**A NOVEL HYBRID APPROACH TO FORECAST CRUDE OIL FUTURES USING INTRADAY DATA**

**Highlights**

* Two novel hybrid methods are proposed based on Machine-Learning and Optimization methods
* Dimension reductions proved their importance in forecasting
* Intraday prediction of future contract prices has certain advantages over daily, weekly or monthly prices of crude oil futures prices

**ABSTRACT**

Prediction of oil prices is an implausible task due to the multifaceted nature of oil markets. This study presents two novel hybrid models to forecast WTI and Brent crude oil prices using combinations of machine learning and nature inspired algorithms. The first approach, MARSplines-IPSO-BPNN, Multivariate Adaptive Regression Splines (MARSPlines) find the important variables that affect crude oil prices. Then, the selected variables are fed into an Improved Particle Swarm Optimization (IPSO) method to obtain the best estimates of the parameters of the Backpropagation Neural Network (BPNN). Once these parameters are obtained, the variables are fed into the BPNN model to generate the required forecasts. The second approach, MARSplines-FPA-BPNN, generates the parameters of BPNN through the Flower Pollination Algorithm (FPA). The forecasting performance of these new models is compared to certain benchmark models. The findings document that the MARSplines-FPA-BPNN model performs better than the other competitive models.

*Keywords*: Crude oil prices; Forecasting; Flower Pollination model; Machine learning model; Particle Swarm Optimization model; Intraday data

*JEL Classification*: Q31; Q47; C53

**1. Introduction**

Crude oil is the world's largest energy source. It is considered as the barometer of global economic trends, while its price fluctuations have a significant impact on economic growth (Zhang et al, 2015). The sale and the profit of the oil business are influenced by crude oil’s price fluctuations and have an effect on capital budgeting decisions (Moshiri and Foroutan, 2006). Crude oil prices clearly indicate whether an economy is in recession or in a booming phase, with the central banks changing their monetary policy stance accordingly (Singh et al, 2019). The prices of crude oil are highly uncertain in the long run (EIA, 2019) as they are inclined by natural disasters, government actions, and demand and supply conditions. There are also other various drivers that determine crude oil prices, such as new oil reserves, weather conditions, and wars. The EIA (2019) has predicted that the price of crude oil could rise to $212 per barrel by 2050. In a shorter period of time, it is highly influenced by its lags, non-linearity and price movements in oil markets, such as in Canada, Dubai and Oman (Chen et al., 2016). The short-term crude oil prices fluctuate considerably over the long run (Polanco-Martínez and Abadie, 2016) due to sudden occurrences, such as an explosion in the plant of the Saudi Aramco Company, Corona virus in China and war conflicts in OPEC economies.

Most of the crude oil trades take place in a futures market where prices fluctuate along with the above-mentioned factors. Policymakers, firms and traders/investors are skeptical about a new factor in the market called algorithmic trading. In the short run, this makes the market to fluctuate more than before. The big giants are taking the help of machine learning to forecast future contract prices. Such trades become a problem to the economy, traders/investors, and organizations involved in crude oil markets. Moreover, it becomes difficult to predict prices in the short run (Zhao et al, 2017) due to the factors mentioned above, as well as to the presence of nonlinear patterns and complexity of nature. A precise forecasting can help the stakeholders to minimize losses and make profits in their transactions. Firms that use crude oil as a raw material and traders who are interested in making profits from the commodity markets have to get a good forecasting model to lock the price beforehand. Recent brunt of coronavirus raises doubts about the potential for crude oil demand. The epidemic has already consumed a substantial portion of global oil demand. Hence, it is highly important for policymakers/companies and traders/investors to forecast crude oil prices (Yu et al, 2016). Therefore, the prediction of crude oil prices is a serious research study due to its significant role in the process of economic growth.

As there are different crude oil types and grades, benchmarks are used as a reference price. The major benchmarks used in the literature are West Texas Intermediate (WTI), Brent Blend and Dubai crude. Another recognized blend includes the OPEC reference basket. From the literature, it is known that there are no single models that can predict crude oil prices accurately. Every model has its own pros and cons. Over the last decades, many forecasting models have been developed, tested and recommended (Fan et al., 2016; Chai et al., 2018a). Traditional economic and statistical models, such as the Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA), the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) and other GARCH-family models have been applied to forecast crude oil prices. To use these traditional models, data should be transformed from nonlinear into a linear form or near to linear, from non-stationary to stationary and pretesting for co-integration. These traditional models forecast the values of the converted data, but not the actual data, while the output is of no use in forecasting the actual values of the series explored. Therefore, to overcome the drawback of these models, the literature has recommended machine learning methods that are capable of dealing with noise and nonlinearity problems by itself (Herrera et al, 2019). Chen and Hao (2017) argue that machine learning methods can handle the irregular, chaotic, and nonlinear data better than the traditional statistical models in terms of forecasting performance. Commonly used machine learning models are those of Backpropagation Neural Network (BPNN), Support Vector Machine (SVM), Random Forest (RF), Fuzzy logic and others (Ahmed and Shabri, 2014; Chen and Hao, 2017; Hamid and Shabri, 2017).

Currently, researchers are building hybrid models by combining two or more models and finding it as a successful process in minimizing the deviation between actual and predicted prices (Azadeh et al., 2015; Chai et al., 2018b; Ding, 2018; Chen et al., 2019). This motivates this paper to build novel hybrid models. There are no studies either that forecast futures prices using the optimization algorithms of IPSO and FPA on BPNN. After reviewing the current literature, we find studies on forecasting daily and monthly spot oil prices that helps policymakers to reach efficient decisions (Yu et al, 2016; Zhao et al., 2017; Li et al, 2019). There also exist a few studies that predict futures market oil prices to analyse the daily and monthly volatility forecasting (Ma et al., 2018; Chatziantoniou et al., 2019; Zhang et al., 2019). There are, however, no studies that forecast every sixtieth minute of crude oil futures prices that companies need in case they want to lock the price on that specific day. Hardly, there are forecasting studies that have used MARSplines as dimension reduction methods (Dai et al., 2013; Chinnathambi et al., 2018; Roy et al., 2018), but it has not been used in crude oil price studies till today.

In a nutshell, the objectives of this study are: a) to build effective hybrid forecasting models to forecast intraday crude oil futures prices, b) to check the efficiency between models with and without dimension reduction methods, c) to verify the effectiveness and superiority of the proposed models with benchmark models. The novelties are as follows: Firstly, the analysis proposes two novel hybrid models [MARSplines-IPSO-BPNN (PM-1) and MARSplines-FPA-BPNN (PM-2)] to forecast one-step ahead crude oil futures prices by combining machine-learning and nature-inspired algorithms. The proposed models are the first of their kind in forecasting crude oil futures. Secondly, it uses every sixtieth-minute crude oil futures contract prices of WTI and Brent to attain incremental predictability. WTI and Brent are chosen because they play an important role in stabilizing global financial markets (Chai et al, 2018b). Thirdly, in the PM-1 process, Multivariate Adaptive Regression Splines (MARSPlines) modelling helps to find the important variables that affect WTI and Brent crude oil futures prices. Then, the selected variables are fed into the Improved Particle Swarm Optimization (IPSO) process to find the best parameters of the BPNN modelling. Once the parameters are obtained, the variables are fed into the BPNN model to forecast oil prices. Fourthly, in PM-2 modelling, the parameters of the BPNN model are obtained through the Flower Pollination Algorithm (FPA) process. The performances of the proposed models are compared with benchmark models. The superiority of the models is checked statistically, while the robustness of the model performance is also checked to explore the efficiency of these models across different time frames.

The remaining part of the study is organized as follows. Section 2 discusses the current strands of the literature on forecasting crude oil prices. In Section 3 we describe and explain the novel models proposed, while methodological issues are discussed in Section 4. In Section 5, we provide the empirical results and the associated discussions, while Section 6 concludes.

