International Journal of Scientific Research in Mathematical and Statistical Sciences
Vol.10, Issue.2, pp.38-45, April 2023
E-ISSN: 2348-4519
Available online at: www.isroset.org

Research Paper

Time Series Analysis of Crude Oil Production in Nigeria between the Years 2010 to 2020

N. F. Adeleye[1*], H.O. Ilo[2], T.O. Gbadamosi[3]

[1,2,3]Department of Statistics/School of Pure and Applied Science, Ogun State Institute of Technology Igbesa, Ogun State Nigeria

**Corresponding Author: leyinton@gmail.com*

Received: 26/Feb/2023; **Accepted**: 27/Mar/2023; **Published**: 30/Apr/2023

Abstract— In Nigeria and other regions worldwide, crude oil is a crucial source of income and energy that necessitates careful attention and technical expertise in its production, refining, and global distribution. Throughout history, crude oil has been the cornerstone of Nigeria's economy, playing a significant role in shaping its political, industrial, and economic future. According to reports, Nigeria's petroleum sector is the nation's biggest industry. This is likely a result of the widespread perception that petroleum is one of the main sources of energy all over the world. In addition to this, petroleum products make a wide range of additional contributions to national development, including the creation of jobs, revenue generation, earnings in foreign currencies, industrialization, and advancements in other economic indicators. The Nigerian National Petroleum Company Limited (NNPC Limited), formerly known as the Nigerian National Petroleum Corporation (NNPC), emerged in April 1977 out of merging the commercial and regulatory responsibilities of the Ministry of Petroleum Resources and the Nigerian National Oil Corporation (NNOC). Nigeria is Africa's top producer of oil and gas. Using the Mann-Whitney test in our analysis section, it was discovered that there is a significant difference in crude oil production between the years 2010 and 2020. After the discovery that the time series data set was non-stationary using the Augmented Dickey-Fuller test, the Detrending method was applied to remove the trend in the dataset and transform it into a stationary dataset that would be suitable for prediction and forecasting. The Augmented Dickey-Fuller test was later used to test if the dataset is truly stationary. The Seasonal Autoregressive Integrated Moving Average (SARIMA) was used to test the model and was also used for the prediction and forecasting of the Crude Oil dataset. The result of the analysis indicates that SARIMA is a good statistical tool that can be used in the prediction of this kind of data (Crude Oil Data).

Keywords— Augmented Dickey-Fuller Test, Mann Whitney Test, Detrending, Nigerian National Petroleum Company Limited (NNPCL), Seasonal Autoregressive Integrated Moving Average (SARIMA), Crude Oil.

1. Introduction

In Nigeria, as well as many other parts of the world, crude oil is a crucial source of energy and revenue. Despite the fact that Nigeria's oil sector was established at the turn of the century, it was not until the end of the civil war, which lasted from 1967 to 1970, that crude oil began to play a prominent role in the country's economy. Nigeria is primarily an agrarian country that heavily relies on exporting primary goods, particularly crude oil products. Since gaining independence in 1960, the country has experienced various conflicts based on ethnicity, region, and religion, which have been further amplified by the significant differences in economic, educational, and environmental progress between the North and South. These might be partially ascribed to the country's significant oil finds, which have an impact on and are influenced by economic and social factors. The main industry in Nigeria is said to be the petroleum sector. This is most likely a result of the widespread perception that petroleum is one of the main sources of energy on the globe. Since petroleum today meets a wide range of energy and related demands, it has been highlighted that the idea that petroleum is versatile is where the scale, worldwide nature, and the role played by the petroleum industry arose. With more than 50% of the world's total commercial energy consumption coming from petroleum, it is the world's most important source of energy. Nigerian crude oil income is a definite asset for spending obligations on a variety of initiatives at the municipal, state, and federal levels. The economic growth and development of Nigeria depend majorly on the production and consumption of petroleum products; hence, the connection between the State and Crude Oil in Nigeria is immeasurable [1], [2]. Oil accounts for about 40 percent of Nigeria's GDP, 70 percent of federal government revenue, and 92 percent of its foreign exchange earnings [3]. Additionally, the daily domestic demand for petroleum products is 530,000 barrels per day (bpd), which is 85,000 bpd greater than the installed refining capacity of 445,000 bpd, which was never produced. Therefore, the availability of petroleum products continues to be a litmus test for Nigeria's succeeding governments. The availability of petroleum products has increased since democracy was established on May 29th, 1999, but not without a cost: periodic price rises for petroleum goods, even while demand for these items is still greater than supply. The discovery of crude oil has had both beneficial and negative effects on Nigeria's economy. Regarding the surrounding communities where the oil wells are exploited, this might be seen negatively. Environmental deterioration still affects some of these communities, which has negative effects on other economic and social variables as well as the ability to support one's living. The relative effects of crude oil on the economy must be assessed even though domestic sales and exports of petroleum products generate significant profits, their influence on the development of the Nigerian economy in terms of returns and productivity is still debatable. Many years ago, crude oil exportation created a robust wealth for Nigeria. Then, the Naira appreciated as foreign exchange influxes offset outflows and Nigeria's foreign reserve assets increased [5], [6]. The global emergence of the crude oil crisis in 2013 affected the Nigerian government's revenue negatively since crude oil exportation is the Nigerian government's major source of revenue [7], [8]. Because Nigeria's economy is solely dependent on the export of crude oil, productivity fell in all other sectors as a result; this caused large-scale urban migration and severe poverty in the rural areas. As a result, the Nigerian labour market has seen very high unemployment, low wages, and pitiful working conditions [9], [10]. From 1970 to 2020, Nigeria's poverty rate increased from 36% to 70%, and unemployment was at 30.7% [11]. Based on this, it is believed that oil revenue did not appear to add value to the Nigerian populace's standard of living or create job opportunities, but rather created social and economic challenges [12], [13]. As if the situation wasn't bad enough, the global economy has been further disrupted by the COVID-19 pandemic, which has caught the entire world off guard since December 2019 [14]. This crashed the world oil price and put the Nigerian government in unchartered waters and the worst recession in 40 years. The government continues to struggle to revive the economy amidst dwindling oil revenues compounded by unemployment, poverty, insurgency, and mismanagement [15], [16]. The oil industry urgently requires a suitable and acceptable production and export policy given the significance of the oil sector to the Nigerian economy. Possible causes of the production and distribution issues include a lack of refineries and subpar management practices. Another issue can be the worry that a specific year's supply will be of lower quality than subsequent years. Despite the significant economic impact of crude oil in Nigeria, the revenue generated has not been effectively managed. With the existence of other economic sectors in the country, the surplus income from the oil industry can be channeled towards diversifying the economy and increasing the overall GDP. The objective of this study, therefore, is to analyze the trend and pattern of crude oil production in Nigeria from 2010 to 2020 and develop a model for predicting the quantity of crude oil production in the country. This will provide insights into the future of crude oil production in Nigeria and aid in proper planning for the growth and development of the nation and its citizens. This research aims to contribute to the existing body of knowledge on time series analysis and Nigerian crude oil production research.

