Asian Journal of Probability and Statistics

Volume 22, Issue 1, Page 1-18, 2023; Article no.AJPAS.97967
ISSN: 2582-0230

Self-exciting Threshold Autoregressive Model with Application to Crude Oil Production in Nigeria

J. N. Onyeka-Ubaka [a*] and O. A. Ebiringa [a]

[a] Department of Statistics, Faculty of Science, University of Lagos, Akoka, Nigeria.

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2023/v22i1473

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/97967

Original Research Article

Received: 22/01/2023
Accepted: 25/03/2023
Published: 03/04/2023

Abstract

In the last two years the price of oil and its fluctuations have reached levels never recorded in the history of international oil markets. The determinants of past, current, and future levels of the price of oil and its fluctuations have been the subject of analysis by academics and energy experts, given the relevance of crude oil in the worldwide economy. The paper, therefore, model and forecast crude oil production (2002-2022) sourced from Central Bank of Nigeria website. The successive least squares estimation and model diagnostics are applied. The results affirmed that the self-exciting threshold autoregressive (SETAR(2,2,1)) model outperformed autoregressive integrated moving average (ARIMA(3,1,3)) model for crude oil production fluctuations based on our diagnostics (AIC, SC, SSR and log-likelihood ratio). The results obtained will greatly assist the government and policy makers in planning for economic development since the long-run success of any country is closely related to how well management is able to foresee the future and develop appropriate strategies.

Keywords: Backward shift; crude oil; SETAR; threshold autoregressive; volatility clustering.

**Corresponding author: Email: jonyeka-ubaka@unilag.edu.ng;*

1 Introduction

Crude oil (Oil) has dominated the economy of Nigeria since it was discovered in commercial quantity. In Nigeria, oil accounts for more than 90 percent of its exports, 25 percent of its gross domestic product (GDP), and 80 percent of its government total revenues, [1]. Thus, a slight oil price fluctuation can have a significant impact on the economic activities of the citizenry. For instance, a US$1 increment in the oil price in the early 1990s increased Nigeria's foreign exchange earnings by about US$650 million (2 percent of GDP) and its public revenues by US$320 million a year. Nigeria's dependence on oil production for income generation obviously has serious consequences for its economy. Secondly, oil is a salient commodity in the economy of any country because it is the main source of energy for domestic and industrial use. Different end products of oil are kerosene, diesel, gasoline, and others while changes in the prices of either the crude oil or any of the end products are expected to have impact on users and the nation at large. Given the incidences of vandalism, oil thefts, and other corrupt practices in the upstream enterprise of the sector, the paper is set to model and investigate variability in the crude oil production through volatility measures, as it was established in the crude oil pricing. In other words, the presence of volatility in crude oil production indicates that the data on crude oil is very volatile, indicating a strong similarity and reliance between the two economic variables as co-causes of the recent economic crisis in Nigeria. Volatility is a measure of how widely values deviate from the central mean value.

Crude oil price fluctuations are affected by global market mechanisms, which include the demand and supply equation and market sentiment. The demand and supply premise is straightforward and derived from the foundations of economics, however market sentiment occasionally reflects the simple expectation that oil demand would either increase or decrease sharply at some point in the future, leading to price changes, [2,3]. The recent changes in oil prices in the global economy are so rapid and unprecedented. This is partly due to activities of Organization of Petroleum Exporting Countries (OPEC) or increased demand of oil by countries like China and India. Some financial and physical reasons contribute to the volatility in oil prices. Physical factors include demand and supply mechanisms, weather events, technological advancements, geopolitical developments, and supply disruptions (including labour strikes, oil spills, vandalism and oil thefts), while financial factors include exchange rates, interest rates, speculations, and financial stress index, [4]. However, the current global economy melt down suddenly counteracted the sky-rocketing oil prices. These recessions are characterized by rapid decline in economic activity while economic expansions are characterized by a more moderate rate of change in economic activity that extends for a longer period of time. At the beginning of the crisis, oil price crashed below $40/b in the world market which had serious consequences on Nigeria fiscal budget which led to the downward review of the budget oil benchmark price, [5]. Today, oil price is oscillating between $60/b and $75/b. This rapid change has become a great concern to everybody including academics and policy makers; therefore a study of this kind is timely. It is a rather widely accepted view that the economy is nonlinear in the sense that major economic variables have nonlinear relationships, [6,7]. A business cycle with such characteristics would be asymmetric and linear models (e.g. autoregressive integrated moving average, ARIMA) lack the ability to capture such asymmetric processes and thus might be a less suitable choice for modelling macroeconomic time series that follows an asymmetric business cycle.

We observed that in the past decades, crude oil production have shown an extreme volatile, asymmetrical and nonlinear behaviours (irreversibility, jumps and limit cycles). These characteristics are collectively referred to as nonlinearity, [8]. If the derivative of the mean function with respect to the parameters relies on at least one other parameter, the model has a nonlinear mean function. Nonlinear models have been categorized into numerous classes: the bilinear models and the threshold models are the two most well-liked classes among them [9,10]. These family of non-linear models seem to capture asymmetries, limit cycles, and jump phenomena in the dynamic structure of climatic, economic, and financial time series. The TAR model developed by Tong [9] has proven to be a helpful and well-liked nonlinear time series modelling technique. Therefore, standard linear time series models, used commonly in econometric analysis, such as autoregressive (AR), autoregressive moving average (ARMA) and generalized autoregressive conditional heteroskedastic (GARCH) models fail to capture the complex dynamic of commodity markets, in particular future oil prices, [11-13]. Moreover, as this dynamic is driven by the economic cycle and seasonal movements, which in turn determine successive contraction and expansion of oil demand in the global economy, switching models, [14,15] are in a better position for the analysis. Also, from the forecasting point of view, there appears to be no clear conclusion as to whether allowing for non-linearity leads to an improvement in forecast performance, [16]. In this paper, we consider the fitting of nonlinear models with parameters changing according to a threshold, known as self-exciting threshold autoregressive (SETAR) models to crude oil production in Nigeria. Self-exciting threshold autoregressive model is a special example of the threshold autoregressive (TAR) model whereby regime switching is being based upon the self-dynamics of the dependent variable(s). It allows greater flexibility in model parameters which involves regime switching behaviour. For the TAR model, there is relationship between the threshold value and the exogenous variable while for the SETAR model, the relationship exists between the threshold value and the endogenous variable. The TAR model assumes that the regime is determined by a variable q_t relative to a threshold value (If q_t is equal to the dependent variable, say y_t, in an autoregressive regression, the model is referred to as self-exciting TAR (SETAR) model, [17].

The threshold technique makes sense because a regime dependent linear model can produce a piecewise linear structure that closely approximates nonlinearity. The TAR models enhance linear approximation by using threshold space. Threshold models' central concept is the use of thresholds to introduce regimes. Under the threshold principle, there exist numerous parametric nonlinear time series models and by breaking down a complicated stochastic system into smaller, simpler subsystems, the approach enables the analysis of the system. The oil and gas industry use TAR models, which have been effectively used to represent nonlinear time series in climate variables like temperature and humidity as well as financial variables like exchange rates, return volatility, and arbitrage trading. For instance, a TAR model of currency rates explains an outer regime of mean reversion with big deviations and an inner regime of slow adjustment for slight disequilibria or modest departures from some long-run equilibrium route or attractor. Handling the asymmetric responses in volatility between positive and negative returns is an essential use of TAR models in volatility. Arbitrage trading in index features and cash prices can be studied using TAR models. In addition to financial variables, TAR models have also been successfully utilized to investigate asymmetries in other macroeconomic variables over the course of the business cycle, such as unemployment, gross national product (GNP), etc. For time series systems that experience periodic shifts as a result of regime transitions, TAR models are particularly well suited. This nonlinear model's distinctive characteristics have made it useful in a variety of application areas.

