information provided by others factors are already included in the essential factors. In the case of worse performance, it may be due to that introducing of extra new delayed inputs brings in nodes that learnt on specific noises. In the interest of space, only the best of the models obtained in the experiment are presented in this paper.

The purely time-delayed method for prediction sometimes leads to prediction that seems to generate a time-delayed time series of the original time series. With the inclusion of some popular indicators used by traders, it might help to remove some of the time-delay characteristics of the prediction.

The father of Dow-Jones, Charles Dow, categorized the trend into three different levels, namely the primary or major trend, the secondary or intermediate trend, and the minor trend [17]. He likened the three levels of trend to the tides, waves and ripples of the sea. The advantage of moving average is that it tends to smooth out some of their irregularities that exist between market days. One disadvantage of this method is that the calculation always lags behind the current market. The primary trend lasts more than one year or longer. The secondary trend usually lasts from three weeks to three months. The minor trend is meaningless in itself. However, a series of three or more minor trends can make up the primary trend or the corrective secondary trend. Following Dow-Jones theory, moving averages are used as inputs to the neural network in the second experiment. MA5, MA10, MA20, MA60, and MA120 are fed to the neural networks to predict the following week's rate. They refer to moving averages for one week, two weeks, one month, one quarter and half a year, respectively. MAs are calculated on the trading days. MA5 stands for five trading days' moving average. The filtered figures may provide more information to the model than the purely delayed data. Further, the forecasts for each of the currencies were

repeated, but with a hybrid of indicators and one time-delay term. The results are presented in Table 4 where the configuration 5-x-1 stands for indicator model while 6-x-1 stands for the hybrid model.

Finally, a more in depth research on one of the currencies, namely CHF, is conducted as the third experiment. Instead of finding just one good solution, data are partitioned equally using bootstrapping rules to find models in different time periods. Specifically, 12 data sets are generated and each contains six years training and validation data while half a year data are used for out of sample testing.

5. Prediction results and discussion

5.1. Weekly exchange rate forecast using purely time delayed time series

The measurements of the forecasting results for purely time-delayed method are shown in Table 3. Figs. 1(a)–(e) are the diagrams showing the predicted and the actual time series of the exchange rates of AUD/USD, CHF/USD, DEM/USD, GBP/USD and JPY/USD for the period of November 1993 to July 1995 (out of sample).

From the table and diagrams, we find that the fitness between the two curves and the NMSE level are quite good. However, the gradients are only a little above 50% which means the forecasts are only slightly better than the chance in tossing a coin. No doubt such results would not be accepted by the practitioners. Thus the second approach is necessary.

5.2. Weekly exchange rate forecast using indicators

The measurements of the forecasting results are shown in Table 4. An additional term of time delay is introduced as input to the 6-x-1 model. Our results however show that this does not contribute significantly to the improvement of hit rate, or D_{stat} , as shown in Table 4. For example as shown in Figs. 2(a) and (b), for Australian Dollar, the results show that the hit rate of 6-4-1 is slightly higher than that of the pure indicator forecasting method with a configuration of 5-3-1. The hit rate of the former is 76.14% and that of the latter is 73.86%. Hence, the diagrams with additional time delay term for the other currencies are not shown in this paper.

The comparative diagrams showing the predicted (NN output) and the actual (XXX target) time series for the period of Nov 1993–July 1995 (out of sample) are shown in Figs. 2(a)–(f) for AUD/USD, AUD/USD (with an additional time-delay input), CHF/USD, DEM/USD, GBP/USD, and JPY/USD.² From the results presented

² The big lap between the actual level and predicted level of JPY/USD is due to the normalization method used in this study and the market changes. The normalized data are calculated using the actual data and the maximum and minimum of training data. The minimum value of JPY/USD before November of 1993 is just below 110, but most of levels after April 1995 are lower than 85. The result of this research shows that the neural networks have only learned the rule of trend and not the level of the exchange rate. So the neural networks can only get better paper profit using Strategy 1. And the graph for JPY/USD is very bad if we only consider the levels.

