Natural Resources Research, Vol. 29, No. 4, August 2020 (© 2019)
https://doi.org/10.1007/s11053-019-09587-1

Original Paper

Improving Adaptive Neuro-Fuzzy Inference System Based on a Modified Salp Swarm Algorithm Using Genetic Algorithm to Forecast Crude Oil Price

Mohamed Abd Elaziz,[1,6] Ahmed A. Ewees,[2,3] and Zakaria Alameer[4,5]

Received 30 May 2019; accepted 29 October 2019
Published online: 16 November 2019

This paper presents a novel forecasting model for crude oil price which has the largest effect on economies and countries. The proposed method depends on improving the performance of the adaptive neuro-fuzzy inference system (ANFIS) using a modified salp swarm algorithm (SSA). The SSA simulates the behaviors of salp swarm in nature during searching for food, and it has been developed as a global optimization method. However, SSA still has some limitations such as getting trapped at a local point. Therefore, this paper uses the genetic algorithm to improve the behavior of SSA. The proposed model (GA-SSA-ANFIS) aims to determine the suitable parameters for the ANFIS by using the GA-SSA algorithm since these parameters are considered as the main factor influencing the ANFIS's prediction process. The results of the GA-SSA-ANFIS are compared to other models, including the traditional ANFIS model, ANFIS based on GA (GA-ANFIS), ANFIS based on SSA (SSA-ANFIS) ANFIS based on particle swarm optimization (PSO-ANFIS), and ANFIS based on grey wolf optimization (GWO-ANFIS). The results show the superiority and high performances of the GA-SSA-ANFIS over the other models in predicting crude oil prices.

KEY WORDS: Salp swarm optimization, Adaptive neuro-fuzzy inference system, Crude oil price, Forecasting, Genetic algorithm.

INTRODUCTION

Crude oil is one of the significant commodities and has a vital influence on the global economy. The volatility of oil price is fundamental to risk management, asset allocation, and asset pricing. Therefore, numerous researches pay attention to forecast its price volatility (Zhang et al. 2019b). Crude oil plays a critical role in society including the technological, political, and economic dimensions (Li et al. 2018). Crude oil is considered as one of the important sources of energy (Lardic and Mignon 2008; He et al. 2010). Interestingly it has been reported that crude oil accounts for 33% of the total world energy consumption (Sun et al. 2018). The price prediction of global crude oil has become a more significant issue (Wang et al. 2018). Several factors raise the uncertainty of crude oil price fluctuations, such as the rising demands imposed by the financial crisis and emerging economies, geopolitical attitudes in

[1]Department of Mathematics, Faculty of Science, Zagazig University, Zagazig, Egypt.
[2]University of Bisha, Bisha, Kingdom of Saudi Arabia.
[3]Department of Computer, Damietta University, Damietta, Egypt.
[4]School of Resources and Environmental Engineering, Wuhan University of Technology, Hubei 430070, People's Republic of China.
[5]Mining and Petroleum Department, Faculty of Engineering, Al-Azhar University, Qena, Egypt.
[6]To whom correspondence should be addressed; e-mail: abd_el_aziz_m@yahoo.com

the Middle East, natural disasters, oil worker strikes, production control, and so on (Yaojie Zhang et al. 2019a).

In the energy market research field, forecasting crude oil prices is considered a challenging issue (Shabri and Samsudin 2017). In the last decade, traditional time series models such as generalized autoregressive conditional heteroscedasticity (GARCH), autoregressive moving average (ARMA), autoregressive models (AR), autoregressive integrated moving average (ARIMA), and vector auto-regression VAR models for forecasting of crude oil have received great attention (Agnolucci 2009; Hou and Suardi 2012; Allegret et al. 2015; Maghyereh 2006; Salisu and Oloko 2015). However, these models have a weakness to capture nonstationary and nonlinearities in the forecasting crude oil prices because these models are linear models.

Despite the availability of numerous forecasting models to forecast the fluctuations of crude oil, an investigation into the manner of improving their performance is worthwhile. For improving the predicting performance of crude oil price, machine learning and deep learning techniques have been extensively employed to predict the fluctuations of crude oil price, such as artificial neural networks (ANN) (Jammazi and Aloui 2012), gene expression programming (GEP) (Mostafa and El-Masry 2016), and support vector machine (SVM) (Yu et al. 2014). Meanwhile, these individual methods of nonlinear artificial intelligence provide better performances than the classic models; such models suffer from some problems of over-fitting and parameter optimization. Therefore, to increase the forecasting accuracy and overcome the shortcomings of solo models, hybrid models have been introduced.

Hybrid models combine a set of individual models to overcome the drawbacks of these models by combining the merits of every model to provide a better capacity (Sun et al. 2018). In the same manner, this study proposes a new hybrid model to increase the accuracy of forecasting crude oil price fluctuations by integrating the adaptive neuro-fuzzy inference system (ANFIS) (Jang 1993) with a modified salp swarm algorithm (SSA) (Mirjalili et al. 2017). In general, these two techniques have several advantages that can enhance the prediction of crude oil price. For instance, the ANFIS offers flexibility in determining nonlinearity in the data of the oil, as well as combining the ANN properties with that of the fuzzy logic system, whereas the traditional SSA method simulates the behaviors of salp swarm during the process of finding the food source.

By comparing the behaviors of SSA with other metaheuristic methods, it has been found that SSA has a good ability to find the global solution in suitable computational time. According to these behaviors, the SSA algorithm has been applied to several applications (El-Fergany 2018; Abbassi et al. 2019; Ibrahim et al. 2018). However, similar to other MH algorithms, SSA still has some limitations (e.g., it can be attractive to a stagnation point). This limitation can affect its convergence rate and the quality of the solution. Therefore, in this paper, the genetic algorithm (GA) is used to improve the performance of the SSA. Then the modified SSA, called GA-SSA, is applied to estimate the ANFIS's parameters to enhance the prediction of crude oil price since these parameters have the largest effect on ANFIS quality.

Moreover, choosing the forecasting variables to build an accurate model is considered as a major challenge. In the current study, ten forecastings (input) variables are used as predictors for the fluctuations of the crude oil price in the future, namely three exchange rates [Canadian dollar (CAD), Euro (EUR), and Chinese yuan (RMB)] as well as coal, natural gas, copper, gold, silver, iron, and past oil prices (West Texas Intermediate crude oil prices). The importance of using foreign exchange rates in the forecasting of the commodity price is considered in the literature. For example, Chen et al. (2010) recorded that exchange rates are the main predictor for forecasting commodity prices. They also reported that commodity prices could depend on the exchange rates of commodity-exporting nations. In addition, several studies examined the relationship between crude oil and all of the Canadian Dollar CAD, Euro EUR, and Chinese yuan RMB such as (Aloui et al. 2013; Bedoui et al. 2018; Ding and Vo 2012; Li et al. 2017b; Aloui and Jammazi 2015; Qin et al. 2015).

Crude oil is also considered as a crucial part of numerous production processes in the world (Zhao et al. 2018). Several studies have reported that oil prices are the leading reason for commodity price changes. The commodity price is applied as a predictor for crude oil price fluctuations because oil is considered as a commonly used source (He et al. 2010). Behmiri and Manera (2015) evaluated the impact of oil price changes on the price fluctuations of metals such as palladium, tin, aluminum, nickel, zinc, copper, gold, lead, platinum, and silver. Mo et al. (2018) investigated the dynamic linkages between the markets of the crude oil and both gold and USD; then they recorded that the relationship between oil and gold is always positive. In addition, Teetranont et al. (2018) applied the interval data in NYMEX and COMEX trading to study the relationship between crude oil prices and gold; their results showed a positive relationship between them. Moreover, the mean motivation is to add coal and natural gas as predictor variables for crude oil because all of them are the most commonly used sources of energy. Several studies documented the relationship between the prices of crude oil, coal, and natural gas such as (Kaufmann and Hines 2018; Tiwari et al. 2019; Brigida 2014; Caporin and Fontini 2017; Ramberg et al. 2017; Li et al. 2017a; Chen and Linn 2017; Guan and An 2017; American Association of Petroleum Geologists 2019).

