Benin Journal of
Statistics
Vol. 1, pp. 45– 60 (2018)

On the performance of GARCH family models using the root mean square error and the mean absolute error

N. G. Emenogu[1]*, M. O. Adenomon[2] and N. O. Nweze[3]
[1]*Federal Polytechnic, Bida, Nigeria*; [2,3]*Nasarawa State University, Keffi, Nigeria*

Abstract. It is a common practice to detect outliers in a financial time series in order to avoid the adverse effect of additive outliers. This paper investigated the performance of GARCH family models (sGARCH; gjrGARCH; iGARCH; TGARCH and NGARCH) in the presence of outliers (small, medium and large) for different time series lengths (250, 500, 750, 1000, 1250 and 1500) using the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE). In a simulation iteration of 1000 times in R environment using rugarch, results revealed that for small size of outliers, irrespective of the length of time series, iGARCH was superior, for medium size of outliers, it was sGARCH and gjrGARCH that were superior irrespective of the time series length, and for a large size of outliers, irrespective of the time series length, gjrGARCH was superior. The study leveled that in the presence of additive outliers, both RMSE and MAE values would increase as the time series length is increased.

Keywords: additive outliers, models, simulation, time series length, R software.

Published by: Department of Statistics, University of Benin, Nigeria

1. Introduction

It is known that response variables are not only affected by exogenous variables but also by themselves from their past behavior (Ali, 2013). On the basis of this theoretical underpinning, autoregressive models have been developed. Box and Jenkins time series modeling is indispensable in analyzing stochastic processes (Ali, 2013). Autoregressive and moving average models are used frequently by many disciplines. The autoregressive framework has very found applications in macroeconomics, such as money supply, interest rate, price, inflation, exchange rates and gross domestic product and in financial time series analysis. The autoregressive heteroskedastic modeling framework is used in financial economics, such as asset pricing, portfolio selection, option pricing, and hedging and risk management (Ali, 2013). Studies abound in the financial literature on modeling the return on stocks. Usually, in the financial market, upward movements in stock prices are followed by lower volatilities, while negative movements of the same magnitude are followed by much higher volatilities (Ali, 2013).

Engle (1982) developed the time varying variance model known as Autoregressive Conditional Heteroskedastcity (ARCH) model. The ARCH model was the first model to assume that volatility is not constant. Bollerslev (1986) extended the model to include the Autoregressive Moving Average (ARMA) structure as the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. Ali (2013) asserted that a number of studies have adopted the Autoregressive Conditional Heteroskedasticity (ARCH) model or the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model to explain volatility of the stock market. Some of these studies have transformed and developed the Engel's basic model to more sophisticated models, such as the integrated GARCH (IGARCH), the threshold GARCH (TGARCH), the exponential GARCH (EGARCH), GARCH-in mean (GARCH-M), etc. (Atoi, 2014; Grek, 2014). However these sophisticated models, in most cases, fail to improve on the forecast accuracy of the original ARCH model.

*Corresponding author. Email: ngelive4jesus@gmail.com

1.1 *Outliers*

In Statistics, an outlier is an observation point that is distant from the other observations. An outlier may be due to variability in the measurement or experimental error. The later is sometimes excluded from the data set. An outlier can cause serious problems in statistical analyses. In the former case, one may wish to discard the outliers or use statistics that are robust to outliers, whereas in the later case they indicate that the distribution has high skewness and that one should be very cautious in using tools or intuitions that assume a normal distribution (Wikipedia, 2017).

There are two types of outliers, namely: innovation outlier (IO) (in which an outlier affects future values of the series), and additive outlier (AO) (in which an outlier affects only the current observation) (McQuarrie and Tsai, 2003). It should be noted that additive outliers affect the forecast performance of GARCH models such that the sum of squares increases as additive outlier increases to a large number (McQuarrie and Tsai, 2003).

This study focuses on the impact of additive outliers on performance of GARCH family models. Consequently, some GARCH models are reviewed and the impacts of additive outliers on the GARCH models are examined. Furthermore the study carried out simulations of the GARCH family models assuming three levels of outliers (small, medium and large) at different time series length. The simulation is replicated 1,000 times for each level of outliers and at different time series length and the performance of the GARCH models is adjudged using the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE).

1.2 *Justification*

In the search for appropriate forecasting models which seek to improve the forecast performance in financial time series, the family of GARCH models become readily most sought after due to the fact that most simple iid innovation models are unable to take into account the relevant information which is available at time t ((Rossi, 2004). Since additive outliers affect forecast performance of GARCH models such that the sum of squares increases as the additive outlier increases, this study will reveal that the GARCH models are more robust in forecasting volatility when additive outliers exist. The aim of this study is to compare the family of GARCH models when the problem of outliers exists in a financial time series.

2. Literature review

2.1 *The GARCH family models*

The Autoregressive Conditional Heteroskedasticity (ARCH) model introduced by Fredrick Engel in 1982 was the first model that assumed that volatility is not constant. ARCH models are commonly employed in modelling financial time series that exhibit time-varying volatility clustering, that is, period of swings interspersed with periods of relative calm (Grek, 2014; Wikipedia, 2017).

