Journal of Mathematics and Statistics Studies
ISSN: 2709-4200
DOI: 10.32996/jmss
Journal Homepage: www.al-kindipublisher.com/index.php/jmss

JMSS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Modeling and Forecasting of Nigeria Crude Oil Production

Acha, Chigozie Kelechi[1] ✉ Amalahu, Christain Chinenye[2] and Eziokwu, C. Emmanuel[3]
[1]Department of Statistics, Michael Okpara University of Agriculture Umudike, Abia State, Nigeria
[2]Department of Mathematics, Faculty of Sciences, University of Agriculture and Environmental Sciences, Umuagwo, Imo State, Nigeria
[3]Departments of Mathematics, College of Physical and Applied Science, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria
Corresponding Author: Acha, Chigozie Kelechi, **E-mail**: acha.kelechi@mouau.edu.ng

| ABSTRACT
This paper assessed comprehensively and systematically the predictive capabilities of the Nigerian Monthly Crude Oil Production forecasting models. To obtain the generality of the empirical results, ARIMA model was used. Some of the frequently used measures of forecast adequacy such as Mean Error (ME), Mean Absolute Error (MAE), Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) were used to evaluate the forecast performance of the chosen models. This study reveals the fact that ARIMA (1, 1, 1) model is the best or optimal model for the period forecasted. The study fitted an appropriate time series models of crude oil production in Nigeria (2005-2022) which provided a useful forecast for quantity of crude oil production and export for the purpose of making reliable budget for the sustenance of the economy. This study reveals the fact that ARIMA (1, 1, 1) model is the best or optimal model for the period forecasted.

| KEYWORDS

Unemployment, Poverty, Insurgency, Autoregressive, Crude oil, Economy, Modeling, Forecasting

| ARTICLE INFORMATION

ACCEPTED: 06 February 2023 **PUBLISHED:** 28 February 2023 **DOI:** 10.32996/jmss.2023.4.1.5

1. Introduction
Crude oil is one of the natural resources to mankind and a major source of energy in Nigeria. It is a vital commodity in world market despite the campaign for green energy and other sources of power. Crude oil is one of the expensive commodities in the international market. Nigerian economy is heavily dependent on crude oil. The volatility in crude price affects all oil-producing countries around the globe either positively or negatively Akinlo and Apanisile, (2015); Abass, Kazeem, and Adedayo (2021). In particular, the global crude oil price crash from 2013 to date affected Nigerian government revenue negatively since crude oil exportation is the Nigerian government major source of revenue (Odupitan, 2017; Adedokun, 2018). But many years back, crude oil exportation created a robust wealth for Nigeria. Then, the Naira appreciated as foreign exchange influxes offset outflows and Nigeria foreign reserves assets increased (Akinyemi et al., 2017; Gylych, et al., 2020). Due to this, productivity declined in all other sectors as the economy of Nigeria solely depends on crude oil exportation and this led to massive migration to cities and widespread poverty in the rural areas. As a result, Nigeria's job market has witnessed a very high degree of unemployment, small wage and pitiable working environments Okoi, (2019). From 1970 to 2020, Nigeria's poverty rate increased from 36% to 70% and in 2020 unemployment is at 30.7%. Based on this, it is believed that oil revenue did not seem to add to the standard of living and create job opportunities but causing social and economic challenges to the Nigerian populace Dapel, (2018). As if the situation is not worse enough, the global economy is further put to a standstill due to the COVID-19 pandemic that cut the whole world unaware since December 2019 till date Otache, (2020). This has crashed the world oil prices and put the Nigerian government in unchartered waters and the worst recession in 40years. The government continues to struggle to revive the economy amidst dwindling oil revenues compounded by unemployment, poverty, insurgency and mismanagement (OECD, 2020).

1.1 Statement of the Problem

It is a known fact that Nigeria crude oil is an important energy source and exceedingly used in all vital sectors of the Nigerian economy with no effective and cost-beneficial alternative, and given that its price dynamics have been comparatively volatile in recent years. A few published studies have been concerned with modeling and forecasting oil price fluctuations and exploring the relationships among the prices of oil and selected macroeconomic variables in developing countries such as Nigeria. In reviewing previous empirical literatures, it can be seen that there is lack of studies on developing countries compared to developed economics. Therefore, this paper seeks to address the gap in this field with updated evidence from Nigeria through modeling and forecasting of Nigeria Crude oil production.

1.2 Objectives of the Study

The main objective of this work is to model and forecast Nigerian Monthly Crude Oil Production: Specifically, the study intends to:

i. estimate the trend and determine the appropriate ARIMA model
ii. estimate the model parameters
iii. make forecast on Bonny light oil production.

2. Literature Review

A research done by Wiri and Tuaneh (2019) Analyzed Intervention models of crude oil prices in Nigeria. The time plot of the series revealed an abrupt increase in the series and this called for intervention models. The knowledge was divided into three classes (actual series, pre-intervention and post-intervention series). The Augmented Dickey-Fuller (ADF) was used to test for unit root on each of the series, and they were all found to be non-stationary at different levels; however, they were non-stationary at the first difference (actual, pre, and post-intervention series). The pre-intervention model that reduced the Akaike Information Criterion (AIC) was the best of the eighteen models that were estimated.

