

Article

A Mutual Information-Based Network Autoregressive Model for Crude Oil Price Forecasting Using Open-High-Low-Close Prices

Arash Sioofy Khoojine ¹, Mahboubeh Shadabfar ^{2,*} and Yousef Edrisi Tabriz ³¹ Faculty of Economics and Business Administration, Yibin University, Yibin 644000, China² School of Economics and Management, Nanjing University of Science and Technology, Nanjing 210094, China³ Department of Mathematics, Payame Noor University, Tehran 19395-3697, Iran

* Correspondence: mahboubeshadabfar@yahoo.com

Abstract: The global financial markets are greatly affected by crude oil price movements, indicating the necessity of forecasting their fluctuation and volatility. Crude oil prices, however, are a complex and fundamental macroeconomic variable to estimate due to their nonlinearity, nonstationary, and volatility. The state-of-the-art research in this field demonstrates that conventional methods are incapable of addressing the nonlinear trend of price changes. Additionally, many parameters are involved in this problem, which adds to the complexity of such a prediction. To overcome these obstacles, a Mutual Information-Based Network Autoregressive (MINAR) model is developed to forecast the West Texas Intermediate (WTI) close crude oil price. To this end, open, high, low, and close (OHLC) prices of crude oil are collected from 1 January 2020 to 20 July 2022. Afterwards, the Mutual Information-based distance is utilized to establish the network of OHLC prices. The MINAR model provides a basis to consider the joint effects of the OHLC network interactions, the autoregressive impact, and the independent noise and establishes an intelligent tool to estimate the future fluctuations in a complex, multivariate, and noisy environment. To measure the accuracy and performance of the model, three validation measures, namely, RMSE, MAPE, and UMBRAE, are applied. The results demonstrate that the proposed MINAR model outperforms the benchmark ARIMA model.

Keywords: crude oil price; west texas intermediate; network autoregressive model; forecasting; mutual information; correlation matrix; multivariate time series

MSC: 62P05; 91B84

Citation: Sioofy Khoojine, A.; Shadabfar, M.; Edrisi Tabriz, Y. A Mutual Information-Based Network Autoregressive Model for Crude Oil Price Forecasting Using Open-High-Low-Close Prices.

Mathematics **2022**, *10*, 3172. <https://doi.org/10.3390/math10173172>

Academic Editors: Olimpia Neagu and Mihaela Simionescu

Received: 3 August 2022

Accepted: 29 August 2022

Published: 3 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Background, Motivations, and Challenges

One of the most significant elements that profoundly influence global financial markets is the price of crude oil, and its price movements always capture the attention of financial experts and market participants [1]. The nonlinearity, nonstationarity, and high volatility of crude oil prices make them a complex and vital macroeconomic indicator to estimate [2,3]. Several elements affect crude oil prices, causing them to be highly nonstationary and nonlinear, including climate change, market fluctuations, innovations, demand for oil, and replacing alternative fuels with energy sources [4]. Moreover, oil prices play a significant role in determining how governments and companies plan for the future [5]. There is considerable interest among policymakers and central banks in many economies in making reliable predictions to ensure they can respond effectively to any fluctuations in oil prices [6].

Several methods have been presented in decades-long research on crude oil price forecasting. However, various conventional econometric models have been developed in

different studies [7,8]. Traditional econometric models can be classified into three categories based on a combination of naive models, exponential smoothing models, and autoregressive family models [9,10], such as the autoregressive integrated moving average (ARIMA) model [11], the autoregressive conditional heteroskedasticity model (ARCH) [12], and the generalized autoregressive conditional heteroskedasticity model (GARCH) [13]. Due to their good performance, autoregressive family models are widely used for modeling and predicting crude oil price volatility [14]. For predicting crude oil prices, traditional econometric models provide a robust framework that is well-established. However, conventional models do not adequately explain crude oil prices nonlinearity [15,16].

Artificial Intelligence (AI) models are another approach to modeling and predicting asset prices, especially crude oil prices [17]. These methods, such as neural fuzzy and artificial neural networks, overcome the limitations of classical and econometric models [18,19]. Many experiments have found that the AI models often had some advantages; however, these models also have shortcomings and disadvantages [20]. Artificial Neural Network (ANN), for instance, is prone to local minima and overfitting [21,22], while other machine learning models, including Support Vector Machine (SVM) and Gaussian processes (GPs), are sensitive to parameter selection [11,23].

Additionally, sentimental models are developed to address nonlinearity issues [24]. However, adding more variables and components improve the sentimental model's performance. For example, the decision tree model performed better on when crude oil demand and supply, monthly GDP, and CPI are factored [7]. It is also suggested the significance of news texts in crude oil price forecasting, and even used Twitter posts regarding US foreign policy as an input variable [25]. Nevertheless, the performance of the sentimental model is highly dependent on the quality of its variables, and selecting the appropriate variables requires much effort [19]. Moreover, many attempts have been made to predict crude oil price by combining the methods mentioned above and signal decomposition analysis [26].

Considering the complexity and nonlinearity of the crude oil price movement, the methods mentioned above are not suitable for modeling it accurately and require a lot of parameters to be estimated, which results in complex and cost-effective calculations [27–29]. In addition, some new methods still require more research and development to determine and choose the correct parameters and features [30].

We have attempted to solve the shortcomings mentioned above with the method we developed in this study. The most crucial challenge in modeling and predicting crude oil prices is selecting a tool that can determine the nonlinearity of price movements without imposing complex calculations [31]. One of these tools, which is a helpful approach in modeling complex systems, is mutual information (MI), which is related to the concept of entropy. Mutual Information describes the expected information contained in a random variable and it is capable of capturing nonlinearity [32–34].

Another issue usually overlooked in oil price modeling is applying the open, high, low, and close prices (OHLC) of a day to model the price fluctuations [35]. In addition, knowing the daily price ranges and the low and high prices of each day enhances our perception of the price market; therefore, we utilize the open, high, and low prices in this study to model the close price.

This paper presents a method for creating a network of open, high, low, and close prices (OHLC) using mutual information and then forecasting the close price using an autoregression method. The model is based on the previous day's closed price, the joint MI correlation of OHLC prices, and the independent noise, which is determined by other factors. Comparatively, with other multivariate time series modeling approaches such as vector autoregressive (VAR), which model each time series individually, relationships across variables are usually lost. The constructed model tries to handle these issues. In addition, unlike multivariate time series models, in which the number of unknown parameters increases as the number of variables increases, the number of unknown parameters in a network autoregressive model remains constant [36,37].

The unique features and the main contributions of this study are as follows:

1. Building a model for forecasting the crude oil price based on the simplicity and non-heavy calculations due to having a few unknown model parameters.
2. Deploying the network of daily OHLC prices to model the actual close price.
3. Using a nonlinear modeling approach by applying mutual information distances.

This paper is structured as follows: The remaining part of Section 1 discusses the analyzed data and its descriptive statistics. Section 2 introduces the methodology and construction of the model and validation methods. An illustration of the modeling results is provided in Section 3. Section 4 discusses choosing an appropriate threshold for the adjacency matrix and comparing the developed model with similar approaches. This paper concludes by summarizing its contributions and suggesting future research directions.

1.2. Data and Descriptive Statistics

In this study, we use the crude oil price historical data from the West Texas Intermediate (WTI) crude oil prices. The WTI is considered one of the leading oil benchmarks for crude oil buyers sourced from the United States. A benchmark in the oil market is a price that serves as a reference price for buyers and sellers of crude oil and plays an essential role in the global oil price [38].

In order to formulate the model used in this study, as was mentioned previously, we need the open, high, low, and close (OHLC) daily prices for conducting the model. We collect the OHLC daily WTI crude oil prices from the US energy information administration (EIA) from the following website: (<http://www.eia.doe.gov/> (accessed on 21 July 2022)). From 1 January 2020 to 20 July 2022, a sampling period of 676 observations is collected. The study divides the sample data into two subsets: train and test sets. As a result, the data between 1 January 2020 and 20 January 2022, has been treated as a train set, with 542 observations accounting for 80% of the total samples used for model training. A total of 134 observations were used in the test set, which covers the period between 21 January 2022 and 20 July 2022.

The collected data is cross-checked by other publicly available resources to ensure correctness and accuracy. A description of the prices can be found in Table 1, which contains the OHLC price summary statistics. The historical prices of OHLC and the return of close data are depicted in Figures 1 and 2.

