

# Neural Network Model vs. SARIMA Model In Forecasting Korean Stock Price Index (KOSPI)

Kyungjoo Lee, Cheju National University, [mrlove@cheju.ac.kr](mailto:mrlove@cheju.ac.kr).  
Sehwan Yoo, University of Maryland Eastern Shore, [syoo@umes.edu](mailto:syoo@umes.edu)  
John Jongdae Jin\*, University of Maryland-Eastern Shore, [jongdaejin@hotmail.com](mailto:jongdaejin@hotmail.com)

## ABSTRACT

*The purpose of this study is to compare the forecasting performance of a neural network (NN) model and a time-series (SARIMA) model in Korean Stock Exchange. In particular, we investigate whether the back-propagation neural network (BPNN) model outperforms the seasonal autoregressive integrated moving average (SARIMA) model in forecasting the Korea Composite Stock Price Index (KOSPI) and its return. Forecasting performance is evaluated by the forecasting accuracy of each model.*

*KOSPI data and its return data over the 390 week (89 month) period extending from January 1999 to May 2006 are analyzed. We find the followings: first, the SARIMA model generally provides more accurate forecasts for the KOSPI than the BPNN model does. This relative superiority of the SARIMA model over the BPNN model is pronounced for the mid-range forecasting horizons. Second, the BPNN model is generally better than the SARIMA model in forecasting the KOSPI returns. However, the difference in forecasting accuracies of the two models is not statistically significant. These results hold for both weekly and monthly data, and are robust across different measures of forecasting accuracy.*

## Keywords:

## INTRODUCTION

The ability to forecast the capital market price index is critical to individual investors and financial analysts as well. Among many forecasting models for stock prices and market price index, the neural network (NN) model has been gaining its popularity in recent years (E.G., Ansari et. al. (1994), Hamid et. al. (2004), Huang et. al. (2005), Kumar et. al. (2006), Malik et. al. (2006), Stansell et. al. (2004), Trinkle et. al. (2005)). Major reasons for the NN model's popularity in capital market forecast are twofold. First, the NN model is data driven method which learns from sample data and hence does not require any underlying assumptions about the data. Thus, the model is known as a universal functional

approximate without severe model misspecification problems due to wrong assumptions (Hornik et. al. (1989)). The model is also outstanding in processing large amount of fuzzy, noisy, and unstructured data. For example, Hutchinson et. al. (1994) examine stock option price data and show that the NN model is computationally less time consuming and more accurate non-parametric forecasting method, especially when the underlying asset pricing dynamics are unknown or when the pricing equation cannot be solved analytically. Second, stock price data are large, highly complex and hard to model because the pricing dynamics are unknown, which suits the NN model.

Korean stock market is considered more volatile than its US counterpart and hence has fuzzier and unstructured price data, which suits the NN model well.

Thus, it is meaningful endeavor to examine how well the NN model performs in forecasting more volatile Korean market data relative to a conventional seasonal autoregressive integrated moving average (SARIMA) model which is one of the most popular forecasting models in capital market studies. The purpose of this study is to compare the ability of the NN model and that of SARIMA model in forecasting Korean Stock Price Index (KOSPI) and its returns. Weekly and monthly data of KOSPI are analyzed in this study.

The remainder of this paper is organized as follows. Sample data and methodology are discussed in the next chapter, which is followed by discussions on empirical results. The concluding remarks are presented in the last chapter.

## INDEX DATA

The data used in this study are KOSPI for closing prices from the Korean Stock Exchange (KSE) data base. The data series span from 4<sup>th</sup> January 1999 to 29<sup>th</sup> May 2006, totaling 390 weeks (89 months) of

observations.<sup>1</sup> The returns data ( $R_t$ ) are defined as the continuously compounded returns on the price in the following way:  $R_t = \ln(P_t/P_{t-1})$ , where  $P_t$  is KOSPI in period  $t$ .

