The Korean Communications
in Statistics Vol. 10 No. 1, 2003
pp. 153-166

Using Evolutionary Optimization to Support Artificial Neural Networks for Time-Divided Forecasting: Application to Korea Stock Price Index[1)]

Kyong Joo Oh[2)]

Abstract

This study presents the time-divided forecasting model to integrate evolutionary optimization algorithm and change point detection based on artificial neural networks (ANN) for the prediction of (Korea) stock price index. The genetic algorithm(GA) is introduced as an evolutionary optimization method in this study. The basic concept of the proposed model is to obtain intervals divided by change points, to identify them as optimal or near-optimal change point groups, and to use them in the forecasting of the stock price index. The proposed model consists of three phases. The first phase detects successive change points. The second phase detects the change-point groups with the GA. Finally, the third phase forecasts the output with ANN using the GA. This study examines the predictability of the proposed model for the prediction of stock price index.

Keywords: Genetic Algorithms; Artificial Neural Networks; Change-Point Detection; Stock Price Index

1. Introduction

Macroeconomic time series data has a series of change points since they are controlled by government's monetary policy (Mishkin, 1995; Oh and Han, 2000). The government takes intentional action to control the currency flow that has direct influence upon fundamental economic indices. For the stock price index, institutional investors play a very important role in determining the direction of its change since they are major investors in terms of marking and volume for trading stocks. They respond sensitively to such economic indices like interest rates, consumer price index, anticipated inflation, etc. Therefore, we can conjecture that the

---

1) This research was financially supported by Hansung University in the year of 2003.

2) Fulltime Lecturer, Department of Business Administration, Hansung University, Seoul 136-792, Korea.
Email: johanoh@hansung.ac.kr

movement of the stock price index also has a series of change points. To reflect these characteristics in stock price index, this study presents a genetic algorithm (GA) approach to change-point detection in artificial neural networks (ANN) for the prediction of stock price index.

A large number of studies on stock market prediction using artificial intelligence (AI) techniques have been performed during the past decades. However, the networks fall into a naive solution such as always predicting the most common output since the training process for ANN tends to be difficult with highly noisy data. Most ANN used in previous research employed the gradient descent algorithm to train the network. The well-known limitations of gradient search techniques applied to ANN have often resulted in inconsistent and unpredictable outputs. Some studies proposed ANN using global search algorithms as a method to solve these problems (Sexton, Alidaee, Dorsey and Johnson, 1998a; Sexton, Dorsey and Johnson, 1998b). This study employs the GA which is a global search algorithm to overcome the limitation of a local search algorithm in ANN.

The basic concept of the proposed model is to obtain intervals divided by change points, to identify them as optimal or near-optimal change-point groups, and to use them in the forecasting of the stock price index. The proposed model consists of three phases. The first phase detects successive change points. The second phase detects the change-point groups with the GA. Finally, the third phase forecasts the output with ANN using the GA. This study then examines the predictability of the proposed models for the prediction of stock price index. To explore the predictability, we divided the stock price index data into two sets -the training data over one period and the testing data over the other period. The predictability of stock price index is examined using the metrics of the root mean squared error (RMSE), the mean absolute error (MAE) and the mean absolute percentage error (MAPE).

The rest of the paper is organized as follows. In Section 2, we outline the research motivation and background. Section 3 describes the proposed integrated neural network model using the GA-based change-point groups. Section 4 describes the research design and experiments. In Section 5, the empirical results of the case study are summarized and discussed. Finally, the concluding remarks are presented in Section 6.

2. Research Background

2.1 Applications of change-point detection

There has been much research interest in the statistical analysis of change-point detection and estimation since it can be encountered in many disciplines such as economics, finance, medicine, psychology, geology, literature, etc., and even in our daily lives. In these applications, the problem is known under different headings, such as quality control, failure and shock detection. These change-point detections have brought to various change-point models, which are classified as likelihood ratio tests, nonparametric approaches, and linear model approaches. Csorgo and Horvath (1997) provide a concise overview and rigorous mathematical treatment of methods for change-point detection and use a number of datasets to illustrate the effectiveness of each technique.

