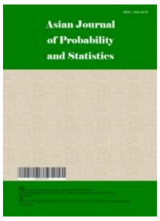


Asian Journal of Probability and Statistics

15(4): 251-275, 2021; Article no.AJPAS.69608
ISSN: 2582-0230

Comparative Performance of ARIMA and GARCH Model in Forecasting Crude Oil Price Data

Atanu, Enebi Yahaya [a*], Ette, Harrison Etuk [b] and Amos, Emeka [b]

[a] Department of Statistics, Federal Polytechnic of Oil and Gas, Bonny Island, Rivers State, Nigeria.
[b]Department of Mathematics/Computer Science, Rivers State University of Science and Technology, Port Harcourt, Nigeria.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2021/v15i430378
Editor(s):
(1) Dr. Dariusz Jacek Jakóbczak, Koszalin University of Technology, Poland.
Reviewers:
(1) Gamaliel Oghenerugba Eweke, Federal University Otuoke, Nigeria.
(2) Bouchemella Abdelhalim, Algeria.
Complete Peer review History, details of the editor(s), Reviewers and additional Reviewers are available here:
https://www.sdiarticle5.com/review-history/69608

Original Research Article

Received 05 April 2021
Accepted 15 June 2021
Published 17 December 2021

Abstract

This study compares the performance of Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroskedasticity models in forecasting Crude Oil Price data as obtained from (CBN 2019) Statistical Bulletin. The forecasting of Crude Oil Price, plays an important role in decision making for the Nigeria government and all other sectors of her economy. Crude Oil Prices are volatile time series data, as they have huge price swings in a shortage or an oversupply period. In this study, we use two time series models which are Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heterocedasticity (GARCH) models in modelling and forecasting Crude Oil Prices. The statistical analysis was performed by the use of time plot to display the trend of the data, Autocorrelation Function (ACF), Partial Autocorrelation Functions (PACF), Dickey-Fuller test for stationarity, forecasting was done based on the best fit models for both ARIMA and GARCH models. Our result shows that ARIMA (3, 1, 2) is the best ARIMA model to forecast monthly Crude Oil Price and we also found GARCH (1, 1) model is the best GARCH model and using a specified set of parameters, GARCH (1, 1) model is the best fit for our concerned data set.

**Corresponding author: Email: eyatanu@gmail.com;*

Keywords: Crude oil; oil price; ARIMA; GARCH; modeling.

1 Introduction

Modelling and forecasting of volatile data have become the major area of interest in financial time series. Volatility in this case refers to a condition where the conditional variance changes between extremely high and low value. In finance, measuring volatility by the conditional variance of return is often adopted as a crude measure of the total risk of the asset. Many values at risk models used for measuring the risk of market require the forecast of the volatility coefficients.

In this study, modelling and forecasting will be carried out using Crude oil price data.

Crude oil prices are volatile time series data because the prices just like any other volatile commodity have huge price swings in periods of oversupply or shortage. The crude oil prices cycle may last over several years responding to demand changes. Crude oil prices give impact to the cost of gasoline, manufacturing, home heating oil and electric power generation. The increase of oil prices will lead to the increase in cost of everything especially food and daily needs. This is because our daily necessities depend on transportation. This high oil prices will finally cause or increase inflation. Crude oil prices affect many related sectors that depend heavily on the usage of crude oil. The inconsistency of crude oil prices makes the modelling and forecasting of crude oil prices an important area of research. Apart from providing the information about the future oil prices to the public, crude oil forecasting is also crucial in determining the world's economic movement.

Crude oil, which is one of the most important commodities that affect the daily life of every one in a number of ways was discovered in commercial quantity in Nigeria on 15 January 1956 by Shell Darcy now known as Shell Petroleum Development Company at Oloibiri community in Bayelsa state [1]. This discovery and subsequent ones made Nigeria one of the major players in international oil trade. In today's world, crude oil is as important as the food that fuels the human body. Oil products are basically used in industries for production of goods and services and they are also used domestically for personal consumption. Oil plays a significant role in the Nigerian economy as the largest contributor in terms of total government revenue but also as the overall contributor in her exports composition. It accounted for about 82.1% of total government revenue during the oil boom in 1974 before reducing to a share of 64.3% by 1986 which was a consequence of the rapid decline in world market price of crude oil. The share of oil revenue in total government revenue still remains substantial as evidenced by the attainment of 85.6% and 86.1% in 2004 and 2005 respectively [2]. Crude oil is a non-renewable commodity but the world consumes it in different ways thus, becomes a challenge for statistician and econometrician to develop a better strategy for understanding the price changing aspect of crude oil. With better strategies, agencies and suppliers in charge of supplying the crude product can take more accurate and up-to date decisions especially for countries like Nigeria where the government yearly budget revolves around crude oil prices. Hence, crude oil price forecasting is very necessary for government agencies and investors to plan their activities in an effective manner. This has opened up research areas where compound and complex nature of the crude oil price is widely researched and most researchers use a variety of different procedures for better forecasting of crude oil price.

Two different techniques are used for crude oil price forecasting. In the first approach, the framework which is used for forecasting is akin to cause and effect, whereas the dependent variable is supposed to be affected by more variables generally called covariates. Sometimes this approach is also called fundamental analysis. This approach is very attractive by placing the reasons for ups and downs in price forecasting. So, many studies including (Ye et al., 2005) have used this technique. They expend the model of crude oil price and examine the nonlinear effect of processing plant utilization, OPEC capability utilization and future environment in markets as independent variables. This method has many limitations e.g. one cannot be sure about a certain explanatory variable that accounts for variations in the crude oil price. It is a difficult task to determine the exact functional form of a variable even if the exact variable is identified. The second approach is the time series modeling. In this approach, we no longer depend on the nature of explanatory variables, rather the predictions for the future values based on the past behaviour of the study variable. Several studies have been conducted using this approach, including [3,4,5,6] for forecasting the crude oil price, they used the well-known Box Jenkins methods while [7,8,9] used the GARCH method in forecasting the crude oil price. Moreover, it is to be observed that time series data act in certain ways because we are not capable to report all the changes of ups and downs based on natural reasoning, economic theory or inventory levels in the crude oil price. For better forecasting, many different approaches are used. In the midst of competing models for obtaining the forecast, selecting an appropriate model is a problem. In such situations, the choice of a model is usually based on the past accuracy, but the problem arises when the differences are statistically significant.

In time series, Autoregressive Integrated Moving Average (ARIMA) model is a generalized form of Autoregressive Moving Average (ARMA) model while Generalized Autoregressive Conditional Heteroskedasticity (GARCH) is a form of Autoregressive Conditional Heteroskedasticity (ARCH). The models are generally referred to as ARIMA (p, d, q) and GARCH (p, q) models where p, d and q are integers greater than or equal to zero and refer to the order of the autoregressive, the integration as the case in ARIMA and moving average. Crude oil price dynamics and evolution can be studied using a stochastic modeling approach that captures the time dependent structure embedded in the time series crude oil price data. The Autoregressive Integrated Moving Average (ARIMA) popularly known as Box-Jenkins Methodology [10] and the autoregressive conditional heteroscedasticity (ARCH) models, with its extension to generalized autoregressive conditional heteroscedasticity (GARCH) models as introduced by Engle, [11] and [12] respectively accommodates the dynamics of conditional heteroscedasticity (the changing variance nature of the data).

2 Literature Review

One objective of analyzing economic data is to predict the future values of certain variables. Time series analysis is an alternative approach that has proved quite successful, especially for short-term forecasting. It uses only the past values of a particular variable to predict its future values [13].

Somarajan S et al. [14] stated that the comprehension of volatility is a crucial concept in analyzing time series data. It is of greater importance for financial data since it furnishes key aspects such as return on investments and helps with effective hedging. The unpredictable nature of volatility causes heteroskedasticity which leads to difficulty in modelling. Consequently, time series models are desirable to predict volatility. Price volatility in the oil market refers to the degree to which crude prices rise or fall over a period of time. In an efficient market, prices reflect known existing and anticipated future circumstances of supply and demand and factors that could affect them. Changes in market prices tend to reflect changes in what markets collectively known or anticipate.

They are plethora of studies related to the oil price volatility, modeling and forecasting. Some of these studies employed either ARIMA or GARCH modeling approaches. The other studies combined ARIMA, GARCH family models and other improved modeling approaches. Only few others have attempted to combine ARIMA and GARCH models in forecasting variant oil prices. Salisu and Fasanya [15], examined crude oil price volatility modeling performance on the daily return of WTI, over the period of January 4, 2000 to March 20, 2012, using a combination of symmetric and asymmetric GARCH models. Sadorsky, [7] considered univariate, bivariate and state-space models where he finds that single-equation GARCH over performs more sophisticated models for forecasting petroleum futures prices. Muhammed and Umar [16], investigated the relevance of GARCH-family models in modeling and forecasting monthly Nigerian Bonny light crude oil prices from April 1986 to December 2015. The GARCH-GED was found to be the parsimonious model and performed better forecast than other GARCH family models and for ARIMA modeling approach, Ahmad [5], undertook a study on modeling and forecasting Oman crude oil prices from September 2000 to August 2010. The study revealed that multiplicative seasonal ARIMA (1, 1, 5) x (1, 1, 1) model is best in forecasting short-term Oman oil prices over the sample periods. Akomolafe and Danladi (2013), examined the application of Box-Jenkins approach (ARIMA) to the Nigerian budgeting for 2013, using monthly crude oil prices from January 1993 to October 2012. The finding from the study indicated that AR (2) is the best fit model. Abiola and Okafor (2013), examined the various forecasting models for the Nigerian crude oil prices from 2005Q1 to 2012Q4. The study discovered that ARIMA (1, 1, 4) model is best fitted forecasting model for predicting Nigerian crude oil price benchmark. Etuk [6], focused on modelling the monthly Nigerian Bonny light crude oil prices from 2006 to 2011, using seasonal ARIMA modelling. The result obtained reveals that ARIMA (0, 1, 1) x (1, 1,1)12 is the best fitted model for the Nigerian monthly oil prices. However, none of the studies reviewed has paid attention to the application of ARIMA and GARCH models in forecasting crude oil price with reference to the Nigerian Bonny light oil price. The present study tries to fill the gap identified from the exiting literatures.