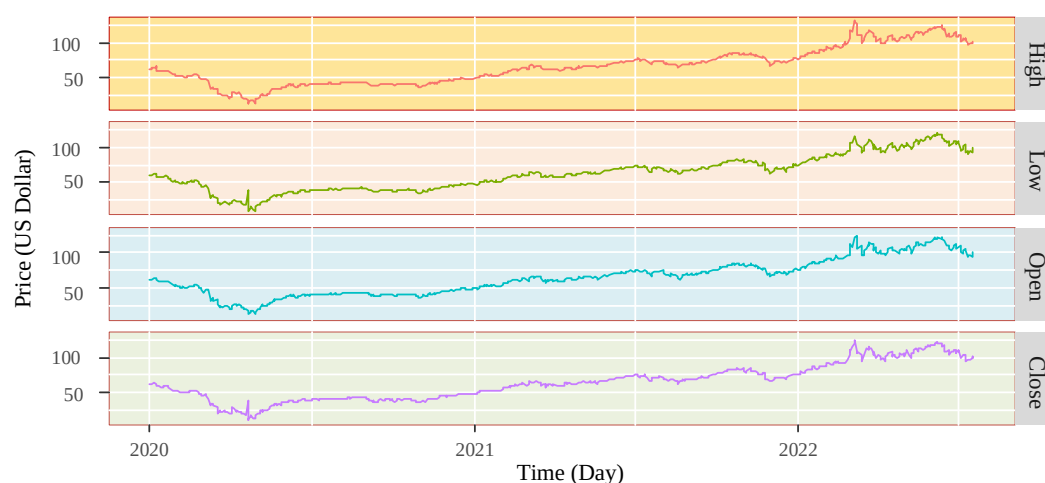
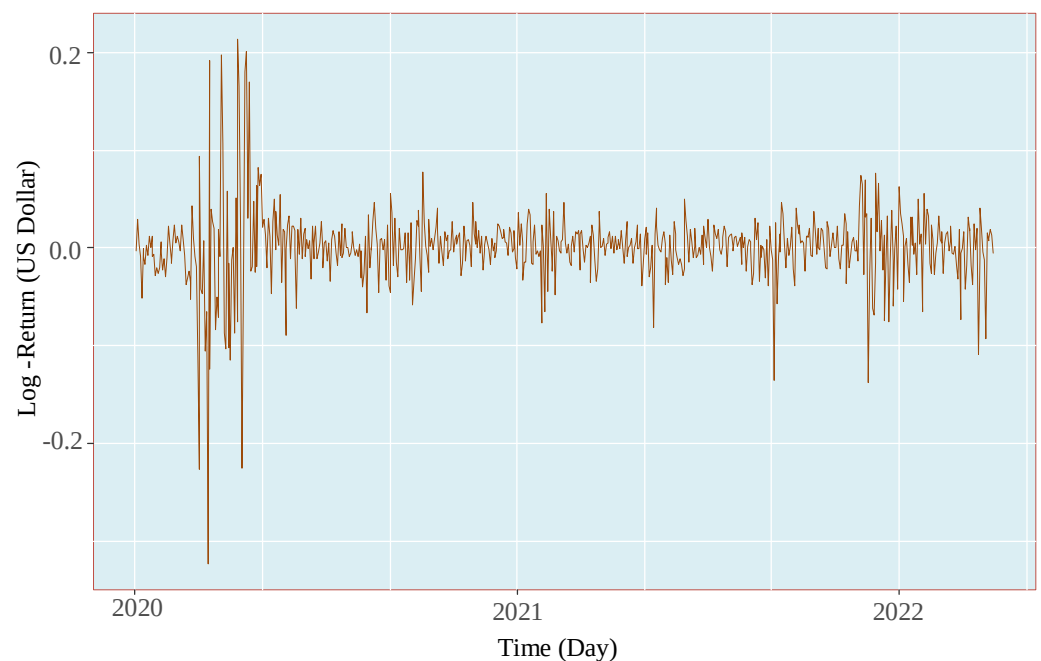


Figure 1. Historical OHLC prices of WTI in US dollar from 1 January 2020 to 20 July 2022.

Table 1. Summary statistics of OHLC crude oil prices of WTI in US dollar from 1 January 2020 to 20 July 2022.

	Open Price	Low Price	High Price	Close Price
Min.	12.96	10.07	13.69	10.01
1st Qu.	41.98	41.47	42.81	41.95
Median	63.39	62.11	64.53	63.38
Mean	64.33	63	65.66	64.41
3rd Qu.	79.48	78.44	80.86	79.41
Max.	123.7	120.79	130.50	123.70

**Figure 2.** Log-Return of the close WTI crude oil price in US dollar from 1 January 2020 to 20 July 2022.

2. Methodology

2.1. Model Construction

In this section, the Network Autoregressive model based on Mutual Information is presented for close oil prices of WTI, hereafter referred to as MINAR, and its characteristics are further described. First, we transform the collected data into a new time series. The log-return for each of the OHLC data sets is calculated as follows:

$$p_{it} = \log \left(\frac{\text{oil price } (i, t)}{\text{oil price } (i, t-1)} \right) \quad (1)$$

where $i = 1, \dots, 4$ refers to close, open, low, and high oil prices, respectively, while t refers to each time spot during the specified period.

Suppose, p_{it} represents the transformed data as calculated in Equation (1) at time t and $i = 1, \dots, 4$ refers to close, open, low, and high prices, respectively, $\mathbf{A} = (a_{ij})_{4 \times 4}$ shows the adjacency matrix of the correlation between the log-returns of transformed OHLC prices, n_i is the sum of the i th row at the adjacency matrix $\mathbf{A}$, $\beta = (\beta_1, \beta_2, \beta_3)$ is the model parameter and ϵ_{it} follows the normal distribution; hence, the MINAR model can be formulated as presented in Equation (2).

$$p_{it} = \beta_1 + \beta_2 n_i^{-1} \sum_{j=1}^4 a_{ij} p_{j(t-1)} + \beta_3 p_{i(t-1)} + \epsilon_{it} \quad (2)$$

2.2. Mutual Information Based Cross-Correlation

Inherently, the movement of oil prices is a nonlinear phenomenon [39], so it is not feasible to apply linear correlations such as Pearson correlations for the constructed OHLC network. We propose to use the transformed data created by Equation (1) in order to identify nonlinear dependencies in oil price movements by using mutual information. To determine the unpredictability of a random variable, Shannon's entropy, which measures the degree of mutual information, be taken into consideration. According to this definition, mutual information can be defined on the basis of two discrete random variables, X and Y , calculated as follows:

$$I(X, Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (3)$$

$p(x, y)$ is the joint probability distribution function of X and Y , and $p(x)$ and $p(y)$ are the marginal probability distributions.

In the case of mutual information, it could be expressed as follows:

$$I(X, Y) = H(X) + H(Y) - H(X, Y) \quad (4)$$

where $H(X)$ is Shannon's entropy, which is defined as:

$$H(X) = - \sum_i p(x_i) \log p(x_i) \quad (5)$$

Similarly, $H(X, Y)$ represents the joint entropy associated with the two variables together. Furthermore, $\log$ represents the logarithm to the base of two.

To estimate Shannon's entropy for practical purposes, we use Schürmann–Grassberger, a Bayesian parametric estimator. Based on the Dirichlet probability distribution, this estimator has been developed. Its density function is defined as follows:

$$p(x; \theta) = \frac{\prod_{i \in \{1, 2, \dots, |\chi|\}} \Gamma(\theta_i)}{\Gamma(\sum_{i \in \{1, 2, \dots, |\chi|\}} \theta_i)} \prod_{i \in \{1, 2, \dots, |\chi|\}} x_i^{\theta_i - 1} \quad (6)$$

θ_i represents the prior probability of x_i , which is the i th element in the set $|\chi|$, and $\Gamma(\cdot)$ is the gamma function.

Using the following equation, one can calculate the entropy of a Dirichlet distribution:

$$\hat{H}^{dir}(X) = \frac{1}{m + |\chi|N} \sum_{x \in \chi} (D_x + N) (\psi(m + |\chi|N + 1) - \psi(D_x + N + 1)) \quad (7)$$

Assume $\psi(z) = \frac{d}{dz} \ln(\Gamma(z))$ is the digamma function, D_x represents the number of data points with value x , $|\chi|$ represents the number of bins from the discretization step, and m represents the sample size.

Then, the mutual information between X and Y is converted into a Euclidean metric, which is the difference between the joint Shannon entropy of X and Y and the mutual information between both of them. This measure is as follows:

$$\delta(X, Y) = H(X, Y) - I(X, Y) \quad (8)$$

Therefore, we can write the cross-correlation matrix of OHLC prices as follows:

$$\delta = \begin{bmatrix} \delta_{11} & \delta_{12} & \delta_{13} & \delta_{14} \\ \delta_{21} & \delta_{22} & \delta_{23} & \delta_{24} \\ \delta_{31} & \delta_{32} & \delta_{33} & \delta_{34} \\ \delta_{41} & \delta_{42} & \delta_{43} & \delta_{44} \end{bmatrix} \quad (9)$$

where δ_{ij} 's are constructed correlation distances in Equation (8), and $i, j = 1, \dots, 4$ are close, open, low, and high prices respectively.

Suppose $\mathbf{A} = (a_{ij})_{(4 \times 4)}$ is the adjacency matrix of the correlations among OHLC prices; by adjusting a threshold as θ , $-1 \leq \theta \leq 1$, the matrix $\mathbf{A}$ is defined as follows [40]:

$$a_{ij} = \begin{cases} 1 & \text{if } i \neq j \text{ and } \delta_{ij} \geq \theta \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

Choosing an appropriate threshold will be discussed in Section 4.

2.3. MINAR Model Features

Consider $\mathbf{A} = (a_{ij})_{(4 \times 4)}$ as the adjacency matrix of OHLC price correlations constructed by Equation (10). An adjacency matrix, $\mathbf{A}$, has two elements, 0 and 1, where 0 indicates no relationship between two prices and 1 indicates a relationship between two prices.

Furthermore, suppose $n_i = \sum_{j=1}^4 a_{ij}$ refers to the i th row sum of the adjacency matrix, and $E(\epsilon_{it}) = 0$ and $\text{var}(\epsilon_{it}) = \sigma^2$.