Over the years there have been several modifications and extensions of the ARCH model. Such modifications have resulted to the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models. The GARCH model was proposed independently by Bollerslev (1986) and Tylor (1986) in order to solve some of the disadvantages of the ARCH model. Some of the disadvantages of the ARCH model include (Tsay, 2005):

i. Assuming model that positive and negative shocks have the same effects on volatility because it depends on the square of the previous shocks, whereas in practice, it is well known that price of a financial asset responds differently to positive and negative shocks.
ii. The ARCH model is rather restrictive. For instance, α_1^2 of an ARCH(1) model must be in the interval $[0, 1/3]$ if the series has a finite fourth moment. The constraint becomes complicated for higher order ARCH models. Which in practice, limits the ability of ARCH models with Gaussian innovations to capture excess kurtosis.
iii. The inability of the ARCH model to provide any new insight for understanding the source of variations of a financial time series. It only provides a mechanical way to describe the behavior of the conditional variance without any indication about what causes such behavior to occur.

iv. ARCH models are likely to overpredict the volatility because they respond slowly to large isolated shocks to the return series.

Further studies on ARCH model are contained in Rossi (2004), Ragnarsson (2011) and Kelkay and Yohannes (2014).

Nelson (1991) proposed the Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH) model proposed. The EGARCH model overcomes some weaknesses of the GARCH model in handling financial time series, such as the violation of non-negativity constraints imposed on the parameters to be estimated. In particular, to allow for asymmetric effects between positive and negative asset returns to be estimated.

The EGARCH(m, s) model, according to Tsay (2005), Dhamija and Bhalla (2010), Jiang (2012), Ali (2013) and Grek (2014), can be written as

$$a_t = \sigma_t \epsilon_t$$

$$\ln(\sigma_t^2) = w + \sum_{i}^{s} \alpha_i \frac{|a_{t-i}| + \theta_i a_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^{m} \beta_j \ln(\sigma_{t-j}^2)$$

which gives EGARCH (1, 1) as

$$a_t = \sigma_t \epsilon_t$$

$$\ln(\sigma_t^2) = w + \alpha\left([|a_{t-1}| - E(|a_{t-1}|)]\right) + \theta a_{t-1} + \beta \ln(\sigma_{t-1}^2)$$

where $|a_{t-1}| - E(|a_{t-1}|)$ are iid and have mean zero. When the EGARCH model has a Gaussian distribution of error term, then $E(|\epsilon_t|) = \sqrt{2/\pi}$, which gives:

$$\ln(\sigma_t^2) = w + \alpha\left(\left[|a_{t-1}| - \sqrt{2/\pi}\right]\right) + \theta a_{t-1} + \beta \ln(\sigma_{t-1}^2)$$

The logarithm of the conditional variance in EGARCH signifies that the leverage effect is exponential and not quadratic. Tsay (2005) asserted that the logarithmic transformation of volatility removes the restriction on the parameters and to guarantee a positive variance.

The Nonlinear Generalized Autoregressive Conditional Heteroskedasticity (NGARCH) model (Higgins and Bera, 1992, Hsieh and Ritchken, 2005, Duan, et al, 2006) is an important modification of the GARCH model. This model exhibits the leverage effect and it has a very attractive feature of stock return.

Another extension of the model includes the Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) model proposed by Glosten et al. (1993). The model assumes a specific parametric form for the conditional heteroskedasticity in the zero mean white noise series. The Threshold GARCH (TGARCH) model by Zakoian (1994) is similar to the GJR GARCH model. The TGARCH model is commonly used to handle leverage effects (Atoi, 2014). This model does not use any restriction on the parameters with a view to guaranteeing that the conditional variance is positive. The Integrated Generalized Autoregressive Conditional Heteroskedasticity (IGARCH) model is a restricted version of the GARCH model, in the sense that the persistent parameters sum up to one, and imports a unit root to the GARCH process. The Quadratic Generalized Autoregressive Conditional Heteroskedasticity (QGARCH) model by Sentana (1995) is used to model asymmetric effects of positive and negative shocks. Hentschel (1995) proposed the family GARCH (fGARCH) model as an omnibus model that nests a variety of other popular symmetric and asymmetric GARCH models including the Asymmetric Power Autoregressive Conditional Heteroskedasticity (APARCH), the Glosten-Jagannathan-Runkle Generalized Autoregressive Conditional Heteroskedasticity (GJR-GARCH), Absolute Value Generalized Autoregressive Conditional Heteroskedasticity (AVGARCH), Nonlinear Generalized Autoregressive Conditional Heteroskedasticity (NGARCH), and so on. The Skew-Generalized Autoregressive Conditional Heteroskedasticity (SGARCH) model was introduced by De Luca, Genton and Loperfido (2005). It is a GARCH structure that takes into account the heteroskedastic nature of financial time series. It allows for parsimonious modeling, which is using as few predictor variables as possible to build a model which is adequate to accomplish the desired level of explanation or prediction, of multivariate skewness. According to De Luca and Loperfido (2012), all its elements are either null or negative, which is consistent with previous empirical and theoretical findings.

3. Methodology

This study focuses on the GARCH models that are robust for forecasting the volatility of financial time series data in the presence of outliers.

