A Study on KOSPI Return Volatility Estimation Using Heterscedastic Time Series Model

Chang-Ho An[1].

[1]Department of Financial Information Engineering, Seokyeong University, Seoul 02713; E-mail: choan@skuniv.ac.kr

Abstracts: In this study, return volatility prediction model was estimated by converting composite stock index into return variables. As for the heteroscedasticity test for the return, the Q-test and the LM-test showed that heteroscedasticity existed at the 2nd and 7th lags. Therefore, the order was determined using the SBC statistic, the parameters were estimated using the ARCH (2) model, and the model fit test was conducted. The parameter estimates of the ARCH(2) model were statistically significant, but the residual analysis showed that autocorrelation existed and did not satisfy the normality test. In the results of applying the GARCH (1,1) model, the parameter estimates, residual analysis, and normality were found to be satisfactory. Therefore, the GARCH (1,1) model was determined as the KOSPI return volatility estimation model and volatility was predicted. Volatility was predicted to show high volatility in the period of January and February 2023, and to stay calm after March for a long-term. Continuing a calm state means that the probability of large volatility is high. Therefore, those in charge of government agencies need to check system improvement or policy establishment and make efforts to identify trends in volatility by market.

Keywords: Rate of Return, Heteroscedasticity, SBC Statistic, GARCH Model, Q-Test And LM-Test.

1. INTRODUCTION

Inflation is expected to continue in the world economy due to the escalating US-China conflict, Russia's invasion of Ukraine, and supply chain instability. Therefore, countries with high trade dependence, such as Korea, Netherlands, and Germany, are very much affected by external shocks such as the continuation of the global economic recession, soaring oil prices, sharp drops in semiconductor prices, soaring exchange rates, and tight monetary policies of major countries. Recently, the volatility of the domestic financial market (stock market, bond market, etc.) is greatly increasing as the US base rate continues to rise and the won-dollar exchange rate continues to rise. Return volatility of composite stock index was found to be affected by composite stock index in the previous day, exchange rate, and individual net purchases [1]. Exchange rate volatility was found to be affected by the exchange rate system, the degree of openness of the financial market, the liquidity of the dollar, and the amount of reserves for short-term foreign debt [2]. Volatility is a measure of the change in the value of financial investment assets (stocks, bonds, exchange rates, etc.), an important measure in determining the price of derivatives, investment strategies, and hedging strategies, and an important factor in explaining the efficiency of financial markets. Volatility has the following characteristics.

Volatility cannot be observed directly, and can be confirmed graphically by estimating the rate of return or the square of the rate of return from the financial asset data. As shown in (Figure. 1.), the characteristics of volatility are that the volatility is different between the time of occurrence and the normal time, that it is divided into high and low, and that, when volatility is high, it remains high for a certain period of time. Higher volatility means greater uncertainty in financial asset prices, and higher volatility indicates that lower returns may occur.

