

International Journal of Statistics and Applied Mathematics



ISSN: 2456-1452
Maths 2020; 5(4): 115-130
© 2020 Stats & Maths
www.mathsjournal.com
Received: 19-05-2020
Accepted: 21-06-2020

Nkemnole E Bridget
Department of Mathematics,
University of Lagos, Nigeria

Wulu J Trarso
Department of Mathematics &
Statistics, University of
Maryland University College,
United States

Comparison between Weibull distribution fitting and limit probability of a Markov chain of import commodities market: Oil price series

Nkemnole E Bridget and Wulu J Trarso

Abstract

The constant concern in commodities market, particularly in oil price, has necessitated accurate models to aid in the generation of relatively good synthetic oil price data. Oil prices are subject to high volatility and fluctuations; and this is a nagging challenge, as it becomes difficult to deduce the changing trends of oil prices in the middle - and long - terms; and, hence, this affects the prediction of oil price levels in the short-term. Basically, traditional methodologies in the estimation of oil volatility have mostly been attentive to the problem of predicting oil price levels and seeking assessment of the best estimates. This research employs Weibull distribution and limit probability of a Markov Chain for a statistical modelling of Brent crude oil prices. Basically, we obtained the pattern and fluctuations of Brent crude oil prices and confirmed that there is an existence of a Markov chain in the observed oil price series. We then used the limit probability of a Markov Chain to deduce the changing trends of the oil prices. Subsequently, we carried out a probability distribution analysis and discovered that the oil price series conform to a Weibull distribution when compared to other probability distributions (e.g., Gamma, Lognormal, and Gumbel Max). Empirical analysis of the Weibull distribution and the limit probability of a Markov Chain of oil price series reveal that there are changing trends of oil prices from the short-term to the middle- and long-terms, respectively. The proposed models does not only indicate specific trend changes in a variety of oil price states over different terms, but also, establishes a consideration for assessing a trend deduction model of fluctuating oil prices.

Keywords: Brent oil price, volatility, limit probability of a Markov chain, weibull distribution, short-term, middle- term, long-term forecast

1. Introduction

Crude Oil is a non-renewable fossil fuel which is distributed randomly all over in the ground, Tuo and Yanbing (2011) [30]. Essentially, the industrialized world depends on crude oil as a chief source of energy supply, both as a major component in many manufacturing processes and for transportation purposes. Without doubt, crude oil represents the highest percentage of the world's energy consumption compared to other energy resources. The price of crude oil is one of the world's most influential global economic indicators and it is precisely observed by policy-makers, producers, consumers and financial market partakers.

It is also open to fluctuations and constant changes mostly due to the interaction between supply and demand forces on the international commodity markets. This instability is up to approximately 25% per annum according to He *et al.* (2012) [15]; and it exerts considerable influence on economic situation of different countries.

Modeling of crude oil prices with exactitude and high performance has been one of the most valuable academic topics explored to solve the problem of fluctuating oil prices. A number of approaches have been utilized to this effect, leading to diverse understandings of market with strong significance to global economic growth. This work aims to advance a method to deduce the trends of fluctuation of oil prices so as to assist investigators to access optimal model to solve this problem. The probability distribution and Markov Chain time dependence, while assuming the presence of time homogeneity without testing is the main tool for this task.

This research focused on modeling of Brent crude oil prices with Weibull distribution and limit probability of a Markov Chain to deduce the changing trends of the oil prices.

Corresponding Author:
Nkemnole E Bridget
Department of Mathematics,
University of Lagos, Nigeria

The changing trends of the oil prices are classified into three levels: Short-term, Medium-term, and Long-term. It comprises five sections, and is structured as follows: Section 1 introduces the work; Section 2 presents a brief review of relevant literature; Section 3 defines the working models with results; Section 4 validates the models applied to a real dataset; Section 5 offers some concluding remarks.

2. Historical Perspectives of Relevant Models

The significance of oil and its influence on the local and international economy has a subject of many research works from different disciplines, leading to numerous perceptions on the same subject.

Kosobud and Stokes (1978) ^[20] verified the pattern of “best market share rules” through the use of a Markov probability model. The work concludes that after the Organization of the Petroleum Exporting Countries (OPEC) has taken shape, the probability of conflicts among suppliers has reduced whereas such a probability among consumers has increased. Abazi (2003) recommended a non-linear stochastic mean and volatility model for crude oil prices dynamics and applied Markov chain Monte Carlo algorithms to estimate the parameters of the model. Holmes and Wang (2003) ^[17] applied a Markov switching model in studying the influence of soaring oil prices on the growth of British GDP, and arrived at the conclusion that during the increasing period of business cycle the soaring oil prices and growth of GDP are asymmetric to various extents.

Dafas (2004) ^[7] suggested a mean-reverting Markov switching jump diffusion model to characterize the stochastic behavior of the crude oil spot prices. Fong and See (2002) looked at the effect of volatility in daily returns on crude oil futures using generalized autoregressive conditional heteroskedasticity (GARCH), regime switching (RS), RSARCH-t and RSGARCH-t models. It discovered that RS models are useful both to financial historians interested in studying the factors behind the evolution of volatility and to oil futures traders interested in using the model to extract short-term forecasts of conditional volatility.

Song (2005) ^[29] conducted the prediction of oil prices by one-state transition matrix without testing the existence of a Markov Chain and calculating the convergence value of transition state matrix. As a result, the outcome is far from the reality, thus concluding that the Markov method could not predict the evolution of the event perfectly. Wei *et al.* (2006) ^[35] divided the time series of oil prices into three states as increasing by a large range, increasing by a small range and decreasing by a large range states. The work established the duration of each state and concluded that the Markov Chain model is superior to an auto-regression model.

Vo, M.T. (2009) ^[31] studied a stochastically fluctuating regime of oil market by a Markov transition model to catch the factors which influence oil market, and concluded that the fluctuation of oil prices is consistent. Larsson and Nossman (2010) ^[22] scrutinized the empirical performance of affine jump diffusion models with stochastic volatility in a time series study of daily West Texas Intermediate (WTI) spot prices. He compares four different models and estimates them using the Markov Chain Monte Carlo method and found that the stochastic volatility model including jumps in both prices and volatility is strong and the model clearly outperforms the others. Haiyan and Zhongxiang (2010) applied Markov chain and Lognormal distribution to model and deduce the trend of fluctuating prices of OPEC basket of crude oils. Following a price-state classification and state transition analysis of changing oil prices from January 2004 to April 2010, the work verifies that the observed crude oil price series follow a Markov Chain. Kordnoori *et al.* (2013) applied the Markov Chain and Gumbel Max distribution to model the prices of Brent crude oils from January 2004 to July 2011 and concluded that the Brent oil prices follow a Markov chain.

Balcilar *et al.* (2013) used the Markov regime switching vector error correlation (MS-VEC) model and a monthly data set from 1859 to 2013 to examine the relationship between US crude oil and stock market prices. They estimate a two regime model that divides the sample into high-and low-volatility regimes based on the variance-covariance matrix of the oil and stock prices. The results revealed that the high-volatility regime occurs when the economy experiences a recession.

Elder *et al.*, (2014) examined price discovery in the prominent market price benchmarks crude oil West Texas Intermediate (WTI) and crude oil Brent and established that no evidence that the dominant role of crude oil WTI in price discovery was diminished by the price spread between WTI crude oil and Brent crude oil that emerged in 2008.

Zlatcu *et al.* (2015) studied the fuel markets of Romanian, Germany, France, Poland and the Czech Republic, examining the price volatility and the response of retail fuel prices to changes in international crude oil prices by estimating univariate and multivariate GARCH models, as well as applying a momentum threshold autoregressive (MTAR) co-integration model. Their results show that the conditional volatility reached its peak during the outbreak of the financial crisis, which offer support for the asymmetric transmission hypothesis. The cointegration analysis detects some signs of asymmetric adjustments, supporting the hypothesis of a faster adjustment above the threshold.

Zhang and Zhang (2015) singled out a Markov switching model (MS(3)-AR(2)) with three regimes for the samples of Brent and WTI crude oil prices before and after the 2008 financial crisis, and discovered three regimes for both crude oil markets before and after the crisis.

Bernard *et al.* (2015) modeled oil futures prices using weekly and monthly data and maturities of one to four months. Their results show that forecast performances improve with longer date-to-maturity futures, suggesting that the role of the convenience yield is greater when physical oil inventories are held for longer durations.

Baum and Zerilli (2016) explored the volatility of crude oil price futures returns from October 2001 to December 2012. The results of the model applied to the intraday data indicate that stochastic volatility models are effective in fitting the volatility of oil price futures returns.