Wiri and Sibeate (2021) modeled Nigeria crude oil prices and compared univariate linear models to univariate nonlinear models. The data for their analysis was gathered from the Central Bank of Nigeria (CBN) Monthly Statistical Bulletin. The upward and downward movement in the series revealed by the time plot suggests that the series exhibit a regime-switching pattern: the cycle of expansion and contraction. At lag one, the Augmented Dickey-Fuller test was used to test for stationarity. For univariate linear ARIMA (p, d, q)) and univariate non-linear MS-AR, seven models were estimated for the linear model and two for the non-linear model. The best model was chosen based on the criterion of least information criterion, AIC (2.006612), SC (2.156581), and the maximum log-likelihood of (-150.5480) for the crude oil prices were used to pick MS-AR (1) for the series. In analysing crude oil prices data, the MS-AR model proposed by Hamilton outperforms the linear autoregressive models proposed by Box- Jenkins. The model was used to predict the series' values over a one-year cycle (12 months).

A research done by Wiri and Tuaneh (2019) analyzed Intervention models of crude oil prices in Nigeria. The time plot of the series revealed an abrupt increase in the series and this called for intervention models. The knowledge was divided into three classes (actual series, pre-intervention and post-intervention series). The Augmented Dickey-Fuller (ADF) was used to test for unit root on each of the series, and they were all found to be non-stationary at different levels; however, they were non-stationary at the first difference (actual, pre, and post-intervention series). The pre-intervention model that reduced the Akaike Information Criterion (AIC) was the best of the eighteen models that were estimated.

Alhassan and Kilishi (2016) modelled Oil price volatility using macroeconomic variables in Nigeria. They used different types of GARCH models to calculate with daily, weekly, and quarterly data in their paper. All of the macroeconomic variables studied (real GDP, interest rate, exchange rate, and oil price) are highly volatile; asymmetric models (TGARCH and EGARCH) outperform symmetric models (GARCH (1, 1) and GARCH – M), and oil price is a major source of macroeconomic instability in Nigeria. The Nigerian economy, as a result, is vulnerable to both internal and external shocks. As a result, they concluded that asymmetric models should be given more weight in dealing with Nigerian macroeconomic volatility, and oil price volatility should be considered an important variable in the study of Nigerian macroeconomic fluctuations.

3. Methodology

A secondary data was collected from Central Bank of Nigeria website to model and forecast Nigerian Monthly Crude Oil Production from the period January 2005 to September, 2022 was used for the analysis.

3.1 Method of Data Analysis

The methodology used in this work is the Box and Jenkins (1976) procedure for fitting autoregressive integrated moving average (ARIMA) model. This method involves a three-stage iterative procedure. These include: model building and model identification, parameter estimation and diagnostic checking of the residuals. The first step in developing a Box-Jenkins model is to determine if the time series is stationary and if there is any significant seasonality (periodicity) in the time series. After stationarity and periodicity have been addressed, the next step is to identify the order of the model. In nonstationary time series, autocorrelation function (ACF) and Partial autocorrelation function (PACF) play an important role in the identification of order of the model to be fitted to the data. Therefore, the major tools used in the identification phase are Plots of the series, autocorrelation function and Partial autocorrelation function (Englama, Duke, Ogunleye and Ismail (2010). Once the ARIMA model parameters (p,d.q) have been identified, they can be used to estimate the ARIMA model that best fit the data using maximum likelihood estimation procedure. The next step is the evaluation (diagnosis) of the model fitted to the observed series.

3.2 Probability Models (Time Domain)

i. Autoregressive process

A stochastic process $X_t, t \in z$ is said to be anautoregressive process of order p,denoted by AR(p),if it satisfies the difference equation.

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + e_t \qquad (3.1)$$

where $\phi_1, \phi_2, \dots, \phi_p$ are constants $e_t, t \in z$is a white noise with zero

mean and variance $\sigma^2 < \infty$. Equation (3.1) may be written as

$$\phi(B)X_t = e_t \qquad (3.2)$$

where,

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - -\phi_p B^p \qquad (3.3)$$

is called the characteristic equation. For stationarity, we require that all

roots of $\phi(B) = 0$ must lie outside the unit circle. An $AR(p)$ process is always invertible but is stationary if $\phi(B) = 0$ has zero outside the unit circle.

The autocovariance function for a stationary autoregressive process of Order p, is given by (Box-Jenkins, 1976) as

$$R(k) = \begin{cases} \phi_1 R(k-1) + \phi_2 R(k-2) + \cdots + \phi_p R(k-p) + \sigma^2, k = 0 \\ \phi_1 R(k-1) + \phi_2 R(k-2) + \cdots + \phi_p R(k-p), k \geq 1 \end{cases} \qquad (3.4)$$

The autocorrelation function of the process is given by

$$\rho(k) = \begin{cases} 1, & k = 0 \\ \phi_1 \rho(k-1) + \phi_2 \rho(k-2) + \phi_p \rho(k-p), & k \geq 1 \end{cases} \qquad (3.5)$$

The second set of equation in (3.5) is called Yule Walker equation for AR(p) process. For an autoregressive process of order p, the partial autocorrelation function ϕ_{kk} will be non zero for k$\leq p$ and zero for k$> p$. In other words, the partial autocorrelation function of pth order autoregressive process has a cut-off after lag p.

ii. Moving average process

A stochastic process $X_t,\ t \in z$ is said to be a moving average process of

order q, denoted by $MA(q)$,if it satisfies the difference equation.

$$X_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \theta_3 e_{t-3} \dots + \theta_q e_{t-q} \qquad (3.6)$$

This can be written as

$$X_t = \sum_{j=1}^{q} \theta_j e_{t-j} + e_t \qquad (3.7)$$

In terms of the backshift operator B, the moving average process of order q is written as

$$X_t = \theta(B) e_t \qquad (3.8)$$

where,

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \theta_3 B^3 + \cdots + \theta_q B^q \quad (3.9)$$