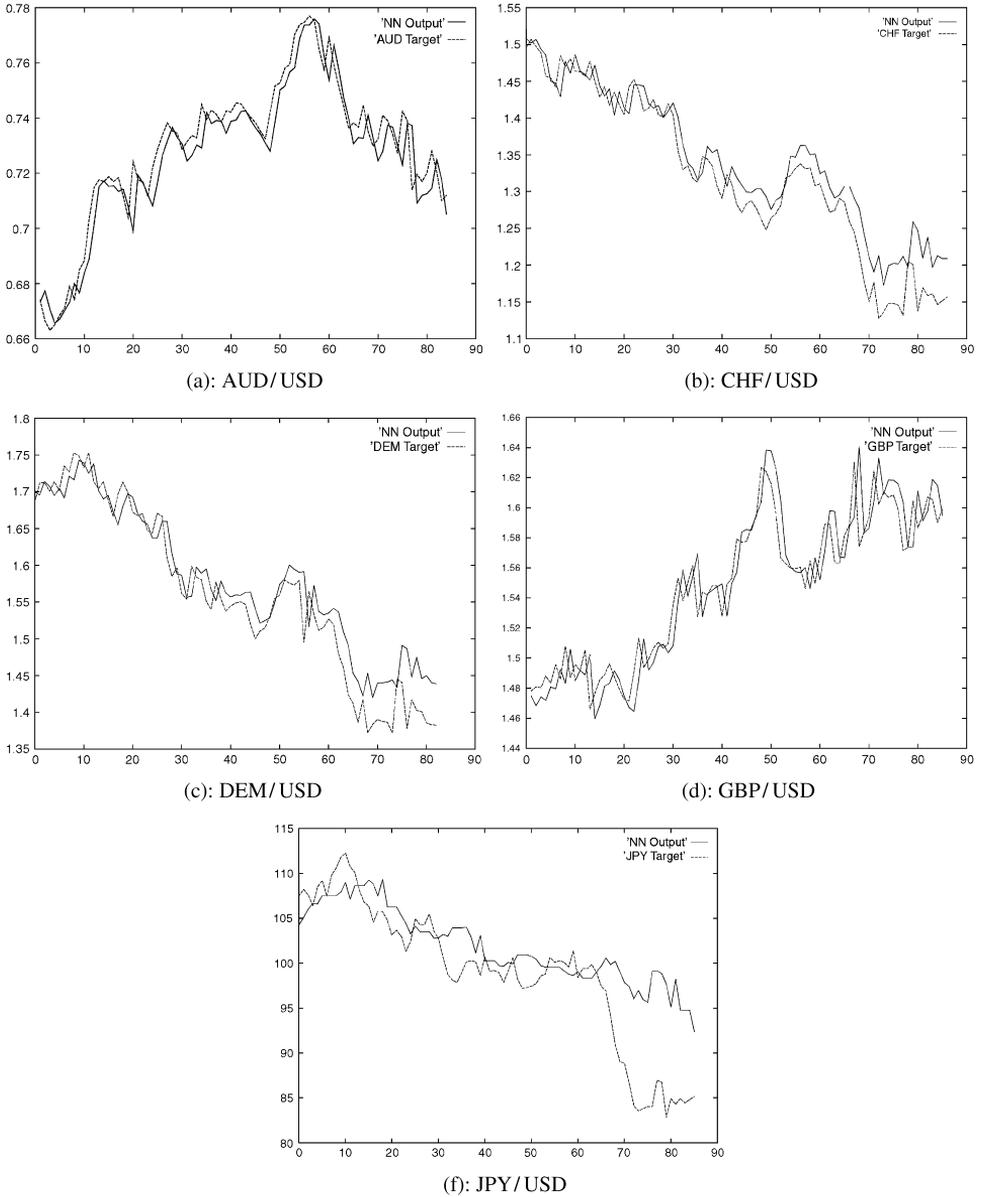


Fig. 1. Purely time-delay model prediction on weekly exchange rates, Nov 1993–July 1995.

here, we can safely conclude that hit rates of approximately 70% can be achieved consistently for AUD, GBP, and somewhat lower for CHF and DEM.