In order to obtain an estimate of the unknown parameters of the model, we should convert Equation (2) into a matrix form as below:

$$(p_{1t}, \dots, p_{4t})^T = (1, 1, 1, 1)^T \beta_1 + (\beta_2 \begin{pmatrix} \frac{1}{n_1} & 0 & 0 & 0 \\ 0 & \frac{1}{n_2} & 0 & 0 \\ 0 & 0 & \frac{1}{n_3} & 0 \\ 0 & 0 & 0 & \frac{1}{n_4} \end{pmatrix} \mathbf{A} + \beta_3 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}) \times (p_{1(t-1)}, \dots, p_{4(t-1)}) + (\epsilon_{1t}, \dots, \epsilon_{4t})^T \quad (11)$$

It can also be written in a concise form as follows:

$$\mathbb{P}_t = \mathbb{B}_1 + \mathbf{G}\mathbb{P}_{t-1} + \epsilon_{it} \quad (12)$$

where,

1. $\mathbb{P}_t = (p_{1t}, \dots, p_{4t}) \in \mathbb{R}^4$,
2. $\mathbb{B}_1 = \beta_1 \mathbf{1}$ wherein $\mathbf{1} = (1, 1, 1, 1)^T$,
3. $\mathbf{G} = \beta_2 \mathbf{W} + \beta_3 \mathbf{I}$,
4. $\mathbf{W} = \text{diag}\{n_1^{-1}, \dots, n_4^{-1}\} \mathbf{A}$,
5. $\mathbf{I}$ is an identity matrix.

Earlier in Section 1, it was mentioned that the model is composed of three main components, each of which represents a distinct feature, as described below in further detail.

1. OHLC network effect

In most of the studies to forecast the crude oil price, only the close price is used as part of the prediction process, and the price fluctuations during the day are not taken into account. However, the price fluctuation that occurs over the course of a day can be affected the close price. As a way to solve this problem, we have utilize the interconnectedness of OHLC pricing on one another to construct a network of them; moreover, we evaluate their effect on the close price. We call this effect as the "OHLC prices network effect".

2. Autoregressive effect

In general, the close price of one day significantly impacts the close price of the following day. This phenomenon is referred to as an "autoregressive effect", and the MINAR model includes it as well.

3. Independent noise effect

As a result of an independent random noise, there should also be a role for the unexplained variation which we call it "model error effect". Equation (2) considers the assumption of a normal distribution for the error term, which has been mentioned

in other studies as well, for instance see [41]. Assuming a normal distribution for the noise term allows us to have a minimum variance. It facilitates prediction and helps the algorithm straightforwardly approximate the Maximum Likelihood Estimator (MLE). For more details, refer to [42].

We illustrate the application of Equation (11) by taking an example. Assuming the close price on the third day is p_{13} where $i = 1$ and $t = 3$, then the following steps need to be taken:

$$\begin{aligned}
 p_{13} &= \beta_1 + \overbrace{\beta_2 n_1^{-1} \sum_{j=1}^4 a_{1j} p_{j2}}^{\text{OHLC network effect}} + \underbrace{\beta_3 p_{12}}_{\text{Autoregressive effect}} + \underbrace{\epsilon_{13}}_{\text{Independent noise}} \\
 &= \beta_1 + \beta_2 \frac{a_{12} p_{22} + \dots + a_{14} p_{42}}{a_{12} + \dots + a_{14}} + \beta_3 p_{12} + \epsilon_{13}
 \end{aligned} \quad (13)$$

The whole process of calculating $\mathbb{P}_t$ is depicted in Figure 3.

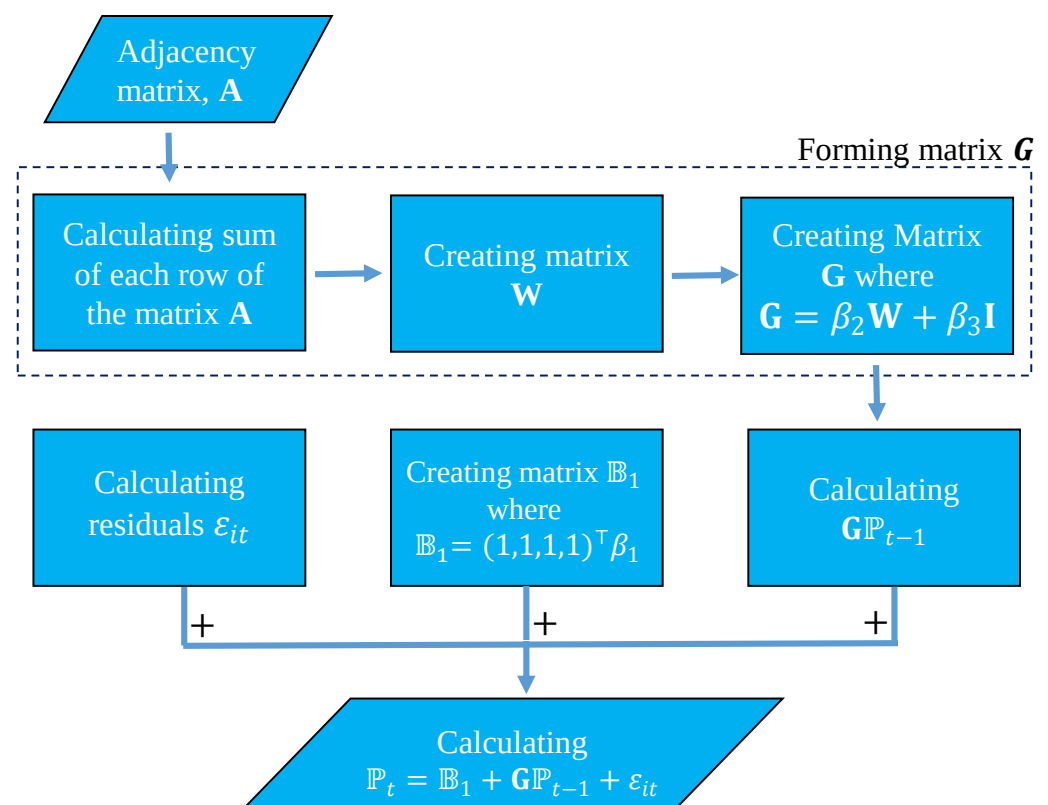


Figure 3. The flowchart of calculating $\mathbb{P}_t$.

2.4. Parameter Estimation

To estimate the unknown parameter, β , Equation (2) is rewritten in the following form:

$$p_{it} = \beta_1 + \beta_2 w_i^\top \mathbb{P}_{t-1} + \beta_3 p_{i(t-1)} + \epsilon_{it} \quad (14)$$

In this case, it would be written as follows:

$$p_{it} = Y_{i(t-1)}^\top \beta + \epsilon_{it} \quad (15)$$

where $Y_{i(t-1)} = (1, w_i^\top \mathbb{P}_{t-1}, p_{i(t-1)})^\top$ and $w_i = (a_{ij}/n_i : 1 \leq j \leq 4)^\top$ indicate the i th row vector of W . Suppose $\mathbb{Y}_t = (Y_{1t}, \dots, Y_{4t})^\top$, then the above model can be written as:

$$\mathbb{P}_t = \mathbb{Y}_{t-1}\beta + \varepsilon_t \quad (16)$$

Then, one can obtain a maximum likelihood estimator (MLE) in the form of a logarithm by taking the following steps:

$$\mathcal{L}(\beta) = -2 \ln 2\pi\sigma^2 - \frac{1}{2\sigma^2} \left(\sum_{t=1}^T \mathbb{P}_t - \sum_{t=1}^T \mathbb{Y}_{t-1}\beta \right)^\top \left(\sum_{t=1}^T \mathbb{P}_t - \sum_{t=1}^T \mathbb{Y}_{t-1}\beta \right) \quad (17)$$

In order to differentiate this expression in accordance with β , the ML estimates will be as follows:

$$\frac{\partial \mathcal{L}}{\partial \beta^\top} = -\frac{1}{2\sigma^2} \left(-2 \sum_{t=1}^T \mathbb{Y}_{t-1}^\top \mathbb{P}_t + 2 \sum_{t=1}^T \mathbb{Y}_{t-1}^\top \mathbb{Y}_{t-1} \beta \right) = 0 \quad (18)$$

Considering the above,

$$\hat{\beta} = \left(\sum_{t=1}^T \mathbb{Y}_{t-1}^\top \mathbb{Y}_{t-1} \right)^{-1} \sum_{t=1}^T \mathbb{Y}_{t-1}^\top \mathbb{P}_t \quad (19)$$

When Equation (16) is substituted into the estimator $\hat{\beta}$ in Equation (19), we obtain the following results:

$$\hat{\beta} = \beta + \left(\sum_{t=1}^T Y_{t-1}^\top Y_{t-1} \right)^{-1} \sum_{t=1}^T Y_{t-1}^\top \varepsilon_t \quad (20)$$

Theorem 1. According to the MINAR model, if $|\beta_2| + |\beta_3| < 1$, then there exists a unique solution that is strictly stationary. The solution takes the form of:

$$\mathbb{P}_t = (\mathbf{I} - \mathbf{G})^{-1} B_1 + \sum_{j=0}^{\infty} \mathbf{G}^j \varepsilon_{t-j} \quad (21)$$

See the proof in Appendix A.

Theorem 2. With the conditions given by Theorem 1, $\hat{\theta}$ is an asymptotic unbiased estimator of θ ; also $\|\hat{\theta} - \theta\| \rightarrow 0$ in probability as $T \rightarrow \infty$.

See the proof in Appendix A.

2.5. ARIMA Model

As a benchmark for the MINAR model, we utilized the ARIMA model in this study. An algorithm developed by Box and Jenkins [41] combines autoregressive (AR) and moving average (MA) models; however, the ARMA(m, n) model is generally used for univariate time series modeling and is a combination of AR(m) and MA(n) models. Accordingly, ARMA(m, n) can be described as follows:

$$p_t = c + \varepsilon_t + \sum_{i=1}^m \lambda_i p_{t-i} + \sum_{i=1}^n \kappa_i \varepsilon_{t-i} \quad (22)$$

where $\kappa_1, \dots, \kappa_n$ and $\lambda_1, \dots, \lambda_m$ are the parameters of the model, and ε is the white noise. To make the series stationary, it is necessary to differentiate it by d ($d = 1, 2, 3, \dots$) times. Such a time series model is called an ARIMA(m, d, n) model.