According to the classification by Csorgo and Horvath (1997), three major change-point detection methods are developed: Likelihood ratio tests (Chernoff and Zacks, 1964; Kander and Zacks, 1966; Gardner, 1969; Sen and Srivastava, 1975; Worsley, 1979; Yao, 1988), Nonparametric approaches (Brodsky and Darkhovsky, 1993; Page, 1955; Parzen 1994; Pettitt, 1979, 1980a), and Model-based approaches (Quandt, 1960; Andrews, 1991; Chu and White, 1992; Siegmund, 1985; Hinkley, 1971; Chow, 1960; Hawkins, 1989). Among these change-point detection methods, Oh and Han (2000) recommended the Pettitt test (Pettitt, 1979, 1980a) as a method to detect the change points in the chaotic time series data since it is free from model assumptions and may provide stable results.

2.2 The Pettitt tests

Consider a sequence of random variables $X_1, X_2, \ldots, X_T$, then the sequence is said to have a change-point at τ if X_t for $t=1, 2, \ldots, \tau$ have a common distribution $F_1(x)$ and X_t for $t=\tau+1, \tau+2, \ldots, T$ have a common distribution $F_2(x)$, and $F_1(x) \neq F_2(x)$. We consider the problem of testing the null hypothesis of *no-change*, $H_0: \tau = T$, against the alternative hypothesis of *change*, $H_A: 1 \leq \tau < T$, using a nonparametric statistic.

An appealing non-parametric test to detect a change would be to use a version of the Mann-Whitney two-sample test. A Mann-Whitney type statistic has remarkably stable distribution and provides a robust test of the change point resistant to outliers (Pettitt, 1980b). Let

$$D_{ij} = sgn(X_i - X_j) \tag{1}$$

where $sgn(x) = 1$ if $x > 0$, 0 if $x = 0$, -1 if $x < 0$, then consider

$$U_{t,T} = \sum_{i=1}^{t} \sum_{j=t+1}^{T} D_{ij}. \tag{2}$$

The statistic $U_{t,T}$ is equivalent to a Mann-Whitney statistic for testing that the two samples $X_1, X_2, \ldots, X_t$ and $X_{t+1}, X_{t+2}, \ldots, X_T$ come from the same population. The statistic $U_{t,T}$ is then considered for values of t with $1 \leq \tau < T$. For the test of H_0: *no change* against H_A: *change*, we propose the use of the statistic

$$K_T = \max_{1 \leq t < T} |U_{t,T}|. \tag{3}$$

The limiting distribution of K_T is $\Pr \cong 2\exp\{-6k^2/(T^2+T^3)\}$ for $T \to \infty$.

The Pettitt test detects a possible change point in the time sequence dataset. Once the change point is detected through the test, the dataset is divided into two intervals. The intervals before and after the change point become basic elements of homogeneous groups which take heterogeneous characteristics from each other, which is explained in Section 3.

There are few artificial intelligence models which consider the change-point detection problems. Most of the previous research has a focus on finding unknown change points for the past, not the forecast for the future (Wolkenhauer and Edmunds, 1997; Li and Yu, 1999). Gorr (1994) and White (1994) have provided the forecasting model to use change-point detection. Oh and Han (2000) proposed the integrated neural network model using the Pettitt test. They obtained intervals divided by change points in the training phase, identified them as change-point groups in the training phase, and forecasted to which group each sample would be assigned in the testing phase. They, however, did not suggest the general method of determining the number of groups. In the step of obtaining significant groups by change points, they determined the number of group through an empirical study.

This study proposes change-point detection supported by the GA in ANN for the prediction of the stock price index. This model detects (near-)optimal change-point groups using the GA. The GA allocates each significant interval to appropriate change-point groups in association with the forecasting model through a specific fitness function.

2.3. The genetic algorithms

The GA is a search algorithm based on survival of the fittest among string structures (Goldberg, 1989). This notion stems from the fact that the members of a species that are best fitted to the environment are likely to survive and produce offspring. The GA has been investigated recently and shown to be effective in exploring a complex space in an adaptive way, guided by the biological evolution mechanisms of reproduction, crossover, and mutation (Adeli and Hung, 1995).

