A Bootstrap Method for Box-Jenkins Models with Application on Brent Crude Oil Prices per Barrel

Adebowale Olusola Adejumo[1*] and James Daniel[2]

[1]University of Ilorin, Ilorin, Nigeria.
[2]National Bureau of Statistics, Abuja, Nigeria.

E-mail: aodejumo@gmail.com*

ABSTRACT

Oftentimes Box-Jenkins methodology is desirable to forecast time series data, appropriate model selection could be a challenge. Even when one selects the best model, estimation of the parameters of the selected model becomes another huddle to jump. One possible solution of these problems is application of bootstrap methods in time series. This method is employed as a means of model validation on some classes of financial time series of which the monthly Brent Crude Oil Price per Barrel belong.

In view of the above problem, this research was set to: explore and explain the behavior of the series; determine the best model; forecast future values of the series; and cross-validate suggested model through bootstrap method. With the aid of S-Plus programming ware, R language, and output by LATEX text ware; *ARIMA* (2*,* 1*,* 2)(2*,* 0*,* 0)[12] was reached as an optimum and parsimonious model for the series after a bootstrap technique was applied. An arithmetic increase in the price was forecast.

(Keywords: autocorrelation, partial autocorrelation, autoregressive integrated moving average, ARIMA, autoregressive moving average, ARMA, autoregressive, AR, moving average, MA, differencing, Brent crude oil, stationary series, bootstrap, block bootstrap, nonoverlapping block bootstrap, moving block bootstrap, circular block bootstrap, stationary block bootstrap, sieve bootstrap)

INTRODUCTION

Interestingly, according to Google's NGram, "bootstrapping" appeared in print for the first time around 1900 and then fell out of use for forty years before, coming back into use (in a different way) around 1945. The use of the term ”bootstrap” was derived from the phrase to pull oneself up by one's bootstraps, widely thought to be based on a phrase of the eighteenth century's ”The Surprising Adventures of Baron Munchausen” by Rudolph Erich Raspe:

> *The Baron had fallen to the bottom of a deep lake. Just when it looked like all was lost, he thought to pick himself up by his own bootstraps.*

It is nether the same as the term ”bootstrap” used in computer science meaning to boot a computer from a set of core instructions, nor ”bootstrap” used in business development environment to mean starting a business venture without ones start-up capital, though the derivations are similar.

Oil is one of the most important commodities in the world. The fluctuation of crude oil price affects global economy, and also affects our daily lives. The oil market is quite complex. Crude oil is a naturally occurring, yellow-to-black liquid found in geologic formations beneath the Earth's surface. It is a fossil fuel which is commonly refined into various types of fuels. Crude oil is distinguishing from petroleum that includes both naturally occurring unprocessed crude oil and petroleum products.

The oil price or the price of oil, generally refers to the spot price of a barrel of benchmark crude oil. The major benchmark oil prices in the world contain Brent crude oil price, WTI (West Texas Intermediate) crude oil price and Dubai/Oman crude oil price. The different types of oil are with different density and sulfur content, that leads to the oil price difference. Crude oil price are commonly measured in USD per barrel. The price of oil is affected by global economic conditions and supply and demand as well as market speculation.

Figure 1: Google Ngram Bootsrap

The International Energy Agency reported that high oil prices generally have a large negative impact on global economic growth. In the United States and Canada, the oil barrel (abbreviated as bbl) is a volume unit for crude oil, it is defined as 42 US gallons, which is equal to 159 liters or 35 imperial gallons. However, Outside the above two countries, volumes of oil are usually reported in cubic meters (m_3) instead of oil barrels. Hasan (2017).

Statistics is a subject of many users and surprisingly few effective practitioners. The traditional road to statistical knowledge is blocked, foremost, by a formidable wall of mathematics. A Bootstrap Method for Box-Jenkins Models with Application on Brent Crude Oil Prices per Barrel avoids that wall. It arms the learner, the users, as well as statisticians with the computational techniques they need to correctly predict and validate time series model. Two of the most important problems in applied statistics are the determination of an estimator for a particular parameter of interest and the evaluation or measurement of the accuracy for such estimator.

The best way to do this consist in using (whenever this is possible) the sampling distribution of the estimator we are concern with. In most cases, this sampling distribution is very difficult to obtain and only an asymptotic approximation is available. Anyway, it is very likely that either the true sampling distribution or its asymptotic approximation depends on some unknown population characteristics. As a result, it is necessary to find some known distribution function that is close (in some sense) to the sampling distribution of the estimator. A possible way to do this is the use of bootstrap method.

As I introduce this thesis, I am particularly motivated by these two problems. Most important is the measurement of accuracy of estimator, particularly when the estimator was complex and standard approximations were either not appropriate or too inaccurate. I am also motivated by the Abraham Waldâava's story of survivor's bias as narrated in the motivation subsection above. The need for estimation of parameter(s) for the model parameter(s) also call for measurement of accuracy of such a model.

Evidences from literature has shown that application of bootstrap in time series is most appropriate, but time series data being what it is, comes with its peculiar challenges of time dependency; serial correlation most times non-stationarity and some more.

A time series is a set of observations usually ordered in equally spaced intervals. The first step in the analysis of any time series is the description of the historic series. It includes the graphical representation of the data. When a time series is plotted, common patterns are frequently found. These patterns might be explained by many possible cause-and-effect relationships. Common components are the trend, seasonal effect, cyclic changes and randomness.

The more interesting and ambitious task is to forecast future values of a series on the basis of its recorded past, and more specifically to calculate prediction intervals (commonly refers to as the error term). So the identification of these components is important in the choice of a forecast model.

The aim of this research work is to present a better model validation scheme through bootstrap method; so as to create a model that best fit the special time series data of which the crude oil price per barrel belong. The specific objectives are to:

i. explore and explain the behavior of the series;

ii. determine the best model;

iii. forecast future values of the series; and

iv. cross-validate suggested model in (ii) through bootstrap method.

LITERATURE REVIEW

Léger and Romano (1990) discussed bootstrap technology and emphasized on modern issues and application such as variable selection, problems with dependent data, and determination of optimal replacement policy in a reliability study depending on the availability of fast computing power.

Davidson and MacKinnon (2006) proposed a weighted bootstrap method, also known in the literature as the wild bootstrap, which results in consistent variance of test statistics even in the presence of heteroscedacity. In this procedure, each observation of the original series is weighted, resampled with reposition from a standard normal distribution. Neumann and Kreiss, (1998) tested the validity of this method, in the contest of time series.

Efron and Tibshirani (1994) used simulation comparison to show that the use of bootstrap bias correction could provide better estimates of classification error rate than cross-validation approach often called leave-one-out which was originally proposed by Burman (1989) and Stone (1974). These results are applicable only in a small sample size. Later several follow-up articles were published that widened the applicability and sueriority of bootstrap (Chermick et al., 1995 1986, 1988; Efron land Tibshirani, 1997; Gong, 1986; Peter and Freedman, 1984).

Politis and Romano (1994) showed that the stationary bootstrap estimate of variance and the moving block estimate of variance are quite close provided that p^{-1} is approximately equal to *l*, where *l* is the block length and is the parameter of the geometric distribution.

Hall et al. (1995) presented a problem of maximum score estimation in estimation and hypothesis test that demonstrated the use and performance of the bootstrap and related resampling techniques provided practical methods for estimating the asymptotic distributions of statistics in carrying out statistical inference. According to them, statistical inference based on first order asymptotic approximation can be highly misleading for example, White (1982) information metric test which is a specification test for parametric models estimated by maximum likelihood is a well-known example of this, but the bootstrap often greatly reduces the error correction probability (ECP) of confidence intervals and error in the rejection probability (ERP) of test, thereby making reliable inference.

Carlstein et al. (1998) proposed sampling blocks according to a data based Markov chain so as to increase the likelihood that consecutive blocks match at their end. They gave conditions under which this matched-block bootstrap (MBB) reduces the bias of a bootstrap estimator of a variance. However, the moving block bootstrap increases the rate of convergence of bias only if the data generation process is a Markov process. The moving block bootstrap does not reduce the variance of the estimator.

Lahiri (1999a) compared the asymptotic minimal values of the mean square error of each of these four methods of the block bootstrap and concluded that the moving block bootstrap and circular block bootstrap are asymptotically equivalent in the sense of mean squared error (MSE). He affirmed that these are advantages in the use of moving block bootstrap and circular block bootstrap in relation to stationary block bootstrap method, even in samples of moderate size.

Berkowitz and Kilian (2000) studied the optimal block bootstrap method can be highly sensitive to the selection of size of the block while Liu and Singh (1992) indicated the stationarity problem of the resampled series by the moving block bootstrapping.

Bühlmann (2002) compared, reviewed and illuminated the theoretical facts about block, sieve and local bootstrap on the finite-sample data. According to Bühlmann (1999) it was shown that the two types of sieve's bootstrap outperform the block bootstrap and that the local bootstrap in some cases exhibits low performance.

Lahiri (2013) discussed block bootstrap procedures and suggested that the use of blocks of random size leads to bigger mean squared error than the ones obtained when blocks with non-random sizes are used. According to Lahiri (1999b) for a given block size *l*, the methods of moving block bootstrap (MBB), circular block bootstrap (CBB), stationary block bootstrap (SBB) and non-overlapping block bootstrap (NBB) Present, asymptotically, the same size of bias, but the variances of the estimators in stationary block bootstrap are Present, asymptotically, the same size of size of bias, but the variances of the estimators in stationary block bootstrap are always at least, twice the variance of the estimators for non-overlapping block bootstrapping and circular block bootstrap. According to Politis and White (2005) it occurs because of the additional randomization generated by bocks random size.

Box et al. (2015) in their book titled "Time Series Analysis, Forecasting and Control" emphasized that for an autoregressive process of order *p*, the Partial Autocorrelation Function (PAF) ϕ_{kk} will be non-zero for *k* less than or equals to p^{th} and zero for k greater than *p*. In another words, the Partial Autocorrelation Function (PAF) of p^{th} order autoregressive process has a cut-off after lag k. Generally, the autoregressive model is given as:

$$X_t = \phi_{11}X_{t-1} + \phi_{22}X_{t-2} + \phi_{33}X_{t-3} + \ldots + \phi_{pp}X_{t-p} + \varepsilon_t \qquad \textbf{(1)}$$

The complexity of such a long model was given relieve by their work which provided a cut off at a point where an autocorrelation coefficient break out of their confidence interval. The Box and Jenkins ARIMA techniques are based on the idea that a time series in which successive values are highly dependent can be regarded as being generated from series of independent shocks.