Crude oil as one of the most important sources of energy and its prices have a great impact on the global economy. Crude oil has metaphorically been referred to as the ‘black gold’ [17]. Therefore, forecasting crude oil prices accurately is an essential task for investors, governments, enterprises and even researchers. However, due to the extreme nonlinearity and nonstationarity of crude oil prices, it is a challenging task for the traditional methodologies of time series forecasting to handle it.

The demand for crude oil will continue to increase, although its pace of growth is expected to slow gradually, according to the British Petroleum (BP) energy outlook 2017. Due to the importance of crude oil, many investors, governments, enterprises and even researchers pay much attention to the crude oil prices. However, a variety of factors such as speculation activities, supply and demand, technique development, geopolitical conflicts and wars can greatly produce effects on the prices of crude oil, making it show high nonlinearity and nonstationarity. Therefore, it is a challenging task to forecast the crude oil prices accurately. Various models have emerged to try to forecast the crude oil prices as accurately as possible in recent years including the autoregressive integrated moving average (ARIMA) and generalized autoregressive conditional heteroskedasticity (GARCH) family models. Wang and Wu (2012) forecasted the volatility of crude oil prices using multivariate and univariate GARCH-class models, and the results indicated that the multivariate models showed better performance than univariate models.

Suleman (2015) examined empirically the best ARIMA and GARCH models for forecasting. The data employed in their study comprise of 189 monthly observations of crude oil price in Nigeria spanning from January, 1998 to September, 2013. At first the stationary condition of the data series was observed by autocorrelation function (ACF) and partial autocorrelation function (PACF) plots, then checked using Kwiatkowski–Phillips–Schmidt–Shin (KPSS) and Augmented Dickey Fuller (ADF) test statistic. It was found that crude oil price is non-stationary. After taking the first difference of logarithmic values of data series, the same types of plots and the same types of statistics show that the data is stationary. The best ARIMA and GARCH models were selected by using the criteria such as AIC, HQC, and SIC. The model for which the values of criteria are smallest was considered as the best model. Hence ARIMA (3, 1, 1) and GARCH (2, 1) were found as the best model for forecasting the crude oil price data series.

Uwilingiyimana et al. [18] also considered the use of ARIMA-GARCH in forcasting inflation rate and had a good model for forecasting Kenya’s inflation rate. The empirical research employs time series analysis, ordinary least square and auto-regressive conditional heteroscedastic to find the estimators. The forecasting inflation analysis was conducted using two models, the ARIMA (1, 1, 12) model was able to produce forecasts based on the stationarity test and history patterns in the data compared to GARCH (1,2) model.

3 Methodology

This study, made use of secondary time series data of monthly Nigerian Bonny light crude oil prices in US$ per barrel from January 2006 to December 2018. The data was sourced from the websites of Central Bank of Nigeria (CBN). Classical time series model in form of Box-Jenkins approach (ARIMA model) and GARCH model are employed. Autoregressive Integrated Moving Average (ARIMA) was developed by Box and Jenkins [10], and often refers to as Box-Jenkins approach. However, ARIMA model has been considered as one of the best forecasting model by most time series scholars. While, Generalised Autoregressive Conditional Heteroscedastic (GARCH) model was developed by Bollerslev (1986), as an extension to Autoregressive Conditional Heteroscedastic (ARCH) which was introduced by Engle [11]. GARCH model explains that the conditional current variance depends on the previous conditional square residuals and the past conditional variance. Consequently, GARCH model became widely acceptable in modeling and forecasting economic and financial series [19].

The study looks at the comparative performance of ARIMA and GARCH models in modelling this data.

According to Elliott et al. (1996), given an observed time series $X_1, X_2, X_3, \dots, X_N$ Dickey and Fuller consider the differential-form autoregressive equations to detect the presence of a unit root (ΔX_t):

$$\Delta X_t = \alpha + \beta_t + \gamma X_{t-1} + \sum_{j=1}^{p} (\delta_j \Delta X_{t-j}) + w_t$$

ARIMA model is the combination of the following processes:

- Autoregressive process (AR)
- Moving Average process (MA)
- Differencing process (d)

3.1 Autoregressive (AR) Process

Autoregressive models are based on the idea that current value of the series, X_t, can be explained as a linear combination of ρ past values, $X_{t-1}, X_{t-2}, \dots, X_{t-\rho}$, together with a random error in the same series. An autoregressive model of order ρ, abbreviated $AR(\rho)$, is of the form:

$$X_t = \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \dots + \emptyset_\rho X_{t-\rho} + w_t = \sum_{i=1}^{\rho} \emptyset_i X_{t-i} + w_t \quad (3.1)$$

where X_t is stationary, $w_t \sim wn(0, \sigma_w^2)$, and $\emptyset_1, \emptyset_2, \dots, \emptyset_\rho$ are model parameters.

3.2 Moving Average (MA) Process

In AR models above, current observation X_t is regressed using the previous observations $X_{t-1}, X_{t-2}, X_{t-3}, \dots, X_{t-p}$, plus an error term w_t at current time point. One problem of AR model is the ignorance of correlated noise structures (which is unobservable) in the time series. { In other words, the imperfectly predictable terms in current time, w_t, and previous steps, $w_{t-1}, w_{t-2}, w_{t-3}, \dots, w_{t-q}$, are also informative for predicting observations.

A moving average model of order q, or MA(q), is defined to be

$$X_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \theta_3 w_{t-3} + \dots + \theta_q w_{t-q} = w_t + \sum_{j=1}^{q} \theta_j w_{t-j} \quad (3.6)$$

Where $w_{t \sim wn(0,\sigma^2),}$ and $\theta_1, \theta_2, \theta_3, \dots, \theta_q$ ($\theta_q \neq 0$) are parameters.

3.3 Arima Model

Many time series data especially crude oil data are nonstationary and so we cannot apply stationary AR, MA or ARMA processes directly. One possible way of handling non-stationary series is to apply *differencing* so as to make them stationary. The first differences, namely $(X_t - X_{t-1}) = (1 - B)X_t$, may themselves be differenced to give second differences, and so on. The dth differences may be written as $(1 - B)dX_t$. If the original data series is differenced *d* times before fitting an ARMA(*p, q*) process, then the model for the original un-differenced series is said to be an ARIMA(*p, d, q*) process where the letter 'I' in the acronym stands for *integrated* and *d* denotes the number of differences taken.

$$\emptyset(B)(1 - B)dX_t = \theta(B)w_t$$

3.4 ARCH (q) Process

In 1982, Engle introduced a new class of stochastic process called the Autoregressive Conditional Heteroskedasticity (ARCH) process, which allows the conditional variance to be time varying as a linear function of lagged errors, leaving the unconditional variance constant over time [11] (Bollerslev, 1986). ARCH was one of the first econometric models that provided a convenient way to model conditional heteroskedasticity in variance. To model an ARCH process, let ε_t denote the disturbance term, which depends on a stochastic component z_t and a time-varying standard devia-tion σ_t (Nelson, 1991). Mathematically, it can be written as

$$\varepsilon_t = \sigma_t z_t \tag{3.14}$$

where z_t ~ i.i.d. $\mathcal{N}(0,1)$. By definition, ε_t is serially uncorrelated with mean zero and conditional variance equal to σ_t^2. The conditional variance, σ_t^2, is modelled as follows:

$$\sigma_t^2 = \omega_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 + \alpha_3\varepsilon_{t-3}^2 + \cdots + \alpha_q\varepsilon_{t-q}^2 = \omega_0 + \sum_{j=1}^{q}\alpha_j\,\varepsilon_{t-j}^2,$$

where $\omega_0 > 0$ and $\alpha_j \geq 0 \ \ \forall_j \in \{1,2,3,\dots,q\}$.

3.5 GARCH Model

GARCH model is known as a model of heterocedasticity which means it's not constant in variance. This model has been used widely in financial and business areas since the data of these areas tend to have variability or highly volatile throughout the time. GARCH model is given as a combination moving average (MA) terms q and p, as the number of autoregressive (AR) terms.

Supposing we have a regression model given as;

$$X_t = X_t' + \varepsilon_t \tag{3.17}$$

where ε_t is the residuals and $\varepsilon_t \sim N(0,\sigma_t)$

Then, GARCH(p,q) model and the variance component is written as:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p}\alpha_i\varepsilon_{t-i}^2 + \sum_{j=1}^{q}\beta_j\sigma_{t-j}^2 \tag{3.18}$$

when p=1 and q=1, then it is considered as a case of GARCH (1,1). Where all the parameters α_o, $\alpha_i, \beta_j \geq 0$; σ_t^2 are the conditional variance, α_o constant term, α_i and β_j are coefficients of the ARCH and GARCH term respectively, σ_{t-i}^2 and ε_{t-j}^2 are the squared errors at lag $_{t-i}$ and *t-j* respectively.

4 Results and Discussions

Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Heterodasticity models were developed from the data collected. This developed model was used to forecast the January, 2019 to December, 2021. The autocorrelation, partial autocorrelation and run sequence plot and Augmented Dickey-Fuller Test were used to identify the model and check the stationarity of the series respectively. ARCH test for volatility was conducted to determine the ARCH effect of the series. The tentative models were ranked and the best model was selected with the lowest Akaike Information Criterion (AIC) value. Ljung-Box statistic was used to check for the randomness of the residual and one-sample Kolmogorov-Smirnov test was used to test for the normality of the residuals of the selected predictive model.

The crude oil price was obtained from CBN statistical bulletin recorded monthly from January, 2006 to December, 2018 as shown in Fig. 1. According to [20], ARIMA models can only be applied to non-stationary time series data, only when the data is transformed into stationary time series data for the purpose of generalization during forecast. Hence, before we perform the analysis of the time series data it is expected that we determine the stationarity of the data. The stationarity test of the crude oil price data is performed by Augmented Dickey-Fuller (ADF) Test of the actual data as presented in Table 1.

