

## A fireworks algorithm based Pi-Sigma neural network (FWA-PSNN) for modelling and forecasting chaotic crude oil price time series

Sarat Chandra Nayak<sup>1,\*</sup>

<sup>1</sup>Department of Computer science and Engineering,  
CMR College of Engineering & Technology, Hyderabad – 501401, India

### Abstract

Capturing the complex correlations among the data on the crude oil price time series is challenging, hence accurate prediction of it is difficult. Contrast to multilayer artificial neural network, Pi-Sigma neural network (PSNN) is characterized with stronger approximation ability, fast learning and higher fault tolerance. Fireworks algorithm (FWA) is a new metaheuristic motivated from the occurrence of fireworks explosion, characterized with fast convergence capability, parallelism and finds global optima. This article intends to achieve the synergetic effect of both on hybridizing them. It uses FWA for optimization of weight and bias of PSNN (FWA-PSNN) and overcomes the limitations of gradient based training. It has single hidden layer structure and trainable weight, hence fast and robust. The FWA-PSNN is evaluated on prediction of crude oil price time series. Extensive simulation results, performance analysis, and statistical significance test suggested the suitability of FWA-PSNN.

**Keywords:** crude oil price forecasting, time series forecasting, higher order neural network, pi-sigma neural network, fireworks algorithm, hybrid model, genetic algorithm.

Received on 16 July 2019, accepted on 04 January 2020, published on 24 January 2020

Copyright © 2020 Sarat Chandra Nayak *et al.*, licensed to EAI. This is an open access article distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/3.0/>), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi: 10.4108/eai.13-7-2018.162803

### 1. Introduction

Crude oil price plays a major function in the economical growth of a country. During last few decades the crude oil price has been facing arbitrary changes in its movement. It depends upon multiple socioeconomically as well as political factors. The random fluctuations and high nonlinearities associated with the crude oil price time series make its forecasting difficult. Nominal change that occurs to the crude oil price gives impact to the petroleum price, crude oil products and the global economy. The volatility in the price series is generally affected by factors like population, demand and supply, political climates, international relationship [1]. Thus, a correct and proficient prediction instrument is crucial for crude oil price forecast. Advances

in computational intelligence techniques including artificial neural networks (ANNs) are used as better alternative to this domain [2]. ANN-based models are applied successfully to forecast the crude oil price [3 - 6].

ANN has the resemblance with the judgment power of human intellect and mimicking it [7 - 9]. The ANN can replicate the process of human actions to crack nonlinear problems. This characteristic made it popular to be used in resolving complex systems. ANNs are measured to be valuable modelling process to map input-output relationship where the data contains both regularities and exceptions such as financial time series. These facts of ANN attract

\*Corresponding author. Email: saratnayak234@gmail.com

financial stake holders to forecast financial time series with ANN based models. Dealing with uncertainty and nonlinearity allied with financial data with ANN-based forecasting method primarily involves identification of patterns in the data and use of such patterns to forecast.

Multilayer neural networks have certain issues like adding more computational complexity, leads to lower learning rate and accuracy. They are characterized with structural complexity, computational overhead and have black box visualization. In contrast to them, higher order neural networks (HONN) have certain characteristics such as stronger approximation capability, speedy learning properties, larger storage and higher fault tolerance capability. Also, they have a prevailing mapping single layer tunable weight set that overcomes the black box nature [10]. Use of higher order terms can amplify the information capacity and help solving intricate nonlinear task even with a petite network without compromising the convergence capability. In the other hand, with the swell in network order, there is intensification in tunable weights which may lead to additional computation cost. To alleviate these drawbacks, another kind of HONN, i.e. PSNN is introduced in [11] which use a lesser amount of connection weights. Few applications of HONN in predicting financial time series are proposed by authors in [12 - 16].

The adjustment of neuron weight and bias of ANN is the key factor of ANN training and requires frequent human interventions. The performance of ANN solely depends upon the adjustment of weight and bias vectors. To circumvent the limitations of gradient descent based ANN training, large number of nature and bio-inspired optimization techniques are proposed and applied [17]. Evolutionary computing techniques are based on the behaviour of nature. Usually, these algorithms are inspired from theory of evolution and termed as evolutionary optimization algorithms or metaheuristic. The ideas of imitating concepts from nature have great potential in developing algorithms to solve engineering problems. In recent past, applications of these techniques have achieved popularity in wide area of engineering, computer science, medicine, economics, finance, social networks and so on. More often, performance of these techniques depends on proper adjustment of several algorithm specific control parameters. There is no distinct technique performing well on all problems. Evolutionary learning techniques such as GA [18], PSO [19], DE [20], Ant colony optimization [21], FWA [22] etc. are more proficient than gradient descent based methods in searching the optimal solution.

FWA is a newly suggested metaheuristic which simulates the occurrence of fireworks explosion at night [22]. Like other nature inspired optimization it is also a population-based evolutionary algorithm. It tries to find the best fit solution in the search space through the explosion of fireworks. Several applications of FWA are found in the literature for solving real data mining problems. Mean time, there are some improved and enhanced version of FWA proposed and their superiority have been established. However its application toward crude oil price time series is limited. Analysis and improvement of fireworks algorithm is

suggested by authors in [23]. A cooperative framework for FWA is suggested in [24]. An enhanced version of FWA is proposed and evaluated on standard benchmark functions [25]. Adaptive FWA and dynamic search in FWA are also introduced by authors in [26 - 27].

The intention of current work is to investigate the potential of FWA based PSNN (FWA-PSNN) hybrid model on forecasting the crude oil price time series. The hybrid model employed FWA to optimize the weight and bias vector of PSNN. Each location (individual) of FWA can be viewed as a possible weight and bias vector for a PSNN. The FWA applies local as well as global search techniques in the form of fireworks explosion to explore the best possible weight and bias vector in the potential search space. The proposed hybrid FWA-PSNN model is evaluated on forecasting crude oil price time series. The performance of the model is compared with that of four other models such as PSO-PSNN, GA-PSNN, gradient descent based PSNN (GD-PSNN) and MLP trained similarly. The major advantages of the proposed FWA-PSNN based forecasting model are as follows.

- Use of FWA which posses the characteristics of fast convergence, parallelism, able to find the global optima, and necessitates lesser number of tunable parameters
- The base ANN, i.e. PSNN possesses single hidden layer architecture and single trainable weight vector, hence structurally simple
- The resulting hybrid model, i.e. FWA-PSNN possesses both the characteristics of FWA and PSNN, hence more fast and robust
- Exhaustive experimentation on real crude oil price time series and comparative study
- Statistical significance of the proposed model

The article is structured in five major sections. Section 1 gives a short description about the crude oil price forecasting, ANN and FWA. An overview of FWA is discussed in Section 2. The FWA-PSNN based forecasting is presented in Section 3. Section 4 summarizes the experimental outcome and analysis. Section 5 gives the concluding remark followed by a relevant list of references.

## 2. Related studies

Crude oil is a precious commodity for economical and industrial development of a country as well as for global economy. Several related and recent studies have been carried out in the domain of crude oil forecasting which includes both linear as well as nonlinear models. The stochastic and nonlinearity behaviour of crude oil prices has been demonstrated in the work [28]. The authors compared statistical and ANN based models and concluded that ANN models produces better forecast. Traditional approaches including linear and parametric models are found not

efficient in capturing the dynamics of financial time series [29, 30].

The first category of forecasting model includes statistical models like autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), generalized autoregressive conditional heteroscedasticity (GARCH) and their hybridization. Many authors proposed this approach for financial and oil market forecasting. Gavriilidis et al. studied the effect of inclusion of oil price shocks from different origin in a set of GARCH-X models [31]. Herrera et al. evaluated the relative performance of various econometrics models using high frequency intra-day volatility data [32]. A hybrid metaheuristic approach based on ANN, ARIMA, and support vector machines (SVM) was proposed by Naderi et al. [33]. A nonlinear metabolic grey model corrected by ARIMA (NMGm-ARIMA) for enhanced accuracy in China's foreign oil dependence has been demonstrated by Wang et al. [34].