2 Models

2.1 Autoregressive (AR) models

Autoregressive models are based on the idea that the current value of the series, X_t, can be explained as a function of *p* past values, $\mathrm{X}_{t-1}, \mathrm{X}_{t-2}, \cdots, \mathrm{X}_{t-p}$, that is, a linear combination of past values of the process plus random shock. An autoregressive model of order *p*, (AR(p)), can be written as;

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \varepsilon_t \tag{1}$$

where X_t is an invertible series, $\phi_1, \phi_2, \cdots, \phi_p$ are parameters and ($\phi_p \neq 0$). Unless otherwise stated, we assume that ε_t is a Gaussian white noise series with zero mean and variance σ^2_{w}.

2.2 Moving average (MA) model

This is linear combination of past errors of the process plus the current shock. Like the autoregressive representation in which the X_t on the left-hand side of the equation are assumed to combine linearly, the moving average model of order q, defining equation are combined linearly to form the data observed. A series X_t is said to follow a moving average of order *q* [M A (*q*)] if

$$\varepsilon_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + a_t \tag{2}$$

where θ_1, $\theta_2, \ldots, \theta_q$, are the MA parameters, MA*(q)* is a stationary series.

2.3 Autoregressive moving average (ARMA)

A general ARMA model can be written as follow:

$$\mathrm{X}_t - \phi_1 \mathrm{X}_{t-1} - \phi_2 \mathrm{X}_{t-2} - \cdots - \phi_p \mathrm{X}_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \tag{3}$$

Equation (3) is stationary and invertible. Re-writing in a backward shift operator B, we obtain

$$\Phi(\mathrm{B})\mathrm{X}_t = \Theta(\mathrm{B})\varepsilon_t \tag{4}$$

2.4 Autoregressive integrated moving average (ARIMA)

The autoregressive integrated moving average (ARIMA(p,d,q)) model, where parameters p, d and q are non-negative integers that refer to the order of the autoregressive, integrated, and moving average parts of the model, respectively. ARIMA models form an important part of the Box-Jenkins [6] approach to time series modelling. ARIMA model is a generalized model in that when one of the three terms (p,d,q) is zero, it is usual to drop to autoregressive (AR), integrated (I) or moving average (MA), respectively. Following Box and Jenkins [18], this can be represented as:

$$\varphi(\mathrm{B})x_t = \phi(\mathrm{B})(1-\mathrm{B})^d f(x_t) = \theta(\mathrm{B})\varepsilon_t \tag{5}$$

Where

$\varphi(\mathrm{B}) = (1-\mathrm{B})^d \phi(\mathrm{B})$ is the generalized autoregressive operator; it is a non-stationary operator with d of the roots of $\varphi(\mathrm{B}) = 0$ equal to unity. B is a backward shift operator such that $\mathrm{B}^d y_t^\lambda = y_{t-d}^\lambda$, and d is a nonnegative integer.

$\phi(\mathrm{B}) = 1 - \phi_1\mathrm{B} - \phi_2\mathrm{B}^2 - \cdots - \phi_p\mathrm{B}^p$, is an autoregressive operator of order p, such that the roots of the polynomial $\phi(\mathrm{B}) = 0$ lie outside the unit circle for stationarity. $\phi_1, \phi_2, \cdots, \phi_p$ are the autoregressive parameters. $f(y_t)$ is the observed time series data.

$\theta(\mathrm{B}) = 1 - \theta_1\mathrm{B} - \theta_2\mathrm{B}^2 - \cdots - \theta_q\mathrm{B}^q$, is a moving average operator of order q, such that the roots of the polynomial $\theta(\mathrm{B}) = 0$ lie outside the unit circle for invertibility. $\theta_1, \theta_2, \cdots, \theta_q$ are the moving average parameters. ε_t is normally independently distributed white noise with mean zero and variance σ_ε^2 .

When d = 0, the model (5) represents a stationary process, that is, the autoregressive moving average (ARMA) models, [19]. Here, the model for X_t is non-stationary because the AR operator on the left hand side has d root on the unit circle. If d is 1, we have a random walk (ARIMA (0,1,0)).

2.5 Threshold autoregressive model (TAR)

Consider a simple AR(p) model for a time series y_t which is a linear combination of past values of the process plus the current shock.

$$y_t = \mu + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \sigma_{\varepsilon t} \tag{6}$$

where $\varphi_i(i = 1, 2, \ldots, p)$ are the AR coefficients of i^{th} regime, $\mathrm{Y}_{t-1}, \mathrm{Y}_{t-2}, \cdots, \mathrm{Y}_{t-p}$ variables are the ones that are believed to influence Y_t dependent variable in each regime, ε_t is the error term $\varepsilon_t \sim N(0,1)$ and σ is the standard deviation of disturbance term. The model parameters $\Phi = (\mu, \varphi_1, \varphi_2, \ldots, \varphi_p)$ and σ are independent of time t and remain constant. To capture non-linear dynamics, TAR models allow the model parameters to change according to the value of a weakly exogenous threshold variable z_t:

$$y_t = X_t\varphi^j + \sigma^{(j)}\varepsilon_{ti}; \qquad r_{j-1} < z_{t-d} < r_j, \quad j = 1,\ 2, \cdots, k \tag{7}$$

where $X_t = (1, y_{t-1}, y_{t-2}, \cdots, y_{t-p})$, d shows the delay parameter, k denotes the number of the regimes in the model, z_{t-d} denotes the threshold variable and r denotes a real number $(-\infty < r < +\infty)$. In essence, the $k - 1$ non-trivial thresholds $(r_1, r_2, \ldots, r_{k-1})$ divide the domain of the threshold variable z_t into k different regimes. In each different regime, the time series, y_t follows a different AR(p) model. For example, a two-variable TAR model, [20] is:

$$\begin{aligned} Y_t &= (\phi_{1,0} + \phi_{1,1}Y_{t-1} + \phi_{1,2}Y_{t-2} + \cdots + \phi_{1,p}Y_{t-p})I(q_{t-1} \le \gamma) + \\ &(\phi_{2,0} + \phi_{2,1}Y_{t-1} + \phi_{2,2}Y_{t-2} + \cdots + \phi_{2,p}Y_{t-p})I(q_{t-1} > \gamma) + e_t \end{aligned} \tag{8}$$

where q_{t-1} denotes the threshold variable, $I(\cdot)$ denotes the indicator function, $q_{t-1} = q(Y_{t-1}, Y_{t-2}, \cdots, Y_{t-p})$ denotes the functional structure, γ denotes the threshold value or threshold parameter, e_t is the martingale difference series. The operation flow on the estimation of the model according to Tsay [21] are:

(i) Determine the level of AR process (p),
(ii) Selecting the d delay parameter,
(iii) Determining the level of the threshold value in the direction of the obtained p and d values.

2.6 Self-exciting threshold autoregressive model (SETAR)

If the threshold variable $z_t = y_{t-d}$, with the delay parameter d being a positive integer, the dynamics or regime of y_t is determined by its own lagged value y_{t-d} and the TAR model is called a self-exciting TAR or SETAR model. For the ease of notation, let SETAR(1) denotes the one regime linear AR model with $k = 1$, SETAR(2) denotes the two-regime TAR model with $k = 2$, and so on. For the one-regime SETAR(1) model, $-\infty < r_1 < \infty$ and the unknown parameters are $\Phi = (\varphi^{(1)}, \sigma^{(1)})$; for the two-regime SETAR (2) model, the unknown parameters include the single threshold $-\infty < r_2 < \infty$ and $\Phi = (\varphi^{(1)}, \varphi^{(2)}, \sigma^{(1)}, \sigma^{(2)})$. The two approaches for testing threshold non-linearity and estimating the unknown parameters in the associated models of SETAR are Tong [9] and Hansen [20], respectively.

