OIL CRUDE PRICE VOLATILITY: A WHITE NOISE STOCHASTIC ANALYSIS

Marvin A. Maulion, Resham Lal Sohal Jr., Romel Mark Irabon, and Exan John Carpio.
Physics Department, Western Mindanao State University, Zamboanga 7000, Philippines

ABSTRACT: *Recently, an abrupt increase in WTI crude oil price was observed in the global commodity market. Existing literature attempted to determine its associated potential harms to the economy, however, most methods did not consider the use of fluctuations as a random variable that changes over time. Thus, to establish a novel and alternative method, the researchers aimed to examine the fluctuations of standard crude oil prices, as recorded in the historical data registry of WTI, using White Noise Analysis. The analysis was carried out by calculating the mean square deviation (MSD) of 2511 data points corresponding to daily crude oil prices from 2012 to 2022. A theoretical framework for the MSD was determined and the resulting probability density function (PDF) exhibited a sinusoidal modulating function, further revealing a non-Markovian stochastic process. The results of this study will serve as baseline information in predicting the most probable crude oil price in the future.*

1. INTRODUCTION

Crude oil is a type of fossil fuel that occurs naturally in the form of petroleum products. These petroleum products are mainly composed of hydrocarbons that can be further developed into marketable byproducts such as gasoline, kerosene, diesel, and other forms of petrochemicals [1,2]. As a matter of fact, crude oil has a vast impact on our daily lives and far-reaching effects are more inclined to sectors such as food production, business, agriculture, and transportation. Thus, it is inevitable for crude oil to become the most crucial and essential gateway for world economic growth, so much that it binds together the first world and third world countries in terms of its nature to both regions.

Being such an important source of energy resource for the whole world, the quality of crude oil is determined sophistically based on its properties that can be dependent on the location and under what condition it was formed. According to the "Introduction to the Oil Industry and O.P.E.C", crude oil is classified depending on its American Petroleum Institute (API) gravity and sulfur content [3]. The parameter API gravity (light or heavy) determines the density which specifies if the crude oil has a higher or lower boiling range and serves as a critical factor in isolating and extracting the different derivatives of the oil. The sulfur content can be classed as sweet or heavy. A good indication of crude oil is sweet which has a lesser amount of sulfur content (0.5% and lower) [3,4]. Crude oil prices may heavily be dependent on the aforementioned quality constraints; thus, the best reference is the West Texas Intermediate (WTI) [5].

West Texas Intermediate is one of the four benchmarks that are prominent in worldwide trading. Crude oil here is regarded as of top-notch quality, sweet, and light that can generate gasoline and diesel fuel for global baseline commodity. The source of WTI originated in mainland Texas, USA. This benchmark posed importance in guiding buyers and sellers for the price of crude oil [5]. With the growing demand for crude oil expenditure in the United States, the likeliness of the volatility for the price of West Texas Intermediate would be seen in the foreseeable future.

In the past years, the value of the WTI (in terms of dollars per barrel) started to rise from \$67.49 in July 2007 to \$133.37 in July 2008, however, this was immediately ended as the value drastically dropped from \$133.37 to \$39.09 in February 2009. Recently, the worst value in the history of West Texas Intermediate was observed in 2020, specifically in April 2020 when it yielded \$16.55 [6]. This unfortunate event was attributed to the start of the global pandemic that occurred in March 2020 and resulted in stooping down the demand for millions of crude oil barrels reserved for that year. With the advent of vaccines, the restrictions are gradually loosening which means that the demand for West Texas Intermediate would return to its pre-pandemic mode. These fluctuations may greatly impact the state of economic growth and a sharp price plunge can devastate cash-strapped oil companies. Moreover, it can affect states that are supporting domestic WTI prices as fluctuations of the international prices can be transferred into the volatility in the actual government stream of expenditures. This oil price volatility can cause many corporations to postpone investments as they wait to see where pricing levels will settle and may cause businesses to have difficulties in adapting to fast-paced cost adjustments and slowing short-run responses.