The contribution of this study can be summarized as follows:

- Propose a modified version of the salp swarm algorithm using a genetic algorithm to improve the performance of the ANFIS to predict the crude oil price
- Use a new model namely GA-SSA-ANFIS for forecasting the price fluctuations of crude oil
- Confirm the power of the following variables in forecasting the price fluctuations of crude oil: coal, natural gas, copper, gold, silver, and iron prices; and CAD, EUR, and RMB exchange rates

The rest of this paper is arranged as follows: the relevant literature review for crude oil forecasting models are listed in the next section. The basics information about ANFIS, GA, SSA, and the proposed method are presented in ''Material and Methods'' section. In ''Experimental Results'' section, the results and discussion are listed while the conclusion and future work are given in the last section.

RELEVANT LITERATURE REVIEW

This section briefly reviews the forecasting methods for crude oil prices. These methods can be divided into three parts, namely conventional mathematical models, the methods of artificial intelligence and machine learning, and hybrid methods.

One of the common conventional mathematical models works in the forecasting domain is called autoregressive integrated moving average (ARIMA). It was used in different domains of forecasting such as economic, engineering, social, stock problems, and energy (Dooley and Lenihan 2005; Parisi et al. 2008; Kriechbaumer et al. 2014; Ediger and Akar 2007). Mohammadi and Lixian (2010) examined the application of ARIMA-GARCH models to forecast the weekly crude oil spot price volatility for the period from 1/2/1997 to 10/3/2009. Zhou and Dong (2012) examined the seasonality of China's crude oil import to help stakeholders with production planning and inventory control. Moreover, Zhao and Wang (2014) used an autoregressive integrated moving average model (ARIMA) to forecast the global crude oil price.

Zhang and Wang (2019) forecasted and estimated the volatility of crude oil price using two regime-switching GARCH (i.e., MRS-GARCH and MMGARCH) and three single-regime GARCH (i.e., GARCH, EGARCH, and GJR-GARCH) models. Hung et al. (2018) used GARCH (1,1), EGARCH (1,1), and GJR-GARCH (1,1) models to forecast the world's oil prices volatility using data of the WTI spot oil price for the period of 01/02/1986 to 25/4/2016. Mirmirani and Cheng Li (2004) applied VAR and ANN methods to forecast the volatility of U.S. oil price.

Artificial neural networks (ANNs), as a machine learning method, are used widely and evaluated for forecasting prices (Lineesh et al. 2010; Khashei and Bijari 2010; Parisi et al. 2008). Baruník and Malinska (2016) used a generalized regression framework based on neural networks to forecast the fluctuations of oil prices. (Mostafa and El-Masry 2016) forecasted fluctuations of oil prices using the artificial neural network (NN) and gene expression programming (GEP) models. Furthermore, Mingming and Jinliang (2012) utilized a real neural network (RNN) to forecast the fluctuations of crude oil prices at different scales. Lean et al. (2017) investigated an empirical study to verify the potentiality and feasibility of support vector machine SVM in the forecasting of crude oil price. Ramyar and Kianfar (2019) investigated the superiority of artificial neural networks over vector autoregressive models for forecasting crude oil prices. Chiroma et al. (2014) predicted monthly prices of West Texas Intermediate crude oil prices using an orthogonal wavelet support vector machine (OSVM) model. Xie et al. (2006) proposed a method to forecast the price of crude oil using support vector machine SVM.

Recently, hybrid models are widely applied to increase the forecasting accuracy and overcome the shortcomings of solo models (Khashei and Bijari 2011). Wu et al. (2019) proposed a novel hybrid model based on long short-term memory (LSTM) and ensemble empirical mode decomposition (EEMD) to forecast crude oil price. Zhu et al. (2019) proposed a hybrid model by integrating optimal combined forecasting model (CFM) and ensemble empirical mode decomposition (EEMD) to forecast the West Texas Intermediate (WTI) crude oil price. Safari and Davallou (2018) combined nonlinear autoregressive NAR neural network, ARIMA, and the exponential smoothing model (ESM) to forecast WTI crude oil spot prices. Moreover, Chai et al. (2018) combined trend decomposition of high-frequency sequences, the possible nonlinearity of model setting, change points, regime-switching, and time-varying determinants to propose a new model to forecast international crude oil price. Cheng et al. (2019) proposed a novel hybrid nonlinear autoregressive neural network and vector error correction (VEC-NAR) model to predict crude oil prices. Md-Khair and Samsudin (2017) introduced a Wavelet-ARIMA model to increase the forecasting precision of the crude oil price.

MATERIAL AND METHODS

Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS is a hybrid model that combines the neural networks and fuzzy logic; therefore, it inherits the advantages of both in its model. It was presented by Jang (1993) . It uses a Takagi-Sugeno inference technique that creates a nonlinear mapping from the input to the output spaces by the fuzzy IF–THEN rules. ANFIS model applies five layers to perform its tasks. These layers, as shown in Fig. 1, are summarized as in the following steps: the inputs x and y are passed to the nodes by Layer 1 to compute the output by the generalized Gaussian membership function μ; Eqs. (1) and (3) define these steps (Jang 1993).

$$O_{1i} = \mu_{A_i}(x),\ i = 1, 2, \tag{1}$$

$$O_{1i} = \mu_{B_{i-2}}(y),\ i = 3, 4 \tag{2}$$

$$\mu(x) = e^{-\left(\frac{x-\rho_i}{\alpha_i}\right)^2}, \tag{3}$$

where A_i and B_i define the membership values of the μ. ρ_i and α_i define the premise parameters set. The next layer uses Eq. (4) to compute the result of each node (the firing strength of a rule). Then the output is normalized in Layer 3 using Eq. (5).

$$O_{2i} = \mu_{A_i}(x) \times \mu_{B_{i-2}}(y) \tag{4}$$

$$O_{3i} = \overline{w}_i = \frac{\omega_i}{\sum_{(i=1)}^{2} \omega_i}, \tag{5}$$

The adaptive nodes at Layer 4 calculates the output using Eq. (6).

$$O_{4,i} = \overline{w}_i f_i = \overline{w}_i(p_i x + q_i y + r_i) \tag{6}$$

where p, q and r define the consequent parameters of the ith node. The last layer has a single node. It calculates the overall output as in the following equation:

$$O_5 = \sum_i \overline{w}_i f_i \tag{7}$$

In many cases, the search space of ANFIS becomes wider and the convergence of training becomes slower as well as getting easily trapped in local optima. Therefore, using a hybrid method is an important task to overcome such a problem.

Genetic Algorithm (GA)

A genetic algorithm (GA) is a type of optimization techniques that mimics the natural biological evolution of species (Ketabchi and Ataie-Ashtiani 2015). It was presented by Holland (1992) to be applied in finding the global optimal solution for a given problem. GA repeats sequential steps to find the optimum result of the problem including initialization, selection, and reproduction to create individuals (Ketabchi and Ataie-Ashtiani 2015). The generated population in each iteration can be called as a generation. Each individual or generation is evaluated by an objective function to take a fitness value. Accordingly, the best one is used by cross-breeding with other generations to improve the population. Thus, a new population is formed by selecting the fit individuals to create a new set of individuals (Alameer et al. 2019b). Algorithm 1 summarizes the entire steps of GA (Ketabchi and Ataie-Ashtiani 2015).

Algorithm 1 Genetic Algorithm (GA)

```
Create the initial population (pop) randomly.
Evaluate each individual in the population
while (termination conditions not met) do
    for each individual repeat do
        Select parents for reproducing
        Crossover
        Mutation
        Generate new population
        Evaluate the new population
        Select the best solution from the population.
    end for
end while
Return the best solution.
```

Salp Swarm Algorithm (SSA)

Salp swarm algorithm (SSA) is one of the recent optimization methods developed by Mirjalili et al. (2017). They tried to mathematically simulate the salp chains behavior of the real salps. Salps are considered as a kind of the Salpidae's family. They look like jellyfishes in their moving behavior and their bodies contain a high water percentage (Henschke et al. 2016).