3.1 *Autoregressive Conditional Heteroskedasticity (ARCH) family model*

Every ARCH or GARCH family model requires two distinct specifications, namely: the mean and the variance equations (Atoi, 2014). The mean equation for a conditional heteroskedasticity in a return series, y_t is given by (Atoi, 2014)

$$y_t = E_{t-1}(y_t) + \epsilon_t \tag{1}$$

where $\epsilon_t = \phi\sigma_t$. The mean equation in equation (1) also applies to other GARCH family models. $E_{t-1}(\bullet)$ is the expected value conditional on information available at time $t-1$, while is the error generated from the mean equation at time t and ϕ_t is the sequence of independent and identically distributed random variables with zero mean and unit variance.

The variance equation for an ARCH(p) model is given by (Grek, 2014)

$$\sigma_t^2 = w + \alpha_1 a_{t-1}^2 + \cdots + \alpha_p a_{t-p}^2 \tag{2}$$

It can be seen in the equation that large values of the innovation of asset returns have bigger impact on the conditional variance because they are squared, which means that a large shock tends to follow another large shock and that is the same way the clusters of the volatility behave. So the ARCH(p) model becomes:

$$a_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = w + \alpha_1 a_{t-1}^2 + \cdots + \alpha_p a_{t-p}^2 \tag{3}$$

where $\epsilon_t \sim N(0,1)$ iid, $w > 0$ and $\alpha_i \leq 0$ for $i > 0$. In practice, ϵ_t is assumed to follow the standard normal or a standardized student-t distribution or a generalized error distribution (Tsay, 2005).

3.2 *Asymmetric power ARCH*

Given that (Rossi, 2004):

$$r = \mu + a_t, \qquad \epsilon_t = \sigma_t \epsilon_{t-1}, \qquad \epsilon_t \sim N(0,1)$$

$$\sigma_t^\delta = w + \sum_{i=1}^{p} \alpha_i \left(|a_{t-i}| - \gamma_i a_{t-i}\right)^\delta + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^\delta \tag{4}$$

where $w > 0$, $\delta \geq 0$, $\alpha \geq 0$, $-1 < \gamma_i < 1$, $\beta_j > 0$, $i = 1, 2, \cdots, p$, $j = 1, 2, \cdots, q$. This model imposes a Box-Cox transformation of the conditional standard deviation process and the asymmetric absolute residuals. This is the model (Equation (4)) is called Asymmetric Power ARCH and it includes seven other models (see Ding, Granger and Engel, 1993). The leverage effect is the asymmetric response of volatility to positive and negative "shocks".

3.3 *Standard GARCH(p, q) (sGARCH) model*

According to Rossi (2004), the asymmetric power ARCH model proposed by Ding, Engel and Granger (1993) given below forms the basis for deriving the GARCH family models. The mathematical model for the GARCH(p,q) model is obtained from equation (4) by letting $\delta = 2$ and $\gamma_i = 0$, $i = 1, 2, \cdots, p$. Thus:

$$a_t = \sigma_t \epsilon_t, \qquad \sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i a_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{5}$$

where $a_t = r_t - \mu_t$ (r_t is the continuously compounded log return series), and $\epsilon_t \sim N(0,1)$ iid, the parameter α_i is the ARCH parameter, β_j is the GARCH parameter, $w > 0$, $\alpha \geq 0$, $\beta \geq 0$ and $\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1$ (Rossi, 2004; Tsay, 2005; and Jiang, 2012).

The restriction on ARCH and GARCH parameters (α_i, β_j) suggests that the volatility (a_i) is finite and that the conditional standard deviation (σ_i) increases. It can be observed that if $q = 0$, then the model GARCH parameter (β_j) becomes extinct and what is left is an ARCH(p) model. To expatiate on the properties of GARCH models, the following representation is necessary. Let $\eta_t = a_t^2 - \sigma_t^2$ so that $\sigma_t^2 = a_t^2 - \eta_t$. By substituting $\sigma_{t-i}^2 = a_{t-i}^2 - \eta_{t-i}$, $(i = 0, \cdots, q)$ into Eq. (5), the GARCH model can be rewritten as

$$\alpha_t = \alpha_0 + \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) a_{t-i}^2 + \eta_t - \sum_{j=1}^{q} \beta_j \eta_{t-j} \tag{6}$$

It can be seen that $\{\eta_t\}$ is a martingale difference series (i.e., $E(\eta_t) = 0$ and $cov(\eta_t, \eta_{t-j}) = 0$, for $j \geq 1$). However, $\{\eta_t\}$ in general is not an iid sequence (Tsay, 2005). A GARCH model can be regarded as an application of the Autoregressive Moving Average (ARMA) idea to the squared series a_t^2. Using the unconditional mean of an ARMA model, results in this

$$E(a_t^2) = \frac{\alpha_0}{1 - \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i)}.$$

provided that the denominator of the prior fraction is positive (Tsay, 2005). When $p = 1$ and $q = 1$, we have GARCH(1, 1) model given by:

$$a_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = w + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \tag{7}$$