Several studies were conducted to determine the volatility of the West Texas Intermediate crude oil prices such as that of Ye et al. [7], who proposed an analysis for the monthly average price of WTI using relative inventories. On the other hand, Yu et al. [8] utilized the method of a neural network to study the daily end price of two benchmarks namely, WTI and Brent crude oil. A similar daily return forecasting was studied by Agnolucci [9] using GARCH and volatility models and lastly, Angela et al. [10] forecasted the volatility of the WTI crude oil daily highs and lows of the price using an empirical model. The fluctuations of WTI crude oil daily prices can be considered as a random variable that evolves through time, and such an event is referred to as a stochastic process. This paper ultimately aims to mathematically model the behavioral pattern of the WTI daily crude oil price volatility using an existing set of calculus-based established methods in white noise analysis.

2. METHODOLOGY

This study will utilize the daily West Texas Intermediate (WTI) Crude oil prices raw data retrieved from fred.stlouisfed.org and sourced from U.S. Energy Information Administration. The data contains the daily average of the closing spot prices (2511 data points) of WTI crude oil in dollars per barrel from 2012 to 2022. The data collected will be analyzed using the method of White-noise analysis for Brownian motion with memory function. The step-by-step highlights of the method [13] can be explained as follows:

1. The raw data of WTI crude oil price, a function of time and denoted as P(t), was plotted using Python software.
2. A best linear fit of the raw data was obtained, and a list of coordinates was generated.

3. The fluctuation of crude oil prices was computed by calculating the difference between the raw data points and the linear fit coordinate points.
4. The Mean Square Deviation (MSD) of the empirical data was quantified and eventually, the points (log(time), log(MSD)) were plotted and named as the graph of empirical MSD.
5. Using a memory function, a theoretical MSD was modeled to fit the empirical MSD data. Moreover, a better fit for the fluctuation behavior was obtained by introducing a modulating function.
6. Lastly, the associated probability density function was determined using the final form of the theoretical MSD.

A stochastic random variable $x(t)$ can be represented as the sum of the initial fixed point x_0 and some Brownian fluctuations $B(t)$ parameterized within random white noise variable $\omega(t)$ given by the expression [11],

$$x(t) = x_0 + B(t) = x_0 + \int_0^T \omega(t)\, dt \qquad (1)$$

To establish a more realistic approach for a real-time scenario like natural processes, we apply a non-Markovian evolution on a fluctuating quantity $x(t)$ by allowing it to have a memory of the past. To achieve this, we incorporate a memory function modulating the ordinary Brownian motion that would affect the path $x(t)$. This time, equation (1) can now be rewritten as [12],

$$x(t) = x_0 + \int_0^T f(\tau - t)\, h(t)\omega(t)\, dt \qquad (2)$$

where $f(\tau - t)$ is the memory function, and $h(t)$ is a function of time. In general, these terms can be determined explicitly depending on the type of phenomenon being modeled. Considering all the possible fluctuation histories, from the path $x(t)$ in between initial point x_0 at $\tau = 0$, and the latest path x_T at $\tau = T$, the probability density function of this fluctuation with strong correlation can be evaluated as the integral of the delta function over the Gaussian white noise measure $d\mu_\omega$ derived in [11],

$$P(X_T, T;\, X_0, 0) = \int \delta(x(T) - x_T) d\mu_\omega \qquad (3)$$

Plugging in equation (2) to equation (3), the resulting probability density function of a non-Markovian process [12] can be written as,

$$P(X_T, T;\, X_0, 0) = \int \delta\left(x_0 - x_T + \int_0^T f(\tau - t) \times h(t)\omega(t)\, dt \right) \qquad (4)$$

Transforming the delta function in terms of its Fourier representation, the probability density function [12] can be expressed as,

$$P(X_T, T;\, X_0, 0) = \frac{1}{2\pi} \int \int_{-\infty}^{+\infty} \exp\big(ik(x_0 - x_T)\big)\, dk \cdot exp\left(ik \int f(\tau - t)\, h(t)\omega(t)\, dt \right) d\mu_\omega \qquad (5)$$