All these aforementioned models have one thing in common; and that is, they used Markov chains to generate good synthetic oil price data. Their successes notwithstanding, the need for a better model is still there. In the attempt to close this gap, this research employs Weibull distribution and limit probability of a Markov Chain for a statistical modeling of Brent crude oil prices.

3. Methods

3.1 Markov Chains Model

A Markov chain is a sequence of random variables X_1, X_2, X_3, \dots , with the Markov property namely that, the conditional probability of any future event, given any past event and the present state, is independent of the past event and depends only on the present state.

Formally,

$$P[X_{t+1} = x | X_t = x_t, \dots, X_1 = x_1] = P[X_{t+1} = x | X_t = x_t]$$

The set of possible values of x is called the state space of the chain.

3.1.1 Transition probabilities

In a Markov chain, the conditional probabilities

$$P\{X_{t+1} = j | X_t = i\}$$

are called *transition probabilities*. In a *time homogeneous Markov Chain*, the transition probabilities are independent of the time index t :

$$P\{X_{t+1} = j | X_t = i\} = P\{X_1 = j | X_0 = i\} \text{ for all } t = 1, 2, \dots$$

Stationary transition probabilities show that transition probabilities do not change over time. Markov Chains can also have *n-step transition probabilities*, which is the conditional probability that the process will be in state j after n -steps provided that it starts in state i at time t .

n-step transition probabilities are defined as the conditional probability

$$P\{X_{t+n} = j | X_t = i\} = P\{X_n = j | X_0 = i\} \text{ for all } t = 0, 1, \dots$$

3.2 Evaluation of Markov Chain Tests

In this section, to ascertain the Markovian assumption that current state depends on that of the previous state, we validate the use of a Markov Chain. To achieve this, the conventional test for independence via chi-square statistic is used.

For testing the goodness of fit of the Markov model, we test the hypothesis:

H_0 : States are independent

H_1 : States are not independent

The Chi-Squared test statistic is given by

$$\chi^2 = \sum_{i,j}^N \frac{(n_{i,j} - E_{i,j})^2}{E_{i,j}} \sim \chi^2(i-1)(j-1), \alpha$$

Where $E_{i,j}$ represents the expected number of states computed using the formula: $\frac{n_i n_j}{n_{..}}$, with n_j representing j th column state total. That is the state that the Markov chain is currently in. n_i is the i th row state total. The row state represents the state that it is transitioning to and $n_{..}$ is the overall state total.

If $\sum_{i,j}^N \frac{(n_{i,j} - E_{i,j})^2}{E_{i,j}}$ is greater than $\chi^2(i-1)(j-1), \alpha$, we reject the null hypothesis that states are independent. Otherwise, we fail to reject the null hypothesis.

3.3 Limit of the transition probability matrix

A transition probability matrix is a square matrix describing the probabilities of moving from one state to another in a dynamic system. A stochastic matrix P is ergodic if $\lim_{n \rightarrow \infty} P^n$ exists, that is if each $p_{ij}^{(n)}$ has a limit as $n \rightarrow \infty$. If we denote the limit matrix, necessarily a stochastic matrix, by L . The components of $X^{(\infty)}$, defined by the equation $X^{(\infty)} = X^{(0)}L$ are the limiting state distributions and signify the approximate proportions of objects in the various states of a Markov chain after a large number of time periods.

3.4 Probability distributions of changing trends of oil prices

This section presents different probabilistic models as well as maximum likelihood method of estimating the unknown parameters of the distributions.

3.4.1 Gamma Distribution

The Gamma distribution is a two-parameter distribution and owing to its moderate skewness, is a useful model in many areas of statistics when the normal distribution is not appropriate.

3.4.1.1 Estimation of the Parameters of Gamma Distribution

The probability density function of the Gamma distribution is given as

$$f(x: \alpha, \theta) = \frac{x^{\alpha-1}}{\Gamma(\alpha)\theta^\alpha} e^{-\frac{x}{\theta}}, > 0, \text{ and } \theta, \alpha > 0,$$

where $\theta > 0$ is the shape parameter and $\alpha > 0$ is the scale parameter. We seek the conditions for the maximum likelihood estimates (MLE) of (θ, α) .

The likelihood function for the Gamma distribution is obtained as follows:

$$L(\alpha, \theta|x) = f_n(x|\alpha, \theta) = \frac{1}{\Gamma^n(\alpha) \cdot \theta^{n\alpha}} \left(\prod_{i=1}^n x_i \right)^{\alpha-1} e^{-\sum_{i=1}^n \frac{x_i}{\theta}}$$

Taking the natural logarithm, we obtain;

$$\ell(\alpha, \theta) = -n \ln(\Gamma(\alpha)) - n\alpha \ln(\theta) + (\alpha - 1) \sum_{i=1}^n \ln(x_i) - \frac{1}{\theta} \sum_{i=1}^n x_i$$

Taking the derivative of the log likelihood with respect to θ and setting it equal to 0:

$$\frac{\partial \ell(\alpha, \theta)}{\partial \theta} = -\frac{n\alpha}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^n x_i = 0$$

Solving for θ , we obtain the following MLE for θ

$$\hat{\theta}|\alpha = \frac{\sum_{i=1}^n x_i}{\alpha \cdot n} = \frac{1}{\alpha} \bar{x}_n$$

Substituting this back into the log likelihood function, taking its derivative with respect to α , and setting the result equal to 0:

$$\frac{\partial \ell(\alpha, \hat{\theta}|\alpha)}{\partial \alpha} = -\frac{n \cdot \Gamma'(\alpha)}{\Gamma(\alpha)} - n \ln\left(\frac{1}{\alpha} \bar{x}_n\right) + \sum_{i=1}^n \ln(x_i) = 0$$

Solving for α , we obtain the following condition for the MLE of α :

$$\ln(\alpha) - \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} = \ln(\bar{x}_n) - \frac{\sum_{i=1}^n \ln(x_i)}{n}$$

Therefore, the MLE values of α and θ must satisfy the following

$$\ln(\hat{\alpha}) - \frac{\Gamma'(\hat{\alpha})}{\Gamma(\hat{\alpha})} = \ln(\bar{x}_n) - \frac{\sum_{i=1}^n \ln(x_i)}{n}$$

$$\hat{\theta} = \frac{1}{\hat{\alpha}} \bar{x}_n$$

3.4.2 Lognormal Distribution

A random variable X has the Lognormal distribution with parameters $\mu \in \mathbb{R}$ and $\sigma \in (0, \infty)$ if $\ln(X)$ has the normal distribution with mean μ and standard deviation σ . The parameter σ is the shape parameter of X while e^μ is the scale parameter of X . The probability density function of the Lognormal distribution with parameters μ and σ is given by

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{[\ln(x)-\mu]^2}{2\sigma^2}}$$

The Lognormal distribution is used to model continuous random quantities when the distribution is believed to be skewed, such as income and lifetime variables.

3.4.2.1 Estimation of the parameters of Lognormal Distribution

To compute the maximum likelihood estimators, we start with the likelihood function. The likelihood function of the Lognormal distribution is derived by taking the product of the probability densities of the individual X which are independent sample drawn from the distribution.

$$L(\mu, \sigma^2 | x) = \prod_{i=1}^n [f(x_i | \mu, \sigma^2)]$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \prod_{i=1}^n x_i^{-1} \left[e^{\left(\frac{\sum_{i=1}^n -\frac{[\ln(x_i) - \mu]^2}{2\sigma^2} \right)} \right]$$

The log-likelihood function of the Lognormal distribution is then derived by taking the natural log of the likelihood function

$$\ell(\mu, \sigma^2 | x) = \ln \left((2\pi\sigma^2)^{-\frac{n}{2}} \prod_{i=1}^n x_i^{-1} \left[e^{\left(\frac{\sum_{i=1}^n -\frac{[\ln(x_i) - \mu]^2}{2\sigma^2} \right)} \right] \right)$$

$$= -\frac{n}{2} \ln(2\pi\sigma^2) - \sum_{i=1}^n \ln(x_i) - \frac{\sum_{i=1}^n \ln(x_i)^2}{2\sigma^2} + \frac{\sum_{i=1}^n \ln(x_i)\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2}$$

We now find $\hat{\mu}$ and $\hat{\sigma}^2$, which maximize $\ell(\mu, \sigma^2 | x)$. To do this, we take the derivative of ℓ with respect to μ and σ^2 and set it equal to 0: with respect to μ ,

$$\frac{\partial \ell}{\partial \mu} = \frac{\sum_{i=1}^n \ln(x_i)}{\sigma^2} - \frac{2n\hat{\mu}}{2\hat{\sigma}^2} = 0$$

$$\Rightarrow \hat{\mu} = \frac{\sum_{i=1}^n \ln(x_i)}{n}$$

With respect to σ^2 ,

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2\hat{\sigma}^2} - \frac{\sum_{i=1}^n (\ln(x_i) - \hat{\mu})^2}{2} (-\hat{\sigma}^2)^{-2}$$

$$\Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^n \left(\ln(x_i) - \frac{\sum_{i=1}^n \ln(x_i)}{n} \right)^2}{n}$$