2. Materials and Methods

Seasonal Autoregressive Integrated Moving Average

Seasonality in a time series is a regular pattern of changes that repeats over time say **S** periods, where **S** defines the number of periods until the pattern repeats. For example, there is seasonality in monthly data for which high values tends to always occur in some particular months and low values tend to occur in other particular months. For this example, if **S = 12** (months per year) is the span of the periodic seasonal behaviour, for quarterly data **S = 4** time periods per year. In a seasonal ARIMA model, seasonal AR and MA terms predict x_t using data values and errors at times with lags that are multiples of **S** (the span of the seasonality). With monthly data (and S = 12), a seasonal first-order autoregressive model would use x_{t-12} to predict x_t. For instance, if we were selling cooling fans we might predict this August's sales using last August's sales. (This relationship of predicting using last year's data would hold for any month of the year.) A seasonal second-order autoregressive model would use x_{t-12} x_{t-24} and to predict x_t. Here we would predict this August's values from the past two Augusts. A seasonal first-order $MA(1)$ model (with S = 12) would use ω_{t-12} as a predictor. A seasonal second-order $MA(2)$ model would use ω_{t-12} and ω_{t-24}.

Differencing

Almost by definition, it may be necessary to examine differenced data when we have seasonality. Seasonality usually causes the series to be non-stationary because the average values at some particular times within the seasonal span (months, for example) may be different from the average values at other times.

Seasonal Differencing

Seasonal differencing is defined as a difference between a value and a value with lag that is a multiple of **S**. With **S = 12** which may occur with monthly data, a seasonal difference is $(1-B^{12})x_t = x_t - x_{t-12}$. The differences (from the previous year) may be about the same for each month of the year giving us a stationary series. With **S = 4**, which may occur with quarterly data, a seasonal difference is $(1-B^4)x_t = x_t - x_{t-4}$. Seasonal differencing removes seasonal trends and can also get rid of a seasonal random walk type of non-stationarity.

Non-seasonal Differencing

If a trend is present in the data, we may also need non-seasonal differencing. Often (not always) a first difference (non-seasonal) will "detrend" the data. That is, we use $(1-B)x_t = x_t - x_{t-1}$ in the presence of a trend.

Differencing for Trend and Seasonality

When both trend and seasonality are present, we may need to apply both a non-seasonal first difference and a seasonal difference. That is, we may need to examine the ACF and PACF of $(1-B^{12})(1-B)x_t = (x_t - x_{t-1}) - (x_{t-12} - x_{t-13})$. Removing the trend doesn't mean that we have removed the dependency. We may have removed the mean (μ_t) part of which may include a periodic component. In some ways, we are breaking the dependency down into recent things that have happened and long-range things that have happened.

Seasonal ARIMA Model

The seasonal ARIMA model incorporates both non-seasonal and seasonal factors in a multiplicative model. One shorthand notation for the model is the $ARIMA(p,d,q)(P,D,Q)S$. With *p* = non-seasonal AR order, *d* = non-seasonal differencing, *q* = non-seasonal MA order, *P* = seasonal AR order, *D* = seasonal differencing, *Q* = seasonal MA order, and *S* = time span of repeating seasonal pattern. Without differencing operations, the model could be written more formally as

$$\Phi(B^S)\phi(B)(x_t - \mu) = \Theta(B^S)\theta(B)\omega_t \tag{1}$$

The non-seasonal components for AR and MA are given respectively:

$$\phi(B) = 1 - \phi_1 B - ... - \phi_p B^p \tag{2}$$

$$\theta(B) = 1 + \theta_1 B - ... - \theta_q B^q \tag{3}$$

The seasonal components for AR and MA are given respectively:

$$\Phi(B^S) = 1 - \Phi_1 B^S - ... - \Phi_P B^{PS} \tag{4}$$

$$\Theta(B^S) = 1 + \Theta_1 B^S + ... + \Theta_Q B^{QS} \tag{5}$$

Note that on the left side of equation (1) the seasonal and non-seasonal AR components multiply each other and on the right side of equation (1) the seasonal and non-seasonal MA components multiply each other. For example, **ARIMA (0,0,0)x(2,1,1)$_{12}$** indicates that no non-seasonal AR(0) term, a seasonal AR(2) term, which indicates $(1-\Theta_1 B^{12} - \Theta_2 B^{12})$, no non-seasonal differencing($d=0$), seasonal differencing $(D=1)$ that is $(1-B^{12})$, no seasonal moving average i.e. MA(0) terms, seasonal moving average MA(1) that means $(1+\Theta_1 B^{12})$ and seasonal period is $S=12$ *S* = 12. Another example is **ARIMA (0, 0, 1)x(0, 0, 1)$_{12}$**, this model includes a non-seasonal MA(1) term, a seasonal MA(1) term, no differencing, no AR terms and the seasonal period is S = 12. The non-seasonal MA(1) polynomial is $\theta(B) = (1+\theta_1 B)$, The seasonal MA(1) polynomial is $\Theta(B^{12}) = 1 + \Theta_1 B^{12}$, the model also becomes

$$(x_t - \mu) = \Theta(B^{12})\theta(B)\omega_t = (1+\Theta_1 B^{12})(1+\theta_1 B)\omega_t \tag{6}$$

If we multiply the two polynomials on the right side, we get

$$(x_t - \mu) = (1 + \theta_1 B + \Theta_1 B^{12} + \theta_1 \Theta_1 B^{13})$$

$$= \omega_t + \theta_1 \omega_{t-1} + \Theta_1 \omega_{t-12} + \theta_1 \Theta_1 \omega_{t-13} \tag{7}$$

Thus the model has MA terms at lags 1, 12, and 13. This leads many to think that the identifying ACF for the model will have non-zero autocorrelations only at lags 1, 12, and 13. There's a slight surprise here, there will also be a non-zero autocorrelation at lag 11. Another example is **ARIMA (1,0,0)×(1,0,0)$_{12}$**, this model includes a non-seasonal AR(1) term, a seasonal AR(1) term, no differencing, no MA terms and the seasonal period is **S = 12**. The non-seasonal AR(1) polynomial is $\phi(B) = 1 + \phi_1 B$, the seasonal AR(1) polynomial is $\Phi(B^{12}) = 1 - \Phi_1 B^{12}$, the model becomes $\omega_t = (1-\phi_1 B)(1-\Phi_1 B^{12})(x_t - \mu)$. If we let $z_t = x_t - \mu$ and multiply the two AR components and also push all but z_t to the right side we get:

$$z_t = \phi_1 z_{t-1} + \Phi_1 z_{t-12} + (-\phi_1 \Phi_1) z_{t-13} + \omega_t \tag{8}$$

This is an AR model with predictors at lags 1, 12, and 13.

Identifying a Seasonal Model

The first step to take is to do a time series plot of the data and examine it for features such as trend and seasonality. If it shows that it is a seasonal data, so look at the pattern across those time units to see if there is indeed a seasonal pattern. The next line of action is to carry out all necessary differencing. If there is seasonality and no trend, then take a difference of lag **S**. For instance, take a 12th difference for monthly data with seasonality. Seasonality will appear in the ACF by tapering slowly at multiples of **S**. If there is a linear trend and no obvious seasonality, then take a first difference. If there is a curved trend, consider a transformation of the data before differencing. If there is both trend and seasonality, apply a seasonal difference to the data and then re-evaluate the trend. If a trend remains, then take the first differences. If there is neither an obvious trend nor seasonality, don't take any differences. For non-seasonal terms, examine the early lags (1, 2, 3,…) to judge non-seasonal terms. Spikes in the ACF at low lags with a tapering PACF indicate non-seasonal MA terms. Spikes in the PACF at low lags with a tapering ACF indicate a possible non-seasonal AR term. For the Seasonal terms, examine the patterns across lags that are multiples of **S**. For example for monthly data, look at lags 12, 24, 36, and so on. Judge the ACF and PACF at the seasonal lags in the same way as the non-seasonal lags, after this estimate the model(s) might be reasonable also don't forget to include any differencing that you did before looking at the ACF and PACF, and lastly examine the residuals to see if the model seems good.

3. Unit Root Test

The Augmented Dickey-Fuller Test and Mann-Whitney U test will be used to analyze the data for this research work. A statistical significance test called the Augmented Dickey-Fuller test provides a p-value that we must use to determine if the time series data is stationary or non-stationary. A simple AR model can be represented as:

$$y_t = \rho y_{t-1} + \mu_t \quad (9)$$

Where y_t is the variable of interest at the time t; ρ is the coefficient that defines the unit root, and μ_t is the error term. If $\rho = 1$ the unit root is present in a non-stationary time series. If a regression model can be represented as

$$\Delta y_t = (\rho - 1) y_{t-1} + \mu_t = \delta y_{t-1} + \mu_t \quad (10)$$

Where Δ is the difference operator and $\delta = \rho - 1$

If is the case, then the differencing will act as the error term, and if the coefficient has any values lower or higher than one, the changes will be detected following previous observations. This may be formally described as follows when ADF is applied to the model:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \ldots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t \quad (11)$$