Furthermore, using the characteristic function [12],

$$C(\xi) = \int \exp\left(i \int_0^T \omega(t)\, \xi(\mathrm{t}) dt \right) = \exp\left(i \int_0^T \xi(\mathrm{t})^2 dt \right) \qquad (6)$$

The probability density function [13] can be expressed as,

$$P(X_T, T;\, X_0, 0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\big(ik(x_0 - x_T)\big)\, dk \times exp\left(-\frac{1}{2} k^2 \int_0^T [\, f(\tau - t)\, h(t)]^2 dt \right) \qquad (7)$$

Equation (7) is a gaussian integral and can further be simplified into a much simpler form. Therefore, the final expression of the probability density function for the non-Markovian process is,

$$P(X_T, T;\, X_0, 0) = \sqrt{\frac{1}{2\pi \int [g(T) f(\tau - t) h(t)]^2 dt}} \times exp\left[\frac{-(x_0 - x_T)^2}{2 \int [g(T) f(\tau - t) h(t)]^2 dt} \right]. \qquad (8)$$

Here, further simplification can still be done after explicitly deriving an equation in terms of the Mean Square Deviation (MSD). MSD is the measure of the degree of deviation from the mean value and is equal to the difference of the second moment $\langle x^2 \rangle$ and the square of the first moment $\langle x \rangle$ given by the formula [12],

$$MSD = \langle x^2 \rangle - \langle x \rangle^2 \qquad (9)$$

where

$$\langle x \rangle = \int_{-\infty}^{+\infty} x\, P(X_T, T;\, X_0, 0)\, dx; \;\; and \qquad \langle x^2 \rangle = \int_{-\infty}^{+\infty} x^2\, P(X_T, T;\, X_0, 0)\, dx \qquad (10)$$

Evaluating the first moment equation (8) will give us $\langle x \rangle = x_0$, while the second moment takes the form,

$$\langle x^2 \rangle = {x_0}^2 + \int [g(\mathrm{T}) f(\tau - t) h(t)]^2 dt \qquad (11)$$

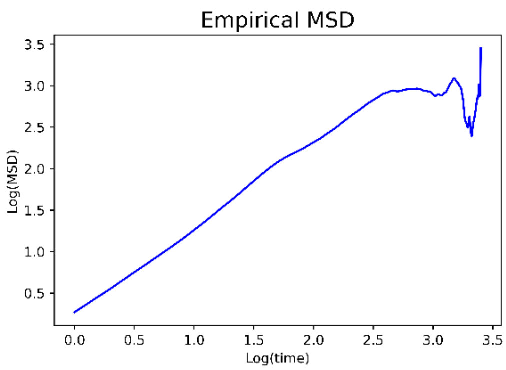
Thus, by equation 9, a generic expression of MSD yields into

$$MSD = \int [g(T)f(\tau - t)h(t)]^2 dt \tag{12}$$

Although equation (8) is the general form of the probability density function $P(X_T, T;\ X_0, 0)$ with memory function $f(\tau - t)$ and $h(t)$, following equation (9), we can simplify it much further in the form,

$$P(X_T, T;\ X_0, 0) = \sqrt{\frac{1}{2\pi MSD} exp\left[\frac{-(x_0 - x_T)^2}{2MSD}\right]} \tag{13}$$

Figure 1. Graphical Representation for Crude Oil Price and



Fluctuation from 2012 – 2022

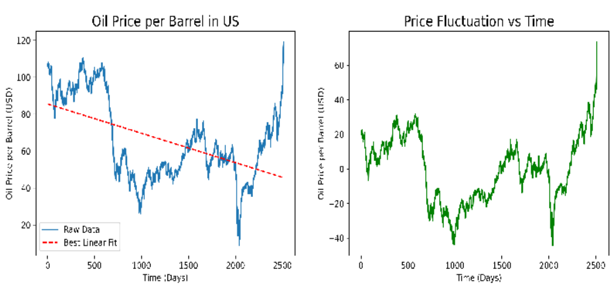


Figure 2. Graph of Crude Oil price vs Time where the blue curve represents the raw market price, the red line represents the best linear fit, and the green line accounts for the price fluctuation