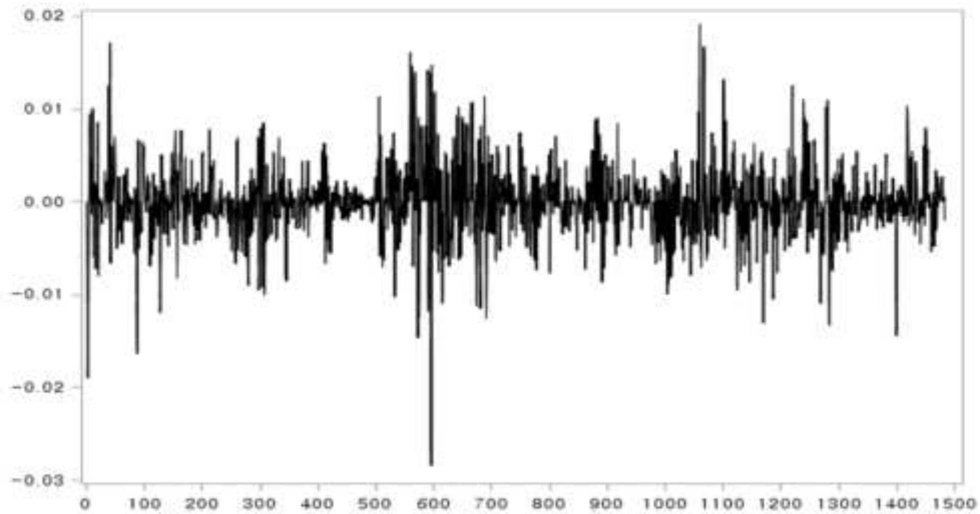

Fig. 1. Volatility clustering and persistence

According to a recent article citing JP Morgan Asset Management's 1st quarter report (2013-2022), it was reported that the return on the domestic stock market was as low as 1.9% and the volatility was as high as 21.3% [3]. Higher volatility of financial assets expands economic uncertainty, which negatively affects stock investment, exports, and productivity. Therefore, efforts to predict and prepare for volatility are an important research task to secure the stability of individual investors and all traders including government and financial institutions, as a means of eliminating uncertainty in investment and trading.

2. Prior Research

Volatility of financial assets is an important measure of uncertainty in financial and economic analysis. Predicting volatility is a very important issue in the business areas of companies, financial institutions, and foreign exchange market traders, for tasks such as investment strategy establishment, asset allocation, and risk management, and its importance is increasing. Volatility prediction research is being actively conducted by applying various models in the financial market, and existing studies on time series models and artificial intelligence techniques that study the volatility of financial assets are as follows.

Bollerslev said that the GARCH(1,1) model is easy to handle and has excellent predictive power among time series models [4], and Chay and Ryu estimated volatility by applying a time series model using data from July 1997 to July 2004. As a result of the estimation, the EGARCH(1,1) model was the most appropriate, and it was confirmed that the fit increased significantly at the expiration date [5]. Wilhelmsson proved the predictive power of the GARCH(1,1) model by analyzing the futures returns of the S&P 500 index [6]. Lee and Kwon applied the GARCH model and the EGARCH model to the excess return of the KOSPI 200 index and analyzed that the impulse persistence in the option market was very strong, as a result of the impulse response test [7]. Ohk and S. G. Lee analyzed that as a result of predicting volatility using the KOSPI 200 index, considering the volatility term structure has high explanatory power for the form of volatility in the option market [8]. Choi and Lee, using the KOSPI 200 index data, SVM, MART, and applying linear regression models, estimated and comparatively analyzed the implied volatility of option prices [9]. Chen et al. predicted and compared stock returns in the Chinese stock market with ANN and LSTM. In the results, it was confirmed that LSTM was more suitable than ANN, and the result of applying LSTM was presented [10]. Kim and Won integrated LSTM and time series models using the KOSPI 200 index and the S&P 500 index. As a result of the study, it was confirmed that the integrated model was suitable, and an integrated LSTM and GARCH-type model was proposed as volatility prediction model [11]. Cho and Kim analyzed the price difference between the regular market and the mini market using daily closing price of KOSPI 200 regular and mini options. As a result, it was confirmed that there existed a difference in price and liquidity between markets [12]. Shin et al. predicted and compared KOSPI 200 option volatility by applying machine learning techniques. As a result of the study, it was found that the performance of the decision model and the random forest algorithm was the best [13]. Kim and Jung applied GARCH models using Bitcoin data from 2017 to 2021 to analyze characteristics of price volatility. In results, it was found that the EGARCH model was most suitable in the Coindesk market and the CGARCH model in the Upbit market [14].

As can be seen in previous studies, there are various models for predicting financial asset volatility, and it can be seen that there are models suitable for each market. Therefore, in this study, we propose a volatility prediction model by applying a GARCH type model after converting KOSPI data into returns.

3. Research Method

Predicting the volatility of financial assets is important because it is directly related to traders, including individual investors and financial institutions. There are several time series models (ARCH, GARCH, AR-GARCH, etc.) that predict the volatility of financial assets, and these models are mainly used to estimate the volatility of the rate of return of financial assets. The models and test methods used in this study are as follows.