The first step of the GA is problem representation. The problems must be represented in a suitable form to be handled by the GA. The GA often works with a form of binary coding. If the problems are coded as chromosomes, the populations are initialized. Each chromosome within the population gradually evolves through biological operations. Larger populations ensure greater diversity but require more computational burden. Once the population size is chosen, the initial population is randomly generated (Bauer, 1994). After the initialization step, each chromosome is evaluated by the fitness function. According to the value of the fitness function, the chromosomes associated with the fittest individuals will be reproduced more often than those associated unfit individuals (Davis, 1994). Crossover allows the search to fan out in diverse directions looking for attractive solutions and permits chromosomal material from different parents to be combined in a single child. In addition, mutation arbitrarily alters one or more components of a selected chromosome. It provides the means for introducing new information into the population. Finally, the GA tends to converge on (near-)optimal solutions (Wong and Tan, 1994).

The GA is usually employed to improve the performance of AI techniques. For ANN, the GA is popularly used to select neural network topology such as input variable selection, determining the optimal number of hidden layers and processing elements. In addition, many studies used the GA to search the weight vector of the network instead of local search algorithms including the gradient descent algorithm (Sexton et al., 1998a; 1998b).

3. Description of the Forecasting Model

In this section, we discuss the architecture and the characteristics of our model. Figure 1 shows the architecture of our model. In Figure 1, I means *the significant interval* made by the Pettitt test. Based on the Pettitt test, the proposed model consists of three phases: The first phase is the change-point detection phase, the second phase is the change-point group detection phase, and the final phase is the forecasting phase.

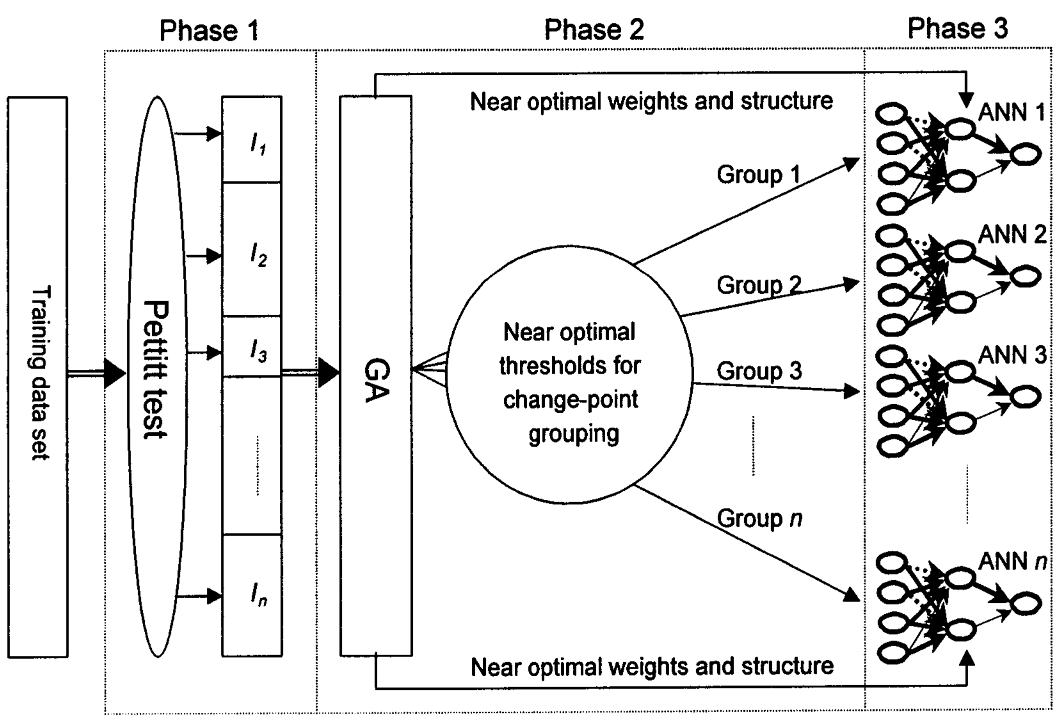

Fig. 1. Architecture of the proposed model

Phase 1. In the first phase, we apply the Pettitt test to the stock price index at time t in the training data. The Pettitt test mentioned in Section 2 is the method for finding just one change point in time series data. Based on this method, multiple change points can be obtained under the binary segmentation method which is explained as follows:

Step 1: Find a change point in X_1, K, X_T by the Pettitt test. If $\hat{\tau}(1)$ is a change point,

$X_1, \mathrm{K}, X_{\hat{\tau}(1)}$ are regarded as I_1 and $X_{\hat{\tau}(1)+1}, \mathrm{K}, X_T$ are regarded as I_2. Otherwise, it is concluded that there does not exist a change point for X_1, K, X_T. $(X_1 \le X_{\hat{\tau}(1)} \le X_T)$

Step 2: Find a change point in $X_1, \mathrm{K}, X_{\hat{\tau}(1)}$ by the Pettitt test. If $\hat{\tau}(2)$ is a change point, $X_1, \mathrm{K}, X_{\hat{\tau}(2)}$ are regarded as I_{11} and $X_{\hat{\tau}(2)+1}, \mathrm{K}, X_{\hat{\tau}(1)}$ are regarded as I_{12}. Otherwise, $X_1, \mathrm{K}, X_{\hat{\tau}(1)}$ are regarded as I_1 like Step 1. $(X_1 \le X_{\hat{\tau}(2)} \le X_{\hat{\tau}(1)})$ Find a change point in $X_{\hat{\tau}(1)+1}, \mathrm{K}, X_T$ by the Pettitt test. If $\hat{\tau}(3)$ is a change point, $X_{\hat{\tau}(1)+1}, \mathrm{K}, X_{\hat{\tau}(3)}$ are regarded as I_{21} and $X_{\hat{\tau}(3)+1}, \mathrm{K}, X_T$ are regarded as I_{22}. Otherwise, $X_{\hat{\tau}(1)+1}, \mathrm{K}, X_T$ are regarded as I_2 like Step 1. $(X_{\hat{\tau}(1)+1} \le X_{\hat{\tau}(3)} \le X_T)$

Step 3: By applying the same procedure of Step 1 and 2 to subsamples, we can obtain several significant intervals under the dichotomy.

After the significant intervals are detected for the whole dataset, each observation is replaced by *the interval mean*, the arithmetic mean of observations which belong to each significant interval. This removes noises from observations, and then makes observations homogeneous within each interval.

Phase 2. In this phase, the GA determines thresholds by the interval means obtained in Phase 1. Then, the whole significant intervals are assigned to some homogeneous groups by these thresholds. This is a process to establish some homogeneous groups into which the whole significant intervals are rearranged. Such groups represent a set of meaningful trends encompassing stock price index. As those trends help to find regularity among the related output values more clearly, the neural network model using the GA has a better ability of generalization for the unknown data. In general, forecasting error is reduced by making the subsampling units within groups homogeneous and the variation between groups heterogeneous (Cochran, 1977). In addition, this phase uses the GA to determine an optimal or near optimal number of change-point groups. The GA detects change-point groups to optimize the fitness function in the forecasting model.

Phase 3. The third phase is built by applying ANN to the data for each change-point group. Once ANN is built, then the sample is used to forecast the stock price index. This study uses the GA to determine the connection weights between layers in ANN. In addition, the GA determines the structure of ANN. That is, the GA selects relevant input variables and (near-)optimal processing elements in the hidden layer for each ANN. The optimization processes of the GA in Phase 2 and 3 are performed simultaneously.

4. Data and Experiments

Research data used in this study comes from the weekly Korea Stock Price Index (KOSPI) from January 1987 to August 1996. The total number of samples includes 502 trading weeks. The training data includes observations from January 1987 to December 1992 and the holdout data includes observations from January 1993 to August 1996. About 38% of the data is used for holdout and 62% for training. The training data is used to search the (near-)optimal parameters and is also employed to evaluate the fitness function. The holdout data is used to evaluate the performance for the unknown data. Initial variables are 13 technical indicators. These variables are selected based on the review of domain experts. The description of input variables is presented in Table 1, which is attached in Apendix (Achelis, 1995; Chang, Jung, Yeon, Jun, Shin and Kim, 1996; Choi, 1995; Edwards and Magee, 1997, Gifford, 1995).