$$y_t = \begin{cases} \lambda_0^{(1)} + \lambda_1^{(1)} y_{t-1} + \varepsilon_t & y_{t-1} \le \tau \\ \lambda_0^{(2)} + \lambda_1^{(2)} y_{t-1} + \varepsilon_t & y_{t-1} > \tau \end{cases} \tag{9}$$

or equivalently,

$$y_t = \lambda_{0,1} + \lambda_{0,2}I_t + \lambda_1 y_{t-1} + \lambda_2 I_t y_{t-1} + \varepsilon_t \tag{10}$$

where $\lambda_1^{(1)}$ and $\lambda_1^{(2)}$ are less than one, ε_t is a white noise process., $\tau \in \mathbb{R}$, $I_t = 0$ if $y_{t1} \le \tau$; $I_t = 1$ if $y_{t1} > \tau$, $\lambda_{0,1} = \lambda_0^1$; $\lambda_{0,2} = \lambda_0^1 - \lambda_0^2$; $\lambda_1 = \lambda_1^2$; $\lambda_2 = \lambda_1^1 - \lambda_1^2$, τ is called the threshold parameter. This is also a "two-regime" SETAR since the value of the autoregressive parameters depend on whether at time t the system is in regime l ($y_{t-1} \le \tau$) or regime 2 ($y_{t-1} > \tau$). More generally, we can imagine an r-regime SETAR such that:

$$y_t = \begin{cases} \lambda_0^{(1)} + \lambda_1^{(1)} y_{t-1} + \varepsilon_t & if \ y_{t-1} \le \tau_1 \\ \lambda_0^{(2)} + \lambda_1^{(2)} y_{t-1} + \varepsilon_t & if \ \tau_1 < y_{t-1} \le \tau_2 \\ \vdots \\ \lambda_0^{(r)} + \lambda_1^{(r)} y_{t-1} + \varepsilon_t & if \ \tau_{r-1} < y_{t-1} \le \tau_r \end{cases} \tag{11}$$

where $-\infty < \tau_1 < \cdots < \tau_r < \infty$

The SETAR model can also be generalized to p^{th} order autoregressive case

$$y_t = \begin{cases} \lambda_0^{(1)} + \lambda_1^{(1)} y_{t-1} + \cdots + \lambda_p^{(1)} y_{y-p} + \varepsilon_t & if \ y_{t-d} \le \tau_1 \\ \lambda_0^{(2)} + \lambda_1^{(2)} y_{t-1} + \cdots + \lambda_p^{(2)} y_{y-p} + \varepsilon_t & if \ \tau_1 < y_{t-d} \le \tau_2 \\ \vdots \\ \lambda_0^{(r)} + \lambda_1^{(r)} y_{t-1} + \cdots + \lambda_p^{(r)} y_{y-p} + \varepsilon_t & if \ \tau_{r-1} < y_{t-d} \le \tau_r \end{cases} \tag{12}$$

where $-\infty < \tau_1 < \cdots < \tau_r < \infty$, $d \in \{1, 2, . . ., p\}$, d is a delay parameter. So the form of the SETAR is determined by three parameters which are the lag length(p), the number of regimes (r), and the delay parameter d, which is sometimes denoted by SETAR(p,r,d). SETAR model with two regimes are specified in the following equation:

$$y_t = \begin{cases} a_0 + \sum_{i=1}^{p} a_i y_{t-i} + \varepsilon_t & if \ y_{t-d} \le \tau \\ \lambda_0 + \sum_{i=1}^{p} \lambda_i y_{t-i} + \varepsilon_t & if \ y_{t-d} > \tau \end{cases} \tag{13}$$

where a_i and λ_i are coefficients; τ is the threshold value; p is the SETAR model order; y_{t-d} is the variable of threshold; d is the parameter of delay and ε_t is a sequence of independent and identically distributed (iid) random variables with $\mu = 0$ and σ^2.

Or using Hansen [20] approach:

$$\begin{aligned} \mathrm{Y}_t &= (\alpha_0 + \alpha_1 \mathrm{Y}_{t-1} + \alpha_2 \mathrm{Y}_{t-2} + \cdots + \alpha_p \mathrm{Y}_{t-p}) \mathrm{I}(\mathrm{Y}_{t-d} \le \gamma) + \\ &(\beta_0 + \beta_1 \mathrm{Y}_{t-1} + \beta_2 \mathrm{Y}_{t-2} + \cdots + \beta_p \mathrm{Y}_{t-p}) \mathrm{I}(\mathrm{Y}_{t-d} > \gamma) + e_t \end{aligned} \tag{14}$$

where p denotes the autoregressive level, γ denotes the threshold parameter. The threshold variable is expressed as Y_{t-d} (d is an integer). Writing Equation (14) in more compact form, we have

$$\mathrm{Y}_t = \mathrm{X}_t(\gamma)'\omega + e_t \tag{15}$$

where $\omega = (\alpha'\beta')'$, $(\alpha = (\alpha_0 \alpha_1 \cdots \alpha_p)'$ and $\beta = (\beta_0 \beta_1 \cdots \beta_p)'$);

$\mathrm{X}_t(\gamma) = (\mathrm{X}_t' \mathrm{I}(\mathrm{Y}_{t-d} \le \gamma) \mathrm{X}_t' \mathrm{I}(\mathrm{Y}_{t-d} > \gamma))'$. Hence,

$$\mathrm{Y}_t = x'\alpha_t \mathrm{I}(\mathrm{Y}_{t-d} \le \gamma) + x_t'\beta \mathrm{I}(\mathrm{Y}_{t-d} > \gamma) + e_t \tag{16}$$

3 Methodology

The paper adapts estimation approaches of Hansen [20], Tong [9] and Tsay [21], where for example, the method for estimating the ω parameter and for the inclusion is the successive least squares method because the model parameters are nonlinear.

$$\hat{\omega} = \left(\sum_{t=1}^{n} \mathrm{X}_t(\gamma)\mathrm{X}_t(\gamma)' \right)^{-1} \left(\sum_{t=1}^{n} \mathrm{X}_t(\gamma)\mathrm{Y}_t \right) \tag{17}$$

$$\hat{e}_t(\gamma) = \mathrm{Y}_t - \mathrm{X}_t(\gamma)'\hat{\omega}(\gamma) \tag{18}$$

and the inclusion variance is

$$\sigma_n^2(\gamma) = \frac{1}{n}\sum_{t=1}^{n} \hat{e}(\gamma)^2 \,. \tag{19}$$

The minimization of Equation (19) is the best principle in the successive least squares estimation of the threshold parameter (γ), where $\hat{\gamma} = \arg\min \hat{\sigma}_n^2(\gamma),\ \ \gamma \in \Gamma = \left[\underline{\gamma}, \bar{\gamma}\right]$. The process will then follow an algorithmic system for selecting the threshold variable Y_{t-d} and the threshold parameter which will minimize the error variance. Equation (9) is estimated using three steps: (i) It is assumed that *d* and γ values are known. Based on these assumptions, the observation values are separated into small sub-groups, the Akaike information criterion (AIC) for each sub-group is calculated with $p_i \ \ (i = 1,\ 2, \cdots, k)$ level, where the maximum level in the regimes may be $L = n^a\,(a < \frac{1}{2})$. This gives

$$AIC(\hat{p}_i) = \min[AIC(k_i)],\ \ i = 1,\ 2 \tag{20}$$

where p_i value of each regime is obtained using Equation (20) in return for the constant *d* and γ values. (ii) The *d* value is kept constant while the threshold parameters that will minimize the AIC data criterion value are tested. That is, the γ value that minimizes the $AIC(d, \hat{\gamma})$ value is selected among the other threshold parameters.

$$AIC(d_0, \hat{\gamma}) = \min[AIC(d_0, \lambda)] \tag{21}$$

(iii) Having determined p_i and γ, the value of is obtained by minimizing the NAIC(*d*) as:

$$NAIC(d) = \frac{AIC(d)}{n - n_d}, \tag{22}$$

since the value of the variable *d* will influence the number of the observations (*n*) in the different regimes. The non-linear Akaike information criterion (NAIC) is used instead of the usual AIC.

