Forecasting Oil Prices: A Comparative Analysis between Neural Network and Regression Models

Jihad el Hokayem[1,*], Joseph Gemayel[1], Dany Mezher[3] and Ale Hejase[2]

[1]*Faculty of Economics, Saint Joseph University of Beirut, Lebanon*.
[2]*Adnan Kassar School of Business, Lebanese American University, Beirut, Lebanon*.
[3]*Faculty of Engineering, Saint Joseph University of Beirut, Lebanon*.

Abstract: After the war between Russia and Ukraine and its implications on various economies, energy security became a trending subject at the international level in 2022. Crude oil is an essential resource that plays a strategic role and its fluctuation has a major impact not only on the firms' profitability but also the stability of several countries. This research examines the possibility of forecasting oil prices where artificial neural network methods in addition to the multiple linear regressions were used in order to attempt to come out with a decent model that can help in forecasting oil prices. A comparison between these two models took place in order to choose the best one. The uniqueness of this research relied on the fact that 26 different variables were used all together, some of them were used for the first time in order to build the forecasting model. The period for constructing and testing both models extended from August 2006 to the beginning of 2019. The neural network model that was built showed to be more promising than the model that used multiple linear regression.

Keywords: Oil Prices, Forecasting, Economics, Regression analysis, Neural Network.

INTRODUCTION

For over several epochs, crude oil has been continuously generated in reservoirs across the globe. As it took millions of years for crude oil to finally reach its ultimate form, countries that are in possession of such a natural resource are considered to be holding on to an extremely powerful economic asset. Crude oil is considered to be an energy product that is a vital driver for most economic movements (Cheong, 2009). Furthermore, crude oil can be extracted and further refined into various types of products, as countries around the world use its derivatives in their daily operations. As the oil and gas drilling sector is demonstrated to constitute three percent of the worldwide gross domestic product (GDP), fluctuations in crude oil prices have been bared to have a substantial influence on the world economy on innumerable levels (Miao et al., 2017). For instance, when the prices of oil fell from their record high of $145 per barrel in 2008 to a low of $29 per barrel in 2014, massive economic stress was placed on several oil-exporting countries, causing a sharp downfall in revenues (Miao et al., 2017). Moreover, with the effect that crude oil prices have on a country's macro economy, it often influences various variables which include inflation, GDP, interest rates, exchange rates, as well as stock market prices (Ferrari et al., 2021). In turn, such an impact would cause the state of the economy of each country to be immensely affected by the oscillations in oil prices with the effect being either constructive or deleterious.

Throughout the years, several events have caused the oil market to witness variations in the prices of its commodities. With the presence of several types of crude oil internationally, the WTI (West Texas Intermediate) crude oil is considered to be the representative of the world crude oil, as it influences the movement of other crude oil prices, such as that of the Brent crude oil (Asianto, 2019). Consecutively, oscillations in the prices of the WTI crude oil eventually end up distressing the economies of many countries. Due to the oil market gaining significant strength globally and becoming a strategic part of many economies, oil prices have become subjected to a series of stimuli (Fan and Xu, 2011). This in turn has caused oil prices to become affected by various financial factors, rather than solely by the conventional supply and demand channels. With oil prices countersigning rapid variations, worldwide economies often perceive a series of turn of events that could be rather appealing or drastic (Byrne et al., 2019). Subsequently, this has led oil prices to witness a series of interactions with a set of variables that could be broken down into financial, macroeconomic, and geopolitical (Cuando et al., 2020). Moreover, with the variations in oil prices, and especially with their increases, fear starts to arise around its supply along with the trepidation about oil scarcity (Fattouh, 2007; Kaufmann and Cleveland, 2001). On account of this, the erratic fluctuation in oil prices has caused numerous researchers and policymakers around the world to study the determinants behind the changes in oil prices as a means to better understand their nature. Cuando et al. (2020) state that the impact of macroeconomic and financial variables has been extensively studied, unlike that of geopolitical factors, which are still yet to be further studied in detail. As a result, tremendous research has been done in order to fully understand the impact of such factors on the variation of oil prices and attempt to construct the connection among them.

Due to the reasons aforementioned, studying the changes in oil prices and the impact it imposes on a country's economy

*Address correspondence to this author at Faculty of Economics, Saint Joseph University, Philadelphia; Cell: 00961(0)3918877,
Tel: +961 (1) 421 644, Fax: +961 (1) 421 649;
E-mail: jihad.hokayem@usj.edu.lb

grows in importance. As oil prices are affected by several economic and political factors, it is crucial to determine how those prices will eventually be altered. Forecasting oil prices has for long been a topic of interest for many researchers as they aim to determine how prices change over time. Moreover, as oil prices are subjected to recurring fluctuations over the different seasons of the year, research has been focused on continuously developing state-of-the-art methods to forecast oil prices. As oil prices have become a point of interest for many economists and decision-makers around the world, tremendous research has been carried on to determine which variables play a role in affecting the forecasts of crude oil prices and the different techniques used to do so (Yin and Yang, 2016). With ongoing technological advancements taking place throughout the years, several techniques that were employed as a tool to forecast crude oil prices began to arise. Ranging from standard regression techniques to more complex machine learning and artificial neural network models, such models have been continuously utilized to forecast oil prices while being compared to one another in order to determine the efficiency of each technique. Moreover, due to the complex characteristics that crude oil often possesses such as its nonlinearity and volatility, various methods have been incessantly being experimented with the aim of studying the changes in oil prices. Nevertheless, as forecasting crude oil prices is affected by the modeling strategy used in terms of data, methodology, and sample period, it is a perplexing task to correctly select which forecasting model to apply in order to obtain the most significantly appealing results (Zhang et al., 2015). Hence, assiduously testing and introducing new models to forecast crude oil prices could remain a point of interest for many researchers and policymakers around the world for years to come.

Based on all the aforementioned oil price determinants, this current research attempts to unify the determinants of oil price fluctuations as well as to provide a model that could assist market players to predict future prices since its large fluctuation affects regional security and stability, economic growth, and the profitability of corporations. Thus, forecasting the future fluctuation of oil prices is essential for being proactive. Moreover, it enables taking the appropriate measures to counter-attack potential risks.

As per the literature, previous models showed mixed signals in terms of oil forecasting and their accuracy was questioned on numerous occasions, especially that oil prices are believed to follow a random walk (Fama, 1995). Our paper will fill the gap by identifying the maximum number of oil price factors in order to build an exhaustive model. A comparison between both approaches, between multiple linear regression and artificial neural network, will be made to choose the best model for the period covered in our paper.