3.1 ARCH(p) model

The ARCH model is a model for modeling and predicting the volatility (conditional variance) at time t, and is called an autoregressive conditional heteroscedasticity model. The ARCH model is the basic model of the volatility estimation model proposed by Engle. If $\varepsilon_t = \phi_1 \varepsilon_{t-1} + \cdots + \phi_m \varepsilon_{t-m} + \nu_t$ is the autoregressive error model, the ARCH(p) model is as follows [15].

$$\nu_t = \sigma_t \, a_t \,, \quad a_t \sim i,i,d \;\; N(0,1)$$

(Equation 1)

$$\sigma_t^2 = \alpha_0 + \alpha_1 \nu_{t-1}^2 + \cdots + \alpha_p \nu_{t-p}^2$$

where, ν_t is the mean equation and σ_t^2 is the variance equation. The parameter estimate of the variance equation, α_0 , must be greater than zero, and α_i must be greater than or equal to zero for $i = 1,2,\cdots,p$. When the return volatility σ_t^2 follows the ARCH model, if the sum of the coefficients is close to 1, it indicates that the volatility of the return continues to be severe.

3.2 GARCH(p,q) model

In the ARCH model, if the numbaer of parameter p is large, the structure of the prediction model is complicated to express and the efficiency of the model is reduced. The proposed model to overcome this disadvantage of the ARCH model is the GARCH model. The GARCH model can bring about the estimation effect of long lag ARCH model even if the number of parameters is small. The GARCH model is a model proposed by Bollerslev. If $\varepsilon_t = \phi_1 \varepsilon_{t-1} + \cdots + \phi_m \varepsilon_{t-m} + \nu_t$ is the autoregressive error model, the GARCH(p,q) model is as follows [16].

$$\nu_t = \sigma_t \, a_t \,, \quad a_t \sim i,i,d \;\; N(0,1)$$

(Equation 2)

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \, \nu_{t-i}^2 + \cdots + \sum_{j=1}^{q} \delta_j \, \sigma_{t-j}^2$$

where, the parameter estimate of the variance equation, α_0 , must be greater than zero, α_i must be greater than or equal to zero for $i = 1,2,\cdots,p$, and δ_j must be greater than or equal to zero for $i = 1,2,\cdots,q$.

3.3 Autocorrelation and heteroscedasticity tests

The Q-statistic, a Portmanteau test statistic that tests the autocorrelation between squared errors using the square of the residual, and the LM-statistic that tests the effect of heteroscedasticity are as follows [17].

$$Q = T(T+2)\sum_{i=1}^{q}\frac{Corr(\widehat{v_t^2}, \widehat{v_{t-i}^2})}{(T-i)} \qquad \text{(Equation 3)}$$

$$LM = \frac{T\,W'\,Z(Z'Z)^{-1}\,Z'\,W}{W'\,W} \qquad \text{(Equation 4)}$$

The Q-test and the LM test are performed using the chi-square statistic of (Equation 3) and (Equation 4) to test the autocorrelation and heteroscedasticity effect.

4. Research results

4.1 Autocorrelation test

In general, there are cases in which the rate of return does not have autocorrelation. Therefore, in order to identify autocorrelation and dependency, squared rate of return is used. The results of identifying autocorrelation and dependency by the Portmanteau test using the squared return data of the KOSPI (2010.01 ~ 2022.12) are shown in (Table 1.), and the results identified as graphs are shown in (Figure. 2.). In the Portmanteau test ($\alpha = 0.05$) of (Table 1.), from 1 to 6 lag, the autocorrelation does not exist because the p-value of the χ^2 statistic is greater than $\alpha = 0.05$. However, from 1 to 12 lag, from 1 to 18 lag, and from 1 to 24 lag, the p-value is smaller than $\alpha = 0.05$, indicating that autocorrelation exists.