This study employs three neural network models. One model, labeled Simple_GANN, uses the GA to select relevant input variables and to determine (near-)optimal number of processing elements in the hidden layer. In addition, the GA also determines the connection weights between layers. But, this model uses only one ANN since it does not have the change-point detection phase. Simple_GANN is recommended by many researchers including Maniezzo (1994). The second model, labeled BPN_GANN, uses backpropagation neural networks (BPN) to predict the change-point group for each given sample. This method is employedby Oh and Han (2000). BPN_GANN trains ANN using the GA for the final output forecasting. Thus, the GA searches (near-)optimal connection weights between layers and network structure. The third model or the proposed model, labeled GA_GANN, employs the GA to determine change-point group detection for each case in addition to network structure optimization and connection weight determination.

This study describes the optimization process of GA_GANN among the above three models. The optimization processes of the other two models are alike GA_GANN except for some factors. As mentioned earlier, the problems must be coded as chromosomes for the experiment. The strings used in this study have the following encoding. This study uses 7 bits to represent the codes for change-point group detection. The values of these bits arethresholds for change-point group detection. Each case is clustered by these bits and is allocated to each ANN model. The possible maximum number of clusters is 8. In addition, each ANN model has 208 bits. The first 169 bits represent the connection weights between the input layer and the hidden layer because each processing element in the hidden layer receives 13 signals from each input variable. The model uses 13 initial input variables and 13 initial processing elements in the hidden layer. These bits are searched from 5 to 5. The following 13 bits indicate the connection weights between the hidden and the output layer because each processing element in the output layer receives a signal from the hidden processing element. These bits also varied between 5 and 5. The next 13 bits represent selection codes for input variables and the following 13 bits are selection codes for processing elements in the hidden layer. These 26 bits are searched between 0 and 1.

The encoded chromosomes are searched to optimize the fitness function. In this study, the fitness function is the MAPE (mean absolute percent error) of the training data. This metric is chosen since it is commonly used (Carbone and Armstrong, 1982) and is highly robust (Armstrong and Collopy, 1992; Makridakis, 1993). The fitness function is as follows:

$$Fitness = \frac{1}{T}\sum_{i=1}^{T}\frac{\left|\hat{X}_i - X_i\right|}{X_i} \tag{4}$$

where $\hat{X}_i$ means the predicted value of observation, X_i for $i = 1, \ldots, T$. The parameters to be searched use only the information about the training data. The GA operates the process of crossover and mutation on initial chromosomes and iterates until the stopping conditions are satisfied. For the control parameters of the GA search, the population size is set at 1000 organisms and the crossover and mutation rates are varied to prevent ANN from falling into a local minimum. The range of the crossover rate is set between 0.5 and 0.7 while the mutation rate ranges from 0.05 to 0.1. The stopping condition of the search is 20,000 trials.

5. Empirical Results

Based on the phase 1 of section 3, the Pettitt test is applied to the stock price index data with 5% significant level. 62 significant intervals are established. For Simple_GANN, the GA selects 8 input variables and 5 processing elements in the hidden layer. BPN_GANN has two ANN models according to the change-point group from BPN. In the first ANN model of BPN_GANN, 6 input variables and 4 processing elements are selected in the hidden layer. For the second ANN model, 5 input variables and 6 processing elements in the hidden layer are recommended by the GA. In addition, GA_GANN has two ANN models according to the change-point group by the GA. In the first ANN model of GA_GANN, 6 input variables and 5 processing elements in the hidden layer are selected. For the second ANN model, 5 input variables and 4 processing elements in the hidden layer are recommended. A series of experiments is tried to obtain the stable results between the training data and the holdout data. Table 2 shows the selected input variables for each model through the evolutionary process of the GA.

Table 3 describes experimental results. In case of MAPE, the outcomes indicate that GA_GANN has higher prediction accuracy than Simple_GANN by 9.23% for the holdout data. In addition, GA_GANN outperforms BPN_GANN by 6.04% for the holdout data. According to RMSE (root mean squared error) and MAE (mean absolute error), GA_GANN outperforms the other two models. From Table 3, we also find that BPN_GANN has lower MAPE than Simple_GANN by 3.18% for the holdout data.