In ARIMA model creation, three steps are involved: identification, parameter estimation, and diagnostic testing. The auto correlation function (ACF) plots are used to determine

the AR and MA terms in the model's identification process after checking the time series stationarity. The ACF is a statistical correlation metric used to determine whether previous values in the time series analysis are related to the latest values. The least squares method is then used to estimate the ARIMA parameters, (m, d, n) . The Akaike information criterion (AIC), the Bayesian information criterion (BIC), and second-order Akaike's information criterion (AICc) are the three main methods for selecting appropriate models. A smaller value of these three criteria makes the given model more efficient. For more detail, see [43].

2.6. Model Validation Methods

In this study, validation techniques are deployed for measuring the accuracy and performance of the MINAR model. To do this, we utilize three measures, two of which are common standard measures, namely, Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error, (MAPE), defined as below:

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (p_t - \hat{p}_t)^2} \quad (23)$$

$$\text{MAPE} = \frac{100}{T} \sum_{t=1}^T \left| \frac{p_t - \hat{p}_t}{p_t} \right| \quad (24)$$

where p_t denotes the actual value of close crude oil price and $\hat{p}_t$ and T are the modeled values of the close crude oil price and the total number of days, respectively.

The third measure is Unscaled Mean Bounded Relative Absolute Error (UMBRAE) index, which is defined by [44]. UMBRAE, with its selectable benchmark, combines the best qualities of various alternative measures without their disadvantages. Based on bounded relative error, it can provide an informative and interpretable result. Compared to other measures, it is less sensitive to forecasting outliers. It is also symmetric and scale-independent [44]. UMBRAE's structure is as follows:

1. Let p_t denote the log-return of close oil price at time t and $\tilde{p}_t$ denote the forecasts of it. Then the forecasting error e_t can be defined as $p_t - \tilde{p}_t$. Let $e_t^* = p_t - p_t^*$ denote the forecasting error at time t obtained by a benchmark model, where p_t^* is the forecast at time t by a benchmark model.

For close oil prices, Mean Bounded Relative Absolute Error (MBRAE) can be defined as:

$$\text{MBRAE} = \frac{1}{T} \sum_{t=1}^T \frac{|e_t|}{|e_t| + |e_t^*|} \quad (25)$$

2. MBRAE can be transformed into an interpretable measure called unscaled MBRAE (UMBRAE) as defined below:

$$\text{UMBRAE} = \frac{\text{MBRAE}}{1 - \text{MBRAE}} \quad (26)$$

3. UMBRAE makes it easy to interpret forecasting method performance in terms of average relative absolute error, as follows:

If UMBRAE is equal to 1, the proposed method performs roughly the same as the benchmark method; at $\text{UMBRAE} < 1$, the proposed method performs roughly $(1 - \text{UMBRAE}) \times \%100$ better than the benchmark method; using the proposed method when $\text{UMBRAE} > 1$, the benchmark method is roughly $(\text{UMBRAE} - 1) \times \%100$ better.

Algorithm 1 shows the entire process of how the MINAR model is applied in order to forecast the crude oil price.

Algorithm 1: The procedure of modeling using MINAR

Data: Collecting WTI OHLC prices and data pre-processing
Result: estimating β .
Transforming data by Equation (1);
Constructing the MINAR model by Equation (2);
Creating correlation matrix using Equation (8);
while setting threshold, θ **do**
 if $\mathbb{Y}_t$ is non-singular **then**
 | Evaluate the model with RMSE, MAPE and UMBRAE;
 end
 Change the threshold and go to the previous step;
end

3. Results and Analysis**3.1. Markov Chain Analysis of Data**

In order to gain a more in-depth understanding of our dataset, we apply the Markov chain analysis method to investigate the movement of close oil prices. Using the same transformation as in Equation (1), we transformed the time series of close oil prices in the WTI to the returns between the closing prices on a daily basis, p_t .

To reduce the dimension of our data, we transform the close prices of oil over a period of time into a sequence of natural numbers, this in turn leads to a new sequence of numbers. As a first step, we split the original data into two groups: training data and test data. Training data are used to train the method, while test data are used to evaluate its accuracy. Consequently, the data between 1 January 2020 and 20 January 2022, have been treated as a train set, with 542 observations accounting for 80 percent of the total samples. Between 21 January 2022 and 20 July 2022, 134 observations were used in the test set. Then we transform the data by Equation (1). Afterward, we calculate the average, $\bar{p}$ and standard deviation, σ of the test data and set the threshold according to the equation below. The whole process of data transformation for Markov chain analysis is depicted in Figure 4. Time series of return are transformed to a sequence of numbers s_t as follows:

$$s_t = \begin{cases} 1 & p_t < \bar{p} - \sigma/2 \\ 2 & \bar{p} - \sigma/2 \leq p_t \leq \bar{p} + \sigma/2 \\ 3 & p_t > \bar{p} + \sigma/2 \end{cases} \quad (27)$$

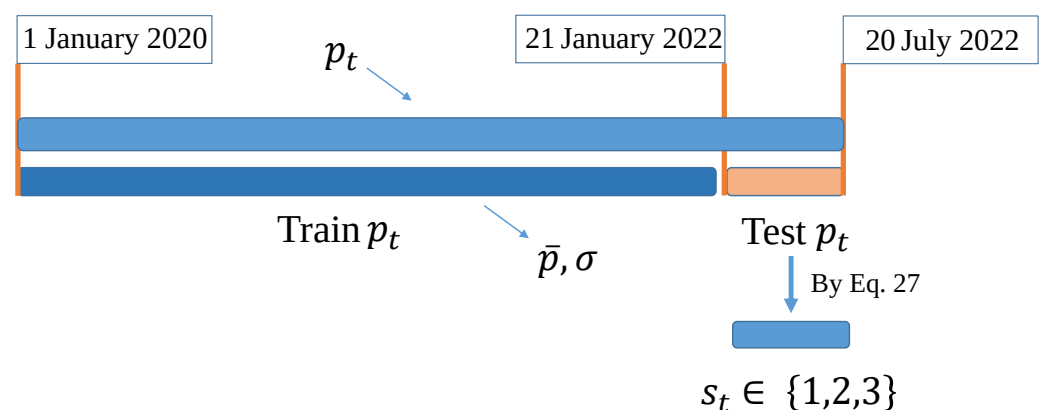


Figure 4. The transformation process of the test data.

Generally, the number one represents the bear day, when the value of the close oil price has declined, the number two represents the stagnant day, when the value of the close oil price has remained the same or has shown a negligible change, and the number three represents the bull day, when the value of the close oil price has risen.

We suppose that $\{s_t, t \geq 1\}$ is a time-homogeneous Markov chain with states $\{1, 2, 3\}$, and the transition probabilities are $P(s_{t+1} = k | s_t = l) = P(s_1 = k | s_0 = l) = p_{lk}$ where $1 \leq l \leq 3$ and $1 \leq k \leq 3$.

For any ergodic Markov chain, $\lim_{N \rightarrow \infty} P_{lk}^N$ exists and is independent of l where N is the time steps. Furthermore,

$$\lim_{N \rightarrow \infty} P_{lk}^N = \pi_k > 0 \quad (28)$$

where the $0 \leq \pi_k \leq 1$ satisfies the following steady-state equations:

$$\pi_k = \sum_{l=1}^3 \pi_l p_{lk}, \quad \text{for } k = 1, 2, 3 \quad \text{and} \quad \sum_{k=1}^3 \pi_k = 1 \quad (29)$$

That is, $\{\pi_k, 1 \leq k \leq 3\}$ is the unique stationary distribution.

As observed in Figure 5, according to Equation (29), Markov chain probabilities are constructed with three states bear day, bull day, and stagnant day. In this figure, each node represents the states, and the arrows between them represent the transition probability. Figure 5 depicts the probability of price movement from one state to another state. The loop represents the movement from one state to itself. Given this Markov Chain graph, the stationary distribution can be described as follows:

$\pi_{\text{close oil price}} = (0.194, 0.601, 0.204)$ where the states are the bear, stagnant, and bull days respectively.