Salps use their swarm behavior (i.e., salp chain) in foraging and moving with fast harmonious changes (Sutherland and Weihs 2017). Food sources are the target of the swarm.

The mathematical model of SSA begins by generating a population then dividing it into leader and followers based on the position. The leader is the front salp of the chain, and the followers are the rest of the salps.

The position is formed in n-dimensions that denote the search space of a given problem, whereas the problem variables are represented by n. The position is frequently updated. The equation below is applied to update the position of the salp leader:

$$x_j^1 = \begin{cases} F_j + c_1((ub_j - lb_j) \times c_2 + lb_j) & c_3 \leq 0 \\ F_j - c_1((ub_j - lb_j) \times c_2 + lb_j) & c_3 > 0 \end{cases} \tag{8}$$

where x_j^1 denotes the leader's position in jth dimension. F_j denotes the food source. ub_j and lb_j are the upper and lower bounds, respectively. c_2 and c_3 are random variables in [0,1] helps in maintaining the search space. c_1 is a coefficient used to balance the exploitation and exploration phases. It is computed as follows:

$$c_1 = 2e^{-\left(\frac{4l}{L}\right)^2}, \tag{9}$$

where l and L indicate the current loop and the max number of loops, respectively.

Subsequently, the position of the followers is updated using the following equation:

$$x_j^i = \frac{1}{2}\left(x_j^i + x_j^{i-1}\right) \tag{10}$$

where x_j^i is the ith follower position and $i > 1$. The entire sequence of the SSA is listed in Algorithm 2.

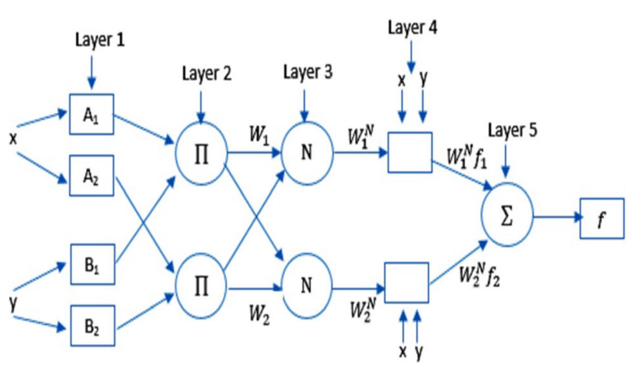

Figure 1. The traditional ANFIS (El Aziz et al. 2017).

Algorithm 2 Salp Swarm Algorithm (SSA)

```
Create a population X.
repeat
    Calculate the objective function for solutions x_i.
    Update the best solution (salp) (F = X^b).
    Update c_1 by Eq. ssa2.
    for i = 1toN do
        if i == 1 then
            Update the salps position of by Eq. (8)
        else
            Update the salps position of by Eq. (10)
        end if
    end for
until (l ≤ L)
Return the best solution F.
```

Proposed Method

Developing an accurate forecasting model for crude oil price fluctuations helps to foresee the market trends in the future, which provides stakeholders with valuable information for making the right decisions to prevent or mitigate risks. Notwithstanding the availability of numerous forecasting models, research on the process of improving the performance of these models continues. In this work, we proposed a new model for forecasting crude oil price fluctuations using a modified salp swarm algorithm (SSA) to improve the performance of the adaptive neuro-fuzzy inference system (ANFIS). However, SSA still has some limitations such as: it is easy to suck at a local point; therefore, we use the genetic algorithm to improve the behavior of SSA. In addition, we argue that the proposed model can be successfully applied for forecasting crude oil price fluctuations. Also, we can confirm that it can be used for future predictions, not only for crude oil prices, but also for all metals, ores, and commodities.

The description of the proposed method is presented in this section (see Fig. 2). The proposed method improves the ANFIS model using GA and SSA algorithms which called GA-SSA-ANFIS. It applies GA and SSA to adjust the parameters of the ANFIS by feeding the best weights between Layer 4 and Layer 5.

The GA-SSA-ANFIS starts by preparing the inputs and dividing the problem into training and testing sets. Then the fuzzy c-mean (FCM) algorithm is used to determine a suitable number of the membership functions by clustering the dataset into different groups. Thereafter, the ANFIS uses these results to start the rest of the steps. The ANFIS's parameters, namely the weights, are adapted using the GA-SSA algorithm, where the GA-SSA searches for the solution in the problem space by exploring various domains.

In the first step in the proposed method, the GA is used to generate the initial population of the SSA. Then the SSA uses this population to start searching for the best weights of the ANFIS. The fitness value of all population is calculated by the following fitness function.

$$\mathit{Objective\ function} = ||obs - pred||^2 \tag{11}$$

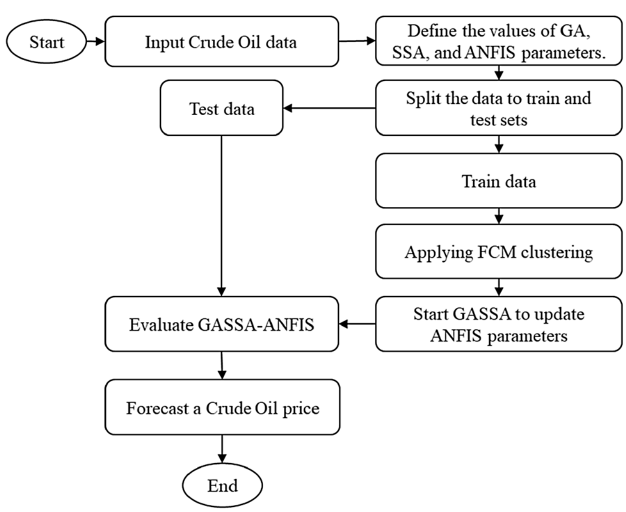

Figure 2. Proposed GA-SSA-ANFIS.

Table 1. The source of the dataset

Variables	Period	Unit	Data source
Oil price	Monthly	\$/barrel	U.S. Energy Information Administration https://www.indexmundi.com/commodities/?commodity=crude-oil&months=360
Iron ore and Copper	Monthly	\$/ton	Thomson Reuters Datastream; World Bank http://www.indexmundi.com/commodities/
Gold price	Monthly	\$/ozt	World Bank http://www.indexmundi.com/commodities/
Silver price	Monthly	\$/ozt	Thomson Reuters Datastream; World Bank http://www.indexmundi.com/commodities/
Coal	Monthly	\$/ton	Coal, Australian thermal coal https://www.indexmundi.com/commodities/?commodity=coal-australian&months=360
Natural Gas	Monthly	\$/Million Metric British Thermal Unit	Thomson Reuters Datastream; World Bank https://www.indexmundi.com/commodities/?commodity=natural-gas&months=360

Therefore, the selected weights are updated based on minimizing the error between the output and the real values in the training phase. These weights are passed to the ANFIS to prepare the output results of a given problem. The GA-SSA works till meeting the stop condition in this paper which is the max number of iterations. After that, the test phase starts, and the best weights are passed to the ANFIS to produce the output.

EXPERIMENTAL RESULTS

In this section, the results of the GA-SSA-ANFIS is compared with other methods to evaluate its performance as a prediction method for iron ore prices.

Data Description

The dataset used in this experiment is collected from the observations of the West Texas Intermediate (WTI) crude oil price presented in (\$/barrel) from January 1989 to December 2018 which represents 360 months. These data were downloaded from different sources, as given in Table 1.