Gregory et al. (2015) Started by approximating a broad class statistics formulated through statistical functionals, They propose a smooth bootstrap by modifying a state-of-the-art (extended) tapered block bootstrap (TBB). Their treatment shows that the smooth TBB applies to time series inference cases not formally established with other TBB versions. Simulations also indicate that smoothing enhances the block bootstrap.

Kumar (2016) used Monte Carlo experiments using the weighted bootstrap, he evaluated the size and power properties in small samples of Chow and Denning's multiple variance ratio test and the automatic variance ratio test of Choi. His results indicate that the weighted bootstrap tests exhibit desirable size properties and substantially higher power than corresponding conventional tests.

METHODOLOGY

Introduction to Time Series Analysis

A time series is a stochastic process in discrete time with a continuous state space.

Notation: $\{X_1, X_2, \ldots, X_n\}$ denotes a time series process, whereas $\{x_1, x_2, \ldots xn\}$ denotes a univariate time series (i.e., a sequence of realizations of the time series process).

$$X_1, X_2, \ldots X_{n-1}, X_n, X_{n+1}$$

Autoregresive (AR)

An Autoregressive (AR) model of order *p*, or an AR(*p*) model, satisfied the equation:

$$X_t = \mu + \sum_{j=1}^{p} \phi_{jj} X_{t-j} + \varepsilon_t$$

$$X_t = \mu + \phi_{11}X_{t-1} + \phi_{22}X_{t-2} + \phi_{33}X_{t-3} + \ldots + \phi_{pp}X_{t-p} + \varepsilon_t$$

$$\text{for} \begin{cases} j : (1)p; \\ t \geq 0; \\ \sum\limits_{k=1}^{p} \varepsilon_t = 0. \end{cases} \qquad \textbf{(2)}$$

AR is a constant. The *p* denotes the order of autoregressive model, defining how many previous values the current value is related to. The model is called autoregressive because the series is regressed on to past values of itself. The error term *εt* in equation 2 refers to the noise in the time series. The error is said to be independently and identically distributed (iid). Commonly, they are also assumed to have a normal distribution.

$\varepsilon_t \sim \mathcal{N}(\mu, \sigma^2).$

For the model in 2 to be of use in practice, the estimator must be able to estimate the value of ϕ_{kk} and *μ*. Note the subscripts are defined so that the first value of the series to appear on the left of the equation is always one.

Moving Average (MA) Models

Moving Average (MA) model of order *q* or an MA (*q*) model is of the form:

$$X_t = \mu + \sum_{k=1}^{q} \theta_{kk}\varepsilon_{t-k} + \varepsilon_t$$

$$X_t = \mu + \theta_{11}\varepsilon_{t-1} + \theta_{22}\varepsilon_{t-2} + \theta_{33}\varepsilon_{t-3} + \ldots + \theta_{qq}\varepsilon_{t-q} + \varepsilon_t$$

$$\text{for} \begin{cases} k : (1)q; \\ t \geq 0; \\ \sum_{k=1}^{q} \varepsilon_t = 0 \end{cases} \qquad \textbf{(3)}$$

MA models imply the time series signal can be expressed as a linear function of previous value on the time series. The error (or noise) term in the equation (3) is the one step and ahead forecasting error. In contrast, MA (*q*) models imply the signal can be expressed as function of previous forecasting errors. It suggests MA (*q*) models make forecast based on the error made in the past, and so one can learn from the error made in the past to improve current forecast.

Autoregressive Moving Average (ARMA) Models

The best model is the simplest model that captures the important features of the data (parsimonious model).

Sometimes, however, neither a simple AR (p) model nor simple MA (q) model exists. In these case, a combination of AR (q) any MA (q) models will almost produce a simple model.

These models are called Autoregressive Moving Average Models, or ARIMA (p, q) models. Once again p is used for number of Autoregressive components, and q is for the number of Moving Average Component. Definition: The Form ARMA (p, q) model is given by the equation:

$$X_t = \mu + \sum_{j=1}^{p} \phi_{jj}X_{t-j} + \varepsilon_t = \mu + \sum_{k=1}^{q} \theta_{kk}\varepsilon_{t-k} + \varepsilon_t$$

$$X_t - \sum_{j=1}^{p} \phi_{jj}X_{t-j} = \mu + \sum_{k=1}^{q} \theta_{kk}\varepsilon_{t-k} + \varepsilon_t$$

$$X_t - \mu + \phi_{11}X_{t-1} + \phi_{22}X_{t-2} + \ldots + \phi_{pp}X_{t-p} + \varepsilon_t =$$

$$\mu + \theta_{11}\varepsilon_{t-1} + \theta_{22}\varepsilon_{t-2} + \ldots + \theta_{qq}\varepsilon_{t-q} + \varepsilon_t$$

$$\text{for} \begin{cases} j : (1)p; \\ k : (1)q; \\ t \geq 0; \\ \sum_{k=1}^{p} \varepsilon_t = 0; \\ \sum_{k=1}^{q} \varepsilon_t = 0. \end{cases} \qquad \textbf{(4)}$$

Where X_t is the observed data point. *μ* is some constants and ϕ_j, Θ_k are defined as for AR and MA model, respectively.

Autoregressive Integrated Moving Average (ARIMA) Model

ARIMA model models, sometimes called the Box-Jenkins models; named after the authors of the iterative Box-Jenkins methodology typically applied to time series data for forecasting consists of three parts: An Autoregressive (AR) part, a Moving Average (MA) part and the difference part.

The model is usually then referred to as the ARIMA models where p is the order of the Autoregressive part, d is the order of difference and q is the order of the moving Average part. For example, a model is referred to have ARIMA (1*,* 1*,* 1) when it has only one Autoregressive parameter called order one and only one moving Average parameter also called order one for the time series data after it was differenced once to attain stationary.

If the model becomes ARMA, which is linearly stationary on its self without being differenced. ARIMA is a linear non-stationary model. If the underlying time series is non-stationary, taking the difference of the series with itself some d-times makes it stationary, and then ARIMA model is applied onto the difference series.

Differencing and Unit Root

In the full class of ARIMA (p,d,q) models, the 'I' stand for integrated. The idea is that one might have a model\ whose terms are the partial sum up to time t, of some ARIMA models. Thus *X* consist of accumulated past shock, that is, shocks to the system, have a permanent effect. Note also that the variance of *Xt* increase without bound as time passes.

A series in which the variance and mean are constant and the covariance between X_t and X_s is a function only of the time difference ($t-s$) is called a stationary series. Clearly these integrated series are non-stationary. Another way to approach this issue is through the model. Again let us consider an autoregressive order 1 model, AR (1):

$$X_t - \mu = \rho(X_{t-1} - \mu) + \varepsilon_t$$

$$X_{t-1} - \mu = \rho(X_{t-1} - \mu) + \varepsilon_{t-1} \qquad \text{if} \begin{cases} |\rho| < 1; \\ t = (t-1) \end{cases}$$

$$X_t - \mu = \rho\varepsilon_{t-} + \rho^2(X_{t-1} - \mu) + \varepsilon_t$$

$$X_t - \mu = \sum_{t=1}^{n} \rho^i + \varepsilon_{t-1} + \varepsilon_t$$

(5)

Which is convergent expression satisfying the stationary conditions. However, if

$$|\rho| = 1$$

the infinite sum does not converge so one require a starting value say

$$X(0) = \mu$$

in which case,

$$X_t = \mu + \sum_{t=1}^{n} \varepsilon_{t-1} + \varepsilon_t$$

is the initial value plus an unweighted sum of or shocks they are permanent and the variance of X grows without bound over time. As a result there is no tendency of the series to return to any particular value and the use of a symbol for the starting value is perhaps a poor choice of symbols in that this is does not really represent a mean of any sort.

This type of series is called a random walk. Many stock series are believed to follow this or some closely related model. The failure of forecasts to return to mean in such a model implies that the best forecast of the future is current value and hence the strategy of buying low and selling high is pretty much eliminated in such a model whether high or low is unknown. Any Autoregressive model like

$$X_t = \phi X_{t-1} + X_{t-1}$$

has associated with a 'characteristic equation' whose root determine the stationary or non-stationary of the series.

AUTOCORRELATION

The correlation between a variable, lagged one or more periods, and itself is called autocorrelation. While the graphical tool that displays the correlations for various lags of time series is called correlogram. The coefficient of autocorrelation is derived bellow:

Autocorrelation Coefficient

$$\begin{aligned}
X_t - X_{t-1} &= (1-\alpha_1-\alpha_2)X_t + \alpha_2 X_{t-1} + X_{t-2} + \varepsilon_t \\
\mathbb{C}\mathrm{ov}(X,Y) &= \mathbb{E}(X-\mu_x)(Y-\mu_y) \\
\text{for } (X,Y) \text{ discrete} &= \sum_x \sum_y (X-\mu_x)(Y-\mu_y) \\
\text{for } (X,Y) \text{ continuous} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (X-\mu_x)dxdy \\
\text{racall that } \mathbb{E}(X-\mu_x)(Y-\mu_y) &= \mathbb{E}(X,Y) \\
\mathbb{E}(X-\mu_x)(Y-\mu_y) &= \mu_x, \mu_y \\
\therefore R(K) &= \mathbb{C}\mathrm{ov}(K_t, X_{t+k}) \\
&= \mathbb{E}(X_t-\mu_x)(Y_{t+k}-\mu_y) \\
&= \mathbb{E}(X_t-\mu_x-Y_{t+k}-\mu_y) \\
\rho_{xy} &= \frac{\mathbb{C}\mathrm{ov}(X,Y)}{\sigma_x\sigma_y} \\
&= \frac{\mathbb{E}(X,Y-\mu_x\mu_y)}{\sigma_x\sigma_y} \\
\mathbb{V}\mathrm{ar}(X_t) &= \sigma_x^2 \\
\rho_k &= \frac{\mathbb{R}(X)}{\mathbb{R}(0)} \\
\rho_K &= \frac{\mathbb{R}(X)}{\mathbb{R}(0)} \\
\therefore \rho_k &= \frac{\mathbb{E}(X_t-\bar{X})(X_t-\bar{X}_{t-k})}{\mathbb{E}(X_t-\bar{X})^2}
\end{aligned}
\qquad
\begin{cases}
\mu \text{ is a constant at time t} \\
K = 0, \pm1, \pm2, \cdots \\
-1 \le \rho_{xy} \le 1
\end{cases}$$

(6)