Also, from the graphs, one can conclude that the forecasts for the first 20 weeks of the pure testing period look “impressive”. The performance on the five currencies

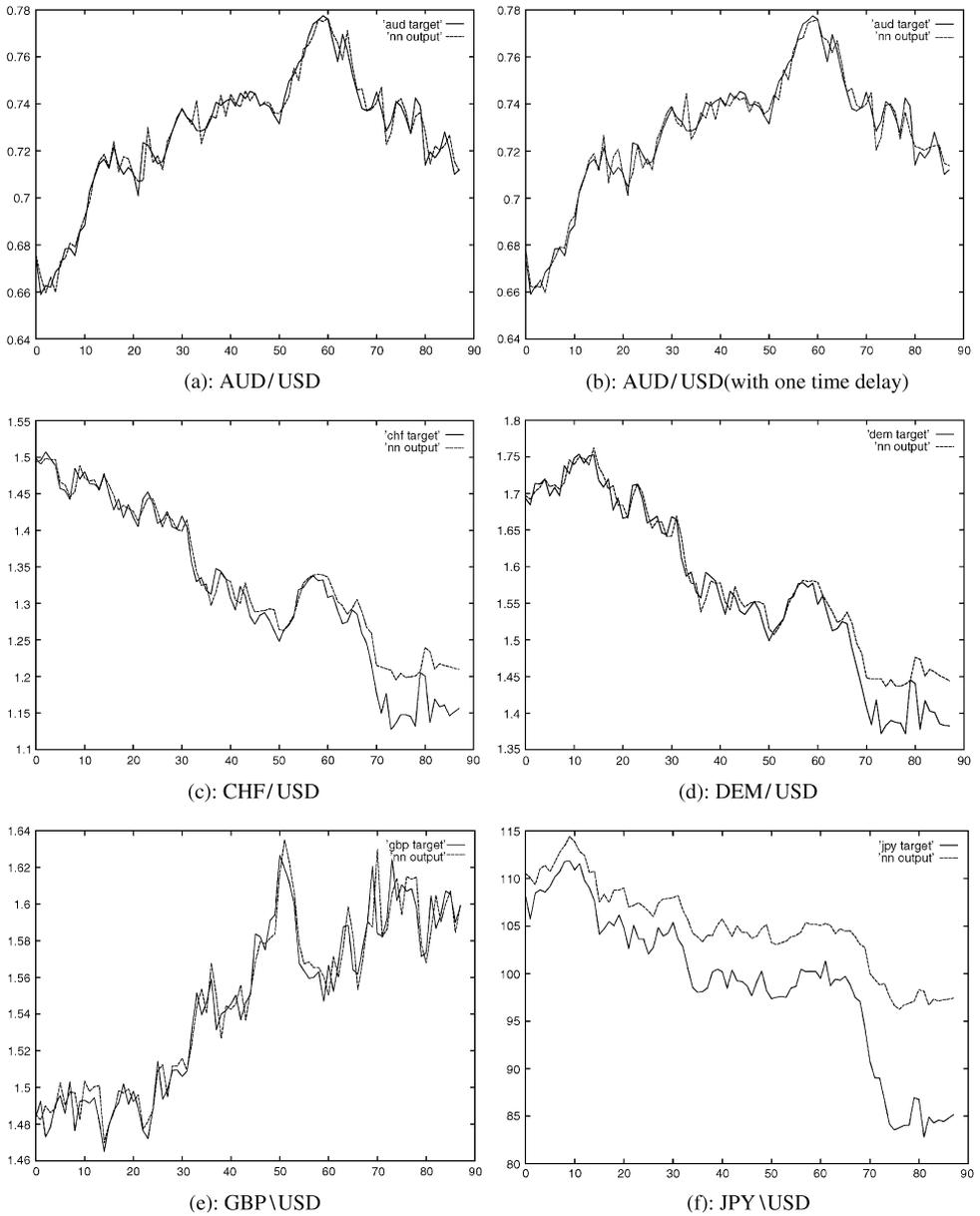


Fig. 2. Indicator model prediction on weekly exchange rates, Nov 1993–July 1995.