**2. Literature review**

***2.1 Statistical/Econometric approaches***

The first strand of the literature associated with this paper is the application of traditional/standard econometric models to predict crude oil futures prices. Xiang and Zhuang (2013) analyse the Brent crude oil price using ARIMA modelling. The results demonstrate good prediction effects. A similar study by Zhao and Wang (2014) uses ARIMA modelling to forecast the average annual price of world crude oil. The findings reveal that the model has the perfect approach capability and predictive efficiency for short-term forecasting. Hou and Suardi (2012) employ GARCH models to characterise the volatility of crude oil prices. The authors document that the nonparametric GARCH model yields a superior performance compared to parametric GARCH models. Wang and Wu (2012) examine the dynamics of energy market volatility using both univariate and multivariate GARCH models. Their results highlight that multivariate models show a better performance than univariate models. Nademi and Nademi, (2018) forecast OPEC, WTI and Brent Crude oil prices using Markov Switching AR-ARCH model. The performance of the proposed model is better than that in ARIMA and GARCH modelling. A study by Gupta and Wohar (2017) forecasts WTI oil prices and S&P500 stock returns using Qualitative Vector Autoregressive (Qual VAR) modelling. Qual VAR allows oil and stock returns to behave as non-linear functions of their own past values during the business cycle. The results show that the Qual VAR model outperforms the random walk model. There are a few studies that have used the methods of Self-Exciting Threshold Auto-regressive (SETAR) models (de Albuquerquemello, 2018), and Dynamic Model Averaging (DMA) (Naser, 2016; Drachal, 2016; Wang et al, 2016) in crude oil price forecasting. Such econometric models are capable of forecasting the prices or their volatility under the assumption of linearity and stationarity characteristics in the original data but they are unable to capture the nonlinearity and the complexity present in crude oil prices. There are studies (Jianwei et al, 2017; Li et al, 2019) that document that the forecasting accuracy of such econometric models is less accurate when compared to machine learning models or any combination of machine learning and standard econometric models.

***2.2 Machine learning models***

When compared to traditional prediction tools, machine learning methods, such as Artificial Neural Networks (ANN) (Movagharnejad et al., 2011; Debnath and Mourshed, 2018), SVM (Xie et al. 2006; Papadimitriou et al., 2014; Fan et al, 2016), and Deep Learning (Zhao et al., 2017) are able to deliver better forecasting accuracy due to their ability of handling nonlinear, non-stationary and complex data structures. However, the application of a single method in forecasting oil prices is affected by the parameter sensitiveness and the over fitting of the data used (Tang et al, 2012). Nowadays, researchers have developed hybrid methods to predict crude oil prices so as to increase the accuracy level of forecasts. Jammazi and Aloui (2012) propose a hybrid model with Harr A Trous wavelet decomposition (HTW) and BPNN to forecast WTI crude oil spot prices. They apply wavelet analysis to decompose the data series and feed those data series into an ANN process. They find that this hybrid method seems to outperform benchmark models. Guo et al. (2012) develop a Genetic Algorithm (GA) based on SVM to predict oil prices; their findings provide supportive evidence for this process against the normal SVM approach. The relationship between oil price and the influencing factors such as market and non-market factors was explored (Zhao et al., 2015) using VAR-SVM modelling. Based on the results from VAR model, SVM model of future crude oil prices was determined. VAR-SVM model was superior in accuracy when compared with other forecasting models. Jianwei et al, (2017) predict crude oil prices by applying the combination of a Variational Mode Decomposition (VMD) and an ARIMA model. Their approach highlights that the VMD-based model performs better than an Ensemble Empirical Mode Decomposition (EEMD)-based models. Lahmiri (2015) reaches similar results when applying VMD-based generalized regression neural network (GRNN) ensemble model over Empirical Mode Decomposition (EMD) models to predict California electricity and Brent crude oil price. Ding (2018) proposes a new hybrid model using Akaike Information Criteria (AIC), EEMD and ANN models. He suggested selection of lag using AIC, decomposing the variable using EEMD and forecast through ANN and ADD (Additional Ensemble Method). His findings demonstrate that the combinations of these methods have a superior performance compared to other models.

To ensure the highest accuracy in forecasting, it is always better to extract the variables which affect oil prices patterns. The normal ANN model can give a better accuracy in forecasting financial time series, if the input variables are carefully selected (Atsalakis and Valavanis, 2009). Futures contracts and spot prices of crude oil are influenced by various determinants, many of which are used as input variable while developing a forecasting model. Thus, it is essential to carefully select the most prominent input, if the machine learning models expect to give an efficient and accurate prediction. There are studies that have used the Principal Component Analysis (PCA) Ding et al, 2017; Sad et al, 2019) and the Independent Component Analysis (ICA) (Fan et al, 2016; Jianwei et al, 2017; Xian et al, 2020) as a dimension reduction method in the stock market, as well as in crude oil markets.

However, there are a few studies which have used the dimension reduction methods in the process of crude oil price forecasting. Wu et al (2019) develop a hybrid model using EEMD and Long Short-Term Memory (LSTM) methods. EEMD methods are used to de-noise the data and divide them into different intrinsic mode functions (IMFs). LSTM is applied to extract the necessary IMFs and to forecast crude oil prices. The authors find that selecting relevant IMFs helps to accurately predict crude oil prices against other competitive models. Zou et al (2020) use VMD to extract the risk factors and model the same with ARMA-GARCH. Furthermore, Convolutional Neural Network (CNN) is employed to assemble the different risk forecast to find out the ensemble forecasted prices. The authors document that their proposed model which reduces the dimensions involved performs better than other competitive models. Tang et al (2020) extract information from search engines (search engines data) to predict crude oil prices. They use an advanced version of EMD modelling, called Multivariate EMD (MEMD), to extract the relevant modes, and then BPNN is applied to forecast crude oil prices. Xian et al (2020) integrate EEMD and ICA to extract the features from the dataset for forecasting WTI crude oil prices.

In particular, certain studies have applied decomposing methods, such as Wavelet (Jammazi and Aloui, 2012; Hamid and Shabri, 2017; Uddin et al, 2019), EMD (Hu et al, 2018; Ding, 2018), and VMD (Jianwei et al, 2017; Lahmiri, 2015; He et al, 2018) to decompose the input data series into subsets of data and then they use optimization methods, namely GA, Artificial Bee Colony (ABC), and Particle Swarm Optimization (PSO) to obtain the optimal parameters of the forecasting methods of ANN and SVM. The recent techniques, namely, EMD is facing mode mixing problems, while the VMD approach has its own problems in setting parameters, leading either to over decomposing the series or to under decomposing the series; this, in turn, implies lower accuracy in the forecasting process (Isham et al, 2019). Though the machine learning methods are efficient in handling nonlinear and nonstationary data, MARSplines methods are capable of finding important input variables for models, such as BPNN, SVM and RF. It can act as a dimension reduction approach, as well as a nonlinear regression method. Therefore, in this paper we propose MARSplines that help in finding the important variable(s) to predict crude oil prices accurately using its nonlinear, irregular and chaotic data handling features. It has been applied as a variable selection tool in stock market studies (Chen at al, 2019; Kao et al, 2013), while no crude oil studies have used this approach to obtain the important variables.

With the machine learning methods, computational algorithms are mixed to find the optimal parameters of the model. Among the computational algorithms, the Particle Swarm Optimization (PSO) approach is highly popular among forecasting studies. This is used in finding the optimal parameters of the ANN, SVR and other Machine Learning modelling approaches. There was perplexity in finding the optimal parameters using normal PSO, while it also ended in local minima (Xiao et al, 2014). Due to this drawback, Alfi and Modares (2011) develop an improved PSO (IPSO) method, called the Adaptive PSO approach, which helps in finding the optimal system and control parameters. Wang and Shoup (2011) develop Poly-hybrid PSO for intelligent parameter adjustment. In this study, we use standard PSO methods by implying crossover and mutation functions of the GA (Xiao et al, 2014) to increase the candidate particle performance. This improved PSO (IPSO) helps to obtain the optimal parameters of BPNN. Another computational algorithm called the Flower Pollination Algorithm (FPA), introduced by Yang (2012), has the ability to optimize the parameters globally and thus to increase the performance of forecasting. This FPA approach is improved compared to GA and PSO in resolving issues related to optimization (Yang, 2012). Chiroma et al (2016) develop a hybrid model by combining FPA and BPNN to forecast the OPEC countries petroleum consumption. This model outperforms other hybrid models by producing lower forecast errors. FPA has been also used in forecasting wind speeds (Zhang et al, 2017) and electricity loads (Pan et al, 2019). This optimization algorithm’s performance is yet to be explored across different fields. In this study, we use FPA to optimize neural network parameters so as to build a hybrid model.