LITERATURE REVIEW

Oil Crack Spread

Crack spread refers to the price difference between crude oil and its refined products (Wang et al., 2015). In other words, as the crack spread decreases, oil prices regularly decrease. Several players are normally concerned with the crack spread. On one side, refineries are apprehensive about the crack spread as they continuously seek to protect the spread. On the other side, as oil consumers are actively exposed to the crack spread, they often tend to take actions that would minimize their risks (Murat and Tokat, 2009). Choi et al. (2015) utilized an ECM and MGARCH model to determine the prices of WTI crude oil, while focusing on the impact of the crack spread. Their results showed that both models based on the crack spread play a significant role in determining the movements in the prices of crude oil. Yin et al. (2018) employed a set of variables, which include the crack spread in a time-varying parameter (TVP) model, to establish their influence on the prices of crude oil. While focusing on the MSPE ratio, their model demonstrated significantly appealing results on all horizons and was able to overcome the random walk. Similarly, Baumeister et al. (2018) implemented a random walk model where they use the crack spread as a means to determine the prices of WTI crude oil and the real US refiners' acquisition cost for crude oil imports. Their results revealed that the crack spread plays a significant role in forecasting the prices of the WTI crude oil and the refiners' acquisition cost, with the MSPE witnessing larger reductions for the refiners' acquisition cost compared to the WTI. Moreover, they proposed how further TVP models could enhance the results associated with the crack spread.

Demand and Supply

With crude oil being a key strategic aspect of the international economic system, the demand and supply channels tend to play a role in influencing its prices and in turn its impact on global policymakers and traders (Wang and Sun, 2017). With forecasts being done to determine the reasons behind the crashes and booms in oil prices, demand and supply are normally associated with such incidents. Kim (2018) conducted a study, where he aimed to determine the reasons behind the deteriorating oil prices in the two periods of 2008 to 2009 and 2014 to 2016. He showed that the 2008 oil price drop was associated with the diminution in real demand associated with the financial crises, while the 2014 oil price drop was associated with the increase in the supply of shale oil (Kim, 2018). In a similar vein, Hamilton (2009) related the variations in the prices of oil to its demand characteristics. For instance, the drop from \$140 to \$60 per barrel in 2008 was associated to the drop in demand from industrial economies, such as China and the Middle East, as well as the low demand price elasticity, which exerted a pressure on the prices of crude oil. Krichene (2002) revealed the impact of the price inelasticity of both demand and supply on the prices of crude oil; he stated that a reduction in the supply and increase in the demand of crude oil would result in an increase in the prices of crude oil, while an increase in the supply and decrease in the demand of crude oil would result in a drop in its prices. Moreover, Breitenfellner et al. (2009) depicted the role of the demand and supply factors in driving the prices of crude oil. By taking various time subsamples, they reveal that the demand and supply factors were a key determinate in influencing the prices of crude oil between the years 2000 and 2008. Furthermore, they stated that the oil price inflation was triggered by the increase in demand during that period, especially with the strategic influence of OPEC.

Exchange Rate

As oil is priced in dollars, alterations in the U.S. dollar exchange rate may influence the prices of oil worldwide, as it may also affect oil exporters and importers (Pirog, 2005). For instance, if the value of the Euro appreciates against the U.S. dollar, an increase in the dollar price of oil would decrease its demand in the United States, however, it would increase its demand in Europe as a result of the Euro appreciation. Brahmasrene et al. (2014) implemented a VAR (Vector Autoregression) model to study the relationship between the exchange rate and variations in oil prices; they revealed that a shock in the exchange rate has a significantly negative effect on the prices of crude oil and changes in the exchange rate affect oil prices in the short run. Similarly, Albulescu and Ajmi (2021) used a VAR model to determine the impact of the exchange rate on oil prices during the 2008 global financial crisis. Their results showed that when the U.S. dollar witnessed phases of depreciation prior to the crisis, increases in international oil prices were witnessed. In parallel, with the outbreak of the 2008 global financial crisis, a plunge in international oil prices was witnessed as a result of the appreciation in the U.S dollar. Furthermore, Breitenfellner and Cuaresma (2008) employed an AR, VAR, and a VEC model to determine whether the U.S dollar exchange rate against the Euro affects oil prices. Their results showed that in the short run, the VAR model was the best in explaining such a relationship, however, in the long run, the VEC model performed better. This provided evidence that the exchange rate plays a role in determining the prices of oil in both the short and long run.

Economic Growth

As a country's economic growth is enhanced and incomes increase, oil consumption eventually begins to increase. This in turn creates a causal relationship between economic growth and fluctuations in crude oil prices (Huntington et al., 2012). Hanabusa (2009) conducted a study where he utilized an AR-EGARCH model to study the relationship between economic growth and oil prices between 2000 and 2008 in Japan. He showed that the relationship between the two is bidirectional, in which changes in oil prices Granger-cause economic growth and economic growth Granger-causes oil price changes as well. In a similar vein, Kilian and Hicks (2012) aimed to determine whether the oil price shocks between 2003 and 2008 were associated with economic growth. Their results revealed that the escalations in oil prices post-2003 were related to a set of factors those of which include changes in the global economic activity. Moreover, they stated that the economic growth resulted in increases in oil prices from 2003 up until mid-2008 before starting to decline as a result of the global financial crisis. In another study conducted by Gong et al. (2020) they utilized a GARCH-MIDAS model to establish the impact of macroeconomic variables including economic growth on the variations in oil prices. Their results showed that economic growth reduces the volatility of crude oil. Moreover, they depicted that economic growth had a significantly negative impact on the fluctuations in oil prices, where, in the short run, high economic growth fluctuations result in high oil price variations.

Speculation

Speculation within the oil market tends to play a role in influencing oil price movements as participants in the oil market are divided into hedgers who seek to minimize their exposure to risk and speculators who seek to make a profit as a result of the movement of oil prices (Huntington et al., 2012). Xiao and Wang (2022) implemented a study in which they aimed to determine the relationship between uncertainty and oil price fluctuations by utilizing the macroeconomic uncertainty (MU) index as an indicator. They employed a quantile regression model where their results showed that the MU index negatively the future returns of crude oil. Moreover, they claimed that the presence of speculation in the crude oil market plays a role in mitigating negative effect imposed by the MU alterations. In contrast, in a study conducted by Fattouh et al. (2013), where they sought to determine whether oil price fluctuations between 2003 and 2008 were driven by speculation, revealed that speculation was not the major factor behind those variations. Rather than speculation in the oil market being the driver behind oil price movements, they claimed that the interaction between the spot and future prices of oil played a crucial role in explaining the variations in oil prices.

