# The usefulness of overfitting via artificial neural networks for non-stationary time series

Jae Joon Ahn*, Kyong Joo Oh**, Tae Yoon Kim***

*Department of Information and Industrial Engineering, Yonsei University (redsamzang@yonsei.ac.kr)

**Department of Information and Industrial Engineering, Yonsei University (johanoh@yonsei.ac.kr)

***Department of Statistics, Keimyung University (tykim@kmu.ac.kr)

**Abstract**

The use of Artificial Neural Networks (ANN) has received increasing attention in the analysis and prediction of financial time series. Stationarity of the observed financial time series is the basic underlying assumption in the practical application of ANN on financial time series. In this paper, we will investigate whether it is feasible to relax the stationarity condition to non-stationary time series. Our result discusses the range of complexities caused by non-stationary behavior and finds that overfitting by ANN could be useful in the analysis of such non-stationary complex financial time series.

Keywords: Non-stationary time series; Overfitting; Artificial neural networks; Asymptotic stationary autoregressive model

## 1. Introduction

Stationarity is usually a desirable assumption in the analysis and prediction of financial time series. Indeed, when one is interested in finding a relation between the present and the past through prediction, the relation must be stationary throughout the evolution of time; otherwise, the prediction would not be possible [3]. In fact, there is a well-defined mathematical definition of stationarity, which imposes certain regular conditions on either a distribution or its moments [1] and various methods are developed to find the stationary relation governing the time series data [2, 4, 5, 9]. Recently, artificial neural networks (ANN) are being used more frequently in the analysis of financial time series as they move from simple pattern recognition to a diverse range of application areas [7]. It is known that the ANN mapping process can cover a greater range of problem complexity and is superior in its generality and practical ease in implementation owing to its powerful and flexible capability [6].

In this short article, we will investigate ANN's capability of extracting a rule governing seemingly non-stationary financial time series data. Indeed, we will investigate the range of complexity due to non-stationarity that ANN might cover and find that overfitting by an ANN might be useful for such complex financial time series analysis. Our discussion will mainly be based on empirical examples. The following section discusses various non-stationary financial time series theoretically and empirically to show that overfitting by ANN could play a significant role in the analysis of non-stationary financial time series. The final section contains the concluding remarks.

## 2. Non-stationary processes

Since a stationary condition usually implies a stable process, a varying process may be considered to be close to stationary if it varies slowly. From this point of view, one might define an asymptotic stationary autoregressive (AR) process $Z_t$ as the following nonlinear AR model:

$$Z_t=f_n(Z_{t-1},Z_{t-2},\dots,Z_{t-p})+\varepsilon_t,\quad t=p+1,2,\ldots,n \qquad (1)$$

where $\lim_n f_n=f$ is termed as an "asymptotic stationary AR function" and $\varepsilon_t$ is the identically and independently distributed (iid) error. Note that $Z_t$ is clearly non-stationary since its AR function depends on $n$. Intuitively speaking, the asymptotic stationary AR process $Z_t$ assumes that it varies slowly enough that its asymptotic stationary AR function $f$ may exist. There are some mathematical extensions of stationarity in this direction by others, see, e.g., asymptotic stationary density [8]. Note that varying degree (and hence complexity) of a process is determined by the convergence speed for $\lim_n f_n=f$. More precisely, the convergence speed may be defined by $\|f_n-f\|=r_n$ for $f_n$, $f\in C(S)$ and a real valued sequence $r_n$ converging to zero. Here $C(S)$ is the space of all continuous real functions on the closed set $S$, metrized by $\|f-g\|=\sup_{t\in S}|f(t)-g(t)|$.

Now, various non-stationary processes (slowly or regularly varying processes) are studied through the Korea stock price index (KOSPI) during the period 1994 ~ 1995 (Figure 1). A rough look at Figure 1 reveals that there are varying degrees of change in each year. For our study, KOSPI Indices of 1995, 1997 and 1999 have been chosen. It is easy to see from Figure 1 that all of them are non-stationary and that their varying degree is increasing in the order of 1995, 1997, and 1999 (this is also clear from sample auto-correlation function (SACF) of KOSPI of each year, see Figure 2). Note that during 1997 ~ 1999, Korea experienced an economic crisis followed by recovery, and as such the KOSPI of those years are characterized by volatile movements. In this paper, SACF is obtained via *STATISTICA version 6*.