$\theta_1, \theta_2, \theta_3, \dots, \theta_q$ are constants and $e_t. t \in z$ is a purely random process with mean zero and variance $\sigma^2 < \infty$. The finite order moving average process $MA(q)$ is always stationary but invertible if all the roots of $\theta(B) = 0$ lies outside the unit circle.

iii. Mixed autoregressive moving average process

A stochastic process X_t , $t \in z$ is said to be an autoregressive moving average process of order (p, q), denoted by $ARMA(p, q)$ if it satisfies an equation of the type

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{j=1}^{q} \theta_j e_{t-j} e_t + e_t \quad (3.10)$$

For every $t \in z$, where $\phi_1, \phi_2, \phi_3, \dots, \phi_p, \theta_1, \theta_2, \theta_3, \dots, \theta_q$ are constants and $e_t, t \in z$ is a purely random process with mean zero, variance $\sigma^2 < \infty$. Equation (3.11) can be written as

$$\emptyset(B) X_t = \theta(B) e_t \quad (3.11)$$

The random process X_t , $t \in z$ satisfying equation (3.11) is stationary if all the roots of $\phi(B) = 0$ is outside the unit circle and invertible if all the roots of $\theta(B) = 0$ is outside the unit circle. The autocovariance function of the process is like that of an autoregressive process after lag q.That is

$$R(k) = \phi_1 R(k-1) + \phi_2 R(k-2) + \cdots + \phi_p R(k-p), k \geq q + 1 \quad (3.12)$$

Hence, the autocorrelation function is

$$\rho(k) = \phi_1 \rho(k-1) + \phi_2 \rho(k-2) + \cdots + \phi_p \rho(k-p), k \geq q + 1 \quad (3.13)$$

or $\phi(B)\rho_k = 0$ (3.14)

iv. Autoregressive integrated moving average process

ARIMA process is specifically designed for modeling non stationary time Series. The general form of ARIMA (p,d,q) process can be expressed as:

$$\left(1 - \phi_1 B - \cdots - \phi_p B^p\right)(1 - B)^d = \left(1 - \theta_1 B - \theta_2 B^2 - \right.$$
$$\left. -\theta_q B^q\right) e_t \quad (3.15)$$

or

$$\phi(B)(1 - B)^d X_t = \theta(B) e_t \quad (3.16)$$

Note:

$a.$ $\phi(B)$ is the autoregressive operator, which is assumed to bestationary if $\phi(B) = 0$ lies outside the unit circle.

$b.$ $\theta(B)$is the moving average operator. It is assumed to be invertible if $\theta(B) = 0$ lies outside the unit circle.

3.3 Model Identification

This entails the use of data and any available information to suggest a subclass of parsimonious models that best describes the data.

The objective of the model identification step is to determine the values of p, d and $q$in the ARIMA(p, d, q) model. When the series exhibits a trend, we may either fit and remove a deterministic trend or difference the series.

Box and Jenkins methodology uses differencing to make a time series stationary when necessary. The first step is to access the autocorrelations and partial autocorrelations plots. Differencing usually reduces the number of large autocorrelation considerably. If the differenced series still does not appear stationary, differencing is done again.

It is often useful to determine the magnitude of a large autocorrelation and partial autocorrelation coefficient. An autocorrelation must be at least$\frac{2}{\sqrt{n}}$ in absolute value to be statistically significant. Note that even though an autocorrelation is statistically significant, it may not be large enough to worry about. By considering the pattern of the autocorrelations and the partial autocorrelations, one can guess a reasonable and parsimonious model for the data. The values of p (autoregressive order),d

(differencing) , and q(moving average order) are determined during the model identification process. These values can be estimated via detailed study of the two autocorrelation plots. The value of p is determined from the partial autocorrelations of the appropriately differenced series. If the partial autocorrelations cut off after a few lags, the last lag with a large value would be the estimated value of p. If the partial autocorrelations do not cut off, you either have a moving average model ($p = 0$) or an ARIMA model with positive $p$and q.

The value of q is found from the autocorrelations of the appropriately differenced series. If the autocorrelation function cut off after a few lags, the last lag with a large value would be the estimated value of q. If the autocorrelations do not cut off, you either have an autoregressive model ($q = 0$) or an ARIMA model with a positive p and q.

3.4 Model Estimation

This is the efficient use of data to make inference about the parameters. It is conditioned on the adequacy of the selected model. It can also be defined as using computation algorithms to arrive at coefficients that best fit the selected ARIMA model.

4. Results and Discussion

This chapter explains explicitly the steps that were adopted in forecasting of Nigerian Monthly Crude Oil Production spanning from January 2005 to September, 2022. We started by ascertaining the time series plot of the original data of the Nigerian Monthly Crude Oil Production.