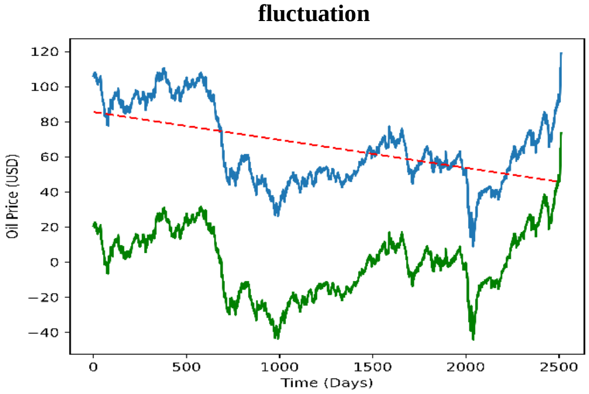


Figure 3. Empirical Log-log plot of Price Fluctuation MSD as a function of Time

3. RESULTS AND DISCUSSIONS

This section presents the entirety of the white noise analysis done for crude oil price per barrel (WTI) from 2012 – 2022 in the United States. The core analysis technique was carried out through an ASCII CSV file using a Python notebook with *NumPy* and *Matplotlib. pyplot* modules in a Jupyter environment.

3.1. Empirical Data Plotting

The set of data points was downloaded from a reputable open access economic domain [6]. A best linear fit was generated in Microsoft Excel and was determined to be **F(t)** = -0.0159**t** + 85.561, where **F(t)** is the price of crude oil per barrel in dollars at a given time **t** in days. On the other hand, the price fluctuation as a function of time was calculated by subtracting **F(t)** from the original set of data points. Because the source file is in .xlsx format, it was converted into a CSV file for easy data source coding and efficient memory processing in Python.

To gain a definite sense of how crude oil price fluctuation and its raw price are related to each other, a graphical representation is shown below for reference. Even a downward linear trend was returned in Excel, as manifested by an abrupt decrease of crude oil price near the 2000th day, a significant increase in the price was observed shortly after it. This is worth considering as the peak price in 2022 exceeded the initial price in 2012.

The MSD for price fluctuations was also calculated. However, a Log-log scale plot was generated to obtain a smoother graph. That is, a simple algebraic translation of (Time, MSD) → (Log(Time), Log(MSD)) was employed, resulting in a change of axes scales.

3.2. Theoretical MSD Modelling

The first few logs (time) that lie within the domain of Figure 3 exhibit a behavior similar to the MSD plot of Bitcoin prices [13]. With such a notion, a good estimate of the memory function combination [12] is

$$f(T - t) = (T - t)^{\frac{\mu - 1}{2}} \tag{14}$$

$$h(t) = \frac{e^{-\beta/2t}}{t^{(\mu+1)/2}} \tag{15}$$

where μ and T are both greater than zero and real. Directly substituting equations (14) and (15) to (12), the theoretical MSD can be modeled as

$$MSD = [g(T)]^2 \int_0^T \left[\frac{(T - t)^{\frac{\mu - 1}{2}} e^{-\beta/2t}}{t^{(\mu+1)/2}}\right]^2 dt \tag{16}$$

This can be written more conveniently as

$$MSD = [g(T)]^2 \int_0^T t^{-(\mu+1)} \ (T-t)^{\mu-1} \times e^{-\frac{\beta}{t}} \, dt \tag{17}$$