The data are divided into two sub-periods, one for the estimation and the other for the forecasting. We use four different forecasting periods to examine the potential impact of forecasting horizons on the forecasting accuracy. Forecasting horizons used are 20% (long range), 13% and 8% (mid range), and 6% (short range) of the total number of observations. For the long range forecasting, the first 313 weekly data are used for model identification and estimation, while the remaining 77 weekly data (about 20% of 390 weeks) are reserved for evaluating the performance of SARIMA model and NN model. Other forecasting periods were defined in the similar way, resulting in mid range (31-50 weeks ahead) and short range (23 weeks ahead) forecasting horizons.<sup>2</sup>

### NEURAL NETWORK FORECASTING

In this study, one of the widely used NN model called the back-propagation neural network (BPNN) is used for time series forecasting. The main reason for adopting BPNN is that BPNN is an efficient way to calculate the partial derivatives of the networks error function with respect to the weights and hence to develop a network model that minimizes the discrepancy between real data and the output of the network model.

BPNN can be trained using the historical data of a time series in order to capture the non-linear characteristics of the specific time series. The model parameters will be adjusted iteratively by a process of minimizing the forecasting errors. For time series forecasting, the relationship between output ( $y_t$ ) and the inputs ( $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ ) can be described by the following mathematical formulae.

<sup>1</sup> KOSPI data is available since the opening of the Korean Options Exchange for stock price index in July 1997. We exclude two year's data (1997-1998) because the Korean stock market had suffered the severe financial crisis called IMF crisis during this period.

<sup>2</sup> For monthly data, we used three forecasting horizons (20%: long range, 13%: mid range, 8%: short range) due to the problem of small observations when applying the ratio of 6%.

$$y_t = a_o + \sum_{j=1}^q a_j f(w_{oj} + \sum_{i=1}^p w_{ij} y_{t-i}) + e_t \tag{1}$$

Where  $a_j$  ( $j = 0, 1, 2, \dots, q$ ) is a bias on the  $j$ th unit, and  $w_{ij}$  ( $i = 0, 1, 2, \dots, p; j = 0, 1, 2, \dots, q$ ) is the connection weights between layers of the model,  $f(\cdot)$  is the transfer function of the hidden layer,  $p$  is the number of input nodes and  $q$  is the number of hidden nodes. The BPNN model performs a nonlinear functional mapping from the past observation ( $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ ), to the future value ( $y_t$ ), i.e.,

$$y_t = \varphi(y_{t-1}, y_{t-2}, \dots, y_{t-p}) + e_t \tag{2}$$

Where  $w$  is a vector of all parameter and  $\varphi$  is a function determined by the network structure and connection weights.

The NeuroSolutions NBuilder toolbox is used to train data for model developments and test data for forecasting accuracy of the models developed. Here the stop criteria for the supervised training of the networks are specified as follows. The maximum epochs specify how many number of iterations (over the training set) will be done if no other criterion kicks in. The training terminates when one of the following four conditions is met (1) mean square error of the validation set begins to rise, indicating that over fitting might be happening.; (2) when the threshold is less than 0.01, i.e., we are on effectively flat ground; (3) when training time has reached 1000 epochs; (4) the goal (difference between output and target is less than 0.01) has been met. After the training was completed, its epochs and the simulation procedure completed successfully, which indicated the network was trained was predicting the output as desired.

According to the principle of Ockham's razor, the simplest networks topology yielding satisfactory results is used. The networks are created as '2-1-1': that is, two input layers, one hidden layer, and one output layer. The network is also trained using various other topologies such as 2-X-1, while X = 2, 3, 4, and 5. However, the best results are obtained when there is one hidden layer (i.e., X=1).

### TIME-SERIES FORECASTING

To obtain the KOSPI forecasts from the SARIMA model, we adopted the Box and Jenkins' method. Basically, Box and Jenkins' method uses the following three-stage approach to select an

appropriate model for the purpose of estimating and forecasting a time-series data.

*Identification:* we used the SARIMA procedure in SAS statistical software to determine plausible models. The SARIMA procedure uses standard diagnostics such as plots of the series, autocorrelation function (ACF), inverse autocorrelation function, and partial autocorrelation function (PACF).

*Estimation:* Each of the tentative models is fit and the various coefficient estimates are examined. The estimated models are compared using standard criteria such as Akaike Information Criteria and the significance level of coefficients.

