INTERNATIONAL JOURNAL OF MATHEMATICAL ANALYSIS AND OPTIMIZATION: THEORY AND APPLICATIONS
VOL. 2019, NO. 2, PP. 571 - 580

State Space Model of Brent Oil Price Dynamics

Nkemnole, E.B [1*] And Agwuegbo, S. O. N. [2]

1 *, Department of Mathematics, University of Lagos, Akoka, Nigeria.
2, Department of Statistics, Federal University of Agriculture, Abeokuta, Nigeria.
Corresponding Author email: enkemnole@unilag.edu.ng

Article Info
Received: 17 November 2018 Revised: 12 August 2019
Accepted: 26 August 2019 Available online: 22 November 2019

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 its impact on economic growth has continued to generate controversies among economic researchers and policymakers. In this paper, a state space model approach was developed to describe the dynamics of Brent crude oil prices. The dynamics were examined as a continuous time stochastic process generalized as an Ornstein-Uhlenbeck equation. The result revealed that the dynamic behaviour of Brent oil price is an Ornstein-Uhlenbeck equation depicting a mean reversion process in crude oil prices. The process is stationary Gauss-Markov process and is the only nontrivial process that satisfies the conditions of allowing linear transformations of the space and time variables. The Ornstein-Uhlenbeck process in this paper is considered as the continuous time analogue of the discrete-time autoregressive process of order one (AR(1)).

Keywords: Stochastic process, Markov process, Random Walk model, Brownian motion, Diffusion process.

MSC2010: 62M10

1 Introduction

Crude Oil is a non-renewable fossil fuel which is distributed randomly all over in the ground [32]. It is arguably the most important driving forces of the global economy and changes in the price of oil have significant effects on economic growth and welfare around the world. 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 [28]. 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. Oil prices have

http://ijmao.unilag.edu.ng/article/view/544

historically exhibited greater levels of volatility than other commodities and assets prices [26]. Issues of oil price volatility and its impact on economic growth have continued to generate controversies among economic researchers and policy makers. Volatility is simply a characterization of price changes over time and in future markets the changes are almost continuous as they occur both within and between trading days. Volatility is not bad for derivative markets and indeed without volatility there would be little interest in many derivative products.

Derivatives have become very important tools for transferring risks from one entity to another. In advanced society, derivative markets have grown so fast in that they can be used for three different purposes such as hedging, speculations, and arbitrage. The derivatives that can be traded include forwards, futures, swaps, options and structured products.

Crude oil has emerged as one of the biggest commodity markets in the world and has been traded like any other commodity in pure physical terms at spot prices. Today crude oil is sold through a variety of contract arrangements such as futures, options, forwards and in spot transactions. Capturing the price behaviour of assets and forecasting future developments is essential in financial management and international policy. [16] and [31], concluded that oil prices move akin to a random walk without a drift term and that the best predictor of the future price of crude oil is the present oil prices.

Random walks are processes with independent increments and processes with independent increments are Markov processes. The essential idea underlying the random walk for real processes is the assumption of mutually independent increments of the order of magnitude for each point of time [15]. Random walk representation may provide another avenue of approach to the solution of the equations characterizing certain Markov processes. Random walk models have been prominent since the 1960 and the earliest random walk models assumed that the accumulation of information occurred at discrete point in time and that each piece of information is either fixed or variable in size [24, 28].

Some researchers Random walks are processes with independent increments and processes with independent increments are Markov processes. [25] verified the pattern of best market share rules' through the use of a Markov probability model. Many researchers in economics and finance have severally developed numerous theories for modelling volatility and such models as autoregressive conditionally heteroscedastic (ARCH) model was introduced by [14]. This model was extended in different directions and the most popular was the general autoregressive conditionally hetroscedastic (GARCH) as proposed by [7]. 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.

Models defined by ARCH [14] and GARCH [7] models are called stochastic volatility models or stochastic variance models and there likelihood functions are more difficult to handle. Stochastic volatility models are those in which the variance of a stochastic process is itself randomly distributed. By assuming that the volatility of the underlying crude oil price is a stochastic process rather than a constant makes it possible and accurate to model crude oil prices as a derivative using either structural models, or financial models.