Properties of the Autocorrelation Function

$$\mathbb{R}(0) = \mathbb{V}\mathrm{ar}(X) = \sigma^2 \implies \rho_0 = \frac{\mathbb{R}(0)}{\mathbb{R}(0)} \qquad \textbf{(7)}$$

$$\mathbb{R}(K) = \mathbb{R}(-K) \implies \rho_K = \rho_{K-1} \qquad \textbf{(8)}$$

$$|\mathbb{R}(K)| \le \mathbb{R}(0) \implies \rho_k \le 1 \implies 1 \le \rho_K \le 1 \qquad \textbf{(9)}$$

HYPOTHESIS TESTING FOR TREAND IN THE SERIES

Actually, the autocorrelation coefficients for all time lags can be tested simultaneously. If the series is truly random, most of the series autocorrelation coefficients should lie within the specified by 0 plus or minus a certain number of standard errors, arranged as the bellow expression. Where z is the standard normal value for a given confidence level and n is the number of observation in the series:

$$0 \pm \left(\frac{1}{\sqrt{t}}\right) z_{\frac{\alpha}{2}}$$

Steps in Hypothesis Testing

In order to test for the presence of stationarity, randomness or seasonality, the following steps are taken:

i. Test Statistic

ii. Hypothesis setting

iii. Decision Rule

iv. Decision

v. Conclusion

Partial Autocorrelation Function (PACF)

In the previous section we have seen how the ACF can be used to identify MA processes, clearly indicating the order q of the process by the number of non-zero terms in the ACF. It would be great if we could similarly identify the order of an AR process. There is another correlation function which allows us to do precisely that: The Partial Autocorrelation Function or PACF. The PACF computes the correlation between two variables *yt* and *yt−k* after removing the effect of all intervening variables *yt*−1*, yt*−2*,* · · · *, yt−k*. We can think of the PACF as a conditional correlation: Another tool that will be needed is the partial autocorrelation function denoted by ϕ_{kk} where k = 1*,* 2*,* · · · *,K* is the set of partial autocorrelation at different or various lag K and it is defined by*:*

$$\phi_{KK} = \frac{\rho_K^*}{\rho_K},$$

and [ρ*K*] is a K by K autocorrelation matrix, and [ρ*K*] with the last column replaced by:

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix}$$

[ρ*K*] is a general form of Yule Walker's equation written as:

$$[\rho_k] = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{n-1} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \cdots & 1 \end{bmatrix}$$

Through which the Yule Walker's equation is formed as shown in the next section.

Parameter Estimation for Autoregressive Process

How does this PACF behave? Well, by convention we set

$$\phi_{11} = \rho_1 \qquad \textbf{(10)}$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} \qquad \textbf{(11)}$$

$$\phi_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} \qquad \phi_{33} = \frac{\rho_1^3 - \rho_1\rho_2^2 - \rho_1^2\rho_3 - 2\rho_1\rho_2 - \rho_3}{2\rho_1^2 - 2\rho_2 - 2\rho_2^2 - \rho_1^2 + 1} \tag{13}$$

etc.

Parameter Estimation for Moving Average Process

$$\phi_{22} = \frac{-\theta^2}{1+\theta^2+\theta^4} \tag{14}$$

$$\phi_{kk} = \frac{-\theta^k(1-\theta^2)}{1-\theta^{2(k+1)}} \tag{15}$$

Measure of Forecasting Accuracy

The forecast error is the difference between the actual value and the forecast value for the corresponding period. $\varepsilon_t = A_t - F_t$ Where ε_t is the forecasting error at period t, A_t is the Actual Value at period t and F_t is the forecast for period t . These are as shown in Table.

Table 1: Formulas for Measuring Forecasting Accuracy.

ERROR MEASUREMENT	FORMULA
Mean Absolute Error (MAE)	$\text{MAE} = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t$
Mean Square Error (MSE)	$\text{MSE} = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2$
Root Mean Squared Error (RMSE)	$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2}$

MODEL BUILDING STRATEGY

The below Figure 2 is the flow chart that depicts steps taken in the series model building:

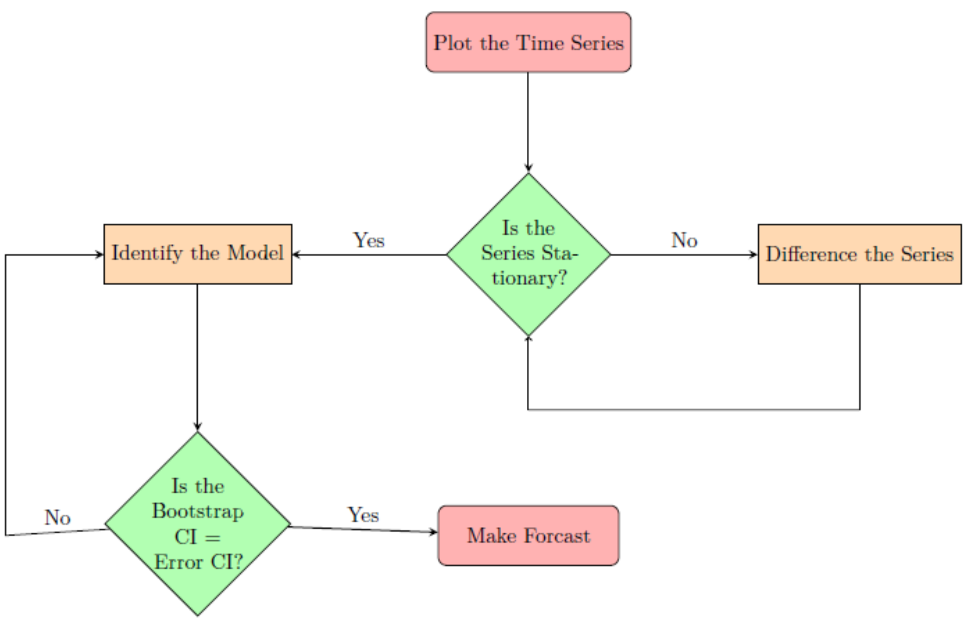


Figure 2: ARIMA MODEL BUILDING FLOWCHART

RESULTS AND DISCUSSION

Data Presentation

381 univariate data set was collected every first trading day of the month from January 1985 to September 2016 chronologically; which I adopted from the Organization of Petroleum Exporting Countries (OPEC) Statistical Bulletin shown in Table 2 below.

Table 2: BRENT CRUDE OIL PRICES PER BARREL($)

Yr/Mth	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
1985	**26.63**	**27.56**	**27.91**	**27.84**	**26.66**	**26.35**	**26.63**	**27.13**	**27.52**	**28.29**	**29.18**	**26.75**
1986	**22.33**	**16.38**	**12.77**	**11.89**	**13.49**	**11.95**	**9.88**	**13.42**	**14.03**	**14.01**	**14.47**	**15.44**
1987	**18.10**	**17.28**	**17.74**	**18.05**	**18.41**	**18.71**	**19.62**	**18.88**	**18.32**	**18.63**	**17.87**	**16.77**
1988	**16.50**	**15.9**	**14.86**	**16.42**	**16.36**	**15.49**	**14.47**	**14.57**	**13.22**	**12.23**	**12.53**	**14.68**
1989	**16.45**	**16.52**	**18.10**	**19.39**	**18.22**	**17.8**	**17.67**	**16.87**	**17.69**	**18.41**	**18.38**	**19.37**
1990	**20.59**	**19.68**	**18.12**	**16.32**	**16.21**	**14.93**	**16.81**	**26.54**	**33.62**	**34.85**	**31.54**	**26.61**
1991	**22.81**	**18.53**	**18.21**	**18.49**	**18.72**	**17.78**	**19.02**	**19.30**	**19.95**	**21.56**	**20.41**	**17.63**
1992	**17.52**	**17.65**	**17.35**	**18.65**	**19.52**	**20.88**	**20.18**	**19.62**	**20.19**	**20.04**	**18.90**	**17.93**
1993	**17.24**	**18.23**	**18.5**	**18.44**	**18.17**	**17.37**	**16.37**	**16.43**	**15.80**	**16.44**	**15.09**	**13.36**
1994	**14.17**	**13.75**	**13.69**	**15.15**	**16.43**	**17.23**	**18.04**	**16.98**	**16.13**	**16.48**	**17.20**	**16.13**
1995	**16.88**	**17.44**	**17.35**	**18.77**	**18.43**	**17.33**	**16.06**	**16.49**	**16.77**	**16.18**	**16.82**	**17.93**
1996	**17.79**	**17.69**	**19.46**	**20.78**	**19.12**	**18.56**	**19.56**	**20.19**	**22.14**	**23.43**	**22.25**	**23.51**
1997	**23.29**	**20.54**	**19.42**	**17.98**	**19.47**	**18.02**	**18.45**	**18.79**	**18.73**	**20.12**	**19.16**	**17.24**
1998	**15.07**	**14.18**	**13.24**	**13.39**	**13.97**	**12.48**	**12.72**	**12.49**	**13.8**	**13.26**	**11.88**	**10.41**
1999	**11.32**	**10.75**	**12.86**	**15.73**	**16.12**	**16.24**	**18.75**	**20.21**	**22.37**	**22.19**	**24.22**	**25.01**
2000	**25.21**	**27.15**	**27.49**	**23.45**	**27.23**	**29.62**	**28.16**	**29.41**	**32.08**	**31.4**	**32.33**	**25.28**
2001	**25.95**	**27.24**	**25.02**	**25.66**	**27.55**	**26.97**	**24.8**	**25.81**	**25.03**	**20.73**	**18.69**	**18.52**
2002	**19.15**	**19.98**	**23.64**	**25.43**	**25.69**	**24.49**	**25.75**	**26.78**	**28.28**	**27.53**	**24.79**	**27.89**
2003	**30.77**	**32.88**	**30.36**	**25.49**	**26.06**	**27.91**	**28.59**	**29.68**	**26.88**	**29.01**	**29.12**	**29.95**
2004	**31.40**	**31.32**	**33.67**	**33.71**	**37.63**	**35.54**	**37.93**	**42.08**	**41.65**	**46.87**	**42.23**	**39.09**
2005	**42.89**	**44.56**	**50.93**	**50.64**	**47.81**	**53.89**	**56.37**	**61.87**	**61.65**	**58.19**	**54.98**	**56.47**
2006	**62.36**	**59.71**	**60.93**	**68.00**	**68.61**	**68.29**	**72.51**	**71.81**	**61.97**	**57.95**	**58.13**	**61.00**
2007	**53.40**	**57.58**	**60.60**	**65.10**	**65.10**	**68.19**	**73.67**	**70.13**	**76.91**	**82.15**	**91.27**	**89.43**
2008	**90.82**	**93.75**	**101.84**	**109.05**	**122.77**	**131.52**	**132.55**	**114.57**	**99.29**	**72.69**	**54.04**	**41.53**
2009	**43.91**	**41.76**	**46.95**	**50.28**	**58.10**	**69.13**	**64.65**	**71.63**	**68.38**	**74.08**	**77.56**	**74.88**
2010	**77.12**	**74.72**	**79.3**	**84.14**	**75.54**	**74.73**	**74.52**	**75.88**	**76.11**	**81.72**	**84.53**	**90.07**
2011	**92.66**	**97.73**	**108.65**	**116.32**	**108.18**	**105.85**	**107.88**	**100.45**	**100.83**	**99.92**	**105.36**	**104.26**
2012	**106.89**	**112.70**	**117.79**	**113.75**	**104.16**	**90.73**	**96.75**	**105.28**	**106.32**	**103.39**	**101.17**	**101.17**
2013	**105.04**	**107.66**	**102.61**	**98.85**	**99.35**	**99.74**	**105.21**	**108.06**	**108.78**	**105.46**	**102.58**	**105.49**
2014	**102.25**	**104.82**	**104.04**	**104.94**	**105.73**	**108.37**	**105.22**	**100.05**	**95.89**	**86.13**	**76.96**	**60.55**
2015	**47.45**	**54.93**	**53.83**	**57.42**	**62.50**	**61.3**	**54.43**	**45.73**	**46.29**	**46.96**	**43.13**	**36.56**
2016	**29.92**	**31.05**	**37.34**	**40.75**	**45.98**	**47.69**	**44.22**	**40.00**	**46.21**			