These data are randomly divided into two sets. The first set, called training, is used during the learning of the parameters of the model to predict the output. The period of this set is taken from January 1989 to December 2009 which represents 70% from the total observation (i.e., 252 observations). Meanwhile, the second dataset, called the

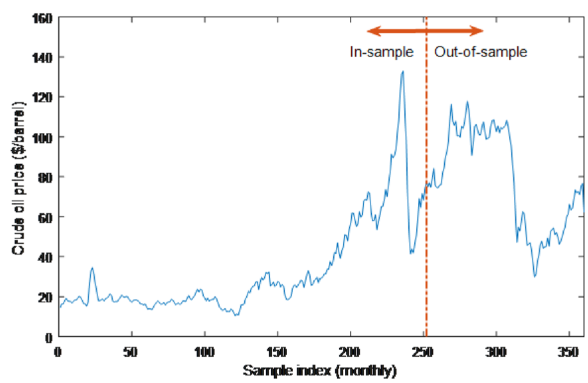

Figure 3. The historical crude oil prices from January 1989 to December 2018.

testing set, is applied to the learned model to evaluate its performance. The period of the testing set is taken from January 2010 to December 2018 which represents 30% from observation (i.e., 108 observations). These two sets are given in Fig. 3

Moreover, to enhance the performance of the GA-SSA-ANFIS, a set of variables is used which have a significant correlation with crude oil prices. The monthly observation data of CAD, EUR and RMB exchange rates, coal, natural gas, copper, gold, silver, and iron prices are collected over the same period of crude oil price.

Since these variables contain multiple financial indexes with different scales, the value of those variables will be normalized to become more homogeneous. This is performed using the following equation, where the original value of all variables X_i is converted to $X_{i(norm)}$.

$$X_{i(norm)} = \frac{X_i - X_{min}}{X_{max} - X_{min}}, i = 1, 2, \ldots, 360, \quad (12)$$

where X_{min} and X_{max} are the minimum and maximum value of the original series, respectively.

Correlation Analysis

In this section, the correlation analysis of the used variables with the crude oil price will be introduced. This aims to determine the relationship between them (Haque et al. 2015) which may be a positive or negative correlation between the variables (Ho 2006).

In any investigation of the relationship between two logically linked variables, correlation analysis is proven to be a useful tool for determining the strength of that relationship. More specifically, Pearson's correlation coefficient is a measure of the linear dependence between two random variables (real-valued vectors). Historically, it is the first formal measure of correlation, and it remains one of the most extensively used measures of relationships. Thus, we conducted this preliminary investigation prior to constructing the proposed model to select the predictor variables by examining the relationship between crude oil prices and predictor variables and confirm any significant correlation between crude oil prices and predictor variables.

Regarding the concept of supply and demand, that can be concluded, as supply increases or demand decreases the price should go down; as supply decreases or demand increases the price should go up. So, major fluctuations in price can have a significant economic impact. Therefore, if the price of any of the power resources is changed, there will be a change in crude oil prices. In addition, if any change happened in the rates of the main currency, there will be also a change in crude oil price. The relation between crude oil and copper, iron, and precious metal (gold and silver) is also present as a nonlinear relationship. However, inflation is directly affecting their prices; therefore, when crude oil prices rise, inflation also rises. This relation was concluded by many previous studies that work in analyzing this relation such as (Šimáková 2011; Jain and Biswal 2016; Dutta et al. 2019).

Moreover, through a literature review, we found that several studies proved relationships between crude oil and the input variables. For instance, (Alameer et al. 2019a, b) employed Pearson's correlation coefficient to explore the correlation between crude oil price and gold, copper, silver, iron ore prices, and USD/RMB. Moreover, the mean motivation to add coal and natural gas as predictor variables for crude oil because all of them are the most commonly used sources of energy. In addition, several studies examined the relationship between crude oil and all of the Canadian Dollar CAD, Euro EUR, and Chinese yuan RMB.

In general, the Pearson's correlation coefficient computes the degree of linear correlation between two variables. This degree is a value in the interval [+1, −1], The value of +1 indicates that there is a completely positive correlation, while 0 refers to lacking a correlation, and −1 represents a completely negative correlation. Figure 4 depicts the pairwise correlation between the variables used in this paper.

It can be observed from this figure that most of the variables high correlation value higher than 0.5 except the nature gas, and USD/RMB which have correlation values 0.47 and 0.19, respectively. In addition, it is notable that all the correlations between crude oil prices and other variables are statically significant since the p value is nearly 0.001. Moreover, Fig. 5 illustrates the relation between each input variable and the crude oil price.

Results and Discussion

The performance of the proposed GA-SSA-ANFIS method for predicting the price of crude oil is evaluated. This process is performed by dividing the dataset into two sets, the first set is the training set which represents 70% of the total sample, whereas the second set is the testing set which consists of 30% of the samples of the dataset. In addition, the proposed GA-SSA-ANFIS is compared with the other nine methods, namely traditional ANFIS, ANFIS based on particle swarm optimization (PSO-ANFIS), ANFIS based on genetic algorithm (GA-ANFIS), ANFIS based on particle swarm optimization (SSA-ANFIS), ANFIS based on grey wolf optimization (GWO-ANFIS), traditional ELM (extreme learning machine), ELM based on PSO (PSO-ELM), ELM based on GA (GA-ELM), and GWO based on ELM (GWO-ELM). These methods are selected since they establish their performance as forecasting methods in different applications.

The comparison results between the proposed GA-SSA-ANFIS approach and the others are given in Table 2. It can be noticed from this table that, in terms of MSE and RMSE the best forecasting method is the proposed GA-SSA-ANFIS which has

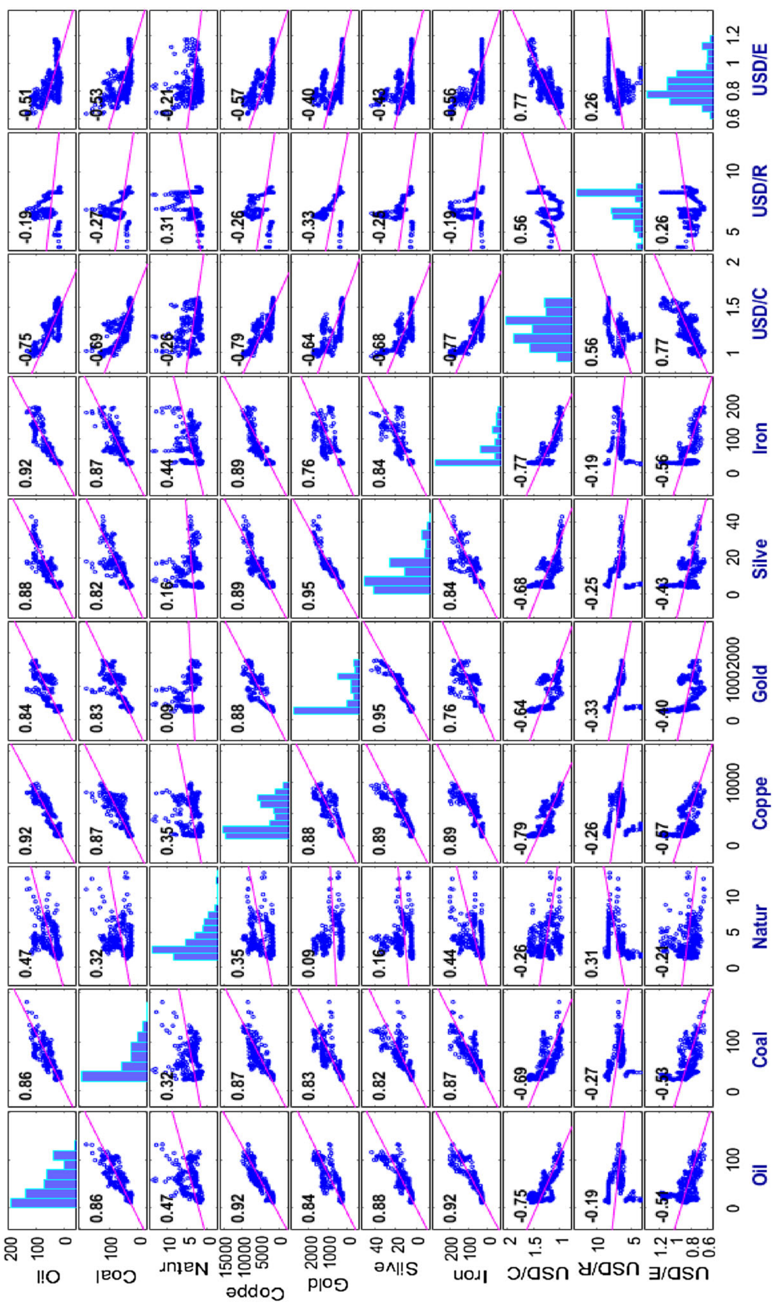

Figure 4. The correlation between the selected variables.

