Application of a Neural Network Approach to Implement a Bivariate Fuzzy Time Series Model

Kunhuang Huarng[1] Hui-Kuang Yu[2] Yu-Fang Lin[3]

[1]Depratment of International Trade, Feng Chia University
[2]Depratment of Public Finance, Feng Chia University
[3]Depratment of Business Administration, Feng Chia University

Abstract

This study applies neural network to establish fuzzy relationships for a bivariate fuzzy time series model. The inputs are the closing price of a stock index and that of a futures index, and the output is the closing price of the stock index of the next day. The forecasting results show that the proposed model outperforms the previous fuzzy time series models.

Keywords: fuzzy time series, neural networks, bivariate, stock index, futures index

1. Introduction

Time series studies have been so popular for a long time. To solve nonlinear problems, fuzzy time series models are believed to be more suitable. Different fuzzy time series studies have been proposed for various applications, such as enrollment (Song and Chissom, 1993), temperature (Chen and Hwang, 2000), stock index (Yu, 2005), etc. However, most of these studies were for one variable only. Recently, some models have been proposed for two variables (Huarng, 2001; Hsu, Tse, and Wu, 2003), which rendered better forecasting results. Hence, this study also proposes a model for two-variable problems. We use the neural network to establish fuzzy relationships. TAIEX (Taiwan Stock Exchange Capitalization Weighted Stock Index) and its corresponding index futures TAIFEX are used as inputs to forecast the TAIEX of the next day.

2. Literature Review

There have been many studies devoted to the forecasting of stock markets. Dickinson (2000) verified the degree to which stock markets are integrated. In that study, it was found that stock prices were affected by a number of key macroeconomic variables. These economic forces, the stock market, and the work on international stock market integration provided a suitable empirical basis on which to investigate the relationship between fundamentals and the stock market. Kwon and Shin (1999) found that the Korean stock market reflected macroeconomic variables on stock market indices. That study reflected that stock price indices were cointegrated with a set of macroeconomic variables - production index, exchange rate, trade balance, and money supply. Grudnitski and Osburn (1993) examined the feasibility of employing neural networks to forecast price changes of Standard & Poor’s 500 Stock Index and gold futures based on past price changes and historical open interest patterns. The neural networks were used to forecast the moving signs of the next month’s price change for S&P and Gold trades. So and Tse (2004) used Hasbrouck and Gonzalo and Granger common-factor models and the M-GARCH model to prove that the movements of three markets, Hang Seng Index, Hang Seng Index futures, and the tracker fund were interrelated. Stoll and Whaley (1990) found that, S&P 500 and Major Market Index futures returns tended to lead stock market returns by about five minutes on average.

3. Data

The data used are the daily closing prices of the stock market and futures market in Taiwan. The data are selected from TEJ Database (Jan. 1999 – Dec. 2003). Closing prices have been used in many previous studies (Mavrides (2003); Peláez (2000); Kwon & Shin (1999)), and so does this study.

4. Methodology

4.1 Fuzzy Time Series Model – One Variable

Chen’s model was proposed for the one-variable problems (Chen, 1996). So the model is taken as the counterpart for comparison. The length of intervals is set as 100. The forecasting process for the year of 1999 is briefly described as follows. Because the smallest stock index in 1999 was 5474.79, the initial value for the universe of discourse is set to 5400. The closing price of 1/5/1999, for example, was 6152.43, which is fuzzified to $A8$. The closing price of next day was

6199.91, which is fuzzified to *A*8. Hence, a fuzzy logic relationship is established as *A*8→*A*8. Then, we follow Chen's model for forecasting.

4.2 Fuzzy Time Series Model – Two Variables

We use the same way for fuzzification and for establishing fuzzy logic relationships. For example, TAIEX of 1/5/1999 was 6152.43, which is fuzzified to *A*8. Similarly, the TAIFEX is fuzzified to *B*7. Hence, a bivariate fuzzy logic relationship is established as *A*8, *B*7→*A*8. Some results are listed in Table 1.