A structural time series model, [19] sets out to capture the salient features of a time series data and can be written as state space model. State space models, [12, 13] and [9] are a widely used tool in time series analysis to deal with processes which gradually change over time. The state space model represents a physical system as first order coupled differential equations and is a fundamental concept in modern control theory. [23] estimated coefficient of a non-linear differential equation using an optimal sequential estimation techniques often referred to as Kalman filter. Kalman's derivation took place within the context of state space models whose core is the recursive least squares estimation. Within the state space notation, the Kalman filter derivation rests on the assumption of normality of the initial state vector, and as well as the disturbances of the system. The state of a system is defined to be a minimum set of information from the present and past such that the future behaviour of the system can be completely described by the knowledge of the present state and the future input. The State space representation is based on the Markov property, which implies that given the present state, the future of a system is independent of its past.

It is clear from historical data that while commodity may rise and fall sharply in the short and medium term, they tend to revert to an average mean over time. Understanding this underlying stochastic process is of clear importance, particularly for crude oil due to its essential role in the world economy [17, 21, 22]. The Mean-Reversion models developed by [8] and [30] are based on the assumption that oil price follows a stochastic process with known mathematical properties. [11] suggested a mean-reverting Markov switching jump diffusion model to characterize the stochastic behaviour of the crude oil spot prices. [5] argued that the mean reversion process has been considered the natural choice for commodities.

In recent years, the Ornstein-Uhlenbeck process has appeared in finance as a model of the volatility of the underlying asset prices. The Ornstein-Uhlenbeck process is a stochastic process that describes the velocity of a massive Brownian particle under the influence of friction. The process is stationary Gauss-Markov process and is the only nontrivial process that satisfies the conditions of allowing linear transformations of the space and time variables.

2 Methods

Random walks are a formalization of a path that consists of a succession of random steps. By letting P_t denote the crude oil spot price at time t, the simplest version of the random walk model without a trend is

$$P_{t+1} = P_t + \varepsilon_{t+1} \tag{2.1}$$

Where ε_{t+1} represents a white-noise term with a mean of zero, constant variance and zero auto-correlation. The changes P_t are independent from each other and can also be described with deterministic trend μ as:

$$P_{t+1} = \mu + P_t + \varepsilon_{t+1} \tag{2.2}$$

It is pertinent to note that in order to facilitate the modelling process, differencing of the model is required. In this case, the change ΔP_t is

$$\Delta P_t = P_t - P_{t-1} \tag{2.3}$$

Furthermore, the data is often seasonally differenced prior to the analysis in order to eliminate seasonal non-stationarity. To account for changing variance, the ARCH/ GARCH family of models is often used in econometrics. [14] provides an autoregressive conditional heteroscedastic (ARCH) model in which the mean and variance can be simultaneously modelled and forecasted. The ARCH (q) model using an AR (p) model for P_t is

$$P_t = \phi_0 + \phi_1 P + \cdots + \phi_p P_{t-p} + \varepsilon_t \tag{2.4}$$

where $\varepsilon_t \sim N(0, \sigma_t^2)$ and $\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_1\varepsilon_{t-2}^2 \cdots + \alpha_q\varepsilon_{t-q}^2 + \beta$
The GARCH (p,q) model using AR (p) model for P_t is

$$P_t = \phi_0 + \phi_1 P_{t-1} + \cdots + \phi_p P_{t-p} + \varepsilon_t \tag{2.5}$$

where $\varepsilon_t \sim N(0, \sigma_t^2)$ and $\sigma_t^2 = \alpha_0 + \alpha_1 P_{t-1} + \cdots + \alpha_q\varepsilon_{t-q}^2 + \beta_1\sigma_{t-1}^2 + \cdots + \beta_p\sigma_{t-p}$
An alternative to ARCH or GARCH models is to assume that σ_t^2 follows a stochastic process and this is usually done by modelling the logarithm of σ_t^2. The log (σ_t^2) follows an AR process with an error component that is independent of the σ_t^2 series in equations (4) and (5). In order to handle this process within the framework of the classical time series analysis, it can be shown that natural ways of trying to define crude oil spot price is as a Markov process . Thus in the study the crude oil spot price $P(t)$ is in a sense, the probabilistic analog of causality representing the first order AR process in continuous time, since differencing in discrete time corresponds to differentiation in continuous time. The first ? order AR process in continuous time, $P(t)$ is such that its derivative is defined as a general equation