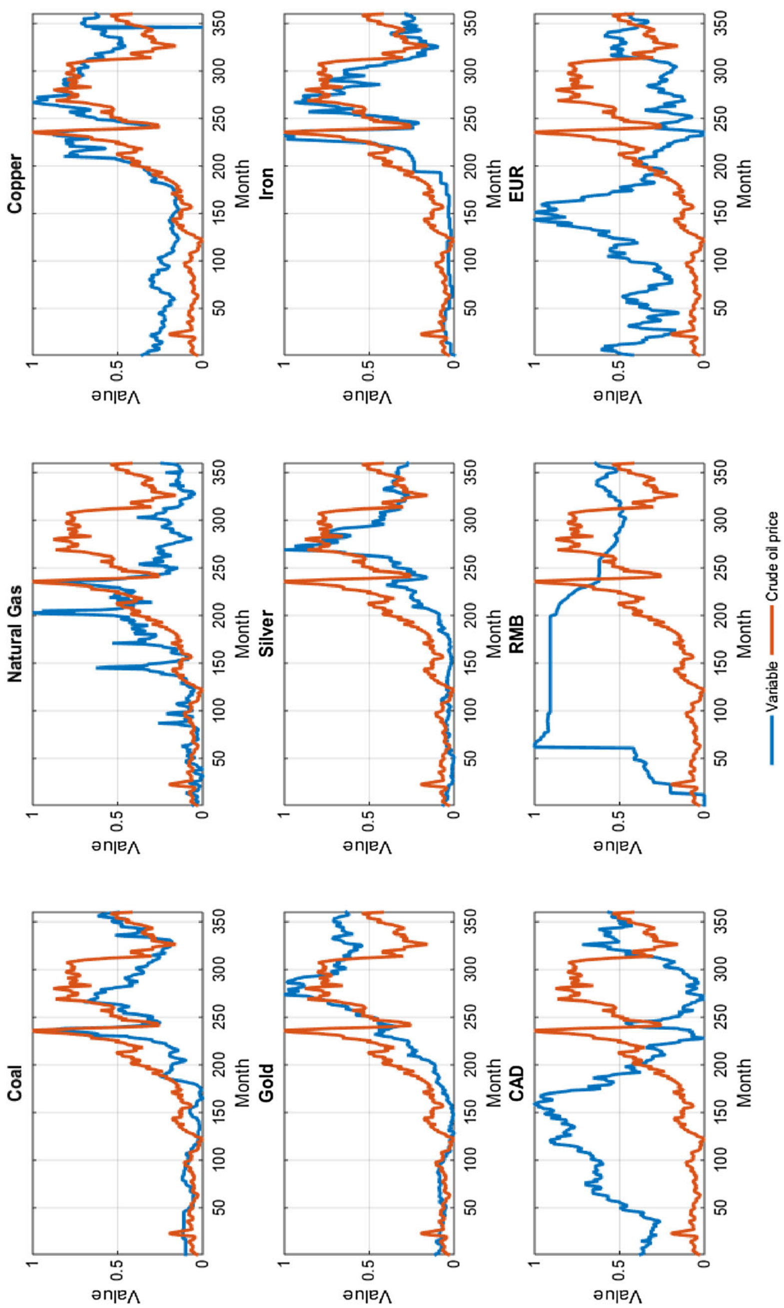

Figure 5. The correlation between the selected variables.

Table 2. Comparison between different methods (boldface refers to th best value)

Model	MSE	RMSE	MAE	STD	R^2
ANFIS	0.0334	0.1737	0.1514	0.0562	0.6204
GWO-ANFIS	0.0284	0.1651	0.1415	0.0341	0.8006
PSO-ANFIS	0.0153	0.1229	0.1061	0.0139	0.8346
GA-ANFIS	0.0150	0.1218	0.1052	0.0125	0.8541
SSA-ANFIS	0.0216	0.1447	0.1244	0.0266	0.8033
GA-SSA-ANFIS	0.0131	0.1145	0.0948	0.0057	0.8818
ELM	0.0248	0.1537	0.1222	0.0349	0.8177
PSO-ELM	0.0158	0.1255	0.0974	0.0100	0.8250
GA-ELM	0.0151	0.1227	0.0958	0.0069	0.8257
GWO-ELM	0.0161	0.1265	0.0980	0.0100	0.8248

the smallest value. Also, GA-ANFIS allocates the second rank, followed by the GA-ELM, PSO-ANFIS, PSO-ELM, and GWO-ELM, respectively, while SSA-ANFIS and ELM have nearly the same performance. In addition, the GWO-ANFIS provides results better than the traditional ANFIS model. In addition, the GA-SSA-ANFIS outperformed the traditional ELM, PSO-ELM, GA-ELM, and GWO-ELM in the above measures. The traditional ANFIS also showed the worst performance.

The GA-SSA-ANFIS is considered as the most stable algorithm among the other methods, as a result of the standard deviation (STD) measure, followed by GA-ELM, PSO-ELM, GWO-ELM, GA-ANFIS, and PSO-ANFIS, respectively. From this measure that can be noticed, the GA algorithm helps in making the proposed method more stable than other methods.

In terms of mean absolute error (MAE), the GA-SSA-ANFIS is ranked first followed by GA-ELM, PSO-ELM, then GWO-ELM. However, the GA-ANFIS is ranked the fifth followed by PSO-ANFIS. The worst algorithm in this measure is the traditional ANFIS. Moreover, by analyzing the results of R^2, it can be observed that the proposed GA-SSA-ANFIS has the highest value which indicates the high performance of the GA-SSA-ANFIS among other methods as a fitted model. Moreover, the ANFIS has the worst results among all tested methods which indicate that the parameters have the largest effect on it and this requires using a suitable optimization method to determine these parameters.

Moreover, Fig. 6 depicts the original price of the oil and its forecasting values using the proposed GA-SSA-ANFIS and the other five methods (i.e., ANFIS, GA-NN, PSO-ANFIS, SSA-ANFIS, and GWO-ANFIS). It can be seen that during the learning process most of these methods, including the GA-SSA-ANFIS, have nearly the same behaviors. However, it can be observed that the proposed GA-SSA-ANFIS approach has a high ability to learn than other methods and this can be noticed at the interval 210–240 when there is high raise in the price of the oil. Also, during the valuation process, it can be seen that the forecasting value for the price of the oil using the proposed GA-SSA-ANFIS is closer to the original value than the other methods (Fig. 7).

Statistical Analysis

In this section, a nonparametric test called Wilcoxon rank-sum (WRS) test is applied to determine whether there is a significant difference between the proposed GA-SSA-ANFIS and the other approaches or not. In general, there are two hypotheses; the first hypothesis assumes that there is no significant difference and this is called the null hypothesis. Meanwhile, the second one is called the alternative hypothesis which assumes that there is a significant difference between the proposed GA-SSA-ANFIS and the others at significant level 0.05.

The results of the WRS test are given in Table 3, and it can be observed that the p value is less than the significant level which indicates that there is a highly significant difference between the proposed GA-SSA-ANFIS and the others models in terms of RMSE and R^2.

From all previous results and discussions, it can be concluded that the proposed GA-SSA-ANFIS has high performance to forecast the price of the oil. This high quality results from different factors. The first factor is the combination between the GA and SSA as a metaheuristic method, which avoids the limitations of SSA and improves the performance of finding the global solutions. The second factor is using the GA-SSA as an optimization method for determining the parameters of the ANFIS model, which avoids the problems facing the traditional method such as back-propagation.