The linearity tests, likelihood ratio test, Augmented Dickey Fuller and Jarque-Bera tests are performed. Afterward, the residuals of the mean equation is used to test for heteroskedasticity effects. If the heteroskedasticity effects are statistically significant, a volatility model is specified and then a joint estimation of the mean and volatility equations is performed. See more details at Akaul and Ozdemir; Caporale; Chan; Dickey and Fuller; Jarque and Bera; Yuehchao and Remya [22-27].

In modelling volatility, choosing an appropriate model from various suitable models is essential. The model selection principle is a criterion to assess if the fitted model suggests an optimum balancing between parsimony and goodness of fit. This paper utilized the frequently used model selection principle, the Akaike Information Criterion (AIC) for the fitted models. Hence, the best model has a lower AIC value and the highest log-likelihood metrics.

4 Results and Discussion

The analysis of the empirical data are performed using JUMPin software and R-Programming language. The time plot and dot plot of the monthly barrels of crude oil data series sourced from Central Bank of Nigeria (CBN) website are shown in Fig. 1 and Fig. 2 below, the trend showed the non-stationarity of the Nigeria crude oil production within the sample period (2002-2022).

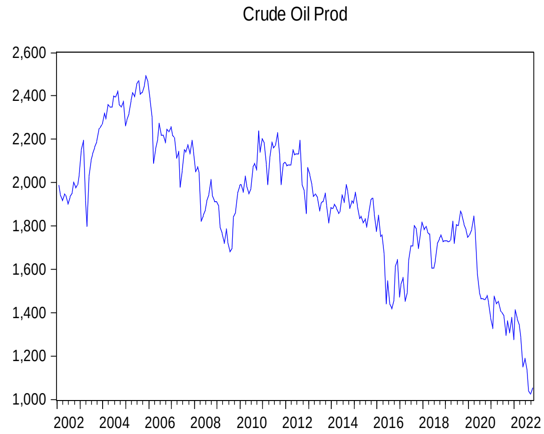

Fig. 1. Time plot of monthly barrels crude oil production

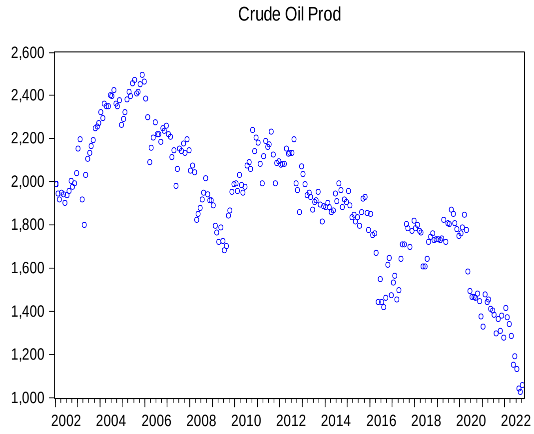

Fig. 2. Dot plot of monthly barrels crude oil production

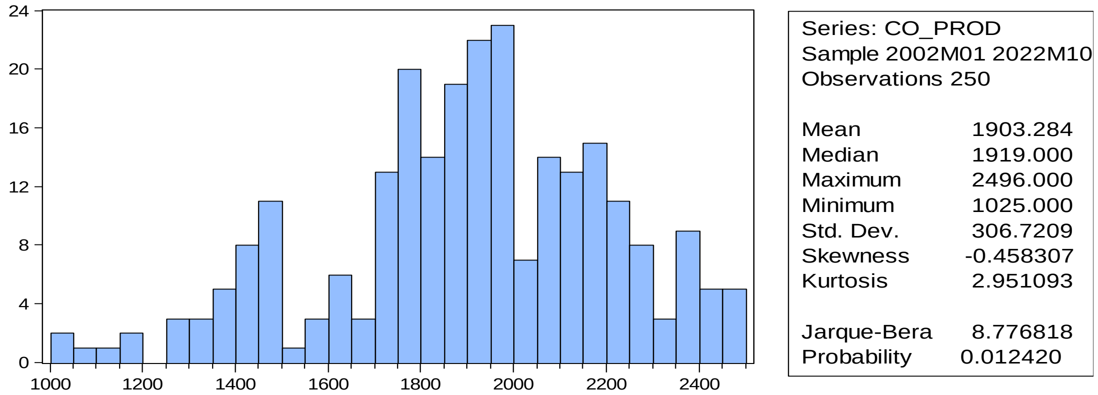

Fig. 3. Histogram of the monthly barrels crude oil productions

The histogram in Fig. 3 also confirms that the crude oil production in Nigeria is not stationary as it is skewed to the left with the value of -0.458; kurtosis of crude oil production gives platykurtic result with 2.951 and the Jarque-Bera statistic denotes that its errors are not normally distributed with $p = 0.012$ for crude oil productions. The basic statistics characterizes significant variations in the prices. The two economic variables appear to disagree with one another when compared to the fundamental statistical metrics in the table. This does not, however, rule out the prospect of crude oil production volatility. The theory of volatility clustering in crude oil output was proposed because fluctuations were thought to be significant during periods of unrest due to operational difficulties brought on by repeated attacks on multinational corporations and low during times of peace. The stationarity test (H_0: Crude oil production data is non-stationary versus H_1: H_0 is not true) obtained using Augmented Dicky Fuller (ADF) at 5% level of significance are presented in Table 1. The test shows that since the probability value of the ADF is greater than 5% (prob. = 0.8531 > 0.05), there is insignificant ADF value. Hence, the crude oil production data in Nigeria is not stationary (has a unit root test), so there is need for differencing.

Table 1. Stationary test of the monthly barrels crude oil production

		t-Statistic	Prob.*	R^2	AIC
Augmented Dickey-Fuller test statistic		-0.660371	0.8531	0.0067	11.43
Test critical values:	1% level	-3.456622			
	5% level	-2.872998			
	10% level	-2.572951			

Following [18,28,29], the non-stationary data can be differenced one or more than one time to achieve stationary. The data differenced and tested again with ADF indicate that the differenced series reached stationarity at the end of the first differenced as presented in Table 2. The probability value of the ADF is lesser than 5% (prob. = 0.0000 < 0.05), this implies that crude oil production data is now stationary.

Table 2. Stationary of the monthly barrels differenced data of crude oil production

		t-Statistic	Prob.*	R^2	AIC
Augmented Dickey-Fuller test statistic		-16.82602	0.000	0.535	11.424
Test critical values:	1% level	-3.456622			
	5% level	-2.872998			
	10% level	-2.572951			

The plot (Fig. 4) shows that the mean is equal to zero, variance approximately constant and presence of volatility clustering. The histogram of the differenced series are leptokurtic confirming our choice of self-exciting autoregressive models for capturing empirical characteristics present in our selected data within the sampled period.

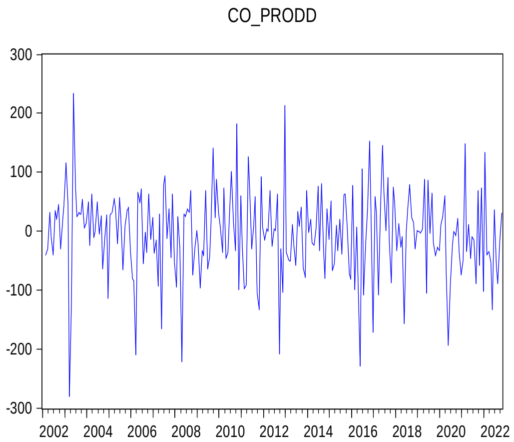

Fig. 4. Time plot of monthly barrels differenced data

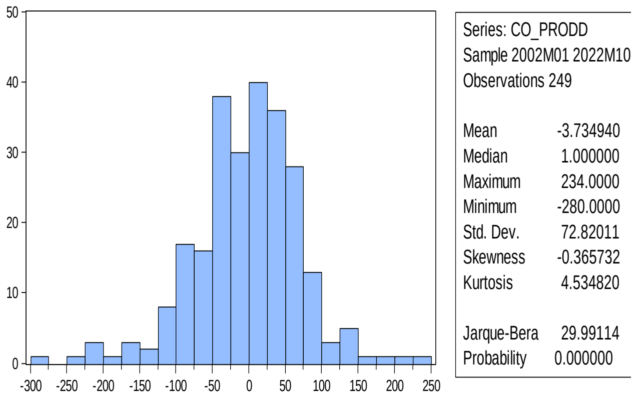

Fig. 5. Histogram of the monthly barrels differenced data

For identification of time domain process, it is more convenient to use the correlogram as model specification tools. Table 3 and graph (Fig. 6) of autocorrelation function (ACF) and partial autocorrelation function (PACF) are used.