shows a degradation in forecasting in about the third quarter after training. This means that, the neural network needs to be retrained, probably every 20 weeks (or half a year) with the latest data to increase the chance of achieving a better forecast. This indicates the presence of ‘recency’ problem of the network, namely, the network did

retain some memory of the history. A half a year forecast period is thus recommended based on the present study. Notice that, in actual application, only validation data sets would be required together with the training data sets. So, we manage to experiment to retrain the neural network model every half a year. In Section 7, the results for Swiss Franc will be presented.

6. Comparison with ARIMA

To compare the neural network forecasting results with other models. The *autoregressive integrated moving average* (ARIMA) model is used as the first benchmark in this research. ARIMA Model was introduced by George Box and Gwilym Jenkins [2] in 1976. The Box–Jenkins methodology provided a systematic procedure for the analysis of time series that was sufficiently general to handle virtually all empirically observed time series data patterns. ARIMA(p, d, q) is the general form of different ARIMA models. Here p stands for the order of the autoregressive process, d represents the degree of differencing involved, and q is the order of the moving average process. To compare the forecasting results of the neural networks, a number of ARIMA models were built. Table 5 is the results of ARIMA models with different trading strategies.

The entire data set was used as fitting data for the ARIMA models. In other words, the data forecast by ARIMA were already used in the fitting stage of ARIMA model building. The results obtained from the ARIMA models should be compared with *Ret1* in Table 4. In other words, the ARIMA models should deliver worse out-of-sample forecasting returns than the ARIMA results indicated in Table 5.

Focusing on the gradients, the ARIMA method can achieve about 50% of correctness while up to 73% of correctness can be achieved using neural network models.

Table 5

The result of using ARIMA (AUD(101) stands for the ARIMA result of AUD using ARIMA(1, 0, 1) model and the same rules are applied to other currencies)

ARIMA model	Gradient (%)	<i>Ret</i> _{1US} (%)	<i>Ret</i> _{2US} (%)
AUD(101)	52.27	1.43	1.36
AUD(202)	54.32	1.53	1.21
CHF(101)	38.64	6.94	– 1.42
CHF(202)	55.86	5.43	0.64
DEM(101)	43.18	3.48	3.49
DEM(202)	44.62	3.48	3.22
GBP(101)	53.41	2.24	3.67
GBP(202)	51.77	2.63	1.32
JPY(101)	44.32	– 1.47	– 0.52
JPY(202)	44.32	– 0.78	0.02

From practitioners' point of view, returns are more important than gradients. With reference to Tables 5 and 4, the differences between ARIMA models and neural network models are significant. The best return for ARIMA models using Strategy 1 regardless of the devaluation and strategies is 6.94%, while for neural network models is 32.36%.

The results show that irrespective of NMSE, gradient or profit, the neural network models are much better than the traditional ARIMA model. In the case of forecasting five major currencies exchange rates, neural network model can work as a viable alternative forecasting tool.

7. Further research on CHF/USA rates forecasting

In this section, we will investigate the consistency of the neural network models. As found in the previous diagrams, the good results can last about half a year. We partition the whole data into 12 segments (namely CHF1 to CHF12) each covering an overlapping period of $6\frac{1}{2}$ years. The twelve $6\frac{1}{2}$ year periods are progressively displaced by $\frac{1}{2}$ year. Thus, the first $6\frac{1}{2}$ year period (for CHF1) spans from January 1984 to June 1990. The next period is slid along the time horizon for half a year, i.e., it spans from July 1984 to December 1990. The last time period spans from July 1989 to November 1995.

