CIRJE-F-165

Can the Neuro Fuzzy Model Predict Stock Indexes Better than its Rivals?

Chin-Shien Lin
Providence University

Haider Ali Khan
University of Denver / CIRJE, University of Tokyo

Chi-Chung Huang
Providence University

August 2002

Can the neuro fuzzy model predict stock indexes better than its rivals?

Chin-Shien Lin
Associate Professor
Department of Finance
Providence University
200 Chungchi Rd., Shalu, Taichung Hsien, 433 Taiwan, R.O.C., e-mail: cslin@pu.edu.tw

Haider A. Khan
University of Denver
Denver
Co. 80208 USA
Tel. 303-871-4461/2324
Fax 303-871-2456
hkhan@du.edu

Chi-Chung Huang
Graduate School of Business Administration
Providence University
Revised, March 2002.

Can the neuro fuzzy model predict stock indexes better than its rivals?

Abstract

This paper develops a model of a trading system by using neuro fuzzy framework in order to better predict the stock index. Thirty well-known stock indexes are analyzed with the help of the model developed here. The empirical results show strong evidence of nonlinearity in the stock index by using KD technical indexes. The trading point analysis and the sensitivity analysis of trading costs show the robustness and opportunity for making further profits through using the proposed nonlinear neuro fuzzy system. The scenario analysis also shows that the proposed neuro fuzzy system performs consistently over time.

Key words: linear, nonlinear, KD indexes, buy and hold, neuro fuzzy

1. Introduction

Accurate predictions of stock market indexes are important for many reasons.Chief among these are the need for the investors to hedge against potential market risks, and the opportunities for market speculators and arbitrageurs to make profit by trading indexes. Clearly, being able to accurately forecast stock market index has profound implications and significance to both researchers and practitioners.

The most commonly used techniques for stock price forecasting are regression methods and ARIMA models (Box and Jenkins, 1970). These models and methods have been used extensively in the past. However, they fail to give an accurate forecast for some series because of their linear structures and some other inherent limitations. Although there are ARCH/GARCH models (Engle, 1982; Bollerslev, 1986) to deal with the non-constant variance, still some series cannot be explained or predicted satisfactorily. Recent research in the area of neural network technique has shown that neural networks possess the properties required for relevant applications, such as nonlinear and smooth interpolation, ability to learn complex non-linear mappings, and self-adaptation for different statistical distributions.

However, neural network cannot be used to explain the causal relationship between the input and output variables. This is because of the essentially black box like nature of the many existing neural network algorithms. A neural network cannot be initialized with prior knowledge. The network usually must learn from scratch. The learning process itself can take very long with no guarantee of success.

On the other hand, the fuzzy expert system approach has been applied to different forecasting problems (Bolloju, 1996; Kaneko, 1996; Al-Shammari and Shaout, 1998), whereby the operator's expert knowledge is used for prediction. Although the fuzzy-logic-based forecasting shows promising results, the process to construct a fuzzy logic system is subjective and depends on somewhat heuristic processes. The choices of membership functions and rule base have to be developed heuristically for each scenario. The rules fixed in this way may not always yield the best forecast, and the choice of membership functions still depends on trial and error.

With these advantages and disadvantages of neural network and fuzzy logic, a neuro-fuzzy framework has emerged by combining the learning ability of the neural network and the functionality of the fuzzy expert system. Its application can be found in the work of Dash et al. (1995), Lie and Sharaf (1995), Studer and Masulli (1997), and Padmakumari et al. (1999). Such a hybrid model is expected to provide humanly understandable ‘fuzzy’ meanings through the creation of more reliable knowledge base through the learning ability of neural network.

Now, some researchers such as Jacobs and Levy (1989) have made the interesting claim that the the stock market is not an ordered system that can be explained by simple rules, nor is it a totally random system for which no predictions are possible. In fact, they claim that the market is a complex system, in which only portions of the system's behavior could be explained and predicted by a set of complex relationships among the variables.

Recognizing the complex characteristics of the stock market leads us to ask if the predictability of the various indexes could be improved by using nonlinear models with endogenous learning capabilities in a fuzzy real world environment. The answer turns out to be ‘yes’, and the specific modeling approach we use demonstrates the advantages of the neuro-fuzzy technique. In particular, our modeling of learning via the neuro fuzzy approach leads to better predictions in the case of utilizing the KD technical indexes to describe the stock index movement. The purpose of this paper is to show this concretely through an investigation of the relative profitability of this proposed KD based neuro-fuzzy trading system.

Specifically, then, the major contributions of this study are (1) to demonstrate and verify the predictability of stock index return by applying neuro fuzzy technique to KD index estimates; (2) to compare the performance of linear and nonlinear models based on KD indexes; (3) to show the robustness of this proposed KD based neuro-fuzzy model; (4) and to show the existence of market opportunities for further profitability via results from this proposed model and its profitability consistency over time.

The remainder of this paper is organized as follows. In section 2, the past work about the technical analysis is reviewed. Section 3 describes how the KD based neuro fuzzy trading system is constructed and how the alternative benchmark models are formulated. The empirical results are shown in section 4. Finally section 5 provides some concluding remarks.

2. Literature Review

In general the approaches to predict stock price could be roughly categorized into two kinds, fundamental analysis and technical analysis. Fundamental analysis is based on macroeconomic data, such as exports and imports, money supply, interest rates, inflationary rates (Fama and Schwert 1977, Campbell 1987, and Fama and French 1988a, 1988b), foreign exchange rates, unemployment figures, and the basic financial status of companies such as dividend yields, earnings yield, cash flow yield, book to market ratio, price-earings ratio, lagged returns, and size. (Basu, 1977; Fama and French, 1992; Lakonishok, Shleifer and Vishny, 1994).

Technical analysis is based on the rationale that history will repeat itself and that the correlation between price and volume reveals market behavior. Prediction is made by exploiting implications hidden in past trading activities, and by analyzing patterns and trends shown in price and volume charts (Epps, 1975; Smirlock and Starks, 1985; Rogalski, 1978; Bohan, 1981; and Brush, 1986).

Basically the test of weak form efficient market is to test whether there exists excess return by using technical analysis. There have been some researches claiming the existence of the weak form of efficient market (Fama 1965; Fama and Blume 1966; Jensen 1967). Also there exist some researches claiming that the weak form efficient market does not exist (Sweeney 1986; Brock, Lakonishok and LeBaron 1992; Bessembinder and Chan 1995). So far the research remains inconclusive.

One of the most commonly used methods in technical analysis is the moving average filter rule. The criterion is that the buying signal happens when the short term moving average line breaks through the long term moving average line from down, and the selling signal happens when the short term moving average line breaks through the long term moving average line from up. The logic behind this rule is to identify periods when expected returns deviate from unconditional means (Bessembinder and Chan 1995). Although Fama and Blume (1966) and Jensen and Benington (1970) concluded that the filter rules are not useful, Brock et al. (1992) and Sweeney (1986) showed the non-trivial ability to predict the price changes by using the filter rules. On the other hand, Bailey et al (1990) and Pan et al. (1991) present evidence that prices in some stock markets exhibit substantial deviations from random walk behavior.