4.3 Neural Networks

The data are divided into training and testing sets. The data of January to October are used as the training instances, while those of November and December are testing instances. The software we use is PCNeuron4.0[1].

PCNeuron provides the back-propagation algorithm. To make it simple, we set 1 layer for the hidden layer to prevent complicated calculations. The number of nodes in the hidden layer is set as $m+n$, where m is the number of inputs in the input layer and n is the number of outputs in the output layer. The structure of the neural network is depicted in Figure 1. Hence, there are two, three, and one nodes in the input, hidden, and output layer, respectively. As for the neural network training, we set the epochs as 1000 and the testing period 10. The setup is summarized in Table 2.

5. Results

The output of the neural network is the forecast stock index of the next day, which is a fuzzy set. Taking the year 1999 as an example, the learning process of convergence is depicted in Figure 2. The horizontal axle of the figure is the number of the learning cycle; the vertical axle is the corresponding root mean square error (RMSE). In that figure, the curve draws close to a horizontal line quickly, which indicates that learning effect is satisfactory.

We also use the scatter plot diagrams to verify the forecasting results. A scatter plot diagram offers a two-dimensional plot of corresponding pairs of variables. The forecasting results of year 1999 are depicted in Figure 3. The horizontal axle of the figure is the actual fuzzy set; the vertical axle is the forecast fuzzy set. All the outputs are plotted on the figure. The closer to the diagonal line, the better the forecasting results. From Figure 3, we see that most of the forecasts (spots) are near diagonal line. This implies that the forecasting results of the model are satisfactory.

We also measure the coefficient correlation to further test the model. The coefficient correlation of forecast and RMSE are listed in Table 3. Each coefficient correlation is larger than 0.76.

After having the forecasting fuzzy sets, we proceed the process of defuzzification. We apply the median of each interval as the defuzzified value for stock index. For example, on November 2, 1999, the forecast stock price fuzzy set is 25, whose interval is between 7800 and 7900. Hence, we use its median 7850 as the defuzzified stock price. The actual closing price of that day was 7721.59. Then, we can compute errors, percentage of error, and squared error. Use those values to compute the RMSE to evaluate the performance for year 1999. The RMSE was 107.9598 and the mean percentage of error was 0.01034. The detail is listed in Table 4.

After getting all the forecast stock index and RMSEs, we can compare the performance of the Chen's model with the proposed bivariate model. The comparison is listed in Table 5. In most years, the performance of the bivariate model is better than Chen's model.

6. Conclusions

Many scholars adopted relevant macroeconomic variables to forecast stock markets. In this study, we choose to use stock index and futures index, which is highly related to the stock index as the input variable to forecast the stock index of the next day. By using the neural network to establish fuzzy relationships, forecast stock index is closer to its corresponding actual index. From the RMSEs, we show that the bivariate model outperforms Chen's model in most years.

Acknowledgments

Special thanks to Professor I-Cheng Yeh, Department of Information Management, Chung Hua University, Taiwan, for providing the neural network software, PCNeuron.

References

1. S.-M. Chen (1996), "Forecasting enrollments based on fuzzy time series," *Fuzzy Sets and Systems*, 81, 311-319
2. S.-M. Chen and J.R. Hwang (2000), "Temperature prediction using fuzzy time series," *IEEE Trans. Syst., Man, Cybern.* Part B., 30(2) 263-275
3. D.G. Dickinson (2000), "Stock market integration and macroeconomic fundamentals: An empirical analysis, 1980-95," *Applied Financial Economics*, 10, 261-276

[1] PCNeuron 4.0 is a software provided with "Applications of neural networks," by I-Cheng Yeh, 1997, (ISBN: 957-652-997-2), in Chinese.