The integral in equation (17) can be easily evaluated by using equation 3.471.3 which can be found in the compilation of definite integrals [14] involving exponential functions. Thus, we get

$$MSD = [g(T)]^2 \ \beta^{-\mu} T^{\mu-1} \Gamma(\mu) \, exp(-\beta/T) \tag{18}$$

$$MSD = [g(T)]^2 \ \frac{\Gamma(\mu) \ T^{\mu-1}}{\beta^{\mu}} \ exp(-\beta/T) \tag{19}$$

To obtain a sinusoidal behavior on the log(time) interval approximately at (3.0, 3.4) with increasing amplitude, a modulation function $g(T)$ was adapted from [13] with slight modifications. An ansatz was determined to be

$$[g(T)]^2 = e^{\frac{-(bT^n+1)\sin(cT)}{\mu}} \tag{20}$$

where b, c, μ are parameters and $n \in \mathbb{Z}^+$. A corresponding polynomial fit for selected n values is shown for comparison.

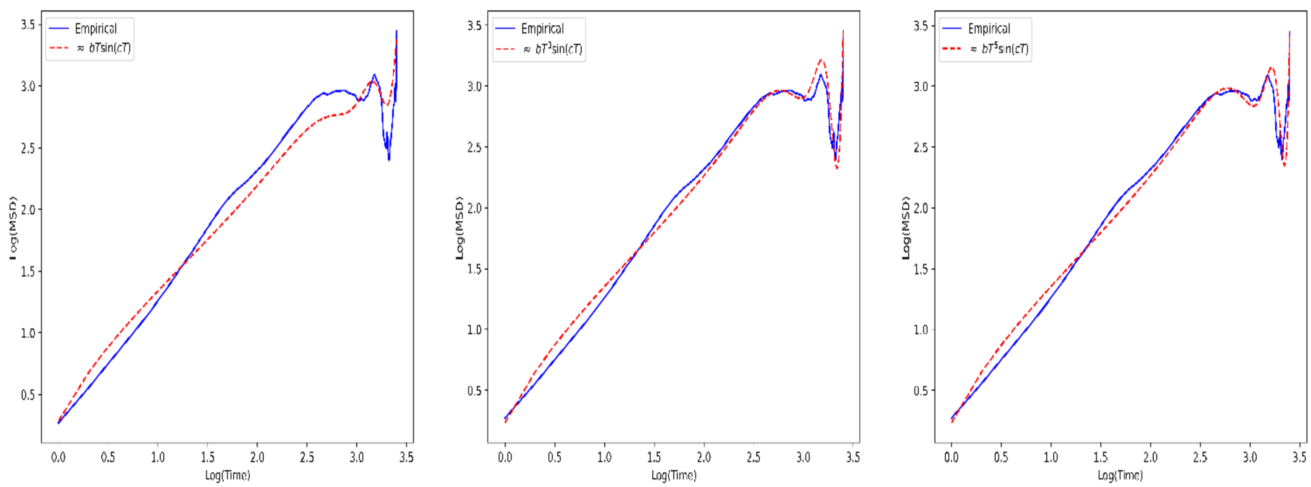


Figure 4. From *Left* to *Right*: Polynomial Ansatz Fit with $n = 1, 3$, and 5 in Equation (20)

From the three polynomial ansatzes in Figure 4, the best fit was observed at $n = 5$, which characterizes a quintic sinusoidal amplitude. From this, the theoretical MSD, by equations (8) and (9), can be written as