The second category includes artificial intelligence, machine learning, soft computing, and hybridization of them. Huang and Wu [35] developed deep multiple kernel learning approach for energy commodity price forecasting. Hybrid linear and non-linear techniques for energy demand forecasting in China and India are proposed by authors in [36]. To study the nonlinear complex nature of crude oil price movement Chen et al. [37] proposed a deep learning based model and achieved improved forecasting accuracy. A GA and fast ensemble empirical mode decomposition (GA-FEEMD) for forecasting crude oil price time series data has been proposed by authors in [38]. They compared the proposed model with ARIMA, and ANN found improvement. A PSO optimized gray Markov model was suggested by Hu et al. [39] and found smaller error and better accuracy compared to others. A hybrid PSO and radial basis functional neural network is proposed in [40] that outperformed other models. A forecasting model based on Grey Wolf Optimizer is proposed for short term forecasting of energy commodity [41]. The model has been compared with model trained with artificial bee colony optimization and differential evolution and found more competitive. A combination of kernel principal component analysis and DE optimized support vector machine is proposed for crude oil price prediction [42]. A survey on computational methods for crude oil price forecasting has been done by authors in [43]. During last decades several nature-inspired metaheuristic algorithms as well as machine learning techniques have shown their capability in application areas like healthcare [44, 46], medical image processing [47], query optimization [48], crude oil price forecasting [49], optimal feature selection and classification [50]. These metaheuristics are population based and able to land at global optima with a reasonable computational cost. FWA is a recent optimization technique which tries to find the best fit solution in the search space through the explosion of fireworks. It has characteristics like fast convergence and reaching global optima. Analysis and improvement of FWA is discussed by authors in [23]. A cooperative framework for fireworks algorithm is proposed

by authors in [24]. Enhanced version of fireworks algorithm is proposed and evaluated on standard benchmark functions by the authors in [25]. Adaptive FWA and dynamic search in FWA are also introduced by authors in [26 - 27]. An icing forecasting system based on FWA and weighted least square support vector machine is proposed in [44] and the result show that the model obtained high prediction accuracy. FWA is also applied for portfolio optimization and found better to other swarm intelligence methods [45]. FWA and adjusted FWA applications for multilevel image threshold and retinal image registration are found in [46, 47].

### 3. Methods

This section describes about the metaheuristic, i.e. FWA used in this article for optimization of weight and bias vector of the PSNN followed by the proposed FWA-PSNN based forecasting. The mathematical detail of FWA is beyond the range of this article. The base articles are cited wherever required.

#### 3.1. FWA metaheuristic

Fireworks algorithm is a recently proposed optimization technique which mimics the explosion process of fireworks [22]. It tries to select certain amount of locations in a search space for explosion of fireworks to produce set of sparks. Locations with qualitative fireworks are considered for the subsequent generation. The procedure continues iteratively up to a desired optimum or reaching the stopping condition. The process mainly comprises three steps: setting off N fireworks at N selected locations, obtaining the locations of sparks after explosion and evaluating them, stop on reaching optimal location or select N other locations for the next generation of explosion. An explosion of fireworks can be visualized as a search process in the local space. According to the basic FWA, for each firework  $x_i$ , the amplitude of explosion ( $A_i$ ) and number of sparks ( $s_i$ ) are defined as follows:

$$A_i = \hat{A} \cdot \frac{f(x_i) - f_{\min} + \epsilon}{\sum_{j=1}^p (f(x_j) - f_{\min}) + \epsilon} \quad (1)$$

$$s_i = \frac{m \cdot f_{\max} - f(x_i) + \epsilon}{\sum_{j=1}^p (f_{\max} - f(x_j)) + \epsilon} \quad (2)$$

where  $\hat{A}$  is the maximum amplitude of explosion.  $f_{\max}$  is the maximum of objective function value and  $f_{\min}$  is the minimum objective function values among the p fireworks. The m is a controlling parameter for total number of sparks generated by a firework and  $\epsilon$  is a constant used to avoid zero division error. Bounds are imposed on  $s_i$  to overcome the devastating effects of marvelous fireworks as follows:

$$s_i = \begin{cases} s_{\max}, & \text{if } s_i > s_{\max} \\ s_{\min}, & \text{if } s_i < s_{\min} \\ s_i, & \text{otherwise} \end{cases} \quad (3)$$

The location of each spark  $x_j$  generated by  $x_i$  is calculated by setting z directions randomly and for each dimension k

setting the component  $x_j^k$  based on  $x_i^k$ , where  $1 \leq j \leq s_i, 1 \leq k \leq z$ .

The setting of  $x_j^k$  can be done in two ways as follows:

- For most sparks, a displacement is added to  $x_j^k$  as:

$$x_j^k = x_i^k + A_i \cdot \text{rand}(-1,1) \quad (4)$$

- For maintaining the diversity, for few specific sparks, an explosion coefficient based on Gaussian distribution is applied to  $x_j^k$  as:

$$x_j^k = x_i^k \cdot \text{Gaussian}(1,1) \quad (5)$$

A new location is mapped to the potential space when it falls out of the search space as:

$$x_j^k = x_{\min}^k + |x_j^k| \% (x_{\max}^k - x_{\min}^k), \quad (6)$$

where % is the modulo operator.

The next step is selection of another N location for the fireworks explosion. This step always keeps the current best location  $x^*$  for the next generation. Remaining N-1 locations are considered on the basis of their distance to other locations. The distance between a location  $x_i$  and other locations (K) can be calculated as the sum of Euclidean distance between them and as follows:

$$\text{Distance}(x_i) = \sum_{j \in K} \|x_i - x_j\|, \quad (7)$$

A location is  $x_i$  selected for the next generation based on a probability value as follows:

$$\text{prob}(x_i) = \frac{\text{Distance}(x_i)}{\sum_{j \in K} \text{Distance}(x_j)} \quad (8)$$

Since the sparks suffer from the power of explosion, they move along z directions simultaneously. This makes FWA to achieve faster convergence. Also, it avoids the premature convergence with the two types of spark generation methods and specific location selection method [22]. The advantages of FWA over standard PSO and its improved variants are demonstrated in the research work [22].

### 3.2. FWA-PSNN based forecasting

The PSNN belongs to the class of HONN and having architecture of double layered as shown in Figure 1. The network is feed forward and fully connected [11]. The first layer is a composed of summing units and the second layer is composed of product units. The network input are feed to input layer consist of summation units and the consequent outputs are feed to the output layer consist of a product unit. The synaptic input-summing layer weights are trainable where as the connection summing-product layer weights are non-tunable and set to a value of unity. Since only one tunable weight set is required, the network achieved a significant decline in training time. Linear activation functions are used at summing layer units and nonlinear activation functions are used by the product units. The order of the network increases by one in addition of each extra

summing unit. The product unit offers higher order capabilities to the network by intensifying the input space from lower to higher dimension. In this way, the network avoids being suffered from exponential swell in weights and simultaneously offering better nonlinear separable capacity to the network.