Overall, the current literature documents that researchers are trying to forecast crude oil prices for a long time using traditional models, while recently they are using single machine learning methods or a combination of two or more of those methods. Due to changing market and trading scenarios, stakeholders in oil markets must be aware about the availability of new methods to forecast crude oil prices. At the same time, researchers are focusing on daily, weekly and monthly crude oil price forecasts, but not on intraday futures contract prices, where most of the crude oil trades occur. With high frequency data, incremental forecasting ability can be achieved (Wang et al., 2010; Degiannakis et al., 2018). Scant attention has hitherto been provided to study the future markets with very high frequency data. In this study, the analysis investigates the combination of machine learning methods with nature inspired algorithms in forecasting every sixtieth minute crude oil futures contract prices.

**3. Methodology of forecasting**

***3.1 Multivariate adaptive regression spline***

MARSplines was developed by Friedman in 1991 (Friedman, 1991); it is a nonlinear and nonparametric method that does not need any assumption or link between the dependent and the independent variables. This method helps in finding the variables that are affecting the course and pattern of the dependent variable. It works on the basis of the Basis Function (BF) for each independent variable as given below:

*Max(0, X-v) or Max(v-X,0)* (1)

where, *v* is a knot, and *X* is an independent variable. The general MARSplines model yields:

(2)

where is the forecasted price, is a constant, is the parameter coefficient of *mth* BF, is *mth* BF, *M* is number of BF terms. The best model is selected using the Generalized Cross Validation (GCV) criterion. It is a goodness of fit test of the MARSplines model which helps to select the optimal independent variables. The best model is the result of two state processes in MARSplines: (i) State 1: while building the model, the MARSplines considers more numbers of BFs using all possible combinations of links among the independent variables so as to overfit the data, (ii) State 2: using the *GCV*, it removes the least contributing combinations of variables. The GCV is calculated as follows:

*GCV = MSEtrain /*  (3)

where *MSEtrain* is the Mean Square Error of training data, is the effective number of parameters, and *n* is the number of observations.

***3.2*** ***Backpropagation Neural Network***

There are various types of ANN, mostly to predict the time series data. In our case, the BPNN approach is used. It has the ability to handle nonlinear and chaotic behaviors of the data through the employment of activation functions and repeatedly revising the weight of the nodes. It has three layers: input, hidden and the output with nodes. The independent variables are fed towards the input layer (i.e., a node for every independent variable), it is processed using a tansig activation function and, then the weights are updated and sent to the next layer called the hidden layer (i.e., multiple nodes). This process repeats in the middle layer and the updated weights are sent to the output layer (i.e., single node). In the output layer, the output value, as well as the actual value of crude oil prices are checked using the Mean Square Error (MSE) criterion. If too much deviations occur, then the last layer neuron returns the error to the hidden layer neurons and the hidden layer neurons, in turn, returns the errors to the input layer so as to recalculate the weights. This process continues until the network output provides the least deviation between the actual and the output value. The basic BPNN structure can be seen in Figure 1.

**H 1**

**I 1**

**I 2**

**H 3**

**H 2**

**O**

Input Hidden Output

**Figure 1.** BPNN architecture with 2-3-1 network topology

The forecasting equation (Equation 1) of BPNN is:

*T* = ƒ(*m*I+*e*) (4)

where *T* is the Network Output, *m* represents the weight of the neuron, and *e* represents the error parameter.

***3.3 Optimization methods***

**3.3.1 Particle Swarm Optimization (PSO)**

The PSO was developed by Eberhart and Kennedy (1995). The idea was generated from flock and swarm behaviors by birds, fishes, and insects for food. The PSO helps to optimize the parameters of ANNs. It is considered better than the GA in terms of the optimization process.

In a normal PSO algorithm, let us assume that a search space has *m* dimensions, while the swarm particles can be denoted by a *n* vectors Ai = (ai1, ai2,…,aim)T. The objective function of the original optimization problem can evaluate particle fitness. Every particle velocity is denoted by *n* dimensional vectors Bi= (bi1, bi2,…,bim)T. The specific best position of the ith particle is represented as Pz = (pz1,pz2, …, pzm)T and the global best position as Gz = (gz1, gz2, …, gzm)T. The new particle’s velocity is expressed as:

bij(t+1) =*w*bij(t) + c1r1[pzj – aij(t)] + c2r2 [gzj – aij(t)] (5)

position: aij(t+1) = aij(t) + bij(t+1), j = 1,2,..,h (6)

bij [bmin, bmax], aij [amin, amax] (7)

where c1 and c2 are considered as acceleration parameters, which are self-thinking of particle and group, respectively, *w* is the inertia weight, and r1 and r2 are random numbers ranging from 0 to 1. Every dimension of particle velocities are fixed to a maximum velocity: bmax. The particles new position is calculated by Equation (5).

***3.3.1.1 Improved PSO (IPSO)***

The general PSO can produce local minima and fails to attain optimal solutions. There are other PSOs which were built to overcome these drawbacks, but they were not able to do it efficiently. Hence, Xiao et al. (2014) develop an improved PSO by incorporating GA in it to solve the local minima and convergence problems.

In stage 1, to improve the balance of the acceleration parameters c1 and c2, dynamic acceleration parameter adjustment strategies are used and these parameters are controlled using increasing and decreasing arccosine functions. This strategy is followed to find the best solution in the initial stages of the optimization process and to overcome the drawbacks of local minima and parameter convergence in the last stages of the optimization process. The right selection of *w*, c1 and c2 provides stability among global and local search (Engelbrecht, 2005). This is represented as:

= + - x (8)

= - (9)

where and are the initial and final values of acceleration parameters’ initial values, respectively, *epoch* is the current *epoch* number, and is the maximum epoch number. Appropriate inertia weights and acceleration parameters makes the balance between local and global search. Thus, based on the fitness value of the particles, adaptive nonlinear adjustment inertia weight strategies are used in this study. A particular strategy yields:

w =  (10)

where wmax and w*min* are the minimum and maximum w, fitness is a current fitness value, *fit*min and *fit*avg are minimum and average fitness values of particles, respectively. From Equation (10), we can infer that the *w* will rise when the particles’ fitness values are stable and it will come down when the values are scattered. Therefore, to preserve their properties, the inertia weights of the superior particles whose fitness values are greater than the average fitness value are lower. In comparison, the inertia weights of weak particles whose fitness values are lower than the average fitness value are greater in order to search for better space.

In stage 2, the Adaptive Genetic Operators (AGO) process is applied to increase the candidate’s particle performance. Based on certain probabilities, the particles will implement crossover and mutation operations. The important operator in GA is crossover; two-point crossover is used in this study which develops offspring by mixing the sections of the parental genome arbitrarily. The crossover rate is planned according to the iteration number and not to the individual fitness. The crossover rate is as follows:

*pct* = *pc,max* x *q(-t/T),* *q*(2,10) (11)

*pc(t)* =  (12)

[*pc,min* < *pc,max*] (0,1) (13)

where *pct* is a calculated variable, *T* is the maximum epoch number, *t* is the iteration number at present, *q* represents crossover probability’s decreasing coefficient, *pc,min* represent minimum crossover probability, *pc,max* is the maximum crossover probability, and *pc(t)* is the crossover probability of *t*th iteration.

To initiate genetic mutations, fractions of off springs are arbitrarily chosen. Then, the mutation operator randomly selects a domicile from a bit-string and turned over its contents. In order to overcome early convergence, an adaptive mutation rate is applied based on the individual fitness, which is based on Equations 14 and 15 below:

 (14)

[*pm,min* < *pm,max*] (0,1) (15)

where *fmax* is the maximum fitness value of current population, *favg* is the average fitness value of every epoch population, *f* is the fitness value of current mutation individual, *pm,min* is the minimum mutation probability, and *pm,max* is the maximum mutation probability. To increase the efficiency of search, the AGO picks the merits of PSO training speed and global search of GA’s. The AGO functions is as follows:

*CPn = 1 - , n = 1,2,….,*  (16)

*rand(0,1) <CPn*(17)

where *n* represents the *n*th epoch number. In every epoch, a number is generated randomly between 0 and 1. When the number is below *CPn*, the present particle initiates the AGO. The AGO will be implemented at a smaller probability when the premature iterations *CPn < 1*, according to Equation (16) and vice versa. During the course of iterations, the AGO increases population search space shrinking to make the particles move out from an earlier optimal search to larger space search. By retaining the multifariousness of the population by particles leads to intensifying the prospects of finding good solutions.