Table 2. Selected input variables for each model

	Simple_GANN	BPN_GANN		GA_GANN	
		ANN1	ANN2	ANN1	ANN2
MA6	*	*	*	*	
RSI	*	*		*	
OBV	*	*	*	*	*
%K					
%D	*				*
Momentum					
Psychology			*		
Disparity6	*		*		
Disparity25	*			*	*
ROC6	*	*		*	
VR		*			
MA25	*		*	*	*
ROC25		*			*

Table 3. Performance results based on RMSE, MAE and MAPE for the holdout data

Model	RMSE	MAE	MAPE (%)
Simple_GANN	165.29	140.79	15.73
BPN_GANN	125.71	104.38	12.55
GA_GANN	113.66	54.81	6.50

We use the pairwise t-test to examine whether the differences exist in the predicted values of models according to the absolute percentage error (APE). Since the forecasts are not statistically independent and not always normally distributed, we compare the APEs of forecast using the pairwise t-test. Where sample sizes are reasonably large, this test is robust to the distribution of the data, to nonhomogeneity of variances, and to statistical dependence (Iman and Conover, 1983). Table 4 shows t-values and p-values. GA_GANN is significantly better than Simple_GANN and BPN_GANN at a 1% significant level. Thus, the proposed model is demonstrated to obtain improved performance using the proposed approach. In addition, BPN_GANN significantly outperforms Simple_GANN at a 1% significant level.

Table 4. Pairwise t-tests for the difference in residuals between the three models for the holdout data based on the APE with the significance level in parentheses.

Model	BPN_GANN	GA_GANN
Simple_GANN	2.93 (0.0037)*	6.80 (0.0000)*
BPN_GANN		5.06 (0.0000)*

*Significant at 1%

In summary, change-point detection supported by the GA in ANN turns out to have a high potential in stock market prediction. This may be attributable to the fact that it categorizes stock market data into homogeneous groups and extracts regularities from each homogeneous group.

6. Concluding Remarks

As mentioned above, Oh and Han (2000) suggested change-point detection to support ANN in interest rate forecasting. However, they did not suggest a general method of determining the number of groups in the step of obtaining significant groups by change points. This study proposes a GA approach to determine the number of change-point groups for multiple ANN models. In this study, the GA searches for an optimal or near-optimal number of groups and simultaneously decides the structure of the ANN model. GA_GANN significantly outperforms Simple_GANN and BPN_GANN. These experimental results imply the GA-based change-point group detection has a high potential for improving the prediction performance. In addition, BPN_GANN significantly outperforms Simple_GANN. This shows the effectiveness of the inclusion of change-point group detection for the prediction of stock price index.

The proposed model shows promise in improving the performance of stock price prediction if further studies focus on the various approaches in the construction of change-point groups. Yao (1988) used a maximum likelihood argument to estimate the number of changes in normal observations. In general, the GA has the limitation of not explaining the background of the optimal model clearly. In this sense, Yao's study can be a guideline for improving the performance of the proposed model even though the GA detects the optimal number of groups. In addition, it is recommended that a comparative study in the GA and other heuristic approaches including simulated annealing and tabu search are performed to obtain the optimal decision of change-point groups, and the proposed model may be applied to other chaotic time series data with nonlinear characteristics.

References

[1] Achelis, S.B. (1995). *Technical Analysis from A to Z*, Probus Publishing, Chicago.

[2] Adeli, H. and Hung, S. (1995). *Machine Learning: Neural Networks, Genetic Algorithms, and Fuzzy Systems,* John Wiley & Sons, New York.

[3] Andrews, D. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica,* 59, 817-858.

[4] Armstrong, J.S. and Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons, *International Journal of Forecasting*, 8, 69-80.

[5] Bauer, R.J. (1994). *Genetic algorithms and investment strategies*, John Wiley & Sons, New York.

[6] Brodsky, B.E. and Darkhovsky, B.S. (1993). *Nonparametric Methods in Change-point Problems*, Kluwer, Dordrecht.

