Proceedings of Joint Conference of The Korean Data Analysis Society and Korean Data And Information Science Society, October 17-18, 2003, pp. 57-72

Using Structural Changes to support the Neural Networks based on Data Mining Classifiers: Application to the U.S. Treasury bill rates

Kyong Joo Oh[1]

Abstract

This article provides integrated neural network models for the interest rate forecasting using change-point detection. The model is composed of three phases. The first phase is to detect successive structural changes in interest rate dataset. The second phase is to forecast change-point group with data mining classifiers. The final phase is to forecast the interest rate with BPN. Based on this structure, we propose three integrated neural network models in terms of data mining classifier: (1) multivariate discriminant analysis (MDA)-supported neural network model, (2) case based reasoning (CBR)-supported neural network model and (3) backpropagation neural networks (BPN)-supported neural network model. Subsequently, we compare these models with a neural network model alone and, in addition, determine which of three classifiers (MDA, CBR and BPN) can perform better. For interest rate forecasting, this study then examines the predictability of integrated neural network models to represent the structural change.

Keywords: Structural Change, Change-Point Detection, Pettitt Test, Backpropagation Neural Networks, Multivariate Discriminant Analysis, Case-Based Reasoning

1. Introduction

The prediction of Interest rate has clear ramifications for cash management in any enterprise. Its movement also affects financing decisions such as capital budgeting and strategic investment. The prediction of interest rate is critical for managing risk in an investment portfolio as well as securing finance for corporate investment. Over the past several decades, statistical techniques and traditional software have been used extensively to model financial markets. Such statistical and software models have been beneficial for understanding market

---

[1]Professor, Division of Business Administration, Hansung University, Seoul 136-792, Korea.
E-mail : johanoh@hansung.ac.kr

behavior. This is especially true when the objective of the model is merely to describe and summarize the characteristics of a market rather than forecast its trajectory. For the task of interest rate forecasting, however, numerous studies in the past have underscored the inadequacy of statistical techniques and simulations based on traditional procedures.

Recently, several studies have demonstrated that artificial intelligence approaches, such as fuzzy theory (Ju et al., 1997) and neural networks (Deboeck and Cader, 1994; Hong and Han, 1996), can be alternative methodologies for chaotic interest rate data (Larrain, 1991; Peter, 1991; Jaditz andSayers, 1995). Previous work in the interest rate forecasting has tended to emphasize statistical techniques and artificial intelligent (AI) techniques in isolation over the past decades. However, an integrated approach, which makes full use of statistical approaches andAI techniques, offers the promise of increasing performance over each method alone (Chatfield, 1993). This study explores the ways in which such technologies may be combined synergistically, and illustrates the approach through the use of MDA, CBR and BPN as a data mining classifier. Up to date, it has been proposed that the integrated neural network model combining two or more models have a potential to achieve a high predictive performance in interest rate forecasting (Kim and Kim, 1996; Kim and Noh, 1997).

The movement of Interest rate is more fluctuatedsensitively by government's monetary policy than other financial data (Gordon and Leeper, 1994; Strongin, 1995; Bernanke and Mihov, 1995; Christiano et al., 1996; Leeper et al, 1996; Bagliano and Favero). Especially, banks play a very important role in determining the supply of money: Much regulation of these financial intermediaries is intended to improve its control. One crucial regulation is reserve requirements, which make it obligatory for all depository institutions to keep a certain fraction of their deposits in accounts with the Federal Reserve System, the central bank in the United States (Mishkin, 1995). It is supposed that government take an intentional action to control the currency flow which has direct influence upon interest rate. Therefore, we can conjecture that the movement of interest rate has a series of change points occurred by the planned monetary policy of government.

We propose three integrated neural network models in terms of data mining classifier based on the inherent characteristics of interest rate: (1) MDA-supported neural network model, (2) CBR-supported neural network model and (3) BPN-supported neural network model. Subsequently, we compare these models witha neural network alone, and determine which of three classifiers (MDA, CBR and BPN) can perform better. This study then examines the predictability of the integrated neural network models for interest rate forecasting using change-point detection and to compare the performance of several data mining classifiers.