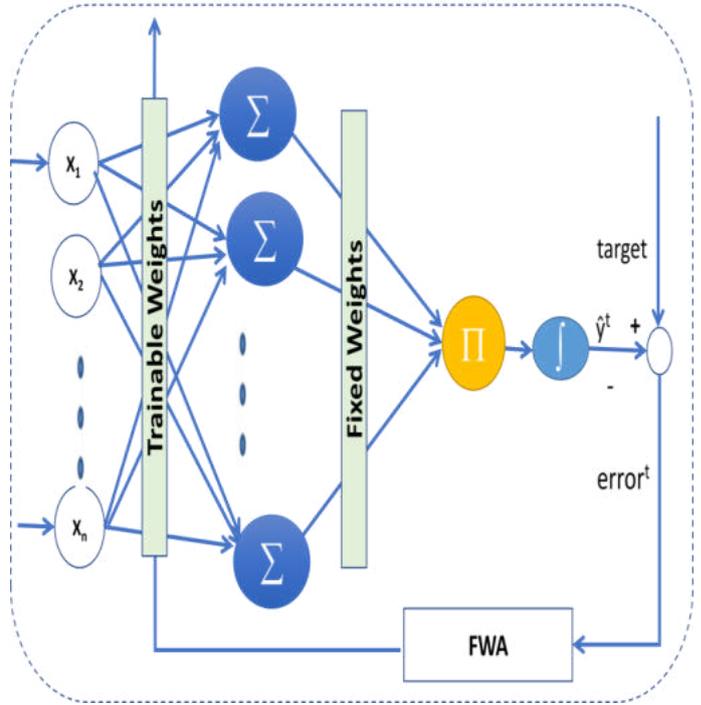


Figure 1. FWA-PSNN based forecasting

The output of FWA-PSNN is calculated as follows. Let the output at  $j^{\text{th}}$  summing component of hidden layer is computed by summing up the products from each input  $x_i$  and corresponding weight  $w_{ij}$  as in Eq. 9.

$$y_j = \sum_{i=1}^n w_{ij} * x_i, \quad (9)$$

where  $n$  = volume of input signals fed to the input neurons. The output unit computes the product of outputs from the summing units of hidden layer. It then forwards this output to a nonlinear sigmoid activation as in Eq. 10.

$$\hat{y} = \sigma(\prod_{j=1}^m y_j), \quad (10)$$

Where,  $m$  is the order of the network and symbol  $\sigma$  is the product operator. Now the target ( $y$ ) is supplied to the output neuron, compared with model estimation ( $\hat{y}$ ) and the error value is calculated as follows.

$$\text{error} = |y - \hat{y}| \quad (11)$$

The objective is to minimize the error function value in Eq. 11. Each candidate solution represents one potential weight set for the model. The candidate with least error value is considered as the best candidate solution. Now the weight set is optimized as by FWA algorithm. A location (individual) of FWA can be viewed as a potential weight and bias vector for the PSNN in the search space. At beginning, a set of such location is initialized and for each

such location, two types of explosion are carried out as discussed in Section 2. The exploration as well as exploitation of the search space is achieved by these explosion methods. The locations are then evaluated in terms of error signal generation. The location with lowest error signal is considered as the best location. The selection process is then carried out with inclusion of this best location and remaining locations as described in Section 2. The above process continues till an optimal location found and the search process then terminates. The best location is the optimal weight and bias vector for the PSNN model. The FWA-PSNN based forecasting is presented by Algorithm1. The FWA parameters are chosen as suggested in [22]. The TrainData and TestData are generated from the original time series using a window of fixed size sliding over it. The training and test datasets are normalized using sigmoid normalization method. This is explained in Subsection 5.2.

```

4.3. Present TestData, and best location x* to PSNN
    Estimate and preserve error
    /*the error will be used to measure the performance of the model*/
End While
    
```

Algorithm 1: FWA-PSNN-based forecasting

```

1. Set FWA parameters
2. Select N locations randomly for fireworks
3. Form TrainData and TestData /*using sliding window from original time series*/
4. While (more TestData)
    4.1. Normalization of TrainData and TestData
        /*using sigmoid method*/
        /*Training phase*/
    4.2. While (stopping criteria == false)
        Set off N fireworks at N locations
        For each firework xi
            Calculate number of sparks si using Eq.3
            Obtain locations of si sparks of firework xi using Eq.4.
        End For
        For k = 1: m
            Select a firework xj randomly
            Generate a specific spark using Eq.5.
        End For
        Present TrainData, location xi to PSNN and evaluate
        Select the best location x* and keep it for next generation
        Select remaining N-1 locations randomly based on a probability using Eq.8.
    End While

/*Test phase*/
    
```

## 4. Experimental results and analysis

This section explains about the experimental data, experimental setup, input selection and normalization, performance metrics, experimental results, comparative study and result analysis.

### 4.1. Experimental data

The crude oil prices (Dollars per Barrel) are extracted from US Department of energy: Energy Information Administration web site: <http://www.eia.doe.gov/> during the period April 1983 to July 2019. The crude oil price series are shown by Figure 2 - 5. The information about the dataset and descriptive statistics are summarized in Table 1 and Table 2 respectively. It can be seen from the Figure 2 – 5 that the crude oil price time series are highly fluctuating and chaotic by nature. There are random picks and falls on the time series. All the experiments are carried out in MATLAB-2015 environment, with Intel ® core TM i3 CPU, 2.27 GHz processing and 2.42 GB memory size.

Table 1 Information about four time series

| S.No. | Crude oil price dataset (Dollars/Barrel) | Period                      | No. of data points |
|-------|--|-----------------------------|--------------------|
| 1     | Daily Crude Oil Price                    | Apr. 04 1983 – July 02 2019 | 9105               |
| 2     | Weekly Crude Oil Price                   | Apr. 08 1983 – June 28 2019 | 1891               |
| 3     | Monthly Crude Oil Price                  | Apr. 1983 – May 2019        | 434                |
| 4     | Annual Crude Oil Price                   | 1983 - 2018                 | 36                 |

Table 2. Descriptive statistics from crude oil price datasets

| Crude oil price dataset (Dollars per Barrel) | Statistics |          |         |          |          |          |
|--|------------|----------|---------|----------|----------|----------|
|  | Min.       | Max.     | Mean    | Std dev. | Skewness | Kurtosis |
| Daily Crude Oil Price                        | 10.4200    | 145.2900 | 42.9828 | 28.5640  | 0.9953   | 2.8829   |
| Weekly Crude Oil Price                       | 11.0900    | 142.4600 | 42.8949 | 28.5241  | 1.0011   | 2.8921   |
| Monthly Crude Oil Price                      | 11.3100    | 134.0200 | 42.8508 | 28.5115  | 0.9969   | 2.8629   |
| Annual Crude Oil Price                       | 14.4000    | 99.7500  | 42.6214 | 28.0382  | 0.8772   | 2.3527   |