PLOT OF THE CRUDE OIL PRICE

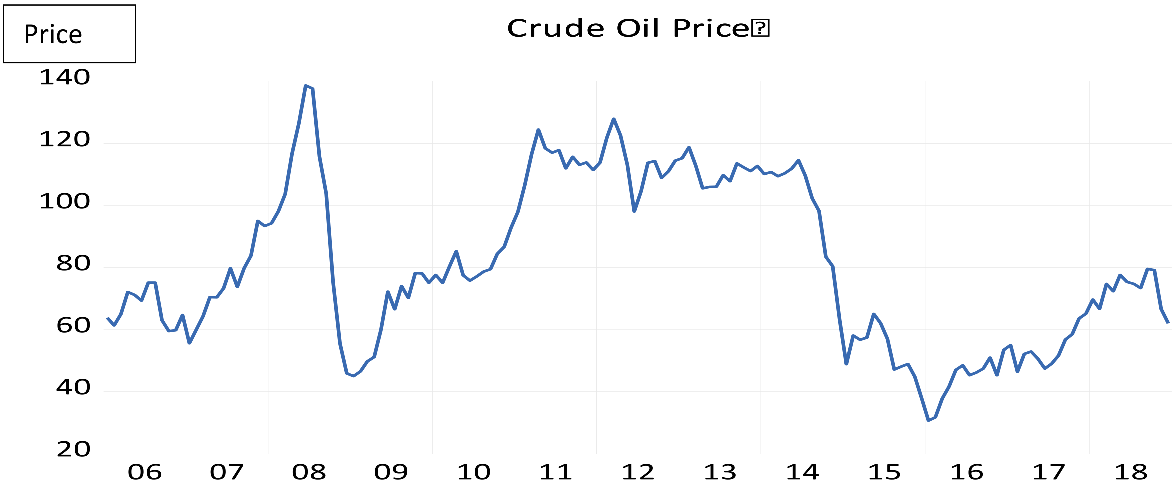


Fig. 1. Time Plot of Crude Oil Price

Table 1. Augmented dickey fuller test (level)

Null Hypothesis: CRUDE_OIL_PRICE has a unit root				
Exogenous: Constant				
Lag Length: 1 (Automatic - based on SIC, maxlag=13)				
			t-Statistic	**Prob.***
Augmented Dickey-Fuller test statistic			-2.296514	0.1744
Test critical values:	1% level		-3.473096	
	5% level		-2.880211	
	10% level		-2.576805	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(CRUDE_OIL_PRICE)				
Method: Least Squares				
Date: 09/18/20 Time: 18:02				
Sample (adjusted): 2006M03 2018M12				
Included observations: 154 after adjustments				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
CRUDE_OIL_PRICE(-1)	-0.043825	0.019083	-2.296514	0.0230
D(CRUDE_OIL_PRICE(-1))	0.373216	0.075569	4.938770	0.0000
C	3.545957	1.623810	2.183726	0.0305
R-squared	0.152928	Mean dependent var		0.004351
Adjusted R-squared	0.141708	S.D. dependent var		6.698478
S.E. of regression	6.205741	Akaike info criterion		6.508115
Sum squared resid	5815.195	Schwarz criterion		6.567276
Log likelihood	-498.1249	Hannan-Quinn criter.		6.532146
F-statistic	13.63052	Durbin-Watson stat		2.080634
Prob(F-statistic)	0.000004			

From the table above, we proceed to take the first difference in order certify that there is no presence of auto correlation in the crude oil price data.

4.1 Stationarity

In checking for stationarity of the crude oil price data, Table 2 and Figure 2 is considered, showing that the data is stationary after the first differencing.

FIRST DIFFERENCED PLOT OF CRUDE OIL PRICE

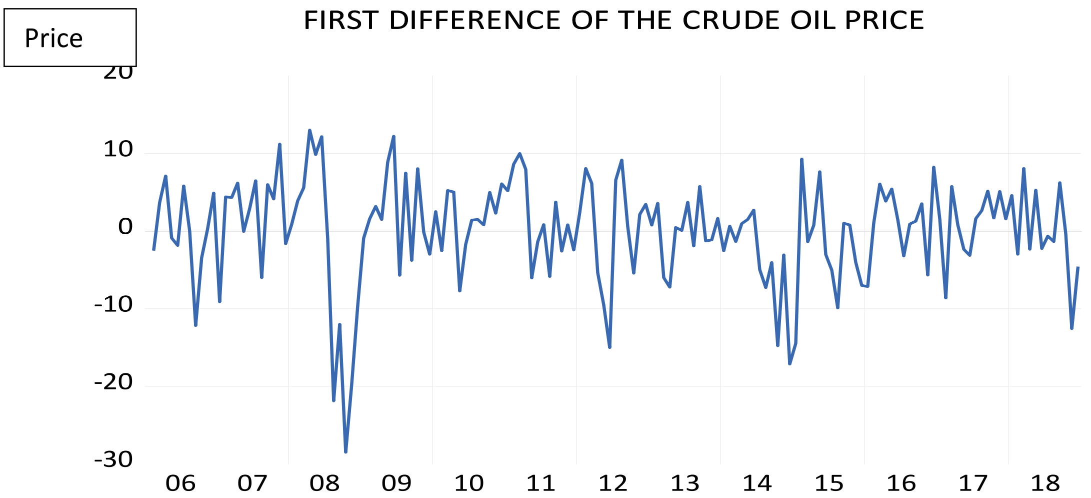


Fig. 2. First Differenced plot of Crude Oil Price

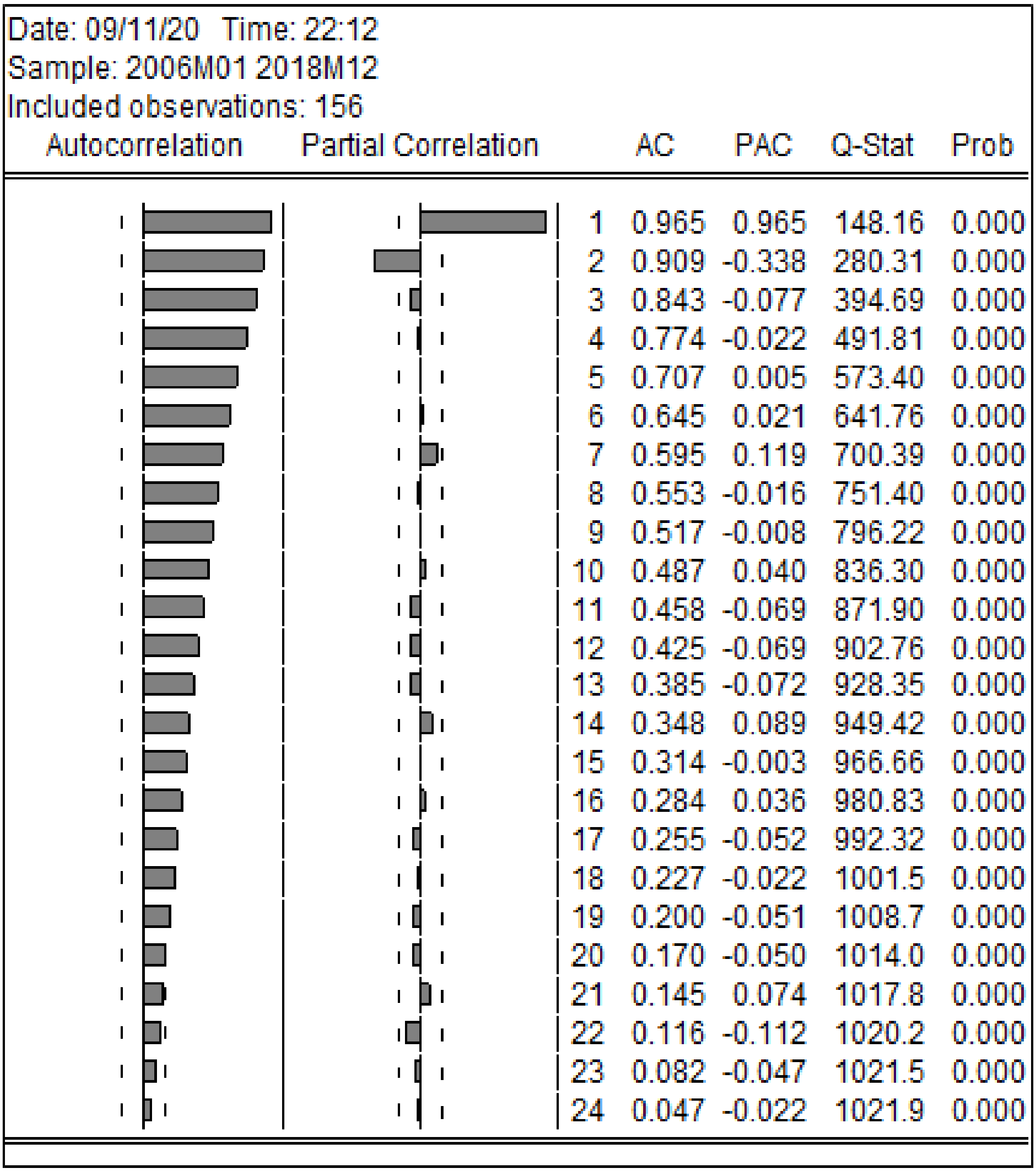
Table 2. First differencing of the crude oil price

Null Hypothesis: D(CRUDE_OIL_PRICE) has a unit root				
Exogenous: Constant				
Lag Length: 0 (Automatic - based on AIC, maxlag=13)				
			t-Statistic	**Prob.***
Augmented Dickey-Fuller test statistic			-8.528964	0.0000
Test critical values:	1% level		-3.473096	
	5% level		-2.880211	
	10% level		-2.576805	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(CRUDE_OIL_PRICE,2)				
Method: Least Squares				
Date: 09/11/20 Time: 22:18				
Sample (adjusted): 2006M03 2018M12				
Included observations: 154 after adjustments				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
D(CRUDE_OIL_PRICE(-1))	-0.648421	0.076026	-8.528964	0.0000
C	-0.001905	0.507057	-0.003756	0.9970
R-squared	0.323673	Mean dependent var		-0.013442
Adjusted R-squared	0.319223	S.D. dependent var		7.626281
S.E. of regression	6.292384	Akaike info criterion		6.529459
Sum squared resid	6018.302	Schwarz criterion		6.568900
Log likelihood	-500.7683	Hannan-Quinn criter.		6.545480
F-statistic	72.74323	Durbin-Watson stat		2.049725
Prob(F-statistic)	0.000000			

4.2 Model Identification

In identifying an appropriate model, we observe the correlogram plot of the crude oil price data, by checking for autocorrelation and partial autocorrelation which shows that Nigeria's crude oil price data in Fig. 3, exhibit a form of autocorrelation. Hence, we will proceed to check the correlogram of the first differenced data in Fig.e 4.

