# Neural Network Modeling supported by Change-Point Detection for the Prediction of the U.S. Treasury Securities

Kyong Joo Oh, Ingoo Han

Graduate School of Management,
Korea Advanced Institute of Science and Technology

**Abstract**

The purpose of this paper is to present a neural network model based on change-point detection for the prediction of the U.S. Treasury Securities. Interest rates have been studied by a number of researchers since they strongly affect other economic and financial parameters. Contrary to other chaotic financial data, the movement of interest rates has a series of change points due to the monetary policy of the U.S. government. The basic concept of this proposed model is to obtain intervals divided by change points, to identify them as change-point groups, and to use them in interest rates forecasting. The proposed model consists of three stages. The first stage is to detect successive change points in the interest rates dataset. The second stage is to forecast the change-point group with the backpropagation neural network (BPN). The final stage is to forecast the output with BPN. This study then examines the predictability of the integrated neural network model for interest rates forecasting using change-point detection.

## 1. Introduction

Interest rates are one of the most closely watched variables in the economy. Their movements are reported almost daily by the news media since they directly affect our everyday lives and have important consequences for the economy. There exist extensive studies in this area using statistical approaches, such as term structure models, vector autoregressive (VAR) models, autoregressive conditionally heteroskedastic (ARCH) - generalized autoregressive conditionally heteroskedastic (GARCH) models and other time series analysis approaches.

Currently, several studies have demonstrated that artificial intelligence (AI) approaches, such as fuzzy theory (Ju et al., 1997) and neural networks (Deboeck and Cader, 1994), can be alternative methodologies for chaotic interest rates data (Larrain, 1991; Peters, 1991; Jaditz and Sayers, 1995). Previous work in interest rates forecasting has tended to use statistical techniques and AI techniques in isolation. However, an integrated approach, which makes full use of statistical approaches and AI techniques, offers the promise of increasing performance over each method alone (Chatfield, 1993). It has been proposed that the integrated neural network models combining two or more models have the potential to achieve a high predictive performance in interest rates forecasting (Kim and Noh, 1997).

In general, interest rates data is controlled by government's monetary policy more than other financial data (Gordon and Leeper, 1994; Strongin, 1995; Christiano et al., 1996; Leeper et al, 1996; Bagliano and Favero, 1999). Especially, banks play a very important role in determining the supply of money. Much regulation of these financial intermediaries is intended to improve their 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). The government takes intentional action to control the currency flow which has direct influence upon interest rates. Therefore, we can conjecture that the movement of interest rates has a series of change points which occur because of the monetary policy of the government.

Based on these inherent characteristics in interest rates, this study suggests the change-point detection for interest rates forecasting. The proposed model consists of three stages. The first stage is to detect successive change points in the interest rates dataset. The second stage is to forecast the change-point group with BPN. The final stage is to forecast the output with BPN. This study then examines the predictability of the integrated neural network models for interest rates forecasting using change-point detection.

Through the discovery of different patterns in the U.S. Treasury securities, the focus then shifts to the change-point detection-assisted modeling of Treasury bill rates with 1 years' maturity and Treasury bond rates with 30 years' maturity. Input variable selection is based on the causal model of interest rates presented by the econometricians. To explore the predictability, we divided the interest data into the training data over one period and the testing data over the next period. The predictability of interest rates is examined using the metrics of the root mean squared error (RMSE), the mean absolute error (MAE) and the mean absolute percentage error (MAPE).

## 2. Application of Change-Point Detection

Financial analysts and econometricians have frequently used piecewise-linear models which also include change-point models. They are known as models with structural breaks in economic literature. In these models, the parameters are assumed to shift − typically once − during a given sample period and the goal is to estimate the two sets of parameters as well as the change point or structural break.

There are few artificial intelligence models to consider the change-point detection problems. Most of the previous research has a focus on the finding of unknown change points for the past, not the forecast for the future (Wolkenhauer and Edmunds, 1997; Li and Yu, 1999). Our model obtains intervals divided by change points in the training phase, identifies them as change-point groups in the training phase, and forecasts to which group each sample is assigned in the testing phase. It will be tested whether the introduction of change points to our model may improve the predictability of interest rates.