The reasons for the different empirical results can come from the different samples, different technical indexes, or different rules. However, in this paper we consider that there exists some relationship between the technical indexes and stock price, the question is that how to use the information to explore this relationship. In other words, the specification of the function form is a difficult problem. A data driven method to construct a model can be an effective way.

Similar to the filter rules, KD technical rules proposed by Lane (1957), is trying to capture the period when expected returns deviate from unconditional means by using K and D indexes instead of the moving averages. Essentially K and D indexes with the advantages of momentum, relative strength, and moving average, and with the consideration of the highest and the lowest prices, are expected to be capable of capturing the short-term variance. However, the KD filter rules could be too simple to be effective. Besides, the parameters for the rules are arbitrary. Therefore, this paper is trying to develop a model based on the knowledge contained in KD technical rules by using neuro-fuzzy. Investment performance is simulated based on the signals produced by the system. Thirty world wide known stock indexes are used as the testing sample. The standard regression model, GARCH-M, and neural network are used to derive comparative results on relative prediction performances.

3. Methodology

3.1 KD Trading System[1]

The commonly used K D indicators are calculated as follows.

$$RSV_t = (C_t - L_9) * 100/(H_9 - L_9) \quad (1)$$

$$K_t = 1/3 * RSV_t + 2/3 * K_{t-1} \quad (2)$$

$$D_t = 1/3 * K_t + 2/3 * D_{t-1} \quad (3)$$

where RSV_t is the raw stochastic value for period t, C_t is the closing price for day t, H_9 and L_9 are the highest price and the lowest prices for the latest nine days respectively, K_t and D_t are the values for K and D on day t. If K and D are not available, 50 are used as the initial values for both in general. The trading rules for the KD trading system are as follows.

Rule 1. If D is greater than 80 and K breaks through D from up then sell out.

---

[1] KD indicators were originally developed by Lange( ); but they have become popular only in recent years. Elder(1987) points out:

The logic of this index is based on the observation that, as prices rise, daily closes tend to occur nearer the high end of their recent range. When prices trend higher or are flat but the daily closes begin to sag lower within that range, they signal internal market weakness and its readiness for a trend reversal to the downside. The opposite occurs in down trends; They are confirmed when the closing prices are near the bottom of the recent range. When closing prices move higher within a range, they show internal strength.

Rule 2. If D is less than 20 and K breaks through D from down then buy in.

Rule 3. When the slope of K is flat, the market trend is likely to change.

Rule 4. When the stock price reaches the new highest (lowest), K and D is not reaching the new highest (lowest), the market trend is likely to change.

This is a so- called an expert system. The parameters, 80 and 20, are just the rule of thumb values. The obvious question that one can ask is: can this expert system beat the market? In addition to implementing the original KD expert system, we also try to fit different models that are commonly used in the literature, namely, regression, ARCH_M, neural network, and neuro fuzzy, to describe the stock index movement by considering it as a pattern recognition problem.

To capture the spirit of the KD trading system, we need to choose the appropriate variables to describe the above KD rules. The cross- over phenomenon when K breaks through D from up is depicted in Figure 2. Let K and D represent the level of K and D, K_D the difference between K and D, and K_D_1 the K_D of the previous day. If the cross over phenomenon happened, as depicted in Figure 2, then K_D_1 would be greater than 0 and K_D would be less than 0. Similarly, when K breaks through D from down, then K_D_1 would be less than 0 and K_D would be greater than 0. Therefore, we use K_D_1 and K_D to describe the cross over

---

See also Murphy(1986).

phenomenon. In other words, K, D, K_D_1 and K_D are used to capture the relationship described in rules one and two.

Let KS represent the slope of K, and KT, DT, and PT indicate the trend of K, D, and P respectively. We use KS to capture the relationship describe in rule 3, and KT, DT, and PT to capture the relationship described in rule 4. Let P_t be the stock price for day t, and $trend_t$ denote the rate of return of day t. Then $trend_t$ is calculated as $(P_t - P_{t-1}) / P_{t-1}$. Totally we have 7 input variables, K, D, K_D, K_D_1, KS, KT, and DT, and one output variable, TREND, to describe the KD system. The independent variables for predicting the index returns are all observable on or before the last day of the day preceding the day to be forecasted. In other words, only observable, but not future, data are used as inputs to the forecasting models.

3.2 The construction of a Fuzzy Logic system

To facilitate the exposition, we only explain the model that describes the first two rules, the crossover phenomenon. The complete system is constructed according to exacly the same logic. A fundamental idea of the fuzzy system is that we no longer say, for instance that "IF K is greater than 80." Instead, we will describe the value of K, for instance, to be very_low, low, medium_low, medium, medium_high, high, very_high. In other words, all the input and output variables will be translated into 'fuzzy' ordinary linguistic terms. Table 1 summarizes the variables and their linguistic terms. K, D, and Trend are described by 7, 7, and 5 terms respectively. K_D, and K_D_1 are both described by 2 terms. Each term is described by a membership function. A membership function, expressed as $u_A(x)$, describes the extent to which an object x belongs to a fuzzy set (term) A.

There are many different kinds of membership functions. Popular ones are Z, S, λ and π (Von Altrock 1997: p. nos?). In our case, we used Z, S, and λ for our experiments.[2] Fig. 1(a), 1(b), 1(c), 1(d), 1(e) shows the membership functions for K, D, K_D, K_D_1, and Trend.

Consider the case where the value of K is 80, D is 60, K_D is 0.6 and K_D_1 is –0.4. It can be found in Figure 1(a) that the degree of K being high is 0.6 and the degree of K being very_high is 0.4. Besides, the degrees of K for other linguistics terms are all 0. The membership function for K equal to 80 can be expressed as follows.

$$u_{very\_low}(80) = 0.0, \quad u_{low}(80) = 0.0, \quad u_{medium\_low}(80) = 0.0, \quad u_{medium}(80) = 0.0,$$

$$u_{medium\_high}(80) = 0.0, \quad u_{high}(80) = 0.6, \quad u_{very\_high}(80) = 0.4$$

Similar to K, the values of the linguistic terms for the other variables can be found from Fig. 1(b), 1(c), and 1(d) as follows.

[2] The Z, S,lambda and pi functional forms were all tried in order to choose the most appropriate one.It

For variable D:

$u_{very\_low}(80) = 0.0$, $u_{low}(60) = 0.0$, $u_{medium\_low}(60) = 0.0$, $u_{medium}(60) = 0.2$, $u_{medium\_high}(60) = 0.8$, $u_{high}(60) = 0.0$, $u_{very\_high}(60) = 0.0$

For variable K_D:

$u_{negative}(1) = 0.0$, $u_{positive}(1) = 1.0$

For variable K_D_1:

$u_{negative}(-1) = 1.0$, $u_{positive}(-1) = 0.0$

After the numeric values have been translated into linguistic values, a much more sophisticated rule, for example, can be obtained as follows:

IF K is high, D is medium, K_D is positive and K_D_1 is negative, then Trend is high_inc. (1)

This is so called an inference rule. Each rule consists of two parts, "IF" part and "THEN" part, describing the extent to which the real object satisfies the condition and the response of this system respectively. The operator proposed by Zimmermann and Thole (1978) to represent logical connectives "and" is the minimum value among all the validity values. The validity of each term for rule (1) is summarized in table 2. Therefore, the validity value of the IF part is equal to min{0.6, 0.2, 1.0, 1.0} = 0.2, which also indicates the degree of validity for the

---

turns out that the last three are the most appropriate. (Why?)