Amongst the $6\frac{1}{2}$ year data, the first 6 years' data, or 312 weeks, were used for training and validation while the remaining half a year's data, or 26 weeks, were used to test the performance of the neural network model. Of the 312 data, 260 were used as training data and 52 are used as validation. For each data segment, a variety of network architectures were experimented. The best architecture, in term of NMSE, for each data segment, is presented in the Table 6.

Table 6

The technical details of chosen architectures and their forecasting potentials^a

Name	Archit.	α	η	NMSE	Grad1 (%)	Grad2 (%)
CHF1	6-2-1	0.9	0.0030	0.170480	64.00	84
CHF2	6-2-1	0.9	0.0050	0.030029	61.54	72
CHF3	6-2-1	0.9	0.0030	0.120893	53.85	76
CHF4	6-2-1	0.9	0.0006	0.123096	61.42	36
CHF5	6-3-1	0.9	0.0006	0.042554	65.38	84
CHF6	6-3-1	0.9	0.0007	0.144654	42.31	76
CHF7	6-4-1	0.9	0.0010	0.662526	53.85	68
CHF8	6-3-1	0.9	0.0008	0.051245	65.38	84
CHF9	6-2-1	0.9	0.0009	0.207124	53.85	72
CHF10	6-2-1	0.9	0.0200	0.106726	69.23	92
CHF11	6-3-1	0.9	0.0003	0.040063	61.53	68
CHF12	6-2-1	0.9	0.0008	0.145288	66.67	82

^a α : learning rate; η : momentum rate; NMSE: normalized mean-squared error; Grad1, Grad2: correctness of gradients for testing data.

Table 7
Benchmark results for different time period

Name	<i>Ret1</i> (%)	<i>Ret1</i> _{USD} (%)	<i>Ret2</i> (%)	<i>Ret2</i> _{USD} (%)	Bench. I (%)	Bench. II (%)
CHF1	18.59	11.20	35.93	27.45	– 6.24	12.59
CHF2	9.31	19.47	12.58	23.05	– 3.00	12.21
CHF3	8.72	29.63	12.49	34.12	19.05	1.27
CHF4	– 7.97	10.97	– 4.96	14.60	32.70	– 9.09
CHF5	47.72	– 0.58	58.99	7.00	25.90	27.99
CHF6	– 10.54	21.62	8.92	48.06	35.94	– 2.45
CHF7	20.34	19.30	26.45	25.35	– 0.87	– 9.69
CHF8	31.85	12.14	67.74	42.66	– 14.95	16.76
CHF9	– 9.91	– 11.57	23.31	15.46	– 6.36	– 13.66
CHF10	20.24	27.80	30.01	38.17	6.28	– 2.11
CHF11	5.23	25.97	19.97	32.84	19.70	– 8.38
CHF12	1.80	13.58	20.71	34.68	11.58	– 3.69

Table 8
Analysis of results for different models^a

Name	<i>Ret1</i>	<i>Ret1</i> _{USD}	<i>Ret2</i>	<i>Ret2</i> _{USD}	Bench. I	Bench. II
Median	9.31	19.30	20.71	27.45	11.58	– 2.11
Average	11.28	14.96	26.01	28.62	9.98	1.81
Portfolio	12.43	11.00	27.03	29.45	9.88	1.88
Acceptability	69.23	84.62	84.62	100	61.54	30.77

^aMedian: median of 12 models; Average: average profit from 12 models; Portfolio: Profit of the sum of 20% of Max, 20% of Min and 60% average of others; Acceptability: Percentage of profit greater than 5% annually.

