InPrime

World Gold Price Forecast using APARCH, EGARCH and TGARCH Model

Yanne Irene, Madona Yunita Wijaya, and Aisyah Muhayani
Program Studi Matematika, Fakultas Sains dan Teknologi
Universitas Islam Negeri Syarif Hidayatullah Jakarta
Email: yanne.irene@uinjkt.ac.id, mywijaya@gmail.com, aisyah.muhayani14@mhs.uinjkt.ac.id

Abstract

Investment is a process of investing money for profit or material result. One investment commodity is gold. Gold is a precious metal in which the value tends to fluctuate over time. This indicates that there is a non-constant variance called heteroscedasticity. The appropriate time-series model to solve this heteroscedasticity problem is ARCH/GARCH. However, this model can't be applied for the financial cases that have an asymmetric effect (the downward and increase tendency in the level of volatility when returns rise and vice versa). Therefore, in this research, we forecast the world gold prices using APARCH, EGARCH, and TGARCH methods. We use the monthly world gold price data from June 1993 until May 2018. The result shows that the best-fitted model to forecasting the world gold prices is EGARCH (1.1). This model has the smallest error than the other models with a Mean Absolute Percentage Error (MAPE) value of 4.66%.

Keywords: return; volatilities; heteroscedasticity; asymmetric effect; APARCH; EGARCH; TGARCH.

Abstrak

Investasi adalah proses menginvestasikan uang untuk keuntungan atau hasil material. Salah satu komoditas investasi adalah emas. Emas adalah logam mulia yang nilainya cenderung berfluktuasi dari waktu ke waktu. Ini menunjukkan bahwa ada varian non-konstan yang disebut heteroskedastisitas. Metode deret waktu yang tepat untuk menyelesaikan masalah ini adalah ARCH/GARCH. Namun model ini tidak dapat digunakan untuk kasus keuangan yang memiliki efek asimetris (kecenderungan menurun dan meningkatnya volatilitas ketika nilai return naik dan sebaliknya). Oleh karena itu, dalam penelitian ini, kami memprediksi harga emas dunia menggunakan metode APARCH, EGARCH, dan TGARCH dengan data harga emas dunia bulanan pada bulan Juni 1993 - Mei 2018. Hasilnya menunjukkan bahwa, di antara ketiga metode itu, model terbaik untuk memprediksi harga emas dunia adalah EGARCH (1.1) dengan nilai *Mean Absolute Percentage Error* (MAPE) sebesar 4,66%.

Kata kunci: return; volatilitas; heteroskedastisitas; efek asimetris; APARCH; EGARCH; TGARCH.

1. INTRODUCTION

The capital market is defined as the market that traded various financial instruments in the long term, both in the form of debt and own capital. Investments are the activities of placing money or funds to obtain additional or certain profits such as buying stocks, bonds, and capital goods like gold. The value of gold is not affected by inflation. Therefore, it will be more profitable than stocks and

tends to be risk-free. The characteristic of the gold movement is influenced by several factors both fundamentally and technically i.e. financial, interest rate, socio-politics, and disaster. The important thing in investing is our attention to the movement of each asset. In the daily asset trading activity, the price of an asset is fluctuating, as is the price of gold either increase or decrease. In other words, the price of the asset is formed based on supply and demand. Because of these fluctuations, the stocks contain heteroscedasticity problems (time-varying variance) i.e. the error variances are not constant over time. The time series models to solve this condition are ARCH (Autoregressive Conditional Heteroscedasticity) that was introduced by Engle in 1982 [1]. Bollerslev [2] had developed its generalized model i.e. GARCH (Generalized Autoregressive Conditional Heteroscedasticity).