Table 3. Correlogram of the monthly barrels differenced data of crude oil production

Autocorrelation	Partial Correlation	Lags	AC	PAC	Q-Stat	Prob
*|. |	*|. |	1	-0.070	-0.070	1.2372	0.266
*|. |	*|. |	2	-0.104	-0.110	3.9985	0.135

<table>
<tr><td>*|. |</td><td>*|. |</td><td>3</td><td>-0.069</td><td>-0.087</td><td>5.2220</td><td>0.156</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>4</td><td>0.049</td><td>0.025</td><td>5.8336</td><td>0.212</td></tr>
<tr><td>.|. |</td><td>*|. |</td><td>5</td><td>-0.061</td><td>-0.073</td><td>6.7822</td><td>0.237</td></tr>
<tr><td>*|. |</td><td>*|. |</td><td>6</td><td>-0.073</td><td>-0.084</td><td>8.1560</td><td>0.227</td></tr>
<tr><td>.|* |</td><td>.|* |</td><td>7</td><td>0.110</td><td>0.091</td><td>11.294</td><td>0.126</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>8</td><td>0.013</td><td>0.000</td><td>11.336</td><td>0.183</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>9</td><td>-0.055</td><td>-0.042</td><td>12.119</td><td>0.207</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>10</td><td>0.023</td><td>0.037</td><td>12.262</td><td>0.268</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>11</td><td>-0.033</td><td>-0.056</td><td>12.551</td><td>0.324</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>12</td><td>0.010</td><td>0.008</td><td>12.580</td><td>0.400</td></tr>
<tr><td>.|* |</td><td>.|* |</td><td>13</td><td>0.104</td><td>0.128</td><td>15.458</td><td>0.280</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>14</td><td>0.069</td><td>0.067</td><td>16.719</td><td>0.271</td></tr>
<tr><td>*|. |</td><td>.|. |</td><td>15</td><td>-0.072</td><td>-0.041</td><td>18.118</td><td>0.257</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>16</td><td>-0.054</td><td>-0.021</td><td>18.906</td><td>0.274</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>17</td><td>0.012</td><td>-0.013</td><td>18.944</td><td>0.332</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>18</td><td>-0.053</td><td>-0.059</td><td>19.706</td><td>0.350</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>19</td><td>-0.033</td><td>-0.019</td><td>20.008</td><td>0.394</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>20</td><td>0.051</td><td>0.016</td><td>20.726</td><td>0.413</td></tr>
<tr><td>.|. |</td><td>*|. |</td><td>21</td><td>-0.047</td><td>-0.091</td><td>21.342</td><td>0.438</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>22</td><td>0.043</td><td>0.055</td><td>21.855</td><td>0.469</td></tr>
<tr><td>*|. |</td><td>*|. |</td><td>23</td><td>-0.086</td><td>-0.086</td><td>23.907</td><td>0.409</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>24</td><td>0.059</td><td>0.028</td><td>24.861</td><td>0.413</td></tr>
<tr><td>.|* |</td><td>.|* |</td><td>25</td><td>0.105</td><td>0.136</td><td>27.942</td><td>0.311</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>26</td><td>0.015</td><td>0.011</td><td>28.004</td><td>0.358</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>27</td><td>-0.002</td><td>0.006</td><td>28.004</td><td>0.411</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>28</td><td>-0.053</td><td>-0.007</td><td>28.804</td><td>0.423</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>29</td><td>0.064</td><td>0.060</td><td>29.979</td><td>0.415</td></tr>
<tr><td>*|. |</td><td>*|. |</td><td>30</td><td>-0.115</td><td>-0.095</td><td>33.754</td><td>0.291</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>31</td><td>0.006</td><td>0.037</td><td>33.763</td><td>0.335</td></tr>
<tr><td>.|. |</td><td>*|. |</td><td>32</td><td>-0.035</td><td>-0.075</td><td>34.116</td><td>0.366</td></tr>
<tr><td>.|. |</td><td>*|. |</td><td>33</td><td>-0.020</td><td>-0.071</td><td>34.234</td><td>0.408</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>34</td><td>-0.026</td><td>-0.026</td><td>34.432</td><td>0.447</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>35</td><td>0.034</td><td>0.013</td><td>34.777</td><td>0.479</td></tr>
<tr><td>.|. |</td><td>.|. |</td><td>36</td><td>-0.029</td><td>-0.065</td><td>35.020</td><td>0.515</td></tr>
</table>

Since the autocorrelation function (ACF) and the partial autocorrelation function (PACF) decay exponentially to zero, either autoregressive model (AR) or moving average (MA) is suspected but the grid search table will be properly used to determine the order *p* or *q*. The spikes after the cut-offs in Fig. 6 indicates that the autoregressive moving average (ARMA) model of order *p*, *q* should be tried as well.

This gives the estimated parameters of autoregressive (p) and moving average (q) of the model using Akaike Information Criterion (AIC) for both self-exciting threshold autoregressive (SETAR) model and autoregressive moving average (ARMA) model.

Out of the AIC values generated from the Table 4, it was found that the most appropriate model in this case should be SETAR (2, 2, 1) with the threshold variable y_{t-2} (p = 2), r = 2 and d = 1 in order to compare the non-linear SETAR model with a linear model which is the best model for crude oil production in Nigeria.

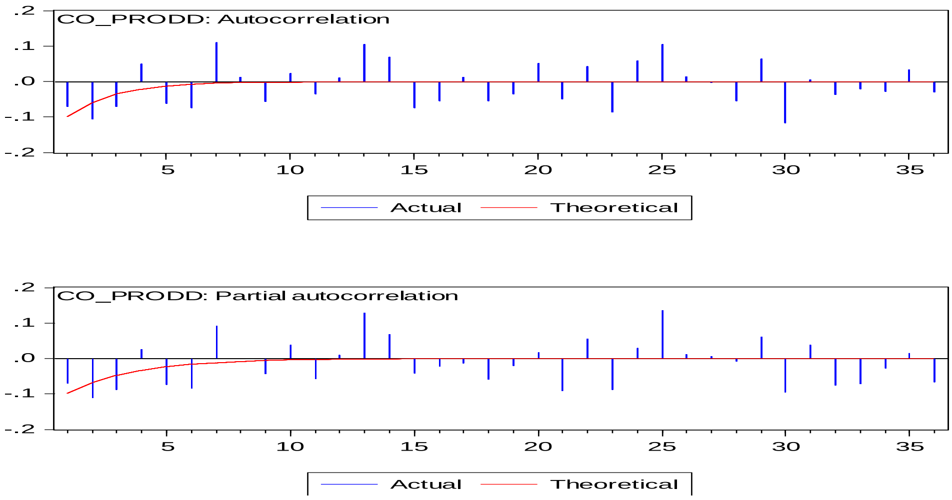

Fig. 6. Graphs of the ACF and PACF respectively, for monthly barrels of crude oil production

Table 4. SETAR AIC values of order *p*

Threshold variable	1	2	3	4	5	6	7	8
y_{t-1}	11.51	11.54	11.55	11.54	11.53	11.55	11.53	11.54
y_{t-2}	11.38	11.40	11.42	11.42	11.43	11.42	11.43	11.44
y_{t-3}	11.42	11.42	11.42	11.43	11.42	11.43	11.44	11.45
y_{t-4}	11.41	11.42	11.42	11.42	11.43	11.42	11.43	11.45
y_{t-5}	11.42	11.42	11.43	11.42	11.43	11.42	11.43	11.45
y_{t-6}	11.43	11.42	11.43	11.42	11.43	11.43	11.4	11.39