Source:Organization of Oil Exporting Countries (OPEC) Statistical Bulletin

Data Exploration

The most obvious reasons for analyzing an important series like this, is to find a way to accurately forecast its future values. However, the analysis process itself sometimes reveals important insight into the series that will help make better decisions.

Histogram

Histogram of Crude Oil Per Barrel is seen in Figure 3.

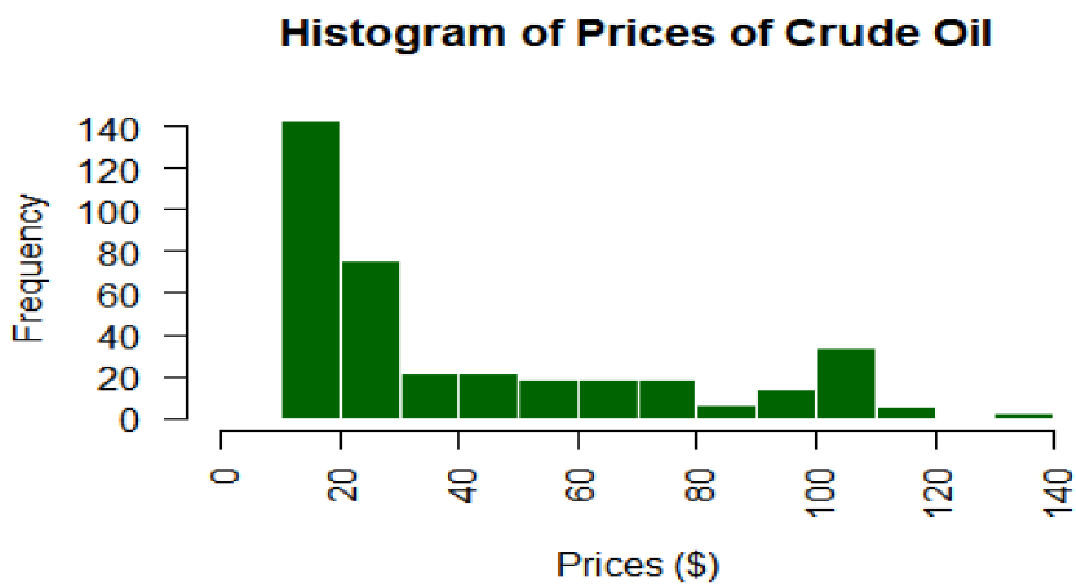


Figure 3: Histogram of Crude Oil Per Barrel

Whisker Box Plot

The oversight way to check for seasonality in time series is through the box plot as shown in Figure 4.

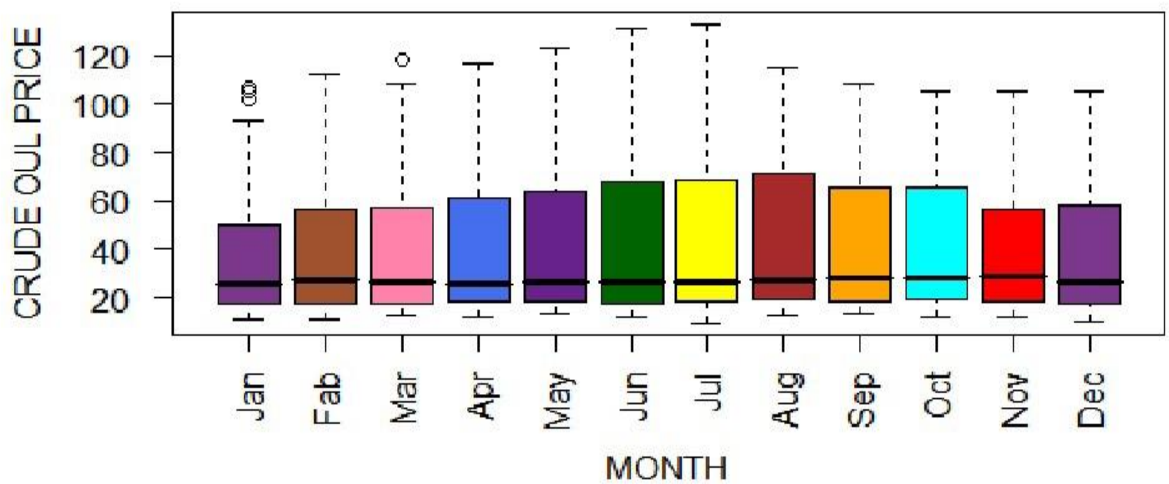


Figure 4: Whisker Box plot of Crude Oil per Barrel.

Pareto Diagram

The Pareto theorem of (80/20)% holds for the price regime of $20 interval. Figure 5 shows that 20% ($1 - $21) and ($41 - $41) of price subgrop experienced 80% of price reign of the crude oil prices.

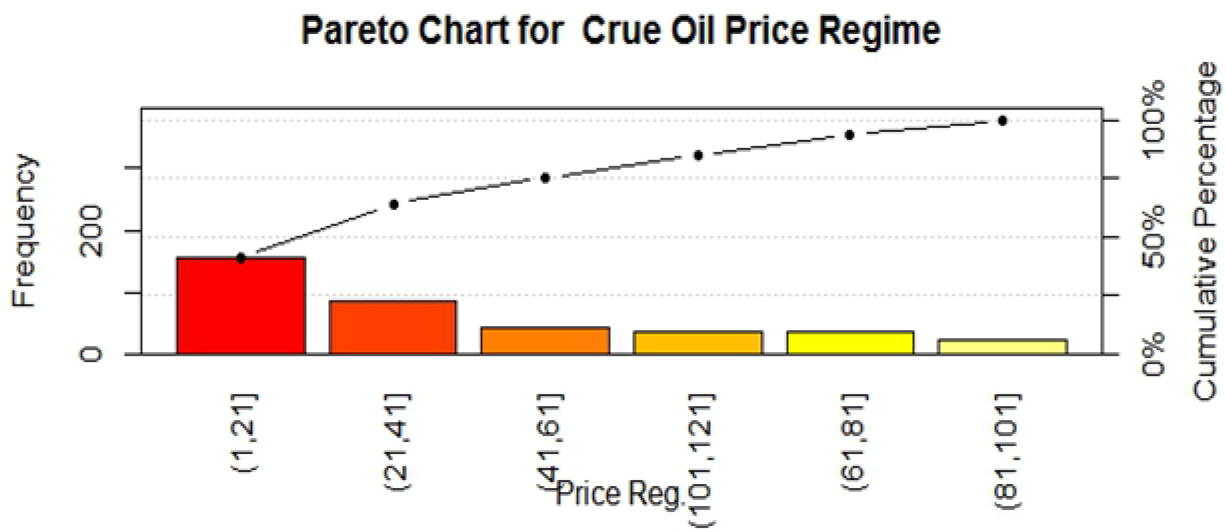


Figure 5: Whisker Box plot of Crude Oil per Barrel.

Time Plot

The first step in any time series analysis is to plot the observations against time. The time plot was achieved through the use of S-plus software.

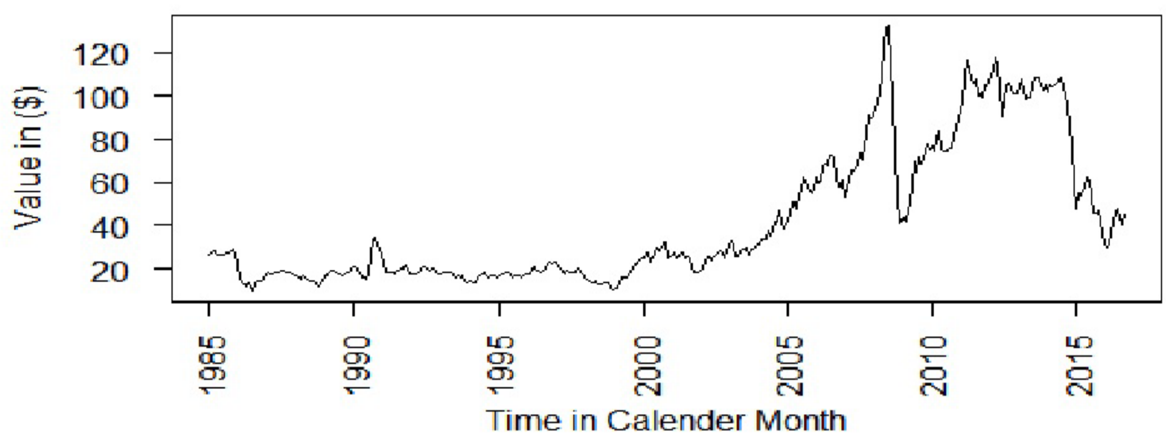


Figure 6: Time plot of Crude Oil per Barrel.