$$aP(t) + \frac{dP(t)}{dt} = Z(t) \tag{2.6}$$

where a is a constant, and $Z(t)$ denotes continuous white noise. However, as $\{Z(t)\}$ cannot physically exist, it is more legitimate to write (6) as

$$dP(t) = -aP(t)dt + dZ(t) \tag{2.7}$$

Equation (7) in the theory of Brownian motion is a Langevin equation. In the physics literature, [27] studied this kind of stochastic differential equation usually referred to as Langevin equation. The Langevin equation of (7) is denoted as

$$P(t) = P(0) + c\int_0^t P(s)ds + \sigma\int_0^t dW_t \qquad t \in (0, T) \tag{2.8}$$

Model (8) is related to the world of time series analysis. In intuitive form (8) can be written as

$$dP(t) = cP(t)dt + \sigma W_t \tag{2.9}$$

and formally setting $dt = 1$. Then

$$P_{t+1} - P_t = cP_t + \sigma(W_{t+1} - W_t) \tag{2.10}$$

or

$$P_{t+1} = \Phi P_t + Z_t \tag{2.11}$$

where $\Phi = c + 1$ is a constant and the random variables $Z_t = \sigma(W_{t+1} - W_t)$ constitute an iid sequence of $N(0, \sigma^2)$. This is an autoregressive process of order. This time series model can be considered as a discrete analogue of the solution to the Langevin equation (8) and the Langevin equation as a linear *Itô* stochastic differential equation. The stochastic process follows a random walk and can be represented as

$$P_t = c + P_{t-1} + a_t \tag{2.12}$$

with a constant c and white noise a_t . If c is not zero then the variables,
$P_t - P_{t-1} = c + a_t$ have a non-zero mean and is called a random walk with a drift. In contrast, the random walk defined here is the boundary case for an $AR(1)$ process and is defined as

$$P_t = P_{t-1} + a_t \tag{2.13}$$

International Journal of Mathematical Analysis and Optimization: Theory and Applications
Vol. 2019, No. 2, pp. 571 - 580

Through recursive substitution

$$P_t - P_{t-1} = a_t \tag{2.14}$$

or equation $(1-B)P_t = a_t$ equation equation $P_t = (1-B)^1 a_t$ equation

$$P_t = a_t + a_{t-1} + \cdots + a_1 \tag{2.15}$$

In order to handle this process within the framework of the classical time series analysis, the observed claim process must be transformed by differencing the process in order to get a stationary process. The transform process is then

$$Z_t = (1-B)^d P_t \tag{2.16}$$

Such a model is called an integrated model because the stationary model that is fitted to the difference data has to be summed or integrated to provide a model for the original non-stationary data. Describing the difference of is said to be an process. By following the [6] three stage model building procedure of identification, estimation, and diagnostic checking, an ARIMA (p,d,q) model is developed. ARIMA (p,d,q) models are typically parsimonious model and the model selection is based on the premise that ACF and the related statistics can be accurately estimated and are stable over time.

In practice, it may turn out that there is more than one plausible model and based on the use of Akaike information criterion (AIC), the goodness of fit of different models is to be compared. The AIC is defined as

The AIC is defined as

$$\begin{aligned} AIC &= -2maximizedlog\text{-}likelihood + 2n \\ &\approx Tln\hat{\sigma}^2 + 2n + const \end{aligned} \tag{2.17}$$

where T is the length of the observed series after any differencing, n is the number of fitted parameters and is the estimated white noise variance. The model with the smallest value of the AIC is judged to be the most appropriate.