Enders [3] suggested that for some financial cases, there is a difference in the magnitude of volatility when the return value occurs, which is called asymmetry. Due to the asymmetric effect on financial data, the ARCH or GARCH models is not appropriate to model in such situation [4]. However, Chen et al. [5] succeed to forecast wind power with asymmetric volatility characteristics using GARCH model. In general, model that is capable to handle the asymmetric effect are the Asymmetric Power ARCH (APARCH) model, Threshold GARCH (TGARCH), and Exponential GARCH (EGARCH). These models have been successfully applied in many fields such as model the pathogens at marine recreational sites [6], model the international tourist arrivals [7], and model the stock market prices [8] [9]. Since the world gold price contains asymmetric effects, APARCH, TGARCH, and EGARCH models are considered in this study.

2. METHOD

We used the monthly data of the world's gold prices from http://www.worldbank.org from June 1993 until May 2018 (264 samples) which is divided into training data (80%) and testing data (20%). The world gold price is transformed into return defined as the rate of return of the asset. Return can be calculated using $r_t = \ln \frac{P_t}{P_{t-1}} = p_t - p_{t-1}$ [10].

The basic of time-series analysis is the stationary condition. The stationary is a condition in which the fluctuations of time series data are around the constant average value and the variation remains constant all the time [11]. One of the tests for checking the current data from a record is the Augmented Dickey-Fuller (ADF) test, with the null hypothesis $H_0: \beta = 0$ (there is a unit root so that the data is not stationary) versus the alternative hypothesis $H_1: \beta < 0$ (no unit root so the data is stationary) [10]. The statistical test used in the ADF test is $ADF - test = \frac{\hat{\beta}}{se(\hat{\beta})}$. If the $ADF - test \leq t_{(n-1,\alpha)}$ or $p - value < \alpha$ then the null hypothesis of non-stationarity is rejected.

If the data is already stationer, we can fit the data using ARMA model that requires several stages, namely the identification of the ARMA model, the estimation of ARMA models, and the diagnostic model test for ARMA residual. To identify the ARMA models, several plots such as ACF, PACF, EACF, and BIC table can be utilized. Least Square method is used to estimate the model parameters. To test the significance of the parameters, we can use $t - test = \frac{estimated\ parameter}{s.e\ estimated\ parameter}$, standard error $\neq 0$. The null hypothesis of insignificant parameter is rejected if $|t_{hitung}| > t_{\frac{\alpha}{2}(n-p-1)}$ or if $p - value < \alpha$, where α is the level of significance.

The next step is to test if the residuals follows white noise process, i.e. no autocorrelation in the model residuals. Ljung-Box test is commonly used to test if the residuals are independent or not [11]. The hypothesis for Ljung-Box test is $H_0 : \rho_1 = \rho_2 = \cdots = \rho_k = 0$ (There is no autocorrelation in the residual) with the test statistic used is $Q = n(n + 2) \sum_{n=1}^{k}(n - k)^{-1} r_k^2$; where $Q \sim \chi_{k-m}^2$; $m = p + q$, k:

maximum length of lag, n: number of sample, and r_k: ACF at lag-k. The null hypothesis is rejected if $Q > \chi^2_{table}$ or if $p - value < \alpha = 0.05$ with level significance is 5%.

The ARCH (p) Model can be formulated as $\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i a_{t-m}^2$ [10]. The effect of ARCH in the data can be investigated through collerogram of the residual square. The GARCH (p, q) model is a generalization of the ARCH (p) model. This model was developed by Bollerslev [2]. GARCH (p, q) model can be expressed as $\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i a_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2$ [12]. The APARCH (p, q) model was introduced by Ding [12]. This model can be written as $\sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i (|a_{t-i}| - \gamma_i a_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta$, $a_t = \sigma_t \epsilon_t$, with $\omega > 0, \delta > 0, \alpha_i > 0, \beta_j > 0, -1 < \gamma_i < 1$. The γ_i is called the leverage effect. According to Ding [12], the parameter δ can be divided into six criteria: (1) If $\delta = 2$, $\beta_j = 0$, for $j = 1,2, \dots q, \gamma_i = 0$ for $i = 1,2, \dots, p$ then APARCH model is an ARCH model, (2) If $\delta = 2$, $\gamma_i = 0 (i = 1,2, \dots, p)$, then APARCH model is a GARCH model, (3) If $\delta = 2$, the APARCH model is a GJR-GARCH model, (4) If $\delta = 1$, the APARCH model is a TARCH model, (5) If $\beta_j = 0, \gamma_i = 0$ for $i = 1,2, \dots, p$ then APARCH model is a NARCH model, (6) If $\delta = \infty$, the APARCH model is a Log-ARCH model.