CORRELOGRAM PLOT OF CRUDE OIL PRICE DATA (LEVEL)


Date: 09/11/20 Time: 22:12
Sample: 2006M01 2018M12
Included observations: 156

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.965	0.965	148.16	0.000
		2	0.909	-0.338	280.31	0.000
		3	0.843	-0.077	394.69	0.000
		4	0.774	-0.022	491.81	0.000
		5	0.707	0.005	573.40	0.000
		6	0.645	0.021	641.76	0.000
		7	0.595	0.119	700.39	0.000
		8	0.553	-0.016	751.40	0.000
		9	0.517	-0.008	796.22	0.000
		10	0.487	0.040	836.30	0.000
		11	0.458	-0.069	871.90	0.000
		12	0.425	-0.069	902.76	0.000
		13	0.385	-0.072	928.35	0.000
		14	0.348	0.089	949.42	0.000
		15	0.314	-0.003	966.66	0.000
		16	0.284	0.036	980.83	0.000
		17	0.255	-0.052	992.32	0.000
		18	0.227	-0.022	1001.5	0.000
		19	0.200	-0.051	1008.7	0.000
		20	0.170	-0.050	1014.0	0.000
		21	0.145	0.074	1017.8	0.000
		22	0.116	-0.112	1020.2	0.000
		23	0.082	-0.047	1021.5	0.000
		24	0.047	-0.022	1021.9	0.000

Fig. 3. Correlogram plot of crude oil price (Level)

CORRELOGRAM OF THE CRUDE OIL PRICE (FIRST DIFFERENCE)

Date: 09/11/20 Time: 22:16
Sample (adjusted): 2006M02 2018M12
Included observations: 155 after adjustments

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.351	0.351	19.414	0.000
		2	0.191	0.078	25.234	0.000
		3	0.061	-0.033	25.823	0.000
		4	-0.047	-0.083	26.186	0.000
		5	-0.078	-0.044	27.173	0.000
		6	-0.177	-0.138	32.317	0.000
		7	-0.117	-0.003	34.565	0.000
		8	-0.131	-0.064	37.388	0.000
		9	-0.105	-0.039	39.231	0.000
		10	-0.025	0.027	39.333	0.000
		11	0.074	0.090	40.251	0.000
		12	0.065	-0.016	40.981	0.000
		13	-0.034	-0.115	41.178	0.000
		14	-0.041	-0.046	41.470	0.000
		15	-0.053	-0.030	41.959	0.000
		16	-0.034	0.005	42.158	0.000
		17	-0.026	0.005	42.274	0.001
		18	0.008	0.027	42.284	0.001
		19	0.032	0.015	42.463	0.002
		20	-0.043	-0.081	42.796	0.002
		21	0.061	0.078	43.472	0.003
		22	0.078	0.028	44.595	0.003
		23	0.056	-0.013	45.182	0.004
		24	-0.066	-0.118	45.992	0.004

Fig. 4. Correlogram of the crude oil price (First Difference)

4.3 Arima Model

The determination of an appropriate ARIMA model in modelling the crude oil data include comparing several models as seen in Table 3 and Table 4, ARIMA (3, 1, 2) was selected using the AIC at 6.489533, BIC at 6.589002, and HQ at 6.529941 which all have the lowest value among compared models as seen in Table 3. The model coefficients are shown in table 5, which are all significant at 5% level of significance and the agreed model is fitted to the actual data in Fig. 5.

Table 3. ARIMA model comparison

Model	LogL	AIC*	BIC	HQ
(3,2)(0,0)	-488.2045	6.489533	6.589002	6.529941
(1,0)(0,0)	-503.654619	6.537479	6.596384	6.561405
(3,1)(0,0)	-500.987271	6.541771	6.659581	6.589623
(2,0)(0,0)	-503.145167	6.543809	6.622349	6.57571
(1,1)(0,0)	-503.270427	6.545425	6.623965	6.577326
(2,2)(0,0)	-501.338481	6.546303	6.664113	6.594155
(0,2)(0,0)	-503.426575	6.54744	6.62598	6.579341
(0,3)(0,0)	-502.826698	6.552603	6.650778	6.592479
(1,2)(0,0)	-502.977576	6.554549	6.652724	6.594426
(1,4)(0,0)	-501.005067	6.554904	6.692349	6.610731
(3,0)(0,0)	-503.05197	6.555509	6.653684	6.595386
(2,1)(0,0)	-503.11499	6.556322	6.654497	6.596199
(4,1)(0,0)	-501.332216	6.559125	6.69657	6.614952
(2,3)(0,0)	-501.335972	6.559174	6.696619	6.615001
(4,2)(0,0)	-500.391303	6.559888	6.716968	6.62369
(4,0)(0,0)	-502.44998	6.560645	6.678455	6.608497
(0,1)(0,0)	-505.758377	6.564624	6.623529	6.58855
(1,3)(0,0)	-502.826667	6.565505	6.683315	6.613357
(0,4)(0,0)	-502.826689	6.565506	6.683316	6.613357
(2,4)(0,0)	-500.828545	6.56553	6.72261	6.629332
(3,3)(0,0)	-500.893321	6.566365	6.723445	6.630168
(4,3)(0,0)	-500.299839	6.571611	6.748326	6.643388
(3,4)(0,0)	-500.30463	6.571673	6.748388	6.64345
(4,4)(0,0)	-500.288525	6.584368	6.780718	6.664121
(0,0)(0,0)	-513.791867	6.655379	6.694649	6.67133

Table 4. The arima model estimation

Dependent Variable: D(CRUDE_OIL_PRICE)
Method: ARMA Conditional Least Squares (Marquardt - EViews legacy)
Date: 09/19/20 Time: 05:42
Sample (adjusted): 2006M05 2018M12
Included observations: 152 after adjustments
Convergence achieved after 66 iterations
MA Backcast: 2006M03 2006M04

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.173187	0.080589	2.149021	0.0333
AR(2)	-0.835803	0.034454	-24.25837	0.0000
AR(3)	0.334680	0.074231	4.508624	0.0000
MA(1)	0.194126	0.028771	6.747186	0.0000
MA(2)	0.966249	0.021738	44.44915	0.0000
R-squared	0.193480	Mean dependent var		-0.066382
Adjusted R-squared	0.171534	S.D. dependent var		6.710991
S.E. of regression	6.108352	Akaike info criterion		6.489533
Sum squared resid	5484.858	Schwarz criterion		6.589002
Log likelihood	-488.2045	Hannan-Quinn criter.		6.529941
Durbin-Watson stat	2.072039			
Inverted AR Roots	.37	-.10+.95i	-.10-.95i	
Inverted MA Roots	-.10+.98i	-.10-.98i		

Table 5. Model coefficients

Coefficient Confidence Intervals
Date: 09/19/20 Time: 05:44
Sample: 2006M01 2021M12
Included observations: 152

		90% CI		95% CI		99% CI	
Variable	**Coefficient**	**Low**	**High**	**Low**	**High**	**Low**	**High**
AR(1)	0.173187	0.039790	0.306585	0.013925	0.332450	-0.037125	0.383499
AR(2)	-0.835803	-0.892835	-0.778772	-0.903893	-0.767714	-0.925718	-0.745889
AR(3)	0.334680	0.211806	0.457553	0.187982	0.481377	0.140960	0.528399
MA(1)	0.194126	0.146501	0.241750	0.137267	0.250984	0.119041	0.269210
MA(2)	0.966249	0.930266	1.002232	0.923289	1.009209	0.909519	1.022979