"THEN" part. In other words, the validity extent of the system's response "TREND is high_inc" is 0.2.

Let us assume that we have, say five, inference rules. Using these five inference rules, we obtain the following inferences:

(1) The validity extent of "TREND is high_dec" is 0.0.

(2) The validity extent of "TREND is small_dec" is 0.3

(3) The validity extent of "TREND is steady" is 0.0

(4) The validity extent of "TREND is small_inc" is 0.2

(5) The validity extent of "TREND is high_inc" is 0.3

Note that there are the following five fuzzy set membership functions: high_dec, small_dec, steady, small_inc, and high_inc. To facilitate discussion, let us denote high-dec, small-dec, steady, small_inc, and high_inc by f_1, f_2, f_3, f_4, and f_5 respectively. For each membership function f_i, let M_i denote the value of TREND which achieves the maximum value of f_i. If the values are within an interval, we choose the medium of the interval. For instance, by consulting Fig. 1(e), we have the following mapping.

M1 =-0.83

M2 =-0.33

M3 = 0.0

M4 = 0.33

M5 = 0.83

Let the validity extent of "TREND belongs to f_i be denoted as U_i, then the value of TREND will be determined by the following formula:

TREND= $\sum_{i=1}^{5} U_i M_i$

Let us assume that U_i's be 0.0, 0.3, 0.0, 0.2, and 0.3. We will have

TREND = $0.0\times(-0.83)+0.3\times(-0.33)+0.0\times(0.0)+0.2\times 0.33+0.3\times(0.83)=0.20$.

This means that predicted trend is 0.20 for the next day.

Buying signals are recognized when the predicted trend is greater than a predetermined threshold value, and selling signals are recognized when it is less than another predetermined threshold value. Usually both threshold values are set equal to 0. Stocks are bought in when signal is greater than 0, and the stocks are held until the trend is less than 0. Buying signal is ignored when there are stocks on hold, and selling signal is ignored when there are no stocks on hold. [3]

---

[3] In this paper, short sell strategy is not considered. Obviously, the buy and sell decisions become more complicated when"shorting" is allowed.

Note, for example, that using the rule “IF K is high, D is medium, K_D is positive and K_D_1 is negative, then Trend is high_inc.”, we will obtain the validity extent of “TREND is high-inc” as 0.2. However, there are 7 terms for K and D, 2 terms for K-D and K-D-1 and 5 terms for Trend. We have 7*7*2*2*5=980 rules to start with. Therefore, we have two problems: (1)How can we eliminate some of the inference rules which are not practical? (2)How can we use the remaining inference rules to obtain a precise value of TREND?

Among all of these 980 rules, some of them are not valid in practical sense. For instance, the following rule obviously makes no sense:

IF D is very_high, K is very_high, K_D is positive and K_D_1 is positive, then Trend is high_inc. (2)

Such a rule must be eliminated. Besides, how can we determine the membership functionthat is appropriate? The training method of neural network can be used to solve both these problems, i.e., to refine the membership functions and to eliminate the irrelevant inference rules.

3.3 Going to a Neuro Fuzzy Formulation

Basically the idea of a neuro-fuzzy system is to find the parameters of a fuzzy system by means of learning methods obtained from neural networks. Many alternative ways of integrating neural nets and fuzzy logic have been proposed (Buckley and Hayashi 1994, Nauck and Kruse 1996, Lin and Lee 1996) which have much in common, but different in implementation aspects. The most common approach is to use so-called Fuzzy Associative Memories (FAMs). A FAM is a fuzzy logic rule with an associated weight. A mathematical framework exists that maps FAMs to neurons in a neural net.[4] This enables the use of a modified *error back propagation algorithm* with fuzzy logic.[5] This approach can help to generate and optimize membership functions and the associated weight of each rule from sample data. In our experiments, we implemented the FAM approach to construct the model.[6]

3.4 Benchmark Models

A linear regression model is based on constructing a linear relationship between dependent and independent variables. GARCH-M is also a linear model, but with the additional nonlinear consideration about the residual variance. Neural network is

---

[4] See for example, Eric B. Baum (1988), "Neural Nets for Economists" and the references therein as well as the references in von Altrock(1996).

[5] For more details on the mathematical foundation and relevant derivations, refer to (Kosko 1992)

[6] Please refer to Tong and Bonissone (1984), Zimmermann (1987), and Klir and Yuan (1995) for the details for the implementation of fuzzy logic, and refer to (Von Altrock 1997) for the details for the implementation of neuro fuzzy.

a model mainly to map the nonlinear relationship among the variables that allows for endogenous learning. Using the terminology introduced before, we can characterize a neuro fuzzy system as an expert system with different weights associated with each rule where the fuzziness of ordinary language is also modelled explicitly.

With the same input and output variables, these models are specified as follows. The linear regression:

$$\begin{aligned} Trend_{t+1} = {} & \beta_0 + \beta_1 \times K\_D + \beta_2 K\_D\_1 + \beta_3 K + \beta_4 D + \beta_5 KS + \beta_6 KT \\ & + \beta_7 PT + \beta_8 DT + \varepsilon_t \end{aligned} \tag{1}$$

where $\varepsilon_t \sim ID(0, \sigma_t^2)$, this is the benchmark model for this research. The buying signal is recognized when Trend > 0.

The GARCH-M (1,1):

$$\begin{aligned} Trend_{t+1} = {} & \beta_0 + \beta_1 \times K\_D + \beta_2 K\_D\_1 + \beta_3 K + \beta_4 D + \beta_5 KS + \beta_6 KT \\ & + \beta_7 PT + \beta_8 DT + \lambda h_t^{1/2} + \varepsilon_t \end{aligned} \tag{2}$$

where ε_t ~ N（0, h_t）and $h_t = \delta_0 + \delta_1 h_{t-1} + \delta_2 \varepsilon_{t-1}^2$.

The neural network model is:

$$Trend_{t+1} = F_2(w_2 F_1(w_1 x)) \tag{3}$$

where F_1 and F_2 are the transfer functions for hidden node and output node, respectively. The most popular choice for F_1 and F_2 are the sigmoid function,

$F(x) = \frac{1}{1+e^{-\alpha x}}$, representing the activation function adopted in the calculation process, w_1 and w_2 are the matrices of linking weights from input to hidden layer and from hidden to output layer, respectively, x is the vector of input variables, K, D, K_D, K_D_1, KS, KT, and DT. Basically a three-layer MLP is implemented in this paper with learning rate equal to 0.1 and momentum equal to 0.7. For details of this procedure and its mathematical background please refer to (Azoff 1994, Beltratti et al. 1996).

The purpose of neural network training is to estimate the weight matrices, w_1 and w_2, in equation (3) such that an overall error measure such as the mean squared errors (MSE) or sum of squared errors (SSE) is minimized. MSE can be defined as

$$MSE = \frac{1}{N}(a_j - TREND_j)^2 \tag{4}$$

where a_j and $TREND_j$ represent the target value and network output for the j th day respectively, and N is the number of days in training data set.