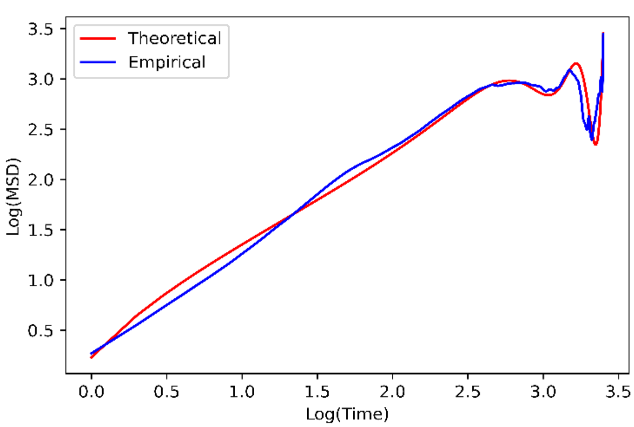


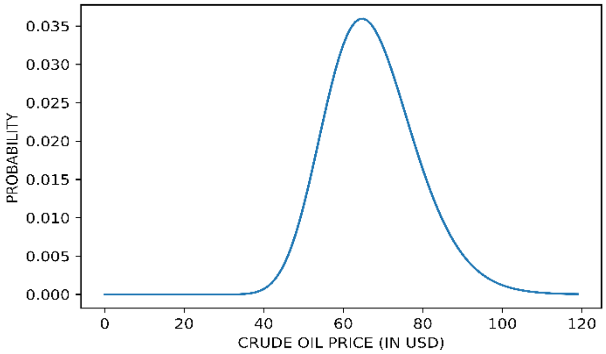
Figure 5. Theoretical MSD Fitting for Crude Oil Price Fluctuation per Barrel with $n = 5$.

$$MSD = \frac{e^{\frac{-(bT^3+1)\sin(cT)}{\mu}} \Gamma(\mu) \ t^{\mu-1}}{\beta^{\mu}} \ exp(-\beta/t) \tag{21}$$

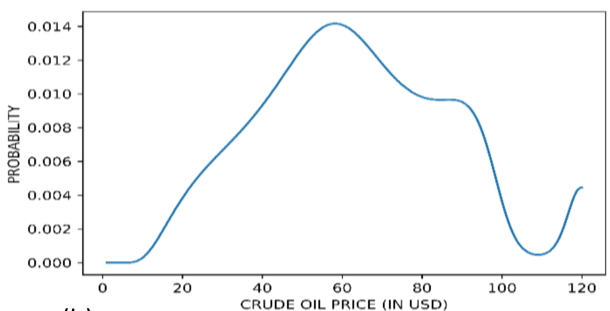
where $\beta = 0.8$, $\mu = 1.8$, $b = -1.0 \times 10^{-16}$, and with $c = -0.0037826$.

From a series of trial and error, the polynomial ansatz fit with $n = 3$ and 5 has a good visual fitting. However, to determine which among the two cases is more accurate, the researchers compared the generated fit (theoretical MSD) with the empirical MSD statistically by calculating the corresponding Standard error of the estimate. For $n = 3$, the error is approximately 0.036 while for $n = 5$, an error of ~0.0145 was noted. Thus, taking into account a relatively lower standard error, $n = 5$ serves as a better fit.

Therefore, the form of the probability density function (PDF) can be modelled as



(a)



(b)

Figure 6. Probability Density Function of Crude Oil Price per barrel for (a) approximately the first quarter of 2012, and (b) ten years (2012-2022), with $x_0 \approx \$65.60$.

$$P(x_T,T;x_o,0) = \frac{\beta^{\frac{\mu}{2}} e^{\frac{\beta}{2T}}}{\sqrt{2\pi\Gamma(\mu)T^{\mu-1}\, exp\left[\left(\frac{-(bT^3+1)\sin(cT)}{\mu}\right)\right]}} \times exp\left(-\frac{\beta^\mu e^{\beta/T}(x_o-x_T)^2}{2\left[\Gamma(\mu)T^{\mu-1}\left(e^{\frac{-(bT^3+1)\sin(cT)}{\mu}}\right)\right]}\right) \quad (22)$$

The PDF in (10) can be used to derive other meaningful quantities such as the first moment or expectation value. If desired, the expression for the expectation value from point a to point b can be determined using

$$< x >= \frac{x_0}{\sqrt{4\pi}}[erf(b-x_o)\,k - erf(a-x_0)\,k] - \frac{1}{4k\pi}\left[e^{-k^2(b-x_0)^2} - e^{-k^2(a-x_0)^2}\right] \quad (23)$$

where $k^2 = \frac{1}{2MSD}$ and erf is the standard mathematical symbol for the error function.