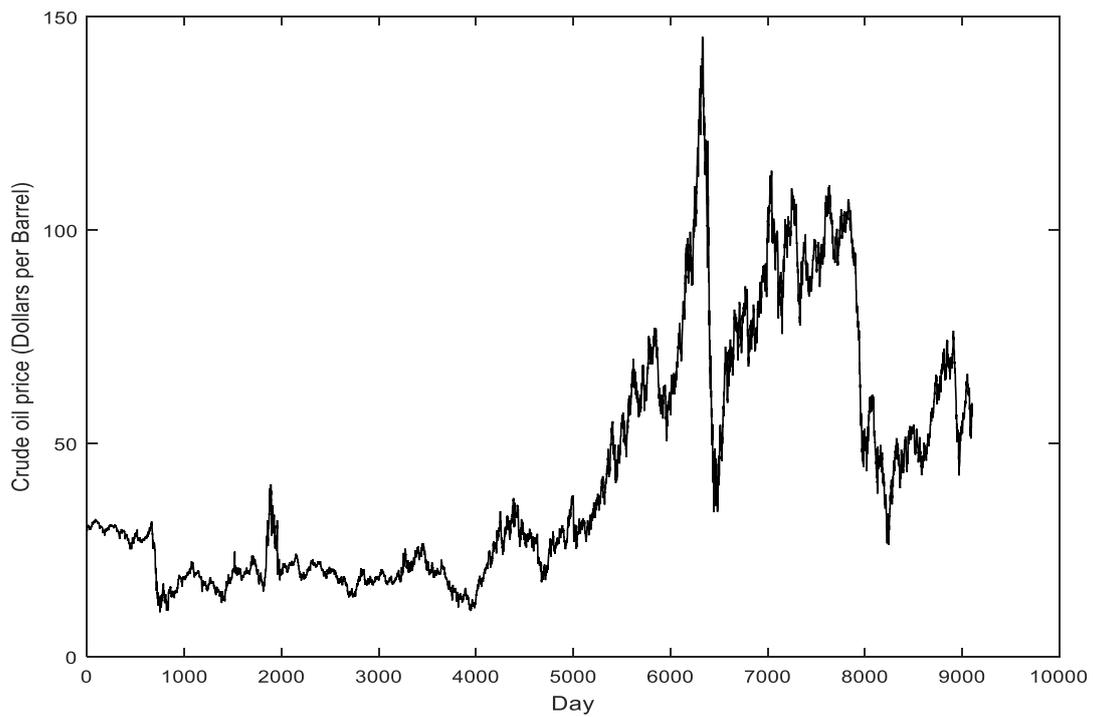


Figure 2. Daily crude oil price data

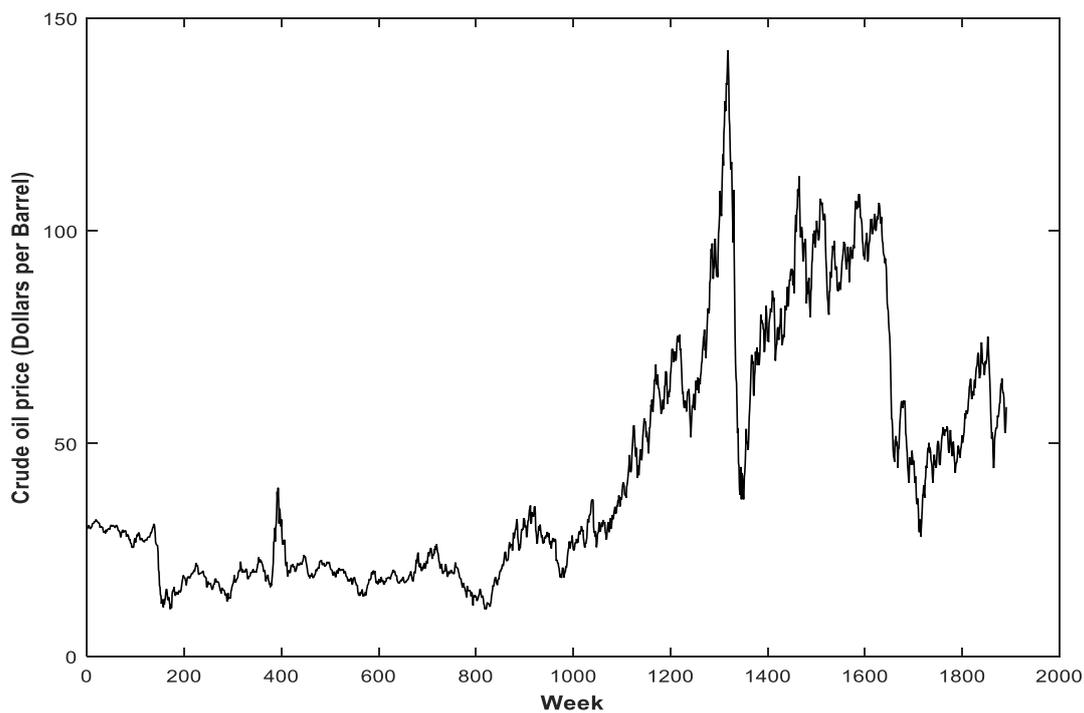


Figure 3. Weekly crude oil price data

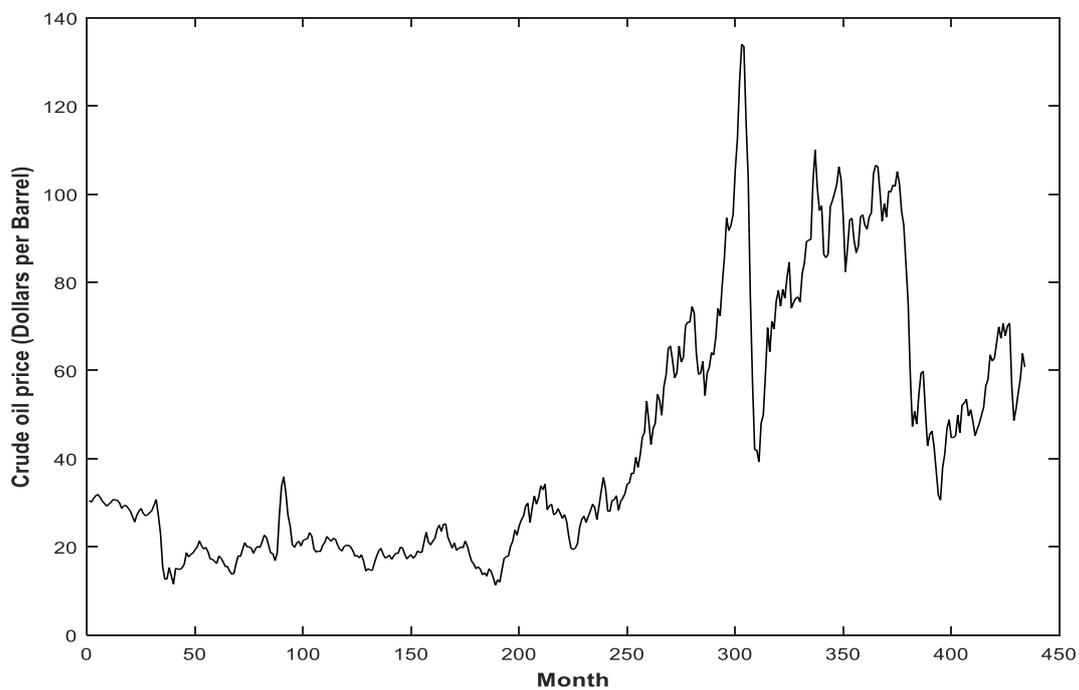


Figure 4. Monthly crude oil price data

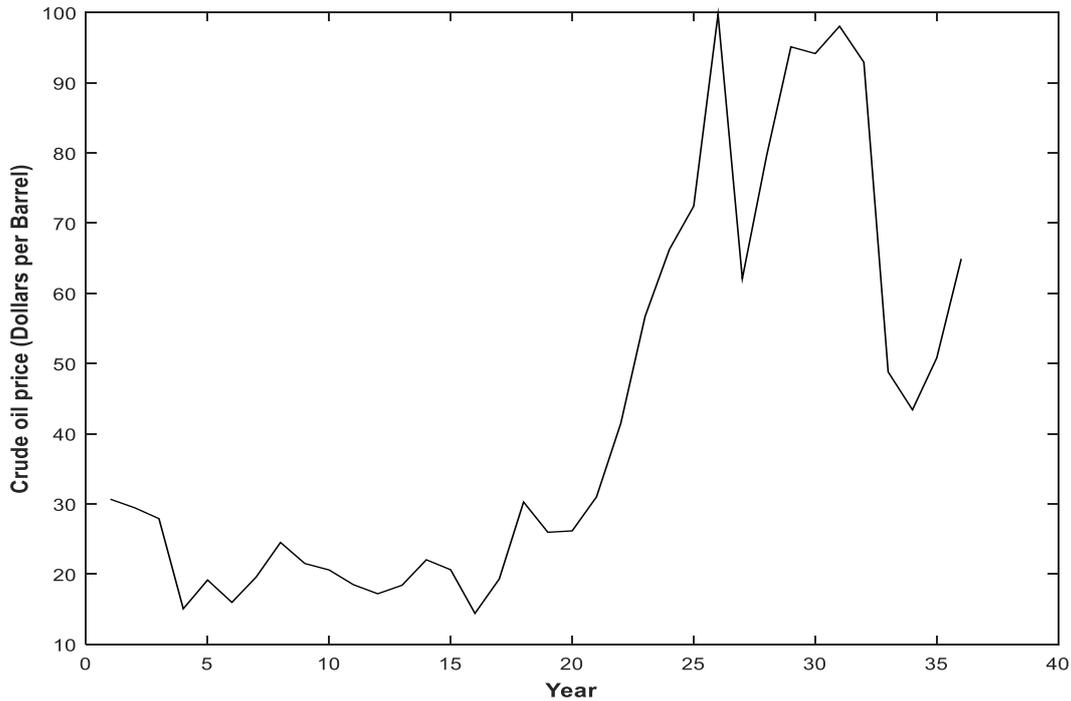


Figure 5. Yearly crude oil price data

#### 4.2. Input selection and normalization

A sliding window of fixed size is used for selecting input for the forecasting models [55]. In this method rather than selecting all data points observed so far, or on some sample, decision is made based only on some recent data points. On each sliding of the window, a new data point is incorporated and the oldest one is discarded. The window moves through whole time series. Selecting the size of window is a matter of experimentation. For an instance, the patterns generated from the time series using a sliding window for one-step-

ahead forecasting is presented below. Here, window size is written as *blen* and *l* is the training length. All the three financial time series are normalized before feeding them to the ANN model [56].