*Diagnostic checking:* SARIMA procedure is used to check if the residuals from the different models are white noise. The procedure uses diagnostics tests such as ACF, PACF, and Ljung-Box Q-statistics for serial correlation.

Applying these steps, SARIMA (110)(12) for the KOSPI price series, and SARIMA(011)(12) for the KOSPI return series are selected as forecasting models for both weekly and monthly data.

## MEASUREMENT OF FORECAST ACCURACY

Forecast error (FE) is determined by subtracting forecasted value from actual value of KOSPI (price or return), and then deflating the difference by the absolute value of actual data as follows:

$$FE_t = (A_t - F_t) / |A_t|,$$

where  $A_t$  = actual value of KOSPI in period  $t$ .

$F_t$  = forecasted value of KOSPI in period  $t$ .

The reason for using the absolute value of actual KOSPI as deflator is to correct for negative values.<sup>3</sup>

Accuracy of a forecast model is measured by sign-neutral forecast error metrics such as absolute forecast errors (AFE) and squared forecast errors (SFE).<sup>4</sup>

## RESULTS

### Using Weekly Data

<sup>3</sup> Although there is no negative price, returns data could have negative values.

<sup>4</sup> AFE assumes that the user's loss function is linear, while SFE assumes quadratic loss function.

Descriptive statistics of forecast errors (FE) from NN model and SARIMA model are presented in Table 1. Panel A of Table 1 shows the descriptive statistics of KOSPI price FE for different forecasting horizons. The SARIMA model provides smaller mean (median) values of FE than the NN model for every forecasting horizon. With respect to FE for KOSPI return, Panel B of Table 1 indicates that the SARIMA model has larger mean and standard deviation, but smaller median than the NN model does. These results suggest that the forecasting ability of each model may not be consistent, but differ depending on whether KOSPI or its return data is forecasted.

<Insert Table 1>

Table 2 presents the results of comparing the models (NN vs. SARIMA) based on their accuracy in forecasting KOSPI price. Panel A of Table 2 summarizes mean AFE (MAE), t-test statistics, Wilcoxon rank sum test statistics, and the associated significance level (p-value) for different forecasting horizons. For every forecasting horizon, the SARIMA model exhibits smaller MAE than the NN model. However, the differences are statistically significant (at  $\alpha < 0.001$ ) only for mid range (31-50 weeks ahead) forecasting horizons, using both the t-test and the nonparametric Wilcoxon test. Panel B of Table 2 provides the results using SFE as a forecasting accuracy metric. The results are basically the same as those in Panel A.

<Insert Table 2>

Overall, these results indicate that the SARIMA model provides more accurate forecasts of KOSPI than the NN model, especially for 31-50 weeks ahead (mid-range) forecasts. These results are consistent with the findings of Wang et. al..

Table 3 shows the results of comparing the model accuracy in forecasting KOSPI returns. Mean values of forecast accuracy metrics (AFE in Panel A and SFE in Panel B), test statistics (t-test and Wilcoxon test), and corresponding significance levels are presented in Table 3.

<Insert Table 3>

The results in Table 3 show that the NN model is better than the SARIMA model in forecasting KOSPI returns though not statistically significant. This finding is robust across different forecasting horizons and forecast error measures.

## Using Monthly Data

Table 4 shows the descriptive statistics of forecast errors (FE) from the NN model and the SARIMA model using KOSPI data (Panel A) and its return data (Panel B). Consistent with the results from weekly data, the SARIMA model exhibits smaller mean values of FE in forecasting KOSPI than the NN model for every forecasting horizon, while the SARIMA model exhibits higher mean values of FE in forecasting KOSPI returns than the NN model for every forecasting horizon. Median and standard deviation also show similar patterns.

<Insert Table 4>

Table 5 presents the results of evaluating the models (NN vs. SARIMA) based on their accuracy in forecasting KOSPI. Panel A of Table 5 summarizes mean AFE (MAE), Wilcoxon rank sum test statistics, and the associated significance level (p-value) for different forecasting horizons.<sup>5</sup> For every forecasting horizon, the SARIMA model exhibits smaller MAE than the NN model. However, the differences are statistically significant (at  $\alpha < 0.01$ ) only for 11 month ahead forecasts. Panel B of Table 5 provides the results using SFE as a forecasting accuracy measure. The results are basically the same as those in Panel A.