Geopolitics

As petroleum resources are concerted in specific regions around the globe, a number of parties tend to initiate geopolitical events as a means to affect the supply of such resources, and in turn, their prices (Fan and Xu, 2011). As oil is a politically sensitive commodity and with several geopolitical events taking place over the years, such events have caused a substantial impact on the prices of oil with the latter being strongly affected by all of the political tensions (Su et al., 2019). In a study instigated by Duan et al. (2021), they implemented a Wavelet analysis to study the causal relationship between geopolitical events and oil prices on the Venezuelan economy. They used monthly data spanning from 2008 to 2019 as a means to highlight the effect of the global financial crisis which initiated in 2008 on oil prices. Their results revealed that the uncertainties associated with the geopolitical risks negatively affected oil prices which in turn further exacerbated the exchange rate as a result of solely relying on oil. In a similar vein, Su et al. (2021) sought to determine the relationship that oil prices hold with geopolitical risk. They constructed a VAR model to study the causal relationship between geopolitical risk and the variation in oil prices. By using monthly data between 1990 and 2018 to test for the Granger causality between the two variables, their results revealed that the causal relationship between geopolitical risk and oil prices is intermittent rather than continuous. They showed that during periods of risk and wars, geopolitics often influence oil prices leading to their upsurge as a result of the decrease in supply, however, when geopolitical risks decrease, oil prices do not necessarily decline. In parallel, Nyangarika (2019) employed an ARIMA model to forecast oil prices through the use of political events as an indicator. In contrast to previous studies, their results depicted that geopolitical risks and military conflicts in particular have an insignificant minimal effect on the fluctuations in oil prices.

Inflation

Inflation was recognized as a main factor in various countries; Baghestani (2014) proved in his paper that the expectations for the inflation of US consumers, accurately predict future oil prices. A study carried out by Osorio and Unsal (2013) claimed that the factors influencing inflation have transitioned over the past two decades in Asia from being caused by monetary and supply shocks to national demand burdens. Along with the influence of the Chinese economy, their results depicted that the increase in the demand for oil from Asian countries often influences the prices of oil and global commodities.

Volatility

Another factor that showed to have an influence on oil prices is volatility; Singleton (2010) related the upsurges in the prices of future oil prices, along with their elevated volatility followed by their strident drop towards the end of 2008, to non-user participants rather than just to the customary supply and demand factors. A different study conducted by Hammoudeh et al. (2004) showed that oil volatility holds a diffusion characteristic, which further played a role in the volatility transmission impact of oil prices on the volatility of oil stocks to vary from day to day. It results in a stable pattern where it starts off insignificant on Monday and finishes by gaining more significance on Friday.

Stock Market

With oil prices being influenced by the worldwide stock markets, several countries, and especially from Asia, witnessed such relationships. Previous studies have stated that stock markets witnessed a spillover phenomenon. Baele (2005) showed that the stock markets and the S&P 500 in particular tend to be the major cause behind those transmissible spillovers, whereas the commodity markets which include those of oil and gold often handle the consequences of such a burden. Moreover, Adam et al. (2015) utilized an LVAR model to study the dynamic relationship between oil prices and the stock market in Indonesia between 2004 and 2013. Their results revealed that during the short and long run, both factors witnessed a significant relationship, which further ascertained their dynamics. On a similar note, Zhu et al. (2014) carried out a study to highlight the relationship between crude oil prices and the stock market in Asia-Pacific countries. By implementing their study between 2000 and 2012, Zhu et al. (2014) reveal that the relationship between the stock and oil prices was not significant prior to the 2008 global financial crisis, whereas succeeding the crisis, the relationship turned into a significant one between the two.

METHODOLOGY

Fahrmeir, Kneib, Lang, and Marx (2007), in their book titled *Regression*, considered regression as the most popular and commonly used statistical methodology in order to analyze empirical problems in social sciences, economics, and life sciences. In line with Fahrmeir, Kneib, Lang, and Marx (2007), Welc and Esquerdo (2018) considered the single-equation linear regression model to be the most commonly applied class of econometric models. In their book *Applied Regression Analysis for Business*, Welc and Esquerdo (2018) defined the regression model as an econometric model and a quantitative analytical tool in which the behavior of certain variables is explained by other variables. A single equation regression model is a mathematical function that has the form of an equation, which quantifies a relationship between a dependent variable, and one or more explanatory variables. The explanatory variables, which are called regressors, have statistical or causal relationships with the dependent variable, which is explained by the model.

The primary elements of such models, according to Welc and Esquerdo (2018), are as follows: First, the coefficients or structural parameters that are associated with individual explanatory variables quantify the direction and impact the size of a given explanatory variable on the dependent variable. Second, the standard errors of individual structural parameters determine the expected ranges into which the true values of individual coefficients fall, as they may not be exactly equal to the estimated ones. Third, T-statistics of individual coefficients, calculated by dividing the value of a given coefficient by its respective standard error, are used to evaluate the statistical significance of explanatory variables and their individual coefficients as per Hejase and Hejase (2013). Thus, one of the key principles of model building is to include only the explanatory variables, which show statistically significant relationships with the dependent variable. Fourth, the coefficient of determination, also called R-squared, measures how good the estimated model to the actual data is. As per (Hejase & Hejase, 2013), the coefficient of determination R^2 or R-squared is computed with the following formula:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} e_i^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

Where "n" is the number of observations used in estimating a given model and "y_i" is the actual observation of the dependent variable, "e_i" is the error or residual difference between the actual value of the dependent variable "y_i" and its estimated value obtained from the regression model; finally, "$\bar{y}$" is the mean value of all actual observations of the dependent variable. The uniqueness of this research relies on the fact that *26 different variables* will be used together, some of them for the first time, in order to build, and then choose, the forecasting model using two different approaches, which are artificial neural network and multiple linear regression. From the concept of single-equation linear regression models, one can further dive into more advanced versions of such models. A subcategory that branches out is the stepwise regression model, which is innovative in its modeling process.

When dealing with stepwise regression, different subsets of independent variables are used for implementing the regressions. Yet, such an aspect is not always beneficial, as the test for significance comes out to be biased at most times along with an insignificant R^2 value (Wilkinson, 1979). Some scholars, such as Ruengvirayudh and Brooks (2016), argue that stepwise regressions may be used to obtain exploratory results and compared to different statistical techniques. A study performed by Miao et al. (2017) implements a stepwise regression to forecast crude oil prices. Through the use of stepwise regression, Miao et al. (2017) are able to eliminate which factors do not affect oil prices by checking their significance. However, their results call for the need to rely on an alternative regression method, such as the Least Absolute Shrinkage and Selection Operator (LASSO), as the stepwise regression may witness issues of multicollinearity. In contrast, Ferrari et al. (2021) utilize a stepwise regression along with a different set of regression techniques to forecast the prices of energy commodities. Their results showed that the stepwise regression performed poorly when forecasting the prices due to sparsity. As studies that used stepwise regression showed to have mixed results where some studies showed that it can be used to forecast time-series data while others depicted otherwise, alternative methods were proposed as a means to enhance the forecasting ability. With the aim of doing so, the ANN model was insinuated as it is considered to give resolute results.