In this study, a series of change points will be detected by the 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). In addition, the Pettitt test is a kind of Mann-Whitney type statistic, which has remarkably stable distribution and provides a robust test of the change point resistant to outliers (Pettitt, 1980b). In this point, the introduction of the Pettitt test is fairly appropriate to the analysis of chaotic interest rates data.

**3. Model Specification**

In this section, we discuss the architecture and the characteristics of our model to integrate the change-point detection and the BPN. Based on the Pettitt test, the proposed model consists of three stages: (1) the change-point detection (CPD) stage, (2) the change-point-assisted group detection (CPGD) stage and (3) the output forecasting neural network (OFNN) stage. The BPN is used as a classification tool in CPGD and as a forecasting tool in OFNN.

**Stage 1: Construction of homogeneous groups**

The Pettitt test is a method to find a change-point in time series data (Pettitt, 1979). It is known that interest rates at time $t$ are more important than fundamental economic variables in determining interest rates at time $t+1$ (Larrain, 1991). Thus, we apply the Pettitt test to interest rates at time $t$ in the training phase. The interval made by the test is defined as the significant interval, labeled $SI$, which is identified with a homogeneous group. Multiple change points are obtained under the binary segmentation method (Vostrikova, 1981).

We, first of all, have to decide the number of change points. 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 which constructs groups as well as maintains the time sequence. In this point, the Stage 1 is distinguished from other clustering methods such as the k-means nearest neighbor method and the hierarchical clustering method which classify data samples by the Euclidean distance between cases without considering the time sequence. In addition, we analyze the characteristics of groups according to descriptive statistics including the mean and the variance, and also observe the density plot of groups since the classification accuracy is highly sensitive to the density of the samples (Wang, 1995).

**Stage 2: Forecast the group with BPN**

The significant intervals in the Stage 1 are grouped to detect the regularities hidden in interest rates. Such groups represent a set of meaningful trends encompassing interest rates. 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, the error for forecasting may be reduced by making the subsampling units within groups homogeneous and the variation between groups heterogeneous (Cochran, 1977). After the appropriate groups hidden in interest rates are detected by the Stage 1, BPN is applied to the input data samples at time $t$ with group outputs for $t+1$ given by CPD. In this sense, CPGD is a model that is trained to find an appropriate group for each given sample.

**Stage 3: Forecast the output with BPN**

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

**4. Data and Variables**

The input variables are M2, CPI, ERIR and IPI. The input data sets in this study consist of the monthly rate of change. Given the data sequence $d_1, d_2, \ldots, d_t$, we form the rate of change at time $t+1$ by dividing the first difference at that time by the datum at time $t$:

$$\frac{d_{t+1} - d_t}{d_t} \tag{4}$$

The data used in this study is monthly yields on the U.S. Treasury securities from January 1977 to May 1999. The forecast of the U.S. Treasury security rates had better not be based on the equivalence alone, but should be performed through individual modeling. In this sense, we build two integrated neural network models for one-year T-bills and thirty-year T-bonds, and establish the experiment interval differently for each model. The motivation for this plan is to see the impact of interval size on the performance and furthermore to demonstrate the generality of the proposed model.

The study employs two neural network models. One model, labeled Pure_NN, involves four input variables at time $t$ to generate a forecast for $t+1$. The second type, labeled BPN_NN, is the two-step BPN model that consists of three stages mentioned in section 3. The first step is the Stage 2 that forecasts the change-point group while the next step is the Stage 3 that forecasts the output. For validation, two learning models are also compared.

**5. Empirical Results**

The Pettitt test is applied to the interest rates dataset. Since the interest dataset is about forty years long for one-year T-bills, it is considered that there exist three or more change points. It is further assumed that there exist two change points because of the small size of data for thirty-year T-bonds.