[7] Carbone, R. and Armstrong, J.S. (1982). Evaluation of extrapolative forecasting methods: Results of academicians and practitioners, *Journal of Forecasting*, 1, 215-217.

[8] Chang, J., Jung, Y., Yeon, K., Jun, J., Shin, D. and Kim, H. (1996). *Technical Indicators and Analysis Methods*, Jinritamgu Publishing, Seoul.

[9] Chernoff, H. and Zacks, S. (1964). Estimating the current mean of a normal distribution which is subjected to changes in time, *Annals of Mathematical Statistics*, 35, 999-1018.

[10] Choi, J. (1995). *Technical Indicators*, Jinritamgu Publishing, Seoul.

[11] Chow, G.C. (1960). Tests of equality between sets of coefficients in two linear regressions, *Economerica*, 28, 591-605.

[12] Chu, C.S.J. and White, H. (1992). A direct test for changing trend, *Journal of Business Economic Statistics*, 10, 289-299.

[13] Cochran, W.G. (1977). *Sampling Techniques*, John Wiley & Sons, New York.

[14] Csorgo, M. and Horvath, L. (1997). *Limit Theorems in Change-Point Analysis*, John Wiley & Sons, New York.

[15] Davis, L. (1994). Genetic algorithms and financial applications. In: Deboeck, G.J. (Ed.), *Trading on the Edge*, John Wiley & Sons, New York, pp. 133-147.

[16] Edwards, R.D. and Magee, J. (1997). *Technical Analysis of Stock Trends*, John Magee, Chicago.

[17] Gardner, L.A. Jr. (1969). On detecting changes in the mean of normal variates, *Annals of Mathematical Statistics*, 40, 116-126.

[18] Gifford, E. (1995). *Investor's Guide to Technical Analysis: Predicting Price Action in the Markets*, Pitman Publishing, London.

[19] Goldberg, D.E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley Publishing, Massachusetts.

[20] Gorr, W. (1994). Research respective on neural networks, *International Journal of*

Forecasting, 10(1), 1-4.

[21] Hawkins, D.L. (1989). An U-I approach to retrospective testing for shift parameters in a linear model, *Communications in Statististics*, 18, 3117-3134.

[22] Hinkley, D.V. (1971). Inference in two-phase regression, *Journal of American Statistical Associations*, 66, 736-743.

[23] Iman, R. and Conover, W.J. (1983). *Modern Business Statistics*, Wiley, New York.

[24] Kander, Z. and Zacks, S. (1966). Test procedures for possible changes in parameters of statistical distributions occurring at unknown time points, *Annals of Mathematical Statistics*, 37, 1196-1210.

[25] Li, H.L. and Yu, J.R. (1999). A piecewise regression analysis with automatic change-point detection. *Intelligent Data Analysis*, 3, 75-85.

[26] Makridakis, S. (1993). Accuracy measures: Theoretical and practical concerns. International, *Journal of Forecasting*, 9, 527-529.

[27] Maniezzo, V. (1994). Genetic evolution of the topology and weight distribution of neural networks, *IEEE Transactions on Neural Networks*, 5 (1) 39-53.

[28] Mishkin, F.S. (1995). *The Economics of Money, Banking, and Financial Markets*, Harper Collins, New York.

[29] Oh, K.J. and Han, I. (2000). Using change-point detection to support artificial neural networks for interest rates forecasting, *Expert Systems with Applications*, 19(2) 105-115.

[30] Page, E.S. (1955). A test for a change in a parameter occurring at an unknown point , *Biometrika*, 42, 523-526.

[31]Parzen, E. (1994). Comparison change analysis approach to change-point estimation, *Journal of Applied Statistics*, 4, 379-402.

[32] Pettitt, A.N. (1979). A non-parametric approach to the change-point problem, *Applied Statistics*, 28(2) 126-135.

[33] Pettitt, A.N. (1980a). A simple cumulative sum type statistic for the change-point problem with zero-one observations, *Biometrika*, 67, 79-84.