Instead of ARIMA model, two strategies serving as new benchmarks are introduced. Benchmark I uses a “buy-and-hold” strategy while Benchmark II uses a “trend-follow” strategy. The Benchmark I’s strategy is to buy the USD at the beginning of the testing period and then sell it at the end of the testing period. The Benchmark II’s strategy is to buy when the market is continually up for two weeks and sell when it is down for one week. Table 7 shows the differences between paper profits and their benchmark profits. All trade on Fridays.

The study of the 12 data segments shows that the neural network models could be applied to future forecasting. Compared with the two benchmarks, the neural network model is better. As shown in Table 8, for the neural network, the worst percentage of acceptable profit, assuming at least 5%, is 69.23 among the four strategies, while the benchmarks can only achieve 61.54 and 30.77 respectively. Referring to Table 8 again, another comparison was made for the forecasting results. Assume that we have a portfolio profit based on 20% of maximum of profit, 20% of minimum profit, and

60% of average profits from other data segments. This will indicate the general performance of each strategy.

Figs. 3(a) and (b) are the graphical presentation of the profit gained for each strategy.

8. Conclusion and further research

The present study has given rise to the following observations:

- (1) The forecasting results are very promising for most currencies except Yen. In using neural networks to perform technical forecasting, i.e., feeding simple technical indicators as inputs to neural network, better results are obtained for Australian Dollar, Swiss Franc, and British Pound and perhaps Deutsch Mark. The results for Japanese Yen are the worst especially when Strategy 2 is used. This is similar to our forecasting results using ARIMA models. There may be some factors related to Yen that need to be studied. One reason could be that the market for Yen is bigger and more efficient than the markets for Australian Dollar, Swiss Franc, and British Pound. The traders of the Yen market may depend more on technical analysis and they may act quickly after any sign appears. Hence, technical analysis may not be a good tool for forecasting the trends of Yen as everybody is well aware of the meaning of technical signals.
- (2) To take more advantage from neural network forecasting results, trading strategies need to be considered in real trading. Strategy 2 has been found more powerful in using the neural network forecasting results, the theoretical study of it is under exploration. Between the two simple trading techniques, the “buy-and-hold” strategy (Benchmark I) is better than the “trend-follow” (Benchmark II) though they are generally worse than neural network models as shown in Fig. 3 and Table 7.
- (3) There is a need to have an effective measure of neural network performance rather than just the goodness of fit such as NMSE. A very small NMSE is not necessarily a general indicator of good performance. The sum of the NMSE of the three parts of data (training, validation and testing) must be kept small, not just the training NMSE alone. Sometimes having small NMSEs for testing and validation is more important than having small NMSE for training. Instead of just using NMSE alone, three levels of judgment, NMSE, gradient and profits are proposed and studied in this research. The hit rate may be a better standard for determining the quality of the forecast. After all, the return depends a lot on the trading strategies and how the forecasting information are being used for trading advantages. However, the level of hit rates and paper profits also depend on the period of forecast. Hence, periodic retraining of the neural networks is necessary. For the practitioners, the levels of exchange rate and trend can be used depending on their expectation of return and risk. In addition to the above-mentioned two strategies, after different trading strategies may be used by them.

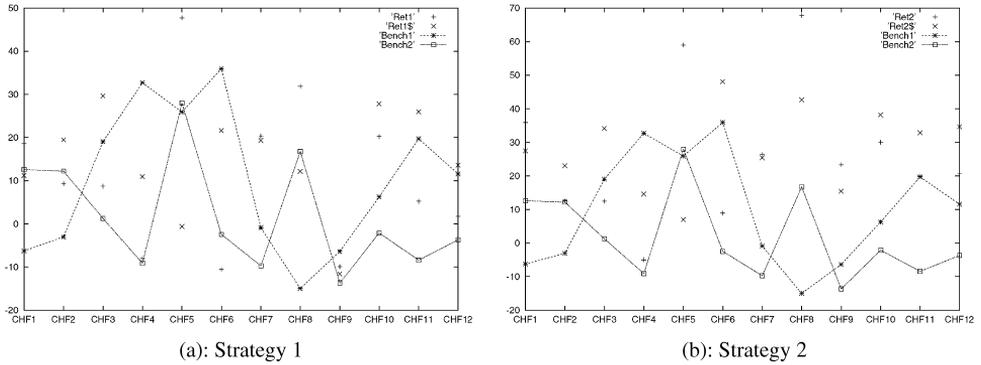


Fig. 3. Each models profits compared with benchmarks. The vertical axis is the annualized profit in percentage, the horizontal axis is the names of different time periods.