3 Results

The empirical 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. The monthly prices of the Brent crude oil prices from January 2009 to June 2017 (US $ per Barrel) is as in Table 1,

The time plot for the crude oil prices is as shown in Figure 1 and the analysis were achieved through the use of R software. The first difference is plotted in Figure 2. The usual identification techniques by examination of plots and autocorrelation functions leave little doubt about the need to difference The strong trend in the original series indicates that the original Brent oil price is nonstationary suggesting a need to differencing. The sample autocorrelation functions of the original functions and the first difference and the partial autocorrelation functions for the first difference is as shown in Table 2. The matching of the first 15 estimated sample autocorrelations and partial autocorrelations of the first difference series suggested an autoregressive model of order 1 as in Table 2. The R software package used the Akaike information criterion to provide the best fit for an autoregressive model to a set of data and is as shown in Table 3. The sample ACF and partial autocorrelation function (PACF) of the difference series suggest an AR (1) model.

International Journal of Mathematical Analysis and Optimization: Theory and Applications
Vol. 2019, No. 2, pp. 571 - 580

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

Month/Year	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	99.85	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	

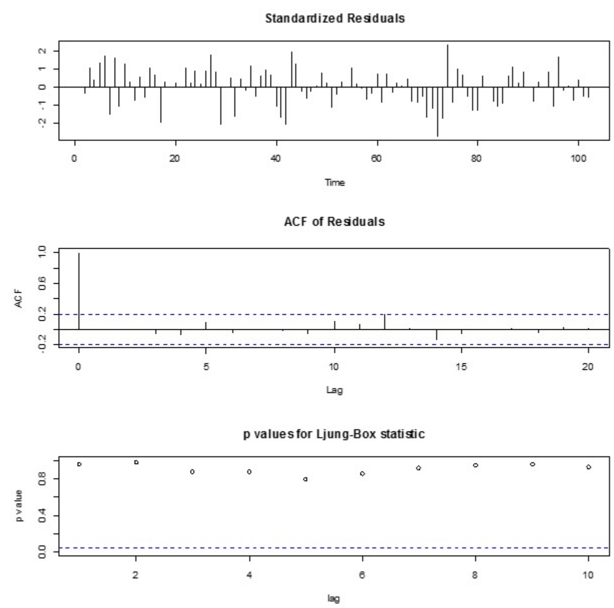

Figure 1:

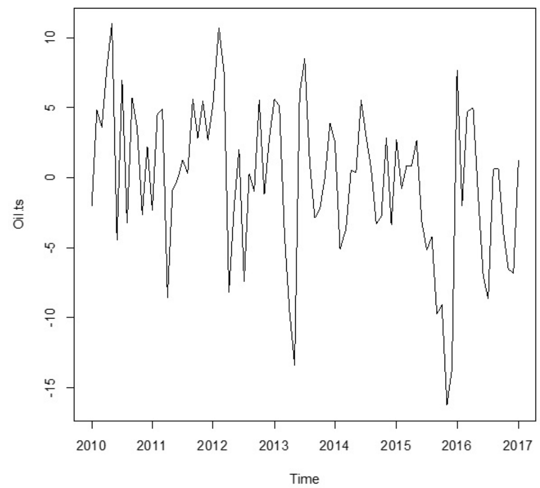

Figure 2:

Table 2: Sample Autocorrelations for P_t and δP_t and Partial autocorrelation for the Brent Oil Prices

Lag K	ACF of P_t	ACF of ΔP_t	PACF of δP_t
1	0.962	0.297	0.297
2	0.913	0.081	0.008
3	0.863	-0.056	-0.085
4	0.818	-0.057	-0.018
5	0.780	0.064	0.104
6	0.745	-0.028	-0.086
7	0.709	-0.012	0.002
8	0.673	-0.027	-0.006
9	0.639	-0.019	-0.007
10	0.606	0.130	0.139
11	0.569	0.166	109
12	0.524	0.235	0.156
13	0.472	0.041	-0.074
14	0.416	-0.137	-0.139
15	0.366	-0.089	0.002

International Journal of Mathematical Analysis and
Optimization: Theory and Applications
Vol. 2019, No. 2, pp. 571 - 580