Where α is the constant' β is the coefficient of time t and p is the lag order of the autoregressive process. Here in the mathematical representation of ADF, we have added the differencing terms that make changes between ADF and the Dickey-Fuller test. The unit root test will then be carried out under the null hypothesis μ = 0 against the alternative hypothesis of μ < 0. Therefore the value for the test statistics becomes

$$DF_\tau = \frac{\hat{\gamma}}{SE(\hat{\gamma})} \quad (12)$$

This may be compared to the applicable critical value for the Dickey-Fuller test. For critical values, the test has a simple distribution known as the Dickey-Fuller table. The critical aspect to remember here is that because the null hypothesis assumes the presence of a unit root, the p-value produced by the test should be smaller than the significance level (say, 0.05) to reject the null hypothesis and conclude that the series is stationary. A non-parametric statistical test that may be used to compare two samples or groups is the Mann-Whitney U test. Because the test is based on ranking the observations in each group, one of the important requirements for the Mann-Whitney U Test is that the variable being compared between the two groups must be continuous. The datasets distributions are presumed to be non-normal. If the datasets are normally distributed, compare the two groups using the unpaired Student's t-test. The data in both groups are not considered to be normal, but rather to be of similar shape across the two groups. A paired samples t-test should be used instead of the former if samples are paired (for instance, two measurements from the same group of participants). A good test requires a sufficient sample size, which is often greater than 5 observations in each group. The Mann-Whitney U Test statistics are represented by the letter U, where

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \quad (13)$$

$$U_1 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2 \quad (14)$$

Where R_1 and R_2 are sum of the ranks for group 1 and group 2 respectively

4. Results and Discussion: Data Presentation and Analysis

The data that was used for this research work was collected from the Nigerian National Petroleum Company Limited (NNPC Limited) website, which covers the total number of barrels of crude oil produced per month from the years 2010 to 2020. The analysis for this research work was done using the Python programming language. Due to the complexity and largeness of the data, the set of data was standardized for this research to reduce both the complexity and the largeness of the data. The data collected shows that there are big differences in the number of barrels of crude oil that were produced between one year and the preceding year. For example, between the years 2019 and 2020, it shows that on average, 61,270,340 barrels of crude oil were produced in 2019, while 53,696,864 barrels of crude oil were produced in 2020. We can see that there is a difference in the production and also that there was a decline in the production of crude oil between the years 2019 and 2020, and statistics show that there was a 12.3% decrease in crude oil production between the years 2019 and 2020.

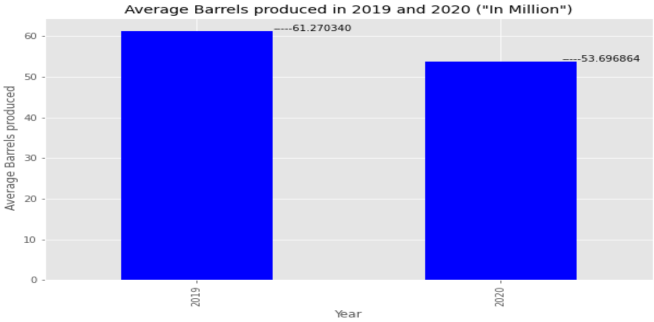

Fig. 1: Crude oil production between the years 2019 and 2020

Due to the non-functioning of Nigerian refineries, the effects on the number of barrels of crude oil produced per day cannot be underestimated, and this will be a great challenge for a country like Nigeria, with crude oil production being her major source of income. From the below chart, we can see that there is a big decline in the production of crude oil in Nigeria from January 2010 until December 2020. This shows that the non-functioning of the Nigerian refineries has affected the production of crude oil in Nigeria.

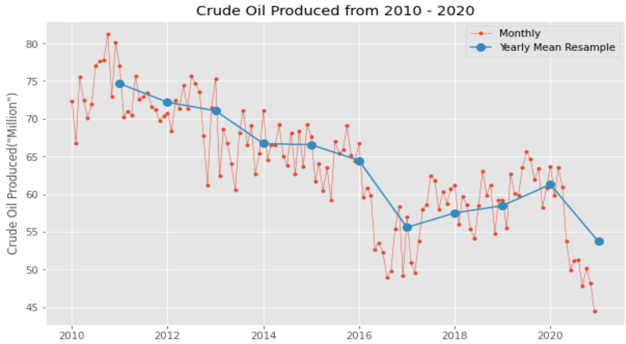

Fig. 2: Crude oil production between the years 2010 and 2020

The next question we need to look at now is how we can predict the future production of crude oil in Nigeria. This question can be answered using time series analysis, and we can use it to get the appropriate method to predict crude oil production in Nigeria. Firstly, we need to know if our data is stationary (i.e., the mean and variance do not vary over time) or not to decide the best approach to this problem. As we can see from the above plot in Figure 2, the crude oil data has a trend, which makes it non-stationary, but we cannot draw any conclusions in this way. We need to decompose the data to check for its behaviour over time.

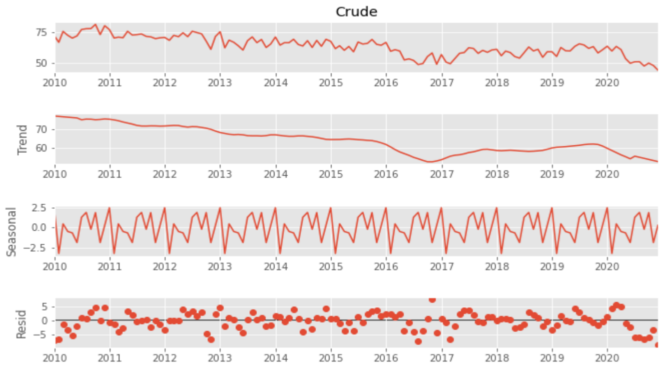

Fig. 3: Crude oil production decomposed data between the years 2010 and 2020

An additive model was used to decompose the time series data, and from the above plot, it can be seen that from the year 2010 to 2020, crude oil production has been going in a downward trend, and as well, the crude oil production is also seasonal, i.e., repetition of data at a certain time. And also, using the mean and variance to verify the stationary gives us the below plot:

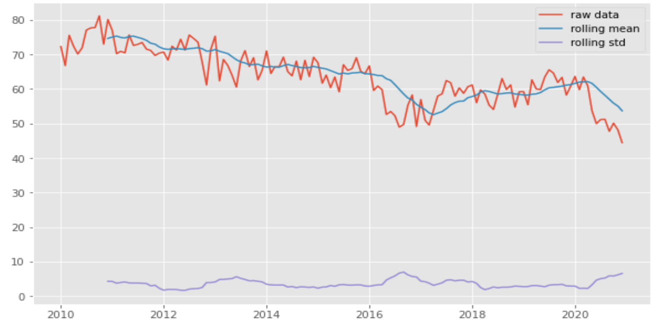

Fig. 4: Crude oil production stationarity data between the years 2010 and 2020

From figure 4 above, we can see that even though the standard deviation varies a little, the mean has so much variability in it that our time series data is stationary.

5. Test of Hypothesis

Hypothesis One:

H_0 : There is no significant difference in crude oil production in Nigeria between the years 2019 and 2020

H_1 : There is a significant difference in crude oil production in Nigeria between the years 2019 and 2020

Using Mann Whitney test.

```
In [108]: import scipy.stats as stats
          stats.mannwhitneyu(train2_test['2019'],train2_test['2020'], alternative = 'two-sided')

Out[108]: MannwhitneyuResult(statistic=116.0, pvalue=0.012022825407617439)
```

From the above using the Mann-Whitney test, the result shows that **Statistic = 116.0, P-value = 0.012022825407617439**

The above result shows that our P-value is 0.012, which is less than the level of significance (0.05). We, therefore, reject the null hypothesis and conclude that there is a significant difference in crude oil production between the years 2019 and 2020.

Hypothesis Two:

H_0 :The time series data is non-stationary

H_1 : The time series data is stationary

Using Augmented Dickey-Fuller Test

Test statistic = -1.045, P-value = 0.736

Table 1: Critical values before Detrending

%	Value	Comment
1%	-3.4860558292	Data is not stationary at 99% C. I
5%	-2.8859430324	Data is not stationary at 95% C. I
10%	-2.5797850694	Data is not stationary at 90% C. I

From the above result, our P-value is 0.736 at 99%, 95%, and 90% confidence Interval, it can be seen clearly that the p-value is greater than the level of significance at 99%, 95%, and 90%; this implies that we do not reject the null hypothesis. We can therefore conclude that our time series data is non-stationary. To make our data stationary and thus conformable for prediction and forecasting, we need to detrend it to remove the trend and later on test the hypothesis all over again. After Detrending the data, this gives us the below plot:

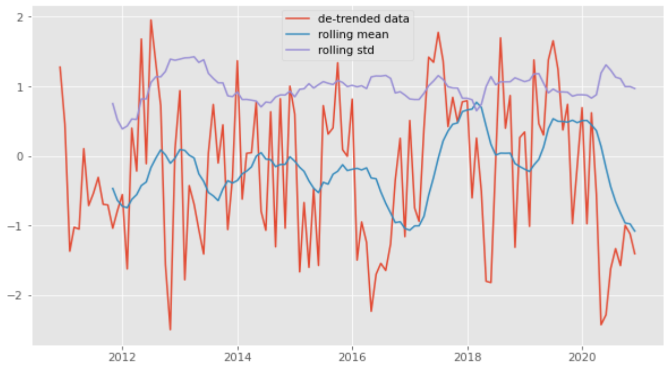

Fig. 5: Detrending data between the years 2012 and 2020

The result from Figure 5 shows that the variability in Mean and Standard Deviation is minimal, but there is still a need for us to confirm if truly the data is now stationary using the Augmented Dickey-Fuller Test.

Hypothesis Two:

H_0 :The time series data is non-stationary

H_1 : The time series data is stationary
From the analysis the test results give us:
Test statistic = -3.552, P-value = 0.007

Table 2: Critical values after Detrending

%	Value	Comment
1%	-3.4924012594	Data is stationary at 99% C. I
5%	-2.8886968193	Data is stationary at 95% C. I
10%	-2.5812552709	Data is stationary at 90% C. I

From the test result above, the P-value is 0.007 at alpha = 0.01, 0.05, and 0.1. The P-value is lower, i.e., 0.007 < 0.05. This implies that we reject the null hypothesis and therefore conclude that the time series data is stationary. This shows that the crude oil data contains seasonality.

The Seasonal Autoregressive Integrated Moving Average (SARIMA) will then be used for modelling, prediction, and forecasting, as this will enable us to know which approach is most appropriate for the prediction of the crude oil data. The evaluation metric we will be using for grid search is the Akaike Information Criterion (AIC) value. We will be using it to pick the combination with the lowest AIC value.
Result:

```
In [119]: sarima_grid_search(y,11)

The set of parameters with the minimum AIC is: SARIMA(0, 1, 1)x(1, 1, 1, 11) - AIC:607.2648982459093
```

From the result above, SARIMA (0, 1, 1) x (1, 1, 1, 11) has the lowest AIC value. We'll then be using this set of parameters to fit our model.