Artificial Neural Networks (ANN) Technique

An Artificial Neuron Network is a model that consists of a set of vastly intertwined nodes, where each node performs a specific task, which generates an output that is carried over to the next node (Vaisla and Bhatt, 2010). Such a process allows ANN to break several barriers in data analysis, rendering it an advantageous tool to be used in econometric analysis. One advantage of using ANN is its application on nonlinear processes; as previous multiple regression models have been capable of predicting linear patterns only, ANN is proficient in modeling nonlinear processes without the need of having previous knowledge about the type of the process (Vaisla and Bhatt, 2010). With the continuous development of ANN and the ongoing training, which ANN go through as a result of the learning processes, ANN started being used as an efficient tool for forecasting. Several studies have tested the efficiency of using ANN to forecast non-linear time series while comparing them to traditional regression methods (Hill et al. 1994; Zhang et al. 2001). Hill et al. (1994) aim to compare the performance of ANN to traditional statistical models in forecasting time-series data. Throughout their study, they reveal that ANN performs similar to the statistical models and rather better in terms of forecasting time-series data. Moreover, in terms of dealing with business models, they claim that ANN may optimistically serve as a worthwhile alternative to forecast such applications. In a similar manner, Zhang et al. (2001) utilize an ANN model to forecast a non-linear time series and compare it to the standard regression models. Their results showed that it is crucial to select the correct number of input nodes for the model where doing so allows for a more accurate predictive ability. Moreover, they depict that ANN possesses a forecasting ability that is superior to ordinary regression models when dealing with nonlinear data. In a study that aims to forecast crude oil prices, Mirmirani and Li (2004) use an ANN model to do so while comparing it to a VAR model. They utilize monthly data that spans between 1980 and 2002, and use the root mean squared error and mean absolute error as criteria for assessment. Their results reveal that the root mean squared error and mean absolute error recorded by the ANN model were 1.2354 and 0.8629, respectively, which are much lower than those recorded by the VAR model, which were 4.69861 and 4.18883, respectively. Moreover, their findings reveal that the ANN model transcends the VAR model in performing forecasting applications. In a similar manner, Kulkarni and Haidar (2009) employ an ANN model to forecast crude oil prices in the short term. They use daily data ranging between September 1986 and August 2007 as it is preferred for short-term forecasting. Their results revealed a root mean squared error and mean absolute error of 0.01922 and 0.00038, respectively, which depicts the efficiency of using an ANN model to forecast short-term crude oil prices.

Twenty-six variables were used to carry on each of the two models, including:

- European Brent oil closing price;
- Delta 1 of the closing price of European Brent oil;
- WTI oil closing price;
- Delta 1 of the closing price of WTI oil;
- Implied volatility of oil options that measures the fear present within the oil market, and is calculated based on the Black and Scholes model;
- Oil over Equity which depicts the movement of US oil companies compared to those of the entire equity market;
- Delta 1 of the Oil over Equity;
- Price of oil in the Dumai port in Indonesia;
- Delta 1 of the price of oil in the Dumai port in Indonesia;
- Oil crack spread which shows the relationship between the prices of gasoline, heating oil and crude oil contracts;
- US index which compares the price of the USD to a basket of various currencies. Usually, oil prices and such an index carry a negatively correlated relationship;
- Inflation of different regions which include those of the US, UK, Euro Zone, Japan, China, and Germany;
- The economic growth (measured by the GDP) of different regions which include those of the US, UK, Euro Zone, Japan, China, and Germany;
- Imbalance between demand and supply;
- Total speculators buyers and sellers of oil;
- Total hedgers buyers and sellers of oil.

Our paper will cover the period between August 4, 2006, and January 4, 2019, which consists of 649 weeks. This period comprises two major economic events which include the 2007-2008 financial crisis and the rise of the shale oil era in 2014.

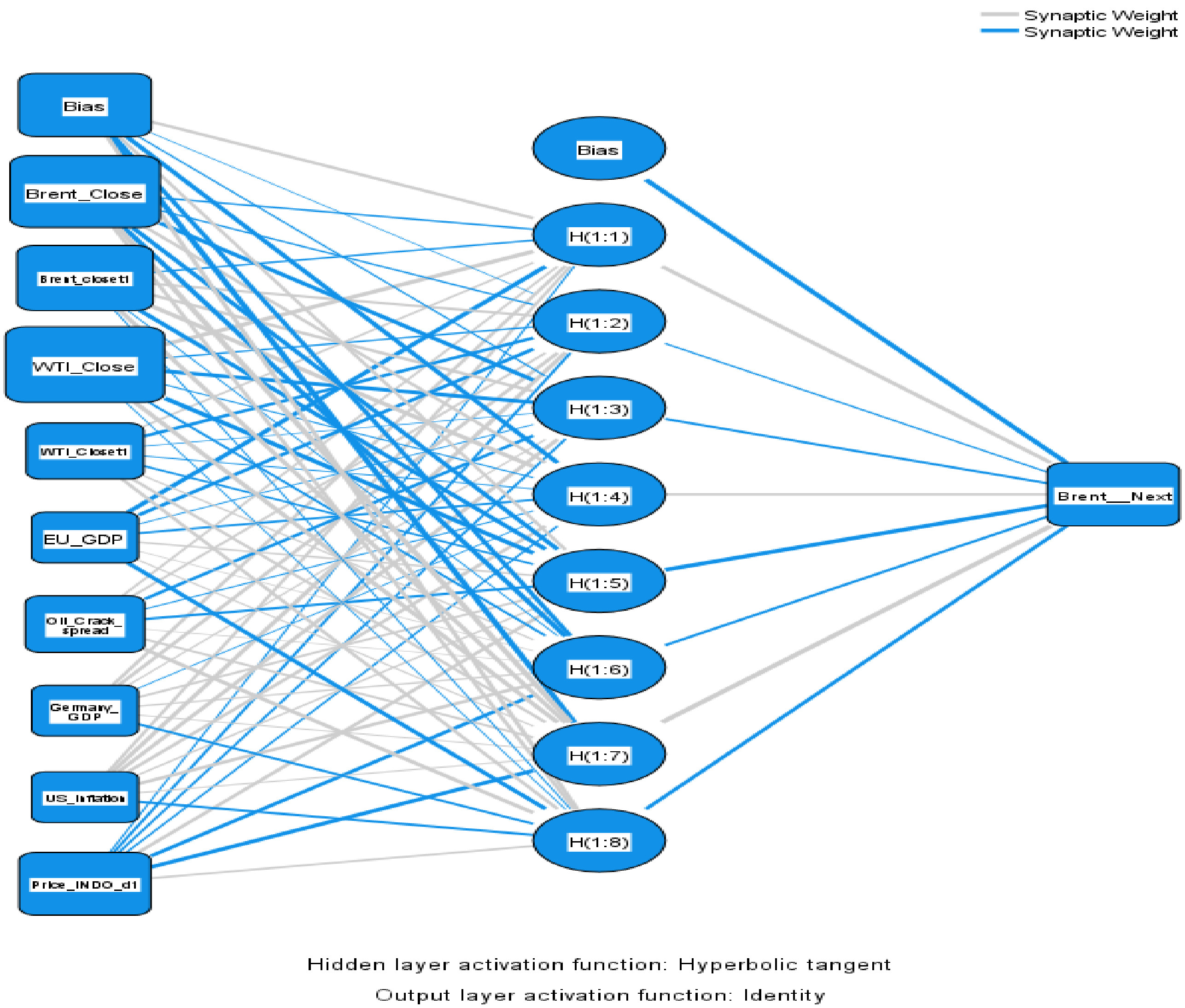

Fig. (1). Proposed ANN model to forecast Brent price for next week.