Table 1. Autocorrelation test of squared return (KOSPI)

Lags	Chi-Square	Pr > ChiSq	Autocorrelation Coefficient					
6	11.53	0.0732	0.164	0.002	0.173	0.011	0.117	0.042
12	25.83	0.0113	0.088	0.245	0.117	-0.004	0.064	-0.032
18	33.77	0.0135	0.003	0.085	-0.037	-0.054	0.102	0.152
24	48.13	0.0024	0.129	-0.024	0.076	0.233	0.033	0.021

As seen in (Figure. 2.), approximately, the range of fluctuations appeared large from 2010 to 2012, continued to be constant after 2012, and again appeared large from 2018 to 2022. The results of (Table 1.) and (Figure. 2.) mean that the return data of the squared KOSPI have autocorrelation and dependency. The results of (Table 1.) and (Figure. 2.) mean that the squared return of the KOSPI have autocorrelation and dependency.

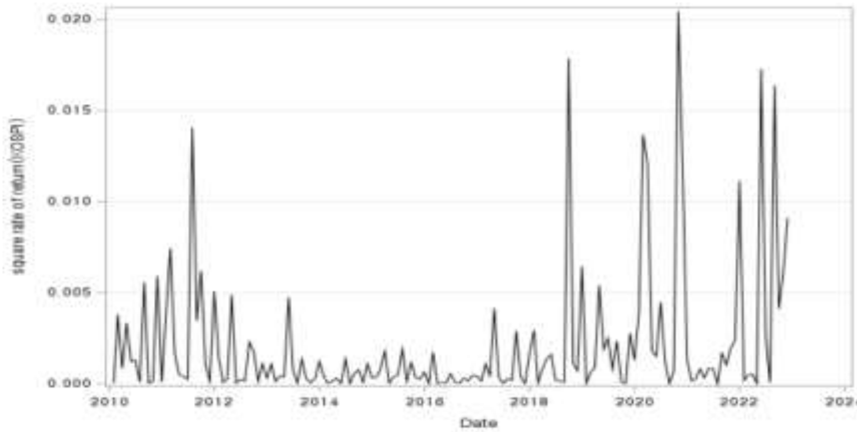

Fig. 2. Volatility trend of squared return (KOSPI)

4.2 heteroscedasticity test

(Table 2.) shows the results of testing the conditional heteroscedasticity of KOSPI return, The existence of heteroscedasticity was confirmed by the Q-test (Portmanteau Q-test) and the LM-test (Lagrange Multiplier-test). In the Q-test, the lags for which the p-value of the Q statistic is greater than $\alpha = 0.05$ are lag 2 (p-value=0.1020) and lag 7 (p-value=0.0656), and in the LM-test, the lag for which the p-value of the LM statistic is greater than is lag 2 (p-value=0.1080). It indicates that conditional heteroscedasticity exists in the return data at lags of 2 and 7.

Table 2. Heteroscedasticity test of return data (KOSPI)

Lags	Q Statistic	Pr > Q	LM Statistic	Pr > LM
1	4.5652	0.0326	4.3594	0.0368
2	4.5660	0.1020	4.4512	0.1080
3	9.6912	0.0214	9.6583	0.0217
4	9.7114	0.0456	9.9475	0.0413
5	12.0855	0.0336	13.3593	0.0202
6	12.3998	0.0536	13.5571	0.0350
7	13.7614	0.0556	16.1956	0.0234
8	24.3782	0.0020	22.9737	0.0034
9	26.8028	0.0015	23.565	0.005
10	26.8055	0.0028	24.0912	0.0074
11	27.5454	0.0038	24.3895	0.0112
12	27.7272	0.0061	26.3011	0.0097

4.3 ARCH(p) model identification and estimation

The SBC (Schwarz Bayessian criterion) statisticstatistic was used to determine the order p in the ARCH(p) model of the KOSPI return. As a result, the value of the statistic was the smallest at lag 2 (p=2), so the ARCH(2) model was estimated by applying the maximum likelihood estimation method. The estimated results are shown in (Table 2.), and the parameter estimates were found to be statistically significant.

Table 3. ARCH(2) model estimation

Parameters by Maximum Likelihood Estimation				
Variable	Estimate	S.E	t -Value	Pr >|t|
ARCH 0	0.000256	0.000057	4.50	<.0001
ARCH 1	0.002620	1.0436E-6	3.21	<.0001
ARCH 2	0.129200	0.136200	1.45	<.0001

4.4 ARCH(2) model fit and normality test

(Table 4.) shows the results of fit test and the Portmanteau test using residuals and residual squares of statistically significant ARCH(2) model. In the Portmanteau test using residuals, the p-value of the χ^2 statistic is greater than $\alpha = 0.05$ at all lags. That is, there is no autocorrelation. However, the result using the residual squares shows that the p-value of the chi-square statistic is smaller than $\alpha = 0.05$ at all lags. That is, there is autocorrelation. Therefore, the ARCH(2) model cannot be said to be fit.