|                |                |          |                       |     |                       |  |
|----------------|----------------|----------|-----------------------|-----|-----------------------|--|
| Train data     |                |          |                       |     | Train target          |  |
| $x(d)$         | $x(d + 1)$     | $\dots$  | $x(d + blen)$         | $:$ | $x(d + blen + 1)$     |  |
| $\vdots$       | $\vdots$       | $\vdots$ | $\vdots$              | $:$ | $\vdots$              |  |
| $x(d + l)$     | $x(d + l + 1)$ | $\dots$  | $x(d + l + blen)$     | $:$ | $x(d + l + blen + 1)$ |  |
| Test data      |                |          |                       |     | Test target           |  |
| $x(d + l + 1)$ | $x(d + l + 2)$ | $\dots$  | $x(d + l + blen + 1)$ | $:$ | $x(d + l + blen + 2)$ |  |

Table 3. Forecasting errors from all models

| Crude oil Price Dataset | Error Statistic    | Forecasting Models |             |             |             |               |
|-------------------------|--------------------|--------------------|-------------|-------------|-------------|---------------|
|                         |                    | MLP                | GD-PSNN     | GA-PSNN     | PSO-PSNN    | FWA-PSNN      |
| Daily Crude Oil Price   | Minimum            | 9.0384e-05         | 8.6386e-05  | 7.6336e-06  | 7.6336e-05  | 7.5590e-08    |
|                         | Maximum            | 0.0502             | 0.0395      | 0.0382      | 0.0376      | 0.0353        |
|                         | Average            | 0.0097             | 0.0095      | 0.0088      | 0.0057      | <b>0.0016</b> |
|                         | Standard Deviation | 0.0087             | 0.0081      | 0.0075      | 0.0075      | 0.0019        |
|                         | Minimum            | 6.9508e-05         | 5.9528e-05  | 3.0527e-06  | 3.0507e-05  | 1.8450e-06    |
| Weekly Crude Oil Price  | Maximum            | 0.0582             | 0.0466      | 0.0432      | 0.0418      | 0.0414        |
|                         | Average            | 0.0128             | 0.0122      | 0.0096      | 0.0094      | <b>0.0034</b> |
|                         | Standard Deviation | 0.0166             | 0.0159      | 0.0055      | 0.0057      | 0.0039        |
|                         | Minimum            | 4.8234e-03         | 4.5224e-05  | 4.1226e-05  | 4.1760e-05  | 1.8193e-05    |
|                         | Maximum            | 0.0867             | 0.0555      | 0.0458      | 0.0452      | 0.0384        |
| Monthly Crude Oil Price | Average            | 0.0180             | 0.0159      | 0.0138      | 0.0098      | <b>0.0078</b> |
|                         | Standard Deviation | 0.0133             | 0.0102      | 0.0082      | 0.0074      | 0.0104        |
|                         | Minimum            | 6.84822e-03        | 5.84879e-04 | 5.74875e-04 | 6.44875e-04 | 4.8149e-04    |
|                         | Maximum            | 0.0473             | 0.0453      | 0.0393      | 0.0399      | 0.0133        |
|                         | Average            | 0.0398             | 0.0293      | 0.0167      | 0.0179      | <b>0.0137</b> |
| Annual Crude Oil Price  | Standard Deviation | 0.0049             | 0.0047      | 0.0042      | 0.0048      | 0.0035        |

### 4.3. Experimental results and discussion

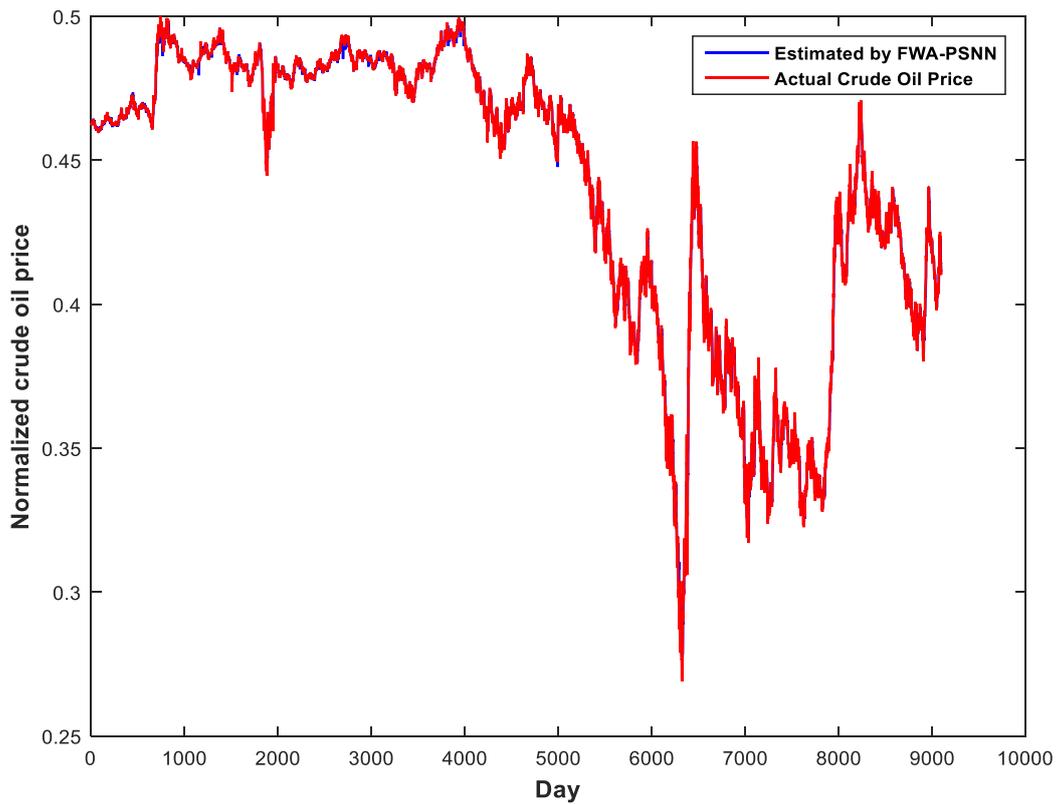
The data after normalization are used to feed the model. To stay away from the stochastic behaviour of ANN based forecast, for each training dataset, we simulated the model for twenty times. The mean error from twenty simulations is considered as the average performance of a model. Adaptive training has been followed as suggested in [13, 15, 55]. During model simulation, different feasible values for the model parameters were experienced and best values are recorded. The acceleration coefficients for PSO are chosen as  $c1 = 2.15$  and  $c2 = 2.15$ . The population size was set to 60 and the selection mechanism followed the global best. The GA used a population of size 70, crossover probability of 0.6 and mutation probability of 0.004. It used elitism method for selection operation (20% better fit individual + 80% binary tournament selection). The learning rate and momentum factor of GD were set to 0.1 and 0.3 respectively. The parameters of FWA were set as suggested in [22], (i.e.  $N = 5$ ,  $m = 50$ ,  $a = 0.04$ ,  $b = 0.8$ ,  $\hat{A} = 40$  and  $\hat{m} = 5$ ). The number of iterations for GA, PSO, FWA, and GD were set to 400, 250, 250, and 600 respectively. The optimal architecture of MLP used is 6-18-1 and that of PSNN is 6-12-1 respectively. The average performance of the models over 20 runs are considered for comparison.

To validate the the proposed model, we developed four other forecasts such as GA-PSNN, PSO-PSNN, GD-PSNN and MLP. All the five models are trained in the similar way.