4. G. Grudnitski and L. Osburn (1993), “Forecasting S&P and gold futures prices: An application of neural networks,” *The Journal of Futures Markets*, 13(6), 631-643
5. Y.-Y. Hsu, S.-M. Tse, and B. Wu (2003), “A new approach of bivariate fuzzy time series analysis to the forecasting of a stock index,” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 11(6), 671-690
6. K. Huarng (2001), “Heuristic models of fuzzy time series for forecasting,” *Fuzzy Sets and Systems*, 123, 137-154
7. C. Jung and R. Boyd (1996), “Forecasting UK stock prices,” *Applied Financial Economics*, 6, 279-286
8. C.S. Kwon and T.S. Shin (1999), “Cointegration and causality between macroeconomic variables and stock market returns,” *Global Finance Journal*, 10(1), 71-81
9. M. Mavrides (2003), “Predictability and volatility of stock returns,” *Managerial Finance*, 29(8), 46-57
10. R.F. Peláez (2003), “Ten-year forecasts of real stock price changes,” *Empirical Economics*, 28, 417-429
11. R.W. So and Y. Tse (2004), “Price discovery in the Hang Seng Index Markets: Index, futures and the tracker fund,” *The Journal of Futures Markets*, 24 (9) 887-907
12. Q. Song and B.S. Chissom (1993), “Forecasting enrollments with fuzzy time series - Part 1,” Fuzzy Sets and Systems, 54, 1-9
13. H.R. Stoll and R.E. Whaley (1990), “The dynamics of stock index and stock index futures returns,” *Journal of Financial and Quantitative Analysis*, 25(4), 441-468
14. H.-K. Yu (2005), “A refined fuzzy time-series model for forecasting,” Physica A, 346 (3-4), 657-681

Table 1. Fuzzy Sets and Fuzzy Logic Relationships

Date	Stock Closing	Fuzzy Set A	Futures Closing	Fuzzy set B	Fuzzy Logic Relationship
1999/1/5	6152.43	8	6120	7	$A8, B7 \rightarrow A8$
1999/1/6	6199.91	8	6245	8	$A8, B11 \rightarrow A11$
1999/1/7	6404.31	11	6510	11	$A11, B11 \rightarrow A11$
1999/1/8	6421.75	11	6452	10	$A11, B10 \rightarrow A11$
1999/1/11	6406.99	11	6435	10	$A11, B10 \rightarrow A10$
1999/1/12	6363.89	10	6390	9	$A10, B9 \rightarrow A10$
1999/1/13	6319.34	10	6352	9	$A10, B9 \rightarrow A9$
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
1999/12/28	8448.84	31	8564	31	
	Lowest of the Year: 5474.79	Initial Value: 5400	Lowest of the Year: 5505	Initial Value: 5500	

Table 2. Neural Network Setup

Number of Nodes in the Input Layer	2
Number of Hidden Layers	1
Number of Nodes in the Hidden Layer	3
Number of nodes in the Output Layer	1
Epochs	1000
Testing Periods	10

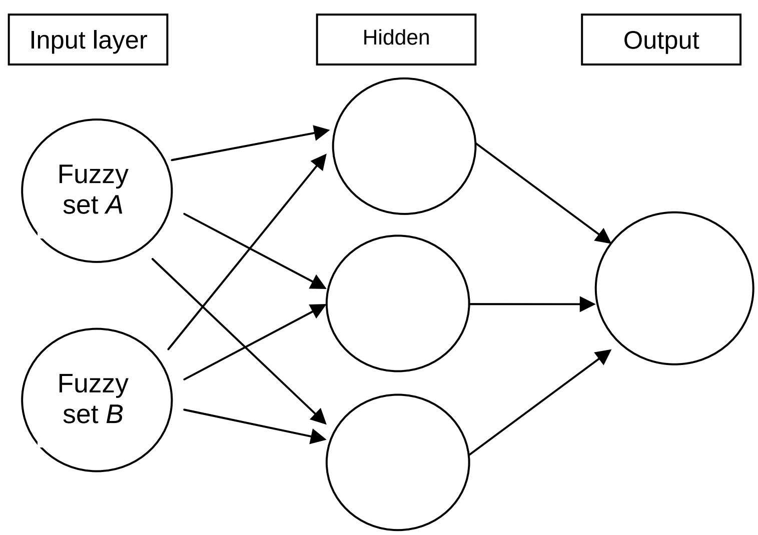

Figure 1. Neural Network Structure

Table 3. Coefficient Correlations and RMSEs of Forecasts

	1999	2000	2001	2002	2003
Corre. Coef.	0.8664	0.9264	0.9771	0.7698	0.8369
RMSE	1.23750	2.59230	1.33330	0.72437	0.59154