Table 4. ARCH(2) model fit test

Lags	Chi-Square	Pr > ChiSq	Autocorrelation Coefficient					
6	3.38	0.7598	-0.021	-0.015	0.131	-0.007	0.043	0.037
12	5.53	0.9380	-0.006	0.015	-0.016	0.003	0.039	-0.103
18	8.77	0.9648	0.004	-0.002	-0.051	0.126	0.006	-0.003
24	11.46	0.9853	-0.005	-0.102	-0.012	-0.056	0.028	-0.02
6	19.48	0.0034	0.31	-0.08	-0.089	-0.065	-0.081	0.043
12	27.59	0.0064	-0.061	-0.082	0.119	0.145	-0.014	-0.053
18	36.3	0.0065	-0.162	-0.142	0.019	0.019	-0.05	0.034
24	42.81	0.0105	0.048	-0.035	-0.107	-0.085	-0.114	-0.02

In addition, a normality test was conducted to test the assumption that the residuals of KOSPI return in the ARCH(2) fitted model are white noise (mean=0, variation=1). It was also confirmed that the results of the normality test using the Shapairo-Wilk statistic (W-statistic) and the Kolmogorov-Smirnov (D-statistic) statistics were not satisfactory. Therefore, it was confirmed that the ARCH(2) model is a model that needs further improvement.

4.5 GARCH(1,1) model estimation

As an alternative model to the ARCH(2) model, the GARCH model proposed by Bollerslev was applied. This is because the GARCH model is a generalized model of the ARCH model, and can describe various volatility and achieve similar effects to the ARCH model. The results of estimating the model by applying the GARCH (1,1) model are shown in (Table 5.).

Table 5. GARCH(1,1) model estimation

Parameters by Maximum Likelihood Estimation				
Variable	Estimate	S.E	t -Value	Pr >|t|
ARCH 0	0.000131	0.000301	0.81	<.0001
ARCH 1	0.1317	0.0692	1.37	0.0001
GARCH 1	0.6396	0.0478	11.31	0.0102

4.6 Fitness and normality test of GARCH(1,1) model

After fitting the estimated model, a Portmanteau test was performed to see if autocorrelation existed using residuals and residual squares, and the results are shown in (Table 6). In the Portmanteau test using the residuals, the p-value of the chi-square statistic was found to be greater than $\alpha = 0.05$ at all lags, The result using the residual squares also showed that the p-value of the χ^2 statistic was greater than $\alpha = 0.05$ at all lags.

Table 6. GARCH(1,1) model fit test

Lags	Chi-Square	Pr > ChiSq	Autocorrelation Coefficient					
6	8.56	0.2000	-0.071	-0.099	0.18	-0.07	0.013	0.035
12	12.87	0.3789	0.107	0.037	-0.041	0.049	-0.034	-0.088
18	19.76	0.3467	-0.037	-0.127	0.055	-0.056	-0.072	-0.103
24	22.13	0.3716	-0.016	0.068	-0.107	-0.061	0.022	0.054
6	8.78	0.1862	0.01	-0.025	0.195	-0.004	0.122	0.026
12	18.09	0.1130	0.047	0.206	0.076	-0.021	0.063	-0.035
18	23.57	0.1695	-0.013	0.104	-0.041	-0.042	0.106	0.075
24	36.07	0.0829	0.039	-0.029	0.016	0.054	-0.009	-0.007

In addition, after fitting the estimated model, a normality test was conducted to test the assumption that the residual of KOSPI return data is white noise (mean=0, variation=1). The test results are shown in (Table 7.), and the p-values of the W statistic, D statistic, W-Sq statistic, and A-Sq statistic were all greater than $\alpha = 0.05$. That is, it satisfies the assumption that the distribution of residuals follows a normal distribution.