Table 5. ARIMA AIC Values of Order p and q

AR(p) / MA(q)	1	2	3	4	5	6	7	8
1	11.418	11.420	11.43	11.43	11.43	11.43	11.42	11.43
2	11.419	11.423	11.418	11.425	11.423	11.422	11.417	11.426
3	11.426	11.420	11.402	11.429	11.428	11.425	11.420	11.432
4	11.429	11.425	11.428	11.434	11.431	11.429	11.422	11.434
5	11.427	11.424	11.427	11.431	11.432	11.427	11.423	11.433
6	11.426	11.423	11.425	11.429	11.427	11.428	11.420	11.431
7	11.419	11.416	11.417	11.420	11.422	11.418	11.422	11.423
8	11.430	11.427	11.431	11.434	11.433	11.431	11.425	11.426

Out of the AIC values generated from the Table 5, the minimum values that is stationary and invertible occur at the order of AR(3) and MA(3) i.e. ARMA(3,3); equivalently to ARIMA(3,1,3) which is the best model for crude oil production in Nigeria. It implies that $p = 3$, $d = 1$ and $q = 3$.

The self-exciting threshold autoregressive (SETAR) model of SETAR (2,2,1) of order 2,2,1 (p, r, d) is given in Table 6.

The results and diagnostics of Tables 6 and 7, respectively affirmed the significance of the SETAR model used to fit the crude oil production with ($p < 0.000$) using the threshold variable with the value of 1476.671 which conform with the results of Patrick [30].

Fig. 7 depicts a plot visualizing the relationship between each threshold and the performance metrics: precision, recall, location and queue rate, with the addition of a single dominating vertex to the graph. That is, a single vertex that is connected to all other vertices of monthly crude oil production within the sample period.

Table 6. SETAR model parameters of crude oil production

Variable	Coefficient	Std. Error	t-Statistic	Prob
Threshold Variable (Linear Part)				
C	-446.9093	353.5963	-1.263897	0.2076
COPROD(-1)	0.842056	0.305558	2.755801	0.0064
COPROD(-2)	0.043863	0.299128	0.146635	0.8836
COPROD(-3)	-0.076106	0.387383	-0.196462	0.8444
COPROD(-4)	0.405556	0.304951	1.329904	0.1850
COPROD(-5)	-0.312596	0.309182	-1.011044	0.3131
COPROD(-6)	-0.020557	0.355418	-0.057839	0.9539
COPROD(-7)	0.806654	0.308709	2.612989	0.0096
COPROD(-8)	-0.412668	0.287937	-1.433188	0.1533
COPROD(-9)	-0.107007	0.287937	-0.376932	0.7066
COPROD(-10)	0.079383	0.377901	0.210064	0.8338
COPROD(-11)	0.053224	0.248051	0.214570	0.8303
Threshold Variables (Nonlinear Part)				
C	457.4781	356.6293	1.282783	0.2010
COPROD(-1)	0.092698	0.313400	0.295782	0.7677
COPROD(-2)	-0.092379	0.314269	-0.293949	0.7691
COPROD(-3)	0.091436	0.399157	0.229072	0.8190
COPROD(-4)	-0.330413	0.319729	-1.033416	0.3026
COPROD(-5)	0.299503	0.324947	0.921698	0.3577
COPROD(-6)	0.010998	0.368688	0.029829	0.9762
COPROD(-7)	-0.743632	0.324134	-2.294215	0.0228
COPROD(-8)	0.441474	0.304911	1.447882	0.1491
COPROD(-9)	0.002817	0.301008	0.009358	0.9925
COPROD(-10)	0.046529	0.390148	0.119261	0.9052
COPROD(-11)	-0.128726	0.258306	-0.498349	0.6188
Slopes				
Slope	4.661832	74297.69	6.27E-05	0.9999
Thresholds				
Threshold	1476.671	5245.681	0.281502	0.7786

Table 7. SETAR model diagnostics of monthly barrel of crude oil production

SETAR Model	Statistics
R^2	0.950866
Akaike Information Criterion (AIC)	11.38059
Schwarz Criterion (SC)	11.10883
Log-likelihood	-1390.812
Sum of Squared Residual (SSR)	1148936
P-Value	0.00000

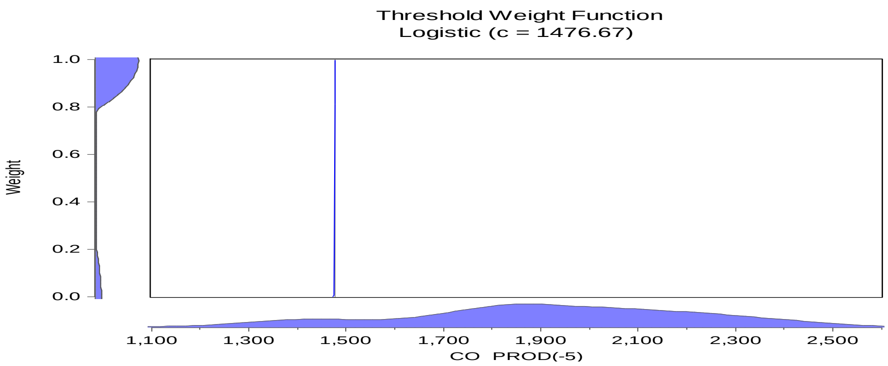

Fig. 7. Graph of threshold smoothing weight

Table 8. SETAR model of the monthly barrels of crude oil production with the order of determination

Variable	Coefficient	Std. Error	z-Statistic	Prob.
Regime 1				
COPRODD(-1)	0.416993	0.356376	1.170093	0.2420
COPRODD(-2)	-1.380890	0.223692	-6.173176	0.0000
Regime 2				
COPRODD(-1)	0.253228	0.091775	2.759229	0.0058
COPRODD(-2)	-0.161217	0.075517	-2.134836	0.0328
Common				
AR(1)	-0.442841	0.106210	-4.169501	0.0000
LOG(SIGMA)	4.114600	0.052856	77.84559	0.0000
Probabilities Parameters				
P1-C	-2.879446	0.429554	-6.703331	0.0000

From the Table 8, the AR(1) model can be written as:

$$\mathrm{X}_t = 4.114600 - 0.442841\mathrm{X}_{t-1} + \varepsilon_t \text{ with } p < 0.05.$$

The autoregressive integrated moving average (ARIMA) model of order 3, 1, 3 (*p*, *d*, *q*) i.e. ARIMA(3,1,3) is given in Table 9.

Table 9. ARIMA model of order determination of the monthly barrels of crude oil production

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.478424	0.262331	-1.823739	0.0694
AR(2)	-0.468634	0.244870	-1.913808	0.0568
AR(3)	0.328620	0.231780	1.417813	0.1575
MA(1)	0.434781	0.238146	1.825688	0.0691
MA(2)	0.391479	0.241679	1.619831	0.1066
MA(3)	-0.573021	0.233765	-2.451270	0.0149
ARIMA Model Diagnostics				
R-squared	0.601596		**AIC**	11.36118
SSR	1270749		**SC**	11.44668
Log likelihood	-1391.425			

From the Table 9, the model can be written as:

$$\mathrm{X}_t + 0.478424\mathrm{X}_{t-1} + 0.468634\mathrm{X}_{t-2} - 0.328620\mathrm{X}_{t-3} = \varepsilon_t - 0.434781\varepsilon_{t-1} - 0.391479\varepsilon_{t-2} + 0.573021\varepsilon_{t-3}$$

Here, adapting [31], model comparison is based on how to establish the better model between the self-exciting threshold autoregressive (SETAR) model and autoregressive integrated moving average (ARIMA) via checking the Akaike Information Criterion (AIC), Schwarz Criterion (SC), Sum of Square Residual (SSR) and Log-likelihood.

