Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 33 (2007) 916–922
Expert Systems with Applications
www.elsevier.com/locate/eswa

Forecasting the volatility of stock price index

Tae Hyup Roh *

Department of Business Administration, Seoul Women's University, 126 Gongreung-Dong, Nowon-Gu, Seoul, Republic of Korea

Abstract

Accurate volatility forecasting is the core task in the risk management in which various portfolios' pricing, hedging, and option strategies are exercised. Prior studies on stock market have primarily focused on estimation of stock price index by using financial time series models and data mining techniques. This paper proposes hybrid models with neural network and time series models for forecasting the volatility of stock price index in two view points: deviation and direction. It demonstrates the utility of the hybrid model for volatility forecasting. This model demonstrates the utility of the neural network forecasting combined with time series analysis for the financial goods.

Keywords: Volatility forecasting; Stock price index; Time series analysis; Neural network; Hybrid model

1. Introduction

Since the inception of KOSPI (Korea Composite Stock Price Index) 200 option market, July 7, 1997, the trading market volume has steeply increased. Investors and financial institutions became worried about risks caused by increasing volatility of KOSPI 200. Accurate volatility forecasting is an essential part in performing risk management, specifically in allocating assets to various portfolios in order to hedge those portfolios' risk efficiently.

There are various models to forecast time series volatilities. Engle suggested the ARCH(p) (Autoregressive Conditional Heteroscedasticity) model that has been used by many financial analysts (Engle, 1982). The GARCH (Generalized ARCH) model is the generalized form of ARCH (Bollerslev, 1986). RiskMetrics by JP Morgan suggested EWMA (Exponentially Weighted Moving Average) (JP Morgan & Reuters, 1996) which is basically a non-stationary GARCH(1, 1) model. By considering the limitation of the GARCH model, leverage effects, the EGARCH (Exponential GARCH) model was proposed (Nelson, 1991). These financial time series models can be analysed with econometrics and be systematically explained based on the market and financial theories. However, due to many noises that are caused by changes in market conditions and environments, financial analysts should consider many market variables. Financial time series models require strict assumptions about distributions of time series, so it is hard to reflect market variables directly in the models. Due to these shortcomings of the financial time series models, ANN (Artificial Neural Network) has been applied to various complex financial markets directly. The ANN model is a nonparametric method and can forecast future results by learning the pattern of market variables without any strict theoretical assumption. Taking advantage of these characteristics, it was revealed that ANN can outperform financial time series models by analysing S&P 500 future index option volatilities (Hamid & Zahid, 2002). The possibility of ANN was demonstrated in forecasting the volatilities of financial time series (Brooks, 1998).

The objective of this study is that the forecasting power in the stock price index domain can be enhanced by integrating financial time series models, such as EWMA, GARCH, EGARCH and ANN models. Prior studies have mainly compared the predicting power between single models. However, this study focuses on two viewpoints:

* Tel.: +82 2 970 5518; fax: +82 2 970 5961.
E-mail address: rohth@swu.ac.kr

0957-4174/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2006.08.001

the deviation and direction of stock price index. In addition, most of the prior studies have adjusted the weight of raw volatilities by repetitive trial and error methods of learning process and found the optimal coefficient of input variables to produce the best results. This study finds the coefficients of input variables by financial time series process and extracts new variables that greatly influence the results.

Having introduced the research background, the remaining sections of this paper are organized as follows: Section 2 provides a brief review of the related work. Sections 3 and 4 present the proposed methodology about the ANN-Financial time series models and experiments. The last two sections state results and the contribution of the research and future consecutive research issues.

2. Related work

2.1. Financial time series volatility forecasting

As a classic financial time series model, ARCH(p) was proposed to model the characteristics of time series that have volatility clustering and fat tail (Engle, 1982). Because ARCH(p) causes time lag p to get much larger in forecasting volatility, generalized ARCH as GARCH(p, q) was suggested (Bollerslev, 1986). GARCH model restricts its parameters to have plus value of conditional variance. These conditions would make the conditional variance process much restrictive over the necessity. The original GARCH model does not consider the negative correlation between future yield and volatility. By the same token, when the market is falling against participant's expectation (negative impact), a negative effect has a bigger influence than the same-sized positive effect. This asymmetric shock is generally called the leverage effect. On the other hand, the general GARCH model results in asymmetric shock without regard to the sign of the impact of conditional volatilities because the square of present yield's residuals has an influence in the future yield's volatility. By considering this leverage effects, the EGARCH model was developed (Nelson, 1991).