Table 3: SARIMA output

Coef	SE	Z	P>/z/	[0.025 0.975]		
Ma.L1	-0.4942	0.082	-5.991	0.000	-0.656	-0.333
Ma.S.L11	-1.0000	2828.403	-0.000	1.000	-5544.568	5542.567
Sigma 2	14.4851	4.1e+04	0.000	1.000	-8.03e+04	8.03e+04

The Root Mean Squared Error of SARIMA with season length = 11 and dynamic = False 4.21
The Root Mean Squared Error of SARIMA with season length = 11 and dynamic = True 13.58
Plotting the residuals

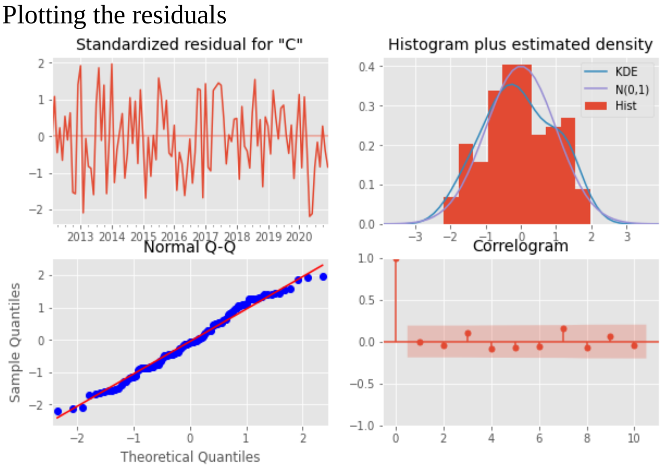

Fig. 6: Standardized Residual plots

The standard residual plot in figure 6 above shows the residuals over time; both the trend and seasonality have been found in our data, and the noise has also been removed. From the histogram plot, we can see that the KDE line follows closely with the N (0,1), which indicates that the residuals are normally distributed. From our QQ plot, the residuals follow the linear trend of the samples taken from a standard normal distribution with N (0, 1). The correlogram plot shows the time series residuals' low correlation. For all these points to have been validated, this indicates that we have found a best-fit model for the crude oil data. Using this prediction, we have:

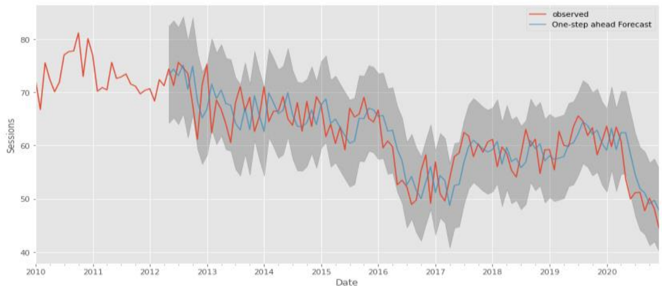

Fig. 7: Prediction for Crude Oil Data plots

The plot in figure 7 above shows the prediction of our crude oil data, the predicted value from May 2012, including its minimum and maximum value. It can be seen that the predicted values lie within the interval that is being predicted, and it also shows that the forecasted result is close to the actual crude oil production. Therefore, the SARIMA method is the best approach to predicting the Nigerian crude oil production data. For forecasting, after we've predicted the data, we also forecast the number of barrels of crude oil that can be produced in the next three years, as shown in the below plot.

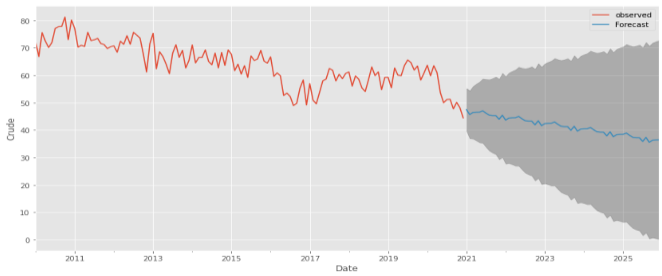

Fig. 7: Forecasted Data for Crude Oil Production for the Next Three Years

Table 4: Three Years Forecast Crude Oil Data

S/N	Date	Predicted Mean	Lower Bound	Upper Bound
1	1-Jan-2023	42.36258378	20.47144075	64.2537268
2	1-Feb-2023	42.46843459	20.12141077	64.81545841
3	1-Mar-2023	42.48718658	19.69339786	65.2809753
4	1-Apr-2023	42.95642981	19.72446617	66.18839345
5	1-May-2023	42.10686529	18.44483949	65.7688911
6	1-Jun-2023	41.39808903	17.31367922	65.48249883
7	1-Jul-2023	41.24193818	16.7424254	65.74145096
8	1-Aug-2023	41.23666082	16.32896207	66.14435957
9	1-Sep-2023	39.90161402	14.59231158	65.21091645
10	1-Oct-2023	41.37168825	15.54302284	67.20035367
11	1-Nov-2023	39.59822104	13.31239089	65.88405119
12	1-Dec-2023	40.33807274	13.62693614	67.04920934
13	1-Jan-2024	40.44392355	13.314147	67.5737001
14	1-Feb-2024	39.72221643	12.92062164	68.00472945