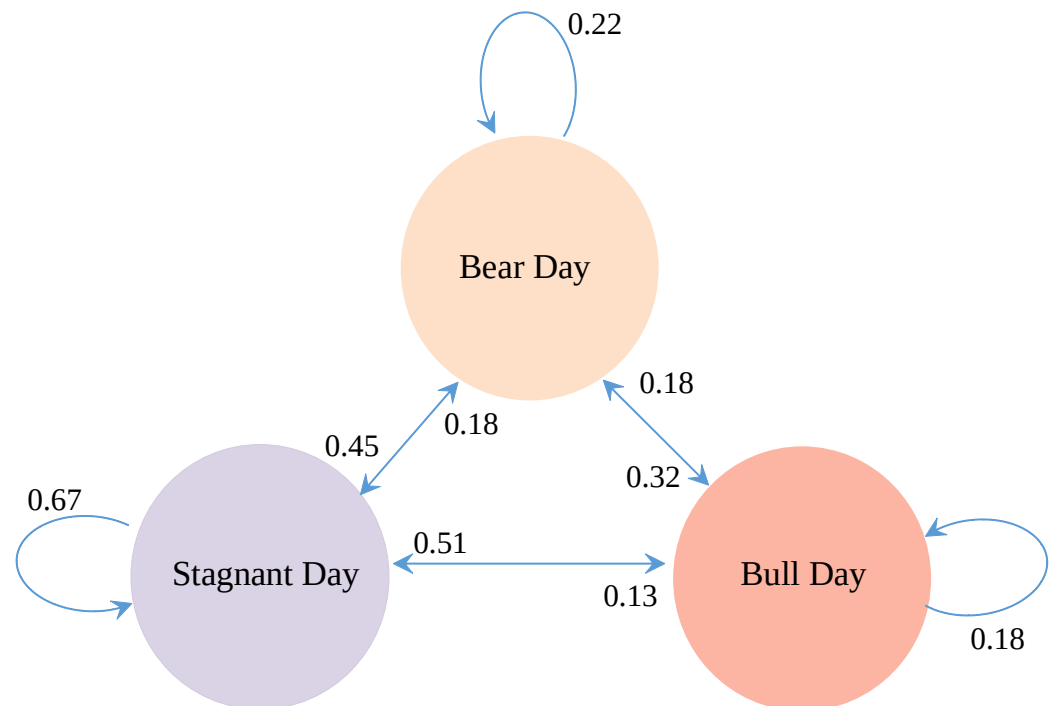


Figure 5. The transition probabilities of close WTI crude oil prices with three different states.

The transition probability indicates that the price movement tends to move from a stagnant day to a stagnant day, which explains why the prices either remain at the same level or move slightly from one level to another over time. There is also a high probability that the bullish trend will turn into a stagnant trend in later days, and prices are slowly rising over the course of time as well. The result of the analysis shows that the price movement is nonlinear and with different probabilities jumps from a state to another. Because of this, standard models such as those used in ARMA do not fully capture the accurate picture of the price movement behind them, and therefore we need a different tool to depict the nonlinear behavior of the system.

3.2. MINAR Model Results

This study uses the ARIMA model as a benchmark for comparing the results of the proposed MINAR method to those obtained by the ARIMA model. The training dataset covers the period from 1 January 2020 to 20 January 2022, while the testing dataset covers the period from 21 January 2022 to 20 July 2022. As for the training part, the ARIMA model is conducted, followed by validating the results using the testing dataset.

The ARIMA model fits the datasets of WTI close oil price to compare them with the proposed model. Additionally, it is essential to note that the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) and the Augmented Dickey–Fuller (ADF) tests also confirm that the training time series are stationary at the 95% significance level. Thus, we apply the differencing method one time to the dataset to achieve the desired results. Later, different models are constructed by adjusting various parameters in the MA and AR components of the ARIMA model, as summarized in Table 2. Moreover, the ARIMA (1,1,1) model also assumes that the AICc criteria should be taken into account. Moreover, it is necessary to verify some assumptions of the chosen model before it can be used. It can be seen in Figure 6 that the residual analysis of the model is presented. In addition, the Ljung–Box test on the residuals and the squared residuals are also statistically significant with a p -value equaling 0.92 at the 95% level. In this case, the model with the lowest AIC, BIC, and AICc is the most appropriate model with the minimum summary measures.

Table 2. Summary measures for AICc in ARIMA close oil prices of WTI model.

Model	AICc
ARIMA(0,1,0)	3232.287
ARIMA(0,1,2)	3449.538
ARIMA(0,1,2)	3225.706
ARIMA(0,1,3)	3225.781
ARIMA(0,1,4)	3298.849
ARIMA(0,1,5)	3451.441
ARIMA(1,1,0)	3464.798
ARIMA(1,1,0)	3303.881
ARIMA(1,1,1)	3224.632
ARIMA(1,1,2)	3225.036
ARIMA(1,1,3)	3298.861
ARIMA(2,1,0)	3226.983
ARIMA(2,1,1)	3225.122
ARIMA(2,1,2)	3298.252
ARIMA(3,1,0)	3227.109
ARIMA(4,1,0)	3299.604
ARIMA(5,1,0)	3452.879

The MINAR model that has been constructed is deployed on WTI's OHLC prices in accordance with Algorithm 1. The dataset split into two subsets in terms of training and testing. In the next step, the MINARI model will be applied to the training dataset of close prices, which will be used for model validation. This is accomplished by transforming the training dataset into a smooth time series, calculated by Equation (1) during the training process. In the following step, the correlation matrix of the OHLC prices is created by Equation (8), as shown in Figure 7. In this way, it is possible to establish the MINAR model for the OHLC prices network based on Equation (2), where the parameter β is estimated as $\beta = (0.040, 0.015, 0.178)$. The comparison between the log-returns of the actual oil close price and the MINAR generated are depicted in Figure 8.

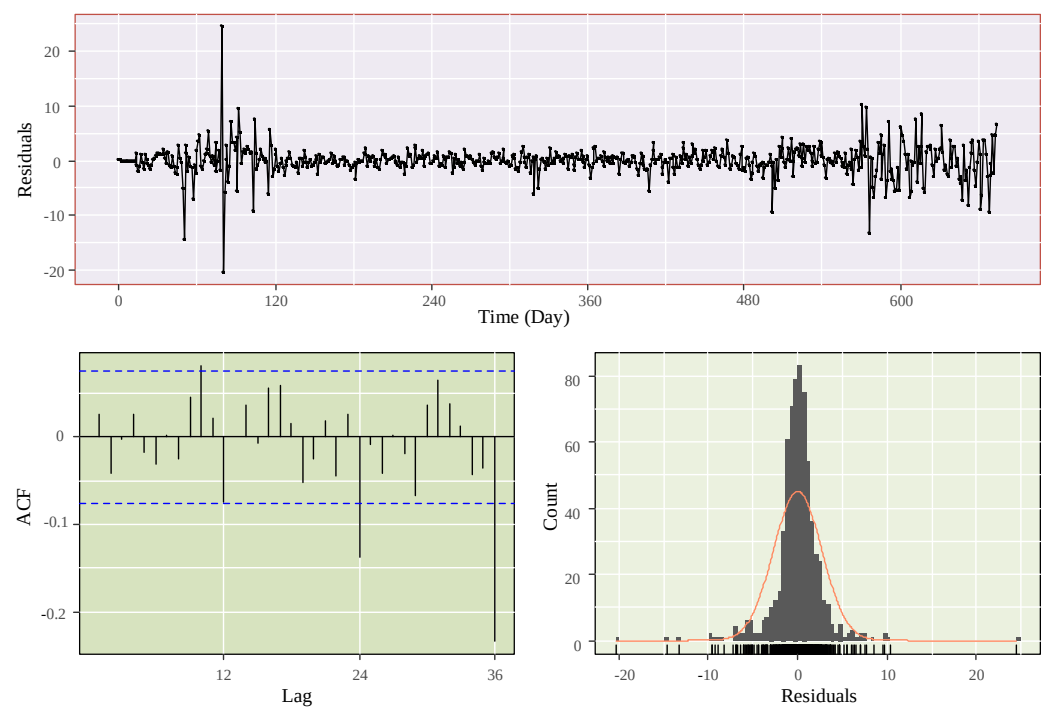


Figure 6. Residuals from ARIMA(1,1,1) for WTI close oil price.

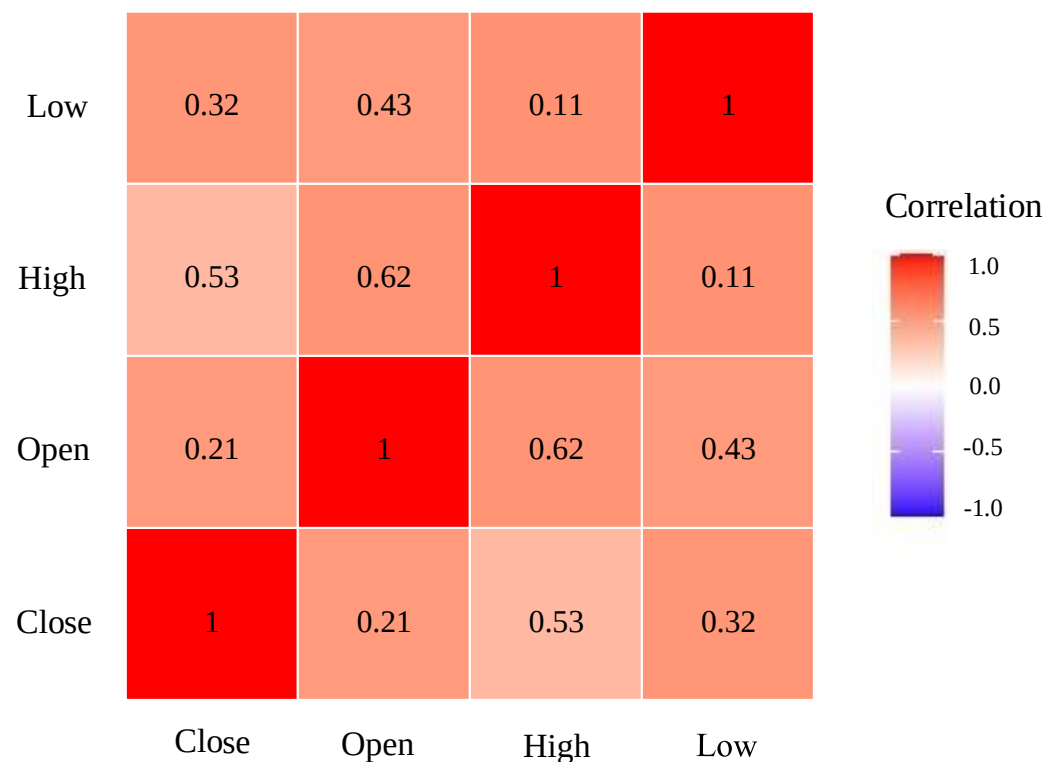


Figure 7. Correlations among the OHLC prices of the WTI using Equation (8).