The error statistics in terms of minimum, maximum, average and standard deviation generated by five models from four datasets are summarized in Table 3. The predicted crude oil prices against actual prices are visualized in Figure 6 – 9. For daily price data all the models are generated lower error values which are acceptable. However, the FWA-PSNN obtains an average error value of 0.0016 which is better than other models. Figure 6 shows the closeness of FWA-PSNN estimated crude oil prices and actual prices. Similar results are observed from weekly and monthly price datasets. In cases of annual crude oil time series the average error generated by FWA-PSNN is bit more compared to other three time series. This might happen due to insufficient training sample, i.e. only 36 data points are available in annual crude oil price time series. The observations from experimental work can be summarized as follows:

- Computational intelligence methods such as ANNs are promising methods for forecasting crude oil prices
- Compared to multilayer ANN, higher order neural network (PSNN here) produced more accurate forecasts

- Nature inspired optimization techniques (GA, PSO and FWA here) are found superior to gradient based method in searching the optimal PSNN model parameters.
- Performance of the proposed model is found better to other hybrid models as well as multilayer artificial neural network model.



**Figure 6.** Actual crude oil prices v/s estimated by FWA-PSNN from daily crude oil price time series

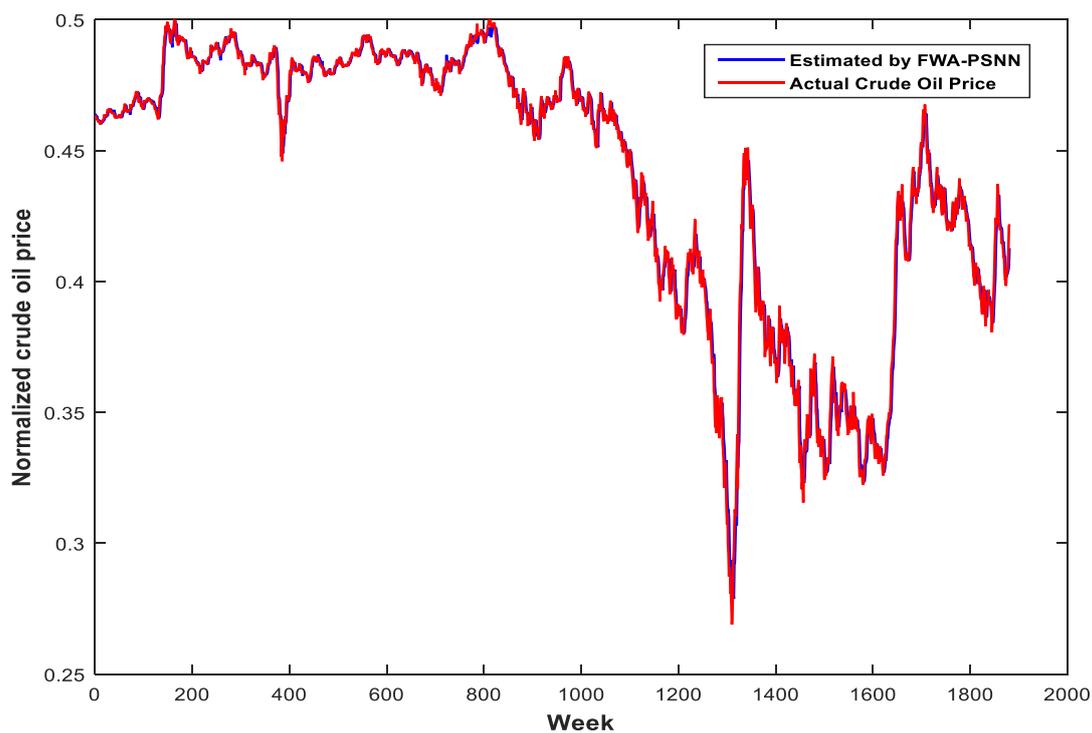


Figure 7. Actual crude oil prices v/s estimated by FWA-PSNN from weekly crude oil price time series

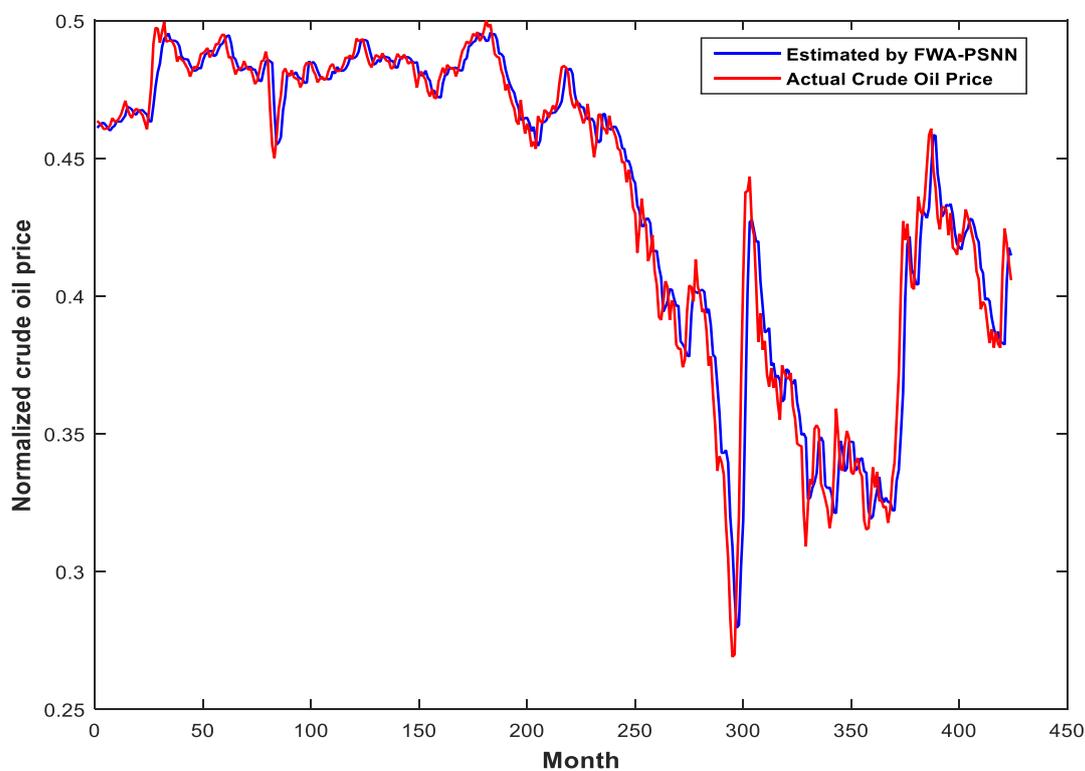


Figure 8. Actual crude oil prices v/s estimated by FWA-PSNN from monthly crude oil price time series

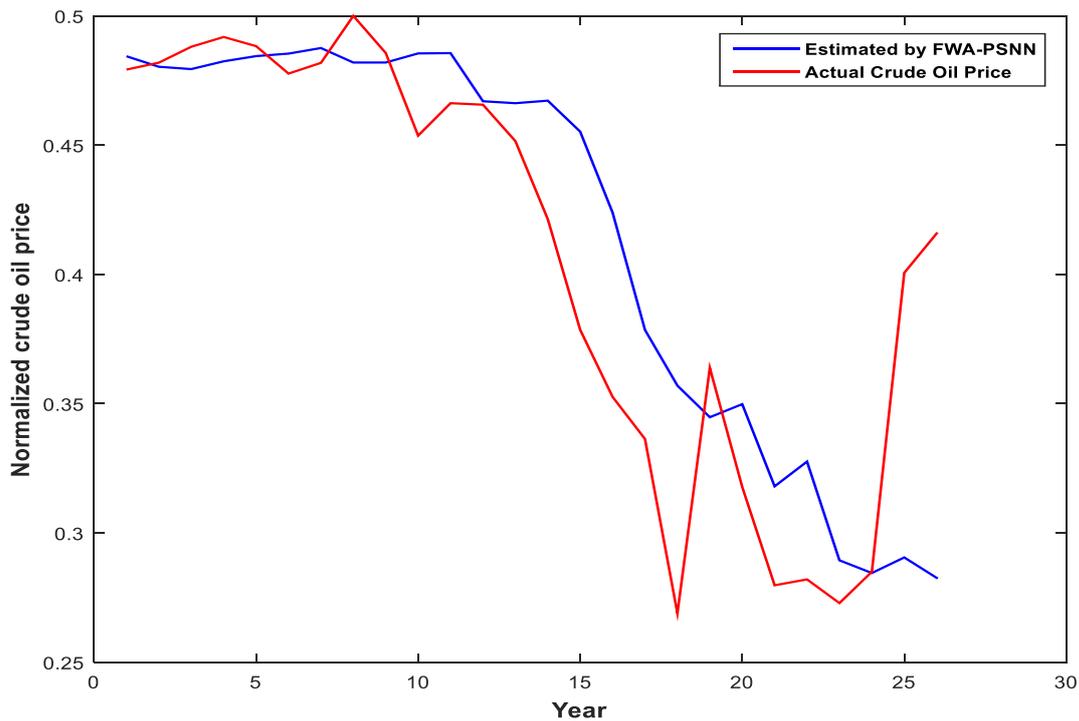


Figure 9. Actual crude oil prices v/s estimated by FWA-PSNN from annual crude oil price time series