RESULTS

Artificial Neural Network Approach

With a series of runs being implemented, their results depicted that several of the recommended covariates caused a slight impact on the Brent Close price for the next week which is the dependent variable under study (A normalized importance above 30% was considered to be relevant). Thus, different runs were done to identify the most relevant covariates and to eliminate explanatory variables that presented minor contributions to the future oil price.

The first group of runs, consisting of all suggested covariates, resulted in only 9 covariates, which could play a role in the forecasting of the future next week's oil price. Those covariates were: WTI close price of the week, WTI close price of the previous week, Brent close price of the week, Brent close price of the previous week, EU GDP, Oil Crack Spread, Germany GDP, US inflation, and delta1 for the price of oil in Dumai port in Indonesia.

In the second group of runs, only the aforementioned 9 covariates were used, but this time the data was split into three parts: 400 weeks for training covering the period from August 4, 2006, till March 28, 2014 (61.63% of the data), 200 weeks for testing covering the period from April 4th, 2014, till January 26th, 2018 (30.81% of the data), and the last 49 weeks of data covering the period from February 2nd, 2018, till January 1st, 2019 (7.55%). The three parts were kept in hold to verify the performance of the optimal forecasting ANN that ended up with a single hidden layer, which includes 8 neurons (excluding the bias unit), as seen in Fig. (**1**).

The best ANN given by SPSS, "Statistical Product and Service Solutions" (Hejase and Hejase, 2013) is established as depicted in Fig. (**1**), which estimates the Brent Close price for the next week given the values of the independent explanatory influencing factors of the current and previous weeks. Moreover, Fig. (**2**) shows the distribution of how well the ANN can match the predicted values to their actual counterparts.

Fig. (2). Predicted versus actual values of Brent oil price.

The Predictors, which includes the WTI close, WTI close of the previous week, Brent close, Brent close of the previous week, EU GDP, Oil Crack Spread, Germany GDP, US inflation, and delta1 for the price of oil in Dumai port in Indonesia resulted in a model relative error of 0.037 for training, 0.027 for testing and 0.197 for holdout. Moreover, the importance percentages for the nine covariates are, as depicted in Table **1**.

Table 1. Independent Variable Importance

	Importance	Normalized Importance
Brent Close (t)	0.246	83.5%
Brent Close (t-1)	0.169	57.5%
WTI Close (t)	0.294	100.0%
WTI Close (t-1)	0.064	21.7%
EU GDP	0.005	1.6%
Oil Crack Spread	0.073	24.7%
Germany GDP	0.003	1.0%
US Inflation	0.005	1.6%
Price INDO delta1	0.143	48.5%

Fig. (**3**) shows the normalized importance of the nine relevant explanatory variables that mainly contribute to forecasting future oil prices.

Fig. (**4**) shows how the ANN predicted values match the actual values, whereas previously mentioned, the model was verified using the data of 49 weeks from the 2nd of February 2018 till the 4th of January 2019. For Fig. (**4**), the Sum of Squares of Errors (SSE) came to be 592.66 and it can be said that the model in general performs very well in tracking the real oil price. In addition, Fig. (**5**) presents the histogram of residuals between actual and predicted oil price values for the holdout period of 49 weeks. The subsequent statistics were calculated for these residuals: Mean = \$2.65, Median = \$1.99, Standard Deviation = \$2.11, Mode = \$0.0, Minimum = \$0.0, and Maximum = \$7.64.

Stepwise Regression Approach

In order to carry on the stepwise regression, 25 regressors were chosen to determine their impact on the closing price of the Brent oil. When implementing the stepwise regression, the results showed to have diversified results. With some variables having a significant effect on the closing price of Brent oil, others turned out to have an insignificant effect. Moreover, some of the variables had a negative influence, while others had a positive influence on the Brent oil price. For instance, a one percent increase in each of the variables imbalance between demand and supply, UK GDP, oil over equity, US inflation, and implied volatility caused an increase in the closing price of Brent oil by 1.90, 0.01, 0.005, 0.01, and 0.0003 percent, respectively. Yet, out of those five variables that caused an increase in the Brent oil price, the effect of implied volatility was insignificant. In parallel, a one percent increase in each of the variables Japan inflation, China GDP, Delta 1 of the closing price of WTI oil, and US index caused a decrease in the closing price of Brent oil by 0.0017, 0.002, 0.004, and 0.0014 percent, respectively. However, among those variables that caused a decrease in the Brent oil price, the effect of Japan inflation, Delta 1 of the closing price of WTI oil, and US index were insignificant. However, it is worth noting that the R-squared of the following regression turned out to be 0.081, which shows that the regressors used in this model only explain 8.1% of the variation in the closing price of Brent oil. As a result, the stepwise regression approach did not turn out to be as a reliable model to determine the prices of crude oil.