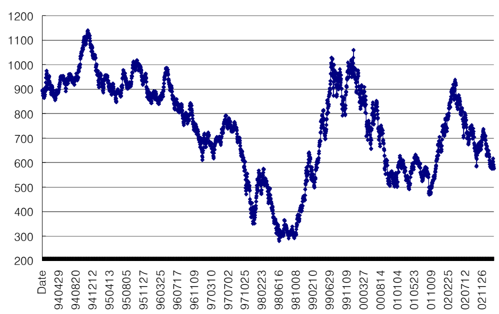
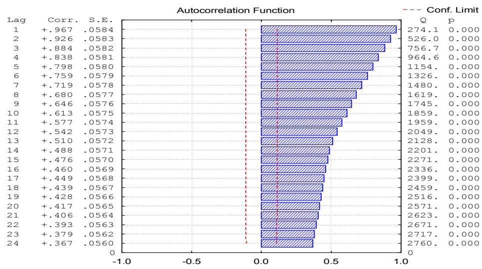
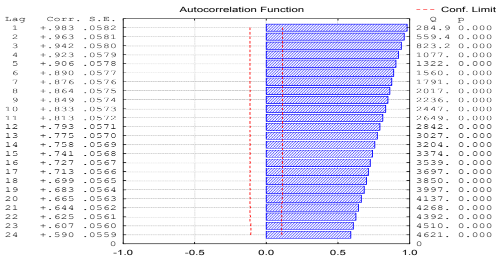



Figure 1. KOSPI from January, 1994 to February, 2003

(a) 1995




(b) 1997




(c) 1999

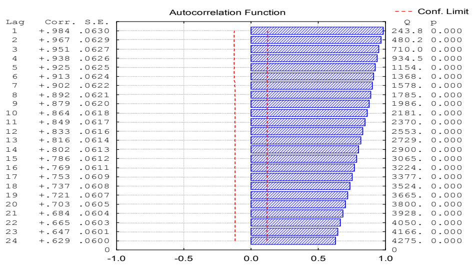



Figure 2. SACF of KOSPI

Since the movement of KOSPI becomes more complex in the order of 1995, 1997 and 1999, we use the model (1) with $p$=3, 4, 5 respectively, i.e.,

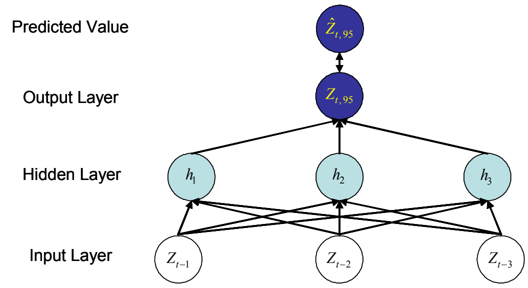
$$Z_{t,95}=f_{n,1}(Z_{t-1},Z_{t-2},Z_{t-3})+\varepsilon_t, \quad (2)$$

$$Z_{t,97}=f_{n,2}(Z_{t-1},Z_{t-2},Z_{t-3},Z_{t-4})+\varepsilon_t, \quad (3)$$

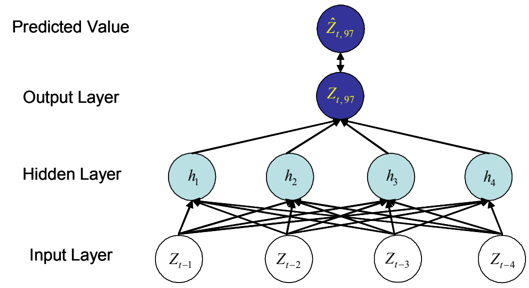
$$Z_{t,99}=f_{n,3}(Z_{t-1},Z_{t-2},Z_{t-3},Z_{t-4},Z_{t-5})+\varepsilon_t \quad (4)$$

where $n$ is the number of observations for each year. Note that the AR order for each year $f_{n,i}$ ($i$=1, 2, 3) has been chosen so that it contains a degree of complexity. Now, $f_{n,i}$ is recovered through a backpropagation neural network (BPN), $p \times p \times 1$ multi-layer ANN for each year which consists of input layer of $p$ units, hidden layer of $p$ units, and output layer of 1 unit. As an activation function, the logistic function is used with learning rate, momentum and initial weights given by 0.1, 0.1 and 0.3, respectively. See Figure 3 for the detailed architecture of the BPN's used here.