Table 3: Estimated Value of ARIMA (p,d,q) Model for the Brent Oil Price

Model	AIC	SE	Log likelihood
ARIMA (1,0,0)	636.27	27.36	-315.13
ARIMA (1,0,1)	629.47	25.06	-310.74
ARIMA (0,1,1)	617.69	25.47	-306.85
ARIMA (0,1,0)	652.54	39.16	-325.27
ARIMA (1,1,0)	616.81	25.25	-306.4
ARIMA (1,1,1)	618.8	25.24	-306.4

From Table 3, for the ARIMA (0,1,1) model, the AIC for the process is at AIC=617.69 which is very close to the ARIMA (1,1,0). Based on the Akaike?s information criteria computed for the models, it appears that ARIMA (1, 1, 0) with an AIC =616.81 is the optimal model. ARIMA (1, 1,0) is simply a random walk model. The random walk defined here is the boundary case for an AR (1). The corresponding fitted autoregressive model is

$$\delta P_t = 0.296 P_{t-1} + a_t$$

$$(0.095)$$

An overall test of model adequacy is provided by Ljung-Box chi-squared statistics. These statistics also known as the Box-Pierce chi-square statistics contain what are known as the portmanteau statistics with their associated p-values. In fitting the AR (1) model to the state space model, none of the chi-square values is significant at the 5% level. The ARIMA model diagnostic is as shown in Figure 3 with various plots produced such as the standardized residuals, the ACF of the residuals, the PACF of the residuals, and the p-values of Ljung-Box Chi-squared statistics.

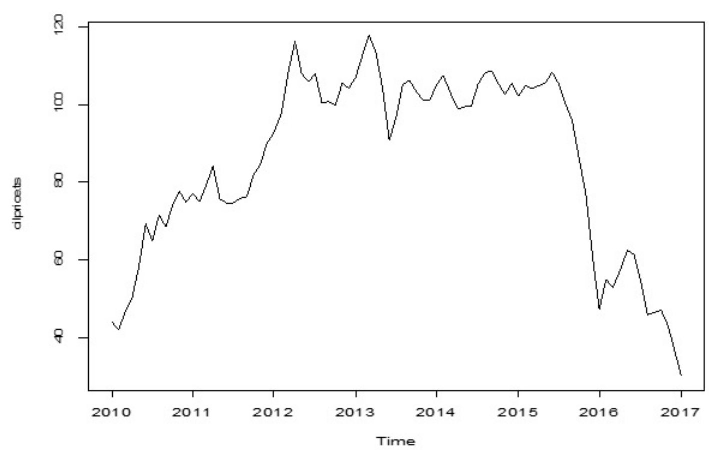

Figure 3:

4 Discussion

This study described the evolution or dynamics of crude oil prices over a given time period as a Ornstein-Uhlenbeck process. The Ornstein-Uhlenbeck process is just like the Brownian motion and is the scaling limit of simple random walk. The study demonstrated that Ornstein-Uhlenbeck processes are a Gaussian process with autocovariance which can be transformed into state space time series model. The state space model demonstrates a bridge between nonlinear stochastic dynamical systems and nonlinear time series model indicating that there is a close relationship between nonlinear time series models and nonlinear stochastic dynamical systems.

The state space model sets out to capture the salient features of the time series and these are apparent from the nature of the series. The state space model can be reduced to an autoregressive integrated moving average (ARIMA) process or an autoregressive moving average (ARMA) process.

References

[1] Gahler, S. 2-Metrische Raume und ihre topologische Struktur. *Mathematishe Nachrichten*, **26**, 115–148, (1963).

[2] Akaike, H. Autoregressive model fitting for the control. *Ann. Inst. Statist. Math.*, **23**, 163–180, (1971).

[3] Akaike, H. A New Look at statistical model identification. *I.E.E.E. Transactions*, , (1974).

[4] Barsky, R.B. & Kilian, L. Oil and the Macroeconomy since the 1970s. *Journal of Economic Perspectives*, **18**, 115-134 (2004).

[5] Benard, J., Khalaf, L., Kichian, M., & McMahon, S. Oil Prices: Heavy Tails, Mean Reversion and the Convenience Yield. *Working Paper, University of Laval and Bank of Canada*. (2008).