Thus, the maximum likelihood estimators are

$$\hat{\mu} = \frac{\sum_{i=1}^n \ln(x_i)}{n} \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^n \left(\ln(x_i) - \frac{\sum_{i=1}^n \ln(x_i)}{n} \right)^2}{n}$$

3.4.3 Gumbel Max Distribution

The Gumbel Max distribution (Generalized Extreme Value distribution Type-I) is used, in probability theory and statistics, to model the distribution of the maximum of a number of samples of various distributions. The Gumbel Max distribution with location parameter μ and scale parameter σ has a probability density function (pdf)

$$f(x) = \frac{1}{\sigma} \exp(-z - \exp(-z))$$

Where $z = \frac{x - \mu}{\sigma}$

The pdf can be written as

$$f(x) = \frac{1}{\sigma} \exp\left(-\frac{x - \mu}{\sigma}\right) \exp\left\{-\exp\left(-\frac{x - \mu}{\sigma}\right)\right\}$$

3.4.3.1 Estimation of the parameters of Gumbel Max Distribution

The probability density function of the Gumbel max distribution is given as:

$$f(x | \mu, \sigma) = \frac{1}{\sigma} \exp\left(-\frac{x - \mu}{\sigma}\right) \exp\left\{-\exp\left(-\frac{x - \mu}{\sigma}\right)\right\}$$

We obtain the MLE of μ and σ as follows

$$L(\mu, \sigma|x) = \prod_{i=1}^n [f(x_i|\mu, \sigma)]$$

$$= \left(\frac{1}{\sigma}\right)^n \prod_{i=1}^n \left[\exp\left(-\frac{x_i - \mu}{\sigma}\right) \exp\left\{-\exp\left(-\frac{x_i - \mu}{\sigma}\right)\right\} \right]$$

The log-likelihood function is then derived by taking the natural log of the likelihood function

$$\ell(\mu, \sigma|x) = \ln \left(\left(\frac{1}{\sigma}\right)^n \prod_{i=1}^n \left[\exp\left(-\frac{x_i - \mu}{\sigma}\right) \exp\left\{-\exp\left(-\frac{x_i - \mu}{\sigma}\right)\right\} \right] \right)$$

$$= -n \ln \sigma - \sum_{i=1}^n \frac{x_i - \mu}{\sigma} - \sum_{i=1}^n \exp\left[-\frac{x_i - \mu}{\sigma}\right]$$

We take the partial derivative of ℓ with respect to μ and σ and set it equal to 0:

$$\frac{\partial \ell(\mu, \sigma|x)}{\partial \mu} = \frac{n}{\sigma} - \frac{\sum_{i=1}^n \exp\left[-\frac{x_i - \mu}{\sigma}\right]}{\sigma} = 0$$

$$\frac{\partial \ell(\mu, \sigma|x)}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma^2}\right) - \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma^2}\right) \exp\left[-\frac{x_i - \mu}{\sigma}\right] - \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma}\right) = 0$$

Solving the system $\frac{\partial \ell(\mu, \sigma|x)}{\partial \mu} = \frac{\partial \ell(\mu, \sigma|x)}{\partial \sigma} = 0$ numerically yields the maximum likelihood estimates of μ and σ as:

$$\hat{\mu} = \hat{\sigma} \left(\ln n - \ln \sum_{i=1}^n \exp\left[-\frac{x_i}{\hat{\sigma}}\right] \right)$$

and

$$\hat{\sigma} = \bar{x} - \frac{\sum_{i=1}^n x_i \exp\left[-\frac{x_i}{\hat{\sigma}}\right]}{\sum_{i=1}^n \exp\left[-\frac{x_i}{\hat{\sigma}}\right]}$$

Where \bar{x} denotes the sample mean

3.4.4 Weibull Distribution

The Weibull distribution is a very flexible life distribution model with two parameters. It has the probability density function

$$f(x, \gamma, k) = \frac{k}{\gamma} \left(\frac{x}{\gamma}\right)^{k-1} e^{-\left(\frac{x}{\gamma}\right)^k}, \quad x > 0,$$

where $k > 0$ is the shape parameter and $\gamma > 0$ is the scale parameter

3.4.4.1 Estimation of the parameters of Weibull Distribution

The probability density function of the Weibull distribution is given as

$$f(x, \gamma, k) = \frac{k}{\gamma} \left(\frac{x}{\gamma}\right)^{k-1} e^{-\left(\frac{x}{\gamma}\right)^k}, \quad x > 0, k, \gamma > 0$$

The likelihood function of the Weibull distribution is obtained as follows

$$L(\gamma, k|x) = \prod_{i=1}^n [f(x_i|\gamma, k)]$$

$$= \left(\frac{k}{\gamma}\right)^n \prod_{i=1}^n \left[\left(\frac{x_i}{\gamma}\right)^{k-1} e^{-\left(\frac{x_i}{\gamma}\right)^k} \right]$$

Taking the natural logarithm gives the log likelihood function

$$\ell(\gamma, k|x) = n \ln\left(\frac{k}{\gamma}\right) + (k-1) \sum_{i=1}^n \ln x_i - n(k-1) \ln \gamma - \sum_{i=1}^n \left(\frac{x_i}{\gamma}\right)^k$$

Differentiating partially with respect to k and γ in turn and equating to zero, we obtain the estimation as follows

$$\frac{\partial \ell(\gamma, k|x)}{\partial k} = \frac{n}{k} + \sum_{i=1}^n \ln x_i - \frac{1}{\gamma} \sum_{i=1}^n x_i^k \ln x_i = 0$$

$$\frac{\partial \ell(\gamma, k|x)}{\partial \gamma} = -\frac{n}{\gamma} + \frac{1}{\gamma^2} \sum_{i=1}^n x_i^k = 0$$

We obtain an estimation of γ as

$$\hat{\gamma} = \frac{1}{n} \sum_{i=1}^n x_i^{\hat{k}}$$

And on substitution, we obtain

$$\frac{1}{\hat{k}} + \frac{1}{n} \sum_{i=1}^n \ln x_i - \frac{\sum_{i=1}^n x_i^{\hat{k}} \ln x_i}{\sum_{i=1}^n x_i^{\hat{k}}} = 0$$

3.5 Different Discrimination Procedure

This section is to succinctly describe methods for choosing the best fitted model to a given data set.

3.5.1 Theoretical Basis

In order to determine a suitable distribution to a data set theoretically, we can consider the Likelihood Ratio Test (LRT).

Generally, given the observed data X_1, \dots, X_n , we can measure the relative plausibility of H_1 to H_0 by the log – likelihood ratio:

$$\ln \left(\frac{f(X_1, \dots, X_n | H_1)}{f(X_1, \dots, X_n | H_0)} \right)$$

The log-likelihood ratio could help us choose which model (H_0 or H_1) is a more likely explanation for the data.

Furthermore, the LRT is a statistical test relying on a test statistics obtained by taking the ratio of the maximum value of the likelihood function.

Let X_1, X_2, \dots, X_n be i.i.d random variables from a known distribution (with p.d.f), if given the likelihood function, then the LRT is defined as:

$$LRT = \frac{L_1(\hat{\theta}_1, \hat{\theta}_2)}{L_2(\hat{\lambda}_1, \hat{\lambda}_2)}$$

Where $L_1(\hat{\theta}_1, \hat{\theta}_2)$ and $L_2(\hat{\lambda}_1, \hat{\lambda}_2)$ are the likelihood function of a known different distributions and $\hat{\theta}_1, \hat{\theta}_2, \hat{\lambda}_1$ and $\hat{\lambda}_2$ are the MLE of $\theta_1, \theta_2, \lambda_1$ and λ_2 respectively.

Now, we obtain the log-likelihood ratio statistic, denoted as T and is defined as the natural logarithm of the ratio of the two maximized likelihood functions:

$$T = \ln(LRT) = \ln \left[\frac{L_1(\hat{\theta}_1, \hat{\theta}_2)}{L_2(\hat{\lambda}_1, \hat{\lambda}_2)} \right]$$

$$T = \ln[L_1(\hat{\theta}_1, \hat{\theta}_2)] - \ln[L_2(\hat{\lambda}_1, \hat{\lambda}_2)]$$

Given more than two models, the log likelihood ratio test applied to each pair of models implies that the model with the largest log likelihood is preferred.

