VICTORIA UNIVERSITY OF TECHNOLOGY

DEPARTMENT OF COMPUTER AND MATHEMATICAL SCIENCES

Forecasting the Yield and Direction of the Australian 10 Year Commonwealth Treasury Bond using Artificial Neural Networks

Callum Scott

(76 ECON 3)

June, 1996

(AMS : 90A20)

TECHNICAL REPORT

VICTORIA UNIVERSITY OF TECHNOLOGY
P O BOX 14428
MELBOURNE CITY MAIL CENTRE
MELBOURNE, VICTORIA, 8001
AUSTRALIA

TELEPHONE (03) 9688 4492
FACSIMILE (03) 9688 4050

Footscray Campus

FORECASTING THE YIELD AND DIRECTION OF THE AUSTRALIAN 10 YEAR COMMONWEALTH TREASURY BOND USING ARTIFICIAL NEURAL NETWORKS

CALLUM SCOTT

Victoria University of Technology.

This paper is concerned with the application of artificial neural networks (ANN) to the forecasting of the time series generated by the 10 Year Commonwealth Treasury Bond yield. A fundamental ANN model is developed and applied to the time series, and demonstrates that forecasting of both the direction and magnitude of the change in yield are possible.

This ANN, unlike several other systems, relies not only on inputs derived directly from the time series itself, but on other indicators that impact on the behaviour of the Bond. Moving and exponential averages of these indicators are used extensively.

The model is able to predict the direction of the yield correctly 68% of the time, and is more accurate 61% of the time when compared to using the previous value of the yield as a prediction for the next. A naive strategy, that of using the previous change in yield as a forecast of the next change, is also compared with the ANN's effort, and is shown to be inferior.

1. INTRODUCTION

The 10 Year Commonwealth Treasury Bond is issued by the Australian Federal Government to raise finance. It is a fixed interest instrument, paying a coupon rate of interest, fixed for the life of the bond, twice yearly. When interest rates move, often due to inflationary or deflationary influences, the rate of interest which this Bond pays remains fixed, but the value, or price, of the bond on the secondary market will vary. The value of these bonds currently on issue is around \$62B, generating a daily turnover of roughly \$4.5B.

Traders and investors in these bonds must attempt to forecast yield behaviour, both in the short and longer term. Given the turnover associated with trade in this Bond, any improvement in yield forecasting would be of great significance to those involved in their purchase or sale. Two main schools of thought exist regarding prediction of price behaviour. Fundamental analysts rely on fundamental data to formulate a prediction. The data used consist of variables like interest rates, Consumer Price Index changes, prices of certain raw materials, etc. They are therefore concerned with the impact of other time series on that being forecast. On the other hand, technical analysts maintain that the price of a commodity or security subsumes anything that can possibly affect it. This can be economic, political, social, psychological or fundamental. Therefore, to predict, the technical analyst is concerned only with price behaviour, and focuses on the behaviour of the time series itself.

This study uses the Bond time series itself, and combines this with time series of other relevant economic, fundamental indicators. Extensive use of moving averages of indicators is used here to enhance the network's predictive power. Examination of the literature indicates that this approach does not appear to have been utilised in the past.

Artificial neural networks (ANNs) offer themselves as a possible tool in the forecasting of financial and economic time series. They operate, and make decisions based on their own experience (Rumelhart et al, 1986). They can be presented with large quantities of historical data from which they can "learn" to make decisions on new data. This behaviour can be compared to that of humans, and in fact, ANNs are loosely based on a crude interpretation of the brain's mode of operation, although they are strictly mathematical in their operation.

The inputs to the network can consist of values taken from the time series itself or values derived from the time series (Chakraborty et al, 1992; Refenes et al, 1993) and other values which influence the behaviour of the time series in question (Kimoto et al, 1990; Barr and Mani, 1994; Grudnitski and Osborn, 1993). The objective is to discern some relationship between the previous behaviour of the series and that of the current behaviour. This would be an indication of non-random behaviour, and thus neural networks have been used in testing the potency of the Efficient Market Hypothesis as indicated below. Any relationship detected is likely to be non-linear.

Several conventional methods for generating forecasts exist (Hall, 1994; Hoff, 1983) and in some cases comparison between them and ANNs have been made (Tang et al., 1991). Others have combined the two techniques in an attempt to increase forecasting accuracy (Lachtermacher and Fuller, 1995). It is claimed that ANNs are useful in detecting dynamic non-linear relationships which, it is postulated, exist in many financial and economic time series (Bosarge 1991). That is, the series are regarded as non-random, and thus some form of forecasting is possible (Azoff, 1994).