The applied study performed in this study consists of the Treasury bill rate of 1 year's maturity in the U.S. from Jan., 1961 to May, 1999. Input variable selection is based on the causal model of interest rate presented by the econometricians. To explore the predictability, we dividedthe interest data into the training data over one period and the testing data over the other period. The predictability of interest rate is examined using the metricsof the root mean squared error (RMSE), the mean absolute error (MAE) and the mean absolute percentage error (MAPE).

We outline the development of change-point detection and its application to the financial economics in Section 2. Section 3 describes the proposed integrated neural network model details through the various data mining classifiers. Section 4 and 5 report the processes and the results of applied study. Finally, the concluding remarks are presented in Section 6.

2. Change-Point Detection

2.1. Existence and Detection of Structural Change for the Financial Economics

The detection and estimation of a structural or parametric change in forecasting is an important and difficult problem. In particular, financial analysts and econometricians have frequently used piecewisemodels which also include change-point models. They are known as models with structural breaks in the economics literature. In these models, the parameters are assumed to shift typically once during a fixed sample period and the goal is to estimate the two sets of parameters as well as the change point or structural break.

In order to detect the structural change, change-point detection methods have been applied to macroeconomic time series. Rappoport and Reichlin (1989) and Perron (1989, 1990) conduct the first study in this field. From then on, several statisticshave been developed which work well in a change-point framework, all of which are considered in the context of breaking the trend variables (Banerjee et al., 1992; Christiano, 1992; Zivot and Andrews, 1992; Perron, 1995; Vogelsang and Perron, 1995). In those cases where only a shift in the mean is present, the statistics proposed in thepapers of Perron (1990) or Perron and Vogelsang (1992) stand out.

In spite of the significant advances by these works, we should bear in mind that some variables do not show just one change point. Rather, it is common for them to exhibit the presence of multiple change points. Thus, it seems advisable to introduce a large number of change points in the specificationsof the models that allow us to obtain the abovementioned statistics. For example, Lumsdaine and Papell (1997) have considered the presence of two or more change points in trend variables. Based on this fact, we also assume the Treasury bill rates have two or more change points in our research model.

Up to date, there are few artificial intelligence models for financial applications to represent the change-point detection problems. Most of the previous research has a focus on the finding of unknown change points for the past, not to forecast for the future(Wolkenhauer and Edmunds, 1997; Li and Yu, 1999). Our model finds change points in the learning phase and forecasts change points in the testing phase. It is demonstrated that the introduction of change points to our model will make the predictability of interest rate greatly improve. In this study, a series of change points will be detected by Pettitt test, a nonparametric change-point detection method since nonparametric statistical property is a suitable match for a neural network model that is a kind of nonparametric method (White, 1992).

2.2. the Pettitt Test: a Nonparametric change-point detection method

In this study, a series of change points will be detected by thePettitt test (Pettitt, 1979, 1980a), a nonparametric change-point detection method, since nonparametric statistical property is a suitable match for a neural network model that is a kind of nonparametric method (White, 1992). *In this point, the introduction of the* Pettitt *test* is *fairly appropriate* for the analysis of chaotic time series data. The Pettitt test is explained as follows.

Consider a sequence of random variables $X_1, X_2, K\ , X_T$, then the sequence is said to have a change-point at τ if X_t for $t = 1, 2, K\ , \tau$ have a common distribution function $F_1(x)$ and X_t for $t = \tau+1, \tau+2, K\ , T$ 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 non-parametric 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

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

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

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

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

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

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

In the time sequence dataset, the Pettitt test detects a possible change point in which the structural change is occurred. Once the structural change is detected through the test, the dataset is divided into two intervals. The intervals before and after the change point form homogeneous groups which take heterogeneous characteristics from each other. This process becomes a fundamental part of the binary segmentation method explained in Section 3.