4. Application

We applied the methods to monthly prices of Brent crude oil from January 2009 to June 2017. Results obtained are presented and discussed below. We also measured performance of methods according to their aptitude to capture the statistical behaviour of the observed data set. Finally, we carried out a comparative analysis of the methods.

4.1 Data

The data used were obtained from the official website of The World Bank, “<http://www.worldbank.org/en/research/commodity-markets>” and it consists of 102 observations. Table 1: gives the monthly prices of the Brent crude oil prices from January 2009 to June 2018 (US\$/Barrel). Table 2 presents some sample distributional statistics for the monthly prices of the Brent crude oil prices. This consists of the mean, variance, skewness, kurtosis etc. Table 3 shows the transformed value of the monthly prices of Brent Crude oil.

Table 1: The monthly prices of the Brent crude oil prices from January 2009 to June 2017 (US\$/Barrel)

	2009	2010	2011	2012	2013	2014	2015	2016	2017
Jan	43.86	77.12	92.69	107.07	105.10	102.10	47.11	29.78	53.59
Feb	41.84	74.76	97.91	112.69	107.64	104.83	54.79	31.03	54.35
Mar	46.65	79.30	108.65	117.79	102.52	104.04	52.83	37.34	50.90
Apr	50.28	84.18	116.24	113.67	98.85	104.87	57.54	40.75	52.16
May	58.15	75.62	108.07	104.09	99.37	105.71	62.51	45.94	49.89
Jun	69.15	74.73	105.85	90.73	99.74	108.37	61.31	47.69	46.17
Jul	64.67	74.58	107.92	96.75	105.26	105.23	54.34	44.13	
Aug	71.63	75.83	100.49	105.27	108.16	100.05	45.69	44.87	
Sep	68.35	76.12	100.82	106.28	108.76	95.85	46.28	45.04	
Oct	74.08	81.72	99.85	103.41	105.43	86.08	46.96	49.29	
Nov	77.55	84.53	105.41	101.17	102.63	76.99	43.11	45.26	
Dec	74.88	90.01	104.23	101.19	105.48	60.70	36.57	52.62	

Table 2: Descriptive statistics of the monthly price of Brent crude oil prices

Minimum	29.780000
First Quartile	52.672500
Median	78.425000
Mean	78.739510
Variance	664.114539
Standard deviation	25.770420
Third quartile	104.077500
Skewness	-0.224953
Kurtosis	-1.483637
Maximum	117.790000
Sample size	102.000000

Figure 1: Shows the trend and stochastic fluctuations of the oil prices which was categorised as five states.

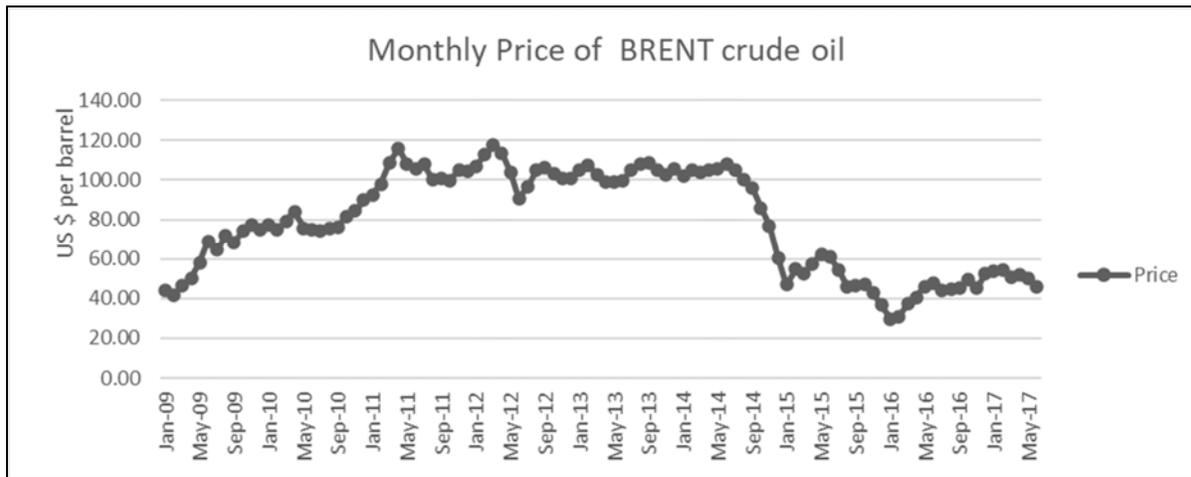
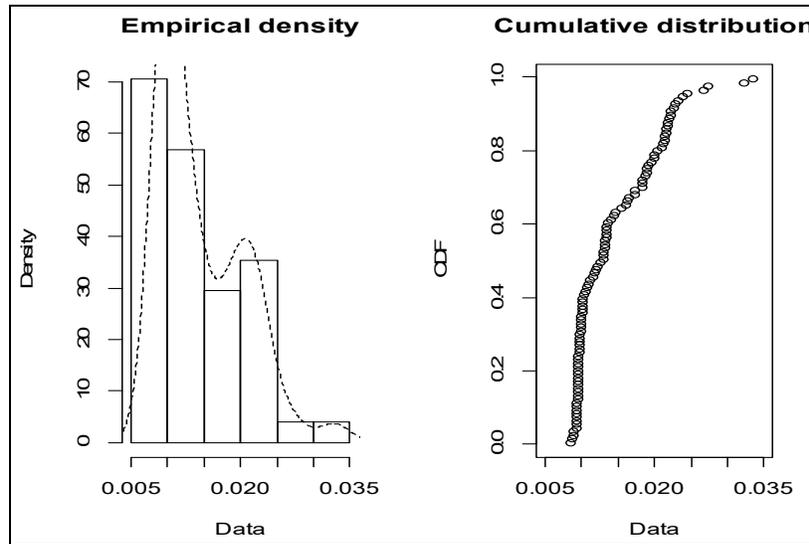


Fig 1: Trend of Brent oil prices from January 2009 to June 2017

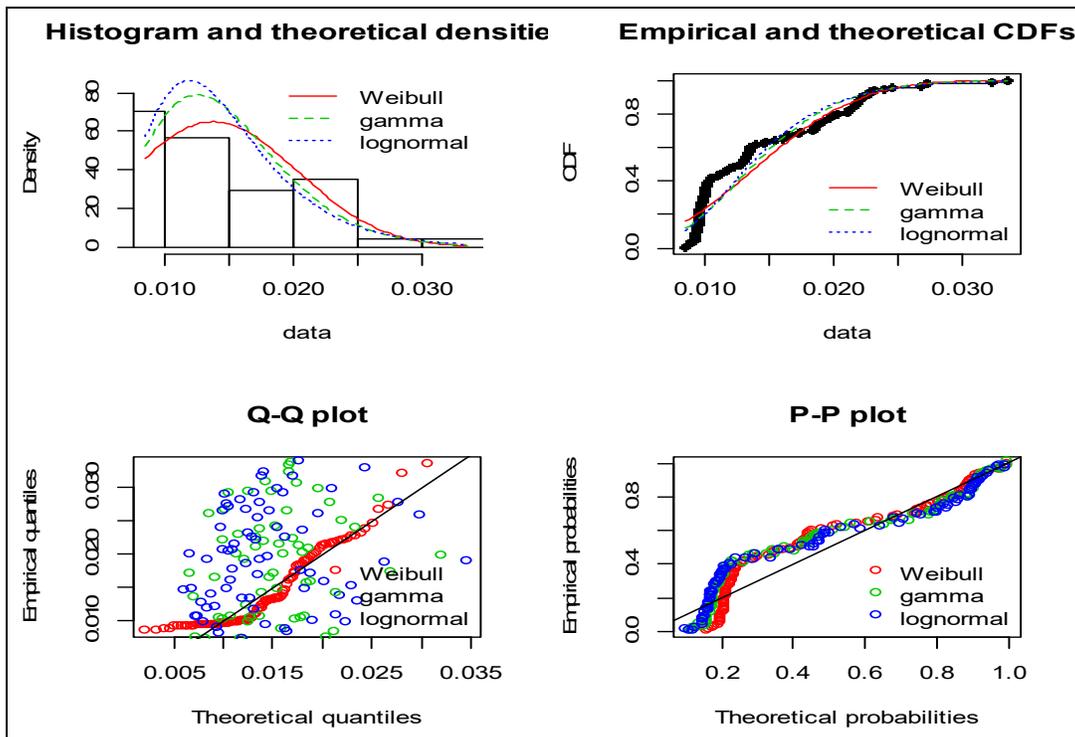
Table 3: Transformed value of the monthly prices of Brent Crude oil