The TGARCH(p, q) is written as $\sigma_t^2 = \omega + \sum_{i=1}^p (\alpha_i + \gamma_i N_{t-i}) a_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2$, $a_t = \sigma_t \epsilon_t$ [12], where $\alpha_i > 0, \beta_j > 0, \gamma_i$ and $N_{t-i} = 1$ for $a_{t-i} < 0$, and $N_{t-i} = 0$ for $a_{t-i} \geq 0$. The EGARCH model, was introduced by Nelson [13]. The TGARCH model is similar to GJR Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) [14]. EGARCH(p, q) model is written as $\ln \sigma_t^2 = \omega + \sum_{j=1}^q \beta_j \ln(\sigma_{t-j}^2) + \sum_{i=1}^p \alpha_i \left|\frac{a_{t-i}}{\sigma_{t-i}}\right| + \sum_{k=1}^r \gamma_k \frac{a_{t-k}}{\sigma_{t-k}}$, $a_t = \sigma_t \epsilon_t$, where $\alpha_i > 0, \beta_j > 0, \omega$, and γ_k are the model parameters.

To estimate the parameters in ARCH, GARCH, APARCH, TGARCH and EGARCH maximum likelihood estimation (MLE) can be used. According to [11], the likelihood function can be written as $(\theta) = f(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta), \theta \in \Omega$. The unction of $L(\theta)$ is maximized by taking the logarithm of the likelihood function $l(\theta) = \log L(\theta) = \sum_{i=1}^n \log f(x_i, \theta), \theta \in \Omega$. The parameter θ can be obtained by taking the first derivative of the log likelihood function equal to zero i.e. $\frac{d(\log(l))}{d\theta} = 0$. To examine the ability of the fitted model in forecasting future data, we use Mean Absolute Percentage Error (MAPE) is calculated using $MAPE = \frac{1}{N} \sum_{t=1}^N \left|\frac{Y_t - \widehat{Y}_t}{Y_t}\right| \times 100\%$, where Y_t is the actual data, $\widehat{Y}_t$ is the predicted data, and N is the number of sample [15].

3. RESULT AND DISCUSSION

The plot of gold price collected on a monthly basis is depicted in Figure 1. The average of monthly gold price monthly $688.5. The highest gold price occurred in May 2011 ($1772.14) and the lowest gold price was observed in July 1999 ($256.08). The plot also shows that the gold price data is not stationary in variance, and thus transformation of the return data with logarithm function is required. The ADF test reveals that the return data after logarithm transformation is already stationer since the p-value is less than 0.05 ($p = 0.01$).