4. Empirical Results

4.1 Sample

The data set was obtained from Taiwan Economic Journal Data Bank (TEJ). Thirty world wide known stock indexes, as listed in table 1, are used for testing the predictive powers of alternative models. The data series include the period from 1992/1/1 to 2000/9/30. It is divided into two periods. The first period (in-sample data set) runs from 1992/1/1 to 1999/9/30 (1600 observations), while the second period (out-sample data set) runs from 1999/10/1 to 2000/9/30 (350 observations). The in-sample data set is used to determine the parameters of the models and the out–sample data set is used for model validation.

4.2 Transaction Simulation

Essentially, the investment simulation implemented here is based on the signal produced by these systems.It is asummed that if the predicted value, TREND, turned out to be higher than the threshold value, a portfolio of stocks interlocked with the stock index was purchased; if the predicted value was lower than the threshold value, the portfolio was sold. This would seem to reflect the rationality embodied in the profit making activities in the stock market adequately.

4.3 Performance Evaluation

In this paper we use RMSE, direction prediction, and rate of return to compare the performance of these different models. RMSE is calculated as follows.

$$\sqrt{\sum_{i=1}^{n}(a_i - TREND_i)^2 \Big/ n} \qquad (4)$$

where n is the number of days in testing data set.

If the predicted value is greater than the threshold value and the actual stock price movement for the next day is up too, then it is counted as *one time correct prediction.* Direction prediction percentage (hit rate) is calculated as follows:

$$\text{Hit\_rate} = \frac{h}{n}, \qquad (5)$$

where h is the number of correct prediction.

The rate of return for each strategy is calculated as follows:

$$R = ((1+r_1)\times(1+r_2)\times\cdots\times(1+r_n))^{\frac{1}{n}} \times n - 1 \qquad (6)$$

where r_i is the daily return for day i. The daily rate of return for cash on deposit is 0.05/250=0.0002 (i.e. 5% for yearly rate of return).

We do the unit root test before constructing the GARCH-M models. The unit root hypothesis is rejected at significance level 0.05 for all thirty series by the Augmented Dickey Fuller (ADF) test. In other words, these thirty series are all stationary. No differencing of any series is therefore necessary when fitting the model. The residuals of the GARH-M model are all white noise, as is readily found by checking the ACF and PACF.[7]

---

[7] On the other hand, many of the assumptions for the linear regression model(for example, the normality assumption, the constant variance assumption, and the assumption of non existence of autocorrelaton)

4.4 Empirical Results

In order to fix notation for facilitating the presentation, we use BH, KD, REG, GM, NN, and NF to denote buy and hold strategy, traditional KD trading system, regression model, Garch_M model, neural network model, and neuro fuzzy model respectively. Based on the predictions of each model on these thirty stock indexes, the paired test of the RMSE is listed in table 3. The number in the cell shows the difference of RMSE of the model at the row and the RMSE of the model at the column. For example, 0.0004 is the difference between GARCH-M and regression. The number in parenthesis is the p-value of the paired test. It can be seen that neuro fuzzy has the biggest RMSE among the methods.

In addition to the RMSE, we also show the direction prediction. Table 4 lists the basic statistics for the correct prediction percentage of all four models for thirty stock indexes. It can be seen that neuro fuzzy has 58.03% correct prediction for the next day's stock movement direction during the test period, neural network 55.77%, GARCH-M 52.83%, and regression 52.47%. Table 5 lists the paired test among these models. The number in the table represents the difference between the correct prediction percentage of model at the row position and the model at the

---

are not satisfied for most of the series. However, we still fit the regression models for each series as our benchmark.

column position. It can be seen that neuro fuzzy has the highest hit rate among all these models.

Based on the signals produced by each model, transactions are implemented for each model and the corresponding rate of return are calculated. Table 6 lists the basic statistics of the yearly rate of return of each model. Neuro fuzzy can achieve yearly rate of return as 27.17%, neural network 19.47%, GARCH-M 12.2%, regression 9.84%, traditional KD 9.56%, and buy and hold 9.35% respectively. Table 7 shows the paired test among these methods. It can be seen that the rate of return of neuro fuzzy is significantly greater than those of the other methods.

In addition to the statistical test of the yearly rate of return, Figure 3 also shows the cumulative wealth for each model for Landon FT 100 Index, which is typical for the other indexes. Neuro fuzzy is significantly the best one among these models.

It is interesting to note that neuro fuzzy is the most profitable model though its RMSE is not the least. This result is similar to Leung, Daouk, and Chen’s work (2000). It implies that the financial forecasters and traders could focus on accurately predicting the direction instead of minimizing the MSE or RMSE.

Since the threshold values of buying and selling signals cane influence the rate of return, we do the sensitivity analysis on the trading points as follows. We divide the signal range from 0 to 1 into 20 points with the interval equal to 0.05 as the alternative buying threshold values. Therefore we have 21 alternative threshold values, 0, 0.05, 0.10, 0.15,…, 1.00. And we do the same processing for the selling signal range from –1 to 0 (-1.0, -0.95, -0.90,…,0). Totally we have 441 (21*21=441) combinations. With the produced signal from the proposed model, we use each combination as the threshold values. For example, one alternative is (0.05,-0.15), which means that if the signal is greater than 0.05, then buy in the stocks. If the signal is less than –0.15, then sell out the stocks. Otherwise, do not do any transactions. Therefore, we can have a rate of return associated with each combination.

Since the profitability of a threshold value combination in the training data set does not promise the profitability in the testing data set, we need to show the robustness of the trading points. We calculate the rate of return for each threshold value combination on training data set and testing data set. The empirical results show that the average is close to 75 percent that if the threshold value combination is profitable in training data set, it will also be profitable in the testing data set. The detailed simulation result is shown in table 8. There are 8 out of 30 series with profitable percentage within 80% to 100%, 18 within 60% to 80%, 3 within 40% to

60%, and only 1 within 20% to 40%, showing the robustness of the proposed model for the parameters.

In addition to the robustness testing, a sensitivity analysis of the influence of transaction cost on the profitability of each model is also conducted. Transaction costs consist of commission fee 0.13% and trading tax 0.3% in Taiwan. The sensitivity analysis is run by setting trading tax equal to 0.0%, 0.1%, 0.3%, 0.5%, and 1.0% respectively. The simulation results are shown in Table 9. It can be seen that neuro fuzzy can consistently beat the traditional KD strategy and buy and hold strategy and the rate of return is decreasing only a little bit as the transaction costs increases.