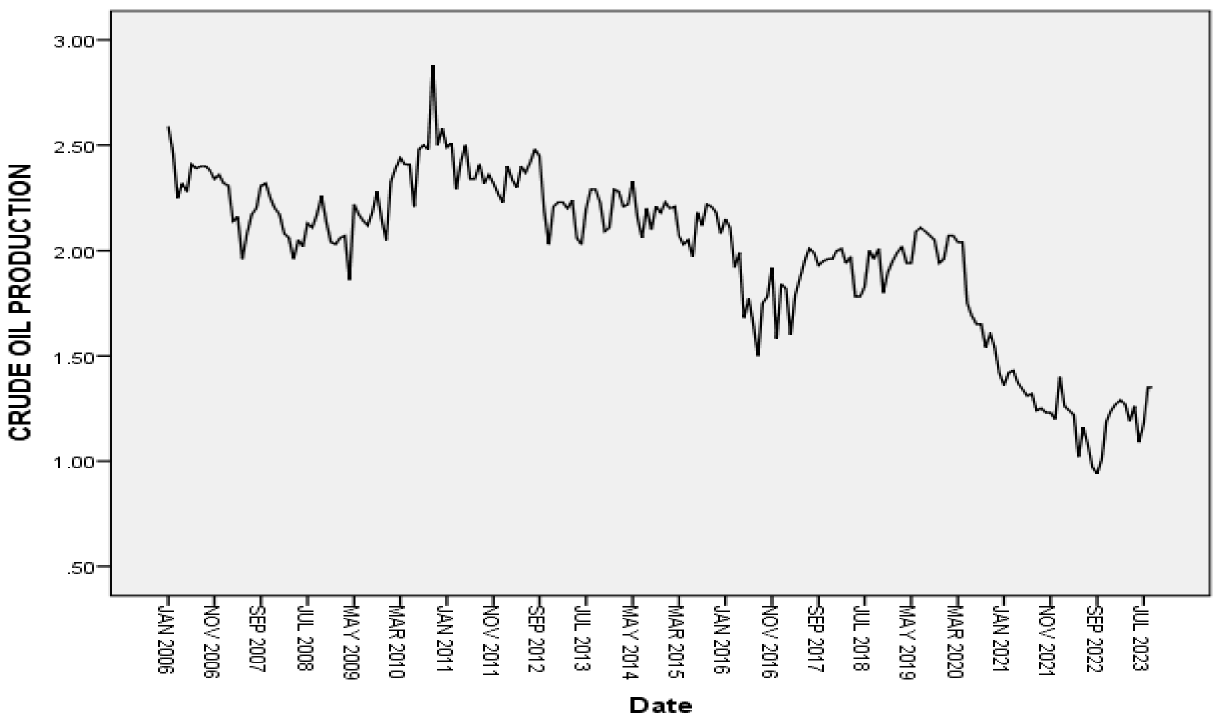

Figure 4.1: Time Series Plot for Nigerian Monthly Crude Oil Production data

4.1 Correlogram: The ACF and PACF of the Original Data

Having observed from the time plot that the mean of the series is changing with time, next we examine the autocorrelation functions (ACF) and partial autocorrelation function (PACF) to see if there exist correlations in data points of the series.