3. Model Specification

3.1. Integrated Neural Network Model based on the Structural Change

Data mining classifiers, change-point detection method and neural network learning methods have been integrated to forecast the Treasury bill rate of 1 year's maturity in the U.S. The advantages of combining multiple techniques to yield synergism for discovery and prediction have been widely recognized (Gottman, 1981; Kaufman et al., 1991). The proposed models are determined by the kind of data mining classifier which is applied to the second phase of model. This section provides the architecture and the characteristics of our research model to involve the change-point detection and the BPN. Based on the Pettitt test, the proposed model is composed of three phases as follows:

Phase 1: Constructing homogeneous groups

Pettitt test is a method to find a change-point in longitudinal data (Pettitt, 1979). It is known that interest rate at time t are more important than fundamental economic variables in determining interest rate at time $t+1$ (Larrain, 1991). Thus, we apply Pettitt test to Treasury bill rates at time t to generate a forecast for $t+1$ in the leaning phase. We, first of all, have to decide the number of change point. If change point is assumed to occur just one in given dataset, only the first step will be perform. Otherwise, all of three steps will be performed successively. The interval made by this process is defined as the significant interval, labeled SI, which is identified with a homogeneous group.

Step 1: Find a change point in $1 \sim N$ intervals by Pettitt test. If r_1 is a change point, $1 \sim r_1$ intervals are regarded as SI_1 and $(r_1+1) \sim N$ intervals are regarded as SI_2. Otherwise, it is concluded that there does not exist a change point for $1 \sim N$ intervals. $(1 \leq r_1 \leq N)$

Step 2: Find a change point in $1 \sim r_1$ intervals by Pettitt test. If r_2 is a change point, $1 \sim r_2$ intervals are regarded as SI_{11} and $(r_2+1) \sim r_1$ intervals are regarded as SI_{12}. Otherwise, $1 \sim r_1$ intervals are regarded as SI_1 like Step 1. $(1 \leq r_2 \leq r_1)$
Find a change point in $(r_1+1) \sim N$ intervals by Pettitt test. If r_3 is a change point, $(r_1+1) \sim r_3$ intervals are regarded as SI_{21} and $(r_3+1) \sim N$ intervals are regarded as SI_{22}. Otherwise, $(r_1+1) \sim N$ intervals are regarded as SI_2 like Step 1. $(r_1 \leq r_3 \leq N)$

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

First of all, the number of structural change should be determined. If just one change point is assumed to occur in a given dataset, only the first step will be performed. Otherwise, all of the three steps will be performed successively. This process plays a role of clustering that constructs groups as well as maintains the time sequence. In this point, Phase 1 is distinguished from other clustering methods such as the k-means nearest neighbor method and the hierarchical clustering method. They classify data samples by the Euclidean distance between cases without considering time sequence.

Phase 2: Group forecasting with data mining classifier

The significant intervals by Phase 1 are grouped to detect the regularities hidden in them and to represent the homogeneous characteristics of them. Such groups represent a set of meaningfultrends encompassing the significant intervals. Since those trends help to find regularity among the related output values more clearly, the neural network model can have a better ability of generalization for the unknown data. This is indeed a very useful point for sample design. In general, theerror for forecasting may be reduced by making the subsampling units within groups homogeneous and the variation between groupsheterogeneous (Cochran, 1977). After Phase 1 detects the appropriate groups hidden in the significant intervals, various classifiers (MDA, CBR and BPN) are applied to the input data samples at time t with group outputs for $t+1$. In this sense, Phase 2 is a model that is trained to find an appropriate group for each given sample.

Phase 3: Forecasting the output with BPN

Phase 3 is built by applying the BPN model to each group. Phase 3 is a mapping function between the input sample and corresponding desired output (i.e. Treasury bill rate). Once Phase 3 is built, then the sample can be used to forecast the Treasury bill rate.

3.2. The proposed models

According to the kind of classifier used in the Phase 2, we proposethree integrated neural network models: (1) MDA-supported neural network model, (2) CBR-supported neural network model and (3) BPN-supported neural network model.