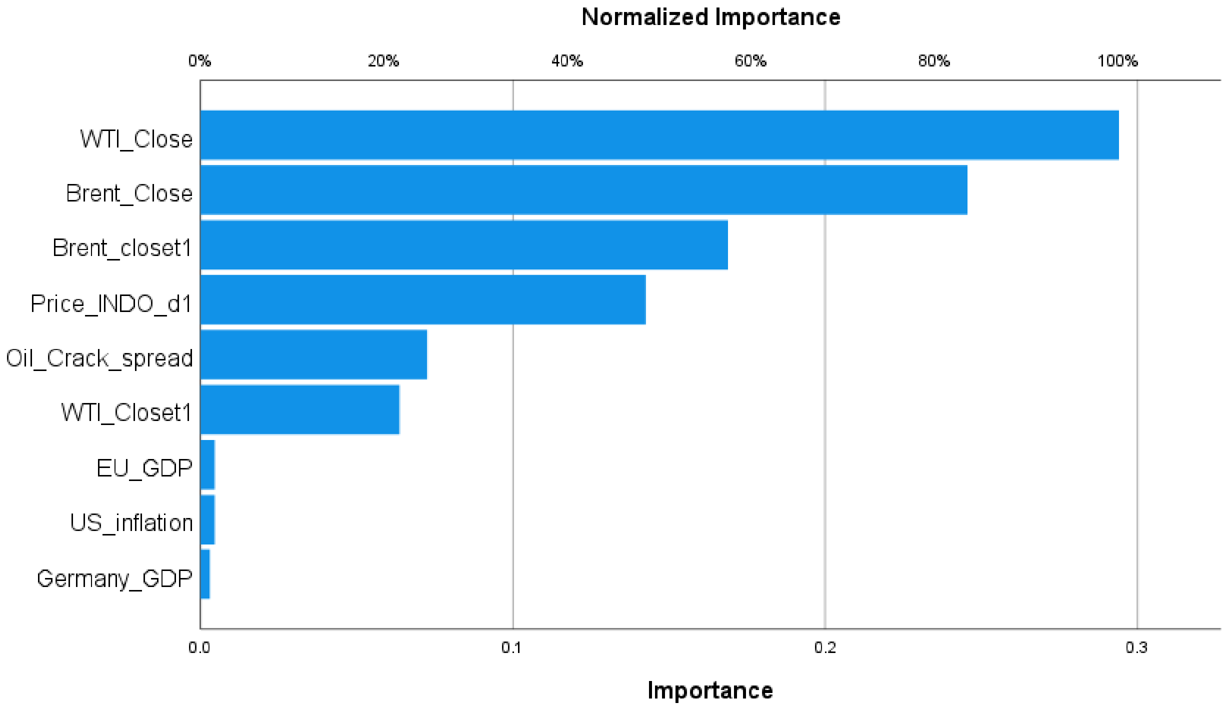

Fig. (3). ANN forecasting model normalized importance of covariates

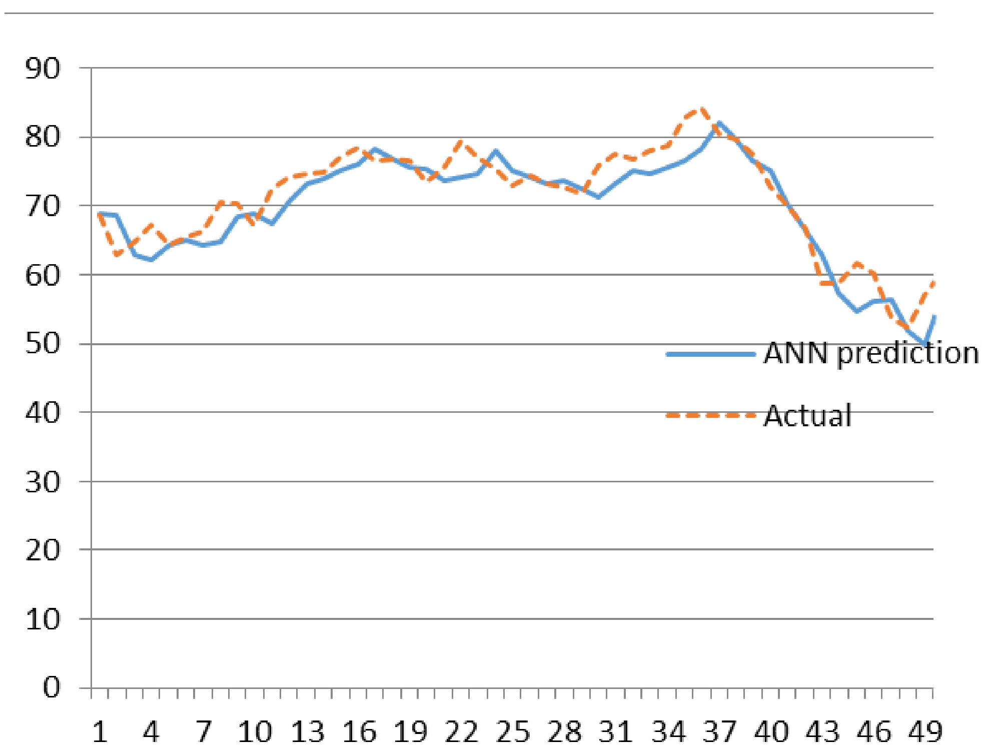

Fig. (4). Comparing the Actual with Predicted oil price values for the 49 weeks holdout data.

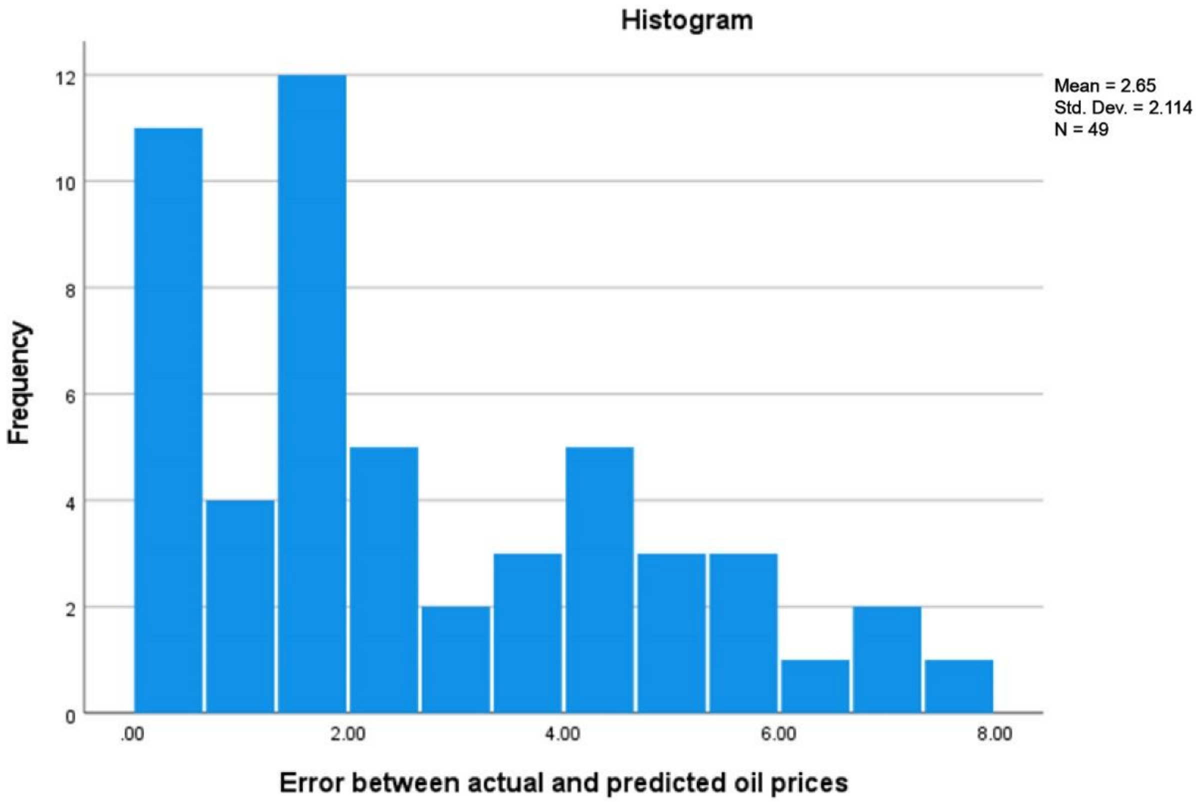

Fig. (5). Histogram for the 49 holdout data errors between actual and predicted oil prices.