<Insert Table 5>

Overall, these results in Tables 4 and 5 indicate that the SARIMA model provides more accurate forecasts of KOSPI than the NN model, especially for mid-range (11 month ahead) forecasts. These results are consistent with those from weekly data and those of Wang et. al..

Table 6 shows the results of comparing the model accuracy in forecasting KOSPI returns. Table 6 provides mean values of forecast accuracy metrics (AFE in Panel A and SFE in Panel B), Wilcoxon test statistics, and corresponding significance levels.

<Insert Table 6>

Like results from weekly data, the results from monthly data in Table 6 show that the NN model is

<sup>5</sup> Due to the problem of small sample observations, the parametric t-test results would be misleading and hence not reported for monthly data. However, the parametric results are basically the same as those from nonparametric Wilcoxon test.

better than the SARIMA model in forecasting KOSPI returns though not statistically significant. This finding is robust across different forecasting horizons and forecast error measures.

## CONCLUSIONS

The purpose of this study is to compare the forecasting performance of a neural network (NN) model and a time-series (SARIMA) model in Korean Stock Exchange. Using KOSPI data and its return data over the 390 week (89 month) period extending from January 1999 to May 2006, we find the followings: first, the SARIMA model generally provides more accurate forecasts for the KOSPI than the BPNN model does. Second, the BPNN model is generally better than the SARIMA model in forecasting the KOSPI returns.

These results may provide useful insight to risk management of KOSPI based mutual funds and derivatives where portfolio formation, hedging, option strategies, and investment decisions are made using predictions of KOSPI and its returns. Due to its ability to discover patterns in nonlinear, uncertain, fuzzy, and volatile data, NN has become a very popular prediction method of stock market variables. The results of this study may add some favorable evidence for NN model in a sense that the model is also working well in predicting certain Korean capital market variables which are more chaotic than those of U.S. capital market.

Since the accuracy of forecasting values is dependent on the developing process of forecasting models, the results of this study are also sensitive to the developing process of the BPNN model and SARIMA model.

## REFERENCES

1. Ansari, A. M. and Frank C. Lin (1994). Forecasting financial trends in stock markets with neural networks. *Institute of Electrical and Electronics Engineers*, pp188-196.
2. Hamid, S. and Z. Iqbal (2004). Using neural networks for forecasting volatility of S & P 500 Index future prices. *Journal of Business Research*, Vol. 57-10, pp 1116.
3. Hornik, K., M. Stinchcombe, & H. White (1989). Multi-layered feedforward networks are universal approximators. *Neural Networks*, Vol. 2, pp.359-366.
4. Huang, W., Y. Nakamori, and S. Wang (2005). Forecasting stock market movement

direction with support vector machine. *Computers & Operations Research* Vol. 32-10, pp 2513.

5. Hutchinson, J., W. Andrew, & t. Poggio (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *Journal of Finance*, Vol. 49. pp. 851-889.
6. Kumar, K. and S. Bhattacharya (2006). Artificial neural network vs. linear discriminant analysis in credit ratings forecast: a comparative study of prediction performances. *Review of Accounting & Finance* Vol. 5-3, pp 216.
7. Malik, F. and M. Nasereddin (2006). Forecasting output using oil prices: a cascaded artificial neural network approach. *Journal of Economics and Business* Vol. 58, pp 168-180.
8. Medeiros, M., T. Terasvirta, and G. Rech (2006). Building neural network models for time series: a statistical approach. *Journal of Forecasting* Vol. 25-1, pp 49.
9. Stansell, S. and S. Eakins (2004). Forecasting the direction of change in sector stock indexes: An application of neural networks. *Journal of Asset Management*, Vol. 5-1, pp 37-48.
10. Tang, y., F. Xu, X. Wan, and Y. Zhang (Working paper). Web-based fuzzy networks for stock prediction. *Georgia State University*.
11. Trinkle, B. (2005). Forecasting annual excess stock returns via an adaptive network-based fuzzy inference system. *Intelligent System in Accounting, Finance, and Management*, Vol. 13-3, pp 165.
12. Wang, W., D. Okunbor, and F. Lin (Working paper). Future trend of the shanghai stock market. *University of Maryland-Eastern Shore*.