FITTING THE ARIMA MODEL

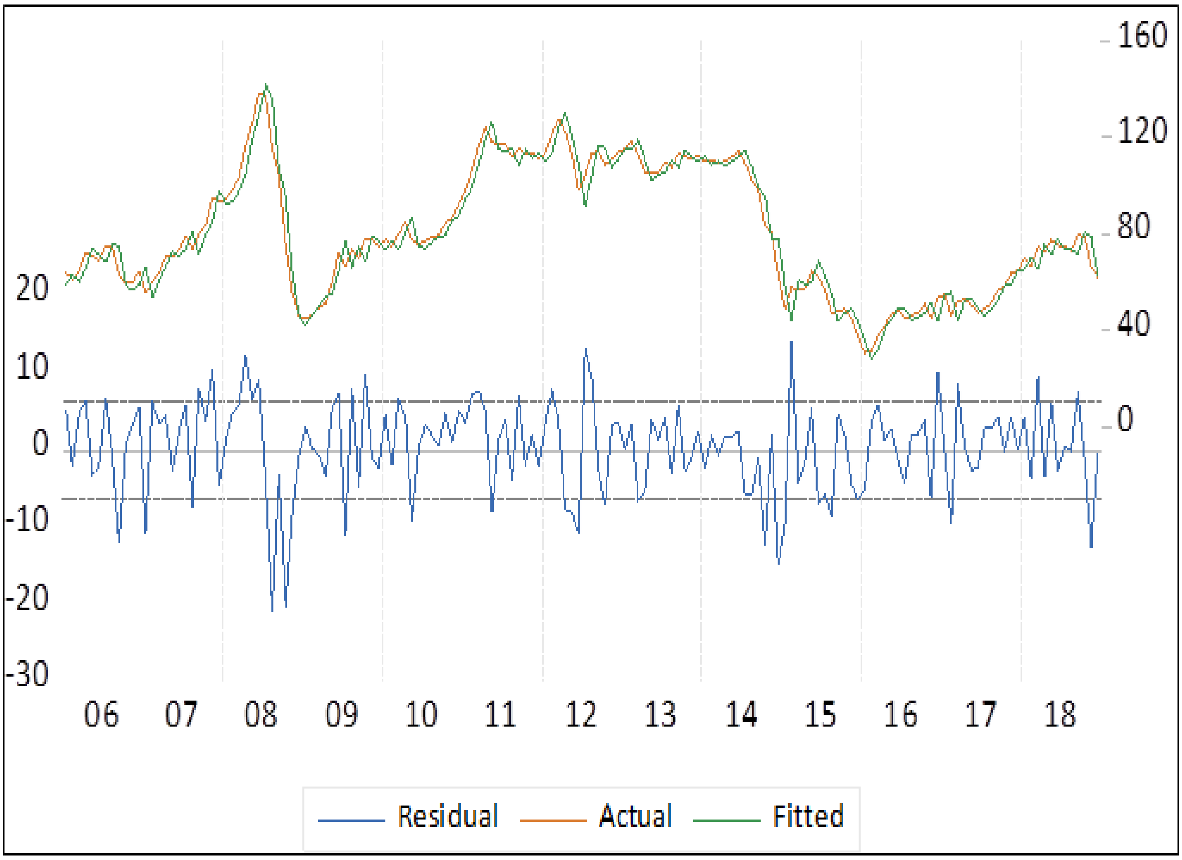


Fig. 5. Fitting the ARIMA model

LJUNG-BOX Q TEST OF THE RESIDUAL

Date: 09/19/20 Time: 16:10
Sample (adjusted): 2006M02 2018M12
Q-statistic probabilities adjusted for 5 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.002	0.002	0.0006	
		2	-0.002	-0.002	0.0013	
		3	0.007	0.007	0.0082	
		4	-0.012	-0.012	0.0296	
		5	0.050	0.051	0.4429	
		6	-0.069	-0.069	1.2200	0.269
		7	0.037	0.038	1.4399	0.487
		8	-0.032	-0.034	1.6073	0.658
		9	-0.055	-0.052	2.1103	0.715
		10	-0.010	-0.014	2.1268	0.831
		11	0.097	0.106	3.7179	0.715
		12	0.069	0.060	4.5334	0.717
		13	-0.069	-0.063	5.3438	0.720
		14	-0.046	-0.049	5.7041	0.769
		15	-0.048	-0.051	6.0989	0.807
		16	-0.020	-0.026	6.1656	0.862
		17	-0.024	-0.020	6.2696	0.902
		18	0.009	0.016	6.2854	0.935
		19	0.037	0.034	6.5260	0.951
		20	-0.111	-0.097	8.7507	0.890
		21	0.040	0.045	9.0438	0.912
		22	0.054	0.039	9.5701	0.921
		23	0.053	0.032	10.090	0.929
		24	-0.111	-0.121	12.361	0.870
		25	0.046	0.079	12.750	0.888
		26	-0.073	-0.091	13.764	0.880
		27	-0.040	-0.020	14.062	0.899
		28	-0.018	-0.029	14.121	0.923
		29	-0.094	-0.090	15.811	0.895
		30	0.052	0.017	16.346	0.904
		31	-0.034	0.011	16.567	0.922
		32	-0.053	-0.059	17.125	0.928
		33	0.049	0.021	17.607	0.936
		34	0.086	0.086	19.084	0.919
		35	-0.050	-0.066	19.597	0.927
		36	-0.062	-0.049	20.376	0.927

Fig. 6. Ljung-Box Q Test

4.4 ARIMA Model Forecast

We use the ARIMA (3, 1, 2) model to forecast the crude oil price data from January, 2018 to December, 2018 comparing it with the actual data. As seen below in Fig. 7 and Table 6, the Thiel inequality coefficient has a low value of 0.192 meaning that our model does have a good forecasting ability.

Model Forecast

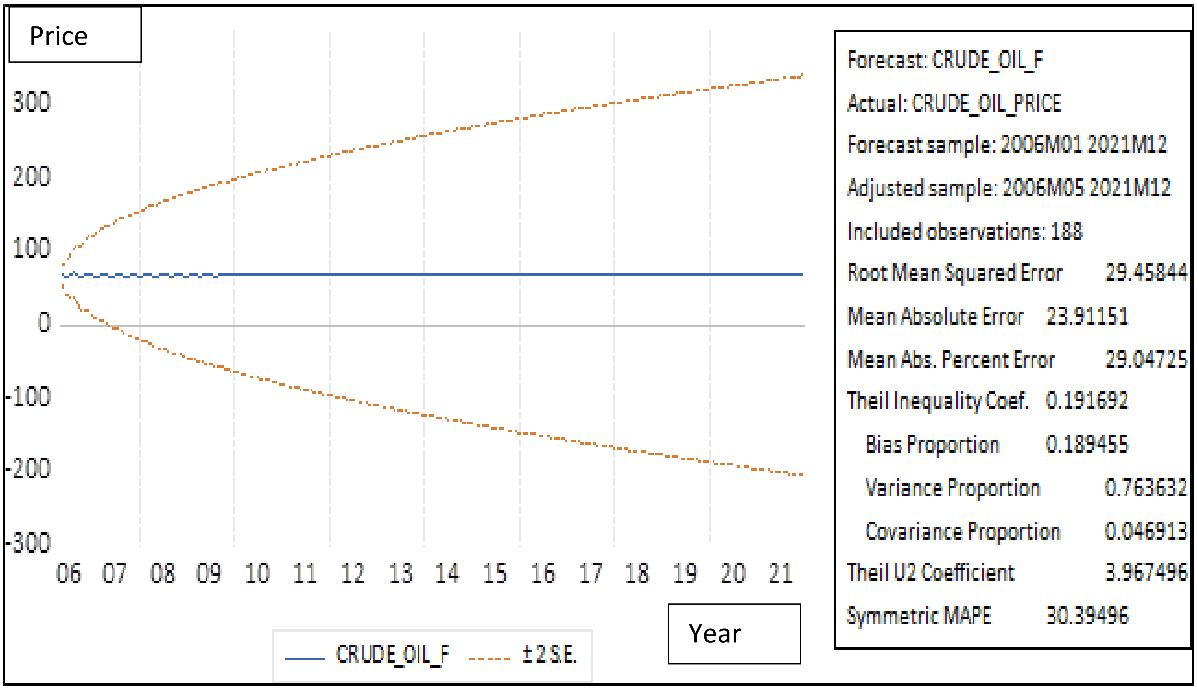


Fig. 7. ARIMA Model Forecast

Table 6. Arima model forecast

Year/ month	F_COP	Year/ month	F_COP	Year/ month	F_COP	Year/month	F_COP
2006M01	63.85	2009M04	51.16	2012M07	104.62	2015M10	48.86
2006M02	61.33	2009M05	60.02	2012M08	113.76	2015M11	44.82
2006M03	65	2009M06	72.24	2012M09	114.36	2015M12	37.8
2006M04	72.09	2009M07	66.52	2012M10	108.92	2016M01	30.66
2006M05	71.18	2009M08	74	2012M11	111.05	2016M02	31.7
2006M06	69.32	2009M09	70.22	2012M12	114.49	2016M03	37.76
2006M07	75.13	2009M10	78.25	2013M01	115.24	2016M04	41.59
2006M08	75.15	2009M11	78.11	2013M02	118.81	2016M05	47.01
2006M09	62.97	2009M12	75.11	2013M03	112.79	2016M06	48.46
2006M10	59.49	2010M01	77.62	2013M04	105.55	2016M07	45.25
2006M11	59.81	2010M02	75.06	2013M05	106	2016M08	46.15
2006M12	64.7	2010M03	80.27	2013M06	106.06	2016M09	47.43
2007M01	55.57	2010M04	85.29	2013M07	109.78	2016M10	50.94
2007M02	59.97	2010M05	77.54	2013M08	107.84	2016M11	45.25
2007M03	64.28	2010M06	75.79	2013M09	113.59	2016M12	53.48
2007M04	70.46	2010M07	77.18	2013M10	112.29	2017M01	55.01
2007M05	70.4	2010M08	78.67	2013M11	111.14	2017M02	46.39
2007M06	73.28	2010M09	79.45	2013M12	112.75	2017M03	52.13
2007M07	79.76	2010M10	84.42	2014M01	110.19	2017M04	52.94
2007M08	73.76	2010M11	86.71	2014M02	110.83	2017M05	50.57

Year/ month	F_COP	Year/ month	F_COP	Year/ month	F_COP	Year/month	F_COP
2007M09	79.76	2010M12	92.79	2014M03	109.47	2017M06	47.42
2007M10	83.86	2011M01	97.96	2014M04	110.41	2017M07	49.01
2007M11	95.05	2011M02	106.57	2014M05	111.9	2017M08	51.64
2007M12	93.4	2011M03	116.56	2014M06	114.6	2017M09	56.79
2008M01	94.26	2011M04	124.49	2014M07	109.63	2017M10	58.46
2008M02	98.15	2011M05	118.43	2014M08	102.33	2017M11	63.56
2008M03	103.73	2011M06	117.03	2014M09	98.27	2017M12	65.11
2008M04	116.73	2011M07	117.86	2014M10	83.5	2018M01	69.68
2008M05	126.57	2011M08	111.99	2014M11	80.42	2018M02	66.67
2008M06	138.74	2011M09	115.73	2014M12	63.28	2018M03	74.72
2008M07	137.74	2011M10	113.12	2015M01	48.81	2018M04	72.37
2008M08	115.84	2011M11	113.92	2015M02	58.09	2018M05	77.64
2008M09	103.82	2011M12	111.46	2015M03	56.69	2018M06	75.38
2008M10	75.31	2012M01	113.81	2015M04	57.45	2018M07	74.72
2008M11	55.51	2012M02	121.87	2015M05	65.08	2018M08	73.35
2008M12	45.87	2012M03	128	2015M06	62.06	2018M09	79.59
2009M01	44.95	2012M04	122.62	2015M07	57.01	2018M10	79.18
2009M02	46.52	2012M05	113.08	2015M08	47.09	2018M11	66.59
2009M03	49.7	2012M06	98.06	2015M09	48.08	2018M12	62