Figure 2 displays the ACF, PACF, EACF, and BIC table of the reutrn data. Based on these plots, MA(1), MA(2), MA(3), MA(4), AR(3), ARMA$(1,1)$, ARMA$(1,2)$, ARMA$(1,3)$, ARMA$(2,2)$, and ARMA$(2,3)$ are selected as the candidate of the fitted model. Among these models, AR(3), MA(1), MA(2), ARMA(1,1), ARMA(1,3) and ARMA(2,3) are models with significant parameters.
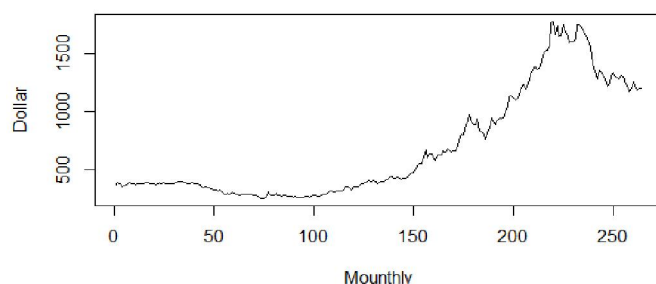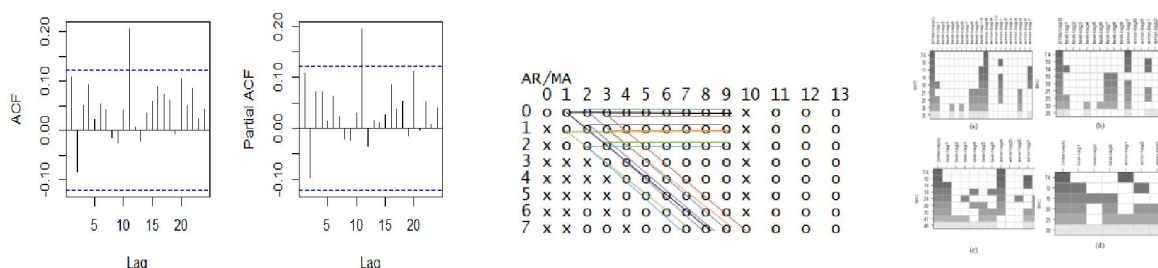
Figure 1. World Gold Price Plot.

Figure 2 Model identification with ACF, PACF, EACF, and BIC tables

Table 2 shows that best fitted model is MA(1) with significant parameter and has a modest order compared to other candidates. The estimated MA(1) model can be written as: $r_t = e_t + 0.1474e_{t-1}$, where r_t is the return of the price of gold at time t and e_t is the residual at time t.

After obtaining the best ARMA model the next step is to test whether the residual is independent or not by using the Ljung-Box test, and the result shows that independent assumption is hold since the p-value is above 0.05 ($\chi^2(1) = 0.15211$, $p = 0.6965$). To test the presence or absence of heteroscedasticity, Ljung-Box test is tested to the residual squared. The test shows that residuals contains heteroscedasticity effect since the p-value is less than $0.05(\chi^2(1)(1) = 11.824, p = 0.000548)$ and thus ARCH/GARCH model is required to model the conditional variance. Table 3 presents the estimated parameters of various GARCH models using maximum likelihood estimation.

Table 2. The estimated parameter without average.

No	Model	Parameter	Parameter Estimate	Standard error	Sign.	AIC
1.	AR (3)	ϕ_1	0.1389	0.0616	Yes	−990.55
		ϕ_2	−0.0971	0.0619	No	
		ϕ_3	0.0871	0.0619	No	
2.	MA (1)	θ_1	0.1474	0.0673	Yes	−991.47
3.	MA (2)	θ_1	0.1407	0.0610	Yes	−990.98
		θ_2	−0.0688	0.0557	No	
4.	ARMA (1,1)	ϕ_1	−0.3387	0.2611	No	−990.82
		θ_1	0.4825	0.2403	Yes	
5.	ARMA (1,3)	ϕ_1	0.9783	0.0198	Yes	−995.55
		θ_1	−0.8758	0.0641	Yes	
		θ_2	−0.1965	0.0770	Yes	
		θ_3	0.1313	0.0582	Yes	

Table 2. (Continued)

6.	ARMA (2,3)	ϕ_1	0.9153	0.1647	Yes	−989
		ϕ_2	−0.7197	0.3170	Yes	
		θ_1	−0.7878	0.1696	Yes	
		θ_2	0.5248	0.3299	Yes	
		θ_3	0.1962	0.0625	Yes	

Table 3 Estimated parameter of GARCH models.