***3.3.1.2 IPSO-BPNN***

The disadvantage of neural network is the manual parameter feeding and confining to local minima. These demerits can be removed by incorporating PSO-GA into the BPNN to optimize the weights (*w*) and biases (*b*). The weights and biases are optimized by using Decimal PSO (DPSO), while the structure of the BPNN is optimized using Binary PSO (BPSO). Let the count of nodes in the input, hidden and output layer are denoted as *L, H* and *O,* respectively.The weight and the error terms are referred as:

*V = [w111,…, w11H, w121,… w12H,… w1L1,….w1LH,b11,… b1H, w211,…, w21O, w221,… w22S,… w2H1,….w1HO,b21,… b2O]*  (18)

*d=L(H+1) + H(O+1)* (19)

where *w1ij (i=1,2,..,L; j=1,2,…H)* denote the input layer to the hidden layer vector’s weight, w2ij *(i=1,2,..,H; j=1,2,…O)* denote the hidden layer to the output layer vector’s weight, *b1i (i=1,2,…H)* denoting bias vector of hidden layer in the BPNN, *b2i (i=1,2,…S)* shows the bias vector of the output layer in BPNN and *d* is the dimension of vector *V*.

The binary coded particle using BPSO is denoting the equivalent hidden layer node’s presence (1) and absence (0). According to Equation 5, the particle velocity is updated. Using the state transition probability, the position of the particle is updated, which depends on the particle velocity. Generally, the nodes in the hidden layer are not more than the sum of twice of the input and output layer nodes. The BPSO formulation is described as follows:

 (20)

*j = 1,2,…h*

*h=2 x (L+H)*

where ρij(t+1) (0,1), j is a random number, sig(.) denotes the sigmoid function, and h denotes the vector’s dimension. When x = 1, it represents that the corresponding nodes in the hidden layer are present, and the weights and bias vector of that node in DPSO is valid, and vice versa.

**3.3.2 Flower Pollination Algorithm**

Like PSO, the Flower Pollination Algorithm (FPA) is a nature inspired algorithm, developed by Yang (2012) on the features of the pollination process of flowering plants using the below set of rules:

a. Biotic pollination and cross-pollination are known to be mechanisms of global pollination with pollinators executing Lévy flights.

b. Local pollination is seen as self-pollination and abiotic.

c. Reproduction probability is seen as flower constancy that is resemblance of two flowers related.

d. The switch probability *p* (0,1) is used to regulate global and local pollinations, because the local pollination can have a fraction *p* in the entire process due to physically close and other factors like wind.

In the global pollination, the pollens can move to longer distances by the pollinators, like insects and birds. This process confirms the pollination and reproduction of the fittest solution which is denoted as u\*. The equation of flower constancy (first rule) is:

*qit+1 = qit + L(qit - u∗)*  (21)

where *qit* is the pollen *i* or solution vector *qi* at iteration *t*, and *u∗* is the present best solution of all solutions at the existing generation/iteration. *L* is the strength of the pollination, which essentially is a step size. There are insects which can travel long distances; to simulate this nature, Lèvy flight is used. Hence L>0 from the Lèvy distribution is drawn as follows:

 (22)

where  refers to the standard gamma function and this particular distribution is valid for large steps s>0. The local pollination given in the second rule and the flower constancy of the flower are mentioned as:

*qit+1 = qit + (qjt – qkt)* (23)

where *qjt* and *qkt* are solution vectors, drawn randomly from the solution set. is a parameter that is drawn from a uniform distribution in the range from 0 to 1. The FPA is recognized for the following: (a) to elude the local landscape and to have a large search space (explore), the insect pollinators trips long distances, and (b) it ensures the convergence to the optimal solution (exploit).

**3.3.1 FPA- BPNN**

The BPNN’s weights and biases are optimized using FPA. The bias in BPNN is represented as the amount of pollen gametes (n). Before initiating the process, the BPNN’s and FPA’s initial parameters have to set due to the sensitivity of the parameter setting in the first step. From the second iteration onwards, the weights and biases are updated using FPA. The output of FPA-BPNN is compared with the given Mean Square Error value (MSE) in every iteration, which is a stopping criterion. If the criterion is not met, again the FPA process starts generating the parameters of BPNN, till it finds the optimal solution. The flow chart (Figure 2) shows the steps to generate final output by FPA-BPNN process. The average sum of error squares are represented as:

MSE*i* = {Vu1(*q*), Vu2(*q*), Vu3(*q*),….., Vun(*q*)} (24)

where the Vu1(*q*), Vu2(*q*), Vu3(*q*),….., Vun(*q*)} show the mean performance of every iteration. The MSE, which is imitated by pollen, is found after the inputs are processed for every pollen population. This results in the pollen gamete q*j* and is calculated through the below Equation 25:

*qj* = Min{Vu1(*q*), Vu2(*q*), Vu3(*q*),….., Vun(*q*)} (25)

The balance pollen gametes are considered by the lasting MSEs. Next, with the help of the Lèvy flight, a new result qit+1 is obtained for global pollination *i* as follows:

*qit+1 = qit + L(qit - u∗)*  (26)

where *qit* is the pollen *i* or solution vector *qi* at iteration *t*, *u∗* is the present best solution of all solutions at the existing generation/iteration, and *L* is the strength of the pollination. A new result *qit+1* local pollination *i* is developed:

*qit+1 = qit + (qjt – qkt)* (27)

The remaining pollen gametes *qi* movements towards *qj*are developed through Equation 28:

 *Q* = (28)

The pollen gamete examines the travels from *qi* towards *qj* through the Lévy flight process, as given below:

 *Qm* = (29)

where *Qm* is the movement of pollen gamete travels from *qi* towards *qj.* The weights (Wxn+1, Wxn) and the biases (Bxn+1, Bxn) of every layer are adjusted based on Equations (30) and (31), respectively:

Wxn+1 = Wxn –*Q*m (30)

Bxn+1 = Bxn –*Q*m (31)

Start

Run BPNN Structure

Initialization of FPA

Enters network weights with every pollen

Training Data

Calculate MSE

FPA-BPNN computed?

No

Yes

Calculate Fitness

Whether fitness computed for all pollens?

No

The current iteration number is added by one

Yes

Update the pollens global optimum fitness based on the fitness of each pollen

Update each nest and position

Whether fitness computed for all pollens?

No

Yes

End

**Figure 2.** Flow chart of FPA-BPNN

**4. Data and forecasting methodology**

Among the different crude oil markets, most contract trades occur in terms of European Brent prices (Brent) and U.S. West Texas Intermediate prices (WTI). This study focuses on WTI and Brent oil futures contract prices (rollover to the nearby futures contract) from the New York Mercantile Exchange and the Intercontinental Exchange, respectively. Good forecasting models can be generated when good predictor variables are used. To forecast the intraday values of KOSPI200 index, Son et al. (2012) use a list of technical indicators (provided in Table 1) as predictor variables. In this study, to forecast the intraday futures contract prices of WTI and Brent, the list of technical indicators considered by Son et al (2012) is used. During April 2019-Sep 2019, several events took place around the globe, such as the US-China trade war, political tensions in the Middle East, Iranian forces seizing a British oil tanker, talks between India and China on the establishment of an “oil buyer’s club, and attacks on Saudi Arabia’s Aramco Company which led to high oil price fluctuations. Moreover, weekly and monthly prices always subsidized these effects due to time lapses. Hence, this study investigates the intraday forecasting of WTI and Brent Crude oil futures prices, respectively. Every sixtieth minute, close prices for crude oil (US$ per Barrel) are obtained from Bloomberg for the period of 6 months (April, 1 to Sep, 30, 2019). Missing data are handled using the mean value of that day and if any given day has few data points, that day’s data are not considered. The raw data are converted into a normalized form using min-max normalization (Han et al., 2011) to reduce the forecasting error. Then, the dataset is divided into 80% (training) and 20% (testing) to build and test the models, respectively.

To forecast oil prices, the analysis introduces two hybrid models, namely MARSplines-IPSO- BPNN (PM-1) and MARSplines-FPA-BPNN (PM-2). The performance of these models is compared with benchmark models, such as IPSO-BPNN, FPA-BPNN, MARSplines-BPNN, MARSplines and BPNN. The optimal model is chosen on the basis of the lowest MSE value across all models used in this study, except for the case of MARsplines, where the lowest GCV is used by default. The forecasted values are converted into a raw form to compare the performances of the models with one another. The statistical measures, namely MSE, Mean Absolute Percentage Error (MAPE), and the Theil U-statistic are used to check the comparison. Moreover, the Diebold

Mariano (DM) test (Diebold and Mariano, 1995) is also used to test the superiority of the proposed model against other competitive models.