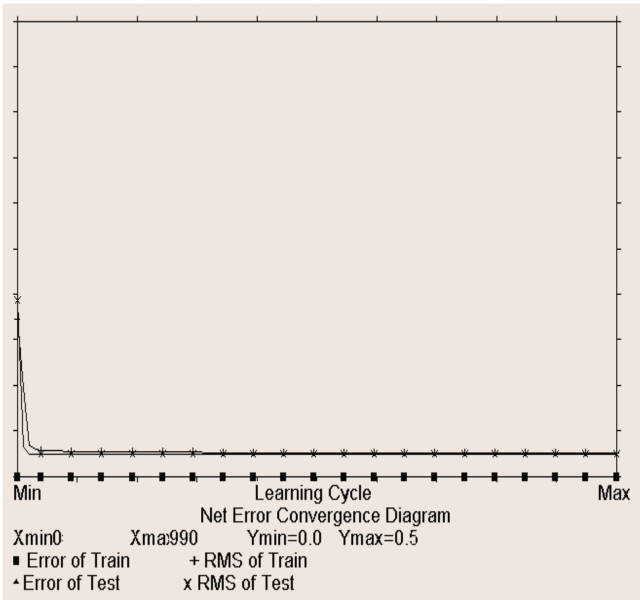

Figure 2. Learning Process for Year 1999

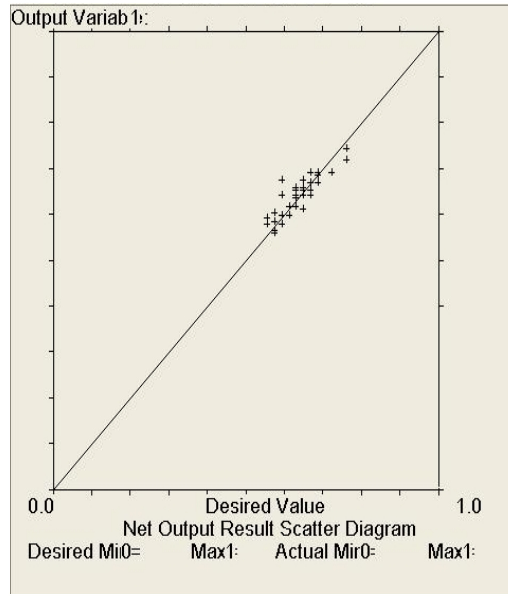

Figure 3. Scatter Plot Diagram for Year 1999

Table 4. Forecasts and Errors

Date	Actual Stock Index	Actual Fuzzy Set	Forecast Fuzzy Set	Forecast Stock Index	Error	Error	Error²
1999/11/2	7721.59	24	25	7850	128.41	0.0166	16489.13
1999/11/3	7580.09	22	25	7850	269.91	0.0356	72851.41
1999/11/4	7469.23	21	23	7650	180.77	0.0242	32677.79
1999/11/5	7488.26	21	22	7550	61.74	0.0082	3811.828
1999/11/6	7376.56	20	22	7550	173.44	0.0235	30081.43
1999/11/8	7401.49	21	21	7450	48.51	0.0066	2353.22
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
1999/12/28	8448.84	31	30	8350	98.84	0.0117	9769.346

Table 5. Comparison of Performance

	1999	2000	2001	2002	2003
Mean Error in Chen's Model	0.012326	0.024273	0.023813	0.017123	0.01037
RMSE in Chen's Model	120.0865	176.3216	147.8381	101.1767	74.4615
Mean Error in Bivariate Model	0.010341	0.042867	0.021686	0.014101	0.007704
RMSE in Bivariate Model	107.9598	259.3887	132.5152	85.09618	57.55638