In this paper, we have deployed three methods to validate the model's efficiency and accuracy, the results of which can be found in the following. For the ARIMA model of the close oil price, the RMSE and MAPE are 6.2 and 9.6, respectively. In addition, the RMSE and MAPE values for the MINAR model are 2.6 and 2.91, respectively. However, the UMBRAE of the MINAR model is 0.943, and its performance is 6.8% compared to the performance of

the ARIMA model as a benchmark. It can therefore be concluded that the MINAR model is a better model in terms of efficiency and accuracy than the ARIMA model.

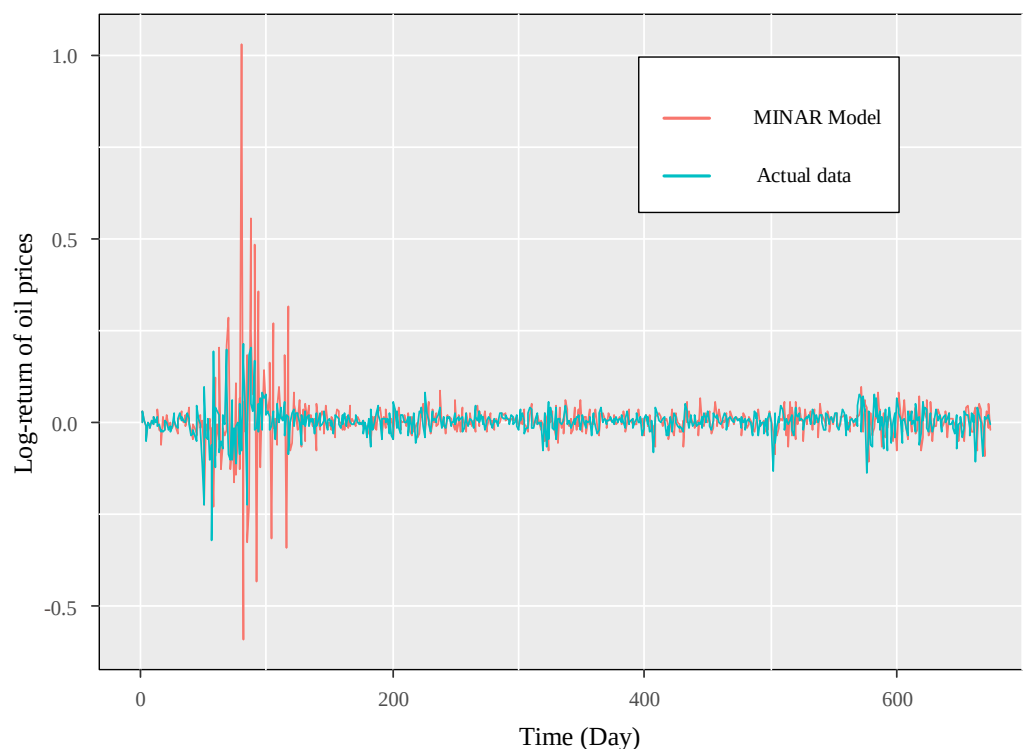


Figure 8. The comparison between the log-returns of the actual oil close price and the MINAR model.

4. Discussion

4.1. Threshold Selection

According to Section 2, the adjacency matrix, representing the OHLC price interaction, is created by comparing the correlation matrix with a selected threshold. In order to establish this threshold and to explain how this process is going to be implemented, additional analysis is needed, which will be discussed in this section.

For an explanation of the methodology used to calculate the correlation threshold, it is necessary to consider the following two conditions. As a matter of first condition, none of the n_i 's in Equation (2) should be zero; this would result in the equation having infinity terms in it. Additionally, a value of n_i 's that minimizes the RMSE, MAPE and UMBRAE is preferable since it enables the Algorithm 1 to perform more accurately.

Whenever these two conditions are met, the threshold value is defined as a selected variable, which can be used to implement an algorithm that meets both conditions. An external loop is then added to the main algorithm to change the value of θ and calculate the corresponding n_i and RMSE, MAPE, and UMBRAE. Table 3 shows the results for different θ values between 0 and 1 with an increment of 0.05. Through the computation of Table 3, there is a parameter, n_4 , which is turned to be 0 from $\theta = 0.45$ onward. As the term n_4 is appeared in the denominator for the calculation of $(p_{1t}, \dots, p_{4t})$ in Equation (11), the results approach infinity, which has no meaning in our context. Thus, the results are valid only for $0 < \theta \leq 0.4$. Further, RMSE, MAPE, and UMBRAE are minimized when $\theta = 0.4$. Thus, it is selected as the best case and applied into the model. The running time of each threshold is seen in Table 3. With increasing increments, the running time associated with each threshold decreases.

Table 3. Different correlation threshold and its corresponding n_i , RMSE, MAPE, UMBRAE, and related running time.

θ	n_1	n_2	n_3	n_4	RMSE	MAPE	UMBRAE Performance (%)	Running Time (s)
0.1	3	3	3	3	3.43	3.63	5.42	101.15
0.15	3	3	2	2	3.28	3.46	5.63	100.8
0.20	3	3	2	2	3.28	3.46	5.63	100.8
0.25	2	2	2	2	3.07	3.32	5.80	100.2
0.30	2	2	2	2	2.94	3.2	6.10	100.2
0.35	1	2	1	3	2.81	3.02	6.30	98.7
0.40	1	2	2	1	2.6	2.91	6.80	98.2
0.45	1	1	2	0	NA	NA	NA	-
0.50	1	1	2	0	NA	NA	NA	-

4.2. Comparison of the MINAR Model with Similar Approaches

In the previous section, we utilized the ARIMA model as a benchmark to evaluate the accuracy and performance of the MINAR model. In this subsection, we evaluate the efficiency and accuracy of the MINAR model in depth; therefore, different methods that can be used in the context of both univariate as well as multivariate time series will be discussed in order to analyze the effectiveness of the MINAR model. As a result, it will be explained whether the proposed model should take precedence over similar approaches. We first briefly introduce two AR and MA models among the family of univariate time series models. Next, we compare them with the MINAR method. Finally, we explain the multivariate time series and vector autoregressive (VAR) model and compare its results with the proposed MINAR model.

The expression $AR(m)$ denotes an autoregressive model of order m . The $AR(m)$ model is formulated as:

$$p_t = c + \sum_{i=1}^m \varphi_i p_{t-i} + \epsilon_t \quad (30)$$

where $\varphi_1, \dots, \varphi_m$ are the parameters of the model, c is a constant, and ϵ_t is white noise.

The $AR(m)$ is utilized the OHLC prices and the results are given in Table 4. Based on this table, both RMSE and MAPE indicators significantly differ from the MINAR indicators. The RMSE and MAPE of the AR model are 9.22 and 12.36, respectively, and for the MINAR model are 2.6 and 2.91, respectively; thus, the MINAR model has the advantage of being more accurate than this model. The AR method is only based on the prices of the past days and is linearly built, which cannot precisely model the movement of crude oil prices.

Table 4. The RMSE and MAPE for OHLC oil prices of the AR model.

Measurement	Open	High	Low	Close
RMSE	9.24	9.26	9.20	9.22
MAPE	12.30	12.38	12.28	12.36

The expression $MA(n)$ denotes an moving average model of order n . The $MA(n)$ model is formulated as:

$$p_t = \mu + \sum_{i=1}^n \theta_i \epsilon_{t-i} + \epsilon_t \quad (31)$$

where μ is the mean of the crude oil prices, the $\theta_1, \dots, \theta_n$ are the parameters of the model and the $\epsilon_t, \epsilon_{t-1}, \dots, \epsilon_{t-n}$ are white noise error terms. The value of n is called the order of the MA model.

The MA(2) is deployed on the OHLC prices, and the results are given in Table 5. According to RMSE and MAPE, the MINAR method is significantly more accurate than the MA method. The RMSE and MAPE of the MA model are 9.64 and 12.78, respectively, and for the MINAR model are 2.6 and 2.91, respectively. There are several problems with the MA method as it relies on the current and the past residual values and cannot predict nonlinear fluctuations in oil prices.

Table 5. The RMSE and MAPE for OHLC oil prices of the MA model.

Measurement	Open	High	Low	Close
RMSE	9.55	9.74	9.47	9.64
MAPE	12.71	12.81	12.74	12.78

Statistically, the vector autoregression is a method of modeling time series data by capturing the relationship between several quantities over time. A VAR model can be thought of as a type of stochastic process model. A VAR model generalizes the single-variable autoregressive model by considering multivariate data.

As in the autoregressive model, each variable has a set of equations describing its evolution over time based on its prior values. An error term and the lagged values of the other variables are included in this equation.

A q -order VAR is denoted VAR(q) and defined as follows:

$$p_t = c + \sum_{i=1}^q A_i p_{t-i} + \epsilon_t \quad (32)$$

The variable c represents the model's intercept as a k -vector. A_i is a time-invariant ($k \times k$)-matrix, and ϵ_t is a k -vector of error terms. Three conditions must be met by the error terms:

1. Every error term has a mean of zero.
2. In error terms, the covariance matrix is positive-semidefinite.
3. There is no correlation in individual error terms.