3.4 *GJR-GARCH(p, q) model*

The Glosten-Jagannathan-Runkle GARCH (GJRGARCH) model, which is a model that attempts to address volatility clustering in an innovation process, is obtained by letting $\delta = 2$. When $\delta = 2$ and $0 \leq \gamma_i < 1$, equation (4) becomes

$$\sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i \left(|\epsilon_{t-i}| - \gamma_i \epsilon_{t-i}\right)^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

$$= w + \sum_{i=1}^{p} \alpha_i \left(|\epsilon_{t-i}|^2 + \gamma_i^2 \epsilon_{t-i}^2 - 2\gamma_i |\epsilon_{t-i}| \epsilon_{t-i}\right) + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2. \tag{8}$$

$$\sigma_t^2 = \begin{cases} w + \sum_{i=1}^{p} \alpha_i^2 (1+\alpha_i)^2 \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, & \epsilon_{t-i} < 0 \\ w + \sum_{i=1}^{p} \alpha_i (1+\alpha_i)^2 \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, & \epsilon_{t-i} > 0 \end{cases}$$

That is,

$$\sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i (1-\gamma_i)^2 \epsilon_{t-i}^2 + \sum_{i=1}^{p} \alpha_i \left\{ (1+\gamma_i)^2 - (1-\gamma_i)^2 \right\} S_i^- \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

$$\sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i (1-\gamma_i)^2 \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{p} 4\alpha_i \gamma_i S_i^- \epsilon_{t-i}^2$$

where

$$S_i^- = \begin{cases} 1, & \text{if } \epsilon_{t-i} < 0 \\ 0, & \text{if } \epsilon_{t-i} \geq 0 \end{cases}$$

Now define $\alpha_i^* = \alpha_i (1-\gamma_i)^2$ and $\gamma_i^* = 4\alpha_i\gamma_i$, then

$$\sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i (1-\gamma_i)^2 \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{p} \gamma_i^* S_i^- \epsilon_{t-i}^2 \tag{9}$$

which is the GJR-GARCH model (Rossi, 2004). But when $-1 \leq \gamma_i < 0$, then

$$\sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i \left(|\epsilon_{t-i}| - \gamma_i \epsilon_{t-i} \right)^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

$$= w + \sum_{i=1}^{p} \alpha_i \left(|\epsilon_{t-i}|^2 + \gamma_i^2 \epsilon_{t-i}^2 - 2\gamma_i |\epsilon_{t-i}| \epsilon_{t-i} \right) + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2.$$

$$\sigma_t^2 = \begin{cases} w + \sum_{i=1}^{p} \alpha_i^2 (1+\alpha_i)^2 \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, & \epsilon_{t-i} < 0 \\ w + \sum_{i=1}^{p} \alpha_i (1+\alpha_i)^2 \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, & \epsilon_{t-i} > 0 \end{cases}$$

$$\sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i (1+\gamma_i)^2 \epsilon_{t-i}^2 + \sum_{i=1}^{p} \alpha_i \left\{ (1-\gamma_i)^2 - (1+\gamma_i)^2 \right\} S_i^+ \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$$

$$\sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i (1+\gamma_i)^2 \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{p} \alpha_i \left\{ 1 + \gamma_i^2 - 2\gamma_i - 1 - \gamma_i^2 - 2\gamma_i \right\} S_i^+ \epsilon_{t-i}^2$$

where

$$S_i^+ = \begin{cases} 1, & \text{if } \epsilon_{t-i} > 0 \\ 0, & \text{if } \epsilon_{t-i} \leq 0 \end{cases}$$

Also define $\alpha_i^* = \alpha_i(1+\gamma_i)^2$ and $\gamma_i^* = -4\alpha_i\gamma_i$, then

$$\sigma_t^2 = w + \sum_{i=1}^{p} \alpha_i^* \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{p} \gamma_i^* S_i^+ \epsilon_{t-i}^2 \tag{10}$$

which allows positive shocks to have a stronger effect on volatility than negative shocks (Rossi, 2004). But when $p = q = 1$, the GJR-GARCH(1,1) model will be written as

$$\sigma_t^2 = w + \alpha\epsilon_t^2 + \gamma S_i \epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{11}$$

3.5 *IGARCH(1, 1) model*

The integrated GARCH (IGARCH) models are unit-root GARCH models. The IGARCH (1, 1) model is specified in Tsay (2005) and Grek (2014) as

$$a_t = \sigma_t \epsilon_t; \qquad \sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + (1-\beta_1) a_{t-1}^2 \tag{12}$$

where $\epsilon \sim N(0,1)$ iid, and $0 < \beta < 1$. Ali (2013) used α_i to denote $1 - \beta_i$. This model is an exponential smoothing model for the $\{a_t^2\}$ series. To see this, rewrite the model as

$$\sigma_t^2 = (1-\beta_1)a_{t-1}^2 + \beta_1\sigma_{t-1}^2$$

$$= (1-\beta_1)a_{t-1}^2 + \beta_1\left[(1-\beta_1)a_{t-2}^2 + \beta_1\sigma_{t-2}^2\right]$$

$$= (1-\beta_1)a_{t-1}^2 + (1-\beta_1)\beta_1 a_{t-2}^2 + \beta_1^2\sigma_{t-2}^2 \tag{13}$$

By repeated substitutions, we have

$$\sigma_t^2 = (1-\beta_1)(a_{t-1}^2 + \beta_1 a_{t-2}^2 + \beta_1^2 a_{t-3}^3 + \cdots) \tag{14}$$

which, according to Tsay (2005), β_1 is the well-known exponential smoothing formation with being the discounting factor.