4. CONCLUSION AND RECOMMENDATION

The primary goal of this study was to develop a novel mathematical method that would analyze the fluctuations of the time-series data involving the daily crude oil price of West Texas Intermediate (WTI). The White- noise analysis for Brownian motion with memory function was utilized in achieving all the significant objectives in this study. The researchers were able to describe the behavioral pattern of the fluctuations of the daily crude price of WTI by graphing the empirical mean square deviation (MSD) against time. The empirical MSD plot was compared to known theoretical MSD (non-Markovian) laid down in [12]. Upon trying several theoretical MSDs listed in [12], it was shown that the empirical MSD showed similarities with theoretical MSD expressed as;

$$MSD = \frac{e^{\frac{-(bT^3+1)\sin(cT)}{\mu}}\Gamma(\mu)\,T^{\mu-1}}{\beta^\mu}\, exp(-\beta/T)$$

where the constant parameters were expressed as $\beta = 0.8$, $\mu = 1.8$, $b = -5 \times 10^{-10}$, and $c = -0.0038$. The probability density function attained in (22) provided us the expectation value from any arbitrary points a to b. For the betterment of the study, it is recommended to determine the diffusion coefficient for the fluctuation to predict behavior on a long-time basis.

REFERENCES

[1] Levine, S., Taylor G., Arthur D., and Tolleth M., “Understanding Crude Oils and Product Markets”, *American Petroleum Institute*, 2014.

[2] Phelan J., “What is crude oil?” Retrieved from https://www.livescience.com/, accessed on March 15, 2022.

[3] OPEC Secretariat Public Relations and Information Department, “I need to know: An Introduction to the Oil Industry & O.P.E.C, 2nd Edition”, *Ueberreuter Print GmbH”* (2013).

[4] Corporate Finance Institution, “Crude Oil Overview”, retrieved from https://corporatefinanceinstitute.com/, accessed on March 15, 2022.

[5] Chen J., “West Intermediate Texas (WTI)”, retrieved from https://www.investopedia.com/, accessed on (March 15, 2022).

[6] Federal Reserve Economic Data, “Crude Oil Prices: West Texas Intermediate (WTI) – Cushing, Oklahoma”, retrieved from https://fred.stlouisfed.org/. Accessed on (March 18, 2022).

[7] Ye M., Zyren J., and Shore J., “A monthly crude oil spot price forecasting model using relative inventories”, *International Journal of Forecasting* **21**, pp. 491-501, (2005).

[8] Yu L., Wang S., and Lai K.K., “Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm”, *Energy Economics* **30(5)**, pp. 2623–2635 (2008).

[9] Agnolucci P., “Volatility in crude oil futures: A comparison of the predictive ability of GARCH and implied volatility models”, *Energy Economics* **31(2)**, pp.316-321 (2008).

[10] He A., Kwok J.T., and Wan A.T., “An empirical model of daily highs and lows of West Texas Intermediate crude oil prices”, *Energy Economics* **32(6)**, pp. 1499-1506 (2010).

[11] Bernido C.C. and Carpio- Bernido M.V., “White noise analysis: Some applications in complex systems, biophysics and quantum mechanics”, *International Journal of Modern Physics* **26(29)**, pp. 3-4. (2012)

[12] Bernido C.C. and Carpio- Bernido M.V., “Methods and Applications of White Noise Analysis in Interdisciplinary Sciences”, *World Scientific*. (2015)

[13] Corcino R., Casas K.P., and Elnar A.R., “Bitcoin: A Non-Markovian Stochastic Process”, *Journal of Science, Engineering and Technology* **6**, pp. 287-298 (2018).

[14] Gradshteyn I.S., and Ryzhik I.M., “Table of Integrals, Series, and Products”, *Academic Press, New York* (1980).