The EWMA model gives more weightage to the recent data than others included in time series. This method is modelled on "RiskMetrics" by JP Morgan in a parametric way. In the case of EWMA, when the objective of volatility forecasting is to catch the short-term movement of it, EWMA is desirable. However, if EWMA places much value only on the recent data, then it reduces the sample size and brings about the result of an increasing possibility of measurement error. EWMA has shortcomings in describing characteristics of the financial time series volatility – 'long-term memories'.

2.2. Stock market applications with data mining techniques

Many studies on stock market prediction using data mining techniques have been performed during the past decade. In early days of these studies, the focus was on estimating the level of the return on stock price index. One of the earliest studies, Kimoto, Asakawa, Yoda, and Takeoka (1990) used several learning algorithms and prediction methods for developing a prediction system for the TSEPI (Tokyo Stock Exchange Prices Index). They used the modular neural network to learn the relationships among various market factors. They concluded that the correlation coefficient produced by their model is much higher than that produced by multiple regression.

Some researchers investigated the issue of predicting the stock index futures market. Trippi and DeSieno (1992) predicted the daily direction of change in the S&P 500 index futures using ANN. They combined the outputs of individual networks using logical (Boolean) operators to produce a set of composite rules. They suggested that their best composite synthesized rule set system achieved a higher gain than the previous research. Recent research tends to hybridize several AI techniques. Nikolopoulos and Fellrath (1994) developed a hybrid expert system for investment advising. In their study, genetic algorithms were used to train and configure the architecture of investor's neural network component. A more recent study of Lee and Jo developed an expert system, which uses knowledge in a candlestick chart analysis (Lee & Jo, 1999). The expert system had patterns and rules, which could predict future stock price movements. The experimental results revealed that a developed knowledge base could provide excellent indicators. In addition, Tsaih, Hsu, and Lai (1998) integrated a rule-based technique and ANN to predict the direction of change of the S&P 500 stock index futures on a daily basis. In addition, some researchers searched the connection weights of ANN using the GA instead of local search algorithms including a gradient descent algorithm. They suggested that the global search techniques including the GA might prevent ANN from falling into a local optimum (Gupta & Sexton, 1999; Kim, 2003; Kim & Han, 2000; Sexton, Dorsey, & Johnson, 1998).

Past studies proved that the ANN model can enhance the predictive power based on real market data and suggested that it is important to integrate the ANN model and financial time series models for future studies (Hamid & Zahid, 2002). Especially, Hu and Tsoukalas (1999) proposed that predictive power can be improved by ANN in which the forecasted results are relearned in the ANN learning process, and he suggested that forecasted results should be studied if they could contribute to the ANN forecasting process. Recently, various studies using ANN have been developed in the fields of forecasting stock index (Armano, Marchesi, & Murru, 2005; Gavrishchaka & Ganguli, 2003; Yao, Li, & Tan, 2002).

These various studies have proposed the basis on which ANN could be applied to financial engineering and nowadays active researches are developed in the fields of finance and forecasting in which parametric models cannot fully explain the characteristics of their market behaviors.

3. Hybrid ANN-time series model

A common shortcoming of ANN in forecasting volatility is that it is not proved econometrically. Also, users get curious about the ANN results because input variables which are the most important factors to influence the forecasted results are determined in the learning process of repetitive trials and errors. In this study, the hybrid ANN-time series model is proposed to solve the difficulties in the ANN learning process and to enhance the predictive power in forecasting volatility.

The hybrid model achieves the efficiency in selecting input variables because they are selected and newly created by the financial time series models. Repetitive trial error process could be eliminated to a one time financial time series process. Input variables that are most weighted on the forecasting are usually contracted to one or two variables. The ANN models have spent most of the time to find out input variables by repetitive trial and errors. This study will prove that hybrid models can improve the predictive power in the framework of both direction and deviation.