3.6 *TGARCH(p, q) model*

The Threshold GARCH model is another model used to handle leverage effects. A TGARCH(p, q) model is given by the following equation:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p}(\alpha_i + \gamma_i N_{t-i})a_{t-i}^2 + \sum_{j=1}^{q}\beta_j\sigma_{t-j}^2 \tag{15}$$

where N_{t-i} is an indicator for negative a_{t-i}, that is,

$$N_{t-i} = \begin{cases} 1, & \text{if } a_{t-i} < 0 \\ 0, & \text{if } a_{t-i} \geq 0 \end{cases}$$

and α_i, γ_i and β_j are nonnegative parameters satisfying conditions similar to those of GARCH models (Tsay, 2015). When $p = 1$, $q = 1$ the TGARCH(1, 1) model becomes:

$$\sigma_t^2 = w + (\alpha + \gamma N_{t-1})a_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{16}$$

3.7 *NGARCH(p, q) model*

The Nonlinear Generalized Autoregressive Conditional Heteroskedasticity (NGARCH) model has been presented in the literature (see Hsieh and Ritchken (2005), Lanne and Saikkonen (2005), Malecka (2014) and Kononovicius and Ruseckas (2015)). The NGARCH model is given as:

$$h_t = w + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{q} \gamma_i \epsilon_{t-i} + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{17}$$

where h_t is the conditional variance, and w, β and α, satisfy $w > 0$, $\beta \geq 0$ and $\alpha \geq 0$. The model can also be expressed as

$$\sigma_t = w + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{q} \gamma_i \epsilon_{t-i} + \sum_{j=1}^{p} \beta_j \sigma_{t-j} \tag{18}$$

3.8 *SGARCH(p, q) model*

The Skew-Generalized Autoregressive Conditional Heteroskedasticity (SGARCH) model can be written as:

$$Y_t = \eta_t \epsilon_t$$

$$\eta_t^2 = \delta_0 + \sum_{i=1}^{q} \delta_i (\eta_{t-i}\epsilon_{t-i})^2 + \sum_{j=q+1}^{q+p} \delta_j \eta_{t+q-j}^2 \tag{19}$$

where Y_t is the leading market return at time t, $\{\epsilon_t\} \sim i.i.d.N(0, 1)$ is the innovation (or shock) of the market, and δ_0 is hypothesized to be Gaussian. η_t^2 is assumed to be positive and the remaining parameters are nonnegative in order to ensure that is positive (De Luca and Loperfido, 2012).

3.9 *Simulation procedure*

The simulation procedure here considers the following equations of GARCH (1,1):

$$\epsilon_t = \sigma_t z_t$$

$$\sigma_t^2 = a_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \tag{20}$$

The case simulated is the case of financial time series where there are outliers at three levels, namely: small outlier values as 0.000005, 0.00006; medium outlier values as 10,50 and large outlier values as 100,500, at the following different time series length: 250, 500, 750, 1000, 1250 and 1500. By the level of the outliers we mean the extreme value that may be small, medium or very large in magnitude that is very different in size compared to other observations in the financial time series. Our choice of simulation is because real life financial time series may not exhibit all the forms of additive outliers characteristics we intend to study in the work. The rugarch package of the R software was used to execute the simulation.

3.10 *Forecast assessment*

The following are the criteria for forecast assessments used:

(1) Mean Absolute Error (MAE) has a formula

$$MAE_j = \frac{\sum_{i=1}^{n} |e_i|}{n}.$$

This criterion measures deviation from the series in absolute terms, and measures how much the forecast is biased. This measure is one of the most common ones used for analyzing the quality of different forecasts (Caraiani, 2010).

(2) The Root Mean Square Error (RMSE) is given as

$$RMSE_j = \sqrt{\frac{\sum_i^n (y_i - y^f)^2}{n}}$$

where y_i is the time series data and y^f is the forecast value of y (Caraiani, 2010).

For the two measures above, the smaller the value, the better the fit of the model (Cooray, 2008). In this simulation study,

$$RMSE = \frac{\sum_{j=1}^{N} RMSE_j}{N}$$

and

$$MAE = \frac{\sum_j^N MAE_j}{N},$$

where $N = 1000$, is the number of iterations or replications in the simulation study. The use of MAE and RMSE in this study rather than Schwarz Information Criterion (SIC) or other information criteria was by convenience, however it has been shown that model selection measures based on median/mean of forecasts are better that measure based on penalizing like information criteria in the study of volatility (Bal, et al, 2016).

4. Results and discussion

4.1 *Results*

The results of the simulation carried out are presented in Table 1 to Table 8 in the Appendix. Using RMSE criterion as presented in Table 9 in the Appendix. iGARCH is suitable when the outlier is small in magnitude while sGARCH and gjrGARCH are suitable when the outlier is medium in magnitude; gjrGARCH is suitable when the outlier is large in magnitude. In the overall, gjrGARCH outperformed all the other GARCH models.