**Table 1: List of predictor variables**

|  |  |
| --- | --- |
| **Technical Indicator** | **Variables** |
| Moving Average (MA) | MA(5), MA(10), MA(20) |
| Exponential Moving Average (EMA) | EMA(5), EMA(10), EMA(20) |
| Disparity (DISP) | DISP(5), DISP(10), DISP(20) |
| EMA Disparity (EDISP) | EDISP(5), EDISP(10), EDISP(20) |
| Relative Strength Index (RSI) | RSI(5), RSI(10), RSI(20) |
| Momentum (MTM) | MTM(5), MTM(10), MTM(20) |
| Relative Difference in Price (RDP) | RDP(1), RDP(2), RDP(3), RDP(5), RDP(10), RDP(20) |
| Price oscillator (OSCP) | OSCP |
| Price oscillator (EOSCP) | EOSCP |

**5. Empirical analysis**

***5.1 Dimension reduction***

The technical indicators are fed into MARSplines for dimension reduction. The list of technical indicators used in this study is listed in Table 1. Among these variables, EMA(5), RDP(1), RDP(3), and EDISP(5) are selected by MARSplines for both energy commodity futures price forecasting based on its GCV value (0.000319 and 0.000345 for WTI and Brent prices, respectively).

***5.2 MARSplines-IPSO-BPNN***

The independent variables selected by MARsplines are fed into BPNN where its parameters are optimized by IPSO. Initially, the parameters of IPSO-BPNN are set as: the size of swarm population = 40, the determined epochs of IPSO = 500, the epochs of BPNN = 2000. Next, in DPSO, the maiden particle positions are set as random numbers between -20 and 20, the maiden particle velocity randomly ranges between -10 to 10, , = 1.05, wmin = 0.25, and wmax = 0.9. In BPSO, the initial particle positions are set as random numbers ranging from -1 to 1, the initial particle velocity randomly ranges from -0.7 to 0.7, c1 = c2 = 1.15 and w = 1. In GA, Pc,min = 0.4, Pc,max = 0.95, Pm,min = 0.01, and Pm,max = 0.7. The MSE is set at 2 x 10-6 or less is reached. The results of the model show that the 9 neurons in hidden layers are used to achieve the least MSE of 0.000016 and 0.000015 for WTI and Brent, respectively (Table 2).

***5.3 MARSplines-FPA-BPNN***

Based on the initial experiments, the pollen gametes numbers are finalized to 22, p = 0.11 (trial and error), dimension = 350 and the iteration = 1000. Similarly, the BPNN input layer’s neurons are set equal to the variables selected by applying MARSplines and one neuron in the output layer. Based on Zhang et al (2001), the analysis sets one hidden layer for the network and according to Berry and Linoff (1997), the hidden layer neurons are set to 13. Hence, the basic network topology used is 4-13-1. The weights and the biases of the BPNN are initialized by FPA. The BPNN’s initial weights are fed using FPA’s weight values at each epoch. The FPA feds’ updated weights to BPNN structure in every epoch. The model convergence is set to stop when the MSE reaches 2 x 10-6 or maximum epochs reach, otherwise, the BPNN continues to train the model till is satisfied with the stopping conditions. The best model occurs when the MSE reaches the minimum level of 0.000002 and 0.000004 for WTI and Brent, respectively (Table 2).

***5.4 IPSO-BPNN, FPA-BPNN and MARSPlines-BPNN***

For IPSO-BPNN and FPA-BPNN, the same parameters are set as given in the sections MARSplines-IPSO-BPNN and MARSplines-FPA-BPNN respectively. From Table 2, it is observed that the lowest MSEs of IPSO-BPNN for WTI and Brent are 0.000608 and 0.000613, respectively. For the case of FPA-BPNN, the MSEs are 0.000597 and 0.000031 for WTI and Brent, respectively. For MARSplines-BPNN, the variables selected by MARSplines are directly fed into BPNN and through a trial and error basis. The lowest MSEs achieved are 0.000715 and 0.000602 for WTI and Brent, respectively.

***5.5 Comparison of results***

The normalized output (Table 2) is denormalized into a raw form to compare the performances of the models using statistical measures, namely MSE, MAPE and Theil U-statistics. The comparison results are reported in Table 3. We can see that the error values obtained by using PM-2 are lower than the other novel model PM-1 and the benchmark models under MSE, MAPE and Theil U- statistic for both WTI and Brent crude oil futures prices.

**Table 2: Comparison of normalized outputs of WTI and Brent crude oil futures prices**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Crude oil | Models | MSE | MAPE | Theil U |
| WTI | PM-1 | 0.000016 | 0.2933 | 0.0050 |
| **PM-2\*** | **0.000002** | **0.0808** | **0.0019** |
| IPSO-BPNN | 0.000608 | 1.7639 | 0.0306 |
| FPA-BPNN | 0.000597 | 1.6917 | 0.0303 |
| MARSplines-BPNN | 0.000715 | 1.9082 | 0.0332 |
| MARSplines | 0.002579 | 3.9518 | 0.0636 |
| BPNN | 0.027066 | 9.1684 | 0.1822 |
|  |  |  |  |  |
| Brent | PM-1 | 0.000015 | 0.3916 | 0.0060 |
| **PM-2\*** | **0.000004** | **0.1722** | **0.0031** |
| IPSO-BPNN | 0.000613 | 1.9747 | 0.0382 |
| FPA-BPNN | 0.000031 | 0.5307 | 0.0087 |
| MARSplines-BPNN | 0.000602 | 1.9898 | 0.0378 |
| MARSplines | 0.000977 | 2.5463 | 0.0482 |
| BPNN | 0.002428 | 4.7320 | 0.0768 |

\*denotes Least errors

**Table 3: Comparison of denormalised outputs of WTI and Brent crude oil futures prices**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Crude oil | Models | MSE | MAPE | Theil U |
| WTI | PM-1 | 0.004111 | 0.0297 | 0.0005 |
| **PM-2\*** | **0.000621** | **0.0079** | **0.0002** |
| IPSO-BPNN | 0.152395 | 0.1776 | 0.0034 |
| FPA-BPNN | 0.149436 | 0.1699 | 0.0034 |
| MARSplines-BPNN | 0.179046 | 0.1924 | 0.0037 |
| MARSplines | 0.645522 | 0.4096 | 0.0071 |
| BPNN | 5.681008 | 1.3029 | 0.0208 |
|  |  |  |  |  |
| Brent | PM-1 | 0.005815 | 0.0322 | 0.0006 |
| **PM-2\*** | **0.001521** | **0.0122** | **0.0003** |
| IPSO-BPNN | 0.228665 | 0.1665 | 0.0038 |
| FPA-BPNN | 0.011880 | 0.0449 | 0.0008 |
| MARSplines-BPNN | 0.224792 | 0.1668 | 0.0038 |
| MARSplines | 0.364451 | 0.2158 | 0.0048 |
| BPNN | 0.905578 | 0.4051 | 0.0076 |

\*denotes Least errors

To statistically check the performances of the models generated by the seven modelling methods, the DM test is performed. Tables 4 and 5 report the results of the DM test for WTI and Brent forecasted futures oil prices, respectively. Through the DM test, it is inferred that when the PM-1 and PM-2 are treated as target models, the p-values are lesser than the significance level of 5%, implying that when MARSplines are used with FPA-BPNN and IPSO-BPNN, the forecasted values are superior against those provided by the other models at a minimum of 95% confidence level for WTI and at 90% confidence level for Brent futures contract oil prediction, respectively. For WTI (Table 4), all tested models are significantly better than the comparable models at a minimum of 90% confidence level, except the performances between IPSO-BPNN and FPA-BPNN, where there is no difference in the predicted values even at the 90% level. In contrast, for the case of Brent prices (Table 5), the forecasting performance of IPSO-BPNN versus MARSplines-BPNN, IPSO-BPNN versus MARSplines and MARSplines-BPNN versus MARSplines are not having any difference (p value > 10%) in prediction.