The time plot in Figure 6 revealed that the monthly crude oil price is non-stationary. The focal assumption of time series analysis is stationarity of the series out of other assumptions like randomness of the series and normality of the series. To be critical in examining the presence trend (non-stationary series) a hypothesis is set up. The time plot is as shown as in Figure 6.

Time series decomposition diagram as shown in Figure 7.

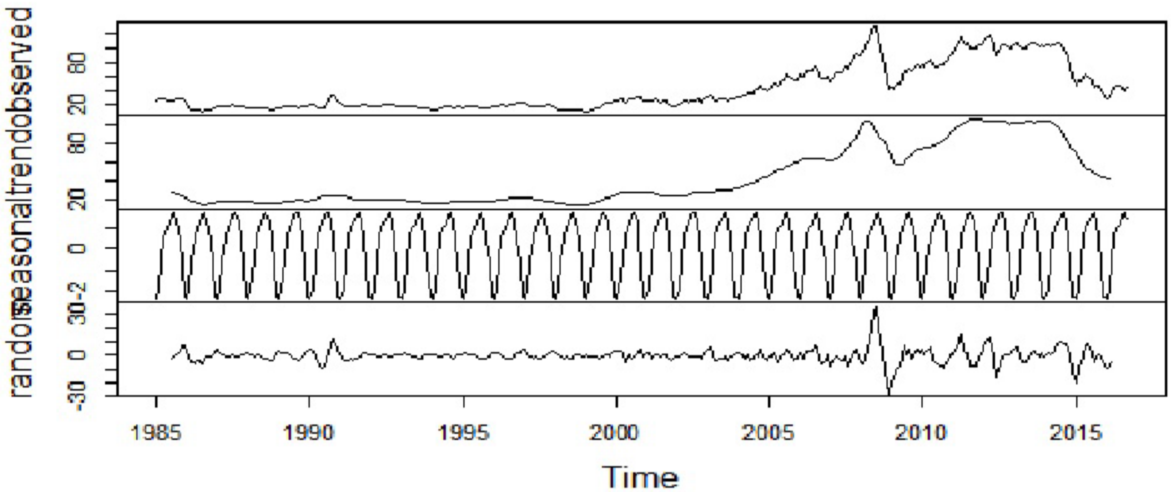


Figure 7: Time Series Decomposition Diagram.

<u>Hypothesis Test for First-Order Stationary</u>

To start the process of hypothesis testing for test of stationarity, a 95% confidence interval is computed as follows:

$$0 \pm (\frac{1}{\sqrt{t}})Z_{\frac{\alpha}{2}} = \pm 0.1318382182 \begin{cases} t = 380 \\ Z_{\frac{\alpha}{2}} = 2.57 \end{cases} \quad = \pm 0.13 \tag{15}$$

Next is the computation of autocorrelation coefficients through S-Plus program with the out-put bellow table:

Table 3: Autocorrelation Coefficients of First Few Lags of Crude Oil Prices

Lag	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ACF	1.0000	0.9913	0.9754	0.9562	0.9368	0.9187	0.9026	0.8897	0.8790	0.8702	0.8630	0.8554	0.8462	0.8355	0.8253	0.8163

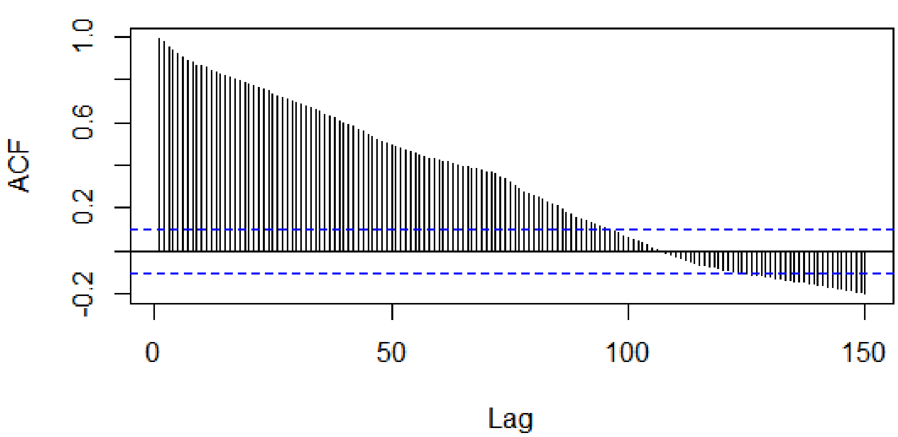


Figure 8: Correlelogram of the Monthly Crude Oil Prices.

Hypothesis Setting for First-Order Stationarity:
*H*0: The series is stationary.
*H*1: The series is not stationary.
At 95% confidence level.

Decision Rule for First-Order Stationarity: Reject *HO*: The series is stationary, if the first several lag value of autocorrelation coefficients are significant and gradually decreases until zero, accept *H*0: The series is stationary, otherwise.
At 95% confidence level.

Decision for First-Order Stationarity: I reject *H*0: The series is stationary, since the first several lag value of autocorrelation coefficients are significant and gradually decreases until zero.

Conclusion for First-Order Stationarity: At 95% significant level, the data do provide sufficient evidence (as seen in Table 3 and Figure 9) to conclude that the series is not stationary.

Hypothesis Test for Randomness

Hypothesis Setting for Randomness:
*H*0: The series is random.
*H*1: The series is not random.
At 95% confidence level.

Decision Rule for Randomness: Reject *HO*: The series is random, if Autocorrelation Coefficients Lag 1 is close to unity. The series is randomness, otherwise.
At 95% confidence level.

Decision for Randomness: I do reject *H*0: The series is randomness, since Autocorrelation Coefficients Lag 1 = 0.9913 and is close to unity.

Conclusion for Randomness: At 95% significant level, the data do provide sufficient evidence (as seen in Table 3 and Figure 9) to conclude that the series is not random.

Table 4: DISPLAY OF FIRST DIFFERENCE OF CRUDE OIL PRICES

	0.93	0.35	-0.07	-1.18	-0.31	0.28	0.50	0.39	0.77	0.89	-2.43
-4.42	-5.95	-3.61	-0.88	1.60	-1.54	-2.07	3.54	0.61	-0.02	0.46	0.97
2.66	-0.82	0.46	0.31	0.36	0.30	0.91	-0.74	-0.56	0.31	-0.76	-1.10
-0.27	-0.60	-1.04	1.56	-0.06	-0.87	-1.02	0.10	-1.35	-0.99	0.30	2.15
1.77	0.07	1.58	1.29	-1.17	-0.42	-0.13	-0.80	0.82	0.72	-0.03	0.99
1.22	-0.91	-1.56	-1.80	-0.11	-1.28	1.88	9.73	7.08	1.23	-3.31	-4.93
-3.80	-4.28	-0.32	0.28	0.23	-0.94	1.24	0.28	0.65	1.61	-1.15	-2.78
-0.11	0.13	-0.30	1.30	0.87	1.36	-0.70	-0.56	0.57	-0.15	-1.14	-0.97
-0.69	0.99	0.27	-0.06	-0.27	-0.80	-1.00	0.06	-0.63	0.64	-1.35	-1.73
0.81	-0.42	-0.06	1.46	1.28	0.80	0.81	-1.06	-0.85	0.35	0.72	-1.07
0.75	0.56	-0.09	1.42	-0.34	-1.10	-1.27	0.43	0.28	-0.59	0.64	1.11
-0.14	-0.10	1.77	1.32	-1.66	-0.56	1.00	0.63	1.95	1.29	-1.18	1.26
-0.22	-2.75	-1.12	-1.44	1.49	-1.45	0.43	0.34	-0.06	1.39	-0.96	-1.92
-2.17	**-0.89**	**-0.94**	**0.15**	**0.58**	**-1.49**	**0.24**	**-0.23**	**1.31**	**-0.54**	**-1.38**	**-1.47**
0.91	**-0.57**	**2.11**	**2.87**	**0.39**	**0.12**	**2.51**	**1.46**	**2.16**	**-0.18**	**2.03**	**0.79**
0.20	**1.94**	**0.34**	**-4.04**	**3.78**	**2.39**	**-1.46**	**1.25**	**2.67**	**-0.68**	**0.93**	**-7.05**
0.67	**1.29**	**-2.22**	**0.64**	**1.89**	**-0.58**	**-2.17**	**1.01**	**-0.78**	**-4.30**	**-2.04**	**-0.17**
0.63	**0.83**	**3.66**	**1.79**	**0.26**	**-1.20**	**1.26**	**1.03**	**1.50**	**-0.75**	**-2.74**	**3.10**
2.88	**2.11**	**-2.52**	**-4.87**	**0.57**	**1.85**	**0.68**	**1.09**	**-2.80**	**2.13**	**0.11**	**0.83**
1.45	**-0.08**	**2.35**	**0.04**	**3.92**	**-2.09**	**2.39**	**4.15**	**-0.43**	**5.22**	**-4.64**	**-3.14**
3.80	**1.67**	**6.37**	**-0.29**	**-2.83**	**6.08**	**2.48**	**5.50**	**-0.22**	**-3.46**	**-3.21**	**1.49**
5.89	**-2.65**	**1.22**	**7.07**	**0.61**	**-0.32**	**4.22**	**-0.70**	**-9.84**	**-4.02**	**0.18**	**2.87**
-7.60	**4.18**	**3.02**	**4.50**	**0.00**	**3.09**	**5.48**	**-3.54**	**6.78**	**5.24**	**9.12**	**-1.84**
1.39	**2.93**	**8.09**	**7.21**	**13.72**	**8.75**	**1.03**	**-17.98**	**-15.28**	**-26.60**	**-18.65**	**-12.51**
2.38	**-2.15**	**5.19**	**3.33**	**7.82**	**11.03**	**-4.48**	**6.98**	**-3.25**	**5.70**	**3.48**	**-2.68**
2.24	**-2.40**	**4.58**	**4.84**	**-8.60**	**-0.81**	**-0.21**	**1.36**	**0.23**	**5.61**	**2.81**	**5.54**
2.59	**5.07**	**10.92**	**7.67**	**-8.14**	**-2.33**	**2.03**	**-7.43**	**0.38**	**-0.91**	**5.44**	**-1.10**
2.63	**5.81**	**5.09**	**-4.04**	**-9.59**	**-13.43**	**6.02**	**8.53**	**1.04**	**-2.93**	**-2.22**	**0.00**
3.87	**2.62**	**-5.05**	**-3.76**	**0.50**	**0.39**	**5.47**	**2.85**	**0.72**	**-3.32**	**-2.88**	**2.91**
-3.24	**2.57**	**-0.78**	**0.90**	**0.79**	**2.64**	**-3.15**	**-5.17**	**-4.16**	**-9.76**	**-9.17**	**-16.41**
-13.10	**7.48**	**-1.10**	**3.59**	**5.08**	**-1.20**	**-6.87**	**-8.70**	**0.56**	**0.67**	**-3.83**	**-6.57**
-6.64	**1.13**	**6.29**	**3.41**	**5.23**	**1.71**	**-3.47**	**-4.22**	**6.21**			