Using MAE criterion as presented in Table 10 in the Appendix. iGARCH is suitable when the outlier is small in magnitude while sGARCH and gjrGARCH are suitable when the outlier is medium in magnitude; gjrGARCH is suitable when the outlier is large in magnitude. In the overall, gjrGARCH outperformed all the other GARCH models.

4.2 *Discussion*

4.2.1 GARCH models performance in the presence of outliers using the Root Mean Square Error (RMSE) from the results of the simulation

When the additive outlier was small, iGARCH outperformed the other models at time series lengths (T) of 250, 750 and 1500, and TGARCH performed better than the other models at time series length (T) of 500 and 1000, while NGARCH performed better than the other models at time series length (T) of 1250. For medium level of additive outliers, it can be clearly seen that the GARCH models that dominated were sGARCH and gjrGARCH. Whereas sGARCH performed better at time series lengths $T = 250$, $T = 500$ and $T = 750$, gjrGARCH outperformed the other models at $T = 1000$, $T = 1250$ and $T = 1500$. For the large level of outliers, gjrGARCH dominated, performing better at time series lengths (T) of 500, 750, 1250 and 1500, while TGARCH performed better at time series length (T) of 250, and sGARCH outperformed the other models at $T = 1000$.

4.2.2 GARCH models performance in the presence of outliers using the Mean Absolute Error (MAE) from the results of the simulation

For the small level of additive outliers, iGARCH dominated as it outperformed the other models at time series lengths (T) of 250, 750 and 1500. TGARCH performed better than the other models at time series length (T) of 500 and 1000, while NGARCH outperformed the other models at time series length (T) of 1250. For the medium level of additive outliers, it can also be seen that sGARCH and gjrGARCH dominated. While sGARCH performed better at time series lengths $T = 250$, $T = 500$ and $T = 750$, gjrGARCH on the other hand outperformed the other models at $T = 1000$, $T = 1250$ and $T = 1500$. For the large level of outliers, gjrGARCH dominated again, performing better at time series lengths (T) of 500, 1250 and 1500, sGARCH performed better than the other models at time series length (T) of 750 and 1000, while TGARCH outperformed the other models at $T = 250$.

5. Conclusion

This study has shown that different models performed better at different levels of outliers and at different time series lengths. This is in line with previous studies. For instance, Atoi (2014) modeled the volatility of stock returns using daily closing data of Nigerian Stock Exchange and found that GARCH (1,1), PGARCH (1,1,1) and EGARCH (1,1) with student's t distribution, and TGARCH with Generalized Error Distribution (GED) were the four best fitted models based on Schwarz Information Criterion. The present study also supports the conclusions in Grek (2014), Chen, Min and Chen (2013), Dijk, Franses and Lucas (1999) and Demos (2000) that different models performed differently under different conditions. Secondly, the study has shown that in the presence of additive outliers, gjrGARCH was superior, especially for medium and large levels of outliers for large time series lengths. Again the study established that for lower time series length, sGARCH was superior irrespective of whether MAE or RMSE was used in the assessment. The study recommends that investors, financial analysts and researchers interested in stock prices and asset return should adapt gjrGARCH and sGARCH when there are outliers in their data.

The contribution of this study to the literature as follows:

i. The fGARCH models in this study revealed their performance varies with the different levels of outliers.
ii. iGARCH is suitable when the outlier is small in magnitude.
iii. sGARCH and gjrGARCH are suitable when the outlier is medium in magnitude.
iv. gjrGARCH is suitable when the outlier is large in magnitude.
v. The Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) criteria produced similar results and conclusion.

References

Ali, G. (2013). EGARCH, GJR-GARCH, TGARCH, AVGARCH, NGARCH, IGARCH, and APARCH models for pathogens at marine recreational sites. Journal of Statistical and Econometric Methods, **2**(3): 57-73.

Atoi, N. V. (2014). Testing volatility in Nigeria stock market using GARCH models. CBN Journal of Applied Statistics, **5**: 65-93.

Bal, C., Demir, S., and Aladag, C. H. (2016). A comparison of different model selection criteria for forecasting EURO/USD exchange rates by feed forward neural network. International Journal of Computing, Communications & Instrumentation Engg., **3**(2): 271-275.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, **31**: 307-327.

Caraiani, P. (2010). Forecasting Romanian GDP using a BVAR model. Romanian Journal of Economic Forecasting, **4**:76-87.

Chen, S., Min, W. and Chen, R. (2013). Model identification for time series with dependent innovations. Statistica Sinica, **23**: 873–899.

Cooray, T. M. J. A. (2008). Applied Time series Analysis and Forecasting. New Delhi: Narosa Publishing House.

De Luca, G., Genton, M. G. and Loperfido, N. (2005). A multivariate skew-GARCH model. Advances in Econometrics, **20**: 33–56.

De Luca, G. and Loperfido, N. (2012). Modeling multivariate skewness in financial returns: a SGARCH approach. The European Journal of Finance, **21**: 1 -19. http://dx.doi.org/10.1080/1351847X.2011.640342.

Demos, A. (2000). Autocorrelation function of conditionally heteroskedastic in mean models. A Paper presented at the 13th ESRC Econometric Study Group Annual Conference, Bristol, England.

Dijk, D. V., Franses, P. H. and Lucas, A. A. (1999). Testing for ARCH in the presence of additive outliers. Journal of Applied Econometrics, **14**(5): 539-562.