As shown in Table 6, the RMSE and MAPE of the close price of crude oil of the VAR model are 5.16 and 8.43, respectively. However, the RMSE and MAPE of the MINAR model close price of crude oil are 2.6 and 2.91, respectively. Therefore, the MINAR model is more accurate than the VAR model in both RMSE and MAPE measurements.

Table 6. The RMSE and MAPE for OHLC oil prices of the VAR model.

Measurement	Open	High	Low	Close
RMSE	4.95	5.42	4.92	5.16
MAPE	8.35	8.51	8.31	8.43

According to this section discussion, methods such as AR and MA do not perform well as the single-variable time series than the MINAR model. According to the results of this section and the results mentioned in the previous section, the MINAR model outperforms than the family of ARIMA models. Additionally, the MINAR model is preferable to the models in similar approaches of multivariate time series. The superior performance of the MINAR model can be summarized by the following principal reasons, as discussed in the section.

1. All the price information available in one day, i.e., OHLC prices in the market, has been used to predict the price of crude oil.
2. Unlike multivariate time series, the interactions between the variables are also considered in addition to considering each time series by building a network of OHLC prices.

3. Using information theory and Shannon's entropy, the shortcomings of linear methods have been solved, and the nonlinear trend of oil price data has been modeled accurately.

5. Summary and Conclusions

In this article, we first discussed the importance of the crude oil price in the global economy and the necessity to forecast it. Then, different forecasting methods and directions of oil price forecasting were reviewed. It is also noted that the high volatility of the oil market and the nonlinearity of the market make crude oil price predicting challenging. As a result, It is discussed that many of these methods are associated with three significant drawbacks that are not easily overcome. In the first place, these methods adopt a linear approach to forecasting; therefore, they can only be applied when there is a linear trend in the data. As a second issue, they have to estimate many parameters for the constructed models, thus making the forecasting process complex to perform. As a third drawback, it should be noted that some new methods still require more research within this area to be developed.

As a result of the importance of predicting the price of oil and how it impacts other economic elements, the method used in this article has a particular significance as it uses the fewest parameters, reduces computational complexity, and is more reliable than similar methods in terms of accuracy. Furthermore, the nonlinear nature of most economic phenomena, such as oil price fluctuations, makes linear estimation methods extremely inaccurate, underscoring the need for nonlinear estimation methods in order to accurately estimate economic phenomena. As a result of the application of information theory-based measures, the MINAR model has overcome this shortcoming. In addition, by utilizing Shannon entropy, we are able to take into account the nonlinear and noisy behavior of our data.

We used the Markov chain analysis of our dataset to explain the nonlinearity of the oil price movement. Considering the complexity of the crude oil price movements, we develop the MINAR model that enables us to forecast the nonlinear nature of price changes by deploying information theory. Furthermore, the MINAR model is constructed using the previous day's close price, the joint MI correlation of OHLC prices, and the independent noise. The advantage of this model over other multivariate time series is that the MINAR model considers the modeling of each time series and also the interactions across the variables. In addition, despite other multivariate time series models, with increasing the number of variables and unknown parameters increasing, the number of unknown parameters in the MINAR models is fixed.

For this purpose, after collecting the OHLC prices of WTI crude oil from 1 January 2020 to 20 July 2022, we transform them into the new log-return prices. Afterward, the correlation matrix among the OHLC is constructed based on mutual information distances. With the method mentioned in the discussion, the appropriate threshold is selected and used in the developed MINAR model, and its parameters are then estimated using the MLE. The accuracy and efficiency of the model are measured using RMSE, MAPE, and UMBRAE methods. According to all three methods, the MINAR model performs better than the benchmark ARIMA model. For the ARIMA model, the RMSE and MAPE are 6.2 and 9.6, respectively. However, for the MINAR model, the RMSE is 2.6, and the MAPE is 2.91. In addition, the UMBRAE method shows that the MINAR model performance is 6.8% better than the ARIMA as a benchmark.

Here we highlight the following main contributions of this study:

1. A crude oil price forecasting model is formulated based on simple, non-heavy calculations due to unknown parameters.
2. Modeling is done of the actual close price using the network of daily OHLC prices in order to overcome the missing information about whole-day price fluctuations.
3. Implemented in this work is a nonlinear modeling approach based on mutual information distances by applying a nonlinear approach to modeling.

To obtain accurate price predictions, further research can be conducted using tick data. A MINAR model should also take into account more external effects that affect the volatility of oil prices in addition to internal effects.

Author Contributions: A.S.K.: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing—original draft, Writing—review and editing; M.S.: Data curation, Investigation, Methodology, Project administration, Resources, Software, Validation, Writing—original draft, Writing—review and editing; Y.E.T.: Data curation, Investigation, Methodology, Validation, Formal analysis, Writing—original draft. All authors have read and agreed to the published version of the manuscript.

Funding: The work described in this paper is supported by the Faculty of Economics and Business Administration of Yibin University, Yibin, China.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data that support the findings of this study are available from the authors upon reasonable request.

Acknowledgments: The first author gratefully acknowledges the Faculty of Economics and Business Administration of Yibin University for the support during the preparation of this paper. The authors would like to thank Mahdi Shadabfar for his valuable suggestions while writing this paper and in responding to the reviewers' comments.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Proof of Theorem 1. It is assumed that $\lambda_i(M)$ is the i th eigenvalue of any arbitrary matrix M . It is first required to verify that Equation (21) produces a solution which meets strict stationary requirements. Because the absolute value of any eigenvalue of the stochastic matrix is less than or equal to 1, as a consequence, $|\lambda_i(\mathbf{W})| \leq 1$, then,

$$\rho = \max_{1 \leq i \leq 4} |\lambda_i(\mathbf{G})| \leq \beta_2 \max_{1 \leq i \leq 4} |\lambda_i(\mathbf{W})| + |\beta_3| < 1 \quad (\text{A1})$$

Accordingly, $\lim_{j \rightarrow \infty} \mathbf{G}_{t-j}^j$ exists and $\mathbb{P}_t$ has the strictly stationary process. In the next step, we prove that the solution is unique.

It is straightforward to verify that $\mathbb{P}_t$ satisfies the MINAR model conditions. Our next step is to determine if the solution of strictly stationary is unique. For any positive arbitrary integer m , assume that $\tilde{\mathbb{P}}_t$ is another strictly stationary solution to the MINAR model with $\sup_{t \geq 0} E\|\mathbb{P}_t\| \leq \infty$ then $\tilde{\mathbb{P}}_t = \sum_{j=0}^{m-1} \mathbf{G}^j(\mathbf{B}_1 + \varepsilon_{t-j}) + \mathbf{G}^m \tilde{\mathbb{P}}_{t-m}$. Thus, by

$$E\|\mathbb{P}_t - \tilde{\mathbb{P}}_t\| = E\left\|\sum_{j=m}^{\infty} \mathbf{G}^j(\mathbf{B}_1 + \varepsilon_{t-j}) + \mathbf{G}^m \tilde{\mathbb{P}}_{t-m}\right\| \leq k\rho^m \quad (\text{A2})$$

where k is a constant independent of t and m . Hence, we have that $E\|\mathbb{P}_t - \tilde{\mathbb{P}}_t\| = 0$, i.e., $\mathbb{P}_t = \tilde{\mathbb{P}}_t$ with probability one and it completes the proof. $\square$

Proof of Theorem 2. Note that

$$\begin{aligned} \|\hat{\theta} - \theta\|^2 &= (\hat{\theta} - \theta)(\hat{\theta} - \theta)^\top \\ &= \left(\sum_{t=1}^T \mathbf{X}_{t-1}^\top \mathbf{X}_{t-1}\right)^{-2} \left(\sum_{t=1}^T \mathbf{X}_{t-1}^\top \varepsilon_t\right) \left(\sum_{t=1}^T \varepsilon_t^\top \mathbf{X}_{t-1}\right) \end{aligned} \quad (\text{A3})$$

Accordingly, by using Theorem 1 and the Chebychev inequality, therefore, we can prove the following:

$$(T^{-1} \sum_{t=1}^T X_{t-1}^{\top} \varepsilon_t)(T^{-1} \sum_{t=1}^T \varepsilon_t^{\top} X_{t-1}) \rightarrow 0 \quad \text{in probability as } T \rightarrow \infty \quad (\text{A4})$$

and

$$(T^{-1} \sum_{t=1}^T X_{t-1}^{\top} X_{t-1})^2 \rightarrow E(T^{-1} \sum_{t=1}^T X_{t-1}^{\top} X_{t-1})^2 \quad \text{in probability as } T \rightarrow \infty \quad (\text{A5})$$

Moreover, we can verify $E(\hat{\theta} - \theta) \rightarrow 0$ as $T \rightarrow \infty$. Thus, $\|\hat{\theta} - \theta\|^2 \rightarrow 0$ in probability as $T \rightarrow \infty$, that is, $\hat{\theta}$ is an asymptotic unbiased estimator of θ . $\square$