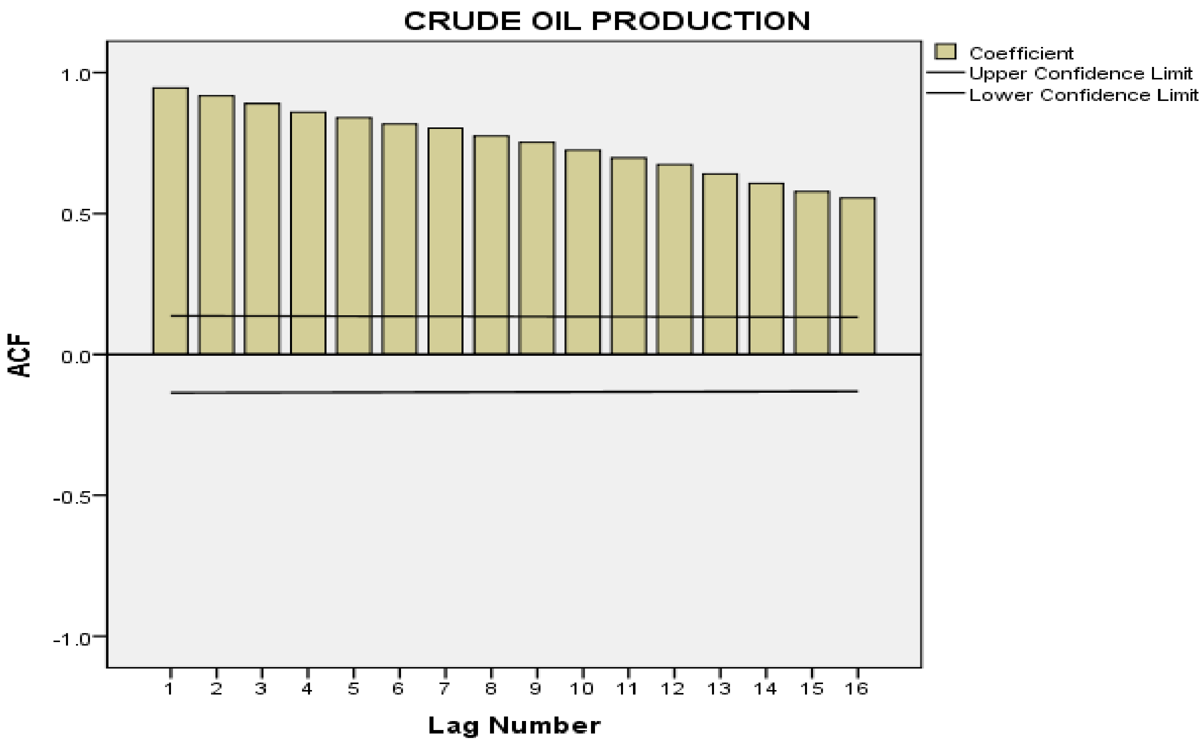

Figure 4.2: Plot of ACF of the raw data

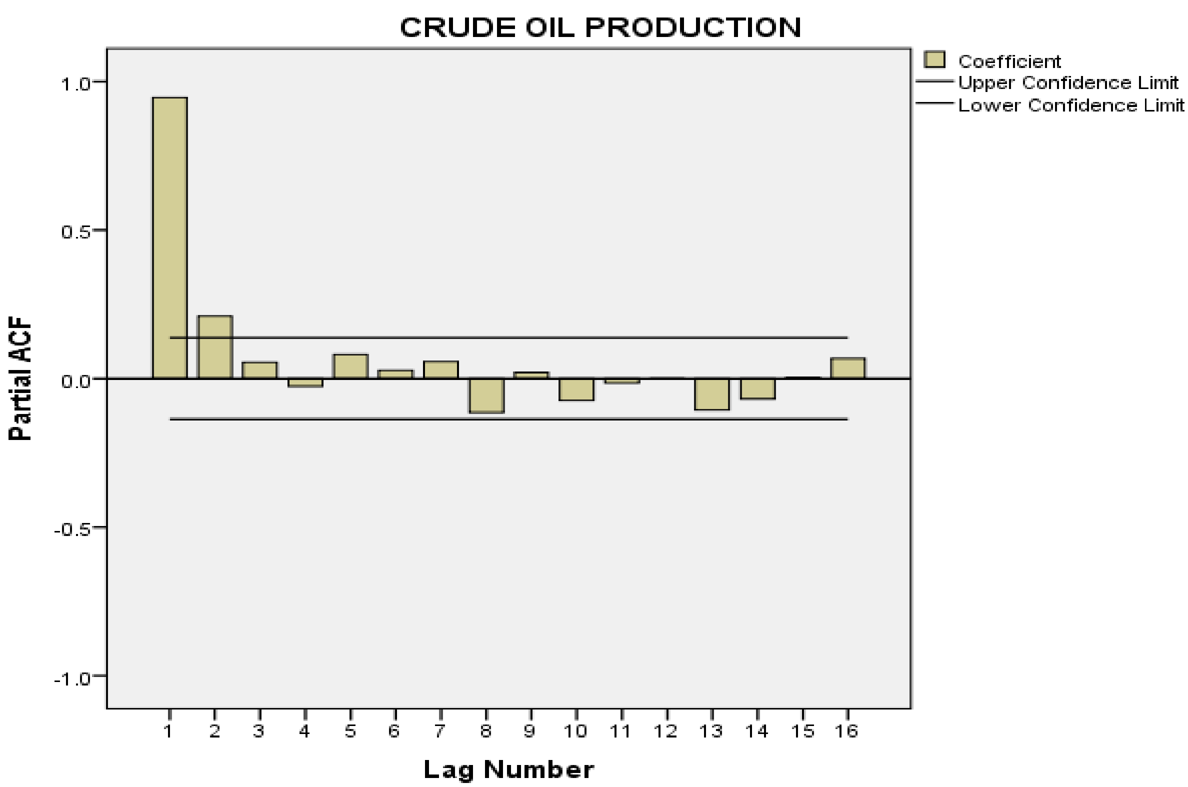

Figure 4.3: Plot of PACF of the raw data

Series is clearly non-stationary from the pattern of the ACF and PACF. We observed that from Figure 4.3 plots that, the series decay exponentially, that is autocorrelation function (ACF) start high and declined slowly, and partial autocorrelation (PACF) dies to zero after the first lag, which also confirms the presence of low order serial dependency in the series. This shows that the series is non-stationary and should be differenced. Next, we perform the unit root test to check the stationary of the series.