Ding, Z., Granger, C. W. J. and Engle, R. F. (1993). A long memory property of stock market returns and a new model. Journal of Empirical Finance, **1**: 83-106.

Duan, J. C., Gauthier, G., Simonato, J. G. and Sasseville, C. (2006). Approximating the GJR-GARCH and EGARCH option pricing models analytically. Journal of Computational Finance, **9**(3): 1 -29.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica, **50**: 987-1007.

Glosten, L. R., Jagannathan, R. and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the national excess return on stock. Journal of Finance, **48**(5): 1779-1801.

Grek, A. (2014). Forecasting accuracy for ARCH models and GARCH(1,1) family which model does best capture the volatility of the Swedish stock market? Statistics Advanced Level Theses, Orebro University.

Hentschel, L. (1995). All in the family Nesting symmetric and asymmetric GARCH models. Journal of Financial Economics, **39**: 71-104.

Higgins, M. L. and Bera, A. (1992). A class of nonlinear ARCH models. International Economic Review, **33**(1): 137–158.

Hsieh, K. C. and Ritchken, P. (2005). An empirical comparison of GARCH option pricing models. Review of Derivatives Research, **8**(3): 129 -150.

Jiang, W. (2012). Using the GARCH model to analyse and predict the different stock markets. Master Thesis in Statistics, Uppsala University Sweden.

Kelkay, B. D. and G. Yohannes, E. (2014). The application of Garch family models to some agricultural crop products in Amhara national regional state. Journal of Economics and Sustainable Development, **5**: 24-35.

Kononovicius, A. and Ruseckas, J. (2015). Nonlinear GARCH model and 1/f noise. Physica A 427: 74-81. DOI: 10.1016/j.physa.2015.02.040

Lanne, M. and Saikkonen, P. (2005). Nonlinear GARCH models for highly persistent volatility. Econometrics Journal, **8**(2): 251-276.

Malecka, M. (2014). GARCH class models performance in context of high market volatility. ACTA Universitatis Lodziensis Folia Oeconmica, **3**: 253–266.

McQuarrie, A. D. and Tsai, C. L. (2003). Outlier detection in autoregressive models. Journal of Computational and Graphical Statistics, **12**(2): 450-471.

MaJose, R.V. (2010). Volatility models with leverage effect. A Doctoral Thesis, Universidad Carlos III De Madrid.

Nelson, D. (1991). Conditional heteroskedasticity in asset returns: A new approach. Econometrica, 59(2): 394-419.

Ragnarsson, F. J. (2011). Comparison of value-at-risk estimates from GARCH models. Master of Science Thesis, Copenhagen Business School, Handelshojskolen, Denmark.

Rossi, E. (2004). Lecture notes on GARCH models. University of Pavia.

Sentana, E. (1995). Quadratic ARCH models. The Review of Economic Studies, **62**(4): 639-661.

Tsay, R. S. (2005). Analysis of Financial Time Series (2nd Edition). New Jersey: John Wiley & Sons.

Taylor, S. (1986). Modelling Financial Time Series. Wiley, Chichster.

Wikipedia (2017). Autoregressive conditional heteroskedasticity (ARCH). https://en.wikipedia.org/wiki. (accessed 2017/07/10)

Zakoian, J. M. (1994). Threshold heteroskedasticity models. Journal of Economic Dynamics and Control, **18**(5): 931-955.

Appendix

Table 1: The RMSE and MAE values from the fGARCH family model at different levels of outlier of 0.000005, 0.00006 at different time series lengths

Outlier	0.000005,0.00006											
Time series length (T)	250		500		750		1000		1250		1500	
Model	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
sGARCH	4.931399	61.202956	7.021094	123.643002	8.618935	185.975292	9.945382	247.901383	11.15621	310.73484	12.21861	372.94698
gjrGARCH	4.91953	61.17128	7.015205	123.499443	8.60339	185.62731	9.967985	248.371364	11.14907	310.66572	12.22246	372.89526
iGARCH	4.890762	60.848528	7.03993	123.91656	8.57163	184.87379	9.971203	248.298866	11.1452	310.4949	12.20729	372.77983
TGARCH	4.936581	61.365738	6.998018	123.271838	8.626176	186.112062	9.928296	247.502235	11.14536	310.59719	12.22121	372.87911
NGARCH	4.90386	60.89776	7.007123	123.314485	8.598005	185.359426	9.977334	248.561108	11.11676	309.78808	12.21001	372.83915

Table 2: The Ranks of the RMSE and MAE values from the fGARCH family model at different levels of outlier of 0.000005, 0.00006 at different time series lengths

Outliers	0.000005,0.00006											
Time series length (T)	250		500		750		1000		1250		1500	
Model	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
sGARCH	4	4	4	4	4	4	2	2	5	5	3	5
gjrGARCH	3	3	3	3	3	3	3	4	4	4	5	4
iGARCH	1	1	5	5	1	1	4	3	2	2	1	1
TGARCH	5	5	1	1	5	5	1	1	3	3	4	3
NGARCH	2	2	2	2	2	2	5	5	1	1	2	2

Table 3: The RMSE and MAE values from the fGARCH family model at different levels of outlier of 10, 50 at different time series lengths