- (4) A backpropagation network used in the present study has proved to be adequate for forecasting. Simple technical indicators such as moving average are enough.

The above observations have confirmed the applicability of neural network for financial forecasting albeit with some problems that need to be addressed. We are seeking an automatic facility of model building to minimize the tedious trial-and-error procedures. The above problems have enlightened us in our effort as our future work in formulating a generic neural network model for forecasting as follows:

- The model should be self-adapting to different situations. Parameters for such adaption will need to be studied and incorporated into one model. Examples of parameters include volatility of current measurement, data size and frequency, etc.
- To provide more statistical analyses in order to find out which kind of data segments will best capture the underlying behavior of the market changes.
- The model should minimize the recency problem by self-detecting the possibility of such a problem and provide its own measure to avoid the problem.
- The model should provide the frequency support. Sampling of data according to volatility need to be applied to the raw data. As Zhou [24] stated, increasing observation frequency does not always help to improve the accuracy of forecasting. In this research, the weekly data are used assuming that they have enough information to capture the “rules”. Due to the volatility of the currency movement [8], a different frequency of data maybe needed than the weekly data. The data could be sampled according to the market character, e.g. bullish, bearish, or trading, etc. In other words, when the market is volatile, we sample more data for training, and vice versa. Non-linear or volatile time scale will be taken into consideration in our further research.
- The model should aim at the financial profit in addition to the usual fitness of forecast and target series. Due consideration should be given to emphasize financial profit, such as, by incorporating profit into neural network training. The forecast-

ing targets such as actual rate, percentage of changes, etc., may be decided automatically by the model.

- The model should recommend the trading rules based on the expectation of returns and risk. More trading strategies are needed in addition the two studied here. To benefit more practitioners, risk-based trading strategies will be taken into consideration in the model.

Acknowledgements

The authors are grateful to the anonymous referees for their helpful comments. We thank Dr. Hean-Lee Poh, Yili Li and Teo Jasic for their friendly cooperation and invaluable discussions.

References

- [1] M. Adya, F. Collopy, How effective are neural networks at forecasting and prediction? A review and evaluation, *J. Forecasting* 17 (1998) 481–495.
- [2] G.E.P. Box, G.M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, CA, 1976.
- [3] H. Green, M. Pearson, Neural nets for foreign exchange trading, in: *Trading on the Edge: Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets*, Wiley, New York, 1994.
- [4] A.C. Harvey, *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge, 1989.
- [5] H.E. Hurst, Long term storage capacity of reservoirs, *Trans. Am. Soc. Civil Eng.* 116 (1951) 770–808.
- [6] R.M. Levich, L.R. Thomas, The merits of active currency risk management: evidence from international bond portfolios, *Financial Anal. J.* 49 (5) (1993) 63–70.
- [7] R. Manger, Using holographic neural networks for currency exchange rates prediction, *Proceedings of the 16th International Conference on Information Technology Interface*, Pula, Croatia, June 1994.
- [8] U. Muller, M. Dacorogna, R. Dave, G. Pictet, R. Olsen, R. Ward, Fractals and intrinsic time – a challenge to econometricians, *Proceedings of the 39th International Conference of the Applied Econometrics Association on Real Time Econometrics*, Luxembourg, October 1993.
- [9] A.M. Rawani, Forecasting and trading strategy for foreign exchange market, *Inform. Decision Technol.* 19 (1) (1993).
- [10] A.N. Refenes, M. Azema-Barac, L. Chen, S. Karoussos, Currency exchange rate prediction and neural network design strategies, *Neural Comput. Appl.* 1 (1) (1993) 46–58.
- [11] A.N. Refenes, A. Zapranis, G. Francis, Stock performance modeling using neural networks: a comparative study with regression models, *Neural Network*, 5 (1994) 961–970.
- [12] E.E. Peters, *Chaos and Order in the Capital markets: A New View of Cycles, Prices, and Market Volatility*, Wiley, New York, 1991.
- [13] H. Pi, Dependence analysis and neural network modeling of currency exchange rates, *Proceedings of First International Workshop on Neural Networks in Capital Markets*, London, 1993.
- [14] T. Plummer, *Forecasting Financial Markets: A Technical Analysis and the Dynamic of Price*, Wiley, New York, 1991.
- [15] A. Poddig, Short term forecasting of the USD/DM exchange rate, *Proceedings of First International Workshop on Neural Networks in Capital Markets*, London, 1993.
- [16] D.E. Rumelhart, J.L. McClelland, *Parallel Distributed Processing: Explorations in the Micro-structure of Cognition*, vol. 1, The MIT Press, Cambridge, MA, 1986, pp. 318–362.