Table 10. Comparison of SETAR and ARIMA models

Models	AIC	SC	SSR	Log-likelihood	Remark
SETAR	**11.38059**	**11.10883**	**1148936**	-1390.812	**Better**
ARIMA	11.36118	11.44668	1270749	**-1391.425**	Good

It can be affirmed from the Table 10 that the SETAR Model outperformed the ARMA model since the model SETAR(2,2,1) had the least values of AIC, SC and sum of squares residual except the Log-likelihood which favoured the ARIMA(3,1,3). The iterative procedure are used to assess model adequacy by checking the basic assumption that ε_t's are uncorrelated random shocks (errors) with zero mean and constant variance.

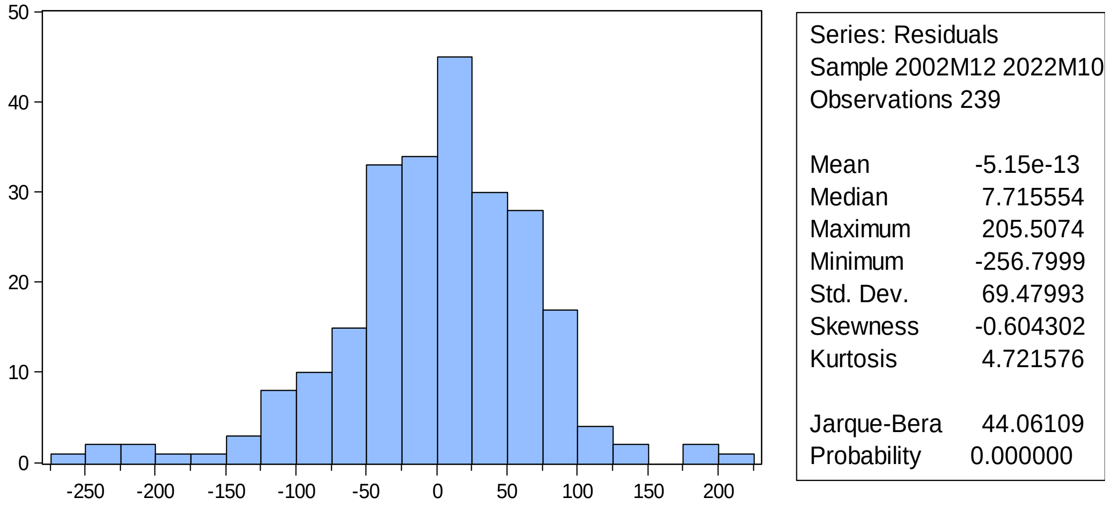

Fig. 8. Graph of crude oil production normality test

From Fig. 8, histogram of normalized residuals ($\varepsilon_t/\sigma^2_\varepsilon$) is established and compared with the normal distribution statistics using the Jarque-Bera test of the probability value of 0.000. For normality of the error terms, histogram of normalized residuals established and compare with the normal distribution using Chi-square (χ^2) goodness of fit test. The hypothesis is stated below:

$$H_0: \text{there is no serial correlation}$$
$$H_1: \text{there is serial correlation}$$

Table 11. Serial correlation test of the monthly barrels of crude oil production with setar model

Breusch-Godfrey Serial Correlation LM Test			
F-statistic	0.759949	Prob. F(2,211)	0.4690
Obs*R-squared	1.709279	Prob. Chi-Square(2)	0.4254

The results give the Breusch-Godfrey Test of serial Autocorrelation. Since P-values (0.4690, 0.4254) > $\alpha(0.05)$, there is no enough evidence to reject the null hypothesis. We therefore conclude at 5% level of significance that there is no serial correlation up to order *p* of the SETAR model. The test results for heteroskedasticity are presented on Table 12 using the hypothesis:

$$H_0: \sigma^2 = \sigma_i^2 \text{ Versus } H_0: \sigma^2 \neq \sigma_i^2$$

Table 12. Test of heteroskedasticity of the monthly barrels crude oil production with SETAR model

Heteroskedasticity Test: ARCH			
F-statistic	1.482140	Prob. F(11,227)	0.1392
Obs*R-squared	16.01517	Prob. Chi-Square(11)	0.1406

Since the P-values (0.1392 and 0.1406) > $\alpha(0.05)$, it implies that there is no enough evidence to reject the null hypothesis and therefore we conclude that the error variances are homoskedastic at 5% level of significance. The model evaluation on crude oil production data using SETAR model showed that none of the time series analysis assumptions is violated at 5% level of significance. Therefore, we move on to forecast future monthly crude oil production values for the next 5 years.

Table 13 gives forecast values of the crude oil production using SETAR (2,2,1) and ARIMA(3,1,3)

Table 13. Predicted monthly barrels of crude oil production in Nigeria

		Forecast Crude Oil Production	
Years	**Months**	**SETAR (2,2,1)**	**ARIMA(3,1,3)**
2022	November	1060.1	1087.48
	December	1067.16	1061.87
2023	January	1073.04	1048.05
	February	1079.04	1079.27
	March	1084.93	1063.24
	April	1090.77	1047.98
	May	1096.53	1072.26
	June	1102.23	1063.21
	July	1107.87	1048.08
	August	1113.44	1066.32
	September	1118.95	1062.14
	October	1124.39	1048.13
	November	1129.78	1061.3
	December	1135.1	1060.37
2024	January	1140.36	1047.99
	February	1145.57	1057.06
	March	1150.71	1058.14
	April	1155.8	1047.58
	May	1160.83	1053.44
	June	1165.8	1055.63
	July	1170.72	1046.89
	August	1175.58	1050.31
	September	1180.38	1052.97
	October	1185.13	1045.92
	November	1189.83	1047.54
	December	1194.47	1050.25
2025	January	1199.06	1044.7
	February	1203.6	1045.05
	March	1208.09	1047.54
	April	1212.52	1043.25
	May	1216.91	1042.74
	June	1221.25	1044.88
	July	1225.53	1041.61
	August	1229.77	1040.57
	September	1233.96	1042.28
	October	1238.11	1039.83
	November	1242.2	1038.47
	December	1246.25	1039.75
2026	January	1250.26	1037.92
	February	1254.21	1036.42
	March	1258.13	1037.3
	April	1262	1035.93
	May	1265.82	1034.4
	June	1269.61	1034.91
	July	1273.35	1033.88
	August	1277.04	1032.38
	September	1280.7	1032.59
	October	1284.31	1031.78
	November	1287.89	1030.36
	December	1291.42	1030.31

		Forecast Crude Oil Production	
Years	**Months**	**SETAR (2,2,1)**	**ARIMA(3,1,3)**
2027	January	1294.91	1029.65
	February	1298.36	1028.34
	March	1301.78	1028.08
	April	1305.15	1027.5
	May	1308.49	1026.3
	June	1311.79	1025.88
	July	1315.05	1025.35
	August	1318.27	1024.24
	September	1321.46	1023.72
	October	1324.61	1023.2
	November	1327.73	1022.18
	December	1330.81	1021.57

It can be affirmed from Table 13 that crude oil production in Nigeria will maintain steady increment from late 2022 to 2027 with the SETAR model while there seem to be fluctuation from late 2022 to 2027 with the ARIMA model. In that regards, SETAR model outperformed the ARIMA model in terms of forecasting with the available data. Therefore, there will be a great increment in the Nigeria crude oil production at gradual process in the next five years.