Theoretically, ANN can approximate any function and financial time series models can be fairly approximated by this flexibility. Therefore, the characteristic of conditional volatility by GARCH(1,1) can be approximated through ANN and the methodology of approximation is achieved by inputting variables obtained through GARCH (1,1) process and learning the conditional volatility pattern through ANN process. ANN can adjust these results by using other market variables realistically to reflect real market behaviours. Input variables can be extracted efficiently by financial time series models and ANN can improve the predictive power by using these variables and other market variables from the perspective of deviation and direction of stock price index.

3.1. NN-EWMA model

The sNN-EWMA model is to give more weightage on recent data and catch the short-term volatility behaviours by extracting variables through the EWMA model. The equation of EWMA is as follows:

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)\varepsilon_{t-1}^2. \tag{1}$$

EWMA gives more weightage on recent data and reduces the shadow effect in which past data severely influences the prediction. NN-EWMA models can be expressed with two variables:

- σ_{t-1}^2: square of volatility at $t-1$,
- ε_{t-1}^2: square of residuals at $t-1$.

Two newly created input variables can be extracted based on the above variables and each variable is adjusted by a smoothing factor (decay factor) λ, $(1-\lambda)$ and then included in the ANN input variables. Newly created variables are as follows:

- $\sigma_{t-1}^{2\prime} = \lambda\sigma_{t-1}^2$,
- $\varepsilon_{t-1}^{2\prime} = (1-\lambda)\varepsilon_{t-1}^2$.

$\sigma_{t-1}^{2\prime}$ and $\varepsilon_{t-1}^{2\prime}$ that are extracted through the EWMA model are similar with variables that are extracted in the GARCH(1,1) model, but two models started from a totally different conditional variance concept, namely, the coefficients of the GARCH(1,1) model are extracted statistically through financial time series, but those of EWMA are determined by the user's discretion based on experience.

3.2. NN-GARCH model

ANN-time series models are used to extract predictive or adjusted input variables by financial time series models. Most of the financial time series models are known to be easily modelled by GARCH(1,1), so in the first place, this paper will try to extract input variables using the GARCH (1,1) model. The GARCH(1,1) process from time series data is as follows:

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2. \tag{2}$$

The GARCH(1,1) model brings about a similar effect like using the long time lag ARCH model even if it uses a small number of parameters. Therefore, it is desirable to use the GARCH(1,1) model to make a time series which has the characteristics of volatility, clustering and fat tail from the perspective of conditional variances. σ_t^2 is the one-period ahead forecast conditional variance. This conditional variance equation is expressed with following three variables:

- α_0: nonconditional volatility coefficient,
- ε_{t-1}^2: residual at $t-1$,
- σ_{t-1}^2: square of volatility at $t-1$.

Consequently, ε_{t-1}^2 and σ_{t-1}^2 which have conditional relations with each other can be extracted and the coefficients of these variables are adjusted to α_1 and β_1. These are included in the input variables for the ANN learning process. The newly extracted variables are as follows:

- $\sigma_{t-1}^{2\prime} = \beta_1\sigma_{t-1}^2$,
- $\varepsilon_{t-1}^{2\prime} = \alpha_1\varepsilon_{t-1}^2$.

3.3. NN-EGARCH model

The NN-EGARCH model is to improve the restrictive conditional variance process of the GARCH(1,1) model, which is like following equation:

$$\ln\sigma_t^2 = \alpha + \beta\ln\sigma_{t-1}^2 + \gamma\left(\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}} - \sqrt{\frac{2}{\pi}}\right|\right) + \omega\frac{\varepsilon_{t-1}}{\sigma_{t-1}}. \tag{3}$$

The EGARCH model can explain the asymmetric shock which means that a large falling in the yield can make the next period volatility greater-leverage effect. This $\frac{\varepsilon_{t-1}}{\sigma_{t-1}}$ term in Eq. (3) explains the leverage effect by the market shock. The EGARCH model can statistically explain an asymmetric shock which cannot be explained by the conditional variance model—GARCH. The EGARCH model is expressed with the following four variables:

- α: nonconditional variance coefficient,
- $\ln\sigma^2_{t-1}$: log value of variance at $t-1$,
- $\left(\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}} - \sqrt{\frac{2}{\pi}}\right|\right)$: Asymmetric shock by leverage effect,
- $\frac{\varepsilon_{t-1}}{\sigma_{t-1}}$: leverage effect.