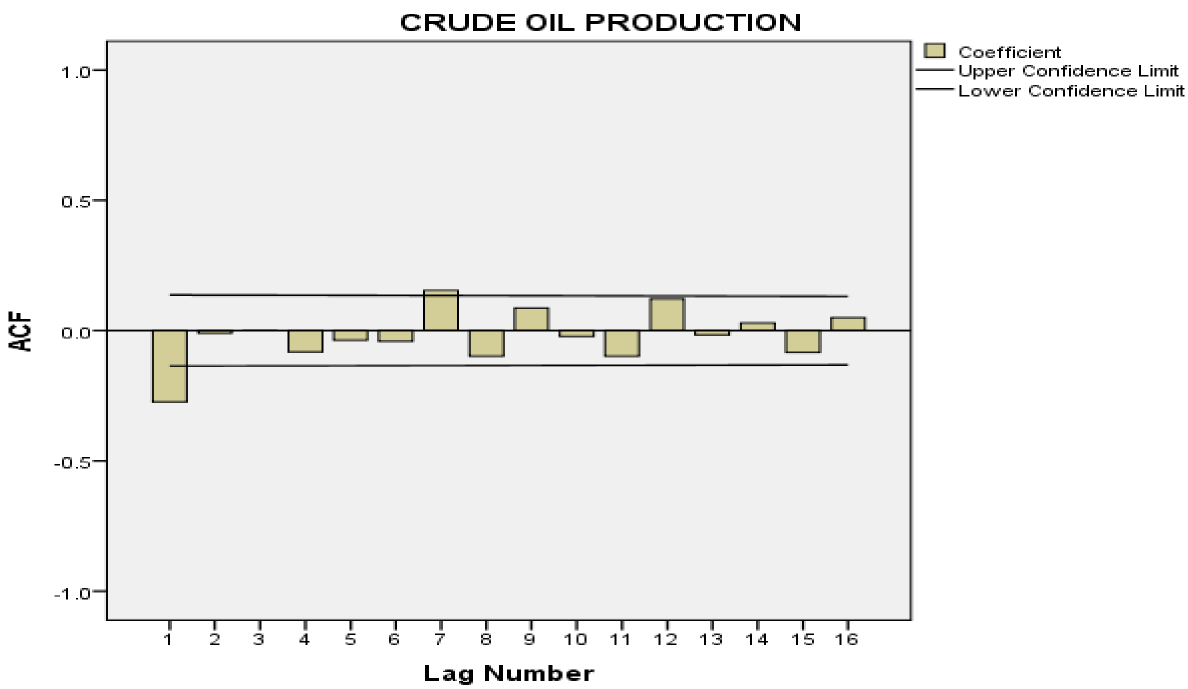

Figure 4.4: ACF Plot of the first differenced series

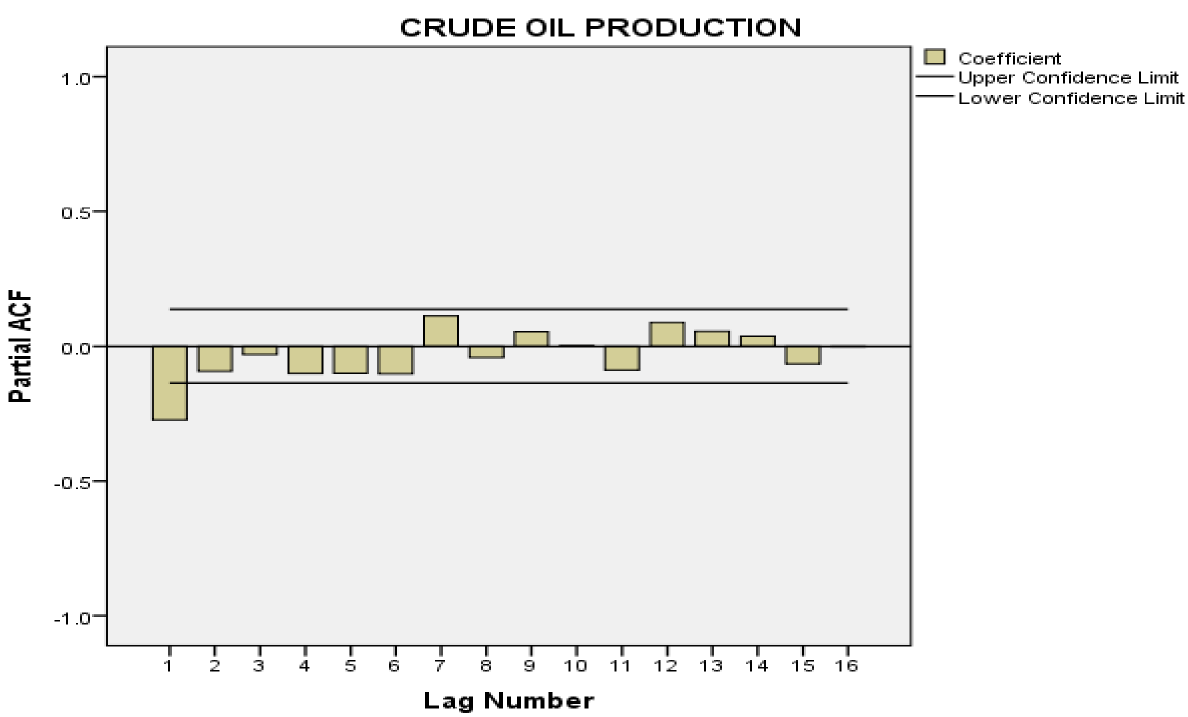

Figure 4.5: PACF Plot of the first differenced series

Figure 4.4 and 4.5 above show no evidence of serial correlation in the residuals and hence the model is adequate.

4.2 Choice of the Best Model

Once stationary have been addressed, the next step is to identify the order (the p, d, and q) of the autoregressive and moving average terms. The primary tools for doing this are: Akaike Stationary R^2, RMSE, MAE, MAPE and Normalized BIC. That is the model that gives minimum values.

i. Arima Model Identification

Table 4.1: Result of ARIMA model identification

MODEL	**ARIMA (1,1,1)**	ARIMA (0,1,1)
STATIONARY R^2	**0.092***	0.086
R^2	**0.922***	0.922
RMSE	**0.110***	0.110
MAE	**0.082***	0.082
MAPE	**4.384***	4.392
NORMALIZED BIC	**-4.336***	-4.360

From table 4.4 we observed that the optimal model is ARIMA (1, 1, 1) that is based on the selection criterion Stationary R^2, R^2, RMSE, MAE, MAPE and Normalized BIC. The asterisks above indicate the best model or optimal model (that is minimized) value of the respective information criteria, Stationary R^2, RMSE, MAE, MAPE and Normalized BIC model or optimal model.

ii. ACF and PACF of the residuals

Next we plot the ACF and PACF of the standardized residuals to visually see if there exist autocorrelation. The graphs are plotted using 20 standardized residuals and in as much as there is no point outside the confidence limit ($\propto$ = 0.05) the residual are uncorrelated.