[6] Box, G.E.P. & Jenkins, G.M Time series analysis, forecasting and control. *(rev.ed), San Franscisco, Holden-Day*, (1976).

[7] Bollerslev, T. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, (1986).

[8] Brennan, M. & Schwartz, E. Evaluating Natural Resource Investments. *The Journal of Business*, 58, 135-157, (1985).

[9] Chatfield, C. *The Analysis of Times Series An Introduction*, 6th ed , Chapman and Hall/CRC,Boca Raton, (2004).

[10] Chung, K.L and Williams, R.J (1990). 'Introduction to Stochastic Integration', 2nd edition, Birkhauser, Basel.

[11] Dafas, P. A. Estimating the parameters of a mean-reverting Markov-switching jump-diffusion model for crude oil spot prices. $25 - 34$, (2004) .

[12] Durbin, J., & S.J. Koopman . Time Series Analysis of Non-Gaussian Observations based on State-Space Models from Both Classical and Bayesian Perspectives (with discussion). *Journal of Royal Statistical Society*, Series B, 62, 3-56, (2000).

[13] Durbin, J. & S. J. Koopman . *Time series analysis by state space methods.* Oxford University Press, Oxford. (2001).

[14] Engle R.F. . Autoregressive Conditional Heteroscedasticity with estimates of the variance of UK inflation. *Econometrica*, 50, 987-1008, (1982).

[15] Franke, J., Hardle, W. & Hafner, C. *Statistics of Financial Markets: An introduction*, Springer-Verlag, New York, (2004) .

[16] Geman, H. *Mean Reversion versions random walk in Oil and Natural Gas Prices*. University of London, United Kingdom, (2007).

[17] Hamilton, J. D. Understanding Crude Oil Prices. *Energy Journal*, 30, pp 179-206, (2008).

[18] Hamilton, J. D. Causes and Consequences of the Oil Shock of 2007-08. *Brookings paper on Economic Activity (Spring)*, 215–261, (2009).

[19] Harvey, A.C. *Forecasting, Structural Time Series Models and Kalman Filter*. Cambridge University Press, Cambridge, (1989).

[20] Hotelling, H. The Economics of Exhaustible Resources. *Journal of Political Economy*, 39, 137-175, (1931).

[21] Huntington, H. G. Oil Shocks and Real U.S. Income. *The Energy Journal*, 28(4), 31-46, (2007).

[22] Jiménez-Rodríguez, R. & Sánchez, M. *Oil price shocks and real GDP growth: empirical evidence from some OECD countries. Applied Economics*, 37, 201-228, (2005).

[23] Kalman, R.E. A New Approach to Linear Filtering and Prediction Problems. Journal of Basic Engineering Transactions Series D. 82: 35-45, (1960).

[24] Laming D.R.J Information Theory of Choice Reaction Time, Wiley, New York, (1968).

[25] Kosobud, R. F. and Stokes, H. H. Oil Market Share Dynamics: A Markov Chain Analysis of Consumer and Producer Adjustments, Empirical Economics, (1978)

[26] Lipsky, J, Economic Shifts and Oil Price Volatility. http:// www.imf.org /np speeches/ 2009/031809. (2009).

[27] Langevin, P. The theory of Brownian movement. Comp. Rend. 146, 530-532, (1908).

[28] Lin, A . Prediction of International Crude oil futures price based on GM(1,1)Gahler, S. 2-Metrische Raume und ihre topologische Struktur. Mathematishe Nachrichten, 26, 115–148, (1963).

[29] Mikosch,T. Elementary Stochastic Calculus With Finance in View, Advanced Series on Statistical Science and Applied Probability,World Scientific, New Jersey. (2008)

[30] Schwartz, E.S. The Stochastic Behaviour of commodity prices implication for valuation and Hedging. The Journal of finance, (1997)

[31] Taylor, J. and van Doren, P. The Energy Security Obsession, Georgetown Journal of Law and Public Policy, (2008)

[32] Tuo, J, and Yanbing. Summary of World Oil Price Forecasting Model, In IEEE Knowledge Acquisition and Modelling (KAM). Fourth International Symposium, (2011)