New input variables can be extracted based on the above variables and each variable is adjusted by β, γ, ω and then gets included in the ANN model:

- $\ln\sigma^{2\prime}_{t-1} = \beta\ln\sigma^2_{t-1}$,
- $\mathrm{LE}(\text{leverage} - \text{effect}) = \gamma\left(\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}} - \sqrt{\frac{2}{\pi}}\right|\right)$,
- $L(\text{leverage}) = \omega\frac{\varepsilon_{t-1}}{\sigma_{t-1}}$.

Existing studies use observable market variables and various others that can be created by financial theories but usually not use newly created variables for ANN. It is the newly created leverage and leverage effect variables that differ from the existing ANN models in the NN-EGARCH model. To specify the efficiency and predictive power by newly created variables, the NN-EGARCH model would make $\ln\sigma^{2\prime}_{t-1}$, LE, L input variable with other market variables and analyse the predictive power during the NN-EGARCH learning process.

4. Experiment

4.1. Dataset and experimental setup

This study conducted experiments to evaluate the proposed model. For experiments, we used KOSPI 200 time series data and option daily trading materials, provided by the KSE KOSPI 200 index database and daily option data. The dataset consists of 930 trading days' indexes for the sample period and 160 indexes for the prediction period.

To verity the appropriateness of financial time series models, this study performs the ADF (Augmented Dickey-Fuller) test and the ARCH LM (Lagrange Multiplier) test. Fig. 1 shows the daily KOSPI 200 index for the experimental dataset.

4.1.1. ADF test

The ADF test is used for verifying stability. Because there is a unit root in the sample data by ADF test, this time series is a nonstationary time series that needs to be differentiating. Therefore, this time series is changed to KOSPI 200 logarithmic yields time series that is stationary.

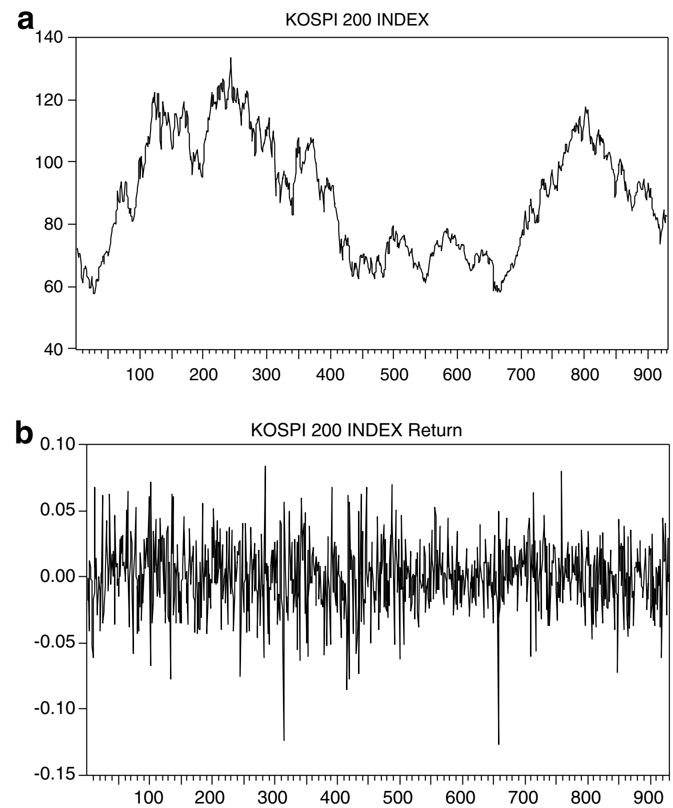

Fig. 1. (a) KOSPI 200 time series; (b) logarithmic time series of KOSPI 200 returns.

Table 1
ADF test of KOSPI 200 index time series datasets

ADF test statistic	−1.929573	1% Critical value*	−3.4401
		5% Critical value	−2.8651
		10% Critical value	−2.5687

The logarithmic value of index returns is calculated. As a result of the ADF test (Table 1), this transformed time series is stationary that is oscillating around a mean value 0.