Besides, the testing data set is arbitrarily divided into two different scenarios, bull market and bear market (or sluggish market) to see the influence of the scenario on the profitability performance. A bull market is defined as the period before the highest point of the testing period. A bear market or a sluggish market is defined as the period after the highest point during the testing period. Since some series have no turning points during the testing period, therefore the sample size for the different scenarios testing will be different. The basic statistics of the simulation results are shown in table 10. Table 11 and 12 show the paired test among the methods when the market is a bear market and a bull market respectively. It can be seen from table

11 that neuro fuzzy could significantly beat buy and hold strategy and traditional KD when the market is a bear market. Table 12 shows that neuro fuzzy is significantly better than the traditional KD but not significantly better than buy and hold when the market is a bull market.[8]

5. Summary and Conclusions

This paper uses 30 well known stock indexes to examine the linear and nonlinear predictability of stock market returns with KD technical trading rules. The empirical results show strong evidence of nonlinear relationships among the key variables in the stock market. Empirically, this is demonstrated most clearly by the nonlinear neuro fuzzy model that was used along with several others to predict returns by using KD technical indexes. The rate of return of the proposed neuro fuzzy system is significantly greater than that of the other methods. In addition to the robustness shown by the trading point analysis, the sensitivity analysis of transaction

[8] A possible explanation for this is that for the extreme case, that is the case with no turning points at all for the testing period, buy and hold will be the worst for the bear market and be the best for the bull market. However, when there are many turning points during the testing period, it is a different matter. Generally, the more turning points there are, the better the Neuro Fuzzy mode will be in prediction performance. For this paper, NF is better than buy and hold but not significantly so during the testing period.

costs also shows the profitability of the proposed system. Finally, the scenario analysis shows that though the rate of return of neuro fuzzy system is not significantly greater than that of buy and hold strategy when the market is a bull market, it is the best in a statistical sense when the market is a bear market or a sluggish market. This conclusion is important for both theory and strategy. Theoretically it shows that the efficient market hypothesis need not hold in the short run, but with learning the possibility of a convergence to the long run efficient market equilibrium can not be ruled out. Strategically, our approach shows that the neuro fuzzy model may allow investors to earn higher returns when there is a bear market which is far from the effient market equilibrium.

References

1. Al-Shammari, M. and Shaout, A., “Fuzzy logic modeling for performance appraisal systems, A framework for empirical evaluation”, Expert Systems with Applications, 14, (1998), 323-238.
2. Azoff, E. M., Neural network time series forecasting of financial markets, 1994, John Wiley & Sons Ltd.
3. Bailey, W., Stulz, R., and Yen, S., 1990, “Properties of daily stock returns from the Pacific Basin stock markets: Evidence and implications”, in: Rhee, S. and Chang, R., eds., Pacific-Basin capital markets research (North-Holland, Amsterdam).
4. Basu, S. , “The investment performance of common stocks in relation to their price -earnings ratios: A test of the efficient market hypothesis,” Journal of Finance 32,(1977) , 663-682.
5. Baum,Eric B. (1988), “Neural Nets for Economists”, in Anderson,P., Arrow, K. and Pines, D., eds. The Economy as an Evolving Complex System, New york: Addison-Wesley.
6. Beltratti, A., Margarita, S., and Terna, P., Neural networks for Economic and Financial Modelling, 1996, Internal Thomson Publishing Inc.
7. Bessembinder, H., and Chan K., “The profitability of technical trading rules in the Asian stock markets”, Pacific-Basin Finance Journal, 3, 1995, 257-284.
8. Bollerslev, T., “ Generalized Autoregressive Conditional Heteroskedasticity ,” Journal of Econometrics, 31(1986) , 307-327.
9. Bolloju, N. , “ Formulation of qualitative models using fuzzy logic,” Decision Support Systems, 17(1996) , 275-298.
10. Box, G. and Jenkins, G., Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco, CA, 1970.

11. Brock, W., Lakonishok, J., and LeBaron B .,"Simple Technical Trading Rules and the Stochastic Properties of Stock Returns," Journal of Finance , Vol.47(1992), 1731-1764.

12. Buckley, J. J. and Hayashi, Y. (1994) "Fuzzy Neural Networks: A Survey", Fuzzy Sets and Systems, 66, 1-13.

13. Campbell, J. (1987). "Stock returns and the term structure", Journal of Financial Economics 18, 373-399.

14. Dash, P.K., Liew, A. C., Rahman, S., and Dash, S.,"Fuzzy and Neuro-fuzzy Computing Models for Electric Load Forecasting", Engng Applic. Artif. Intell. Vol. 8, No. 4, 1995, 423-433.

15. Engle, R. F., "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," Econometrica, 50(1982), 987-1008.

16. Fama, E. F., "The Behavior of Stock Market Prices," Journal of Business , Vol.33, January(1965),34-105.

17. Fama, E. F. and Blume,M.E., "Filter Rules and Stock Market Trading," Journal of Business, Vol.39, January(1966), 226-241.

18. Fama, E., & French, K. "Permanent and temporary components of stock prices", Journal of Political Economy 96, 1988a, 246-273.

19. Fama, E., & French, K. "Dividend yields and expected stock returns", Journal of Financial Economics, 1988b, 22, 3-25.

20. Fama, E., and French,K., "The cross-section of expected stock returns," Journal of Finance ,vol.47(1992), 427-465.

21. Fama, E., & Schwert, W., "Asset returns and inflation", Journal of Financial Economics, 1977, 5, 115-146.

22. Jensen, M.C., "Random Walks : Reality or Myth—Comment," Financial Analysts Journal,vol.23(1967),77-85.

23. Jensen, M.C. and Benington, G.A. , "Random Walks and Technical Theories : Some Additional Evidence ," Journal of Finance , vol.25(1970),49-54.

24. Jocobs, B., and K. Levy. ,"The complexity of the Stock Market," The Journal of portfolio Management ,16 (1989),69-76.

25. Kaneko, Takaomi "Building a financial diagnosis system based on fuzzy logic production system," Computers ind.Engng, Vol.31,No.3/4(1996), 743-746.

26. Klir,G.J., and Yuan , B. , Fuzzy sets and fuzzy logic:Theory and Applications, Upper Saddle River ,NJ:Prentice Hall , 1995.

27. Kosko, B., Neural Networks and Fuzzy Systems. A Dynamical Systems Approach to Machine Intelligence. NJ :Prentice Hall, Englewood Cliffs, ,1992.

28. Lakonishok, J.,Shleifer, A.,and Viahmy , R. W. "Contrarian investment , extrapolation, and risk," Journal of finance, 49(1994),1541-1578.

29. Lane, G., " ",

30. Leung, M. T., Daouk, H., and Chen, A-S., "Forecasting stock indices: a comparison of classification and level estimation models", International Journal of Forecasting, 16, 2000, 173-190.

31. Lie, T. T. and Sharaf, A. M., "A novel neuro-fuzzy based self-correcting online electric load forecasting model", Electric Power Systems Research, 34, (1995), 121-125.

32. Lin, C.-T and Lee, C. G. (1996) Neural Fuzzy Systems. A Neuro-Fuzzy Synergism to Intelligent Systems. Prentice Hall, New York.

33. Nauck, D. and Kruse, R. (1996) "Neuro-Fuzzy Systems Research and Applications outside of Japan", Fuzzy-Neural Networks, Soft Computing series, pp. 108-134. Asakura Publ., Tokyo.

34. Padmakumari, K., Mohandas, K.P., and Thiruvengadam, S., "Long term distribution demand forecasting using neuro fuzzy computations", Electrical

Power and Energy Systems, 21, (1999), 315-322.

35. Pan, M., Chiou, J., Hocking, R. and Rim, H., 1991, “An examination of mean-reverting behavior of stock prices in Pacific-Basin stock markets”, in: Rhee, S. and Chang, R., eds., Pacific-Basin capital markets research, Vol. II (North-Holland, Amsterdam).