(a) 1995




(b) 1997




(c) 1999

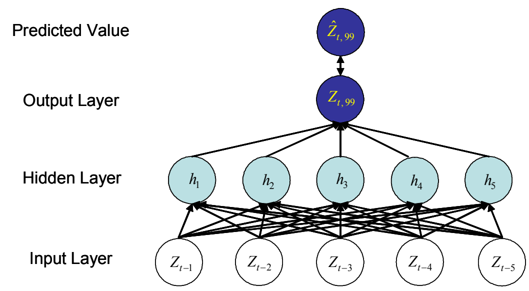
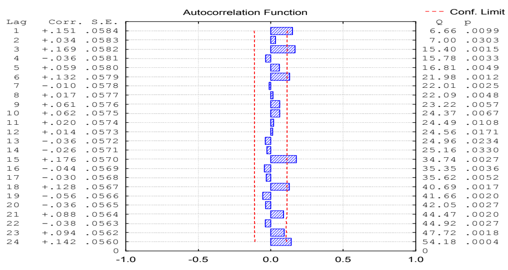


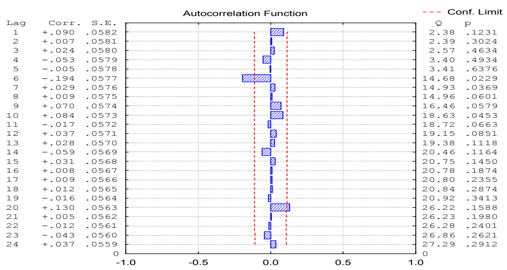

Figure 3. Architectures of Artificial Neural Networks. $h_i$ ($i$=1, 2, K ,5) means hidden node.

To check the effects of overfitting by ANN for complex problems, an ANN with validation set and an ANN without validation set are trained for each year. Since validation set is set up to avoid the overfitting in the training data, an ANN without validation set is likely to be overfitted to the training data. SACF of residuals for each year are given in Figure 4 for the ANN with validation set and in Figure 5 for the ANN without validation set. It is easy to notice that residual SACF's are improved significantly by an ANN without validation set across years. Indeed, when one compares residual SACF of an ANN without validation set with that of an ANN with validation set for each year, residual SACF of an ANN without validation set looks more independent than its counterpart. In particular, 1999 – the most complex year – reports the most noticeable improvement. Another measure of fitting capacity, monthly residual squared errors (RSE) for each year are given in Figure 6. Again improvement by ANN without validation set could be noticed, though the improvement is relatively mild, compared to SACF. Also, 1999 stands out in monthly RSE improvement by ANN without validation set though its absolute magnitude exceeds other years. In Figure 6, "w/ Val." denotes "with validation set" and "w/o Val." "without validation set".

(a) 1995




(b) 1997




(c) 1999

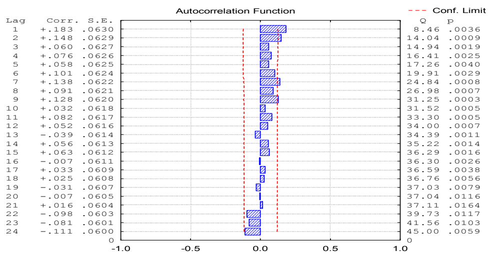
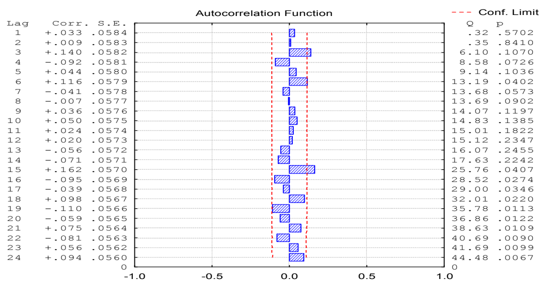
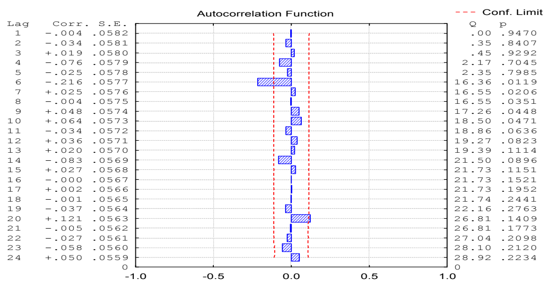