15	1-Mar-2024	39.48570111	12.98366855	68.880169
16	1-Apr-2024	39.24918579	11.39955411	68.43098118
17	1-May-2024	39.01267047	10.78171737	68.11700516
18	1-Jun-2024	38.77615515	10.16388063	68.35030484
19	1-Jul-2024	38.53963983	9.54604389	68.72934004
20	1-Aug-2024	38.30312451	8.92820715	67.773666
21	1-Sep-2024	38.06660919	8.310370411	69.7307709
22	1-Oct-2024	37.83009387	7.672533671	68.39232104
23	1-Nov-2024	37.59357855	7.074696932	69.53373728
24	1-Dec-2024	37.35706323	6.456860192	70.03605288
25	1-Jan-2025	37.12054791	5.839023453	70.44635927
26	1-Feb-2025	38.90740774	5.221186713	71.30242458
27	1-Mar-2025	38.05784322	4.603349973	70.83511737
28	1-Apr-2025	37.34906695	4.193942381	70.50419152
29	1-May-2025	37.19291611	3.664199016	70.72163319
30	1-Jun-2025	37.18763875	3.289446258	71.08583124
31	1-Jul-2025	35.85259194	1.588907942	70.11627594
32	1-Aug-2025	37.32266618	2.592888704	72.05244365
33	1-Sep-2025	35.54919897	0.398372864	70.70002507
34	1-Oct-2025	36.28905066	0.75222781	71.82587351
35	1-Nov-2025	36.39490148	0.476229595	72.31357336
36	1-Dec-2025	36.41365347	0.11714949	72.71015744

6. Conclusion and Recommendation

The production of crude oil in 2019 and 2020 showed a significant decline of about 12.3%. Records from the NNPC Limited website reveal that the Nigeria Refinery had been non-functional since June 2019, which contributed to this decline. The plotted data in research question two clearly indicates this fall in crude oil production during the specified years. After decomposing the crude oil data using an additive model, it was found to have a trend and seasonality. This was confirmed through the Augmented Dickey-Fuller Test.

After De-trended the time series data, it was discovered that the variability between the mean with time and standard deviation with time is minimal, which also makes the time series data stationary. This was also confirmed using the Augmented Dickey-Fuller test. We modelled the De-trended data using the Seasonal Autoregressive Integrated Moving Average (SARIMA), and it was deduced that the model SARIMA (0, 1, 1) x (1, 1, 1, 11) is the best fit for the data with the lowest AIC. We were also able to conclude that we got the best set of parameters for our data because we divided our data into a test and train dataset during analysis. The prediction plot for the test data was plotted, and the result shows that our predicted value follows the test data closely, and this was used to make a forecast. Seasonal Autoregressive Integrated Moving Average (SARIMA) could be regarded as a good statistical tool for modeling, predicting, and forecasting time series data because it gives accurate results when fitting a model.

It is highly recommended that the Nigerian Government takes steps to repair its existing refineries, construct new ones, and develop plans for maintaining both old and new refineries. The shutdown of refineries has a detrimental impact on crude oil production in Nigeria. Additionally, the government should provide more opportunities for private sector involvement in the oil industry. All data used in this study was obtained from the Nigeria National Petroleum Company Limited website. However, the research faced some limitations, such as financial constraints as the study was solely funded by the researchers and limited access to materials.

Data Availability
The data used for this research work is accessible through the Nigeria National Petroleum Company Limited website as secondary data. Also, any reader of this research work can have access to the data supporting the conclusions of this study through the NNPCL website, while some data are unavailable due to the operational privacy of the company and the Nigerian government. The main part of the study's limitations is the lack of funds

Conflict of Interest
No conflict of interest was experienced by the three authors during the research work of this journal.

Funding Source
This research was fully funded by the authors personally.

Authors' Contributions
N. F. Adeleye researched the literature part and also conceived the study. H. O. Ilo and N. F. Adeleye worked on the data analysis. While T. O. Gbadamosi wrote the first draft of the manuscript. N. F. Adeleye, H. O. Ilo and T. O. Gbadamosi all reviewed and edited the manuscript and approved the final version of the manuscript.

Acknowledgements
We will also like to use this medium to acknowledge the support of Dr. Ogunde of the department of Statistics, University of Ibadan, Nigeria and Mr. Fatoki Olayode of the department of Statistics, Ogun State Institute of Technology for their moral and academic support throughout the period of this research work.

References

[1] C. Okunroumu, "Niger's refinery now supplies fuel to Northern states" **This day Newspaper October 2004.**

[2] S. Falegan, & D. Okah, "Pipelines are safest for transportation of oil and gas," Manhattan Institute for Policy Research Issue Brief, **No. 23, 1-10, 1980.**

[3] J. Wurthmann, "Consumer gasoline prices: An empirical investigation." Monthly Labour Review**, 1-10. 2006.**

[4] D. Sabo, & H. Igwo, "Vertical integration by oil-exporting countries" OPEC Intereconomics, September/October, **pp 211-215, 2007.**

[5] A. A. Akinyemi, O. S. Enilolobo, & N. I. Theodore, "Impact of Oil Revenue and Exchange Rate Fluctuation on Economic Growth in Nigeria 1981-2015." Journal of Management and Administration, **2: 77-104, 2017.**

[6] J. Gylych, A. A. Jibrin, B. Celik, & A. Isik, "Impact of Oil Price Fluctuation on the Economy of Nigeria, the Core Analysis for Energy Producing Countries, Global Market and Global Trade," Intech Open, **pp. 1-28, 2020**.