Dependent Variable: BRENT_CLOSE				
Method: Stepwise Regression				
Date: March 9, 2022 Time: 13:44				
Sample (adjusted): Aug 4, 2006 – Feb 2, 2018				
Included observations: 614 after adjustments				
Number of always included regressors: 1				
Number of search regressors: 25				
Selection method: Stepwise forwards				
Stopping criterion: p-value forwards/backwards = 0.5/0.5				
Variable	Coefficient	Std. Error	t-Statistic	Prob.*
C	0.008533	0.011814	0.722250	0.4704
JAPAN_INFLATION(-1)	-0.001681	0.001757	-0.957013	0.3389
IMBALANCE_DEMAND_SUPP(-1)	1.891244	0.505765	3.739369	0.0002
UK_GDP	0.007603	0.003545	2.144796	0.0324
D1_OIL_ETF_TO_S&P(-1)	0.004688	0.002286	2.050559	0.0407
CHINA_GDP(-1)	-0.002045	0.000964	-2.122223	0.0342
US_INFLATION(-1)	0.013465	0.006680	2.015770	0.0443
D1_WTI(-1)	-0.003801	0.002358	-1.611997	0.1075
Implied Volatility(-1)	0.000266	0.000281	0.944899	0.3451
DOLLAR_INDEX(-1)	-0.001414	0.001593	-0.887634	0.3751

R-squared	0.081014	Mean dependent var	2.12E-05
Adjusted R-squared	0.057963	S.D. dependent var	0.045497
S.E. of regression	0.044159	Akaike info criterion	-3.376341
Sum squared resid	1.166094	Schwarz criterion	-3.261162
Log likelihood	1052.537	Hannan-Quinn criter.	-3.331550
F-statistic	3.514487	Durbin-Watson stat	2.163182
Prob(F-statistic)	0.000008		

Fig. (6). Results of the stepwise regression model.

DISCUSSION

This paper was designed to predict the prices of Brent crude oil through the use of two different approaches with the aim of determining which one performs better. The results showed that the stepwise regression model was not as efficient in determining crude oil prices when compared to the ANN approach. In parallel, the ANN performed relatively well in matching the predicted values of the oil prices to the actual prices. The results of the stepwise regression are similar to the findings of Fama (1995), who states that oil prices are said to follow a random walk hypothesis. As a low percentage of the regressors played a role in explaining the variation of the Brent oil closing price (R-squared= 8%), the stepwise regression was inefficient in predicting the prices of crude oil.

In parallel, the results of the ANN model align with those of Zhang et al. (2001), who stated that the ANN model possessed a better ability in determining oil prices than standard regression models. That was the case in this study as in fact, the ANN did perform better than the stepwise regression model. Furthermore, the results were in line with those of Kulkarni and Haidar (2009), whose results showed that the ANN played a significant role in forecasting crude oil prices in the short term. Similarly, our results depicted how the ANN provided significant results when determining the prices of the Brent oil price. Consecutively, it can be claimed that the ANN is a promising tool when dealing with crude oil prices. This would allow researchers and policymakers to make the right decisions in order to achieve the most efficient results and returns.

CONCLUSION

Crude oil is still a major subject that interests businesses and impacts the global economy. As its fluctuation is subject to various determinants, which tend to either positively or negatively alter prices, the profitability of corporations, and the state of the economy of several countries will be affected. Thus, understanding and forecasting oil prices is still the main topic for researchers and decision-makers in the public and private sectors. Moreover, by developing such an understanding, policy-makers will be able to take the appropriate decisions to manage price fluctuations.

This paper constructed an artificial neural network to forecast oil prices using a set of 26 distinct variables. Furthermore, it aimed to draw a comparison between neural network and regression models. Such a comparison was executed as a means to know which method is more reliable to forecast oil prices, executed in the period from August 2006 until the beginning of 2019, and assisted with weekly observations using the previously listed variables. With the results of the stepwise regression showing insignificant effects for most of the selected variables, the implementation of the artificial neural network demonstrated to be a more reliable method to forecast the Brent oil prices using the combination of the variables over the period under study. As the Russia-Ukraine war emerged recently, forecasting the direction of oil prices grew in importance, especially after it triggered a surge in inflation at the global level, putting energy and food security at risk for various countries.

It would also be interesting if other researchers use other methods and variables on the same data to compare the results with the current findings.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

REFERENCES

Adam, Pasrun, Usman Rianse, Edi Cahyono, and Manat Rahim. 2015. “Modeling of the dynamics relationship between world crude oil prices and the stock market in Indonesia”. *International Journal of Energy Economics and Policy 5*(2): 550.

Albulescu, Claudiu Tiberiu Ahdi Noomen Ajmi. 2021. “Oil price and US dollar exchange rate: Change detection of bi-directional causal impact”. *Energy Economics, 100*, 105385.

An, Jaehyung, Alexey Mikhaylov, and Nikita Moiseev. 2019. “*OIL PRICE PREDICTORS: MACHINE LEARNING APPROACH”. International Journal of Energy Economics and Policy 9(5): 1–6*.doi:10.32479/ijeep.7597

Asianto, Abitur. 2019. “The Price Determinants of West Texas Intermediate Crude Oil”. *Jurnal Aplikasi Bisnis dan Manajemen (JABM) 5*(1): 153-153.

Baele, Lieven. 2005. “Volatility spillover effects in European equity markets”. *Journal of Financial and Quantitative Analysis*, 373-401.

Baghestani, Hamid. 2014. “Inflation expectations and energy price forecasting”. *OPEC Energy Review* 38(1): 21-35.

Baumeister, Christiane, Lutz Kilian, and Xiaoqing Zhou. 2018. “Are product spreads useful for forecasting oil prices? An empirical evaluation of the Verleger hypothesis”. *Macroeconomic Dynamics 22*(3): 562-580.

Brahmasrene, Tantatape, Jui-Chi Huang, and Yaya Sissoko. 2014. “Crude oil prices and exchange rates: Causality, variance decomposition and impulse response”. *Energy Economics 44*: 407-412.

Breitenfellner, Andreas and Jesus Crespo Cuaresma. 2008. “Crude oil prices and the USD/EUR exchange rate”. *Monetary Policy & the Economy* (4).

Breitenfellner, Andreas, Jesus Crespo Cuaresma, and Catherine Keppel. 2009. “Determinants of crude oil prices: supply, demand, cartel or speculation”. *Monetary Policy & the Economy 4*(4): 111-136.

Byrne, Joseph Paul, Marco Lorusso, and Bing Xu. 2019. "Oil prices, fundamentals and expectations". *Energy Economics 79*: 59-75.

Cheong, Chin Wen. 2009. "Modeling and forecasting crude oil markets using ARCH-type models". *Energy policy 37*(6): 2346-2355.

Choi, Hankyeung, David J. Leatham, and Kunlapath Sukcharoen. 2015. "Oil price forecasting using crack spread futures and oil exchange traded funds". *Contemporary Economics 9*(1): 29-44.

Cunado, Juncal, Rangan Gupta, Chi Keung Lau, and Xin Sheng. 2020. "Time-varying impact of geopolitical risks on oil prices". *Defence and Peace Economics 31*(6): 692-706.