	2009	2010	2011	2012	2013	2014	2015	2016	2017
Jan	0.0228	0.012967	0.010789	0.009339	0.009515	0.0098	0.0212	0.0336	0.0187
Feb	0.023898	0.013376	0.010213	0.008874	0.00929	0.0095	0.0183	0.0322	0.0184
Mar	0.021438	0.012611	0.009204	0.00849	0.009754	0.0096	0.0189	0.0268	0.0196
Apr	0.019889	0.011879	0.008603	0.008798	0.010116	0.0095	0.0174	0.0245	0.0192
May	0.017196	0.013224	0.009253	0.009607	0.010064	0.0095	0.0160	0.0218	0.0200
Jun	0.014461	0.013382	0.009448	0.011022	0.010026	0.0092	0.0163	0.0210	0.0217
Jul	0.015464	0.013409	0.009266	0.010335	0.0095	0.0095	0.0184	0.0227	
Aug	0.013961	0.013188	0.009952	0.009499	0.009246	0.0100	0.0219	0.0223	

Sep	0.014631	0.013138	0.009919	0.009409	0.009195	0.0104	0.0216	0.0222	
Oct	0.013499	0.012237	0.010015	0.00967	0.009485	0.0116	0.0213	0.0203	
Nov	0.012895	0.01183	0.009487	0.009884	0.009744	0.0130	0.0232	0.0221	
Dec	0.013354	0.01111	0.009594	0.009882	0.00948	0.0165	0.0273	0.0190	



Histogram and CDF plots of an empirical distribution for a continuous variable (transformed value of the monthly prices of Brent Crude oil) as provided in R software by the plotdist function. Four goodness-of-fit plots for various distributions fitted to continuous data (Weibull, Gamma and Lognormal distributions fitted to transformed value of the monthly prices of Brent Crude oil) as provided in R software by functions denscomp, qqcomp, cdfcomp and ppcomp.



4.2 The Oil Price Transition States

In this section, we categorized the oil prices as five states: low state, middle - low state, middle state, middle - high state and high state. These states constitute the following full space for stochastic events of crude oil prices:

(0, 40) [40, 60) [60, 80) [80, 100) [100,120)

The average of oil prices is US\$ 78 per barrel. Thus, we treat the [60, 80) as a middle state of fluctuating oil prices, the [40, 60) interval as a middle - low state, and the [80, 100) interval as a middle - high state. The three states are broadly termed as the middle state. By contrast, we treat the (0, 40) interval as a low state of oil prices and the [100, 120) interval as a high state. Let A represent the (0, 40) interval of low-state oil price, B the [40, 60), C the [60, 80), D the [80, 100) the middle state interval, and E the [100, 120) high state interval. Table 4 shows the state transition frequency matrix of oil price five state transition chains.

Table 4: State Transition Frequency of Oil Prices from January 2009 to June 2017

	A(0,40)	B[40,60)	C[60,80)	D[80,100)	E[100,120)	n_i
A(0,40)	3	1	0	0	0	4
B[40,60)	1	25	2	0	0	28
C[60,80)	0	2	15	2	0	19
D[80,100)	0	0	2	8	4	14
E[100,120)	0	0	0	4	32	36
n_j	4	28	19	14	36	$n=101$

4.2.2 Tests for the Existence of a Markov Chains

We test whether the five state transition chain follows a Markov chains or not. by chi-squae test. To test this, we use the formula (Raheem and Ezepeue (2016); Xu, (2001)).

$$\chi^2 = \sum_{i=1}^m \sum_{j=1}^m \frac{(n_{ij} - n_i.n_j/n)^2}{(n_i.n_j/n)}$$

where n_i , ($i = 1, 2, \dots, m$) and n_j ($j = 1, 2, \dots, m$) are the frequency of state i and state j respectively. This has χ^2 distribution with $(m - 1)^2$ degrees of freedom, where m refers to the number of states. The test results are reported in Table 5.

Table 5: χ^2 Testing Results of Crude Oil Price Transition Chain

$\frac{(n_{ij} - n_i.n_j/n)^2}{(n_i.n_j/n)}$	j = 1 (A)	j = 2 (B)	j = 3 (C)	j = 4 (D)	j = 5 (E)
i = 1 (A)	50.9709	0.0107	0.7525	0.5545	1.4257
i = 2 (B)	0.0107	38.279	2.0267	3.8812	9.9802
i = 3 (C)	0.7525	2.0267	36.5244	0.1525	6.7723
i = 4 (D)	0.5545	3.8812	0.1525	18.9202	0.1964
i = 5 (E)	1.4257	9.9802	6.7723	0.1964	28.6342
$\sum_{i=1}^m$	53.7123	54.1778	46.2284	23.7048	47.0088
$\sum_{i=1}^m \sum_{j=1}^m$				224.8321	

With $m = 5$, the degrees of freedom $(m - 1)^2 = (5 - 1)^2 = 16$. Using a 5% significance level and referring to the χ^2 tables with 16 degrees of freedom, we find that $\chi^2_{(16,\alpha=0.005)} = 26.296$. The observed value of the sample statistics χ^2 is 224.8321, which is higher than $\chi^2_{(16,\alpha=0.005)} = 26.296$. Thus, we reject the null hypothesis that states are independent. As a result, it confirms that a state transition chain of Brent crude oil prices from January 2009 to June 2017 follows a Markov Chain.

4.3 Limit Probability of a Markov Chain for Changing Trends of Brent Oil Prices

Let P represent the first – stage transition matrix. Having each element n_{ij} of respective row from Table 3 divided by the sum of its row (n_i), then we obtain the first – stage transition matrix P:

$$P = \begin{bmatrix} 0.7500 & 0.2500 & 0.0000 & 0.0000 & 0.0000 \\ 0.0357 & 0.8929 & 0.0714 & 0.0000 & 0.0000 \\ 0.0000 & 0.1053 & 0.7895 & 0.1053 & 0.0000 \\ 0.0000 & 0.0000 & 0.1429 & 0.5714 & 0.2857 \\ 0.0000 & 0.0000 & 0.0000 & 0.1111 & 0.8889 \end{bmatrix}$$

Secondly, we can derive the second – stage transition matrix, ..., and the convergence state transition matrix as follows:

$$P^2 = \begin{bmatrix} 0.5714 & 0.4107 & 0.0179 & 0.0000 & 0.0000 \\ 0.0587 & 0.8137 & 0.1201 & 0.0075 & 0.0000 \\ 0.0038 & 0.1772 & 0.6459 & 0.1433 & 0.0301 \\ 0.0000 & 0.0150 & 0.1945 & 0.3733 & 0.4172 \\ 0.0000 & 0.0000 & 0.0159 & 0.1622 & 0.8219 \end{bmatrix}$$

$$P^5 = \begin{bmatrix} 0.2834 & 0.6060 & 0.0978 & 0.0108 & 0.0021 \\ 0.0866 & 0.6709 & 0.1947 & 0.0344 & 0.0135 \\ 0.0206 & 0.2870 & 0.4090 & 0.1477 & 0.1359 \\ 0.0031 & 0.0688 & 0.2004 & 0.2121 & 0.5156 \\ 0.0002 & 0.0105 & 0.0717 & 0.2005 & 0.7171 \end{bmatrix}$$

.....

$$P^{250} = \begin{bmatrix} 0.0396 & 0.2772 & 0.1881 & 0.1386 & 0.3564 \\ 0.0396 & 0.2772 & 0.1881 & 0.1386 & 0.3564 \\ 0.0396 & 0.2772 & 0.1881 & 0.1386 & 0.3564 \\ 0.0396 & 0.2772 & 0.1881 & 0.1386 & 0.3564 \\ 0.0396 & 0.2772 & 0.1881 & 0.1386 & 0.3564 \end{bmatrix}$$

The row – vector, which has been converged to the same value, is the Markov chain’s limit of oil price series denoted by L . With the initial state probability vector, $X^{(0)} = [0 \ 1 \ 0 \ 0 \ 0]$, we have

$$X^{(\infty)} = [0 \ 1 \ 0 \ 0 \ 0] \begin{bmatrix} 0.0396 & 0.2772 & 0.1881 & 0.1386 & 0.3564 \\ 0.0396 & 0.2772 & 0.1881 & 0.1386 & 0.3564 \\ 0.0396 & 0.2772 & 0.1881 & 0.1386 & 0.3564 \\ 0.0396 & 0.2772 & 0.1881 & 0.1386 & 0.3564 \\ 0.0396 & 0.2772 & 0.1881 & 0.1386 & 0.3564 \end{bmatrix}$$

$$X^{(\infty)} = [0.0396 \ 0.2772 \ 0.1881 \ 0.1386 \ 0.3564]$$

- $X^{(\infty)}$ is the limiting state probability vector

The implication is that the probability distribution of states becomes stabilized after a continuous transition process. This convergence process is the changing trends of crude oil prices, and the row – vector as the limit probability is the ultimate state of oil price series. It is expressed as follows:

	A(0,40)	B[40,60)	C[60,80)	D[80,100)	E[100,120)
Limit Probability Value of Markov Chain	0.0396	0.2772	0.1881	0.1386	0.3564

This limit probability vector indicates the ultimate probability or proportion of five states in the crude oil price series. The probability of low level state A(0,40) is 0.0396, meaning that the proportion in the series is 3.96%; the probability of middle – low level state B[40,60) is 0.2772, the proportion is 27.72%, the probability of middle level state C[60,80) is 0.1881, the proportion is 18.81%, the probability of middle – high level state D[80,100) is 0.1386, the proportion is 13.86%, the probability of high level state E[100,120) is 0.3564, the proportion is 35.64%, respectively.