Year/month	F_COP	Year/month	F_COP
2019M01	60.92066752	2020M09	59.43122685
2019M02	59.95943544	2020M10	59.27630292
2019M03	59.15889232	2020M11	59.49532437
2019M04	59.46241886	2020M12	59.59316357
2019M05	59.86237762	2021M01	59.37519931
2019M06	59.41003135	2021M02	59.32897836
2019M07	59.098988	2021M03	59.53589356
2019M08	59.55704988	2021M04	59.53741205
2019M09	59.74496037	2021M05	59.34926539
2019M10	59.29055449	2021M06	59.38466194
2019M11	59.20810495	2021M07	59.54855399
2019M12	59.63650955	2021M08	59.4843846
2020M01	59.62753497	2021M09	59.34813623
2020M02	59.24032447	2021M10	59.4330241
2020M03	59.32414384	2021M11	59.54012625
2020M04	59.65928854	2021M12	59.44212585
2020M05	59.51768333		
2020M06	59.24109666		
2020M07	59.42371561		
2020M08	59.63912259		

4.5 Test for ARCH Effect

The ARCH effect test is presented in Table 7 showing the presence of ARCH at lag 2.

RESIDUAL TEST FOR ARCH EFFECT OF THE ARIMA MODEL AT LAG (2)

Table 7. ARCH TEST

Heteroskedasticity Test: ARCH

F-statistic	12.31252	Prob. F(2,147)	0.0000
Obs*R-squared	21.52224	Prob. Chi-Square(2)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 04/25/20 Time: 17:49
Sample (adjusted): 2006M07 2018M12
Included observations: 150 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	22.34000	6.434762	3.471768	0.0007
RESID^2(-1)	0.013983	0.076373	0.183084	0.8550
RESID^2(-2)	0.379979	0.076751	4.950793	0.0000

R-squared	0.143482	Mean dependent var	36.43912
Adjusted R-squared	0.131828	S.D. dependent var	67.79828
S.E. of regression	63.17155	Akaike info criterion	11.14938
Sum squared resid	586624.7	Schwarz criterion	11.20960
Log likelihood	-833.2037	Hannan-Quinn criter.	11.17384
F-statistic	12.31252	Durbin-Watson stat	1.949202
Prob(F-statistic)	0.000011		

Table 7, shows the arch effect test using chi-square which is statistically significant with the LM statistic at 21.52224 and p-value of 0.0000. The coefficient at lag 2 (b_2) is also statistically significant at 1% level. Hence, we reject the null hypothesis and conclude that ARCH effect is present.

4.6 GARCH Model

The GARCH model is shown in Table 8 with significant parameters and an Akaike info criterion value of 6.405837 and the GARCH model coefficients are shown in Table 9 and fitted to the actual data as seen in Fig. 8.

Table 8. GARCH model estimation

Dependent Variable: D(CRUDE_OIL_PRICE)
Method: ML ARCH - Normal distribution (Marquardt / EViews legacy)
Date: 09/19/20 Time: 05:55
Sample (adjusted): 2006M05 2018M12
Included observations: 152 after adjustments
Convergence achieved after 24 iterations
MA Backcast: 2006M03 2006M04
Presample variance: backcast (parameter = 0.7)
GARCH = C(6) + C(7)*RESID(-1)^2 + C(8)*GARCH(-1)

Variable	**Coefficient**	**Std. Error**	**z-Statistic**	**Prob.**
AR(1)	0.164278	0.100787	1.629950	0.1031
AR(2)	-0.879695	0.029410	-29.91126	0.0000
AR(3)	0.211267	0.099401	2.125408	0.0336
MA(1)	0.045181	0.011552	3.910959	0.0001
MA(2)	0.969345	0.008821	109.8947	0.0000
	Variance Equation			

C	5.338590	3.272477	1.631361	0.1028
RESID(-1)^2	0.303731	0.127579	2.380741	0.0173
GARCH(-1)	0.576761	0.136306	4.231360	0.0000
R-squared	0.158620	Mean dependent var		-0.066382
Adjusted R-squared	0.135725	S.D. dependent var		6.710991
S.E. of regression	6.238966	Akaike info criterion		6.405837
Sum squared resid	5721.931	Schwarz criterion		6.564989
Log likelihood	-478.8436	Hannan-Quinn criter.		6.470490
Durbin-Watson stat	1.739239			
Inverted AR Roots	.24	-.04+.95i	-.04-.95i	
Inverted MA Roots	-.02+.98i	-.02-.98i		

Table 9. GARCH model coefficients

Coefficient Confidence Intervals
Date: 09/19/20 Time: 05:55
Sample: 2006M01 2021M12
Included observations: 152

Variable	Coefficient	90% CI		95% CI		99% CI	
		Low	High	Low	High	Low	High
AR(1)	0.164278	-0.002576	0.331131	-0.034935	0.363490	-0.098817	0.427372
AR(2)	-0.879695	-0.928384	-0.831007	-0.937827	-0.821564	-0.956468	-0.802923
AR(3)	0.211267	0.046709	0.375826	0.014794	0.407741	-0.048209	0.470744
MA(1)	0.045181	0.026056	0.064305	0.022347	0.068015	0.015024	0.075337
MA(2)	0.969345	0.954742	0.983948	0.951910	0.986780	0.946319	0.992370
C	5.338590	-0.079010	10.75619	-1.129707	11.80689	-3.203891	13.88107
RESID(-1)^2	0.303731	0.092525	0.514938	0.051563	0.555900	-0.029300	0.636763
GARCH(-1)	0.576761	0.351105	0.802416	0.307341	0.846180	0.220947	0.932575

4.7 GARCH model fitting

Fig. 8. Fitting the GARCH model

4.8 ARCH LM test

Table 10. ARCH LM test

Heteroskedasticity Test: ARCH				
F-statistic	1.330058	Prob. F(1,149)		0.2506
Obs*R-squared	1.335986	Prob. Chi-Square(1)		0.2477
Test Equation:				
Dependent Variable: WGT_RESID^2				
Method: Least Squares				
Date: 09/20/20 Time: 15:42				
Sample (adjusted): 2006M06 2018M12				
Included observations: 151 after adjustments				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
C	1.091387	0.144493	7.553198	0.0000
WGT_RESID^2(-1)	-0.094185	0.081667	-1.153281	0.2506
R-squared	0.008848	Mean dependent var		0.996731
Adjusted R-squared	0.002196	S.D. dependent var		1.462922
S.E. of regression	1.461315	Akaike info criterion		3.609708
Sum squared resid	318.1809	Schwarz criterion		3.649671
Log likelihood	-270.5329	Hannan-Quinn criter.		3.625943
F-statistic	1.330058	Durbin-Watson stat		1.946103
Prob(F-statistic)	0.250642			

From Table 10 above, it is shown that the result is not statistically significant and hence, we conclude that there is no presence of autocorrelation in the residual.

4.9 GARCH forecast

The forecast using the GARCH (1, 1) model is given Fig. 9 and Table 11, and it's considered to perform well in forecasting the crude oil price, given a Theil inequality coefficient of 0.191436.

4.10 Model out of sample forecast

using eviews 11 for each month in 2019, 2020, 2021 is presented below for the model GARCH (1, 1) with the average forecast staying a little above 68 dollars per barrel.

4.11 Model comparison

Diebold-Mariano test, shown in Table 12 and Fig. 11, was used in comparing the ARIMA and GARCH model and with RMSE, MAE, SMAPE, and, Theil U1 as our measure of accuracy of forecast, we have shown that the GARCH model performs better than the ARIMA model in forecasting crude oil prices.

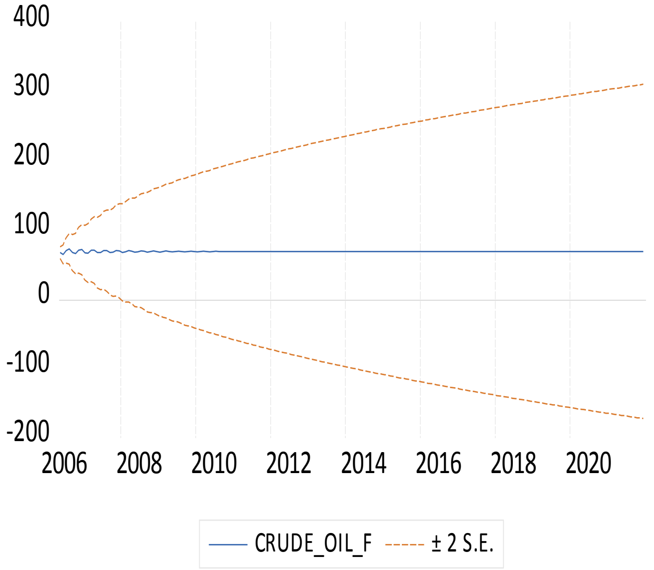
From Table 12, comparing RMSE, MAE, SMAPE, Theil U1 and AIC of both GARCH and ARIMA models and using these measures to determine the accuracy for our forecast, it is shown that future forecast is best carried out with the GARCH model in forecasting Crude Oil Prices which agrees with other researches in this field including, Yaziz et al. [21].