4.1.2. ARCH LM test

The ARCH LM test is needed to verify the heteroscedasticity in time series. The null hypothesis that there is no heteroscedasticity to 5 time lags is rejected by a 5% significance level, so there is a heteroscedasticity within 5 time lags. At the same time, joint significance of lagged squared residual to 5 time lags is accepted by F-statistics. Therefore, the ARCH model is appropriate for modelling the volatility of KOSPI 200 return, and the GARCH(1, 1) model is

Table 2
Results of 5-lag ARCH-LM test

ARCH test: 5-lag			
F-statistic	2.345315	Probability	0.039580
Obs* R-squared	11.65435	Probability	0.039844

Table 3
Expected learning effects of extracted input variables

Models	Extracted variables and expected learning effects
NN	• σ^2_{t-1}: Simple volatility learning at $t-1$ without repetitive trial and error
NN-GARCH	• $\sigma^{2\prime}_{t-1} = 0.836611\sigma^2_{t-1}$: Learning the pattern of conditional volatility between one period time intervals • $\varepsilon^{2\prime}_{t-1} = 0.064785\varepsilon^2_{t-1}$: Learning the residual effects to learn the conditional volatility
NN-EGARCH	• $\ln\sigma^{2\prime}_{t-1} = 0.900850\ln\sigma^2_{t-1}$: Learning the pattern of conditional volatility between one period time intervals • $L(\text{leverage}) = -0.062328\frac{\varepsilon_{t-1}}{\sigma_{t-1}}$: Learning the leverage effects • $\text{LE}(\text{leverage}-\text{effect}) = 0.138036\left(\left\lvert\frac{\varepsilon_{t-1}}{\sigma_{t-1}}-\sqrt{\frac{2}{\pi}}\right\rvert\right)$: Learning the asymmetric shock by leverage effects
NN-EWMA	• $\sigma^{2\prime}_{t-1} = 0.97\sigma^2_{t-1}$: Learning the pattern of volatility adjusted by decay factor between one period time intervals • $\varepsilon^{2\prime}_{t-1} = 0.03\varepsilon^2_{t-1}$: Learning the residual effects adjusted by decay factor between one period time intervals

recommended because the GARCH(1,1) model can be changed to ARCH(∞) by the repetitive input process (see Table 2).

To verify the forecastibility of the hybrid ANN-time series model, this paper inspects the reasonability of the financial time series model by analysing KOSPI 200 time series. Then, NN-GARCH(1,1) is analysed to prove that integration of ANN and GARCH can enhance the predictive power. After analysing the NN-GARCH model, this part will suggest which model (GARCH, EGARCH, EWMA) has the most predictive power in forecasting volatility of KOSPI 200 index.

The process of verification can be classified into two ways. The ANN, GARCH(1,1) and NN-GARCH(1,1) models are compared with one another from the points of deviation (MAE) and direction (Hit ratio). To assess the volatility, this study uses the realized volatility defined as true volatility in the market and used to compare predictive power between each models. Therefore, in this study, realized volatility is defined as a 22-day true standard deviation of the logarithmic return of the KOSPI 200 index. The volatility of yields is as follows:

$$\text{TRV}_\tau = \sqrt{\frac{1}{\tau_t}\sum_{k=1}^{\tau_t}(r_{t,k}-\bar{r}_t)^2},$$

where TRV: true realized volatility, τ_t: 22 days after t, $\bar{r}_t = \frac{1}{\tau_t}\sum_{k=1}^{\tau_t} r_{t,k}$: mean of realized return during 22 days after t. $r_{t,k}$: return of KOSPI 200 index at month t, day k.

Computed realized volatility is expressed as a 22-day volatility, $\sigma_{22\,\text{days}} = \sigma_{1\,\text{day}} \times \sqrt{22}$, to be compared with the forecasted 22-day volatility of the KOSPI 200 index.

4.2. Variable coefficient estimation

Table 3 summarizes the expected learning effects of extracted input variables by each model. The NN model is added to compare the results and to analyse the learning effects by integration. The NN model is analysed for the purpose of comparison among the proposed models. Table 4 exhibits the relative contribution factors produced by the ANN learning process. The contribution factors of input variables extracted by time series models are adjusted as if it were the coefficients of financial time series models.