The Efficient Market Hypothesis (EMH) (Fama, 1965; Fama, 1970; Haugen, 1993), however, postulates that forecasting of this nature is impossible, and maintains that using yesterday's yield as a forecast for today's is likely to be as accurate as any forecasting tool. Another possible naive strategy would be to use the change in yield from yesterday to today as a forecast of tomorrow's change. Much work has been done in this area, both addressing the challenge to forecast and to refute the EMH (White, 1988; Bosarge, 1991).

ANNs require historical data to enable training to take place. This data can be presented to the ANN as a moving window along the time series being predicted and that of any other leading indicator used in the prediction process. This study uses a different approach. Instead of moving windows, appropriate moving and exponential averages are used as inputs to the network. The most effective averages are obtained by examining the underlying economic fundamentals driving the behaviour of the inputs, and by an experimental process of sensitivity analysis. It should be noted that many technical analysts, attempting to predict the behaviour of stock and bond prices, use moving and exponential averages in their forecasting methods. Preprocessing of this nature could be likened to the addition of another layer to the network. By adopting the method explored in this study, the network is rendered more parsimonious and facilitates its convergence to a meaningful result.

In this study, to detect any non-random behaviour, and thus to enable forecasting using ANNs, a comparison of the results obtained using yesterday's yield as a forecast for today's and that of an ANN forecast is made. A comparison between the prediction of an ANN and the use of the previous change in yield as a prediction for tomorrow's change is also made. In both the case of using the previous yield, and the previous change, it was found that ANNs obtained a superior forecast, both in the direction of movement, and in closeness of value to the actual result. This points to the existence of non-randomness in at least the financial time series under investigation, and to the possible use of ANNs in forecasting the behaviour of other such series.

2. DATA

Monthly yields for the Bond are used from February 1986 to March 1995. Similarly, Consumer Price Index (CPI), 90 Day Bill yields and USA Government securities (ten years or over) yields are used. These are all also monthly, except for CPI, where only quarterly figures were available.

The 90 Day Bill yield, which trades daily, is an average of the assessed daily market yield for the week ended last Wednesday of the month, using the mid-point of the bid-offer spread. The 10 Year Commonwealth bond yield is the closing yield for the last business day of the month. Again, the mid-point of the bid-offer spread is used. The USA Government securities yield is an unweighted average of securities callable on ten years or more as defined by the Federal Reserve, New York.

The above data are made available by the Australian Bureau of Statistics(ABS) on the Internet.

These data are chosen as the most relevant in directing the behaviour of the Bond yield (Valentine, 1991). However, the combination and/or weighting of the influence of the various data presents a problem in using it to predict the Bond yield movement. This is the task of the ANN.

3. DATA PREPROCESSING

Some preprocessing of raw data before being used as input for a neural network is deemed useful, and in many cases essential (Azoff, 1994). This assists the ANN to formulate any relationship, and to speed this process, which is referred to as "training". The nature of the preprocessing required can be determined from any existing knowledge of yield behaviour, and in some cases, experimentation.

It is necessary to input historical data to the ANN so that training can occur. One method of inputting this is to use a "moving window" of data, with input vectors consisting of a fixed number of previous months' relevant variables which is moved along the time series. Here, the alternative of determining relevant moving averages to capture short and long term memory is used. This means that the number of inputs to the ANN is reduced, as is the effect of "noise". This reduces the amount of training required, and also helps the network to converge more readily.

The latest yield of the Bond would, of course, have significant influence on its next movement, and this data are left in this form for input to the ANN. The recent past behaviour of the Bond yield series could also be relevant to future behaviour, therefore some input which contains this short-term memory would be necessary. Moving averages of the yield were generated and sensitivity analysis, using a neural network, was carried out. It was found that the two month moving average had the greater effect in improving a forecast. A similar approach was used to determine what indicator would best reflect any influential long-term memory effect. An exponential moving average of the bond, using a weight of 0.3, was determined to be the most effective, again using sensitivity analysis on an ANN.

The difference in yield on the 10 Year Commonwealth Bond and that of the 90 Day Bill was used as an indication of the difference between short and long term interest rates, representing a rough term structure of interest rates. This is also recognised as a marked influence on the behaviour of the yield in question (Peirson et al, 1995). The two month moving average of this value, and its exponential moving average (weight = 0.3) were found, like those of the Bond, to be the most effective indicators in representing the influence of recent and more distant historical behaviour.

The change in value of CPI is a relevant influence on the Bond yield behaviour (Valentine, 1991). This was used as input, as was the two month moving average of this change, and the exponential moving average (weight = 0.3) of the change. Again, they were deemed the most suitable after testing on an ANN.