Table 11. GARCH model forecast

Year/ month	**F_COP**	**Year/month**	**F_COP**	**Year/month**	**F_COP**	**Year/month**	**F_COP**
2006M01	N/A	2009M01	67.57287047	2012M01	68.36148556	2015M01	68.40593938
2006M02	N/A	2009M02	68.38002161	2012M02	68.49753211	2015M02	68.39819796
2006M03	N/A	2009M03	69.12660561	2012M03	68.4112126	2015M03	68.37844919
2006M04	N/A	2009M04	68.34969371	2012M04	68.29540596	2015M04	68.38679909
2006M05	66.69138094	2009M05	67.73582293	2012M05	68.3810586	2015M05	68.40390818
2006M06	64.14242724	2009M06	68.47615228	2012M06	68.47876744	2015M06	68.39520118
2006M07	69.97071656	2009M07	68.97365476	2012M07	68.39500443	2015M07	68.38048409
2006M08	72.02992353	2009M08	68.27442818	2012M08	68.31338566	2015M08	68.38934051
2006M09	66.70257605	2009M09	67.8783179	2012M09	68.39430614	2015M09	68.40190246
2006M10	65.24726572	2009M10	68.5334582	2012M10	68.46170281	2015M10	68.39306592
2006M11	70.1296762	2009M11	68.84181554	2012M11	68.38434582	2015M11	68.38243466
2006M12	71.08648109	2009M12	68.23246276	2012M12	68.32944514	2015M12	68.39111557
2007M01	66.64116999	2010M01	67.99950912	2013M01	68.40271549	2016M01	68.40002705
2007M02	66.10070285	2010M02	68.56243064	2013M02	68.446705	2016M02	68.39160841
2007M03	70.12457671	2010M03	68.7310978	2013M03	68.3778772	2016M03	68.38422002
2007M04	70.32190572	2010M04	68.21439117	2013M04	68.34335261	2016M04	68.39229482
2007M05	66.70035689	2010M05	68.10005916	2013M05	68.40752203	2016M05	68.39834227
2007M06	66.78194184	2010M06	68.57145519	2013M06	68.43389367	2016M06	68.39067145
2007M07	70.02289268	2010M07	68.64030901	2013M07	68.37448249	2016M07	68.38579733
2007M08	69.71842292	2010M08	68.2127807	2013M08	68.35508047	2016M08	68.39302224
2007M09	66.83459278	2010M09	68.18156765	2013M09	68.40972835	2016M09	68.39687627
2007M10	67.31339217	2010M10	68.56708122	2013M10	68.42322199	2016M10	68.39012394
2007M11	69.86461511	2010M11	68.5675476	2013M11	68.37326619	2016M11	68.3871507
2007M12	69.25326685	2010M12	68.22189549	2013M12	68.36473461	2016M12	68.39341649
2008M01	67.00969224	2011M01	68.24614879	2014M01	68.4101297	2017M01	68.39563481
2008M02	67.71791375	2011M02	68.5543001	2014M02	68.41453826	2017M02	68.3898591
2008M03	69.67876244	2011M03	68.51056189	2014M03	68.37352619	2017M03	68.38828259
2008M04	68.90387256	2011M04	68.2374214	2014M04	68.37250115	2017M04	68.39357313
2008M05	67.20125054	2011M05	68.29612917	2014M05	68.40934226	2017M05	68.39460888
2008M06	68.01747833	2011M06	68.53681346	2014M06	68.40763164	2017M06	68.3897919
2008M07	69.48564563	2011M07	68.46700184	2014M07	68.37472512	2017M07	68.38920716

Year/ month	F_COP	Year/month	F_COP	Year/month	F_COP	Year/month	F_COP
2008M08	68.64909232	2011M08	68.2562076	2014M08	68.37860747	2017M08	68.39356739
2008M09	67.39256804	2011M09	68.33384053	2014M09	68.40783156	2017M09	68.3937804
2008M10	68.23223716	2011M10	68.51727964	2014M10	68.40226506	2017M10	68.38985618
2008M11	69.29879783	2011M11	68.43458729	2014M11	68.37646253	2017M11	68.38994531
2008M12	68.46989426	2011M12	68.2760336	2014M12	68.38329468	2017M12	68.39345707

Year/month	F_COP	Year/month	F_COP	Year/month	F_COP	Year/month	F_COP
2018M01	68.39312651	2019M01	68.39196793	2020M01	68.39156206	2021M01	68.39149533
2018M02	68.39000176	2019M02	68.39060657	2020M02	68.39113299	2021M02	68.39146159
2018M03	68.39052114	2019M03	68.39151666	2020M03	68.39184295	2021M03	68.39187933
2018M04	68.39328546	2019M04	68.39267218	2020M04	68.39217695	2021M04	68.39187975
2018M05	68.39262252	2019M05	68.39177379	2020M05	68.39151661	2021M05	68.39150522
2018M06	68.39019158	2019M06	68.39080198	2020M06	68.39126431	2021M06	68.39153157
2018M07	68.39095943	2019M07	68.39167676	2020M07	68.39187432	2021M07	68.39186546
2018M08	68.393084	2019M08	68.39248557	2020M08	68.39205697	2021M08	68.39181801
2018M09	68.39224396	2019M09	68.39164358	2020M09	68.39149705	2021M09	68.39152205
2018M10	68.39039922	2019M10	68.39097857	2020M10	68.39137327	2021M10	68.39158572
2018M11	68.39128399	2019M11	68.39178089	2020M11	68.39188408	2021M11	68.39184651
2018M12	68.39287468	2019M12	68.39231981	2020M12	68.39195859	2021M12	68.39177081



Forecast: CRUDE_OIL_F
Actual: CRUDE_OIL_PRICE
Forecast sample: 2006M01 2021M12
Adjusted sample: 2006M05 2021M12
Included observations: 188

Root Mean Squared Error	29.43081
Mean Absolute Error	23.91251
Mean Abs. Percent Error	29.08211
Theil Inequality Coef.	0.191436
Bias Proportion	0.188017
Variance Proportion	0.763376
Covariance Proportion	0.048607
Theil U2 Coefficient	3.969711
Symmetric MAPE	30.39763

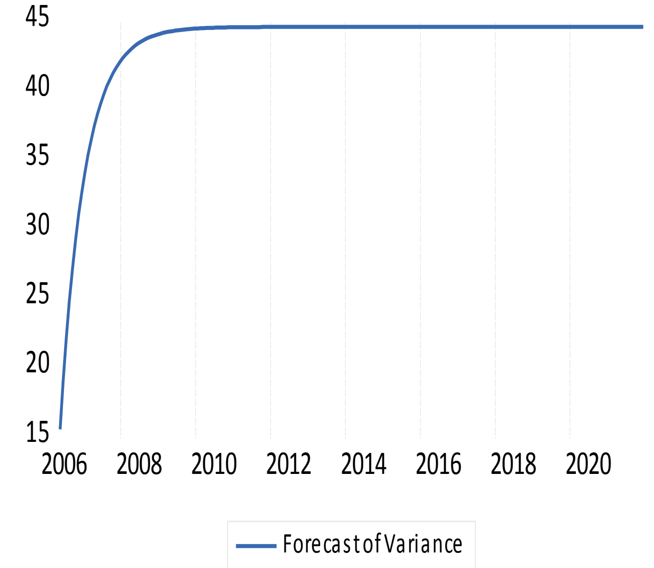


Fig. 9. GARCH Model Forecast

Table 12. Forecast evaluation

Date: 09/19/20 Time: 18:15
Sample: 2006M01 2021M12
Included observations: 192
Evaluation sample: 2006M01 2021M12
Training sample: 2006M05 2019M06
Number of forecasts: 7

Combination tests

Null hypothesis: Forecast i includes all information contained in others

Forecast	F-stat	F-prob
CRUDE_OIL_F_ARIMA	0.032203	0.8578
CRUDE_OIL_F_GARCH	0.070271	0.7913

Diebold-Mariano test (HLN adjusted)
Null hypothesis: Both forecasts have the same accuracy

Accuracy	**Statistic**	**<> prob**	**> prob**	**< prob**		
Abs Error	2.932408	0.0039	0.9981	0.0019		
Sq Error	5.346955	0.0000	1.0000	0.0000		
Evaluation statistics						
Forecast	RMSE	MAE	MAPE	SMAPE	Theil U1	Theil U2
CRUDE_OIL_F_ARIMA	29.03768	23.23381	28.16163	29.57402	0.190077	3.922870
CRUDE_OIL_F_GARCH	28.20304	22.75043	28.31242	28.90870	0.182148	3.995576
Simple mean	28.60349	22.98214	28.22282	29.22602	0.185977	3.955267
Simple median	28.60349	22.98214	28.22282	29.22602	0.185977	3.955267
Least-squares	28.66389	23.01833	28.21403	29.27591	0.186550	3.950116
Mean square error	28.59133	22.97480	28.22461	29.21591	0.185861	3.956331
MSE ranks	28.46619	22.90239	28.24910	29.11643	0.184670	3.967835