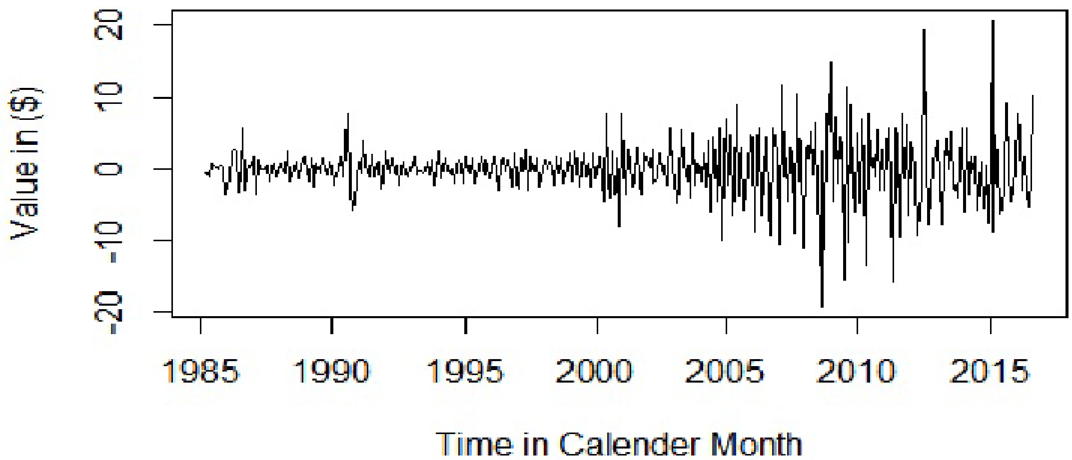


Figure 9: Time Plot of First Difference of the Monthly Crude Oil Prices.

It is noted that the autocorrelation coefficients of the first several lags are significantly different from zero for instance, autocorrelation coefficient of lag 1, lag 2, lag 3, · · · , lag 20 = 0.9913, 0.9913, 0.9754, · · · , 0.7820, respectively which individually less than the ±0.13 (the confidence bound for autocorrelation) and the autocorrelation coefficients gradually regressed towards zero rather than dropping exponentially. This shows that a trend exists in the series (that is the series is non stationary) or there is presence of serial correlation in the series.

The ACF for the series decays very slowly indicating that it is non-stationary. Non-stationary stochastic processes tend to generate series whose estimated autocorrelation function fail to die out rapidly; that is, the estimated autocorrelations for non-stationary processes tend to persist for a large number of lags. Persistently large values of them indicate that the time series is non-stationary, or serially correlated and that transformation of the series through at least one difference is needed.

Since the theoretical autocorrelations and partial autocorrelations are only independent of time for stationary processes, it is necessary to difference the original series until it can be assumed to be a realization of a stationary process. The differencing is achieved through S-Plus program and the output is presented in the Table 4.

Table 5: Autocorrelelogram for the First Difference of the Crude Oil Prices.

Lag	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
ACF	1.0000	0.9913	0.9754	0.9562	0.9368	0.9187	0.9026	0.8897	0.8790	0.8702	0.8630	0.8554	0.8462	0.8355	0.8253	0.8163	0.8081	0.8010	0.7945	0.7885

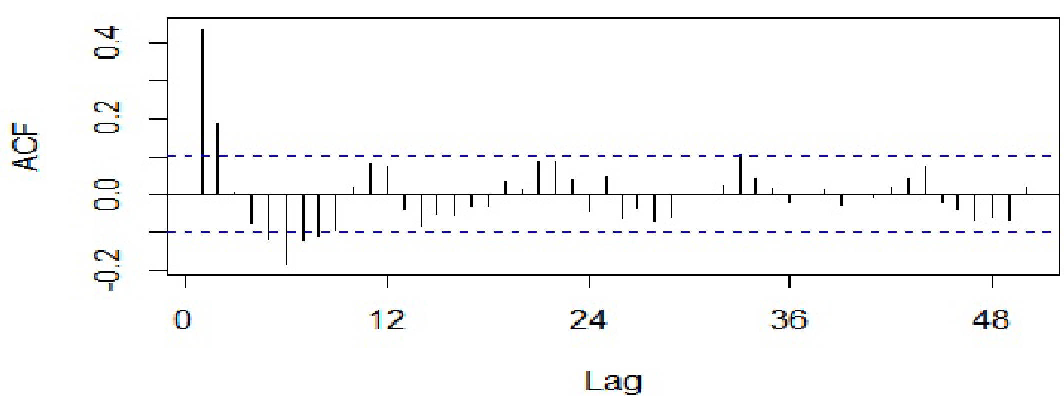


Figure 10: Autocorrelogram of First Difference of the Monthly Crude Oil Prices,

Hypothesis Test for Second-Order Stationary

Having achieved the first differenced series; a test for stationary of the series is necessary and will be conducted again to check if stationarity has been reached. The first step on this is to plot the autocorrelogram of the first differenced series to view how the first differenced time series behaves. The autocorrelogram obtained from Table 5 through S-Plus is presented in Figure 10.

Hypothesis Setting for Second-Order Stationarity

Hypothesis Setting for First Order:
*H*0: The series is stationary.
*H*1: The series is not stationary.
At 95% conidence level.

Decision Rule for Second-Order Stationarity: Reject *HO*: The series is stationary, if the first several lag value of autocorrelation coefficients are significant and gradually decreases until zero, accept *H*0: The series is stationary, otherwise.
At 95% confidence level

Decision for Second-Order Stationarity: I reject *H*0: The series is stationary, since the first several lag value of autocorrelation coefficients oscillate around zero.

Conclusion for Second-Order Stationarity: At 95% significant level, the data do provide sufficient evidence (as seen in Table 3 and Figure 9) to conclude that the series is stationary. Ordinarily, one would have expect that decision rule for test of seasonality be made on autocorrelation coefficient the restriction to this is that strength of serial correlation will overpower the autocorrelogram in such a way that spike in it will hardly be noticed. at this stage of the analysis that the trend have been satisfactorily removed, I can now test for seasonality as follows:

Hypothesis Test For Seasonality

Hypothesis for Seasonality:
*H*0: The series is seasonal.
*H*1: The series is not seasonal.
At 95% confidence level.

Decision Rule for Seasonality: Reject *HO*: The series is seasonal, if a spike in Autocorrelation Coefficients after differencing at a regular lag interval. The series is seasonal, otherwise. At 95% confidence level.

Decision for Seasonality: I do not reject *H*0: The series is seasonal, since there is a noticeable spike in Autocorrelation Coefficients 6 has interval as can be seen in Table 5, and Figure 8.

Conclusion for Seasonality: At 95% significant level, the data do provide sufficient evidence (as seen in Table 5 and Figure 8) to conclude that the series is not seasonal. Haven achieved detrend, and deseasonalised the series; model identification and parameter estimation is now discussed in the sections below.

Model Strategy

The steps demonstrated in Figure 2, for easy step-taking in this thesis, the model strategy involves model identification and parameter estimation as shown in the subsections below.

Model Identification

The model for the monthly crude oil prices is achieved by estimating the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) of the differenced series. The selection of a time series model is frequently accomplished by matching estimated autocorrelations coefficients with the theoretical autocorrelation. The matching of the first 50 estimated autocorrelations and partial autocorrelations of the underlying stochastic processes suggested that the series were stationary, with the ACF, PACF. The estimated ACF and PACF are as shown in Table 6 (6a and 6b, respectively).

Table 6: ACF and PACF Table.

Lag	0	1	2	3	4
ACF	1.00	0.44	0.19	0.01	-0.08

(a) ACF

Lag	0	1	2	3	4
PACF	-0.27	-0.14	-0.16	-0.15	-0.08

(b) PACF

With the matching of the *ACF* with the *PACF*, the model is identified as an ARIMA (2*,* 1*,* 2) model. Plot of the Autocorrelogram and Partial Autocorrelogram is shown in Figure 11 (11a and 11b respectively) and their corresponding tables in Table 6 (6a and 6b, respectively). P

lot of the Autocorrelation Function and Partial Autocorrelation Function greatly assisted in the understanding of the model. From the ACF and PACF, the functions both had a sharp cut off(indicating no additional non-seasonal terms to be included in the model) after lag 2 after the first difference suggest that the model may include in its formation ARIMA (2*,* 1*,* 2).

In addition I had tested the seasonality of the series to be positive, thus the need to include seasonal component in the series. A seasonal ARIMA model is formed by including additional seasonal terms in the ARIMA model as demonstrated in the Equation 16 below:

$$\text{ARIMA} \quad \underbrace{(p, d, q)}_{\text{Non-Seasonal Part}} \quad \underbrace{(P, D, Q)}_{\text{Seasonal Part}} [12] \qquad \textbf{(16)}$$

Parameter Estimation

The likely estimates of parameter that will be needed to form appropriate model is summarized in Table 7. The estimates are obtained by **S-plus** program with the method on Equation 10, 11, 12 and 13.