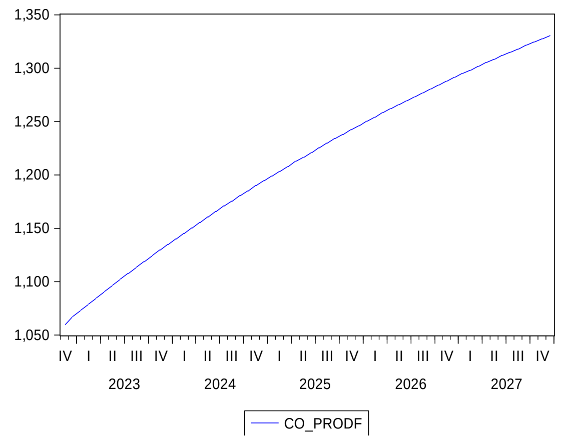

Fig. 9. Forecasting with SETAR model

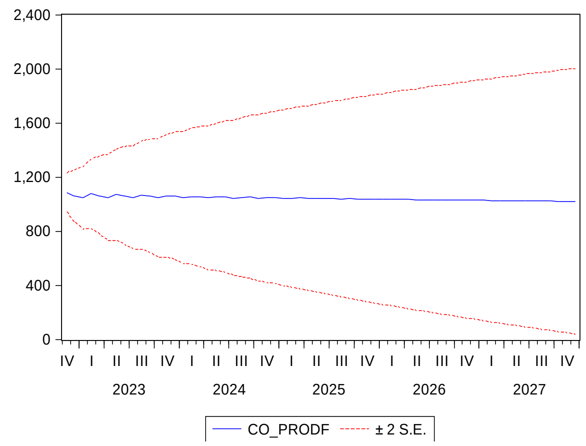

Fig. 10. Forecasting with ARIMA model

Figs. 9 and 10 affirmed that forecast of SETAR model performs better than ARIMA model for the years of forecasting, all other things being equal. Importantly, the authors submit that the forecast results will be of immense assets to the government, policymakers and citizenry in general if and only if the government will step up in eradicating or minimizing the activities of pipeline vandalism, crude oil theft and corruptions among workers and executives in oil sector.

5 Conclusion

The paper was consciously designed to achieve the following objectives: to fitting the univariate time series models for crude oil production, modelling the linear and asymmetries of crude oil production behaviour, comparing SETAR and ARIMA models and to forecast crude oil production for the next five years. To demonstrate the relationship and pattern between the crude oil production, exploratory data analysis (time plots and descriptive statistics), stationarity test via Augmented Dickey Fuller (ADF), model specification tools (autocorrelation function and partial autocorrelation function), Akaike information criterion (AIC), Schwarz criterion (SC), SETAR model, ARIMA model, normality test, serial correlation test, heteroskedasticity and forecast evaluation were carried out. The results affirmed that the best model for crude oil production is SETAR (2,2,1) based on our disgnostics (AIC, SC, SSR and log-likelihood). There is no co-integration on the series since the crude oil production series is stationary at integration order d = 1, from the Augmented Dickey-Fuller unit root test for the series. The output of this study will be beneficial to government, policymakers and citizenry in general. Setting and attaining the macroeconomic goals of the government requires accurate projection of the future behaviour of crude oil output. The benchmarking of crude oil production, which has an impact on price in the budgeting process, is one of the primary issues that Nigeria policymakers must deal with. Although benchmarking crude oil output and prices have been incorporated as a budgetary tool to protect the government from the ambiguities in the volatile international crude oil market knowing fully well that benchmarking could have different effects on the economy.

Acknowledgements

The Authors acknowledged all scholars whose works are cited in this paper.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] CBN annual statistical bulletin, Central Bank of Nigeria. Available from: http://www.cbn.gov.ng. Vol. 201. Abuja Retrieved from Central Bank of Nigeria. 2022;21.

[2] Agwuegbo SON, Onugha EE, Ezepue PO, Onyeka-Ubaka JN, Akintunde AA. Diffusion model for crude oil price dynamics. J Niger Assoc Math Phys. March 2017;40:215-22.

[3] Abiola AG, Okafor HO. Searching for appropriate crude oil price benchmarking method in the Nigerian budgeting process. Int Knowl Sharing Platform. 2013;3(12).

[4] Algia H, Abdelfatteh B. The volatility of oil prices: what Factors. Arab J Bus Manag. 2015;5:1-8.

[5] Central Bank of Nigeria (CBN) annual. Stat Bull.

[6] Onyeka-Ubaka JN, Okafor RO. Modelling crude oil price time-varying volatility using jump-diffusion model. Unilag Journal of Medicine. Sci Technol:135-52.

[7] Clements MP, Smith JP. Evaluating the forecast densities of linear and non-linear models: applications to output growth and unemployment. J Forecasting. 2000;19:255-76.

[8] Granger CWJ, Terasvirta T. Modelling nonlinear economic relationships. New York: Oxford University Press.

[9] Tong H. Nonlinear time series*:* A dynamical system approach. Oxford: Clarendor Press; 1990.

[10] Tong H, Lim KS. Threshold autoregressive limit cycles and cyclical data. J R Stat Soc B. 1980;42(3):245-92.

[11] Bollerslev T, Engle RF, Wooldridge JM. A capital asset pricing model with time-varying covariances. J Pol Econ. 1988;96(1):116-31.

[12] Deaton A, Laroque G. On the behaviour of commodity prices. Rev Econ Stud. 1992;59(1):1-23.

[13] Engle RF. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica. 1982;50(4):987-1007.

[14] Emenike KO. The Interrelationship between Crude oil price volatility and money market rate volatility in developing oil-producing economy. East Eur Bus J. 2017;3(1):28-47.

[15] Hamilton JD. A new approach to the economic analysis of nonstationarity time series and the business cycle. Econometrica. 1989;57(2):357-84.

[16] De Gooijer JG, Kumar K. Some recent developments in nonlinear time series modelling, testing and forecasting. Int J Forecasting. 1992;8:135-56.

[17] Zapata HO, Gauthier WM. Threshold models in theory and practice Selected Paper prepared for presentation at the Southern Agricultural Economics Association Annual Meeting, Mobile, AL, Feb 1-5, 2013.

[18] Box GP, Jenkins GM. Time Series Analysis, forecasting and control. San Francisco: Holden-day; 1976.

[19] Onyeka-Ubaka JN. 'A Modified BL-GARCH Models for Distributions with Heavy Tails,' A [thesis] submitted to the School of Postgraduate Studies. University of Lagos; 2013.

[20] Hansen BE. Inference in TAR models. Q J. 1997;2(1).

[21] Tsay RS. Testing and modeling threshold autoregressive processes. J Am Stat Assoc. 1989;84(405):231-40.

[22] Akaul I, Ozdemir S. Modelling and forecasting of turkey current account imbalance. 2014;52:207-32.

[23] Caporale GM. Modelling volatility of cryptocurrencies using Markov-Switching GARCH models. Res Int Bus Fin. 2019;48:143-55.

[24] Chan KS. Percentage points of likelihood ratio tests for threshold autoregression. J Int Money Fin. 1991;20:133-48.

[25] Dickey DA, Fuller WA. Distribution of the estimators for autoregressive time series with a unit root. J Am Stat Assoc. 1979;74:427-31.

[26] Jarque CM, Bera AK. Efficient tests of normality, homoscedasticity and serial independence of regression residuals. Econ Lett. 1980;6(3):255-9.

[27] Yuehchao W, Remya K. Time Series Analysis of Apple Stock Prices Using GARCH Models; 2017. Available: https://rstudio-pubs-static.s3.amazonaws.com

[28] Box GEP, Pierce DA. Distribution of residual autocorrelations in autoregressive moving average time series models. J Am Stat Assoc. 1970;65(332):1509-26.

[29] Box GP, Jenkins GM, Reinsel GC. Time series analysis: forecasting and control Wiley series. 4th ed. CA: Holden-day; 2008.

[30] Patrick M. Properties of the power envelope for tests against both stationary and explosive alternatives: the effect of trend. J Time S Anal. 2020;41(1):146-53.

[31] Boero G, Marrocu E. The performance of SETAR models: A regime conditional evaluation of point, interval and density forecast. Cen Ricerche Econ He NordSud; 2002.

© 2023 Onyeka-Ubaka and Ebiringa; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://www.sdiarticle5.com/review-history/97967