Trimmed mean could not be calculated due to insufficient data

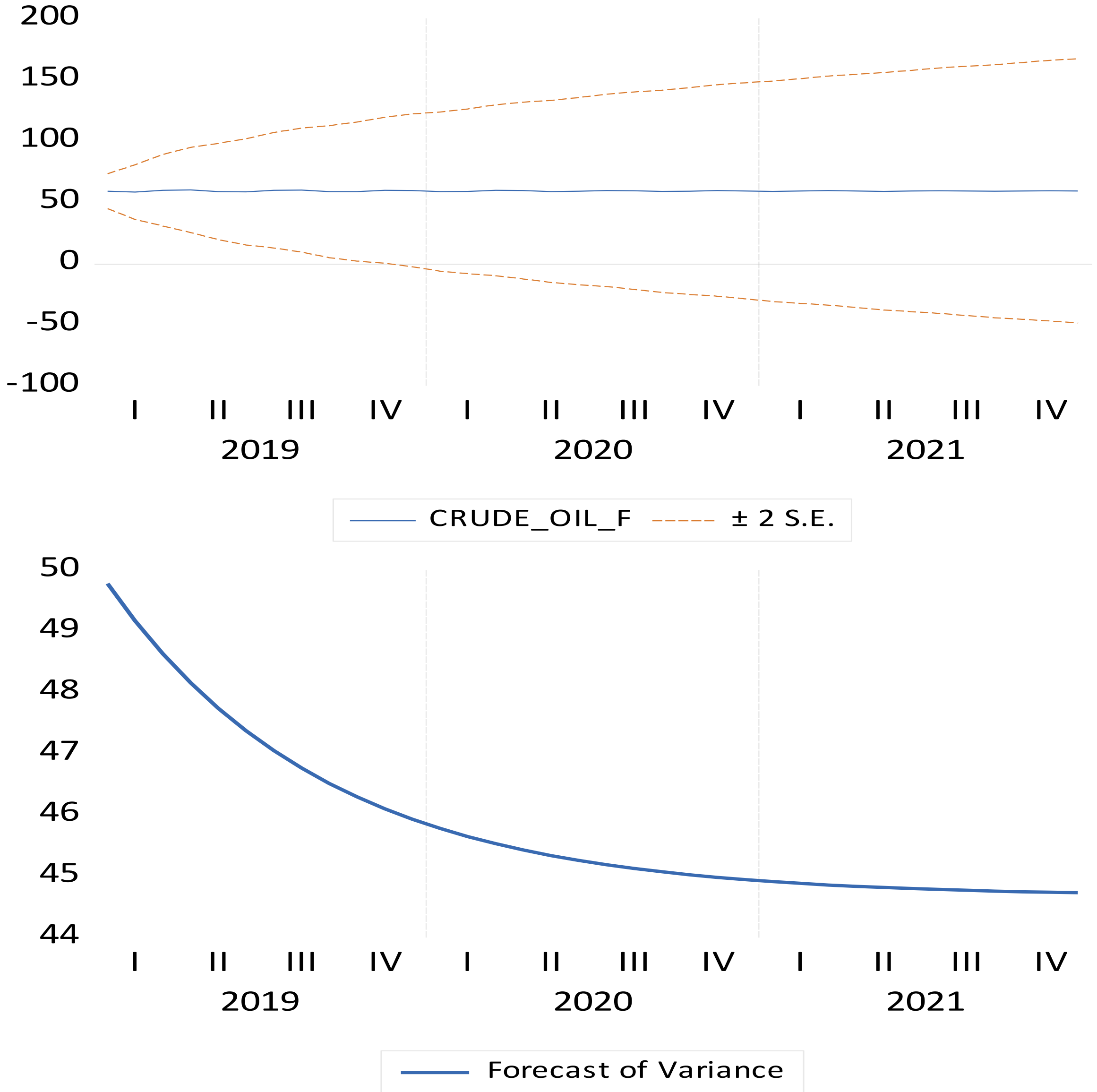


Fig. 10. Model Out of Sample Forecast

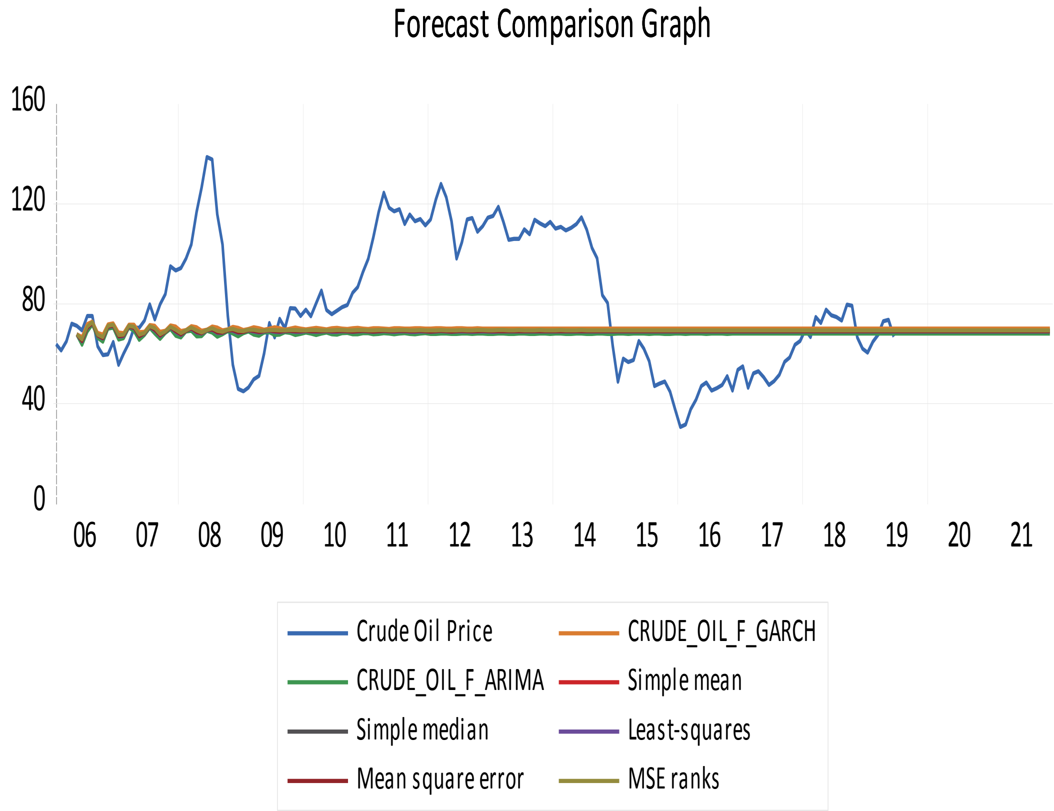


Fig. 11. Forecast comparison graph

5 Conclusion

In order to forecast the Nigeria's crude oil price, we fit two-time series models. In finding out which of these model is better, a question arises "Do the two models give equal forecasting performance"? To get the answer of this question we use Diebold-Mariano tests to compare the forecasting models and relaxing of all the assumptions. Table 12, shows the test values and critical values. The tests show that there is no evidence to reject the null hypothesis that the two models perform equally at 5% level of significance. That is, ARIMA(3, 1, 2) model and GARCH(1, 1) model have same forecasting performance. However, using Table 12, RMSE, MAE, SMAPE, Theil U1 and AIC as our measure of the accuracy for our forecast, it is shown that future forecast is best carried out with the GARCH model in forecasting Crude Oil Prices which agrees with other researches in this field including, Yaziz et al. [21], and Shabri and Samsudin [22].

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Dharam PG. The political economy of oil gas in Africa. The case of Nigeria, Lagos, Nigeria: Taylor and Francis Publishers; 1991.

[2] Akpan EO. Oil Price Shocks and Nigeria's Macro economy. paper presented at the Annual Conference of case. Economic Development in Africa, 22nd-24th March, Oxford; 2009.

[3] Liu LM. Journal of Forecasting. 1991;10 (5):521-547.

[4] Agnolucci P. Volatility in crude oil futures: A comparison of the predictive ability of GARCH and implied volatility models. Energy Economics. 2009;31(2):316-321.

[5] Ahmad MI. Modeling and forecasting Oman crude oil prices using Box-Jenkins techniques. International Journal of Trade and Global Markets. 2012;5:24-30.

[6] Etuk, Ette Harrison. Seasonal ARIMA modelling of Nigerian monthly crude oil prices. Asian Economic and Financial Review. 2013;3(3):333-340.

[7] Sadorsky P. Modeling and forecasting petroleum futures volatility. Energy Economics. 2006;28(4):467-488.

[8] Hou A, Suardi S. A Nonparametric GARCH Model of Crude Oil Price Return Volatility. Energy Economics. 2012;34(2):618-626.

[9] Ahmed RA, Shabri AB. Fitting GARCH Models to Crude Oil Spot Price Data. Life Science Journal. 2013;10(4).

[10] Box GEP, Jenkins GM. "time series analysis: forecasting and control," Holden-day, san Francisco; 1976.

[11] Engle RF. Autoregressive conditional heteroskedasticity with estimates of the variance of u.k. Inflation. Econometrica. 1982;50:987–1008.

[12] Borellslev T. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics. 1986; 31:307–327.

[13] Judge GG, Hill RC, Griffiths WE, Gotkepohl H, Lee T. Introduction to the theory and practice of econometrics.1988;1024 p. Edition:2. Ed.

[14] Somarajan S, Shankar M, Sharma T, Jeyanthi R. Modelling and analysis of volatility in time series data. In: Wang J., Reddy G., Prasad V., Reddy V. (eds) Soft Computing and Signal Processing. Advances in Intelligent Systems and Computing. 2019;898. Springer, Singapore. DOI:https://doi.org/10.1007/978-981-13-3393-4_62

[15] Salisu AA, Fasanya IO. Comparative performance of volatility models for oil price. International Journal of Energy Economics and Policy. 2012;2(3):167-183.

[16] Muhammed G, Umar Bashir F. The relevance of GARCH-family models in forecasting Nigerian oil price volatility. Bullion. 2018;42(2). CBN publication, April-June.

[17] Bamisaye OA, Obiyan AS. Policy analysis of oil sector in Nigeria. European Journal of Social Sciences. 2006.
Available: http://eurojournals.com/ejss%203%201.pdf#page=43.

[18] Uwilingiyimana, Charline, Tu, Joseph, Harerimana Jean. Forecasting inflation in Kenya using ARIMA - GARCH models. International Journal of Management and Commerce Innovations. 2015;3:2348-758515.

[19] Gencer GH, Musoglu Z. Volatility modeling and forecasting of Istanbul gold exchange (IGE). International Journal of Financial Research. Published by Sciedu Press. 2014;5(2):87-101. Available:www.sciedu.ca/ijfr

[20] Yahaya AE, Etuk EH, Kingdom N, Chimee NW. ARIMA model for gross domestic product (GDP): Evidence from Nigeria. Archives of Current Research International. 2020;20(7):49-61.

[21] Yaziz SR, Ahmad MH, Nian LC, Muhammad N. A comparative study on Box-Jenkins and GARCH Models in forecasting crude oil prices. Journal of Applied Science. 2011;11(7):1129-1135. ISSN 1812-5654.

[22] Shabri A, Samsudin R. Crude oil price forecasting based on hybridizing wavelet multiple linear regression model, particle swarm optimization techniques and principal component analysis. Science World Journal. 2014;May 8.
DOI:http://dx.doi.org/10.1155/2014/854520

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/69608