No	Model	Parameter	Parameter Estimate	Standard Error	$p-value$	Sign.	AIC
	GARCH (1,1)	θ_1	0.213015	0.076982	0.005656	Yes	−3.8618
		α_0	0.000162	0.000082	0.049625	Yes	
		α_1	0.198758	0.072355	0.006015	Yes	
		β_1	0.688515	0.094746	0.000000	Yes	
	GARCH (1,2)	θ_1	0.212713	0.077678	0.06174	No	−3.8550
		α_0	0.000179	0.000098	0.068592	No	
		α_1	0.235928	0.102768	0.021691	Yes	
		β_1	00491746	0.286982	0.086619	No	
		β_2	0.151220	0.209911	0.471279	No	
	GARCH (1,4)	θ_1	0.243015	0.102672	0.017937	Yes	−3.9021
		α_0	0.000081	0.000070	0.245665	No	
		α_1	0.386124	0.118848	0.001159	Yes	
		β_1	0.000001	0.075666	0.999990	No	
		β_2	0.000000	0.062328	1.000000	No	
		β_3	0.603167	0.156508	0.000116	Yes	
		β_4	0.000000	0.106985	1.000000	No	
	GARCH (1,5)	θ_1	0.250642	0.100922	0.013009	Yes	−3.8953
		α_0	0.000080	0.000056	0.150931	No	
		α_1	0.395920	0.101552	0.000097	Yes	
		β_1	0.000000	0.095940	0.999997	No	
		β_2	0.000000	0.768424	1.000000	No	
		β_3	0.599680	0.041261	0.000000	Yes	
		β_4	0.000000	0.094865	1.000000	No	
		β_5	0.000000	0.678657	1.000000	No	

Based on Table 3, the best GARCH model is GARCH $(1,1)$ which can be expressed as: $r_t = e_t + 0.213015e_{t-1}$, $\sigma_t^2 = 0.000162 + 0.198758a_{t-1}^2 + 0.688515\sigma_{t-1}^2$. Figure 3 shows the correlogram to test the asymmetric effect. It shows that some bars exceed the significance limit which means that there is an asymmetric effect on volatility in the return of the world gold price data. Due to the asymmetric effect on the global gold price return data, the ARCH or GARCH models cannot be used. Therefore, APARCH, TGARCH and EGARCH models are considered to take into account the asymmetric effect.

Table 4 shows the estimated parameters of APARCH, TGARCH, and EGARCH models and Table 5 shows the forecasted gold price using APARCH $(1,1)$, TGARCH $(1,1)$, and EGARCH $(1,1)$ models. The plot of these forecasted values is displayed in Figure 4. Based on Table 5, the EGARCH (1.1) have the smallest MAPE value of 4.66%. Therefore, we conclude that the best fitted model to forecast the world gold price in the next 36 period is EGARCH $(1,1)$ model with conditional mean

and variance $r_t = e_t + 0.250517e_{t-1}$, $\ln \sigma_t^2 = -1.463484 + 0.145181 \left| \frac{a_{t-1}}{\sigma_{t-1}} \right| + 0.7882832 \ln(\sigma_{t-1}) +$
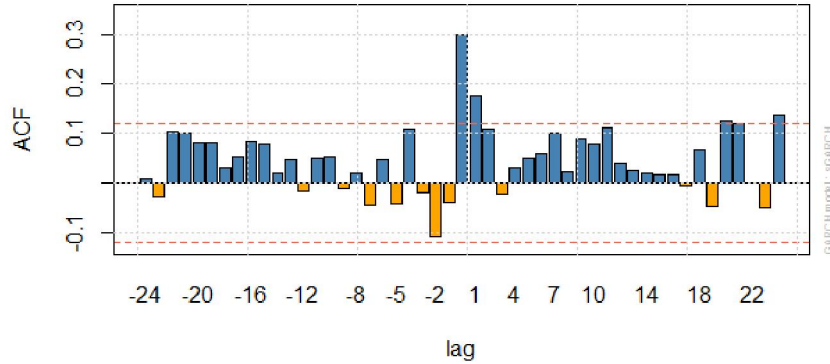
$0.387589 \frac{a_{t-1}}{\sigma_{t-1}}$.