Outlier	10,50											
Time Series Length (T)	250		500		750		1000		1250		1500	
Model	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
sGARCH	70.80231	118.59578	70.46127	118.58901	69.83468	118.25637	70.16223	119.31764	69.81693	119.30811	70.51907	121.11153
gjrGARCH	71.69847	120.10907	70.6713	118.9951	69.96143	118.38869	68.57628	116.64949	67.85942	116.01714	69.43985	119.28377
iGARCH	72.18174	123.33328	72.60727	124.69260	72.2639	123.0646	72.10861	122.67465	72.04126	123.14981	72.10363	123.82272
TGARCH	71.9265	121.2603	72.07572	121.51990	72.07916	122.15585	71.73488	122.01710	72.03237	123.12899	72.10291	123.90638
NGARCH	71.67376	120.60028	71.60875	120.43185	70.08789	118.50482	70.44886	119.71692	69.80573	119.24967	69.94543	120.09876

Table 4: The Ranks of the RMSE and MAE values from the fGARCH family model at different levels of outlier of 10, 50 at different time series lengths

outliers	10,50											
Time series length (T)	250		500		750		1000		1250		1500	
Model	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
sGARCH	1	1	1	1	1	1	2	2	3	3	3	3
gjrGARCH	3	2	2	2	2	2	1	1	1	1	1	1
iGARCH	5	5	5	5	5	5	5	5	5	5	5	4
TGARCH	4	4	4	4	4	4	4	4	4	4	4	5
NGARCH	2	3	3	3	3	3	3	3	2	2	2	2

Table 5: The RMSE and MAE values from the fGARCH family model at different levels of outlier of 100, 500 at different time series lengths

OUTLIER	100,500											
	250		500		750		1000		1250		1500	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
SGARCH	723.8332	1223.1663	717.2956	1197.7994	711.0493	1185.5396	709.5818	1183.6321	708.9052	1183.6610	704.5643	1176.5473
GJRGARCH	730.0609	1276.7807	711.0275	1184.9110	704.8452	1195.3627	713.1035	1192.8623	703.415	1175.799	698.0259	1165.7791
IGARCH	767.9728	1500.6277	721.7786	1225.9165	735.1428	1377.8925	768.0416	1503.2967	726.2934	1237.5534	738.1277	1342.4379
TGARCH	718.2462	1211.7168	718.904	1210.781	720.2065	1205.1220	721.221	1227.972	720.7309	1204.4912	721.0888	1204.4408
NGARCH	871.6781	1600.8879	871.6815	1601.6008	721.7176	1241.8651	721.1639	1203.0861	721.1041	1203.5297	721.1119	1204.3342

Table 6: The Ranks of the RMSE and MAE values from the fGARCH family model at different levels of outlier of 100, 500 at different time series lengths

outliers	100, 500											
Time series length (T)	250		500		750		1000		1250		1500	
Model	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
sGARCH	2	2	2	2	2	1	1	1	2	2	2	2
gjrGARCH	3	3	1	1	1	2	2	2	1	1	1	1
iGARCH	4	4	4	4	5	5	5	5	5	5	5	5
TGARCH	1	1	3	3	3	3	4	4	3	4	3	4
NGARCH	5	5	5	5	4	4	3	3	4	3	4	3

Table 7: The Performances of the fGARCH family models at different levels of outliers and at different time series lengths using RMSE

Forecast Statistics: RMSE						
Size of Outlier	Time series length (T)					
	250	500	750	1000	1250	1500
Small	iGARCH	TGARCH	iGARCH	TGARCH	NGARCH	iGARCH
Medium	sGARCH	sGARCH	sGARCH	gjrGARCH	gjrGARCH	gjrGARCH
Large	TGARCH	gjrGARCH	gjrGARCH	sGARCH	gjrGARCH	gjrGARCH

Table 8: The Performances of the fGARCH family models at different levels of outliers and at different time series lengths using MAE

Forecast Statistic: MAE						
Size of Outlier	Time series length (T)					
	250	500	750	1000	1250	1500
Small	iGARCH	TGARCH	iGARCH	TGARCH	NGARCH	iGARCH
Medium	sGARCH	sGARCH	sGARCH	gjrGARCH	gjrGARCH	gjrGARCH
Large	TGARCH	gjrGARCH	sGARCH	sGARCH	gjrGARCH	gjrGARCH

Table 9: The Overall Performances Rating of the fGARCH family models at different levels of outliers and at different time series lengths using RMSE

	Small	Medium	Large	Total
sGARCH	-	16.6%	5.6%	22.2%
gjrGARCH	-	16.6%	22.2%	38.8%
iGARCH	16.7%	-	-	16.7%
TGARCH	11.1%	-	5.6%	16.7%
NGARCH	5.6	-	-	5.6%

Table 10: The Overall Performances Rating of the fGARCH family models at different levels of outliers and at different time series lengths using MAE

	Small	Medium	Large	Total
sGARCH	-	16.6%	11.1%	27.7%
gjrGARCH	-	16.6%	16.7%	33.3%
iGARCH	16.7%	-	-	16.7%
TGARCH	11.1%	-	5.6%	16.7%
NGARCH	5.6	-	-	5.6%