Table 1

Descriptive Statistics for Forecast Errors: Using weekly data

Panel A: KOSPI				
Forecasting Horizon (weeks)	Model	Mean	Standard Deviation	Median
77	NN	0.121	0.093	0.109
	SARIMA	0.093	0.071	0.102
50	NN	0.171	0.073	0.201
	SARIMA	0.107	0.056	0.123
31	NN	0.099	0.027	0.089
	SARIMA	0.038	0.048	0.054
23	NN	0.106	0.027	0.108

	SARIMA	-	0.074	-0.084
	A	0.090		

Panel B: KOSPI Return

Forecasting Horizon (weeks)	Model	Mean	Standard Deviation	Median
77	NN	0.184	1.134	0.891
	SARIMA	0.246	1.759	0.505
50	NN	0.104	1.304	0.655
	SARIMA	0.035	1.811	0.430
31	NN	0.075	1.413	0.580
	SARIMA	0.147	1.606	0.442
23	NN	0.705	1.743	-0.606
	SARIMA	1.577	2.781	-1.091

- 1) Forecast errors are defined as:  $(A-F)/|A|$ , where A=actual value of KOSPI or return, F=forecast value of KOSPI or return.
- 2) The number of weeks used to estimate forecasting models (NN, SARIMA) is different depending on forecasting horizons. For example, 313 weeks (367 weeks) of data were used to obtain forecasts for the 77 weeks (23 weeks) ahead. Each forecasting horizon represents roughly 20%, 13%, 8% and 6% of total number of observations (390 weeks).
- 3) NN= Back-propagation neural network model without hidden layer. SARIMA= Best-fitted SARIMA model: (110)(12) for index and (011)(12) for return.

Table 2

Comparison of Forecasting Accuracy for KOSPI: Using weekly data

Panel A: Mean Absolute Errors (MAE)				
Model	Forecasting Horizons (Weeks)			
	77	50	31	23
NN	0.123	0.171	0.099	0.106
SARIMA	0.099	0.108	0.055	0.095
t-stat (p-value)	1.940 * (0.055)	4.890*** (0.001)	6.500 *** (0.001)	0.730 (0.469)
z-stat (p-value)	1.489 (0.137)	4.236*** (0.001)	5.125 *** (0.001)	1.274 (0.203)

Panel B: Mean Square Errors (MSE)

Model	Forecasting Horizons (Weeks)			
	77	50	31	23
NN	0.023	0.034	0.010	0.012
SARIMA	0.014	0.014	0.004	0.013
t-stat (p-value)	3.020 *** (0.004)	5.510 *** (0.001)	5.990 *** (0.001)	0.340 (0.734)
z-stat (p-value)	1.489 (0.137)	4.236 *** (0.001)	5.125 *** (0.001)	1.274 (0.203)

- 1) MAE=  $1/n \sum |FE|$ . Forecast errors (FE) are defined as:  $(A-F)/|A|$ , where A=actual value of KOSPI index or return, F=forecast value of KOSPI or return.
- 2) MSE=  $1/n \sum (FE)^2$ . Forecast errors (FE) are defined as:  $(A-F)/|A|$ , where A=actual value of KOSPI index or return, F=forecast value of KOSPI or return.
- 3) Wilcoxon rank sum test statistic.

\* : significant at  $\alpha < 0.10$ ; \*\* : significant at  $\alpha < 0.05$ ;  
\*\*\*: significant at  $\alpha < 0.01$

Table 3 Comparison of Forecasting Accuracy For KOSPI Returns: Using weekly data

Panel A: Mean Absolute Errors (MAE)

Model	Forecasting Horizons (Weeks)			
	77	50	31	23
NN	1.082	1.110	1.098	1.342
SARIMA	1.243	1.309	1.205	1.953
t-stat (p-value)	1.070 (0.287)	1.000 (0.324)	0.430 (0.667)	1.030 (0.313)
z-stat (p-value)	1.328 (0.184)	0.093 (0.926)	0.028 (0.978)	0.527 (0.598)

Panel B: Mean Square Errors (MSE)

Model	Forecasting Horizons (Weeks)			
	77	50	31	23
NN	1.304	1.678	1.962	3.404
SARIMA	3.114	3.217	2.518	9.885
t-stat (p-value)	1.650 (0.104)	1.430 (0.158)	0.480 (0.634)	1.210 (0.237)
z-stat (p-value)	1.328 (0.184)	0.093 (0.926)	0.028 (0.978)	0.527 (0.598)

- 1) Return= $\ln(P_t/P_{t-1})$ , where  $P_t$ =KOSPI in week t.