[34] Pettitt, A.N. (1980b). Some results on estimating a change-point using nonparametric type statistics, *Journal of Statistical Computation and Simulation*, 11, 261-272.

[35] Quandt, R.E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes, *Journal of American Statistical Associations*, 55, 324-330.

[36] Sen, A.K. and Srivastava, M.S. (1975). Some one-sided tests for change in level, *Technometrics*, 17, 61-64.

[37] Sexton, R.S., Alidaee, B., Dorsey, R.E. and Johnson, J.D. (1998a). Global optimization for artificial neural networks: A tabu search application, *European Journal of Operational Research*, 106(2-3) 570-584.

[38] Sexton, R.S., Dorsey, R.E. and Johnson, J.D. (1998b). Toward global optimization of neural networks: A comparison of the genetic algorithm and backpropagation, *Decision Support Systems*, 22(2) 171-185.

[39] Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals*, Springer-Verlag, New York.

[40] White, H. (1994). Can neural networks forecast in the big leagues? Comparing network forecasts to the pros? In: *Proceedings of International Symposium of Forecasting*, Stockholm, Sweden, p.24.

[41] Wolkenhauer, O. and Edmunds, J.M. (1997). Possibilistic testing of distribution functions for change detection, *Intelligent Data Analysis*, 1, 119-127.

[42] Wong, F. and Tan, C. (1994). Hybrid neural, genetic, and fuzzy systems. In: Deboeck, G.J. (Ed.), *Trading on the Edge*, John Wiley & Sons, New York, 243-261.

[43] Worsley, K.J. (1979). On the likelihood ratio test for a shift in location of normal populations, *Journal of American Statistical Associations*, 74, 365-367.

[44] Yao, Y.C. (1988). Estimating the number of change-points via Schwarz's criterion, *Statistics and Probabilty Letters*, 6, 181-189.

[Received August 2002, Accepted February 2003]

Appendix: Table 1. Description of input variables (C_t: Closing price at time t)

Variable Name	Formular
MA6 (6-day moving average)	$\frac{1}{6}\sum_{i=1}^{6} C_{t-i+1}$
RSI (Relative strength index)	$100-\frac{100}{1+\frac{\sum_{i=1}^{6} UP_{t-i+1}/6}{\sum_{i=1}^{6} DOWN_{t-i+1}/6}}$ where UP: Upward price change, and $DOWN$: Downward price change
OBV (On balance volume)	$OBV_{t-1}+kV_t$ $k=1$ if $C_t > C_{t-1}$, -1 if $C_t < C_{t-1}$, 0 if $C_t = C_{t-1}$ where V_t means the volume at time t
%K (Stochastic %K)	$\frac{C_t - LL_{t-6}}{HH_{t-6}-LL_{t-6}}\times 100$ where HH_n and LL_n mean the highest high and the lowest low in the last n days respectively
%D (Stochastic %D)	$\frac{\sum_{i=1}^{6} \%K_{t-i+1}}{6}$
Momentum	$C_t - C_{t-5}$
Psychology	$\frac{1}{6}\sum_{i=1}^{6} UD_{t-i+1}$ where $UD_i=1$ if $C_t > C_{t-1}$, 0 if $C_t \leq C_{t-1}$
Disparity6 (6-day disparity)	$\frac{C_t}{MA_6}\times 100$
Disparity25 (25-day disparity)	$\frac{C_t}{MA_{25}}\times 100$
ROC6 (6-day price rate-of-change)	$\frac{C_t}{C_{t-6}}\times 100$
VR (Volume ratio)	$\frac{\left\{\sum_{i=1}^{6} VU_{t-i+1}+\frac{1}{2}\sum_{i=1}^{6} VS_{t-i+1}\right\}}{\left\{\sum_{i=1}^{6} VD_{t-i+1}+\sum_{i=1}^{6} VS_{t-i+1}\right\}}\times 100$ where VU_n, VD_n and VS_n mean the volume of an up, down and steady day of the stock price index, respectively, for n days
MA25 (25-day moving average)	$\frac{1}{25}\sum_{i=1}^{25} C_{t-i+1}$
ROC25 (25-day price rate of change)	$\frac{C_t}{C_{t-25}}\times 100$