**Table 4: DM test results on WTI**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| TESTED MODEL | COMPARED MODELS | | | | | |
| PM-2 | IPSO-BPNN | FPA-BPNN | MARSplines  -BPNN | MARSplines | BPNN |
| PM-1 | 2.1452  (0.0319)\*\* | -2.4158  (0.0156)\*\* | -2.3281  (0.0199)\*\* | -2.3235  (0.0201)\*\* | -2.6293  (0.008)\*\*\* | -2.7199  (0.0065)\*\*\* |
| PM-2 |  | -2.4071  (0.016)\*\* | -2.3238  (0.0201)\*\* | -2.3198  (0.0203)\*\* | -2.6269  (0.0086)\*\*\* | -2.7211  (0.0065)\*\*\* |
| IPSO-BPNN |  |  | 0.6277  (0.5301) | -1.8888  (0.0589)\* | -2.6858  (0.0072)\*\*\* | -2.6729  (0.0075)\*\*\* |
| FPA-BPNN |  |  |  | -2.0357  (0.0417)\*\* | -2.6886  (0.0071)\*\*\* | -2.673  (0.0075)\*\*\* |
| MARSplines-BPNN |  |  |  |  | -2.7182  (0.0065)\*\*\* | -2.6619  (0.0077)\*\*\* |
| MARSplines |  |  |  |  |  | -2.5155  (0.0118)\*\* |

\*\*\*p-value<0,01; \*\*p-value<0,05; \*p-value<0,1

**Table 5: DM test results on BRENT**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| TESTED MODEL | COMPARED MODELS | | | | | |
| PM-2 | IPSO-BPNN | FPA-BPNN | MARSplines  -BPNN | MARSplines | BPNN |
| PM-1 | 2.5812  (0.0098)\*\*\* | -1.9555  (0.0505)\* | -2.0046  (0.045)\*\* | -1.9371  (0.0527)\* | -1.7677  (0.0771)\* | -2.2678  (0.0233)\*\* |
| PM-2 |  | -1.9661  (0.0492)\*\* | -2.2057  (0.0274)\*\* | -1.948  (0.0514)\* | -1.7749  (0.0759)\* | -2.2693  (0.0233)\*\* |
| IPSO-BPNN |  |  | 1.9535  (0.0507)\* | 0.9845  (0.3248) | -1.5061  (0.132) | -2.3626  (0.0181)\*\* |
| FPA-BPNN |  |  |  | -1.9343  (0.053)\* | -1.7634  (0.0778)\* | -2.2695  (0.0232)\*\* |
| MARSplines-BPNN |  |  |  |  | -1.5465  (0.1219) | -2.368  (0.0178)\*\* |
| MARSplines |  |  |  |  |  | -2.6523  (0.0079)\*\*\* |

\*\*\*p-value<0,01; \*\*p-value<0,05; \*p-value<0,1

To check the forecasting robustness of the models used, the data sample is divided into training and testing in the ratio 60:40, 70:30, 80:20 (original) and 90:10. The robustness evaluation of WTI and Brent’s one-step ahead future contract’s denormalized predicted values are presented in Tables 6 and 7. From these Tables it can be inferred that the PM-2 model, i.e. MARSplines-FPA-BPNN, forecast WTI and Brent futures crude oil prices with the least errors compared to the other models. This highlights the importance of using dimension reduction methods in forecasting studies. The next model which is closely forecasting WTI and Brent futures contracts prices was the MARSplines-IPSO-BPNN. FPA-BPNN performs better than IPSO-BPNN without the application of dimension reduction method, probably because of the presence of demerits of BPNN, its forecasting performance (Table 6 and 7) is lower than other models in this study.

**Table 6: Robustness evaluation of models on forecasting WTI crude oil futures prices**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Ratio | Models | MSE | MAPE | Theil U |
| 60:40 | PM-1 | 0.006276 | 0.0375 | 0.0009 |
| **PM-2\*** | **0.000950** | **0.0083** | **0.0007** |
| IPSO-BPNN | 0.185243 | 0.1925 | 0.0050 |
| FPA-BPNN | 0.124397 | 0.1824 | 0.0050 |
| MARSplines-BPNN | 0.370705 | 0.2413 | 0.0519 |
| MARSplines | 0.789793 | 0.5442 | 0.0085 |
| BPNN | 9.516864 | 1.7615 | 0.0400 |
|  |  |  |  |  |
| 70:30 | PM-1 | 0.005141 | 0.0345 | 0.0005 |
| **PM-2\*** | **0.000766** | **0.0081** | **0.0003** |
| IPSO-BPNN | 0.249132 | 0.3670 | 0.0039 |
| FPA-BPNN | 0.171215 | 0.1345 | 0.0035 |
| MARSplines-BPNN | 0.372334 | 0.2349 | 0.0038 |
| MARSplines | 0.851140 | 0.4921 | 0.0073 |
| BPNN | 15.468355 | 1.8084 | 0.0324 |
|  |  |  |  |  |
| 80:20 | PM-1 | 0.004111 | 0.0297 | 0.0005 |
| **PM-2\*** | **0.000621** | **0.0079** | **0.0002** |
| IPSO-BPNN | 0.152395 | 0.1776 | 0.0034 |
| FPA-BPNN | 0.149436 | 0.1699 | 0.0034 |
| MARSplines-BPNN | 0.179046 | 0.1924 | 0.0037 |
| MARSplines | 0.645522 | 0.4096 | 0.0070 |
| BPNN | 5.681008 | 1.3029 | 0.0208 |
|  |  |  |  |  |
| 90:10 | PM-1 | 0.003149 | 0.0318 | 0.0005 |
| **PM-2\*** | **0.000696** | **0.0080** | **0.0002** |
| IPSO-BPNN | 0.171143 | 0.2259 | 0.0037 |
| FPA-BPNN | 0.157811 | 0.1616 | 0.0034 |
| MARSplines-BPNN | 0.407063 | 0.2858 | 0.0041 |
| MARSplines | 0.557972 | 0.4758 | 0.0068 |
| BPNN | 8.464817 | 1.5967 | 0.0211 |

\*denotes Least errors

**Table 7: Robustness evaluation of models on forecasting Brent crude oil futures prices**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Ratio | Models | MSE | MAPE | Theil U |
| 60:40 | PM-1 | 0.002833 | 0.0370 | 0.0061 |
| **PM-2\*** | **0.002602** | **0.0955** | **0.0028** |
| IPSO-BPNN | 0.023574 | 0.0462 | 0.0082 |
| FPA-BPNN | 0.008276 | 0.0464 | 0.0074 |
| MARSplines-BPNN | 0.396756 | 0.2461 | 0.0096 |
| MARSplines | 0.732565 | 0.3146 | 0.0156 |
| BPNN | 15.14008 | 0.4597 | 0.0386 |
|  |  |  |  |  |
| 70:30 | PM-1 | 0.00591 | 0.0247 | 0.0004 |
| **PM-2\*** | **0.001425** | **0.0117** | **0.0002** |
| IPSO-BPNN | 0.283973 | 0.1965 | 0.0059 |
| FPA-BPNN | 0.010098 | 0.0389 | 0.0007 |
| MARSplines-BPNN | 0.295423 | 0.1255 | 0.0034 |
| MARSplines | 0.889441 | 0.3107 | 0.0087 |
| BPNN | 9.563925 | 1.9545 | 0.0107 |
|  |  |  |  |  |
| 80:20 | PM-1 | 0.005815 | 0.0322 | 0.0006 |
| **PM-2\*** | **0.001521** | **0.0122** | **0.0003** |
| IPSO-BPNN | 0.228665 | 0.1665 | 0.0038 |
| FPA-BPNN | 0.011880 | 0.0449 | 0.0008 |
| MARSplines-BPNN | 0.224792 | 0.1668 | 0.0038 |
| MARSplines | 0.364451 | 0.2158 | 0.0048 |
| BPNN | 0.905578 | 0.4051 | 0.0076 |
|  |  |  |  |  |
| 90:10 | PM-1 | 0.005933 | 0.0137 | 0.0008 |
| **PM-2\*** | **0.002475** | **0.0134** | **0.0003** |
| IPSO-BPNN | 0.314854 | 0.2734 | 0.0023 |
| FPA-BPNN | 0.047117 | 0.0530 | 0.0013 |
| MARSplines-BPNN | 0.38021 | 0.3264 | 0.0084 |
| MARSplines | 0.435798 | 0.1923 | 0.0022 |
| BPNN | 3.914398 | 0.2579 | 0.0131 |

\*denotes Least errors

**6. Conclusion**

In the current state of globalization, crude oil prices play a vital role in economic growth. Therefore, it is important to forecast these prices to reach better economic and investment decisions. However, forecasting crude oil prices is not an easy task, because of its volatility nature. Though researchers have used statistical methods to forecast oil prices, it still remains a difficult task. Previous literature focuses on nonlinear methods, such as neural networks, support vector machines, and fuzzy logic for forecasting oil prices. Due to the complex dynamic behavior of crude oil prices, this study proposed two new hybrid approaches namely, MARSplines-IPSO-BPNN and MARSplines-FPA-BPNN to forecast every sixtieth minute of WTI and Brent crude oil futures contract prices (rollover). The normalized WTI and Brent close prices were considered as dependent and the normalized technical indicators as the predictor variables. Statistical measures like MSE, MAPE and Theil U-statistics were used to check the forecasting performance of the models.