Table 4
Input variables and relative contribution factors

Input variables	NN	NN-EWMA	NN-GARCH	NN-EGARCH
KOSPI200 yield square	0.0518	0.0532	0.0567	0.0393
Promised volume	0.0545	0.0452	0.0421	0.0422
KOSPI200 at $t-1$	0.0568	0.0489	0.0504	0.0516
KOSPI200 yield	0.0583	0.0626	0.0602	0.0534
3-Month government bond price	0.0596	0.0555	0.0567	0.0614
1-Year government bond yield	0.0605	0.0519	0.0561	0.0489
Open interest volume	0.0633	0.0555	0.0564	0.0528
Premium average	0.0654	0.0623	0.0693	0.0606
Contract volume	0.0667	0.0591	0.0593	0.0566
1-Year government bond price	0.0674	0.0593	0.0623	0.0576
3-Month government bond yield	0.0685	0.0556	0.0549	0.0558
KOSPI 200 at t	0.0731	0.0722	0.0648	0.0680
σ^2_{t-1}	0.2542			
$\sigma^{2\prime}_{t-1} = 0.97\sigma^2_{t-1}$		0.2244		
$\varepsilon^{2\prime}_{t-1} = 0.03\varepsilon^2_{t-1}$		0.0946		
$\sigma^{2\prime}_{t-1} = 0.836611\sigma^2_{t-1}$			0.2144	
$\varepsilon^{2\prime}_{t-1} = 0.064785\varepsilon^2_{t-1}$			0.0963	
$\ln\sigma^{2\prime}_{t-1} = 0.900850\ln\sigma^2_{t-1}$				0.2176
LE(leverage − effect)				0.0778
L(leverage)				0.0565
Total	1.0000	1.0000	1.0000	1.0000

The statistical meanings of extracted input variables are expressed by relative contribution allocation. These results can be used as a statistical and neural network's basis to propose desirable models.

5. Results

A deviation comparison is done through MAE. The sequence of the smallest MAE is NN-EWMA > NN > NN-GARCH > NN-EGARCH. The MAE of NN-EWMA is greater than NN, so this result can specify that the adjustment by the decay factor λ in NN-EWMA model is not desirable in the same manner as in the pure EWMA models. Therefore, this step can confirm whether the financial time series model that is not successful in volatility forecasting is also not successful in hybrid models, for example NN-EWMA. This outcome can be confirmed by the results of the NN-EGARCH model that shows the smallest MAE, because the EGARCH model is very successful in forecasting volatilities. The NN-EGARCH model can produce new input variables, such as leverage and leverage effect. These newly created variables improve prediction of volatilities with the help of logarithmic conditional variance equations. Table 5 shows this result more clearly by the comparison of the forecastability rising ratios.

NN-GARCH and NN-EGARCH models are inferred to enhance the predictive power in the point of MAE when compared with the NN model. Especially, the NN-EGARCH model (29.43% rising) is a much better model than NN-GARCH (7.67% rising) in MAE. To guarantee this result statistically, a one-way ANOVA is implemented in Table 6.

Table 5
Rising of precision accuracy of hybrid models compared with NN model

Accumulated forecasting days	NN-EWMA (%)	NN-GARCH (%)	NN-EGARCH (%)
20	−52.53	−15.87	36.94
40	−23.61	−5.15	8.70
60	−16.00	−0.19	9.99
80	−7.26	5.45	9.42
100	−0.96	13.03	12.57
120	−0.51	16.91	13.16
140	−4.16	14.86	18.21
160	−12.42	7.67	29.43
Formula	$\frac{(\mathrm{MAE}_{\mathrm{NN}} - \mathrm{MAE}_{\mathrm{comparison\_model}})}{\mathrm{MAE}_{\mathrm{NN}}}$		

Table 6
One way ANOVA for comparing MAE

	NN	NN-EWMA	NN-GARCH	NN-EGARCH
NN		0.5638	0.8543	0.0098***
NN-EWMA	0.5638		0.1492	0.0000***
NN-GARCH	0.8543	0.1492		0.1000*
NN-EGARCH	0.0098***	0.0000***	0.1000*	

* 10% significance level.
** 5% significance level.
*** 1% significance level.

Table 7
Result of hit ratio by each model

Accumulated forecasting days	NN	NN-EWMA (%)	NN-GARCH29 (%)	NN-EGARCH15 (%)
10	30.00	30.00	100.00	100.00
20	40.00	35.00	80.00	85.00
40	32.50	32.50	62.50	57.50
60	41.67	43.33	61.67	58.33
80	47.50	45.00	58.75	58.75
100	46.00	45.00	60.00	58.00
120	44.17	43.33	57.50	58.33
140	42.86	42.14	58.57	59.29
160	43.75	43.13	60.00	60.63

A direction comparison is done through the hit ratio analysis, and the hit ratio of the NN model can be increased by the NN-GARCH and NN-EGARCH models. Table 7 shows that NN-GARCH (60.00%) and NN-EGARCH (60.63%) increase the hit ratio when compared with NN (43.75%). Especially, Hit ratios of hybrid models are greater in the early stages of the forecasted interval. Therefore, short-term (under 30 days) volatility forecast by the hybrid model is more excellent.