CONCLUSIONS

This study presents an alternative forecasting model for the crude oil price fluctuations since it has the largest effect on the economy of the world not

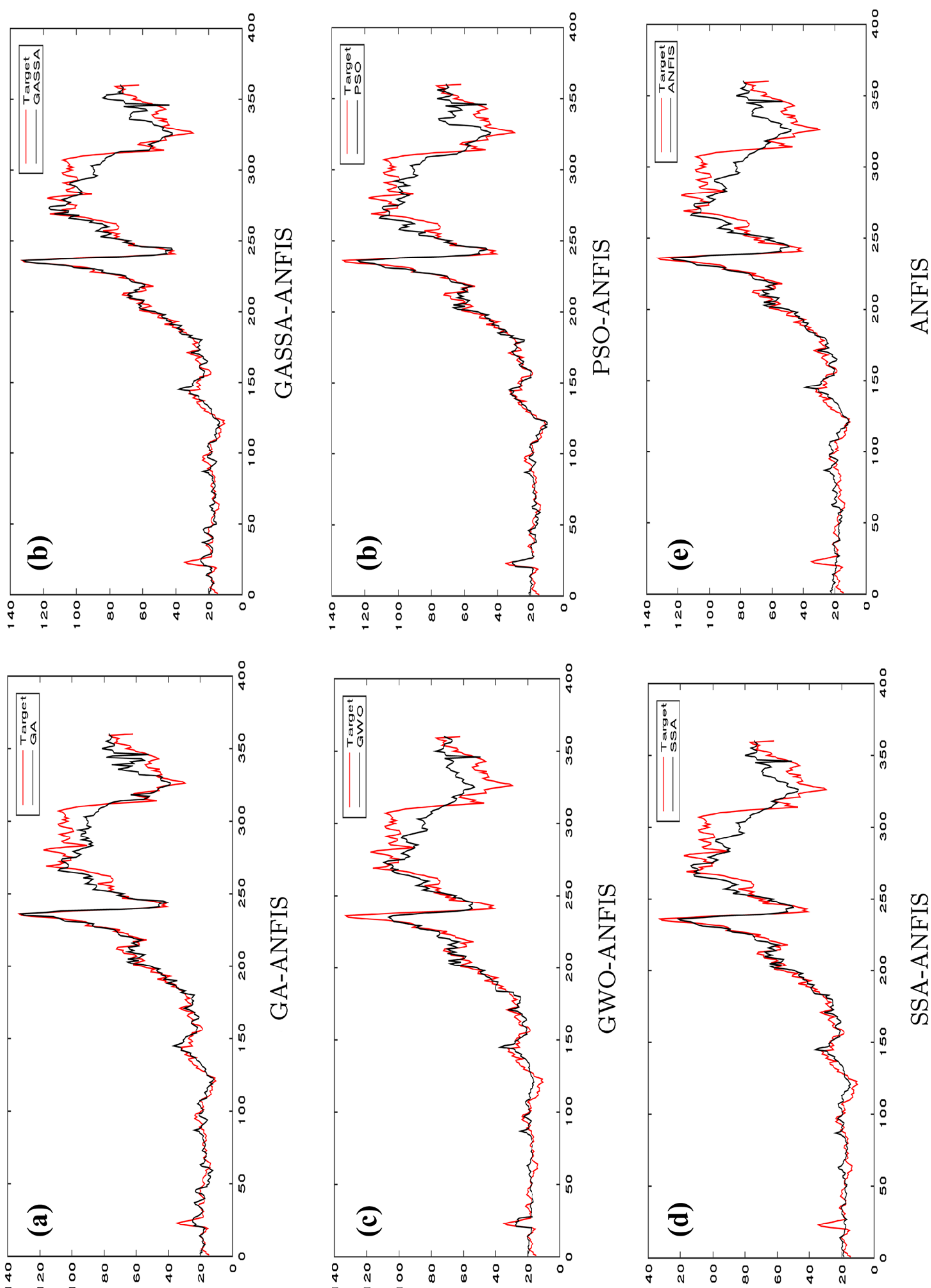

Fig 6. The original crude oil price and its prediction value using GA-ANFIS, GA-SSA-ANFIS, GWO-ANFIS, PSO-ANFIS, SSA-ANFIS, and ANFIS algorithms.

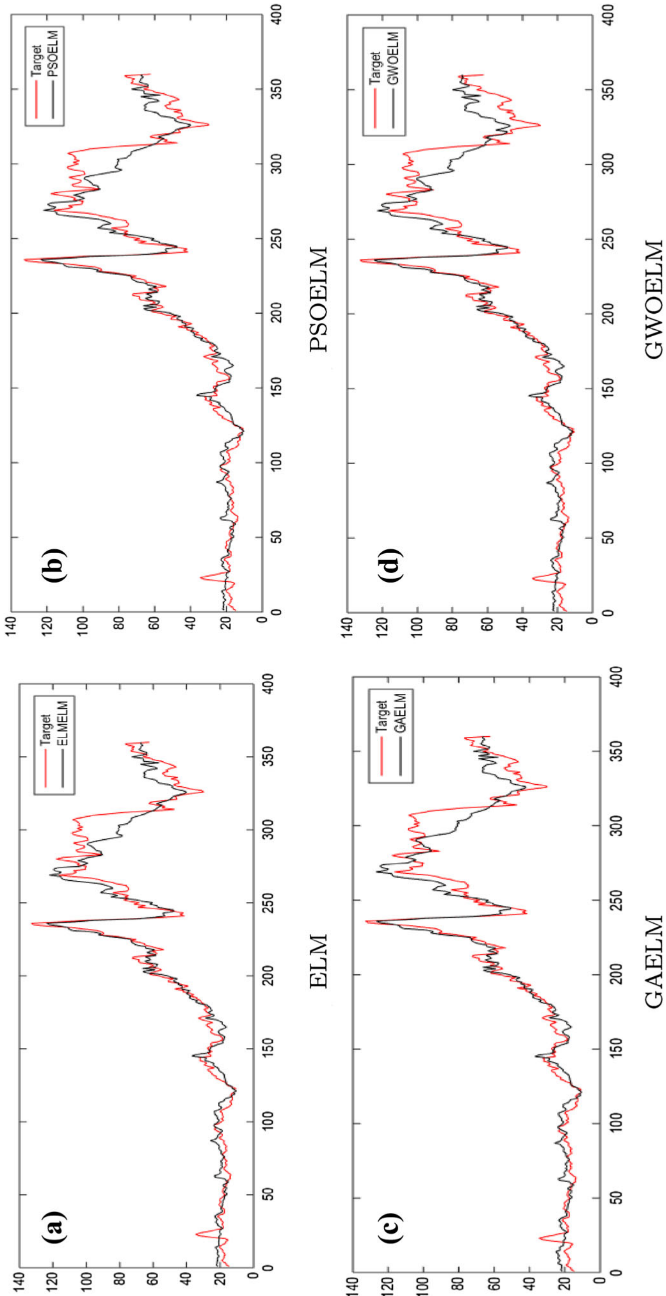

Figure 7. The original crude oil price and its prediction value using ELM, PSO-ELM, GA-ELM, and GWO-ELM algorithms.

Table 3. Comparison between GA-SSA-ANFIS and different methods using WRS measure

	RMSE		R2	
Algorithm	*p* value	*h*	*p* value	*h*
GWO	7.90E-08	1	1.23E-07	1
PSO	2.96E-07	1	3.42E-07	1
GA	1.04E-06	1	1.41E-05	1
ANFIS	3.99E-08	1	4.00E-08	1
SSA	2.56E-07	1	9.17E-08	1
PSO-ELM	5.47E-06	1	6.80E-08	1
GA-ELM	2.36E-06	1	4.00E-08	1
GWO-ELM	8.90E-07	1	6.80E-08	1
ELM	5.69E-07	1	5.50E-08	1

only a specific country. Therefore, improving an efficient and accurate forecasting method is a very important task that allows experts to make the right decisions. This is achieved through providing valuable information, about the crude oil price, to the stakeholders. In general, there are many forecasting models, whereas the ANFIS model is considered one of the most popular method. However, the performance of the ANFIS depends on its parameters so when these parameters are not determined by a suitable method, the performance of the ANFIS is degraded. Therefore, the combination of the genetic algorithm and the salp swarm algorithm was applied to determine the parameters of ANFIS. This combination, called GA-SSA, aims to enhance the ability of SSA to find a global solution (i.e., the parameters of the ANFIS). According to the GA-SSA and ANFIS, a new forecasting crude oil price fluctuations model was developed. Notably, several researchers proved the high level of accuracy of traditional ANFIS, PSO-ANFIS, GA-ANFIS, SSA-ANFIS, and GWO-ANFIS, and they used these algorithms for improving many problems. Moreover, these algorithms are highly popular and strong techniques for forecasting various commodity prices. The comparison results showed that the GA-SSA-ANFIS has a high ability for forecasting crude oil price fluctuations against the other models. Moreover, the results proved that GA-SSA-ANFIS is better than the other five models in terms of RMSE, MSE, STD, and R^2 measures. In addition, this superiority of the proposed GA-SSA-ANFIS approach was further confirmed by using a nonparametric test called the Wilcoxon rank-sum test, which established that there is a significant difference between GA-SSA-ANFIS and the compared models. Finally, the GA-SSA-ANFIS model provides promising results for forecasting commodity prices with high quality. According to these results, the proposed GA-SSA-ANFIS can be applied in the future works to estimate the future mining, refining, and fixed costs of mining projects.