Duan, Wenqi, Adnan Khurshid, Abdur Rauf, Khalid Khan, and Adrian Cantemir Calin. 2021. "How geopolitical risk drives exchange rate and oil prices? A wavelet-based analysis". *Energy Sources, Part B: Economics, Planning, and Policy 16*(9): 861-877.

Fama, Eugene F. 1995. "Random walks in stock market prices". *Financial analysts journal 51*(1): 75-80

Fan, Ying, & Jin-Hua Xu. 2011. "What has driven oil prices since 2000? A structural changeperspective". *Energy Economics 33*(6): 1082-1094.

Fattouh, Bassam. 2007. "The drivers of oil prices: the usefulness and limitations of non-structural model, the demand-supply framework and informal approaches". *Oxford Institute for Energy Studies.*

Fattouh, Bassam, Lutz Kilian, and Lavan Mahadeva. 2013. "The role of speculation in oil markets: What have we learned so far?". *The Energy Journal 34*(3).

Ferrari, Davide, Francesco Ravazzolo, and Joaqin Vespignani. 2021. "Forecasting energy commodity prices: A large global dataset sparse approach". *Energy Economics 98,* 105268.

Gong, Xu, Mingchao Wang, and Liuguo Shao. 2020. "The impact of macro economy on the oil price volatility from the perspective of mixing frequency". *International Journal of Finance & Economics*.

Gabralla, Lubna, Rania Jammazi, and Ajith Abraham. 2013. "Oil price prediction using ensemble machine learning". *International Conference on Computing, Electrical and Electronic Engineering (ICCEEE):* 674-679

Hamilton, James D. 2009. "Understanding crude oil prices". *The Energy Journal 30*(2): 179-206.

Hammoudeh, Shawkat, Selahattin Dibooglu, and Eisa Aleisa. 2004. "Relationships among US oil prices and oil industry equity indices". *International Review of Economics & Finance* 13(4): 427-453.

Hanabusa, Kunihiro. 2009. "Causality relationship between the price of oil and economic growth in Japan". *Energy Policy 37*(5): 1953-1957.

Hejase, Ale J. and Hussin J. Hejase. 2013. "Research Methods: A Practical Approach for Business Students". *Masadir Inc*

Hill, Tim, Leorey Marquez, Marcus O'Connor, and William Remus. 1994. "Artificial neural network models for forecasting and decision making". *International journal of forecasting 10*(1): 5-15.

Huntington, Hillard, Zhuo Huang, Saud M. Al-Fattah, Michael Gucwa, and Ali Nouri. 2012. "Oil markets and price movements: A survey of determinants". *SSRN Electronic Journal 2274543*.

Kaufmann, Robert K. and Cutler Cleveland. 2001. "Oil production in the lower 48 states: economic, geological, and institutional determinants". *The Energy Journal 22*(1).

Kilian, Lutz and Bruce Hicks. 2013. "Did unexpectedly strong economic growth cause the oil price shock of 2003–2008?". *Journal of Forecasting 32*(5): 385-394.

Kim, Myung Suk. 2018. "Impacts of supply and demand factors on declining oil prices". *Energy 155*: 1059-1065.

Krichene, Noureddine. 2002. "World crude oil and natural gas: a demand and supply model". *Energy economics 24*(6): 557-576.

Kulkarni, Siddhivinayak and Imad Haidar. 2009. "Forecasting model for crude oil price using artificial neural networks and commodity futures prices". *International Journal of Computer Science and Information Security* 2(1).

Miao, Hong, Sanjay Ramchander, Tianyang Wang, and Dongxiao Yang. 2017. "Influential factors in crude oil price forecasting". *Energy Economics 68*: 77-88.

Mirmirani, Sam and His Cheng Li. 2004. "A comparison of VAR and neural networks with genetic algorithm in forecasting price of oil. Applications of artificial intelligence in finance and economics". *Emerald Group Publishing Limited.*

Murat, Atilim and Ekin Tokat. 2009. "Forecasting oil price movements with crack spread futures". *Energy Economics 31*(1): 85-90.

Nyangarika, Anthony. 2019. "Oil price factors".

Osorio, Carolina and D. Filiz Unsal. 2013. "Inflation dynamics in Asia: Causes, changes, and spillovers from China". *Journal of Asian Economics* (24): 26-40.

Pirog, Robert. 2005. "World oil demand and its effect on oil prices". *Library of Congress Washington DC Congressional Research Service.*

Singleton, Kenneth J. 2010. "The 2008 boom/bust in oil prices".

Su, Chi-Wei, Khalid Khan, Ran Tao, and Moldovan Nicoleta-Claudia. 2019. "Does geopolitical risk strengthen or depress oil prices and financial liquidity? Evidence from Saudi Arabia". *Energy 187*, 116003.

Su, Chi-Wei, Meng Qin, Ran Tao, and Moldovan Nicoleta-Claudia. 2021. "Is oil political? From the perspective of geopolitical risk". *Defence and Peace Economics 32*(4): 451-467.

Vaisla, Kunwar Singh and Ashutosh Kumar Bhatt. 2010. "An analysis of the performance of artificial neural network technique for stock market forecasting". *International Journal on Computer Science and Engineering 2*(6): 2104-2109.

Wang, Yudong, Li Liu, Xundi Diao, and Chongfeng Wu. 2015. "Forecasting the real prices of crude oil under economic and statistical constraints". *Energy Economics 51*: 599-608.

Wang, Qingfeng and Xu Sun. 2017. "Crude oil price: Demand, supply, economic activity, economic policy uncertainty and wars–From the perspective of structural equation modelling (SEM)". *Energy 133*: 483-490.

Xiao, Jihong and Yudong Wang. 2022. "Macroeconomic uncertainty, speculation, and energy futures returns: Evidence from a quantile regression". *Energy 241*, 122517.

Yin, Xuluo, Jiangang Peng, and Tian Tang. 2018. "Improving the forecasting accuracy of crude oil prices". *Sustainability 10*(2): 454.

Yin, Libo and Qingyuan Yang. 2016. "Predicting the oil prices: do technical indicators help?". *Energy Economics 56*: 338-350.

Zhang, G. Peter, B. Eddy Patuwo, and Micheal Y. Hu. 2001. "A simulation study of artificial neural networks for nonlinear time-series forecasting". *Computers & Operations Research 28*(4): 381-396.

Zhang, Jin-Liang, Yue-Jun Zhang, and Lu Zhang. 2015. "A novel hybrid method for crude oil price forecasting". *Energy Economics 49*: 649-659.

Zhao, Yang, Jianpiang Li, and Lean Yu. 2017. "A deep learning ensemble approach for crude oil price forecasting". *Energy Economics 66*: 9-16.

Zhu, Hui-Ming, Rong Li, and Sufang Li. 2014. "Modelling dynamic dependence between crude oil prices and Asia-Pacific stock market returns". *International Review of Economics & Finance 29:* 208-223.

Received: Jun 14, 2022 Revised: Jun 30, 2022 Accepted: Sep 12, 2022