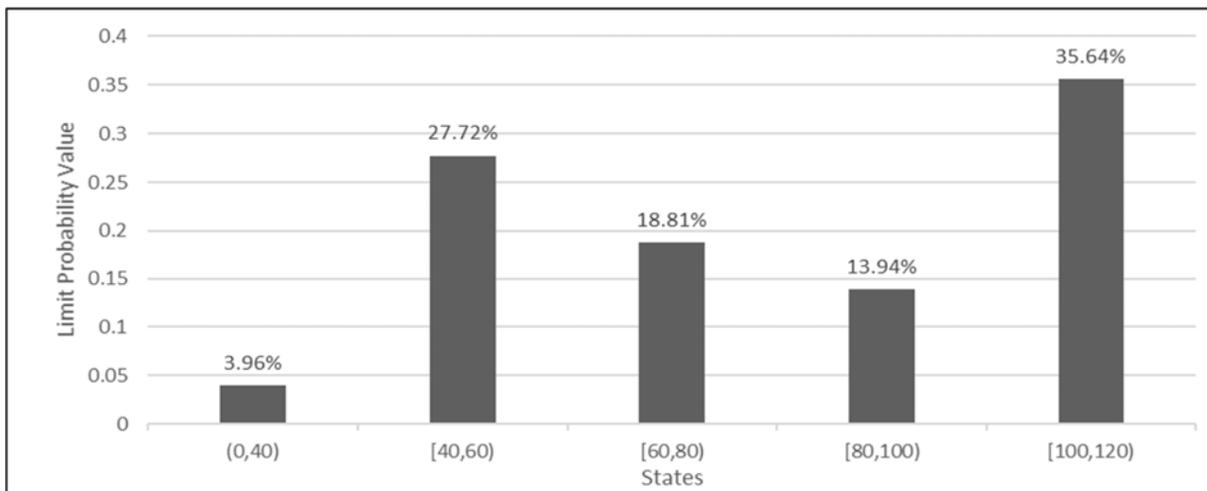


Fig 2: The Limit Probability of Brent monthly oil prices

The limit probability of oil - price state transition chain as a Markov Chain is the ultimate state of oil price series. It approximates the changing trends of oil price in the medium- and long-terms, but not in the short-term.

4.4 The Probability Distribution of the Changing Trends of Oil Prices

In this research work, four distributions have been considered, namely, Gamma, Lognormal, Gumbel Max and Weibull distribution. We carry out some statistical tests among the competing models to determine which distribution best fits the oil price data.

4.4.1 Parameter Estimates of the Distributions

The maximum likelihood estimates of each of the parameters of the distributions are obtained. Table 6 shows parameter estimates for the distributions.

Table 6: Parameter estimates of the Distribution

Distributions	Estimated Parameters
GAMMA $GA(\alpha, \theta)$	$\hat{\alpha} = 9.5030657$ $\hat{\theta} = 8.2859390$
LOGNORMAL $LN(\mu, \sigma^2)$	$\hat{\mu} = 4.3045803$ $\hat{\sigma} = 0.3649262$
GUMBEL MAX $GB(\mu, \sigma)$	$\hat{\mu} = 65.72135$ $\hat{\sigma} = 23.94365$
WEIBULL $WE(\gamma, k)$	$\hat{\gamma} = 87.816001$ $\hat{k} = 3.587463$

4.4.2 Log – Likelihood Ratio Values of the Distributions

The log-likelihoods form the basis of the selection of the distribution that fits the data. The distribution with the maximum log – likelihood value is considered the most suitable. The log – likelihood values are provided in Tables 7.

Table 7: Log – likelihood values of the Distributions

Distributions	Gamma $GA(\alpha, \theta)$	Lognormal $LN(\mu, \sigma^2)$	Gumbel Max $GB(\mu, \sigma)$	Weibull $WE(\gamma, k)$
Log- Likelihood	-478.0142	-480.9768	-481.3798	-473.236

The log – likelihood ratio value is gotten by finding the difference between likelihood values of competing distributions

The Weibull distribution arises as the most appropriate distribution for the data as it has the highest log – likelihood value.

4.4.3 Goodness-of-Fit Test

We employed the Kolmogorov - Smirnov (see Kolmogorov (1941) and Smirnov (1948)) and Anderson-Darling (1952) goodness-of-fit test to measure the "distance" between the data and the distribution we are testing, and compare that distance to some threshold value. Since the goodness-of-fit test statistics indicate the distance between the data and the fitted distributions, it is obvious that the distribution with the lowest statistic value is the best fitting model.

Table 8: Goodness-of-fit test

	Gamma	Lognormal	Gumbel Max	Weibull
Kolmogorov-Smirnov	0.17975	0.18196	0.20851	0.10540
Anderson- Darling	4.2883	5.4835	7.3024	4.0169

The Weibull distribution indicated to have the best fit of the four distributions. The Kolmogorov – Smirnov and Anderson - Darling statistics have the least figure under the Weibull distribution.

These results show that in comparison to Gamma, Lognormal and Gumbel Max distribution, the Weibull distribution is a more acceptable fit for the Brent crude oil price.

4.4.4 Comparison of Model

We compare the suitability of the distribution models. Specification measures such as Akaike Information Criteria (AIC) and Bayes Information Criteria (BIC) were employed. We noticed that the AIC and BIC for the Weibull distribution is smaller than those of the Gamma, Lognormal and Gumbel Max distribution. Thus, the Weibull distribution is a better fit for the Brent oil Price.

Table 9: Comparison of the results of the models

	Gamma	Lognormal	Gumbel Max	Weibull
AIC	960.0283	965.9536	966.7596	950.4721
BIC	965.2782	971.2035	972.0095	955.7220

From Table 6 – 9, it can be concluded that the oil price series conform to a Weibull distribution.

The function of the Weibull distribution is as follows:

$$f(x, \gamma, k) = \frac{k}{\gamma} \left(\frac{x}{\gamma}\right)^{k-1} e^{-\left(\frac{x}{\gamma}\right)^k}, \quad x > 0, \text{ where } k > 0 \text{ is the shape parameter and } \gamma > 0 \text{ is the scale parameter.}$$

The estimated parameters are: $k = 3.587$ and $\gamma = 87.816$

Substituting $k = 3.587$ and $\gamma = 87.816$ in the Weibull distribution, we have

$$\begin{aligned} f(x) &= \frac{3.587}{87.816} \left(\frac{x}{87.816}\right)^{3.587-1} e^{-\left(\frac{x}{87.816}\right)^{3.587}} \\ &= 0.041 \left(\frac{x}{87.816}\right)^{2.587} e^{-\left(\frac{x}{87.816}\right)^{3.587}} \end{aligned}$$

Hence, the fitted probability density function (pdf) is given as:

$$f(x) = 0.041 \left(\frac{x}{87.816}\right)^{2.587} e^{-\left(\frac{x}{87.816}\right)^{3.587}}$$

4.5 Probability Value of Observed Oil Prices

The fitted probability density function, $f(x) = 0.041 \left(\frac{x}{87.816}\right)^{2.587} e^{-\left(\frac{x}{87.816}\right)^{3.587}}$

We calculate the probability value of each interval as follows:

$$p_i = \int_{\text{the lower limit of interval}}^{\text{the upper limit of interval}} \frac{k}{\gamma} \left(\frac{x}{\gamma}\right)^{k-1} e^{-\left(\frac{x}{\gamma}\right)^k} dx, x = 0, 20, \dots, 120, \dots$$

Substituting the end – point value of each interval,

$$p_i = \int_{x_i}^{x_i+20} 0.041 \left(\frac{x}{87.816}\right)^{2.587} e^{-\left(\frac{x}{87.816}\right)^{3.587}} dx, x_i = 0, 20, \dots, 120, \dots$$

Table 10: Probability values of observed oil prices

Observed interval	(0,20)	[20,40)	[40,60)	[60,80)	[80,100)	[100,120)	>120
Probability value of the interval	0.0049	0.0529	0.1673	0.2861	0.2856	0.1565	0.0467
Observed prices	20	40	60	80	100	120	>120
Accumulated probability value	0.0049	0.0578	0.2251	0.5112	0.7968	0.9533	1.0000

As shown in Table 10, the probability of oil prices falling in the interval of US\$ 20-40 per barrel is 0.0529, 0.1673 in the interval of US\$ 40-60 per barrel, 0.2861 in the interval of US\$ 60-80 per barrel, 0.2856 in the interval of US\$ 80-100 per barrel respectively, whereas the probability of oil prices higher than US\$ 100 per barrel is 0.2032.