The proposed models (MARSplines-IPSO-BPNN and MARSplines-FPA-BPNN) are compared with the benchmark models (IPSO-BPNN, FPA-BPNN, MARSplines-BPNN, MARSplines and BPNN) and the results obtained indicate that the MARSplines-FPA-BPNN group of models outperforms the other models. The forecasting performance of MARSplines-FPA-BPNN is shown to be quite promising than other competitive models, namely IPSO-BPNN and FPA-BPNN developed by Xiao et al (2014) and Chiroma et al (2016), respectively. Next to MARSplines-FPA-BPNN, MARSplines-IPSO-BPNN outperforms the compared models. Since the proposed models reduce the input variables using MARSplines, it is easy for the optimization algorithms to optimize the BPNN parameters to obtain the final result. This helps the proposed model to illustrate lowest errors and a better forecasting accuracy when compared to other models. Among the proposed models, it has been identified that the properties of the FPA optimization methods work better than IPSO. Statistically, to ensure the obtained outcome, the DM test was also performed. To confirm the robustness of the model, the dataset was divided into 60:40, 70:30, 80:20 and 90:10; the new findings confirmed that the MARSplines-FPA-BPNN was robust to forecast WTI and Brent crude oil futures prices with lowest MSE, MAPE and Theil U statistic values.

Overall, the results provide solid support to the proposed MARSplines-FPA-BPNN model, which demonstrates its effectiveness among the models considered. On the whole, this study emphasized the merits of using dimension reduction methods and optimization algorithms to get better results in forecasting crude oil prices. This study coincides with Jammazi and Aloui (2012), Safari and Davallou (2018) and Li et al (2019) who was supporting the hybrid models over single models in forecasting crude oil prices.

Theoretically, this study contributes to the development of forecasting models through the dimension reduction method to achieve lower prediction errors. The application of a nonlinear regression model, i.e. MARSplines as a dimension reduction tool, helps to handle the noisy data, determining the important variables by adjusting its splines which helps the hybrid models to make lower prediction errors. This is the first time that the nature inspired optimization algorithms (IPSO and FPA) are mixed with MARSplines and BPNN in forecasting. Like other nonlinear forecasting models, the newly proposed models are capable of handling chaotic, irregular and noisy data, so the proposed models can be used in financial market data, as well as in other time series forecasting studies.

The results of this research are expected to help traders, investors, firms and policy makers to make swift decisions. Besides crude oil, the proposed model can be used for forecasting other energy commodities, stocks, or economic indicators. By using dimension reduction methods, decision makers can learn to comprehend the significant variable which considerably influences the path of futures crude oil prices. It also helps to save the running time of the model that helps decision makers to reach decisions in time, which may result in reducing the expected costs and losses associated with the forecasting process. This forecasting model can be also used by hedge fund managers and asset managers to build long-short strategies to make sufficient profits. Moreover, OPEC and OECD countries can also use the recommended models to frame policies in relevance to global crude oil prices. Finally, this study contributes to the development of more reliable forecasting models and anticipates that the findings will be valuable in supporting policymakers and aiding their decision making process in international energy markets.

**References**

Ahmed, R. A., & Shabri, A. B. (2014). Daily crude oil price forecasting model using arima, generalized autoregressive conditional heteroscedastic and support vector machines. American Journal of Applied Sciences, 11(3), 425.

Alfi, A., & Modares, H. (2011). System identification and control using adaptive particle swarm optimization. Applied Mathematical Modelling, 35(3), 1210-1221.

Atsalakis, G. S., & Valavanis, K. P. (2009). Surveying stock market forecasting techniques–Part II: Soft computing methods. Expert Systems with Applications, 36(3), 5932-5941.

Azadeh, A., Asadzadeh, S. M., Mirseraji, G. H., & Saberi, M. (2015). An emotional learning-neuro-fuzzy inference approach for optimum training and forecasting of gas consumption estimation models with cognitive data. Technological Forecasting and Social Change, 91, 47-63.

Berry, M. J., & Linoff, G. (1997). Data Mining Techniques, John Wiley & Sons. Newyork

Chai, J., Lu, Q., Hu, Y., Wang, S., Lai, K. K., & Liu, H. (2018a). Analysis and Bayes statistical probability inference of crude oil price change point. Technological Forecasting and Social Change, 126, 271-283.

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

Chatziantoniou, I., Degiannakis, S., & Filis, G. (2019). Futures-based forecasts: How useful are they for oil price volatility forecasting?. Energy Economics, 81, 639-649.

Chen, H., Liao, H., Tang, B. J., & Wei, Y. M. (2016). Impacts of OPEC's political risk on the international crude oil prices: An empirical analysis based on the SVAR models. Energy Economics, 57, 42-49.

Chen, Y., & Hao, Y. (2017). A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction. Expert Systems with Applications, 80, 340-355.

Chen, Y. J., Lin, J. A., Chen, Y. M., & Wu, J. H. (2019). Financial Forecasting With Multivariate Adaptive Regression Splines and Queen Genetic Algorithm-Support Vector Regression. IEEE Access, 7, 112931-112938.

Chinnathambi, R. A., Campion, M., Nair, A. S., & Ranganathan, P. (2018, October). Investigation of Price-Feature Selection Algorithms for the Day-Ahead Electricity Markets. In 2018 IEEE Electrical Power and Energy Conference (EPEC) (pp. 1-6). IEEE.

Chiroma, H., Khan, A., Abubakar, A. I., Saadi, Y., Hamza, M. F., Shuib, L., & Herawan, T. (2016). A new approach for forecasting OPEC petroleum consumption based on neural network train by using flower pollination algorithm. Applied Soft Computing, 48, 50-58.

Dai, W., Shao, Y. E., & Lu, C. J. (2013). Incorporating feature selection method into support vector regression for stock index forecasting. Neural Computing and Applications, 23(6), 1551-1561.

de Albuquerquemello, V. P., de Medeiros, R. K., da Nóbrega Besarria, C., & Maia, S. F. (2018). Forecasting crude oil price: Does exist an optimal econometric model?. Energy, 155, 578-591.

Debnath, K. B., & Mourshed, M. (2018). Forecasting methods in energy planning models. Renewable and Sustainable Energy Reviews, 88, 297-325.

Degiannakis, S., & Filis, G. (2018). Forecasting oil prices: High-frequency financial data are indeed useful. Energy Economics, 76, 388-402.

Diebold, F. X., & Mariano, R. S. (1995). Com paring Predictive Accuracy. Journal of Business and Economic Statistics, 13(3), 253-263.

Ding, Y. (2018). A novel decompose-ensemble methodology with AIC-ANN approach for crude oil forecasting. Energy, 154, 328-336.

Ding, Z., Liu, Z., Zhang, Y., & Long, R. (2017). The contagion effect of international crude oil price fluctuations on Chinese stock market investor sentiment. Applied energy, 187, 27-36.

Drachal, K. (2016). Forecasting spot oil price in a dynamic model averaging framework—Have the determinants changed over time? Energy Economics, 60, 35-46.

EIA, E.I.A., 2019. Annual Energy Outlook 2019 with projections to 2050. Technical Report. US Department of Energy: Washington, DC, pp 32. <https://www.eia.gov/outlooks/aeo/pdf/aeo2019.pdf> (accessed on 12 November, 2019).

Engelbrecht A. P., Computational Swarm Intelligence. Chichester, GB: Wiley, 2005.

Fan, L., Pan, S., Li, Z., & Li, H. (2016). An ICA-based support vector regression scheme for forecasting crude oil prices. Technological Forecasting and Social Change, 112, 245-253.

Friedman, J. H. (1991). Multivariate adaptive regression splines. The annals of statistics, 19(1), 1-67.

Guo, X., Li, D., & Zhang, A. (2012). Improved support vector machine oil price forecast model based on genetic algorithm optimization parameters. Aasri Procedia, 1, 525-530.

Gupta, R., & Wohar, M. (2017). Forecasting oil and stock returns with a Qual VAR using over 150 years off data. Energy Economics, 62, 181-186.

Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. Elsevier. 3rd Edition, San Francisco.

Hamid, M. H., & Shabri, A. (2017, May). Wavelet regression model in forecasting crude oil price. In AIP Conference Proceedings (Vol. 1842, No. 1, p. 030019). AIP Publishing.

He, K., Tso, G. K., Zou, Y., & Liu, J. (2018). Crude oil risk forecasting: New evidence from multiscale analysis approach. Energy Economics, 76, 574-583.

Herrera, G. P., Constantino, M., Tabak, B. M., Pistori, H., Su, J. J., & Naranpanawa, A. (2019). Long-term forecast of energy commodities price using machine learning. Energy, 179, 214-221.

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

Hu, J., Du, L., & Hei, Y. (2018, August). Forecasting Crude Oil Price Based on EMD-Wavelet-GARCH Model. In International Conference on Management Science and Engineering Management (pp. 543-554). Springer, Cham.

Isham, M. F., Leong, M. S., Lim, M. H., & Zakaria, M. K. (2019). A Review on Variational Mode Decomposition for Rotating Machinery Diagnosis. In MATEC Web of Conferences (Vol. 255, p. 02017). EDP Sciences.

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

Jianwei, E., Bao, Y., & Ye, J. (2017). Crude oil price analysis and forecasting based on variational mode decomposition and independent component analysis. Physica A: Statistical Mechanics and its Applications, 484, 412-427.

Kao, L. J., Chiu, C. C., Lu, C. J., & Chang, C. H. (2013). A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting. Decision Support Systems, 54(3), 1228-1244.

Lahmiri, S. (2015). Comparing variational and empirical mode decomposition in forecasting day-ahead energy prices. IEEE Systems Journal, 11(3), 1907-1910.

Li, J., Zhu, S., & Wu, Q. (2019). Monthly crude oil spot price forecasting using variational mode decomposition. Energy Economics, 83, 240-253.

Ma, F., Zhang, Y., Huang, D., & Lai, X. (2018). Forecasting oil futures price volatility: New evidence from realized range-based volatility. Energy Economics, 75, 400-409.

Moshiri, S., & Foroutan, F. (2006). Forecasting nonlinear crude oil futures prices. The Energy Journal, 27(4), 81-95.

Movagharnejad, K., Mehdizadeh, B., Banihashemi, M., & Kordkheili, M. S. (2011). Forecasting the differences between various commercial oil prices in the Persian Gulf region by neural network. Energy, 36(7), 3979-3984.

Nademi, A., & Nademi, Y. (2018). Forecasting crude oil prices by a semiparametric Markov switching model: OPEC, WTI, and Brent cases. Energy Economics, 74, 757-766.

Naser, H. (2016). Estimating and forecasting the real prices of crude oil: A data rich model using a dynamic model averaging (DMA) approach. Energy Economics, 56, 75-87.

Pan, L., Feng, X., Sang, F., Li, L., Leng, M., & Chen, X. (2019). An improved back propagation neural network based on complexity decomposition technology and modified flower pollination optimization for short-term load forecasting. Neural Computing and Applications, 31(7), 2679-2697.

Papadimitriou, T., Gogas, P., & Stathakis, E. (2014). Forecasting energy markets using support vector machines. Energy Economics, 44, 135-142.

Polanco-Martínez, J., & Abadie, L. (2016). Analyzing crude oil spot price dynamics versus long term future prices: A wavelet analysis approach. Energies, 9(12), 1089. (https://doi.org/10.3390/en9121089 - 2 December, 2019)

Roy, S. S., Roy, R., & Balas, V. E. (2018). Estimating heating load in buildings using multivariate adaptive regression splines, extreme learning machine, a hybrid model of MARS and ELM. Renewable and Sustainable Energy Reviews, 82, 4256-4268.

Sad, C. M., da Silva, M., dos Santos, F. D., Pereira, L. B., Corona, R. R., Silva, S. R., ... & Lacerda Jr, V. (2019). Multivariate data analysis applied in the evaluation of crude oil blends. Fuel, 239, 421-428.

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

Singh, V. K., Kumar, P., & Nishant, S. (2019). Feedback spillover dynamics of crude oil and global assets indicators: A system-wide network perspective. Energy Economics, 80, 321-335.

Son, Y., Noh, D. J., & Lee, J. (2012). Forecasting trends of high-frequency KOSPI200 index data using learning classifiers. Expert Systems with Applications, 39(14), 11607-11615.

Tang, L., Yu, L., Wang, S., Li, J., & Wang, S. (2012). A novel hybrid ensemble learning paradigm for nuclear energy consumption forecasting. Applied Energy, 93, 432-443.

Tang, L., Zhang, C., Li, L., & Wang, S. (2020). A multi-scale method for forecasting oil price with multi-factor search engine data. Applied Energy, 257, 114033. (https://doi.org/10.1016/j.apenergy.2019.114033).

Uddin, G. S., Gençay, R., Bekiros, S., & Sahamkhadam, M. (2019). Enhancing the predictability of crude oil markets with hybrid wavelet approaches. Economics Letters, 182, 50-54.

Wang, P. C., & Shoup, T. E. (2011). A poly-hybrid PSO optimization method with intelligent parameter adjustment. Advances in Engineering Software, 42(8), 555-565.

Wang, Y., Ma, F., Wei, Y., & Wu, C. (2016). Forecasting realized volatility in a changing world: A dynamic model averaging approach. Journal of Banking & Finance, 64, 136-149.

Wang, Y., & Wu, C. (2012). Forecasting energy market volatility using GARCH models: Can multivariate models beat univariate models?. Energy Economics, 34(6), 2167-2181.

Wang, T., & Yang, J. (2010). Nonlinearity and intraday efficiency tests on energy futures markets. Energy Economics, 32(2), 496-503.

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.

Xian, L., He, K., Wang, C., & Lai, K. K. (2020). Factor analysis of financial time series using EEMD-ICA based approach. Sustainable Futures, 2, 100003. (https://doi.org/10.1016/j.sftr.2019.100003)

Xiang, Y., & Zhuang, X. H. (2013). Application of ARIMA model in short-term prediction of international crude oil price. In Advanced Materials Research (Vol. 798, pp. 979-982). Trans Tech Publications.

Xiao, Y., Xiao, J., Lu, F., & Wang, S. (2014). Ensemble ANNs-PSO-GA approach for day-ahead stock e-exchange prices forecasting. International Journal of Computational Intelligence Systems, 7(2), 272-290.

Xie, W., Yu, L., Xu, S., & Wang, S. (2006, May). A new method for crude oil price forecasting based on support vector machines. In International Conference on Computational Science (pp. 444-451). Springer, Berlin, Heidelberg.

Yang, X. S. (2012, September). Flower pollination algorithm for global optimization. In International conference on unconventional computing and natural computation (pp. 240-249). Springer, Berlin, Heidelberg.

Yu, L., Dai, W., & Tang, L. (2016). A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting. Engineering Applications of Artificial Intelligence, 47, 110-121.

Zhang, Y., Ma, F., & Wei, Y. (2019). Out-of-sample prediction of the oil futures market volatility: A comparison of new and traditional combination approaches. Energy Economics, 81, 1109-1120.

Zhang, G. P., Patuwo, B. E., & Hu, M. Y. (2001). A simulation study of artificial neural networks for nonlinear time-series forecasting. Computers and Operations Research, 28(4), 381-396.

Zhang, J. L., Zhang, Y. J., & Zhang, L. (2015). A novel hybrid method for crude oil price forecasting. Energy Economics, 49, 649-659.

Zhang, W., Qu, Z., Zhang, K., Mao, W., Ma, Y., & Fan, X. (2017). A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting. Energy Conversion and Management, 136, 439-451.

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, Berlin, Heidelberg.

Zhao, L., Cheng, L., Wan, Y., Zhang, H., & Zhang, Z. (2015). A VAR-SVM model for crude oil price forecasting. International Journal of Global Energy Issues, 38(1-3), 126-144.

Zhao, Y., Li, J., & Yu, L. (2017). A deep learning ensemble approach for crude oil price forecasting. Energy Economics, 66, 9-16.

Zou, Y., Yu, L., Tso, G. K., & He, K. (2020). Risk forecasting in the crude oil market: A multiscale Convolutional Neural Network approach. Physica A: Statistical Mechanics and its Applications, 541, 123360. (https://doi.org/10.1016/j.physa.2019.123360).