REFERENCES

Abbassi, R., Abbassi, A., Heidari, A. A., & Mirjalili, S. (2019). An efficient salp swarm-inspired algorithm for parameters identification of photovoltaic cell models. *Energy Conversion and Management*, *179*, 362–372.

Agnolucci, P. (2009). Volatility in crude oil futures: A comparison of the predictive ability of garch and implied volatility models. *Energy Economics*, *31*(2), 316–321.

Alameer, Z., Elaziz, M. A., Ewees, A. A., Ye, H., & Jianhua, Z. (2019a). Forecasting copper prices using hybrid adaptive neuro-fuzzy inference system and genetic algorithms. *Natural Resources Research*, *28*(4), 1385–1401.

Alameer, Z., Elaziz, M. A., Ewees, A. A., Ye, H., & Jianhua, Z. (2019b). Forecasting gold price fluctuations using improved multilayer perceptron neural network and whale optimization algorithm. *Resources Policy*, *61*, 250–260.

Allegret, J.-P., Mignon, V., & Sallenave, A. (2015). Oil price shocks and global imbalances: Lessons from a model with trade and financial interdependencies. *Economic Modelling*, *49*, 232–247.

Aloui, C., & Jammazi, R. (2015). Dependence and risk assessment for oil prices and exchange rate portfolios: A wavelet based approach. *Physica A: Statistical Mechanics and its Applications*, *436*, 62–86.

Aloui, R., Safouane Ben Aïssa, M., & Nguyen, D. K. (2013). Conditional dependence structure between oil prices and exchange rates: A copula-garch approach. *Journal of International Money and Finance*, *32*, 719–738.

Baruník, J., & Malinska, B. (2016). Forecasting the term structure of crude oil futures prices with neural networks. *Applied energy*, *164*, 366–379.

Bedoui, R., Braeik, S., Goutte, S., & Guesmi, K. (2018). On the study of conditional dependence structure between oil, gold and usd exchange rates. *International Review of Financial Analysis*, *59*, 134–146.

Behmiri, N. B., & Manera, M. (2015). The role of outliers and oil price shocks on volatility of metal prices. *Resources Policy*, *46*, 139–150.

Brigida, M. (2014). The switching relationship between natural gas and crude oil prices. *Energy Economics*, *43*, 48–55.

Caporin, M., & Fontini, F. (2017). The long-run oil-natural gas price relationship and the shale gas revolution. *Energy Economics*, *64*, 511–519.

Chai, J., Xing, L.-M., Zhou, X.-Y., Zhang, Z. G., & Li, J.-X. (2018). Forecasting the wti crude oil price by a hybrid-refined method. *Energy Economics*, *71*, 114–127.

Cheng, F., Li, T., Wei, Y., & Fan, T. (2019). The vec-nar model for short-term forecasting of oil prices. *Energy Economics*, *78*, 656–667.

Chen, F., & Linn, S. C. (2017). Investment and operating choice: Oil and natural gas futures prices and drilling activity. *Energy Economics*, *66*, 54–68.

Chen, Y.-C., Rogoff, K. S., & Rossi, B. (2010). Can exchange rates forecast commodity prices? *The Quarterly Journal of Economics*, *125*(3), 1145–1194.

Chiroma, H., Abdul-Kareem, S., Abubakar, A., Zeki, A. M., & Usman, M. J. (2014). Orthogonal wavelet support vector machine for predicting crude oil prices. In *Proceedings of the first international conference on advanced data and information engineering (DaEng-2013)* (pp. 193–201). Springer: Berlin.

Ding, L., & Vo, M. (2012). Exchange rates and oil prices: A multivariate stochastic volatility analysis. *The Quarterly Review of Economics and Finance*, *52*(1), 15–37.

Dooley, G., & Lenihan, H. (2005). An assessment of time series methods in metal price forecasting. *Resources Policy*, *30*(3), 208–217.

Dutta, A., Bouri, E., & Roubaud, D. (2019). Nonlinear relationships amongst the implied volatilities of crude oil and precious metals. *Resources Policy*, *61*, 473–478.

Ediger, V. Ş., & Akar, S. (2007). Arima forecasting of primary energy demand by fuel in turkey. *Energy policy*, *35*(3), 1701–1708.

El Aziz, M. A., Hemdan, A. M., Ewees, A. A., Elhoseny, M., Shehab, A., Hassanien, A. E., & Xiong, S. (2017). Prediction of biochar yield using adaptive neuro-fuzzy inference system with particle swarm optimization. In *PowerAfrica, 2017 IEEE PES* (pp. 115–120). IEEE.

El-Fergany, A. A. (2018). Extracting optimal parameters of pem fuel cells using salp swarm optimizer. *Renewable Energy*, *119*, 641–648.

Energy Minerals Division American Association of Petroleum Geologists. (2019). Unconventional energy resources: 2017 review. *Natural Resources Research*, *28*(4), 1661–1751.

Guan, Q., & An, H. (2017). The exploration on the trade preferences of cooperation partners in four energy commodities' international trade: Crude oil, coal, natural gas and photovoltaic. *Applied Energy*, *203*, 154–163.

Haque, M. A., Topal, E., & Lilford, E. (2015). Relationship between the gold price and the australian dollar-us dollar exchange rate. *Mineral Economics*, *28*(1–2), 65–78.

Henschke, N., Everett, J. D., Richardson, A. J., & Suthers, I. M. (2016). Rethinking the role of salps in the ocean. *Trends in Ecology & Evolution*, *31*(9), 720–733.

He, Y., Wang, S., & Lai, K. K. (2010). Global economic activity and crude oil prices: A cointegration analysis. *Energy Economics*, *32*(4), 868–876.

Ho, R. (2006). *Handbook of univariate and multivariate data analysis and interpretation with SPSS*. London: Chapman and Hall/CRC.

Holland, J. H., et al. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. Cambridge: MIT press.

Hou, A., & Suardi, S. (2012). A nonparametric garch model of crude oil price return volatility. *Energy Economics*, *34*(2), 618–626.

Hung, N. T., & Thach, N. N. et al. (2018). Garch models in forecasting the volatility of the world's oil prices. In *International econometric conference of Vietnam* (pp. 673–683). Springer.

Ibrahim, R. A., Ewees, A. A., Oliva, D., Elaziz, M. A., & Lu, S. (2018). Improved salp swarm algorithm based on particle swarm optimization for feature selection. *Journal of Ambient Intelligence and Humanized Computing*, *10*(8), 3155–3169.

Jain, A., & Biswal, P. C. (2016). Dynamic linkages among oil price, gold price, exchange rate, and stock market in india. *Resources Policy*, *49*, 179–185.

Jammazi, R., & Aloui, C. (2012). Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling. *Energy Economics*, *34*(3), 828–841.

Jang, J.-S. R. (1993). Anfis: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, *23*(3), 665–685.

Kaufmann, R. K., & Hines, E. (2018). The effects of combined-cycle generation and hydraulic fracturing on the price for coal, oil, and natural gas: Implications for carbon taxes. *Energy Policy*, *118*, 603–611.