Figure 3. Cross-correlation of residual squares versus lag residual.

Table 4. Estimated parameters APARCH, TGARCH, and EGARCH model.

No	Model	Parameter	Parameter Estimate	Standard error	$p-value$	Sign.	AIC
1.	APARCH (1,1)	θ_1	0.243072	0.078005	0.001833	Yes	−3.8638
		ω	0.000670	0.001623	0.679683	No	
		α_1	0.207740	0.088090	0.018361	Yes	
		β_1	0.609322	0.137287	0.000009	Yes	
		γ_1	−0.379492	0.223690	0.089789	No	
		δ	1.697095	0.679985	0.012568	Yes	
2.	TGARCH (1,1)	θ_1	0.257441	0.074442	0.000544	Yes	−3.865
		ω	0.007895	0.003623	0.029330	Yes	
		α_1	0.212035	0.071981	0.003222	Yes	
		β_1	0.610148	0.131541	0.000004	Yes	
		γ_1	−0.519816	0.211938	0.014180	Yes	
3.	EGARCH (1,1)	θ_1	0.250517	0.076174	0.001006	Yes	−3.8684
		ω	−1.463484	0.681506	0.031759	Yes	
		α_1	0.145181	0.082258	0.077572	No	
		β_1	0.7882832	0.100649	0.000000	Yes	
		γ_1	0.387589	0.122557	0.001564	Yes	

Table 5. Results forecast and accuracy of world gold price forecast with each method

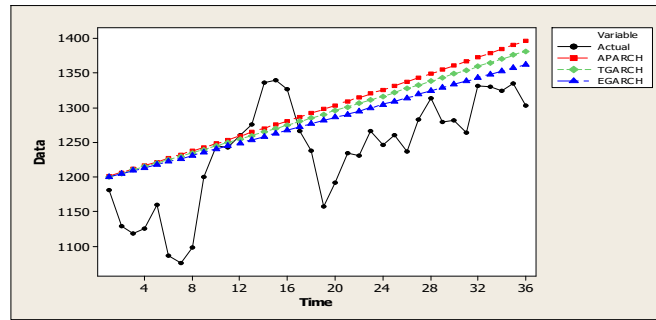| No.	The Actual Price	The forecast using		
		APARCH (1,1)	TGARCH (1,1)	EGARCH (1,1)
1.	1181.5	1200.928	1200.357	1200.127
2.	1128.31	1206.117	1205.2	1204.5
3.	1117.93	1211.329	1210.062	1208.888
⋮	⋮	⋮	⋮	⋮
34.	1324.66	1384.571	1370.913	1353.168
35.	1334.76	1390.554	1376.444	1358.098
36.	1303.45	1396.563	1381.997	1363.046
	MAPE (%)	5.68	5.226734	4.66

Figure 4. The price forecast using APARCH $(1,1)$, TGARCH $(1,1)$, and EGARCH $(1,1)$ models.

Figure 4 shows that the actual gold prices and the forecasted gold prices has different movements. The actual gold price tends to have an increasing trend causing the forecasting results to have a rising trend as well. In addition, the actual price in the middle of the year of 2015 to 2018 was impacted by economic problems where the exchange rate in different countries tend to be unstable. When the dollar strengthened, the price of commodity goods tended to move lower and when the dollar weakened, commodity goods were more likely to rise. This is what causes commodity goods such as gold to be fluctuating up and down from time to time causing the results between the actual and the forecast gold price somewhat different than that because in this research we do not take into account other factors but only the price of gold itself. Table 6 shows that the monthly world gold price forecast from June 2018 until December 2018 increases over time.