[7] E. Odupitan, "Effects of crashing crude oil prices on oil producing countries: Nigeria's Perspective." Central University of Applied Sciences, Kokkola, Finland. **2017.**

[8] A. Adedokun, "The effects of oil shocks on government expenditures and government revenues nexus in Nigeria: With erogeneity restrictions," Future Business Journal, **4(2): pp 219-232, 2018.**

[9] O. S. Adesina, “Unemployment and Security Challenges in Nigeria: International Journal of Humanities and Social Science,” **3(7): pp 146-156, 2013.**

[10] E. G. Okoro, “Oil Price Volatility and Economic Growth in Nigeria: a Vector Auto-Regression (VAR) Approach,” ACTA UNIVERSITATIS DANUBIUSŒCONOMICA, 10(1): **pp 70-82, 2014.**

[11] S. A. Aniche, “Rising price of premium motor spirit (PMS) and widening poverty gap in Nigeria,” International Journal of Management Sciences and Humanities**, 1(1): pp12-25, 2015.**

[12] A. Englama, O. O. Duke, T. S. Ogunleye, & F. U. Ismail, “Oil Prices and Exchange Rate Volatility in Nigeria: An Empirical Investigation. Central Bank of Nigeria Economic and Financial Review,” 48(3), **pp 31-48, 2010.**

[13] Z. Dapel, “Poverty in Nigeria: Understanding and Bridging the Divide between North and South,” Center for Global Development Conference, **2018.**

[14] I. Otache, "The Effects of the Covid-19 Pandemic on Nigeria's Economy and Possible Coping Strategies," Asian Journal of Social Sciences and Management Studies, **7(3): pp 173-179, 2020.**

[15] J. Raji, G. Aitalohi, G. S. Loris, O. H. Antonio, T. Emilija, O. J. Orinya, S. M. Isa, M. A. Asfaw, I. Ifeyinwa, K. Masami, R. A. M. Tawfick, C. Mariano, Z. Siegfried, S. M. Rudibert, L. S. Craig, B. O. William, V. Tara, L. J. William, V. Julia, B. A. Peter, E. Jakob, L. Woori, M. E. Wamboka, D. Supriyo, Y. N. Desiree, A. Samik, & A. A. Gafar, “Nigeria in Times of COVID-19: Laying Foundations for a Strong Recovery (English),” Nigeria Economic Update Washington, D.C. World Bank Group, **2020.**

[16] OECD Policy Response to COVID-19, "The impact of coronavirus (COVID-19) and the global oil price shock on the fiscal position of oil-exporting developing countries," **2020.**

[17] O. I. Erasmus, M. E. Wamboka, D. Supriyo, Y. N. Desiree, “Deregulation of the downstream oil sector in Nigeria: History and fallacies,” Journal of Public Administration and Policy Research, **3(11): pp 128-139, 2011.**

[18] O. Fatoki, H. O. Ilo, & P. O. Ugwoke, “Time Series Analysis of Crude Oil Production in Nigeria,” Advances in Multidisciplinary & Scientific Research Journal, **Vol. 3. No.3, Pp 1-10, 2017.**

[19] G. E. Sirakayaa, & D. S. Woodside, “Optimal scheduling of petroleum products distribution in Nigeria,” Nigerian Journal of Technology, **26(1): pp 67-80, 2005.**

[20] E. A. Smallman, & K. I. Moore, “Militant oil agitations in Nigeria’s Niger Delta and the economy,” International Journal of Humanities and Social Science, **1(5): 140-145, 2011.**

[21] F. Yoo, & M. Chon, “Vertical Integration by oil-exporting countries,” OPEC Intereconomics, October, **pp 211-215, 2008.**

[22] N. F. Adeleye, & O. Fatoki, “Modelling the Reliability of a Complex Repairable System (Aircraft Engine) Using Weibull Distribution Model and Gamma Distribution Model (A Case Study of Arik Airline Aircraft Engine),” International Journal of Scientific Research in Mathematical and Statistical Sciences **Vol. 9, Issue-4, pp.23-32 2022.**

[23] B. Jabir, “Evaluation of Average Outgoing Quality for Reliability Acceptance Sampling Plan Based on the Exponentiated Rayleigh Distribution.” International Journal of Scientific Research in Mathematical and Statistical Sciences, Vol.8, Issue.6, pp.**40-45, 2021**.

AUTHORS PROFILE

N. F. Adeleye obtained his B.Sc. and M.Sc. in Statistics from the University of Lagos, Nigeria, in 2007 and 2012 respectively. After his graduation from the University, he has been working as a lecturer in the Department of Statistics at Ogun State Institute of Technology, Igbesa, Ogun State, Nigeria since the year 2010. He has been a member of the Nigerian Statistical Association (NSA) since 2015 and he is also a member of the Professional Statistical Society of Nigeria (PSSN) in the year 2022. He has published some notable research papers in reputed international journals and conferences that are available online. He has over 13 years of teaching experience and 9 years of research experience.

H. O. Ilo obtained his B.Sc. and M.Sc. in Mathematics and Statistics from Lagos State University, Ojo Lagos State, Nigeria, and the University of Lagos, Nigeria in the year 2004 and 2009 respectively. After his graduation from the university, he has been working as a lecturer in the Department of Statistics at Ogun State Institute of Technology, Igbesa, Ogun State, Nigeria, since 2007 till date. He has been a member of the Nigerian Statistical Association (NSA) since the year 2015 and a member of the Professional Statistical Society of Nigeria (PSSN) since 2001. He has published some notable research papers in reputed international journals and conferences that are available online. He has over 15 years of teaching experience and over 12 years of research experience.

T. O. Gbadamosi obtained his Ordinary National Diploma (OND) and Higher National Diploma (HND) in Statistics from the Ogun State Institute of Technology, Igbesa, Ogun State, Nigeria, in the years 2019 and 2022, respectively. He is a hardworking, dedicated, and goal-oriented statistician whose ambition is to become a major player in the world of statistics.