The probability of oil prices below 20 US\$/barrel is merely 0.49%. This probability distribution of oil prices puts the probability of recent oil prices in the range of US\$ 20 to US\$ 120 per barrel at 0.9484. It can be considered as an inevitable event. The probability of oil prices in the interval of US\$ 20-100 per barrel is 0.7919.

The probability of oil prices over 120 US\$/barrel is 0.0467. It is a fairly low probability event. The probability of oil prices over 200 US\$/barrel is 0. It is an impossible event. However, it should be pointed out that these probability distribution fittings are based on the recent statistics so that they only reveal recent changes in oil prices. Thus, this deduction is only meaningful for the recent changing trends of oil prices. In the next section, we discuss deducing changing trends of oil prices in the middle- and long-terms.

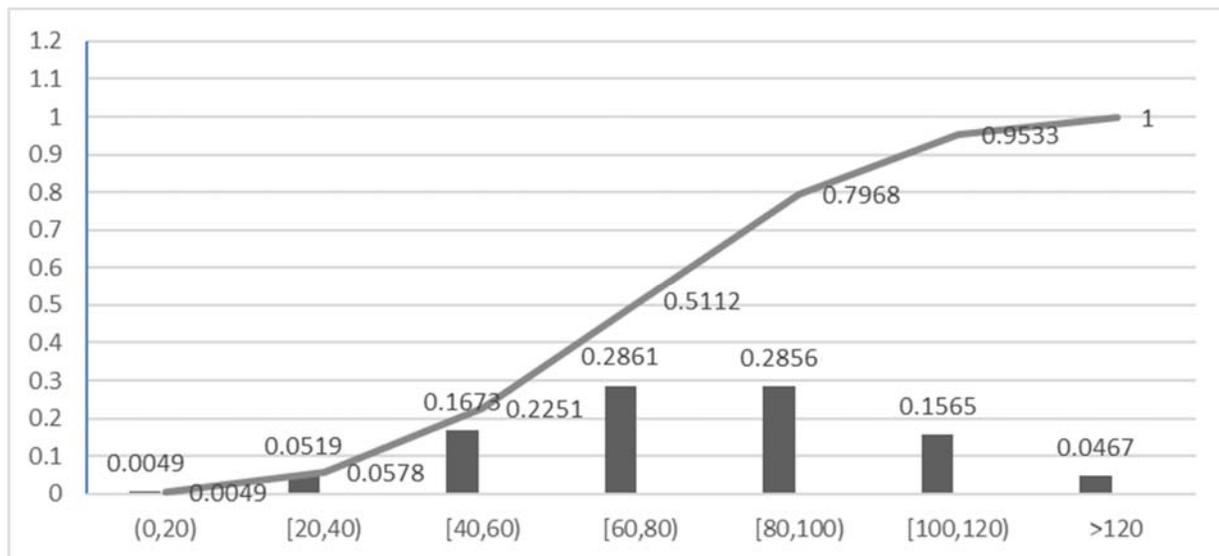


Fig 3: The probability Distribution and ccumulative probability values of Brent oil prices

From the general distribution of oil prices, the probability of oil prices below 120US\$/barrel is 0.9533, indicating that this price or less would prevail in the market.

4.6 A Deduction Model of Integrating the Limit Probability of a Markov Chain With a Probability Distribution

In the previous sections, we discuss the Markov chain model and the probability function model, separately. Here, we integrate the two models, in order to infer the changing trends of Brent oil prices from the short – term to the middle – and long – terms.

Table 11: Comparison between Weibull Distribution and Limit Probability of a Markov chain of Brent oil prices

Price Interval	(0,40)	[40,60)	[60,80)	[80,100)	[100,120)
Probability value of Weibull Distribution	0.0578	0.1673	0.2861	0.2856	0.2032*
Limit probability value of Markov chain	0.0396	0.2772	0.1881	0.1386	0.3564
The difference between the two probabilities	-0.0182	0.1099	-0.0980	-0.1470	0.1532
The percentage of the above difference (%)	-31.49	65.69	-34.25	-51.47	75.39

*0.2032 is the probability of the crude oil prices higher than US\$ 100 per barrel.

As shown in Table 11, the probability of oil prices being 40 US\$/barrel or less is 0.0578 in the short - term, while such a probability is 0.0396 in the middle - and long - terms, 31.49% less than that in the short-term. The probability of oil prices being in the [60,80) and [80,100) intervals are reduced from 0.2861 and 0.2856 in the short-term to 0.1881 and 0.1386 in the middle - and long - terms by 34.25% and 51.47% respectively. By contrast, the probability of oil prices falling in the [40, 60) and [100,120) intervals are increased by 65.69% and 75.39%.

4.6.1 Theoretical Frequency of a Weibull Distribution and the Limit Frequency of a Markov Chain

Suppose that the length of a deduction period is the same as 102 months of the sample statistic. Multiplying the probability of a Weibull distribution and the limit probability of a Markov Chain in Table 11 by the sample observations of 102, we have the theoretical frequency of a Weibull distribution and the limit frequency of a Markov chain of oil prices, as shown in Table 12.

Table 12: A Comparison between the Theoretical Frequency of a Weibull Distribution and the Limit Frequency of a Markov Chain

	(0,40)	[40,60)	[60,80)	[80,100)	[100,120)
Theoretical frequency of Weibull distribution	5.8956	17.0646	29.1822	29.1312	20.7264*
Limit frequency of a Markov chain	4.0392	28.2744	19.1862	14.1372	36.3528
The difference between the two frequencies	-1.8564	11.2098	-9.9960	-14.994	15.6264
The percentage of the above difference (%)	-31.49	65.69	-34.25	-51.47	75.39

*20.7264 is the theoretical frequency of a Weibull distribution of crude oil prices higher than US\$ 100 per barrel.

It is clear from Table 12 that the number of monthly oil prices falling in the (0,40), [60,80) and [80,100) interval are decreasing. Specifically, the number of monthly oil prices in the (0,40) interval has been reduced from 5.8956 to 4.0392 by 1.8564, the number in the [60, 80) interval has fallen from 29.1822 to 19.1862 by 9.9960 and the number of monthly oil prices in the [80, 100) interval has also been reduced from 29.1312 to 14.1372 by 14.994. They altogether drop by 26.8464 \approx 27. By contrast, the number of monthly oil prices in the [40,60) and [100,120) intervals has increased by 11.2098 and 15.6264, respectively. The sum of the total increased number is 26.8362 \approx 27, which equals to the total decreased number. But the increasing range of each interval is different.

4.7 Time Homogeneity

One important aspect of Markov probabilities is the assumption of time homogeneity. That is, the probability of moving from a given state to the next state is assumed to remain constant over time. In the case of Brent crude oil price movements we speculate that this is a requirement that is not strictly adhere to. This study tested time dependence, while assuming the presence of time homogeneity without testing. The Brent oil prices exhibits periods of greater and lesser volatility. It does not demonstrate time homogeneity. This will be subject to the next paper.

5. Summary

Crude oil prices have been unstable over time and with huge disparities. It is this instability in oil price series that makes deduction of the changing trends of oil prices in the middle- and long- terms problematic and hence renders the prediction of their price levels in the short-term challenging. We certified and corroborated the use of a Markov chain to ascertain the Markovian assumption that current state depends on that of the previous state by using the conventional test for independence via chi – square statistic. The observed value of the sample statistics χ^2 , 224.8321 is much higher than $\chi^2_{(16,\alpha=0.005)} = 26.296$. Consequently, it verified that there is an existence of a Markov chain in the observed Brent crude oil price series from January 2009 to June 2017. We then modelled the oil prices series using the limit probability of a Markov chain and observed at the middle and long – term, that the probability of low level state is 0.0396, middle – low level state is 0.2772, the middle level state is 0.1881, middle – high level state is 0.1386 and high level state is 0.3564, having the proportion is 3.96%, 27.72%, 18.81%, 13.86% and 35.64%, respectively in the series.