Ketabchi, H., & Ataie-Ashtiani, B. (2015). Evolutionary algorithms for the optimal management of coastal groundwater: A comparative study toward future challenges. *Journal of Hydrology*, *520*, 193–213.

Khashei, M., & Bijari, M. (2010). An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Systems with Applications*, *37*(1), 479–489.

Khashei, M., & Bijari, M. (2011). A novel hybridization of artificial neural networks and arima models for time series forecasting. *Applied Soft Computing*, *11*(2), 2664–2675.

Kriechbaumer, T., Angus, A., Parsons, D., & Casado, M. R. (2014). An improved wavelet–arima approach for forecasting metal prices. *Resources Policy*, *39*, 32–41.

Lardic, S., & Mignon, V. (2008). Oil prices and economic activity: An asymmetric cointegration approach. *Energy Economics*, *30*(3), 847–855.

Lean, Y., Zhang, X., Wang, S., et al. (2017). Assessing potentiality of support vector machine method in crude oil price forecasting. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(12), 7893–7904.

Li, H., Chen, L., Wang, D., & Zhang, H. (2017a). Analysis of the price correlation between the international natural gas and coal. *Energy Procedia*, *142*, 3141–3146.

Lineesh, M. C., Minu, K. K., & John, C. J. (2010). Analysis of nonstationary nonlinear economic time series of gold price: A comparative study. *International Mathematical Forum*, *5*(34), 1673–1683.

Li, X., Shang, W., & Wang, S. (2018). Text-based crude oil price forecasting: A deep learning approach. *International Journal of Forecasting*, *35*(4), 1548–1560.

Li, X.-P., Zhou, C.-Y., & Chong-Feng, W. (2017b). Jump spillover between oil prices and exchange rates. *Physica A: Statistical Mechanics and its Applications*, *486*, 656–667.

Maghyereh, A. (2006). Oil price shocks and emerging stock markets: A generalized var approach. In *Global stock markets and portfolio management* (pp. 55–68). Springer.

Md-Khair, N. Q. N., & Samsudin, R. (2017). Forecasting crude oil prices using wavelet arima model approach. In *International conference of reliable information and communication technology* (pp. 535–544). Springer.

Mingming, T., & Jinliang, Z. (2012). A multiple adaptive wavelet recurrent neural network model to analyze crude oil prices. *Journal of Economics and Business*, *64*(4), 275–286.

Mirjalili, S., Gandomi, A. H., Mirjalili, S. Z., Saremi, S., Faris, H., & Mirjalili, S. M. (2017). Salp swarm algorithm: A bio-inspired optimizer for engineering design problems. *Advances in Engineering Software*, *114*, 163–191.

Mirmirani, S., & Cheng Li, H. (2004). A comparison of var and neural networks with genetic algorithm in forecasting price of oil. In *Applications of artificial intelligence in finance and economics* (pp. 203–223). Emerald Group Publishing Limited.

Mohammadi, H., & Lixian, S. (2010). International evidence on crude oil price dynamics: Applications of arima-garch models. *Energy Economics*, *32*(5), 1001–1008.

Mo, B., Nie, H., & Jiang, Y. (2018). Dynamic linkages among the gold market, us dollar and crude oil market. *Physica A: Statistical Mechanics and its Applications*, *491*, 984–994.

Mostafa, M. M., & El-Masry, A. A. (2016). Oil price forecasting using gene expression programming and artificial neural networks. *Economic Modelling*, *54*, 40–53.

Parisi, A., Parisi, F., & Díaz, D. (2008). Forecasting gold price changes: Rolling and recursive neural network models. *Journal of Multinational financial management*, *18*(5), 477–487.

Qin, J., Xinsheng, L., Zhou, Y., & Ling, Q. (2015). The effectiveness of china's rmb exchange rate reforms: An insight from multifractal detrended fluctuation analysis. *Physica A: Statistical Mechanics and its Applications*, *421*, 443–454.

Ramberg, D. J., Chen, Y. H. H., Paltsev, S., & Parsons, J. E. (2017). The economic viability of gas-to-liquids technology and the crude oil–natural gas price relationship. *Energy Economics*, *63*, 13–21.

Ramyar, S., & Kianfar, F. (2019). Forecasting crude oil prices: A comparison between artificial neural networks and vector autoregressive models. *Computational Economics*, *53*(2), 743–761.

Safari, A., & Davallou, M. (2018). Oil price forecasting using a hybrid model. *Energy*, *148*, 49–58.

Salisu, A. A., & Oloko, T. F. (2015). Modeling oil price–us stock nexus: A varma–bekk–agarch approach. *Energy Economics*, *50*, 1–12.

Shabri, A., & Samsudin, R. (2017). Hybridizing wavelet and multiple linear regression model for crude oil price forecasting. In *Proceedings of the international conference on computing, mathematics and statistics (iCMS 2015)* (pp. 157–164). Springer.

Šimáková, J. (2011). Analysis of the relationship between oil and gold prices. *Journal of Finance*, *51*(1), 651–662.

Sun, S., Sun, Y., Wang, S., & Wei, Y. (2018). Interval decomposition ensemble approach for crude oil price forecasting. *Energy Economics*, *76*, 274–287.

Sutherland, K. R., & Weihs, D. (2017). Hydrodynamic advantages of swimming by salp chains. *Journal of The Royal Society Interface*, *14*(133), 20170298.

Teetranont, T., Chanaim, S., Yamaka, W., & Sriboonchitta, S. (2018). Investigating relationship between gold price and crude oil price using interval data with copula based garch. In *International conference of the Thailand econometrics society* (pp. 656–669). Springer.

Tiwari, A. K., Mukherjee, Z., Gupta, R., & Balcilar, M. (2019). A wavelet analysis of the relationship between oil and natural gas prices. *Resources Policy*, *60*, 118–124.

Wang, J., Athanasopoulos, G., Hyndman, R. J., & Wang, S. (2018). Crude oil price forecasting based on internet concern using an extreme learning machine. *International Journal of Forecasting*, *34*(4), 665–677.

Wu, Y.-X., Wu, Q.-B., & Zhu, J.-Q. (2019). Improved eemd-based crude oil price forecasting using lstm networks. *Physica A: Statistical Mechanics and its Applications*, *516*, 114–124.

Xie, W., Yu, L., Xu, S., & Wang, S. (2006). A new method for crude oil price forecasting based on support vector machines. In *International conference on computational science* (pp. 444–451). Springer.

Yaojie Zhang, Y., Wei, Y. Z., & Jin, D. (2019a). Forecasting oil price volatility: Forecast combination versus shrinkage method. *Energy Economics*, *80*, 423–433.

Yu, L., Zhao, Y., & Tang, L. (2014). A compressed sensing based ai learning paradigm for crude oil price forecasting. *Energy Economics*, *46*, 236–245.

Zhang, Y.-J., & Wang, J.-L. (2019). Do high-frequency stock market data help forecast crude oil prices? Evidence from the midas models. *Energy Economics*, *78*, 192–201.

Zhang, Y.-J., Yao, T., He, L.-Y., & Ripple, R. (2019b). Volatility forecasting of crude oil market: Can the regime switching garch model beat the single-regime garch models? *International Review of Economics & Finance*, *59*, 302–317.

Zhao, C. L., & Wang, B. (2014). Forecasting crude oil price with an autoregressive integrated moving average (arima) model. In *Fuzzy information & engineering and operations research & management* (pp. 275–286). Springer.

Zhao, L.-T., Wang, Y., Guo, S.-Q., & Zeng, G.-R. (2018). A novel method based on numerical fitting for oil price trend forecasting. *Applied Energy*, *220*, 154–163.

Zhou, Z., & Dong, X. (2012). Analysis about the seasonality of china's crude oil import based on x–12-arima. *Energy*, *42*(1), 281–288.

Zhu, J., Liu, J., Wu, P., Chen, H., & Zhou, L. (2019). A novel decomposition-ensemble approach to crude oil price forecasting with evolution clustering and combined model. *International Journal of Machine Learning and Cybernetics*, *10*(12), 3349–3362.