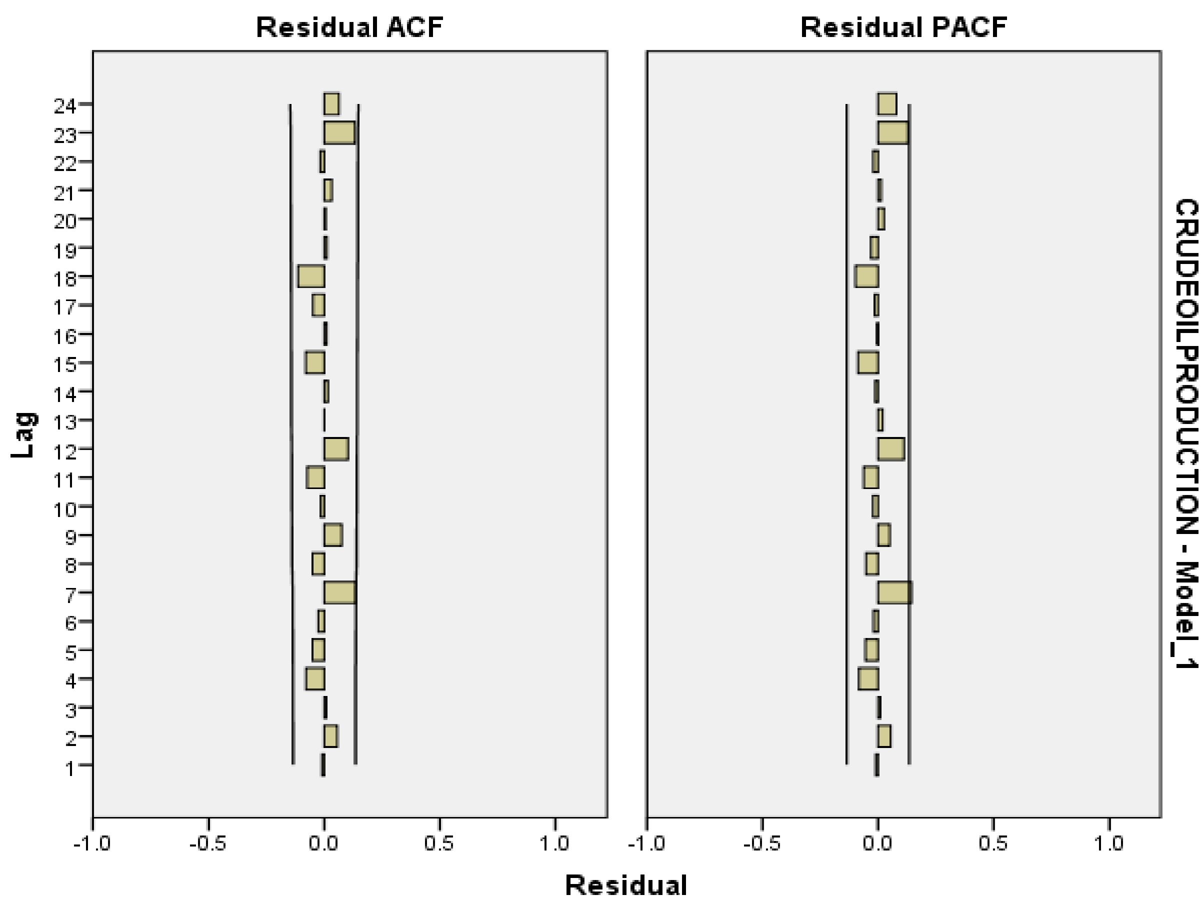

Figure 4.6: ACF and PACF of the residuals.

Figure above show no evidence of serial correlation in the residuals and hence the model is adequate.

4.3 ARIMA and ARMA Model Forecast Performance

We compute forecast error statistics in order to assess the performance of the forecast. For us to know how close the forecast graph is to the actual graph we need to plot the forecast against the actual. The results of the computations are presented in the figure below:

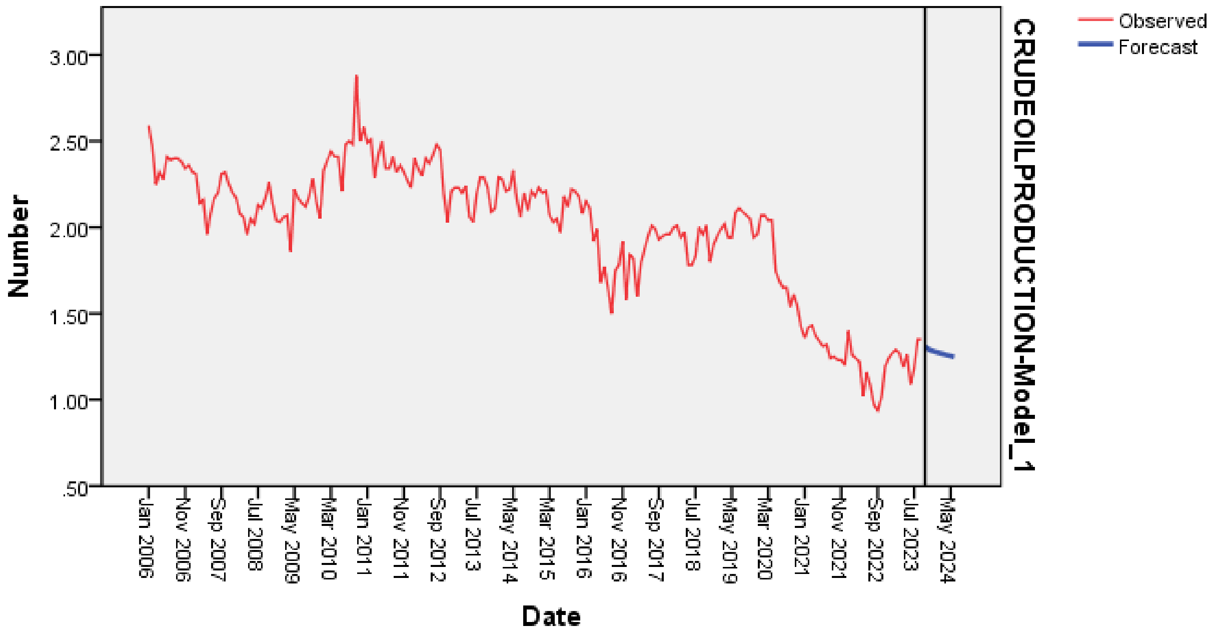

Table 4.2: The table above shows the out-sample actual versus predicted values from January 2023 to June, 2024. From the table, there is no much deviation from the actual value.