36. Studer, L. and Masulli, F., “Building a neuro-fuzzy system to efficiently forecast chaotic time series”, Nuclear Instruments and Methods in Physics Research, 389, (1997), 264-267.

37. Sweeny, R.J., “ Beating The Foreign Exchange Market, “ Journal of Finance ,Vol. 41(1986), 163-182.

38. Tong, R. M. and Bonissone, P.P.”Linguistic Solutions to Fuzzy Decision problems,” in Zimmermann, H.-J., Zadeh, L. A., and Gaines, B. R. (Eds.), Fuzzy Sets and Decision Analysis, Amsterdam (1984), p. 323-334.

39. von Altrock , C., Fuzzy Logic & NeuroFuzzy Applications In Business & Finance, Upper Saddle River, N.J: Prentice Hall PTR. 1997.

40. Zimmermann, H.-J., Fuzzy Sets, Decision Making, and Expert Systems, Boston: Kluver Academic Publisher (1987). ISBN 089838-149-5.

41. Zimmerman, H.-J. and Thole, U. “On the Suitability of Minimum and Product Operators for the Intersection of Fuzzy Sets,” Fuzzy Sets and Systems, Vol. 2 (1978) , 173-186.

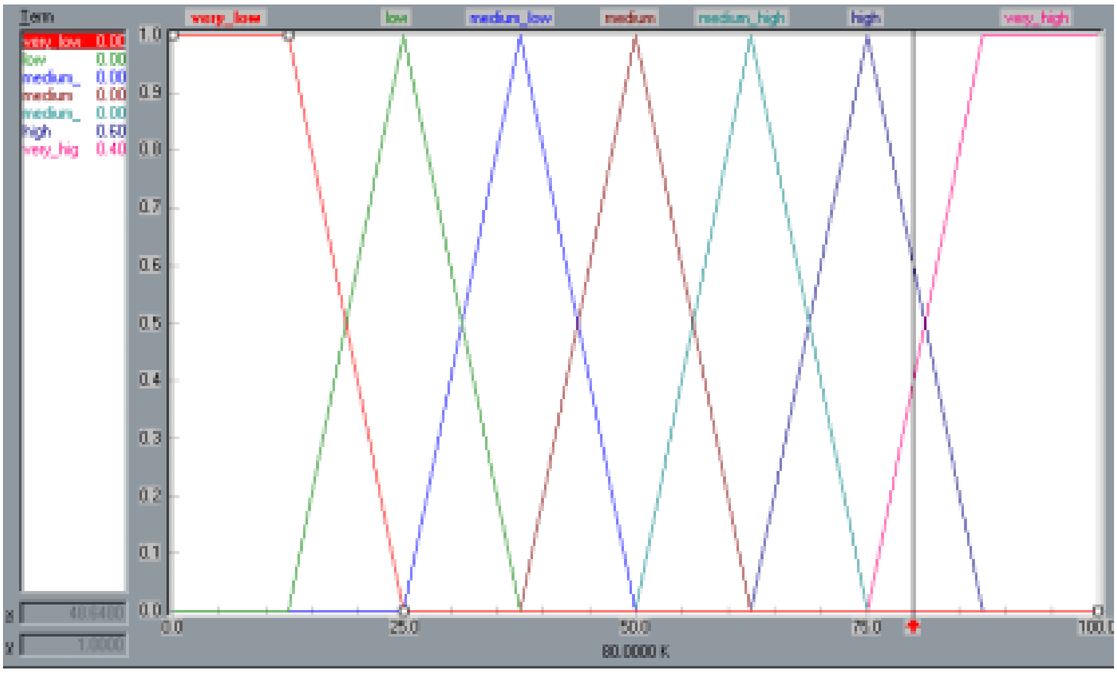

Figure 1(a). Membership function of "K"

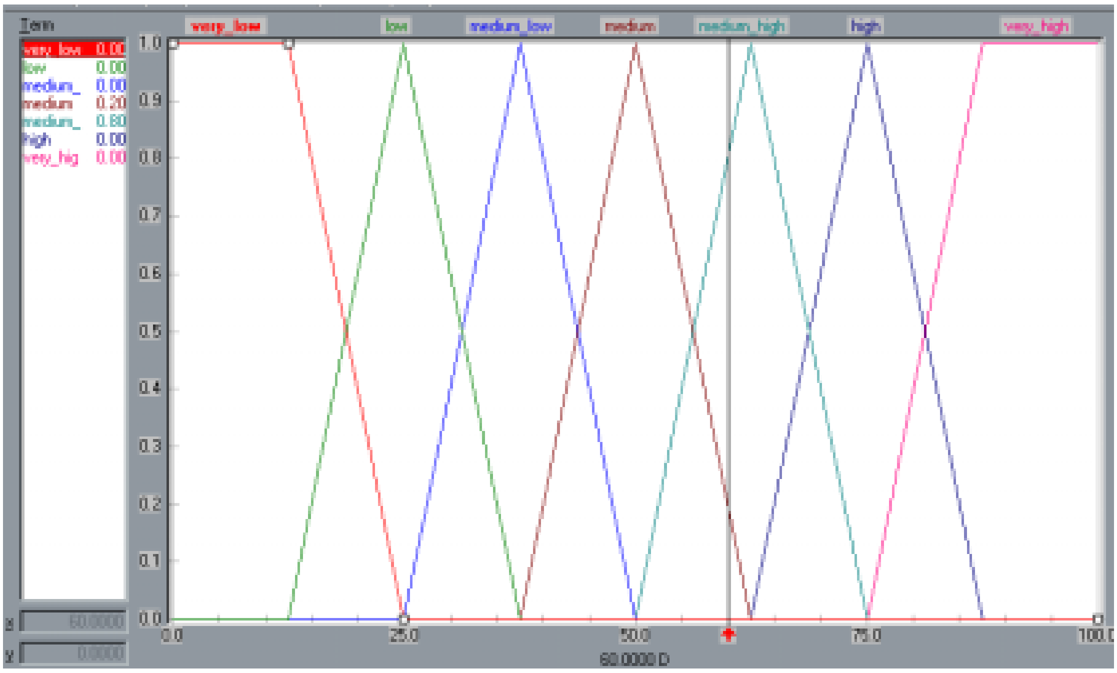

Figure 1(b). Membership function of "D"