Table 6. World gold price forecast for June – December 2018

Month	Price ($)
June 2018	1368.013
July 2018	1372.997
August 2018	1378
September 2018	1383.021
October 2018	1388.06
November 2018	1393.118
December 2018	1398.194

4. CONCLUSION

APARCH, EGARCH, and TGARCH models used to forecast the price of gold because these models able to model the volatility of gold prices with the presence of asymmetric effect. Several models of candidates which can model the world's gold price data are APARCH $(1,1)$, EGARCH $(1,1)$ dan TGARCH $(1,1)$. Of the three models, the best model to forecast the world's gold prices is EGARCH $(1,1)$ with conditional mean MA(1): $r_t = e_t + 0.250517 e_{t-1}$, $\ln \sigma_t^2 = -1.463484 + 0.145181 \left| \frac{a_{t-1}}{\sigma_{t-1}} \right| + 0.7882832 \ln(\sigma_{t-1}) + 0.387589 \frac{a_{t-1}}{\sigma_{t-1}}$, where $r_t = \ln \frac{P_t}{P_{t-1}}$ and P_t is the world gold price at month t. This best model has succeeded in forecasting the gold price for some future period with the prediction of the smallest error compared to other models with MAPE 4.66%. This suggest that the EGARCH model is able to explain the effects of volatility in gold prices well.

REFERENCES

[1] R. F. Engle, "Autoregressive Conditional Heteroscedasticity with Estimate of the Variance of United Kingdom Inflation," *Econometrica,* pp. 987-1008, 1982.

[2] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics,* vol. 31, pp. 307-327, 1986.

[3] W. Enders, Applied Econometric Time Series Fourth Edition, New York: John Wiley and Sons, INC, 1995.

[4] E. A. Akpan and I. U. Moffat, "Detection and Modelling of Asymmetric GARCH Effects in a Discrete-Time Series," *International Journal of Statistics and Probability,* vol. 6, no. 6, pp. 111-119, 2017.

[5] H. Chen, J. Zhang, Y. Tao and F. Tan, "Asymmetric GARCH Type Models for Asymmetric Volatility Characteristics Analysis and Wind Power Forecasting," *Prot Control Mod Power Syst ,* vol. 4, no. 29, 2019, https://doi.org/10.1186/s41601-019-0146-0.

[6] G. Ali, "EGARCH, GJR-GARCH, TGARCH, AVGARCH, NGARCH, IGARCH and APARCH Models for Pathogens at Marine Recreational Sites," *Journal of Statistical and Econometric Methods,* vol. 2, no. 3, pp. 57-73, 2013.

[7] D. Chikobvu and T. Makoni, "Statistical modelling of Zimbabwe's international tourist arrivals using both symmetric and asymmetric volatility models," *Journal of Economic and Financial Sciences,* vol. 12, 2019, 10.4102/jef.v12i1.426.

[8] S. O. Ibrahim, "Forecasting the volatilities of the Nigeria stockmarket prices," *CBN Journal of Applied Statistic,* vol. 8, no. 2, pp. 23-45, 2017.

[9] S. Kumari and S. Priyanka, "Modelling Stock Return Volatility in India," *Munich Personal RePEc Archive, Paper No. 86674,* 2018.

[10] R. S. Tsay, Analysis of Financial Time Series Second Edition, Canada: A John Wiley and Sons, INC Publication, 2005.

[11] J. D. Cryer and K. S. Chan, Time Series Analysis with Application with R, New York: Springer Science+Business, LLC, 2008.

[12] D. Ding, "Modeling of Market Volatility with APARCH Model," Uppsala University, Swedia, 2011.

[13] B. D. Nelson, "Conditional heteroskedasticity in asset returns: a new approach," *Econometrica,* vol. 59, p. 347–70, 1991.

[14] J. M. Zakoian, "Threshold heteroskedasticity models," *Journal of Economic Dynamics and Control,* vol. 18, no. 5, pp. 931-955, 1994.

[15] S. Makridakis, "Accuracy measures: theoretical and practical concerns," *International Journal of Forecasting,* vol. 9, no. 4, pp. 527-529 doi:10.1016/0169-2070(93)90079-3, 1993.