Table 4. Computed DM statistic values

| Crude oil price dataset                     |          | PSO-PSNN | GA-PSNN | GD-PSNN | MLP     |
|---|----------|----------|---------|---------|---------|
| Daily Cushing, OK Crude Oil Price Dataset   |          | 1.9845   | -2.2650 | 2.2363  | 2.7746  |
| Weekly Cushing, OK Crude Oil Price Dataset  | FWA-PSNN | 2.1035   | 2.2505  | -3.1507 | -3.2362 |
| Monthly Cushing, OK Crude Oil Price Dataset |          | -2.3165  | 2.0163  | 2.1553  | 2.7058  |
| Annual Cushing, OK Crude Oil Price Dataset  |          | 1.9965   | 1.9775  | 2.1105  | -3.1528 |

Further, to find the precise advantage of the proposed model we conducted a statistical significance test, i.e. Deibold-Mariano test [55, 58]. It compares two or more forecasts pair wise. The computed DM statistics are presented in Table 4. It is observed that the statistics obtained are beyond the critical range. Hence, the null hypothesis of no difference between the proposed and other models is rejected.

To check the performance of the forecasts considered we computed the relative worth (RW) of a model [57]. It considers the average reductions% in the forecast error of

the best performing model over all datasets. Then,  $RW_j$  of the  $j^{th}$  combination method is defined as in Eq. 12

$$RW_j = \frac{1}{D} \sum_{i=1}^D \left[ \left( \frac{err_i^w - err_{ij}}{err_i^w} \right) * 100 \right], \forall j = 1, 2, \dots, n \tag{12}$$

where,  $err_{ij}$  is the forecast error from  $j^{th}$  model on  $i^{th}$  dataset,  $err_i^w$  is the error of worst individual model for the same dataset. D is the total number of datasets. The

computed RW values are summarized in Table 4. It can be observed that the relative worth value of the proposed model is highest, i.e. 69.7968% followed by PSO-PSNN and GA-PSNN with values 42.0950% and 28.9129%. The GD-PSNN model obtained the lowest worth value of 11.1994%.

Table 5. Relative worth values of forecasting models

| Forecasting model | Relative worth value |
|-------------------|----------------------|
| FWA-PSNN          | 69.7968              |
| PSO-PSNN          | 42.0950              |
| GA-PSNN           | 28.9129              |
| GD-PSNN           | 11.1994              |

## 5. Conclusions

In order to incarcerate the uncertainties coupled with the crude oil prices, this paper proposed a hybrid forecasting model termed as FWA-PSNN. The model uses the fast and effective learning ability of FWA and better generalization property of PSNN for modelling and forecasting crude oil prices efficiently. FWA is used to explore the most favourable weight and bias vector of PSNN based forecast and quite able to minimize the forecasting error. The proposed FWA-PSNN based forecasting is validated on predicting one-step-ahead oil price of four real crude oil price datasets. The performance of FWA-PSNN is compared with that of four other forecasts such as PSO-PSNN, GA-PSNN, GD-PSNN and MLP trained similarly. From extensive simulation studies and result analysis it is observed that FWA-PSNN performed superior to others. The current work may be extended by applying the proposed model to other data mining problems. Exploration of other nature inspired optimization techniques can be another direction.

## References

- [1] Nelson, Y., Stoner, S., Gemis, G., & Nix, H. D. (1994). Results of Delphi VIII survey of oil price forecasts. *Energy report, California Energy Commission*.
- [2] Mahdiani, M. R., & Khamehchi, E. (2016). A modified neural network model for predicting the crude oil price. *Intellectual Economics*, 10(2), 71-77.
- [3] Chiroma, H., Abdul-kareem, S., Shukri Mohd Noor, A., Abubakar, A. I., Sohrabi Safa, N., Shuib, L., ... & Herawan, T. (2016). A review on artificial intelligence methodologies for the forecasting of crude oil price. *Intelligent Automation & Soft Computing*, 22(3), 449-462.
- [4] Hamdi, M., & Aloui, C. (2015). Forecasting crude oil price using artificial neural networks: a literature survey. *Econ Bull*, 3(2), 1339-1359.
- [5] Wang, J., & Wang, J. (2016). Forecasting energy market indices with recurrent neural networks: Case study of crude oil price fluctuations. *Energy*, 102, 365-374.
- [6] Chiroma, H., Abdulkareem, S., & Herawan, T. (2015). Evolutionary Neural Network model for West Texas Intermediate crude oil price prediction. *Applied Energy*, 142, 266-273.
- [7] S. Haykin, *Neural Networks and Learning Machine*, PHI, ISBN -978-81-203-4000-8, 2010.
- [8] V. Kecman, *Learning and Soft Computing*, Pearson Education, ISBN 81-317-0305-3, 2006.
- [9] Rajasekaran, S., & Pai, G. V. (2003). *Neural networks, fuzzy logic and genetic algorithm: synthesis and applications (with cd)*. PHI Learning Pvt. Ltd.
- [10] Wang, Z., Fang, J. A., & Liu, X. (2008). Global stability of stochastic high-order neural networks with discrete and distributed delays. *Chaos, Solitons & Fractals*, 36(2), 388-396.
- [11] Ghosh, J., & Shin, Y. (1992). Efficient higher-order neural networks for classification and function approximation. *International Journal of Neural Systems*, 3(04), 323-350.
- [12] Nayak, S. C., Misra, B. B., & Behera, H. S. (2016). Fluctuation prediction of stock market index by adaptive evolutionary higher order neural networks. *International Journal of Swarm Intelligence*, 2(2-4), 229-253.
- [13] Nayak, S. C. (2017). Development and Performance Evaluation of Adaptive Hybrid Higher Order Neural Networks for Exchange Rate Prediction. *International Journal of Intelligent Systems and Applications*, 9(8), 71.
- [14] Nayak, S. C., Misra, B. B., & Behera, H. S. (2015). A pi-sigma higher order neural network for stock index forecasting. In *Computational Intelligence in Data Mining-Volume 2* (pp. 311-319). Springer, New Delhi.
- [15] Nayak, S. C., Misra, B. B., & Behera, H. S. (2017). Adaptive Hybrid Higher Order Neural Networks for Prediction of Stock Market Behavior. In *Nature-Inspired Computing: Concepts, Methodologies, Tools, and Applications* (pp. 553-570). IGI Global.
- [16] Nayak, S. C., Misra, B. B., & Behera, H. S. (2018). On Developing and Performance Evaluation of Adaptive Second Order Neural Network With GA-Based Training (ASONN-GA) for Financial Time Series Prediction. In *Advancements in Applied Metaheuristic Computing* (pp. 231-263). IGI Global.
- [17] N. Shadbolt, "Nature-inspired computing," *IEEE Intell Systems*, 19(1), pp. 2-3, 2004.
- [18] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*, Addison-Wesley, Reading, MA, USA, 1989.