Table 7: PARAMETER/ESTIMATE TABLE

Parameter	ϕ_{11}	ϕ_{22}
Estimate	1.48	-0.59

(a) AR

Parameter	θ_{11}	θ_{22}
Estimate	-1.04	0.14

(b) MA

Parameter	α_{11}	α_{22}
Estimate	0.12	-0.16

(c) SAR

I recommend the following models:

$$\text{ARIMA}(2, 1, 2) \qquad \textbf{(17)}$$

$$\text{ARIMA}(2, 1, 2)(1, 0, 0)[12] \qquad \textbf{(18)}$$

$$\text{ARIMA}(2, 1, 2)(2, 0, 0)[12] \qquad \textbf{(19)}$$

and,

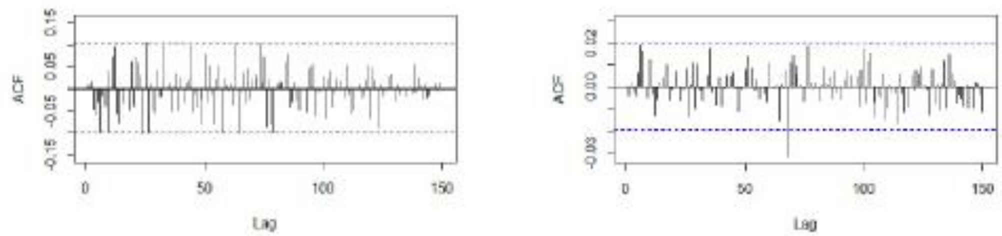
$$\text{ARIMA}(2, 1, 2)(3, 0, 0)[12] \qquad \textbf{(20)}$$

BOOTSTRAP METHOD FOR DIAGNOSTIC CHECK

Almost all model validity test are based on the examination of the residuals,

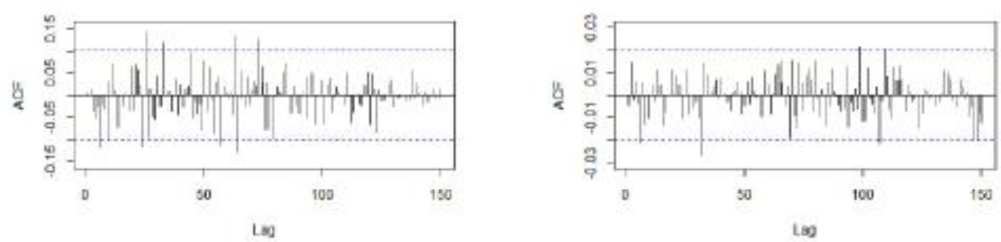
$$\varepsilon_t = X_t - \hat{X}_t \text{ where } \hat{X}_t$$

is the fitted value, or some functions of the residuals. In some cases where some financial time series defile the use of ACF of the residuals to determine the validity of a time series model, thus the search for a more robust test for model validity. The use of ACF of bootstrap residuals in comparison with ACF of the residual. I therefore present the Autocorrelogram residuals with their corresponding Autocorrelogram bootstrap residuals of models presented in Equation 17, Equation 18 Equation 19 and Equation 20 are presented in the figures below:



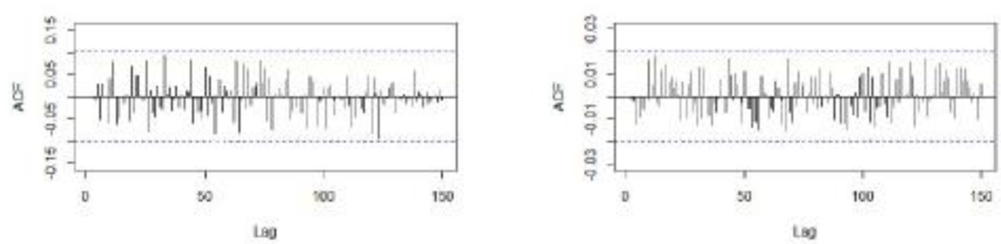
(a) ACF ERROR (b) ACF BOOT ERROR

Figure 12: VALIDATION FOR ARIMA(2, 1, 2)



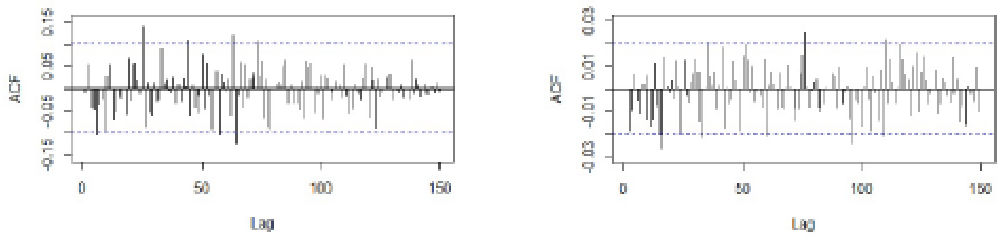
(a) ACF ERROR (b) ACF BOOT ERROR

Figure 13: VALIDATION FOR ARIMA(2, 1, 2)(1, 0, 0)[12]



(a) ACF ERROR (b) ACF BOOT ERROR

Figure 14: VALIDATION FOR ARIMA(2, 1, 2)(2, 0, 0)[12]



(a) ACF ERROR

(b) ACF BOOT ERROR

Figure 15: VALIDATION FOR **ARIMA**(2, 1, 2)(3, 0, 0)[12]

Table 8: FORECAST-ACCURACY METRIC TABLE

MODEL	*ARIMA(2,1,2)*	*ARIMA(2,1,2)(1,0,0)[12]*	*ARIMA(2,1,2)(2,0,0)[12]*	*ARIMA(2,1,2)(3,0,0)[12]*
AIC	**2086.870**	**2085.060**	**2069.690**	**2080.190**
Sigma^2	**13.810**	**13.710**	**13.080**	**13.430**
Log Likelihood	**-1038.430**	**-1036.530**	**-1027.840**	**-1032.090**
ME	**0.023**	**0.018**	**0.073**	**0.026**
RMSE	**3.691**	**3.673**	**3.588**	**3.626**
BIC	**2106.570**	**2108.700**	**2097.280**	**2111.710**
MAE	**2.494**	**2.473**	**2.471**	**2.490**
MASE	**0.728**	**0.722**	**0.712**	**0.721**

Table 9: POINT AND INTERVAL FORECAST TABLE

	Point Foreast	***Lo 80***	***Hi 80***	***Lo 95***	***Hi 95***
Oct-16	50.33	45.69	54.96	43.24	57.42
Nov-16	52.81	44.66	60.96	40.34	65.27
Dec-16	55.35	44.14	66.55	38.21	72.49
Jan-17	56.85	43.11	70.59	35.84	77.86
Feb-17	55.75	39.98	71.52	31.63	79.86
Mar-17	56.42	39.04	73.79	29.84	82.99
Apr-17	55.93	37.29	74.57	27.42	84.43
May-17	55.41	35.77	75.05	25.37	85.45
Jun-17	55.48	35.02	75.94	24.20	86.76
Jul-17	55.86	34.74	76.99	23.55	88.18
Aug-17	56.50	34.79	78.21	23.30	89.70
Sep-17	56.99	34.78	79.21	23.02	90.97
Oct-17	57.27	34.47	80.06	22.40	92.13
Nov-17	58.08	34.69	81.48	22.30	93.87
Dec-17	59.37	35.37	83.37	22.67	96.07

To be double sure, I present a more conventional method of testing accuracy of a model: A forecast-accuracy metric table to measure accuracies of suggested models presented in Equation 17, Equation 18, Equation 19 and Equation 20. The accuracy-metric Table is presented on Table 8. Log Likelihood and Mean Error (ME) places a greatest metric value on a best forecast model to be selected, while others like The Akaike information criterion (AIC), the Bayesian information criterion (BIC) and the rest place lowest metric value on best model to be selected.

Obviously, ARIMA (2, 1, 2)(2, 0, 0)[12] stands out to be the best model which is in agreement with my Bootstrap Method presented on Figure 12, Figure 13, Figure 14 and Figure 15 above.

$$X_t = \phi_{11}X_{t-1} + \phi_{22}X_{t-2} + \theta_{11}\varepsilon_{t-1} + \theta_{22}\varepsilon_{t-2} + \alpha_{11}X_{t-1} + \alpha_{22}X_{t-2} + \varepsilon_t \quad \textbf{(21)}$$

Thus, Equation 21 becomes the best, parsimonious, consistent and the most representative model for the series. Also, from the Table of Parameter Estimate in Table 7 the parameters are being fed to Equation 21 which is given as:

$$X_t = 1.48X_{t-1} - 0.59X_{t-2} - 1.04\varepsilon_{t-1} + 0.14\varepsilon_{t-2} + 0.12X_{t-1} - 0.16X_{t-2} + \varepsilon_t \quad \textbf{(22)}$$

With the Equation 21 and Equation 22 the following are forecast the month of October 2016, November 2016, December 2016 and January 2017 and presented in Table 9. The forecast is not just on point forecast alone, but also on 80% and 90% interval as can be seen on the same table. To make it eye-friendly, a pictorial diagram of forecast plot is presented below:

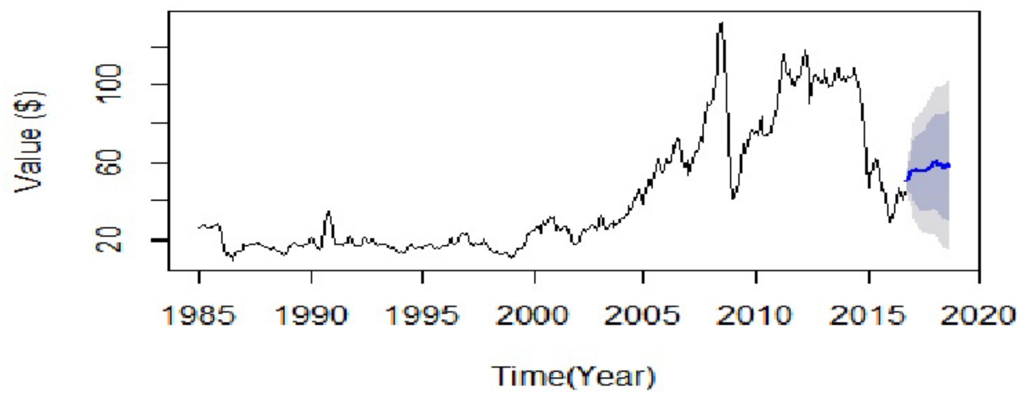
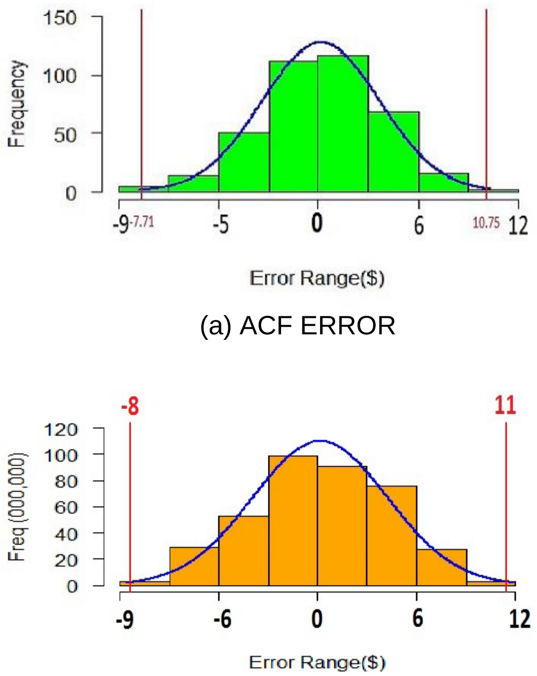


Figure 16: Forecast Plot of ARIMA (2, 1, 2)(2, 0, 0)[12].