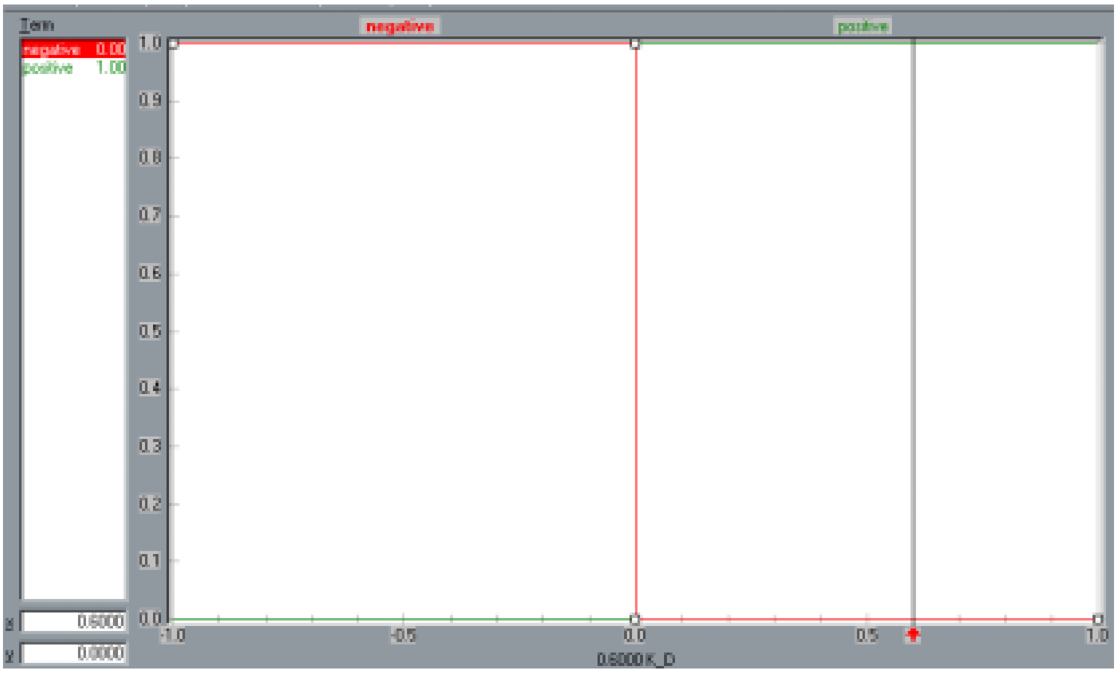

Figure 1(c). Membership function of “K_D”

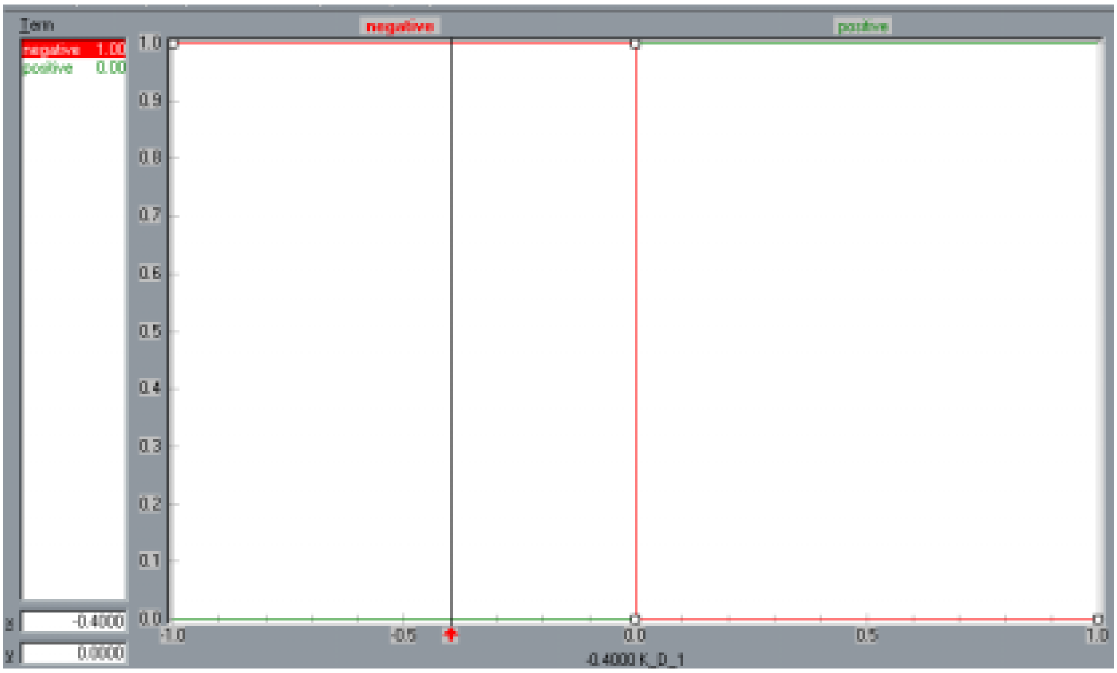

Figure 1(d). Membership function of “K_D_1”

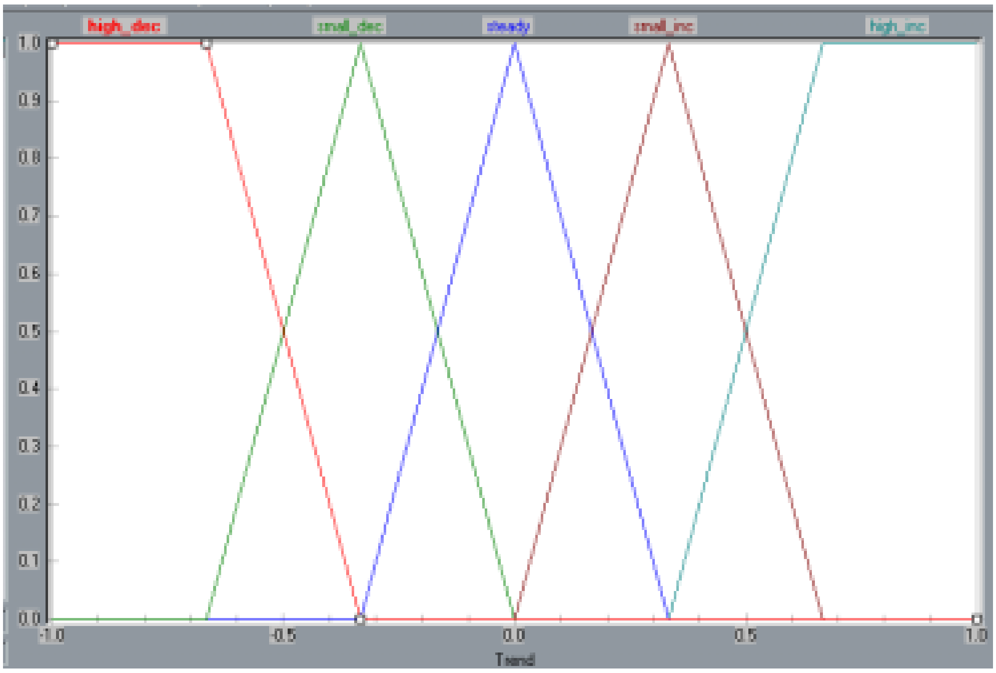

Figure 1(e). Membership function of "Trend"