3.2.1. MDA-supported neural network model

MDA is used to classify individuals into one of two or more alternative groups on the basis of a set of measurements. Theoretically, This method is based on the Fisher's linear discriminant function by maximizing the ratio of between-groups and within-groups variances (Fisher, 1936). The groups are knownto be distinct, and each individual belongs to one of them. The MDA can be used to identify which variables contribute to making the classification (McLachlan, 1992; Hair, et al., 1995). This model applies MDA to forecast the change-point group in the second stage.

3.2.2. CBR-supported neural network model

CBR canbe an effective tool for predicting temporal variables. By extending the conceptof a case to include several immediate neighbors and by seeking multiple exemplars from the case base, a CBR system can forecast a trajectory with a remarkable accuracy. In fact, such a system can exceed the performance of neural networks in some events, a methodology which has been widely applied to the task of forecasting (Deboeck and Cadar, 1994; Trippi and Turban, 1992; Kolodner, 1991, 1993). This model uses CBR to forecast the change-point group in the second stage.

3.2.3. BPN-supported neural network model

The neural network methodology has been applied extensively to solve practical problems following the publication of the backpropagation algorithm for the multi-layer perceptron (Rumelhart, 1986). The algorithm was developed for the perceptron model, a simple structure to simulate a neuron (Rosenblatt, 1957). Today the BPN is the most widely used neural networks algorithm in science, engineering, finance and other fields. This model uses the BPN two times. In the second stage, BPN is applied to forecast the change-point group. Finally, BPN also forecast the desired output.

4. Data and Variables

The input variables used in this study are M2, consumer price index, expected real inflation rates and industrial production index. They are used in both Phase 2 and Phase 3. The lists of variables used in this study are summarized in Table 1. They are those which were found significant in interest rates forecasting by previous study (Oh and Han, 2000). To obtain stationary and thereby facilitate forecast, the input data were transformed by a logarithm and a difference operation. Moreover, the resulting variables were standardized to eliminate the effects of units.

Table 1. Description of Variables.

TBILL	Treasury Bill with 1 year's maturity	Output
M2	Money Stock	Input
CPI	Consumer Price Index	Input
ERIR	Expected Real Interest rate	Input
IPI	Industrial Production Index	Input

The training phase included observations from January1961 to August 1991 while the testing phase runs from September 1991 to May 1999. The interest rate data are presented in Figure 1. Figure 1 shows that the movement of interest rates is highly fluctuated during the last forty years.

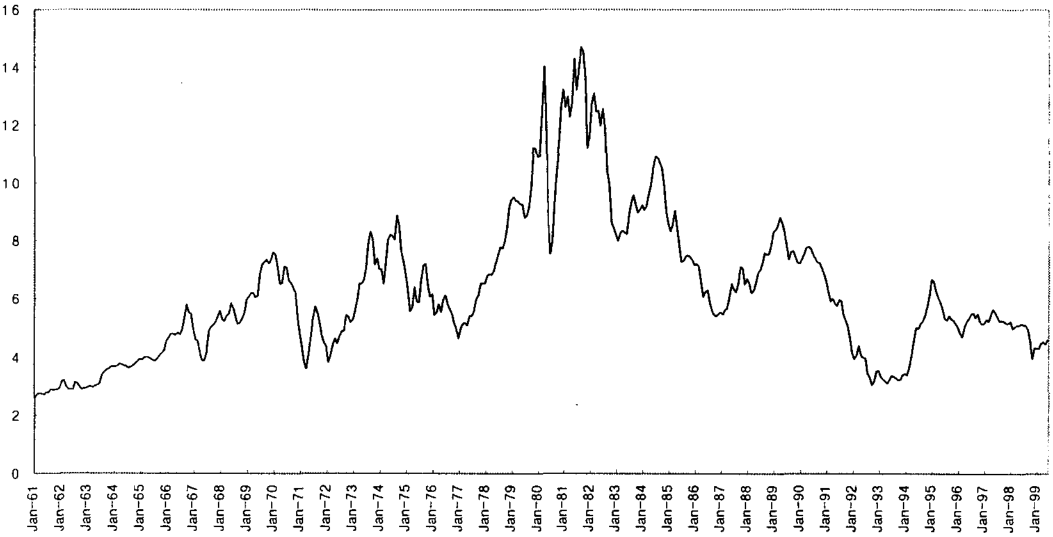

Figure 1. Yield of US Treasury bills with a maturity of 1 year from Jan. 1960 to May 1999