Table 4  
Descriptive Statistics for Forecast Errors: Using monthly data

Panel A: KOSPI

Forecasting Horizon (months)	Model	Mean	Standard Deviation	Median
17	NN	0.140	0.094	0.133
	SARIMA	0.115	0.065	0.124
11	NN	0.191	0.065	0.211
	SARIMA	0.107	0.056	0.127
7	NN	0.119	0.025	0.130
	SARIMA	0.111	0.027	0.102

Panel B: KOSPI Returns

Forecasting Horizon (months)	Model	Mean	Standard Deviation	Median
17	NN	0.129	1.101	-0.662
	SARIMA	0.238	1.429	-0.004
11	NN	0.307	1.166	0.416
	SARIMA	0.313	1.465	-0.004
6	NN	0.630	1.151	0.845
	SARIMA	1.607	2.118	1.159

- 1) Forecast errors are defined as:  $(A-F)/|A|$ , where A=actual value of KOSPI or return, F=forecast value of KOSPI or return.
- 2) The number of weeks used to estimate forecasting models (NN, SARIMA) is different depending on forecasting horizons. For example, 72 months (78 months) of data were used to obtain forecasts for the 17 months (11 months) ahead. Each forecasting horizon represents roughly 20%, 13% and 8% of total number of observations (89 months).
- 3) NN= Back-propagation neural network model with one hidden layer.
- 4) SARIMA= Best-fitted SARIMA models(110)(12) for index and (011)(12) for returns

Table 5  
Comparison of Forecasting Accuracy for KOSPI: Using monthly data

Panel A: Mean Absolute Errors (MAE)

Model	Forecasting Horizons (Months)		
	17	11	7
NN	0.144	0.191	0.119
SARIMA	0.115	0.109	0.111
z-stat (p-value)	1.033 (0.302)	2.561 ** (0.010)	0.511 (0.609)

Table 6  
Comparison of Forecasting Accuracy for KOSPI  
Returns: Using monthly data

Panel B: Mean Square Errors (MSE)

Model	Forecasting Horizons (Weeks)		
	17	11	7
NN	0.028	0.040	0.015
SARIMA	0.017	0.014	0.013
z-stat (p-value)	1.033 (0.302)	2.561 ** (0.010)	0.511 (0.609)

\*: significant at  $\alpha < 0.10$ ; \*\*: significant at  $\alpha < 0.05$ ;  
\*\*\*: significant at  $\alpha < 0.01$

- 1) NN= Back-propagation neural network model with one hidden layer.
- 2) SARIMA= Best-fitted SARIMA models(110)(12) for index and (011)(12) for returns

Panel A: Mean Absolute Errors (MAE)

Model	Forecasting Horizons (Months)		
	17	11	7
NN	1.025	1.090	1.110
SARIMA	1.187	1.228	1.927
z-stat (p-value)	0.827 (0.408)	0.722 (0.470)	0.400 (0.689)

Panel B: Mean Square Errors (MSE)

Model	Forecasting Horizons (Weeks)		
	17	11	7
NN	1.158	1.329	1.500
SARIMA	1.979	2.049	6.321
z-stat (p-value)	0.827 (0.408)	0.722 (0.470)	0.400 (0.689)

\*: significant at  $\alpha < 0.10$ ; \*\*: significant at  $\alpha < 0.05$ ;  
\*\*\*: significant at  $\alpha < 0.01$

- 1) NN= Back-propagation neural network model with one hidden layer.
- 2) SARIMA= Best-fitted SARIMA models(110)(12) for index and (011)(12) for returns