- [17] R. Teweles, E. Bradley, T. Teweles, *The Stock Market*, 6th Edition, Wiley, New York, 1992, pp. 374–379.
- [18] A.S. Weigend, Generalization by weight-elimination applied to currency exchange rate prediction, *Proceedings of the IEEE International Joint Conference on Neural Networks*, Singapore, November 1991.
- [19] A.S. Weigend, N.A. Gershenfeld (Eds.), *Time Series Prediction: Forecasting the Future and Understanding the Past*, Proceedings of the NATO Advanced Research Workshop on Comparative Time Series Analysis, Santa Fe, NM, May 14–17, 1992, Addison-Wesley, Reading, MA, 1994.
- [20] H. White, Learning in artificial neural networks: a statistical perspective, *Neural Comput.* 1 (1989) 425–464.
- [21] B.K. Wong, A bibliography of neural network business applications research: 1988–September 1994, *Expert Systems* 12 (3) (1995) 253–262.
- [22] B.K. Wong, T.A. Bodnovich, Y. Selvi, Neural-network applications in business: a review and analysis of the literature(1988–1995), *Decision Support Syst.* 19 (4) (1997) 301–320.
- [23] J.T. Yao, H.-L. Poh, T. Jasic, Foreign exchange rates forecasting with neural networks, *Proceedings of the International Conference on Neural Information Processing*, Hong Kong, September 1996, pp. 754–759.
- [24] B. Zhou, Estimating the variance parameter from noisy high frequency financial data, MIT Sloan School Working Paper, No. 3739, 1995.
- [25] X.R. Zhang, Non-linear predictive models for intra-day foreign exchange trading, *Int. J. Intelligent Syst. Accounting, Finance Manage.* 3 (4) (1994) 293–302.



J. Yao is a Senior Lecturer in the Department Information Systems, College of Business, Massey University. (J.T.Yao@massey.ac.nz) Jingtao Yao was a Teaching Assistant in the Department of Information Systems, School of Computing, National University of Singapore. His research interests include financial forecasting, neural networks, software engineering, and business process reengineering. His research has been published in international refereed journals and conference proceedings. He obtained a B.Eng. (1983), a M.Sc. (1988) in Computer Science from Xi'an Jiaotong University, and a Ph.D. (1999) in Information Systems from the National University of Singapore.



Chew Lim Tan is an Associate Professor in the Department of Computer Science, School of Computing, National University of Singapore. His research interests are expert systems, neural networks, computer vision and natural language processing. He obtained a B.Sc. (Hons) degree in Physics in 1971 from the University of Singapore, an M.Sc. degree in Radiation Studies in 1973 from the University of Surrey in UK, and a Ph.D. degree in Computer Science in 1986 from the University of Virginia, USA.