The study employed two neural network models. One model, labeled Pure_NN, involve input variables at time t to generate a forecast for $t+1$. The input variable is M2, CPI, ERIR and IPI. The second type has two-step forecasting models which consist of three phases mentioned in section 3. The first step is Phase 2 that forecasts the change-point group while the next step is Phase 3 that forecasts the desired output. The classifiers used in the model were as shown in Table 2. For validation, four learning models were also compared.

Table 2. Models and their associated classifiers for US Treasury bill rate forecasting.

Pure_NN	None
CBR_NN	Case Based Reasoning
BPN_NN	Backpropagation Neural Network
MDA_NN	Multivariate Discriminant Analysis

5. Empirical Results

The Pettitt test is applied to the interest rate dataset. Since the interest dataset is about forty years long, it is considered that there exist two or more change points. Therefore, we obtain 4 significant intervals as like the result of Table 3. Table 3 also presents descriptive statistics including the mean and the variance. Group 1 is the stable interval that has low variance. Group 2 and 3 are more fluctuated intervals than Group 1 in term of the variance. Group 4 is highly fluctuated. Skewness and kurtosis show that four groups have similar attributes in the distribution.

Table 3. Period and descriptive statistics of groups for the learning phase, Jan. 1961 - Aug. 1991

	Group 1	Group 2	Group 3	Group 4
Periods	Jan. 61 - Nov. 65	Dec. 65 Feb. 73	Mar. 73 May 78	Jun. 78 Aug. 91
Minimum	2.720	3.600	4.640	5.260
Maximum	4.230	7.610	8.880	14.700
Range	1.510	4.010	4.240	9.440
Mean	3.378	5.419	6.507	8.654
Variance	0.219	0.938	1.008	5.240
Standard Deviation	0.468	0.969	1.004	2.289
Skewness	0.147	0.496	0.363	0.781
Kurtosis	-1.544	-0.361	-0.575	-0.135

To highlight the performance due to various models, the actual values of Treasury bill rates and their predicted values are shown in Figure 2. The predicted values of pure BPN model (i.e. PURE_NN) and CBR-supported neural network model get apart from the actual values in some intervals. Numerical values for the performance metrics by predictive model are given in Table 4. Figure 3 presents histograms of RMSE, MAE and MAPE of predictionsfor each learning model. According to RMSE, MAE and MAPE, the outcomes indicate that BPN-supported neural network model and MDA-supported neural network model are superior to the pure BPN model and CBR-supported neural network model.

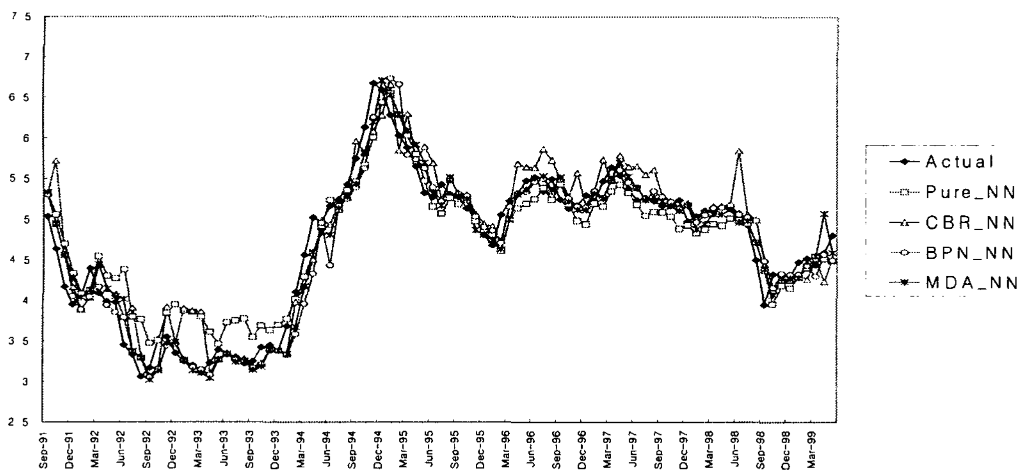

Figure 2. Actual vs predicted values due to various models for TBILL

Table 4. Performance results in the case of US Treasury bill rate forecasting based on the root mean squared error (RMSE), the mean absolute error (MAE) and the mean absolute percentage error (MAPE)