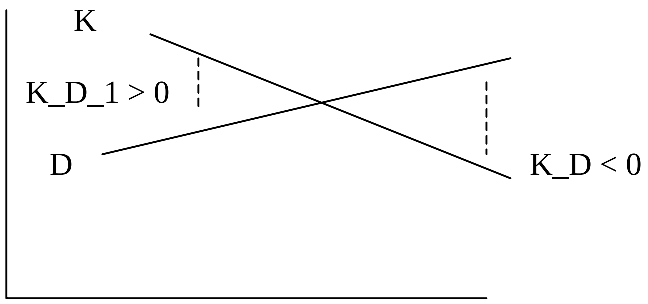

Figure 2. The cross over phenomenon

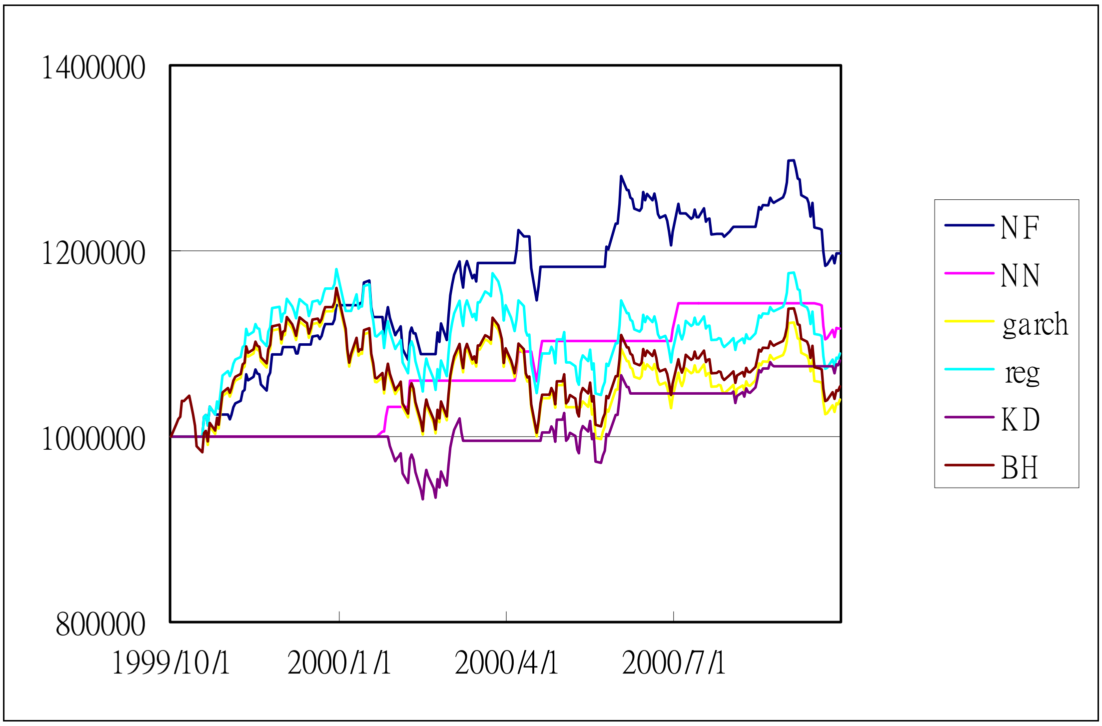

Figure 3. The equity curve for each model for Landon FT100 Index

Table 1. Properties of linguistic variables and their terms

Variable Name	Type	Minimum	Maximum	Term Names
K	Input	0	100	very_low, low, medium_low, medium, medium_high, high, very_high
D	Input	0	100	very_low, low, medium_low, medium, medium_high, high, very_high
K_D	Input	-1	1	negative, positive
K_D_1	Input	-1	1	negative, positive
Trend	Output	-1	1	high_dec, small_dec, steady, small_inc, high_inc

Table 2. The corresponding validity extent of each term for rule (1).

Variable	Values	Membership function	Validity
K	80	K is high	0.6
D	60	D is medium	0.2
K_D	0.6	K_D is positive	1.0
K_D_1	-0.4	K_D_1 is negative	1.0

Table 3. Paired test of RMSE

	REG	GM	NN	NF
REG				
GM	0.0004 (0.201)			
NN	-0.017 (0.049) *	-0.021 (0.006) *		
NF	0.036 (0.002) *	0.032 (0.001) *	0.0526 (0.000) *	

（*: α ＝0.05）

Table 4. Basic statistics for the correct prediction percentage.

	N	Minimum	Maximum	Average	Standard Deviation
REG	30	.45	.58	.5247	0.024
GM	30	.48	.58	.5283	0.029
NN	30	.49	.64	.5577	0.0035
NF	30	.52	.66	.5803	0.0038

Table 5. Paired test for the correct prediction percentage.

	REG	GM	NN	NF
REG				
GM	0.004 (0.423)			
NN	0.037 (0.000) *	0.029 (0.001) *		
NF	0.056 (0.000) *	0.052 (0.000) *	0.023 (0.000) *	

（*: α ＝0.05）

Table 6. Basic statistics of the yearly rate of return

	N	Minimum	Maximum	Average	Standard Deviation
BH	30	-0.32	0.50	0.093	0.239
KD	30	-0.27	0.38	0.096	0.152
REG	30	-0.31	0.49	0.098	0.174
GM	30	-0.22	0.50	0.122	0.165
NN	30	-0.05	0.53	0.194	0.159
NF	30	-0.05	0.73	0.271	0.200

Table 7. Paired tests of the yearly rate of return

	BH	KD	REG	GM	NN	NF
BH						
KD	0.002 （0.935）					
REG	0.004 （0.868）	0.002 （0.905）				
GM	0.028 （0.409）	0.026 （0.365）	0.023 （0.096）			
NN	0.101 (0.001) *	0.099 （0.00） *	0.096 （0.00） *	0.072 （0.00） *		
NF	0.1782 （0.00） *	0.1761 （0.00） *	0.1733 （0.00） *	0.1498 （0.00） *	0.077 (0.001) *	

（*: α ＝0.05）

Table 8 The percentage of the trading point combinations that is also profitable in the testing data

Percentage	No. of series
80%~100%	8
60%~80%	18
40%~60%	3
20%~40%	1

Table 9. Rate of return under different transaction costs

	Investment strategy		
Transaction cost	NF	BH	KD
0.00	27.17	9.56	9.35
0.10	24.87	8.33	9.35
0.30	22.65	8.05	9.32
0.50	20.53	7.22	9.32
1.00	18.28	5.98	9.31

Table 10. Basic statistics of the rate of return under different market scenario

Bear market

	N	Minimum	Maximum	Average	Standard Deviation
BH	30	-0.40	0.27	-0.170	0.174
KD	30	-0.30	0.12	-0.073	0.107
NF	30	-0.18	0.21	-0.002	0.088

Bull market

	N	Minimum	Maximum	Average	Standard Deviation
BH	30	0.15	0.73	0.329	0.149
KD	30	-0.01	0.38	0.156	0.087
NF	30	0.09	0.86	0.340	0.189

Table 11. Paired test of the yearly rate of return when the market is a bear market

	Paired difference		t	Degree of freedom	significance (two-tailed)
	mean	Standard deviation			
NF - KD	0.0710	0.1008	3.307	21	0.003*
NF - BH	0.1673	0.1240	6.329	21	0.000*
KD - BH	0.0962	0.1300	3.474	21	0.002*

（*: α ＝0.05）

Table 12. Paired test of the yearly rate of return when the market is a bull market

	Paired difference		t	Degree of freedom	significance (two-tailed)
	mean	Standard deviaiton			
NF - KD	0.1848	0.1551	5.955	24	0.000*
NF - BH	0.0111	0.0959	0.582	24	0.566
KD - BH	-0.1736	0.1205	-7.202	24	0.000*

（*: α =0.05）