- [19] J. Kennedy, RC. Eberhart, Swarm intelligence. Morgan Kaufmann, San Francisco, 2001.
- [20] K. Price, R. Storn, J. Lampinen, "Differential evolution: a practical approach to global optimization," Springer, Berlin, 2005.
- [21] M. Dorigo, T. Stutzle, Ant colony optimization. The MIT Press, Cambridge, MA, USA, 2004.
- [22] Tan, Y., & Zhu, Y. (2010, June). Fireworks algorithm for optimization. In International conference in swarm intelligence(pp. 355-364). Springer, Berlin, Heidelberg.
- [23] Li, X. G., Han, S. F., & Gong, C. Q. (2017). Analysis and improvement of fireworks algorithm. *Algorithms*, 10(1), 26.
- [24] Zheng, S., Li, J., Janecek, A., & Tan, Y. (2017). A cooperative framework for fireworks algorithm. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 14(1), 27-41.
- [25] Zheng, S., Janecek, A., & Tan, Y. (2013, June). Enhanced fireworks algorithm. In 2013 IEEE Congress on evolutionary computation (pp. 2069-2077). IEEE.
- [26] Li, J., Zheng, S., & Tan, Y. (2014, July). Adaptive fireworks algorithm. In 2014 IEEE Congress on evolutionary computation (CEC) (pp. 3214-3221). IEEE.
- [27] Zheng, S., Janecek, A., Li, J., & Tan, Y. (2014, July). Dynamic search in fireworks algorithm. In 2014 IEEE Congress on evolutionary computation (CEC) (pp. 3222-3229). IEEE.
- [28] Moshiri, S., & Foroutan, F. (2006). Forecasting nonlinear crude oil futures prices. *The Energy Journal*, 81-95.
- [29] Atsalakis, G. S., & Valavanis, K. P. (2010). Surveying stock market forecasting techniques-Part I: Conventional methods. *Journal of Computational Optimization in Economics and Finance*, 2(1), 45-92.
- [30] Fan, L., & Li, H. (2015). Volatility analysis and forecasting models of crude oil prices: A review. *International Journal of Global Energy Issues*, 38(1-3), 5-17.
- [31] Gavriilidis, K., Kambouroudis, D. S., Tsakou, K., & Tsouknidis, D. A. (2018). Volatility forecasting across tanker freight rates: The role of oil price shocks. *Transportation Research Part E: Logistics and Transportation Review*, 118, 376-391.
- [32] Herrera, A. M., Hu, L., & Pastor, D. (2018). Forecasting crude oil price volatility. *International Journal of Forecasting*, 34(4), 622-635.
- [33] Naderi, M., Khomehchi, E., & Karimi, B. (2019). Novel statistical forecasting models for crude oil price, gas price, and interest rate based on meta-heuristic bat algorithm. *Journal of Petroleum Science and Engineering*, 172, 13-22.
- [34] Wang, Q., Li, S., & Li, R. (2018). China's dependency on foreign oil will exceed 80% by 2030: Developing a novel NMGM-ARIMA to forecast China's foreign oil dependence from two dimensions. *Energy*, 163, 151-167.
- [35] Huang, S. C., & Wu, C. F. (2018). Energy commodity price forecasting with deep multiple kernel learning. *Energies*, 11(11), 3029.
- [36] Wang, Q., Li, S., & Li, R. (2018). Forecasting energy demand in China and India: Using single-linear, hybrid-linear, and non-linear time series forecast techniques. *Energy*, 161, 821-831.
- [37] Chen, Y., He, K., & Tso, G. K. (2017). Forecasting crude oil prices: a deep learning based model. *Procedia computer science*, 122, 300-307.
- [38] Akrom, N., & Ismail, Z. (2017). A Hybrid GA-FEEMD for Forecasting Crude Oil Prices. *Indian Journal of Science and Technology*, 10, 31.
- [39] Hu, H., Zhai, X., & Guan, X. (2017, November). Crude oil output forecasting based on PSO of unbiased gray Markov model. In 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS) (pp. 644-647). IEEE.
- [40] CHANDAR, S. K., SUMATHI, M., & SIVANANDAM, S. (2015). CRUDE OIL PREDICTION USING A HYBRID RADIAL BASIS FUNCTION NETWORK. *Journal of Theoretical and Applied Information Technology*, 72(2).
- [41] Yusof, Y., & Mustaffa, Z. (2015, March). Time series forecasting of energy commodity using grey wolf optimizer. In *Proceedings of the International Multi Conference of Engineers and Computer Scientists (IMECS'15)* (Vol. 1, No. 1).
- [42] Hu, H., Fan, L., & Guan, X. (2017, September). The research on modeling and simulation of crude oil output prediction based on KPCA-DE-SVM. In 2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA) (pp. 93-97). IEEE.
- [43] Gabralla, L. A., & Abraham, A. (2013). Computational modeling of crude oil price forecasting: A review of two decades of research. *International Journal of Computer Information Systems and Industrial Management Applications*, 5, 729-740.
- [44] Firdaus, H., Hassan, S. I., & Kaur, H. (2018). A Comparative Survey of Machine Learning and Meta-Heuristic Optimization Algorithms for Sustainable and Smart Healthcare. *African Journal of Computing & ICT*, 11(1), 4, 1-17.
- [45] Kaur, P., & Sharma, M. (2019). Diagnosis of Human Psychological Disorders using Supervised Learning and Nature-Inspired Computing Techniques: A Meta-Analysis. *Journal of medical systems*, 43(7), 204.

- [46] Gautam, R., Kaur, P., & Sharma, M. (2019). A comprehensive review on nature inspired computing algorithms for the diagnosis of chronic disorders in human beings. *Progress in Artificial Intelligence*, 1-24.
- [47] Suguna, S. K., Ranganathan, R., Sangeetha, J., Shandilya, S., & Shandilya, S. K. (2019). Application of Nature—Inspired Algorithms in Medical Image Processing. In *Advances in Nature-Inspired Computing and Applications* (pp. 61-100). Springer, Cham.
- [48] Sharma, M., Singh, G., & Singh, R. (2019). A review of different cost-based distributed query optimizers. *Progress in Artificial Intelligence*, 8(1), 45-62.
- [49] Naderi, M., Khamsehchi, E., & Karimi, B. (2019). Novel statistical forecasting models for crude oil price, gas price, and interest rate based on meta-heuristic bat algorithm. *Journal of Petroleum Science and Engineering*, 172, 13-22.
- [50] Faris, H., Hassonah, M. A., Ala'M, A. Z., Mirjalili, S., & Aljarah, I. (2018). A multi-verse optimizer approach for feature selection and optimizing SVM parameters based on a robust system architecture. *Neural Computing and Applications*, 30(8), 2355-2369.
- [51] Ma, T., & Niu, D. (2016). Icing forecasting of high voltage transmission line using weighted least square support vector machine with fireworks algorithm for feature selection. *Applied Sciences*, 6(12), 438.
- [52] Bacanin, N., & Tuba, M. (2015, May). Fireworks algorithm applied to constrained portfolio optimization problem. In *2015 IEEE Congress on evolutionary computation (CEC)* (pp. 1242-1249). IEEE.
- [53] Tuba, M., Bacanin, N., & Alihodzic, A. (2015, April). Multilevel image thresholding by fireworks algorithm. In *2015 25th International Conference Radioelektronika (RADIOELEKTRONIKA)* (pp. 326-330). IEEE.
- [54] Tuba, E., Tuba, M., & Dolicanin, E. (2017). Adjusted fireworks algorithm applied to retinal image registration. *Studies in Informatics and Control*, 26(1), 33-42.
- [55] Nayak, S. C., Misra, B. B., & Behera, H. S. (2018). ACFLN: artificial chemical functional link network for prediction of stock market index. *Evolving Systems*, 1-26.
- [56] Nayak, S. C., Misra, B. B., & Behera, H. S. (2014). Impact of data normalization on stock index forecasting. *Int. J. Comp. Inf. Syst. Ind. Manag. Appl*, 6, 357-369.
- [57] Nayak, S. C. (2019). Escalation of Forecasting Accuracy through Linear Combiners of Predictive Models. *EAI Endorsed Transactions of Scalable Information system*.
- [58] Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, 20(1), 134-144.