Pure_NN	0.0973	0.2506	5.969%
CBR_NN	0.1015	0.2489	5.553%
BPN_NN	0.0584	0.1745	3.746%
MDA_NN	0.0481	0.1733	3.784%

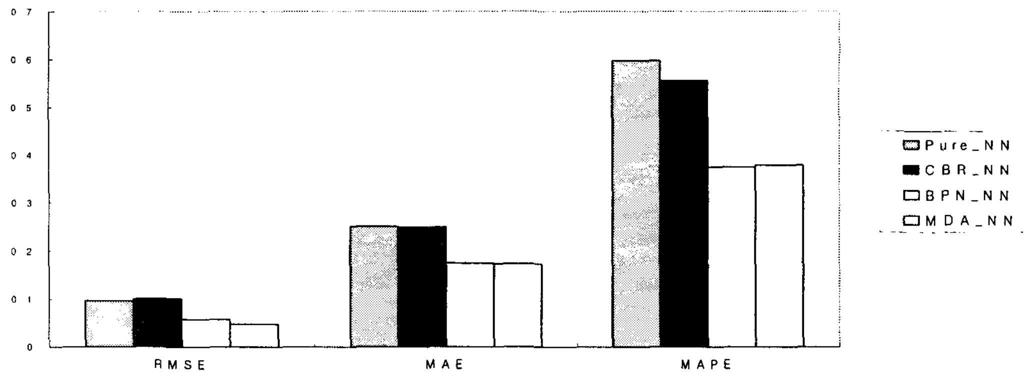

Figure 3. Histogram of RMSE, MAE and MAPE resulting from forecasts of TBIILL

Our approach to integration involves a multistrategy technique which may be called second-order learning. CBR has provided the good results with the second order learning process and the integrated approach (Kim and Noh, 1997; Kim and Joo, 1997). However, CBR did not perform well in this study. In the second-order learning, the forecast from the superior method is selected on a case-by-case basis to determine the output of overall model. In other words, the second step (BPN in this study) serves as a metalevel process to determine which of three elementary modules (CBR, BPN and MDA in this study) perform better. In this point, CBR is not a good metalevel predictive method. Thus, we will choose BPN or MDA as the elementary module for real application of model.

We use the pairwise t-test to examine whether there exist the differences in the predicted values of models according to the absolute percentage error (APE). This metric is chosen since it is commonly used (Carbone and Armstrong, 1982) and is highly robust (Armstrong and Collopy, 1992; Makridakis, 1993). Since the forecasts are not statistically independent and not always normally distributed, we compare the forecast's APEs 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 5 shows t-values and p-values when the predictionaccuracies of the left-vertical methods are compared with those for the right-horizontal methods. Mostly, the neural network models using change-point detection perform significantly better than the pure BPN model at a 1%

significant level except the analysis of CBR-supported neural network model. Therefore, our research model is demonstrated to obtain the improved performance through the change-point detection approach.

Table 5. Pairwise t-tests for the differences in residuals for US interest rate prediction based on the absolute percentage error (APE) with the significance level in parentheses.

Model	BPN_NN	CBR_NN	Pure_NN
MDA_NN	-0.14 (0.882)	3.29 (0.001)***	3.75 (0.000)***
BPN_NN		3.25 (0.001)***	3.43 (0.000)***
CBR_NN			0.72 (0.467)

*** Significant at 1%

The neural network models using change-point detection turn out to have a high potential in interest rate forecasting. This is attributable to the fact that it categorizes the input data samples into homogeneous group and extracts regularities from each homogeneous group. Therefore, the neural network models using change-point detection can cope with the noise or irregularities more efficiently than the pure BPN model. In addition, BPN and MDA perform very well as a tool in interest rate forecasting.