Both the yield on USA securities similar to the Bond and the USA - Australian currency exchange rate influence the Bond behaviour (Valentine, 1991). They were combined as input for the ANN. The yield was multiplied by the current exchange rate (\$Aust/\$US). The two month moving average and the exponential moving average of this value were again determined to be the most useful as further inputs. Thus twelve inputs for the ANN to predict the Bond yield were determined.

4. METHOD

An ANN consisting of 12 input nodes (to take the variables described in the previous section), 17 nodes in a single hidden layer and one node in the output layer, to deliver the ANN's prediction of the Bond Yield, was created. The number of nodes in the hidden layer was calculated using the following formula (Azoff, 1994; Neuroshell2, 1995): number of hidden neurons = 1/2(inputs + outputs) + square root of number of patterns in training file.

This method of calculating the number of hidden nodes ensures an adequate number to allow the ANN to function properly without "overfitting", where the network memorises rather than generalises. When this happens prediction using new data does not occur. It should be said, however, that an adequate quantity of data is necessary for any meaningful training to occur (Azoff, 1994).

The ANN was trained using standard back-propagation, with the hidden layer using a logistic threshold function. A threshold function of this type introduces nonlinearity to the network, which enables the network to predict. Depending on the threshold function used, the inputs have to be scaled either between 0 and 1 or between -1 and 1. In this case, where a logistic, or sigmoid function is used, the inputs were scaled between 0 and 1. Scaling is achieved by using the maximum and minimum of the data set for each input. In this instance, the maximum and minimum used were set at the values 2.5 standard deviations from the mean. Data outside this range were set to 1 and 0 respectively. Output from the network is initially in the same range as the scaled input. The above process is reversed to obtain results in the required range.

Training vectors were presented to the ANN until no further error reduction was obtained after 20000 vectors had been presented. The training vectors consisted of the 12 variables described in the Data Preprocessing section for each month, with the next month's Bond yield as the output. For the experiments performed, training took approximately two minutes on a 75mhz Pentium processor. The ANN was tested on approximately 10% of the data, randomly selected, and trained on the remainder. This was repeated ten times, using the same data set, but with different random selections each time.

5. RESULTS

The ANN made a prediction of the next month's yield having been given the inputs, as previously described, for the preceding month. The absolute error of this prediction was calculated. The ANN was given a trend score, 1 if it predicted the direction of yield movement correctly, 0 if wrong. An accuracy score was also allocated to the prediction, where 1 was scored if the prediction was nearer the actual yield than that of the previous month's yield.

A trend score was calculated for using the previous month's change in yield as a forecast, enabling a comparison of this with that of the ANN's score. An accuracy score was also calculated where the accuracy of the ANN's prediction was compared with that of using the previous month's change in yield.

These results are summarised in Table 1, where the ANN's prediction is compared with that of using the last previous yield as a forecast. Table 1 also includes the average absolute error of the ANN forecast incurred in each sample. In Table 2, the accuracy of the ANN's prediction is compared with that of using the last previous change in yield to predict the next month's yield. A trend score is also shown for using the change as a trend predictor.

Table 1

Comparison of ANN Prediction with Previous Bond Yield

No. in Sample	*Trend Score*	*Accuracy Score*	*Ave Error %*
19	12	11	3.93
18	13	12	3.10
12	8	6	2.82
12	9	9	2.90
7	4	4	1.56
15	11	11	2.59
16	10	9	2.89
13	9	7	4.00
12	9	8	2.42
11	7	6	2.54

Probability of trend score success = 0.6815
Standard deviation = 0.0401
95% confidence interval = 0.7601 to 0.6029

Probability of accuracy score success = 0.6148
Standard deviation = 0.0419
95% confidence interval = 0.6969 to 0.5327

Table 2

Comparison of ANN Prediction with Previous Change in Bond Yield

No. in Sample	*Accuracy Score (ANN)*	*Trend Score (Change)*
19	11	8
18	12	10
12	6	5
12	9	4
7	4	1
15	8	7
16	11	6
13	7	7
12	8	4
11	7	4

Probability of accuracy score success = 0.6148
Standard deviation = 0.0419
95% confidence interval = 0.6969 to 0.5327

Probability of trend (chge) score success = 0.4148
Standard deviation = 0.0424
95% confidence interval = 0.4979 to 0.3317

6. DISCUSSION

Equally likely outcomes of success or failure in predicting both trend and yield itself can be regarded as the null hypothesis. If the outcomes are equally likely, the probability of each occurring would be 0.5. Given that the distribution is binomial, the data yields a standard deviation of 0.0430

The results of the comparison between the ANN and using the previous change in the Bond's yield as a predictor, shown in Table 2, also points to non-random behaviour of the time series under examination. The ANN was shown to be more accurate on approximately 61% of occasions, with a 95% confidence interval of 69.69% to 53.27%.