We also confirmed that the Weibull distribution models of the oil price series performed better among the selected probability distributions using some test criterion. From our investigation, coupled with the results in Tables 6 – 9, we discovered that the oil price series can best be modelled using the Weibull distribution. From the table, it is apparent that the Weibull distribution has the highest log – likelihood function of -473.236 and also has the lowest AIC and BIC values of 950.4721 and 955.7220 respectively. This indicates that among the chosen statistical distributions, the Weibull distribution stands a better chance in providing a good

fit for the oil price data. The maximum likelihood estimates of the parameters of the distributions were also obtained. It was observed at the short – term that the probability of low level state is 0.0578, middle – low level state is 0.1673, middle level state is 0.2861, middle–high level state is 0.2856, and high level state is 0.1565.

Our work has, correspondingly, established a trend deduction model of fluctuating oil prices. This model integrates the probability distribution of oil price series with the limit probability distribution of a Markov Chain. The difference between them indicates specific changes in a variety of oil price states from the short – term to the middle – and long – terms.

The probability of oil prices below 40 US\$/barrel has been reduced by 31.49% while that of oil prices over 40 US\$/barrel has increased markedly.

6. Conclusion

Our analysis and the above summary undoubtedly show that there is an existence of a markov chain in the oil price data i.e. oil price state are clearly not independent and the Weibull distribution would provide a good fit to the oil price data when compared with the Gamma distribution, Lognormal distribution and Gumbel max distribution.

We then integrated the Markov chain model and Weibull distribution to model the oil price series and obtained results which shows that the oil prices will be over 40 US\$/barrel.

However, this deduction is only meaningful for the recent changing trends of oil prices. A lot of factors determine the changing trends of oil prices such as current supply and output, future supply and reserves, demand from major countries, political events and crises.

A change in the oil-price level as well as its variability poses an arduous challenge to a diverse range of actors in the global market. The overall effect of changes and uncertainties in oil prices on the economy as a whole depends on many parameters, including a country's position in the oil market and her degree of competitiveness.

7. References

- Balakrishnan N, Kateri M. "On the Maximum Likelihood Estimation of Parameters of Weibull Distribution Based On Complete and Censored Data", *Statistics and Probability Letters*. 2008; 78(17):2971-2975
- Bartlett MS. "The Frequency Goodness of Fit Test for Probability Chains", *Cambridge Philosophical Proceedings*. 1951; 47:86-94.
- Bierens HJ. *Information Criteria and Model Selection*, Pennsylvania: Pennsylvania State University, 2006.
- Bright OO, Agatha A. *Comparative Analysis of Some Distributions on the Capital Requirement Data for the Insurance Company*. 2013; 1(4):37-47.
- Burbidge J, Harrison A. "Testing for the Effects of Oil Price Rise using Vector Autoregressions", *International Economic Review*. 1984; 25:459-484.
- Castaneda LB, Arunachalam V, Dharmaraja D. *Introduction to Probability and Stochastic Processes with Applications*, New Jersey: John Wiley & Sons, 2012
- Dafas PA. Estimating the parameters of a mean-reverting Markov-switching jump-diffusion model for crude oil spot prices, 2004, 25-34.
- Davou NC, Samuel NE, Gokum TK. *Markov Chain Model Application on Share Price Movement in Stock Market*. 2013; 4(10):84-95.
- Douc R, Moulines E, Soulier P. "Computable Convergence Rates for Subgeometric Ergodic Markov Chains", *Bernoulli*, 2007; 13:831-848.
- Fisz M. *Probability Theory and Mathematical Statistics*, 3rd Edition, New York: Krieger Publishing, 1980.
- Garg VK, Singh JB. "Markov Chain Approach on the Behavior of Rainfall", *International Journal of Agricultural and Statistical Sciences*. 2010; 6(1):157-162.
- Goldfeld SM, Quandt RE. "A Markov Model for Switching Regressions", *Journal of Econometrics*. 1973; 1(1):3-15.
- Guure CB, Ibrahim NA. "Methods for estimating the 2-parameter Weibull Distribution with Type 1 Censored Data", *Research Journal of Applied Sciences, Engineering and Technology*. 2013; 5(3):689-694.
- Hamidreza M, Shaghayegh K, Mohammadmohsen O. "Modelling the Fluctuations of Brent Oil Prices by a Probabilistic Markov Chain", *Journal of Computations & Modelling*. 2011; 1(2):17-26.
- He K, Yu L, Lai KK. "Crude Oil Price Analysis and Forecasting using Wavelet Decomposed Ensemble Model", *Energy*. 2012; 46:564-574,
- Hoel PG. "A Test for Markoff Chains", *Biometrika*. 1954; 41:430-433.
- Holmes MJ, Wang P. "Oil Price Shocks and the Asymmetric Adjustment of UK Output: A Markov-switching Approach", *International Review of Applied Economics*. 2003; 17(2):181-192.
- Javier JG, Alberto R. "Use of Gumbel and Weibull Functions to Model Extreme Values of Diameter Distributions in Forest Stands", *Annals of Forest Science*. 2014; 71(7):741-750.
- Jing Y, Wang S, Yuan Q. *The Application of Markov Process in Geology*, Beijing: Geological Publishing House, 1986.
- Kosobud RF, Stokes HH. "Oil Market Share Dynamics: A Markov Chain Analysis of Consumer and Producer Adjustments", *Empirical Economics*. 1978; 3(4):253-275.
- Kundu D, Manglick A. "Discriminating between the Weibull and Log-Normal Distributions", *Naval Research Logistics (NRL)*. 2004; 51(6):893-905
- Larsson K, Nossman M. "Jumps and Stochastic Volatility in Oil Prices", *Time Series Evidence*, 2010, 53-65.
- Lu D. *Stochastic Process and Application*, Beijing: Tsinghua Press, 1987.
- Michael K. "Markov Chains: Models, Algorithms and Applications", Springer, 2005, 1-18.
- Oluwatomisin MO, Paul O, Adeyemi AO. *Oil Price and Exchange Rate Volatility in Nigeria*, 2014, 01-09.

26. Raheem MA, Yahya WB, Obisesan KO. "A Markov Chain Approach on Pattern of Rainfall Distribution", Journal of Environmental Statistics. 2015; 7(1):1-13.
27. Raheem MA, Ezepue PO. A Three- State Markov Model for Predicting Movements of Asset Returns of a Nigerian Bank. CBN Journal of Applied Statistics. 2016; 7(2):77-99.
28. Simeyo O, Edgar OO, Robert NN. "Application of Markov Chain to Model and Forecast Stock Market Trend: A Study of Safaricom Shares in Nairobi Securities Exchange, Kenya", International Journal of Current Research. 2015; 7(4):14712-14721.
29. Song D. "The Applications of Markov Law in Predicting the Price of Petroleum", Value Engineering. 2005; 12:105-106.
30. Tuo J, Yanbing. "Summary of World Oil Price Forecasting Model", In. IEEE Knowledge Acquisition and Modeling (KAM), 2011 Fourth International Symposium on, 2011, 327-330,
31. Vo MT. "Regime-switching Stochastic Volatility: Evidence from the Crude Oil Market", Energy Economics. 2009; 31(5):779-788.
32. Wang DQ, Martiz JS. "Note on Testing a Three-State Markov Chain for Independence", Journal of Statistics Computation and Simulation, 1990; 37:61-68.
33. Wang D, Cao L, Gao X, Li T. "Data Mining in Oil Price Time Series Analysis", Communications of the IIMA. 2006; 6(3):111-115
34. Wang Z. Basics of Probability Theory and its Application, Beijing: China Science Press, 1979.
35. Wei W, Chen Z, Wang J. "A Research on the Three States Markov-Switching Mode - An Application in the Analysis of World Oil Price Fluctuation", The Study of Finance and Economics. 2006; 6:120-131
36. Xie W, Yu L, Xu S, Wang S. "A New Method for Crude Oil Price Forecasting Based on Support Vector Machines", Springer, 2006, 444-451
37. Xu Haiyan and Zhang Zhong Xiang A trend deduction model of fluctuating oil prices, 2010, 1 – 19.
38. Xu H. "The Application of Markov Chain in the Research of Import Commodities Market in Kazakstan", in M. M. Shyngysbaeva (ed.) Kazakstan Economic Reform: Experience, Theory and Practice, Kazakhstan: Almaty, 2001, 129-139
39. Zakariya YA. "Using Maximum Likelihood Ratio Test to Discriminate Between the Inverse Gaussian and Gamma Distributions", International Journal of Statistical Distributions and Applications. 2015; 1(1):27-32.