6. Concluding Remarks

This study has suggested the integrated neural network models in the interest rate *forecasting using change-point detection*. The basic concept of proposed model is toobtain significant intervals by change-point detection, to identify them as change-point groups, and to involve them in interest rate forecasting. We propose integrated neural network models which consist of three phases. In the first phase, we conduct the nonparametric statistical test for the change-point detection to construct the homogeneous groups. In the second phase, we apply several kinds of classifiers to forecast the change-point group. In thefinal phase, we apply BPN to forecast the desired output.

The neural network models using change-point detection perform significantly better than the pure BPN model at a 1% significant level except the analysis of CBR-supported neural network model. Experimental results showed that the neural network models using change-point detection outperform the pure BPN model significantly, which implies the high potential of involving thechange-point detection in the model. BPN and MDA performed very well as data mining classifiers while CBR did not. Our integrated neural network models are demonstrated to be useful intelligent data analysis methods with the concept of change-point detection. In conclusion, we have shown that the proposed models improve the predictability of interest rate significantly.

The proposed model has the promising possibilities to improve the performance if further studies are to focus on the various approaches in the construction and the prediction of change-point group. In final phase of the model, other intelligent approaches can be used to forecast the final output besides BPN. In addition, the proposed models may be applied to other chaotic time series data, such as stock market prediction and exchange rate prediction. By the extension of these points, future research is expected to provide more improved neural network models with superior performances.

References

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

Bagliano, F.C. and C.A. Favero, "Information from financial markets and VAR measures of monetary policy," *European Economic Review,* 43 (1999), 825-837.

Banergee, A., R. Lumsdaine and J. Stock, "Recursive and sequential tests of the unit root and trend break hypothesis: Theory and international evidence, *Journal of Business and Economic Statistics,* 10 (1992), 271-287.

Bernanke, B.S., Mihov, I., "Measuring monetary policy," Working Paper No. 5145, NBER, Cambridge, MA, 1995.

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

Chatfield, C., "Neural networks: Forecasting breakthrough or passing fad?" *International Journal of Forecasting,* 9 (1993), 1-3.

Christiano, L.l., Eichenbaum, M., Evans, C.L., "The effects of monetary policy shocks: Evidence from the flow of funds," *Review of Economics and Statistics,* 78 (1996), 16-34.

Cochran, W.G., *Sampling Techniques,* Third edition, John Wiley & Sons, New York, 1977.

Deboeck G. and M. Cader, *Trading on the Edge*, John Wiley & Sons, New York, 1994.

Fisher, R.A., "The use of multiple measurements in taxonomic problems, *Annals of Eugenics,* 7 (1936), 179-188.

Gordon, D., Leeper, E.M., "The dynamic impacts of monetary policy: An exercise in tentative identification," *Journal of Political Economy,* 102 (1994), 1228-1247.

Gottman, J.M., *Time-Series Analysis,* Cambridge University Press, New York, 1981.

Hair, J.F. Jr., R.E. Anderson, R.L. Tatham, and W.C. Black, *Multivariate Data Analysis with Readings,* Fourth edition, Prentice Hall, New York, 1995.

Hong, T. and I, Han, "The prediction of interest rate using artificial neural network models," *First asia-pacific decision sciences institute conference* (1996), 975-984.

Iman, R. and W.J. Conover, *Modern Business Statistics,* Wiley, New York, 1983.

Jaditz, T. and Sayers C.L., " Nonlinear in the interest rate risk premium," In R.R. Trippi, *Chaos and Nonlinear Dynamics in the Financial Markets* (pp. 335-357), Irwin, 1995.

Ju Y.J., C.E. Kim, and J.C. Shim, "Genetic-based fuzzy models: interest rate forecasting problem," *Computers and Industrial Engineering,* Vol. 33, No. 3-4 (1997), 561-564.