Numerical values for the performance metrics by the predictive model are given in Table 1. According to RMSE, MAE and MAPE, the outcomes indicate that the proposed neural network model is superior to the pure BPN model for both of the interest rates.

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). Table 2 shows t-values and p-values. The neural network models using change-point detection perform significantly better than the pure BPN model at a 1% significant level. Therefore, the proposed model is demonstrated to obtain improved performance using the change-point detection approach.

Table 1(a) Performance results of one-year Treasury bill rate forecasting based on the root mean squared error (RMSE), the mean absolute error (MAE) and the mean absolute percentage error (MAPE)

| Model | RMSE | MAE | MAPE |
|---|---|---|---|
| Pure_NN | 0.0973 | 0.2506 | 5.969% |
| BPN_NN | 0.0584 | 0.1745 | 3.746% |

(b) Performance results of thirty-year Treasury bond rate forecasting based on the RMSE, the MAE and the MAPE

| Model | RMSE | MAE | MAPE |
|---|---|---|---|
| Pure_NN | 0.3001 | 0.4610 | 7.836% |
| BPN_NN | 0.0709 | 0.2132 | 3.542% |

Table 2 Pairwise t-tests for the difference in residuals between the pure BPN model and the proposed neural network model for one-year T-bills and thirty-year T-bonds based on the absolute percentage error (APE) with the significance level in parentheses.

| Interest Rates | Test Value |
|---|---|
| One-year T-bills | 3.43 (0.000)*** |
| Thirty-year T-bonds | 8.17 (0.000)*** |

*** Significant at 1%

## 6. Concluding Remarks

The neural network models using change-point detection perform significantly better than the pure BPN model at a 1% significant level. These experimental results imply the change-point detection has a high potential to improve the performance. Our integrated neural network model is demonstrated to be a useful intelligent data analysis method with the concept of change-point detection. In conclusion, we have shown that the proposed model improves the predictability of interest rates significantly.

The proposed model has the promising possibility of improving the performance if further studies are to focus on the optimal decision of the number of change point and the various approaches in the construction of change-point groups. In the Stage 3, other intelligent techniques besides BPN can be used to forecast the output. In addition, the proposed model may be applied to other chaotic time series data such as stock market prediction and exchange rate prediction.

## References


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

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

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

Cochran, W.G.; *Sampling techniques*. New York: John Wiley & Sons. 1977.

Deboeck, G.J., & Cader, M.; Trading U.S. treasury notes with a portfolio of neural net models, In G.J. Deboeck (Ed.), *Trading on the Edge* (pp. 102-122). New York: John Wiley & Sons. 1994.

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

Jaditz, T., & Sayers, C.L.; Nonlinearity in the interest rate risk premium. In R.R. Trippi (Ed.), *Chaos & nonlinear dynamics in the financial markets* (pp. 335-357). Boston, MA: Irwin. 1995.

Ju, Y.J., Kim, C.E., & Shim, J.C.; Genetic-based fuzzy models: Interest rates forecasting problem. *Computers and Industrial Engineering, 33*, 561-564. 1997.

Kim, S.H., & Noh, N.H.; Predictability of interest rates using data mining tools: A comparative analysis of Korea and the US. *Expert Systems with Applications, 15(1)*, 85-95. 1997.

Kim, S.K., & Park, J.I.; A structural equation modeling approach to generate explanations for induced rules. *Expert Systems with Applications, 10(3-4)*, 403-416. 1996.

Larrain, M.; Empirical tests of chaotic behavior in a nonlinear interest rate model. *Financial Analysts Journal, 47*, 51-62. 1991.

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

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

Mishkin, F.S.; *The economics of money, banking, and financial markets*, New York: Harper Collins. 1995.

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

Peters, E.E.; *Chaos and order in the capital markets*. New York: John Wiley & Sons. 1991.

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

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

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

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

Vostrikova, L.J.; Detecting "disorder" in multidimensional random process, *Sov. Math. Dokl.*, 24, 55-59. 1981.

Wang, S.; The unpredictability of standard backpropagation neural networks in classification applications. *Management Science, 41*, 555-559. 1995.

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

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