Using the previous change in Bond yield as a guide to the direction of the next change is clearly an inferior method, with a success rate of only around 41%, with a 95% confidence interval of 49.79% to 33.17%.

When the above results are compared to the null hypothesis, which requires the probabilities of success to lie in the interval 0.5842 to 0.4158, it is clear that all the probabilities of success obtained by the ANN lie outside and above the 95% confidence interval of the null hypothesis. It is therefore rejected. The ANN is clearly superior to a hold strategy, and to the naive strategy of using yesterday's change as a predictor for that of tomorrow.

The model used in this study could be of use as another tool to predict the direction of the Bond yield. Investors or traders can buy or sell on a direction prediction, and would probably find the model of use in assisting a decision with respect to hedging, in order to, for instance, protect current holdings.

7. FURTHER RESEARCH

Much remains to be researched here. The ANN used in this study is of the simplest type, using back-propagation and one hidden layer. The effect of varying threshold functions should be examined. The topology of the ANN should also be varied to see if improvements in forecasting are obtained by using a different number of hidden layers. Varying the number of hidden neurons should also be attempted. Back-propagation is used here, which is probably the most popular learning method for ANNs. Others could be examined. A recurrent ANN, where feedback is used may also be useful in improving performance. Comparison of the method used here with that of using moving windows of the time series used as input would be a useful exercise. Other times series, the behaviour of which may be similar to that of the Bond examined here, should also be investigated.

8. ACKNOWLEDGMENTS

The author would like to thank Dr. Don Watson and Dr. George Nassios for their comments and participation in many useful discussions.

8. REFERENCES

E.M. Azoff. *Neural Network Time Series Forecasting of Financial Markets.* John Wiley and Sons Ltd., Chichester, U.K.1994.

D. Barr and G. Mani. Using Neural Nets to Manage Investments. *AI Expert*, February, 1994, pp. 16 - 21.

W.E. Bosarge Jr. Adaptive Processes to Exploit the Nonlinear Structure of Financial Markets. *Santa Fe Institute of Complexity Conference: Neural Networks and Pattern Recognition in Forecasting Financial Markets.* 1991.

K. Chakraborty, K. Mehrotra, C.K. Mohan and S. Ranka. Forecasting the Behavior of Multivariate Time Series Using Neural Networks. *Neural Networks*,1992, Vol. 5, No. 6, pp. 961 - 970.

E. Fama. The Behaviour of Stock Market Prices. *Journal of Business,* January, 1965.

E. Fama. Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, May, 1970, pp. 383 - 417.

G. Grudnitski and L. Osburn. Forecasting S&P and Gold Futures Prices: An Application of Neural Networks. *The Journal of Futures Markets.* 1993, vol. 13, No. 6, pp. 631 - 643.

S. Hall (Ed.). *Applied Economic ForecastingTechniques.* Harvester Wheatsheaf, Hemel Hempstead, U.K., 1994.

R.A. Haugen. *Modern Investment Theory.* 3rd ed., chpt 23, 24, Prentice - Hall, Englewood Cliffs, New Jersey, 1993.

J.C. Hoff. *A Practical Guide to Box-Jenkins Forecasting.* Lifetime Learning Publications, Belmont, California, 1983.

T. Kimoto, K. Asakawa, M. Yoda and M. Takeoka. Stock Market Prediction System with Modular Neural Networks. *Proceedings of the IEEE International Joint Conference on Neural Networks, San Diego, C.A.* 1990, ppI1 - I6.

G. Lachtermacher and J.D. Fuller. Backpropagation in Time Series Forecasting. *Journal of Forecasting*, 1995, Vol. 14, 381-393.

NeuroShell2. Ward Systems Group, Inc., Frederick, MD. 1995.

G. Peirson, R. Bird, R. Brown and P. Howard. *Business Finance*, 6th Edition, McGraw-Hill, Roseville, NSW, 1995.

A.N. Refenes, M. Azema-Barac, L.Chen and S.A. Karoussos. Currency Exchange Rate Prediction and Neural Network Design Strategies. *Neural Computing and Applications,* 1993, vol. 1, pp. 46 - 58.

D.E. Rumelhart, G.E. Hinton and R.J. Williams. *Parallel Distributed Processing,* MIT Press, Cambridge, MA., 1986.

Z. Tang, C. de Almeida and P.A. Fishwick. Time Series Forecasting Using Neural Networks vs. Box-Jenkins Methodology. *Simulation,* November 1991, pp303-310.

H. White. Economic Prediction Using Neural Networks: The Case of IBM Daily Returns. *Proceedings of the IEEE International Conference on Neural Networks,* July, 1988, pp. II451 - II458.

T. Valentine. *Interest Rates and Money Markets in Australia.* Financial Training and Analysis Services Pty. Ltd. 1991.