Table 7. Normality test of GARCH(1,1) model

test	statistic		p-value	
Shapiro-Wilk	W	0.988496	Pr < W	0.2339
Kolmogorov-Smirnov	D	0.031653	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.017259	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.210815	Pr > A-Sq	>0.2500

As a result of model diagnosis, the estimated GARCH(1,1) model has no autocorrelation and satisfies the assumption of normality, so it can be said to be fit as a predictive model for KOSPI return volatility. Therefore, estimation equation of the prediction model proposed in this study is (Equation 5).

$$\sigma_t^2 = 0.000131 + 0.1317a_{t-1}^2 + 0.6396\sigma_{t-1}^2 \qquad \text{(Equation 5)}$$

4.7 Volatility prediction by GARCH(1,1) model

The result of predicting KOSPI volatility by (Equation 5) is shown in (Figure. 3.). In the graph, the line after the reference line (straight line) on the x-axis is the predicted volatility. In the forecast results, it is predicted that volatility is high between January and February 2023, and will be small for a considerable period after March 2023.

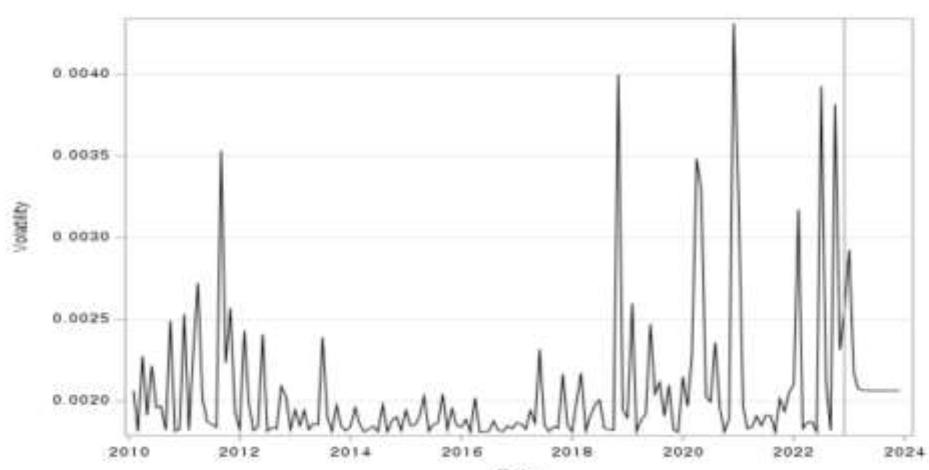

Fig. 3. KOSPI volatility prediction: GARCH(1,1)

CONCLUSION

In this study, the KOSPI return volatility prediction model was estimated using the conditional heteroscedastic time series model. As for the data, the KOSPI closing price (provided by ECOS) from January 2010 to December 2022 was converted into a return data. To estimate the volatility model, the heteroscedasticity effect of the return was first tested. As a result, the return data had a heteroscedasticity effect, so it was possible to apply the conditional heteroscedastic time series model. Therefore, as a result of applying the ARCH (2) model, the parameter estimates were significant, but the residuals had autocorrelation and were not satisfied with the normality test. However, in the results of applying the GARCH (1,1) model, autocorrelation did not exist and the normality test was satisfied, so KOSPI return volatility was predicted using the GARCH (1,1) model. The prediction results showed high volatility between January and February 2023 and a long period of calm after March. A long period of calmness in the prediction results indicates that the probability of large volatility increases. When stock market volatility expands, it causes inflation, which leads to an economic recession, and increases uncertainty in the entire economy, which has a great impact on the real economy. Therefore, government agencies should strive to improve systems and establish policies to stabilize market volatility and protect all traders. Predicting the future volatility of financial assets accurately is an important yet difficult problem. This is because the volatility models for each financial market are different and may differ from time to time. Therefore, volatility forecasting research such as artificial intelligence techniques including time series models and integrated models of time series and artificial intelligence should be conducted continuously.

Acknowledgments

This Research was supported by Seokyeong University in 2023.