Kaufman, K.A., R.S. Michalski and L. Kerschberg, "Mining for knowledge in databases: Goals and general description of the INLEN system," In G. Piatetsky-Shapiro and W.J. Frawley (Eds), *Knowledge Discovery in Databases* (pp. 449-462), AAAI/MIT Press, Cambridge, MA, 1991.

Kolodner, J., "Improving human decision making through case-based decision aiding," *AI Magazine* 12 (1991), 52-68.

Kolodner, J., *Case-Based Reasoning,* Morgan Kaufmann, San Mateo, CA, 1993.

Kim S.H. and H.J. Noh, "Predictability of interest rate using data mining tools: a comparative analysis of Korea and the US," *Expert Systems with Application,* 15 (1997), 85-95.

Kim, S.H. and J. Joo, " Second-order learning for complex forecasting tasks: case study of Video-On-Demand," *Korean Journal of Expert System,* 3 (1997), 31-45.

Kim, S.H. and K.M. Kim, "Integrating multivariate statistics and neural networks for financial prediction: case study in interest rate forecasting," *First Asia-Pacific Decision Sciences Institute Conference* (1996), 995-1003.

Larrain, M., "Empirical tests of chaotic behavior in a nonlinear interest rate model," *Financial Analysts Journal,* Sep.-Oct. (1991), 51-62.

Leeper, E.M., "Narrative and VAR approaches to monetary policy: Common identification problems," *Journal of Monetary Economics*, 40 (1997), 641-657.

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

Lumsdaine, R.L. and D.H. Papell, "Multiple trends and the unit root hypothesis," *The Review of Economics and Staistics,* 79 (1997), 212-218.

McLachlan, G.J., *Discriminant Analysis and Statistical Pattern Recognition,* John Wiley & Sons, New York, 1992.

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

Mishkin, F.S., *The economics of money, banking, and fiancial markets,* Fourth edition, Harper

Collins, New York, 1995.

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

Perron, P., "The great crash, the oil price shock, and the unit root hypothesis," *Econometrica*, 57 (1989), 1361-1402.

Perron, P., "Testing for a unit root in time series with a changing mean," *Journal of Besiness and Economic Statistics*, 8 (1990), 153-162.

Perron, P., "Further evidence on breaking trend functions in macroeconomic variables," manuscript, Universit de Montreal, Canada, 1995.

Perron, P. and T. Vogelsang, "Nonstationarity and level shifts with an application to purchasing power parity," *Journal of Business and Economic Statistics*, 10 (1992a), 301-320.

Peters E.E., *Chaos and Order in the Capital Markets*, John Wiley & Sons, New York, 1991.

Pettitt, A.N., "A non-parametric approach to the change-point problem," *Applied Statistics*, 28, No.2 (1979), 126-135.

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

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

Rappport, P. and L. Reichlin, "Segmented trends and non-stationary time series," *The Economic Journal*, 99 (1989), 168-177.

Rosenblatt, F., *Principles of Neurodynamics*, Spartan, New York, 1962.

Rumelhart, D.E., G.E. Hinton and R.J. Williams, "Learning internal representations by back propagation: In D.E. Rumelhart, J.L. McClelland and PDP research group (Eds), *Parallel Distributed Processing*, Volume 1, MIT Press, Cambridge, MA, 1986.

Strongin, S., "The identification of monetary policy disturbances. Explaining the liquidity puzzle," *Journal of Monetary Economics*, 35 (1995), 463-497.

Trippi, R.R. and E. Turban, *Neural Networks in Finance and Investing*, Second edition, Irwin, Chicago, 1996.

Vogelsang T. and P. Perron, "Additional tests for a unit root allowing for a break in the trend function at an unknown time, Manuscript, Department of Economics, Ithaca, New York, 1995.

White, H., "Connectionist nonparametric regression: Multilayer feedforward networks can learn arbitrary mappings," in H. White (Ed.), *Artificial Neural Networks: Approximations and Learning Theory*, Blackwell, Oxford, UK, 1992.

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

Zivot, E. and D.W.K. Andrews, "Further evidence on the great crash, the oil-price shocks, and the unit-root hypothesis," *Journal of Business and Economic Statistics,* 10 (1992), 251-270.