Figure 16 depicts Table 9 and an In-Series forecast of the series, while Table 10 present the In-Series errors from January 1985 to September 2016.

Table 10: IN-FORECAST ERROR TABLE

Mth/Yr	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
1985	0.03	0.80	-0.08	-0.22	-1.06	0.32	0.51	0.39	0.15	0.58	0.55	-2.77
1986	-3.26	-3.86	-0.87	0.60	1.61	-2.93	-2.12	3.85	-1.47	-0.95	0.07	0.84
1987	2.41	-1.60	0.78	0.05	-0.12	0.57	1.15	-1.50	0.03	0.88	-0.74	-1.23
1988	-0.48	-0.80	-1.05	2.03	-0.63	-1.48	-1.12	1.27	-1.67	-0.74	0.75	2.03
1989	0.88	-1.22	1.67	0.42	-1.56	0.37	0.48	-0.84	1.42	0.60	-0.55	0.71
1990	0.83	-1.35	-1.36	-0.69	0.91	-1.36	2.24	8.97	2.19	-2.24	-3.39	-2.51
1991	-0.73	-1.89	2.28	0.71	-0.47	-1.12	1.20	-1.59	0.20	1.51	-1.60	-1.85
1992	-0.47	0.95	0.11	-0.39	-0.35	-1.02	-0.28	0.44	-0.78	1.03	-1.92	-1.58
1993	-0.47	0.95	0.11	-0.39	-0.35	-1.02	-0.28	0.44	-0.78	1.03	-1.92	-1.58
1994	1.72	-1.02	-0.15	1.56	0.53	0.27	0.23	-1.52	-0.10	0.69	0.66	-1.32
1995	1.01	0.58	-0.39	1.26	-1.03	-1.04	-0.82	1.34	0.06	-0.74	0.52	0.79
1996	-0.59	-0.23	1.89	0.61	-2.05	0.40	1.59	-0.07	1.63	0.71	-1.68	1.65
1997	-0.25	-2.44	0.04	-0.56	2.31	-2.32	0.66	0.25	-0.50	1.15	-1.30	-1.66
1998	-1.32	0.38	-0.37	0.54	-0.34	-1.73	0.76	-0.48	1.48	-1.41	-1.42	-0.47
1999	1.54	-1.53	2.25	1.60	-0.79	-0.25	2.62	0.55	1.41	-0.56	2.28	0.14
2000	-0.21	2.39	-0.47	-3.99	6.19	0.87	-2.59	2.04	2.53	-1.67	1.04	-7.21
2001	4.39	1.01	-2.45	2.32	0.78	-1.66	-1.30	1.85	-1.22	-3.99	0.04	1.62
2002	0.00	0.36	3.24	-0.93	-0.06	-0.99	1.76	0.71	1.62	-1.09	-2.02	3.36
2003	2.30	1.01	-4.15	-3.23	3.51	1.82	-0.73	0.95	-3.57	2.98	-0.42	0.32
2004	1.02	-0.72	3.34	-0.43	3.54	-4.02	3.84	3.53	-1.46	5.10	-6.71	-0.03
2005	6.07	0.65	4.96	-3.28	-2.38	8.73	-0.02	4.44	-2.33	-2.85	-0.06	4.03
2006	5.35	-5.09	2.38	7.37	-1.19	-1.75	5.59	-1.86	-8.99	2.38	1.88	2.44
2007	-8.80	8.03	1.82	1.46	-2.35	4.32	3.81	-4.80	9.42	2.33	6.69	-5.06
2008	4.99	1.96	8.17	5.41	11.42	3.50	-1.16	-16.01	-7.25	-17.39	-5.75	-2.51
2009	5.25	-4.52	2.65	-1.52	2.29	5.73	-10.03	9.08	-4.51	9.35	2.84	-4.70
2010	3.27	-1.75	6.74	4.06	-9.30	3.27	1.94	-1.87	-0.31	2.00	-0.59	4.25
2011	1.08	3.86	9.16	2.94	-8.84	3.12	3.10	-6.02	3.63	0.31	6.41	-4.35
2012	3.83	4.15	2.68	-5.83	-7.51	-7.95	12.39	7.26	-4.04	-3.01	-1.52	2.29
2013	3.65	0.90	-5.23	-0.17	1.68	1.62	4.32	-1.63	0.35	-3.11	-0.04	3.92
2014	-4.32	4.64	-0.61	0.40	-0.98	0.99	-2.89	-2.77	-2.04	-7.95	-4.96	-12.96
2015	-5.38	11.91	-7.04	1.08	1.79	-5.57	-6.54	-6.15	3.71	-0.21	-5.17	-4.00
2016	-4.72	2.13	4.93	-1.02	2.07	-0.66	-4.70	-2.78	7.32			



(a) ACF ERROR

(b) ACF BOOT ERROR

Figure 17: Validation for ARIMA (2, 1, 2).

Figure 17 further shows the normality stand of the error term especially to readers who are not so friendly with the Box-Jenkins terminology, and also conformity of confidence interval of both error term and the bootstrap error term. Having used the bootstrap method as a validation tool for Box-Jenkins model which provides accurate forecast for the series, the next step is the summary of findings and conclusion.

CONCLUSION

From the investigation performed in this study, the following conclusions can be drawn:

i. Box-Jenkins methodology was used for model identification which indicated that ARIMA (2, 1, 2)(2, 0, 0)[12] is the best fit for the monthly crude oil prices (January 1985 to September 2016) among set of most fitting model presented in Equation: 17, 18, 19 and 20 in no order.

ii. A Bootstrap Method for Model Validation was used to check the validity of the model also affirmed by forecast-accuracy metrics like Akaike information criterion (AIC), Bayesian information criterion (BIC), among others, to be the best model as can be seen in Table 8 and Figure 17.

iii. The forecast for the price of crude oil per barrel (Brent being the benchmark) for October 2016, November 2016, December 2016 and January 2017 are $50.33, $52.81, $55.35, and $56.85, respectively, thus the crude oil prices are forecast to increase up to the December, 2017 to the ton of $59.37 as a point forecast and ($35.37, $96.07) as 95% interval forecast as can be seen in Table 9 and Figure 16.

REFERENCES

1. Berkowitz, J. and L. Kilian. 2000. "Recent Developments in Bootstrapping Time Series". *Econometric Reviews*. 19(1):1–48.

2. Box, G.E., G.M. Jenkins, G.C. Reinsel, and G.M. Ljung. 2015. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons: New York, NY.

3. Brockwell, P.J. and R.A. Davis. 2016. *Introduction to Time Series and forecasting*. Springer: Berlin, Germany.

4. Bühlmann, P. 2002. "Bootstraps for Time Series". *Statistical Science*. 52–72.

5. Carlstein, E., K-A Do, P. Hall, T. Hesterberg, and H.R. Künsch. 1998. "Matched-Block Bootstrap for Dependent Data". *Bernoulli*. 305–328.

6. Chatfield, C. 2016. *The Analysis of Time Series: An Introduction*. CRC Press: Boca Raton, FL.

7. Davidson, R. and J.G. MacKinnon. 2006. "Bootstrap Methods in Econometrics".

8. Duchesne, P. and B. Rémillard. 2005. "Statistical Modeling and Analysis for Complex Data Problems, Volume 1". Springer Science & Business Media: London, UK.

9. Ediger, V.S. and S. Akar. 2007. "ARIMA Forecasting of Primary Energy Demand by Fuel in Turkey". *Energy Policy*. 35(3):1701–1708.

10. Efron, B. and R.J. Tibshirani. 1994. *An introduction to the Bootstrap*. CRC Press: Boca Raton, FL.

11. Gregory, K.B., S.N. Lahiri, and D.J. Nordman. 2015. “A Smooth Block Bootstrap for Statistical Functionals and Time Series”. *Journal of Time Series Analysis*. 36(3):442–461.

12. Hall, P., J.L. Horowitz, and B.-Y. Jing. 1995. “On Blocking Rules for the Bootstrap with Dependent Data”. *Biometrika*. 82(3):561–574.

13. Hasan, J. 2017. “OPEC Basket Price”. (accessed February 3, 2017).

14. Hwarng, H.B. and H. Ang. 2001. “A Simple Neural Network for ARMA (p, q) Time Series”. *Omega*. 29(4):319–333.

15. Kobayashi, M. 2014. Resources for Studying Statistical Analysis of Biomedical Data and R”. In: *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics*.183–195. Springer: Berlin, Germany.

16. Kreiss, J.-P., E. Paparoditis, et al. 2003a. ”Autoregressive-Aided Periodogram Bootstrap for Timeseries”. *The Annals of Statistics*. 31(6):1923–1955.

17. Kreiss, J.-P., Paparoditis, E., et al. 2003b. ”Autoregressive-Aided Periodogram Bootstrap for Timeseries”. *The Annals of Statistics*. 31(6):1923–1955.

18. Kumar, D. 2016. “Weighted Bootstrap Approach for the Variance Ratio Tests: A Test of Market Efficiency”. *Theoretical Economics Letters*. 6(03):426.

19. Lahiri, S.N. 1999a. “Theoretical Comparisons of Block Bootstrap Methods”. *Annals of Statistics*. 386–404.

20. Lahiri, S.N. 1999b. “Theoretical Comparisons of Block Bootstrap Methods”. *Annals of Statistics*. 386–404.

21. Lahiri, S.N. 2013. *Resampling Methods for Dependent Data*. Springer Science & Business Media: London, UK.

22. Léger, C. and J.P. Romano. 1990. ”Bootstrap Choice of Tuning Parameters”. *Annals of the Institute of Statistical Mathematics*, 42(4):709–735.

23. Liu, R.Y. and K. Singh. 1992. “Moving Blocks Jackknife and Bootstrap Capture Weak Dependence”. *Exploring the Limits of Bootstrap*. 225:248.

24. Politis, D.N. and J.P. Romano. 1994. “Large Sample Confidence Regions Based on Subsamples under Minimal Assumptions”. *The Annals of Statistics*. 2031–2050.

SUGGESTED CITATION

Adejumo, A.O. and J.Daniel. 2017. “A Bootstrap Method for Box-Jenkins Models with Application on Brent Crude Oil Prices per Barrel”. *Pacific Journal of Science and Technology*. 18(2):88-109.

Pacific Journal of Science and Technology