		95% Limits		
Period	**Forecast**	**Lower**	**Upper**	**Actual**
1/2023	1.23265	0.76985	1.69545	1.27000
2/2023	1.22574	0.74767	1.70381	1.29000
3/2023	1.21883	0.72595	1.71170	1.27000
4/2023	1.21191	0.70467	1.71915	1.19000
5/2023	1.20500	0.68378	1.72621	1.26000
6/2023	1.19809	0.66326	1.73291	1.09000
7/2023	1.19117	0.64308	1.73926	1.18000
8/2023	1.18426	0.62321	1.74531	1.35000
9/2023	1.17734	0.60363	1.75106	1.35000
10/2023	1.31	0.09	1.53	
11/2023	1.29	0.56528	1.76175	
12/2023	1.28	0.54648	1.76673	
1/2024	1.28	0.96	1.60	
2/2024	1.27	0.93	1.61	
3/2024	1.26	0.93	1.61	
4/2024	1.26	0.88	1.64	
5/2024	1.25	0.85	1.66	
6/2024	1.25	0.83	1.67	

5. Conclusion

In this paper, ARIMA model was used to estimate the data that best describes the Nigerian Monthly Crude Oil Production. The data set is Nigerian Monthly Crude Oil Production for the period of January, 2005 – September, 2022. Result analysis revealed that the series became stationary at first difference. Further analysis showed that among all the class of ARIMA model based on Stationary R^2, R^2, RMSE, MAE, MAPE and Normalized BIC, the best or optimal model was ARIMA (1, 1, 1). The performance of the model for out-of-sample shows that ARIMA (1, 1, 1) has the minimum ME, MSE, MAE, RMSE which indicate that ARIMA (1, 1, 1) model is the best or optimal model for the period forecasted.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
ORCID iD: 0000-0002-2998-351X
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Abass I. T., Kazeem A. A. and Adedayo, F. A. (2021). Time Series Analysis of Nigerian Monthly Crude Oil Price. *FUPRE Journal of Scientific and Industrial Research. 5*(1).
[2] Abdulkareem, A and Abdulkareem, K A. (2016). Analysing Oil Price- Macroeconomic Volatility in Nigeria, *CBN Journal of Applied Statistics (JAS)*, Article 1. Available at: https://dc.cbn.gov.ng/jas/vol7/iss1/1
[3] Adedokun, (2018). The effects of oil shocks on government expenditures and government revenues nexus in Nigeria (with exogeneity restrictions), *Future Business Journal, 4*(2), 219-232, ISSN 2314-7210, https://doi.org/10.1016/j.fbj.2018.06.003.
[4] Akinlo, T., and Apanisile, O. T. (2015). The impact of volatility of oil price on the economic growth in sub-Saharan Africa. *British Journal of Economics, Management and Trade*, 5, 338–349. doi:10.9734/BJEMT/2015/12921
[5] Akinyemi, A., Oluwafemi E. and Ikenna, N. (2017). Impact of Oil Revenue and Exchange Rate Fluctuation on Economic Growth in Nigeria. 10.5281/Zenodo.10201872.
[6] Box, G. E. P. and Jenkins G.M. (1976).Time Series Analysis: Forecasting and Control, rev. ed. Oakland, CA: Holden Day.
[7] Dapel Z. (2018). Poverty in Nigeria: Understanding and Bridging the Divide between North and South. Center for Global Development.
[8] Englama, A, Duke, O. O., Ogunleye T. S. and Ismail F. U., (2010). Oil Prices and Exchange Rate Volatility in Nigeria: An Empirical Investigation. *Central Bank of Central Bank of Nigeria Economic and Financial Review, 48*(3), 31-48.
[9] Gylych, J., Jbrin, A. A., Celik, B., and Isik, A. (2020). The Effect of Oil Price Fluctuation on the Economy of Nigeria. *International Journal of Energy Economics and Policy*, *10*(5), 461–468. Retrieved from https://www.econjournals.com/index.php/ijeep/article/view/9493
[10] Odupitan, A. (2017). Understanding the distribution channels of fast moving consumable goods (FMCG) of food industry in Nigeria.
[11] OECD (2020). Policy Response to COVID-19, 2020. The impact of coronavirus (COVID19) and the global oil price shock on the fiscal position of oil-exporting developing countries. Accessed 15 May 2021
[12] Okoi, O. (2019). The paradox of Nigeria's oil dependency. African Portal. Available at https://www.africaportal.org/features/p aradox-nigerias-oil-dependency/
[13] Otache I. (2020). The Effects of the Covid-19 Pandemic on the Nigeria's Economy and Possible Coping Strategies, *Asian Journal of Social Sciences and Management Studies, 7*(3), 173-179.
[14] Wiri L. and Sibeate P. (2021). Linear and non-linear modeling of Nigerian Crude oil prices.
[15] Wiri L. and Tuaneh, G. B. (2019). Modelling Nigeria Crude Oil Prices Using ARIMA, Pre-Intervention and Post- Intervention model. *Asian Journal of Probability and Statistics. 3*(1).