Figure 4. Residual SACF of KOSPI (ANN with validation set)

(a) 1995




(b) 1997




(c) 1999

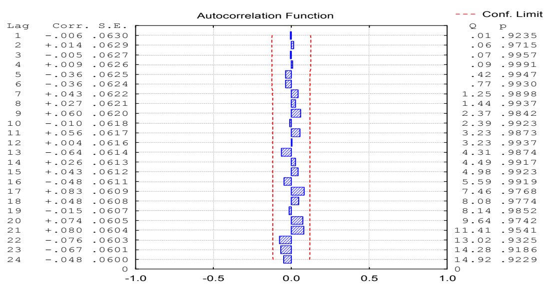
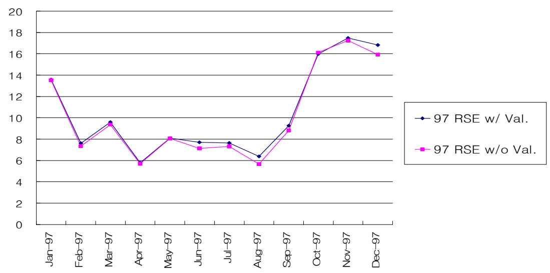



Figure 5. Residual SACF of KOSPI (ANN without validation set)

(a) 1995


(b) 1997




(c) 1999

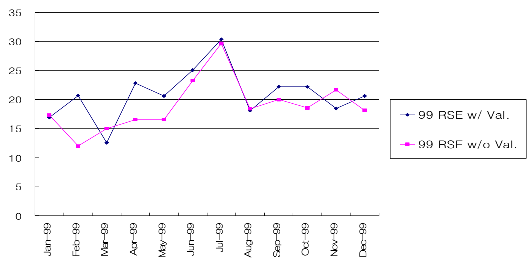



Figure 6. Monthly RSE

From our experiments, it appears that the complex feature of financial time series data may be traced reasonably well by allowing overfitting by an ANN which could be effectively done by employing a rather complicated auto-regressive model and removing the validation set in the training process. In summary our finding is that though fitting is hard for data with complex movements, overfitting may ease the problem significantly.

### 3. Concluding Remarks

In the analysis of financial time series data, stationarity is a desirable assumption. We focus on the case that this assumption does not hold any more (i.e., complex time series) and consider an ANN in the analysis of such cases. Using the empirical example of KOSPI, we find an intuitive and interesting result that overfitting by an ANN (ANN without validation set combined with a rather complicated auto-regressive model) might be a key to success for complex financial time series analysis.

## References

Anderson, T.W. (1971), *The Statistical Analysis of Time Series*, Wiley, New York.

Burgess, A.N., & Refenes, A-P,N. (1999), Modeling non-linear moving average processes using neural networks with error feedback: An application to implied volatility forecasting, *Signal Processing*, 74(1), 89-99.

Deco, G., Neuneier R., & Schurmann, B. (1997), Non-parametric data selection for neural learning in non-stationary time series, *Neural Networks*, 10(3), 401-407.

Ho, S.L., Xie, M., & Goh, T.N. (2002), A comparative study of neural network and Box-Jenkins ARIMA modeling in time series prediction, *Computers & Industrial Engineering*, 42(2-4), 371-375.

Lee, K.C., & Oh, S.B. (1996), An intelligent approach to time series identification by a neural network-driven decision tree classifier, *Decision Support Systems*, 17(3), 183-197.

Patterson, D.W. (1996), *Artificial Neural Networks*, Prentice Hall, New Jersey.

Trippi, R.R., & Turban, E. (1996), *Neural Networks in Financial and Investing*, Irwin, Chicago.

Yu, B. (1993), Density estimation in the $L^{\infty}$ norm for dependent data with applications to the Gibbs sampler, *The Annals of Statistics*, 21(2), 711-735.

Zhang, G.P. (2003), Time series forecasting using a hybrid ARIMA and neural network model, *Neurocomputing*, 50, 159-17