By the analysis of MAE and hit ratio, the NN-GARCH and NN-EGARCH models show a good performance when compared with the NN model. By the same token like this analysis, the extraction of new input variables like leverage effect by financial time series model can enhance the predictive power in the overall perspectives.

6. Conclusions

This study proposed the hybrid model between the ANN and financial time series models to forecast volatilities of stock price index. It specified that ANN-time series models can enhance the predictive power for the perspective of deviation and direction accuracy. Most of the prior studies have adjusted the weight of raw volatilities by repetitive trial and error of learning process and found that the optimal coefficient of input variables to produce the best results. This study found the coefficients of input variables by financial time series process and extracted new variables that greatly influence the results through analysing stock market domain.

Experimental results showed that the proposed hybrid NN-EGARCH model could be improved in forecasting volatilities of stock price index time series. Of course, there are still many tasks to be done for the hybrid ANN-time series model. These NN-time series models should be further tested for robustness by applying them to other problem domains.

Acknowledgements

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (KRF-2006-332-B00095).

References

Armano, G., Marchesi, M., & Murru, A. (2005). A hybrid genetic-neural architecture for stock indexes forecasting. *Information Sciences, 170*, 3–33.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics, 31*, 307–327.

Brooks, C. (1998). Predicting stock index volatility: can market volume help? *Journal of Forecasting, 17*, 59–80.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimator of the variance of United Kingdom inflation. *Econometrica, 50*(4), 987–1008.

Gavrishchaka, V., & Ganguli, S. B. (2003). Volatility forecasting from multi-scale and high dimensional market data. *Neurocomputing, 55*, 285–305.

Gupta, J. N. D., & Sexton, R. S. (1999). Comparing backpropagation with a genetic algorithm for neural network training. *Omega, 27*(6), 679–684.

Hamid, S. A., & Zahid, I. (2002). Using neural networks for forecasting volatility of S&P 500 index futures prices. *Journal of Business Research, 5881*, 1–10.

Hu, M. Y., & Tsoukalas, C. (1999). Combining conditional volatility forecasts using neural networks: An application to the EMS exchange rates. *Journal of International Financial Markets, Institution and Money, 9*, 407–422.

JP Morgan and Reuters. (1996). *RiskMetrics – Technical Document* (4th ed.), New York.

Kim, K. (2003). Financial time series forecasting using support vector machines. *Neurocomputing, 55*, 307–319.

Kim, K., & Han, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert Systems with Applications, 19*(2), 125–132.

Kimoto, T., Asakawa, K., Yoda, M., & Takeoka, M. (1990). Stock market prediction system with modular neural network. In *Proceedings of the international joint conference on neural networks, San Diego, California* (pp.1–6).

Lee, K. H., & Jo, G. S. (1999). Expert systems for predicting stock market timing using a candlestick chart. *Expert Systems with Applications, 16*(4), 357–364.

Nelson, D. B. (1991). Conditional heterosdasticity in asset returns: a new approach. *Econometrica, 59*(2), 347–370.

Nikolopoulos, C., & Fellrath, P. (1994). A hybrid expert system for investment advising. *Expert Systems, 11*(4), 245–250.

Sexton, R. S., Dorsey, R. E., & Johnson, J. D. (1998). Toward global optimization of neural networks: a comparison of the genetic algorithm and backpropagation. *Decision Support Systems, 22*(2), 171–185.

Trippi, R. R., & DeSieno, D. (1992). Trading equity index futures with a neural network. *The Journal of Portfolio Management, 19*, 27–33.

Tsaih, R., Hsu, Y., & Lai, C. C. (1998). Forecasting S&P 500 stock index futures with a hybrid AI system. *Decision Support Systems, 23*(2), 161–174.

Yao, J. T., Li, Y., & Tan, C. L. (2002). Option price forecasting using neural networks. *Omega, 28*, 455–466.