REFERENCES

[1] [1] Engle, R. F., Kroner, K. F., Multivariate Simultaneous Generalized ARCH, Econometric Theory 11 (1995), 122-150. https://doi.org/10.1017/s0266466600009063

[2]https://www.bok.or.kr/portal/bbs/B0000347/view.do?nttId=10076916&menuNo=201106&pageIndex=1

[3] https://contents.premium.naver.com/geriforum/gericenter/contents/230223012514549qj

[4] Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. The review of economics and statistics, (1987), 542-547. https://doi.org/10.2307/1925546

[5] Chay, J. B and Ryu, H. S., Intraday Stock Market Volatility around Expiration Days in Korea, Korean Corporation Management Review, 12 (2005), 213-224.

[6] Wilhelmsson, A., Garch Forecasting Performance under Different Distribution Assumptions. Journal of Forecastingt. 25 (2006), 561-578. https://doi.org/10.1002/for.1009

[7] Jam, F. A. (2019). CRYPTO CURRENCY–A NEW PHENOMENON IN MONETARY CIRCULATION. Central Asian Journal of Social Sciences and Humanities, 4(1), 39-46.

[8] Lee, J. W. and Kwon, T. H., The Information Content and Volatility Transfer Effect of Implied Volatility on KOSPI 200 Return. The Korean Journal of Financial Engineering, 5 (2006), 41-59.

[9] Ohk, K. Y. and S. G. Lee, S. G., Deterministic Model of Volatility Surface in the KOSPI200 Options Market. Journal of the Korean Data Analysis Society, 14 (2012), 1633-1643.

[10] Choi, J. and Lee, J. T., An estimation of implied volatility for KOSPI200 option. Journal of the Korean Data and Information Science Society, 25 (2014), 513-522. 10.7465/JKDI.2014.25.3.513

[11] Chen, K., Zhou, Y. & Dai, F., A LSTM-based method for stock returns prediction: A case study of China stock market. In Big Data (Big Data). 2015 IEEE International Conference on (pp. 2823-2824). IEEE. (2015). https://doi.org/10.1109/bigdata.2015.7364089

[12] Kim, H. Y. and Won, C. H., Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. Expert Systems with Applications 103 (2018), 25-37. https://doi.org/10.1016/j.eswa.2018.03.002

[13] Cho, Y. and Kim, D., The Market Efficiency of the KOSPI200 Index Options and Mini Options. Journal of Industrial Economics and Business, 34 (2021), 319-340. https://doi.org/10.22558/jieb.2021.4.34.2.319

[14] Shin, S. H., Oh, H. Y. and Kim, J. H., Estimation of KOSPI200 Index option volatility using Artificial Intelligence. Journal of the Korea Institute of Information and Communication Engineering, 26 (2022), 1423-1431. https://doi.org/10.37727/jkdas.2022.24.2.665

[15] Libao, M. F. . (2023). LP-Conjugation: Unraveling Time Series Dynamics through Chaotic Map Convolution . International Journal of Membrane Science and Technology, 10(3), 977-983. https://doi.org/10.15379/ijmst.v10i3.1642

[14] Kim, H. B. and Jung, D. S., A Study on the Characteristics of Bitcoin Price Change Using the GARCH Model. Journal of the Korean Data

Analysis Society, 24 (2022), 665-681. http://doi.org/10.37727/jkdas.2022.24.2.665

[15] Engle, R. F., Autoregressive Conditional Heteroscedasticity with estimates of the Variance of United Kingdom Inflation, Econometrica, 50 (1982), 987-1008. https://doi.org/10.2307/1912773

[16] Bollerslev, T. P., Generalized Autoregressive Conditional Heteroscedasticity " , Journal of Econometrics, 31 (1986), 307- 327. https://doi.org/10.1016/0304-4076(86)90063-1

[17] McLeod, A. I. and Li, W. K., Diagnostic Checking ARMA Time Series Models using Squared_Residual Autocorrelations, Journal of Time Series Analysis, 4 (1983), 269-273. https://doi.org/10.1111/j.1467-9892.1983.tb00373.x

DOI: https://doi.org/10.15379/ijmst.v10i4.1890