References

- Shabri, A.; Samsudin, R. Daily Crude Oil Price Forecasting Using Hybridizing Wavelet and Artificial Neural Network Model. *Math. Probl. Eng.* **2014**, *2014*, 201402. [\[CrossRef\]](#)
- Yu, L.; Dai, W.; Tang, L. A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting. *Eng. Appl. Artif. Intell.* **2016**, *47*, 110–121. [\[CrossRef\]](#)
- Safari, A.; Davallou, M. Oil price forecasting using a hybrid model. *Energy* **2018**, *148*, 49–58. [\[CrossRef\]](#)
- Wu, J.; Chen, Y.; Zhou, T.; Li, T. An Adaptive Hybrid Learning Paradigm Integrating CEEMD, ARIMA and SBL for Crude Oil Price Forecasting. *Energies* **2019**, *12*, 1239. [\[CrossRef\]](#)
- Zhang, Y.J.; Wei, Y.M. The dynamic influence of advanced stock market risk on international crude oil returns: An empirical analysis. *Quant. Financ.* **2011**, *11*, 967–978. [\[CrossRef\]](#)
- Wu, J.; Miu, F.; Li, T. Daily Crude Oil Price Forecasting Based on Improved CEEMDAN, SCA, and RVFL: A Case Study in WTI Oil Market. *Energies* **2020**, *13*, 1852. [\[CrossRef\]](#)
- Huang, Y.; Deng, Y. A new crude oil price forecasting model based on variational mode decomposition. *Knowl.-Based Syst.* **2021**, *213*, 106669. [\[CrossRef\]](#)
- Abdollahi, H.; Ebrahimi, S.B. A new hybrid model for forecasting Brent crude oil price. *Energy* **2020**, *200*, 117520. [\[CrossRef\]](#)
- Abosedra, S.; Baghestani, H. On the predictive accuracy of crude oil futures prices. *Energy Policy* **2004**, *32*, 1389–1393. [\[CrossRef\]](#)
- Coppola, A. Forecasting oil price movements: Exploiting the information in the futures market. *J. Futures Mark.* **2008**, *28*, 34–56. [\[CrossRef\]](#)
- Xie, W.; Yu, L.; Xu, S.; Wang, S. A New Method for Crude Oil Price Forecasting Based on Support Vector Machines. In *International Conference on Computational Science*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 444–451. [\[CrossRef\]](#)
- Morana, C. A semiparametric approach to short-term oil price forecasting. *Energy Econ.* **2001**, *23*, 325–338. [\[CrossRef\]](#)
- Agnolucci, P. Volatility in crude oil futures: A comparison of the predictive ability of GARCH and implied volatility models. *Energy Econ.* **2009**, *31*, 316–321. [\[CrossRef\]](#)
- Li, J.; Zhu, S.; Wu, Q. Monthly crude oil spot price forecasting using variational mode decomposition. *Energy Econ.* **2019**, *83*, 240–253. [\[CrossRef\]](#)
- Tularam, G.A.; Saeed, T. Oil-Price Forecasting Based on Various Univariate Time-Series Models. *Am. J. Oper. Res.* **2016**, *6*, 226–235. [\[CrossRef\]](#)
- Bisoi, R.; Dash, P.; Mishra, S. Modes decomposition method in fusion with robust random vector functional link network for crude oil price forecasting. *Appl. Soft Comput.* **2019**, *80*, 475–493. [\[CrossRef\]](#)
- Parvini, N.; Abdollahi, M.; Seifollahi, S.; Ahmadian, D. Forecasting Bitcoin returns with long short-term memory networks and wavelet decomposition: A comparison of several market determinants. *Appl. Soft Comput.* **2022**, *121*, 108707. [\[CrossRef\]](#)
- Wang, S.; Yu, L.; Lai, K.K. A Novel Hybrid AI System Framework for Crude Oil Price Forecasting. In *Chinese Academy of Sciences Symposium on Data Mining and Knowledge Management*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 233–242. [\[CrossRef\]](#)
- Chiroma, H.; Abdul-kareem, S.; Shukri Mohd Noor, A.; Abubakar, A.I.; Sohrabi Safa, N.; Shuib, L.; Fatihu Hamza, M.; Ya'u Gital, A.; Herawan, T. A Review on Artificial Intelligence Methodologies for the Forecasting of Crude Oil Price. *Intell. Autom. Soft Comput.* **2016**, *22*, 449–462. [\[CrossRef\]](#)
- Sehgal, N.; Pandey, K.K. Artificial intelligence methods for oil price forecasting: A review and evaluation. *Energy Syst.* **2015**, *6*, 479–506. [\[CrossRef\]](#)
- Mostafa, M.M.; El-Masry, A.A. Oil price forecasting using gene expression programming and artificial neural networks. *Econ. Model.* **2016**, *54*, 40–53. [\[CrossRef\]](#)
- Ding, Y. A novel decompose-ensemble methodology with AIC-ANN approach for crude oil forecasting. *Energy* **2018**, *154*, 328–336. [\[CrossRef\]](#)
- Yu, L.; Zhang, X.; Wang, S. Assessing Potentiality of Support Vector Machine Method in Crude Oil Price Forecasting. *EURASIA J. Math. Sci. Technol. Educ.* **2017**, *13*, 7893–7904. [\[CrossRef\]](#)

24. Bristone, M.; Prasad, R.; Abubakar, A.A. CPPCNDL: Crude oil price prediction using complex network and deep learning algorithms. *Petroleum* **2020**, *6*, 353–361. [\[CrossRef\]](#)
25. Oussalah, M.; Zaidi, A. Forecasting Weekly Crude Oil Using Twitter Sentiment of U.S. Foreign Policy and Oil Companies Data. In Proceedings of the 2018 IEEE International Conference on Information Reuse and Integration (IRI), Salt Lake City, UT, USA, 6–9 July 2018; IEEE: Salt Lake City, UT, USA, 2018; pp. 201–208. [\[CrossRef\]](#)
26. Reboredo, J.C.; Rivera-Castro, M.A. A wavelet decomposition approach to crude oil price and exchange rate dependence. *Econ. Model.* **2013**, *32*, 42–57. [\[CrossRef\]](#)
27. Wu, C.; Wang, J.; Hao, Y. Deterministic and uncertainty crude oil price forecasting based on outlier detection and modified multi-objective optimization algorithm. *Resour. Policy* **2022**, *77*, 102780. [\[CrossRef\]](#)
28. Miao, H.; Ramchander, S.; Wang, T.; Yang, D. Influential factors in crude oil price forecasting. *Energy Econ.* **2017**, *68*, 77–88. [\[CrossRef\]](#)
29. Miao, H.; Ramchander, S.; Wang, T.; Yang, J. The impact of crude oil inventory announcements on prices: Evidence from derivatives markets. *J. Futures Mark.* **2018**, *38*, 38–65. [\[CrossRef\]](#)
30. Cortazar, G.; Ortega, H.; Valencia, C. How good are analyst forecasts of oil prices? *Energy Econ.* **2021**, *102*, 105500. [\[CrossRef\]](#)
31. Zhang, J.L.; Zhang, Y.J.; Zhang, L. A novel hybrid method for crude oil price forecasting. *Energy Econ.* **2015**, *49*, 649–659. [\[CrossRef\]](#)
32. Archer, E.; Park, I.; Pillow, J. Bayesian and Quasi-Bayesian Estimators for Mutual Information from Discrete Data. *Entropy* **2013**, *15*, 1738–1755. [\[CrossRef\]](#)
33. Baudot, P.; Tapia, M.; Bennequin, D.; Goillard, J.M. Topological Information Data Analysis. *Entropy* **2019**, *21*, 869. [\[CrossRef\]](#)
34. Khoojine, A.S.; Dong, H. Topological Structure of Stock Market Networks during Financial Turbulence: Non-Linear Approach. *Univers. J. Account. Financ.* **2019**, *7*, 106–121. [\[CrossRef\]](#)
35. Mann, A.D.; Gorse, D. A New Methodology to Exploit Predictive Power in (Open, High, Low, Close) Data. In *International Conference on Artificial Neural Networks*; Springer: Cham, Switzerland, 2017; pp. 495–502. [\[CrossRef\]](#)
36. Khoojine, A.S.; Han, D. Stock price network autoregressive model with application to stock market turbulence. *Eur. Phys. J. B* **2020**, *93*, 133. [\[CrossRef\]](#)
37. Sioofy Khoojine, A.; Shadabfar, M.; Hosseini, V.R.; Kordestani, H. Network Autoregressive Model for the Prediction of COVID-19 Considering the Disease Interaction in Neighboring Countries. *Entropy* **2021**, *23*, 1267. [\[CrossRef\]](#)
38. Déès, S.; Karadeloglou, P.; Kaufmann, R.K.; Sánchez, M. Modelling the world oil market: Assessment of a quarterly econometric model. *Energy Policy* **2007**, *35*, 178–191. [\[CrossRef\]](#)
39. Zhu, J.; Long, H.; Deng, J.; Wu, W. Risk measurement of oil price based on Bayesian nonlinear quantile regression model. *Alex. Eng. J.* **2021**, *60*, 5567–5578. [\[CrossRef\]](#)
40. Khoojine, A.S.; Han, D. Network analysis of the Chinese stock market during the turbulence of 2015–2016 using log-returns, volumes and mutual information. *Phys. A Stat. Mech. Its Appl.* **2019**, *523*, 1091–1109. [\[CrossRef\]](#)
41. Box, G.E.; Jenkins, G.M.; Reinsel, G.C. *Time Series Analysis, Forecasting and Control*; John Wiley & Sons: Hoboken, NJ, USA, 1998.
42. Giada, L.; Marsili, M. Algorithms of maximum likelihood data clustering with applications. *Phys. A Stat. Mech. Its Appl.* **2002**, *315*, 650–664. [\[CrossRef\]](#)
43. Paoletta, M.S. *ARMA Model Identification*; Springer Science & Business Media: Cham, Switzerland, 2018; pp. 405–442. [\[CrossRef\]](#)
44. Chen, C.; Twycross, J.; Garibaldi, J.M. A new accuracy measure based on bounded relative error for time series forecasting. *PLoS ONE* **2017**, *12*, e0174202. [\[CrossRef\]](#)