Hybridization Of Stdl With Optimal Kernel Extreme Learning Machine (Okelm) Based Short Term Crude Oil Price Forecasting In Commodity Futures Market

V. Veeramanikandan, M. Jeyakarthic

Abstract: Futures markets offer contemporary price quotations for a group of contracts with maturities twelve or more months in the future. It coherently determines price levels and price variations suitable for contract temporal definitions. Crude oil is an essential kind of energy and its prices have a huge influence on worldwide financial system. So, it is highly needed to predict the price of crude oil precisely for investors, government, and also academicians. But, because of the non-linearity and non-static prices of crude oil prices, it is a challenging task for the conventional models of time series forecasting for managing it. This study proposes a new hybridization of seasonal-trend decomposition procedures based on loess (STDL) and optimal kernel extreme learning machines (OKELM) for short-term and mid-term forecasting of every day close price of CRUDE OIL index. The parameter tuning of ELM takes place by the use of grey wolf optimization algorithm (GWO) to further enhance the prediction performance of the ELM. The validation of the proposed work takes place by the use of two performance measures namely MASE and SMAPE. The experimental outcome showed the superiority of the proposed model over the compared methods.

Index Terms: Crude oil, Prediction model, Grey wolf optimization, Seasonal Trend Decomposition by Loess (STDL), Kernel Extreme Learning Machine (KELM).

1 INTRODUCTION

Crude oil is a primary commodity for global economy. It is an important element for the financial growth of developed and developing countries. Besides, political relations, weather conditions, variation in financial market also plays an important role in crude oil market that raises the volatile nature of price levels in oil market index. The impact of variations in oil prices reaches to massive number of goods and services which straightaway affects the financial status of the country and people. Therefore, to minimize the negative effect of the price differences, it is highly essential to predict the price levels. Although precise forecast of closed prices of crude oil index offers more details for users, only less number of studies have worked on it. [1] examined the results of three classical parametric models for crude oil prices and made recommendations to significant users who focused on futures market watching. [2] presented a restricted evolving spline approach for capturing the adaptive process of the variation in seasonal patterns of tomato prices and predict the future prices. [3] makes use of back-propagation neural network (BPNN) to predict the future price of tomatoes. [4] presented many ANN- based predictive approaches to simulate and predict the varying price levels of tomatoes in US. [5] designed a predictive technique depending upon radial basis function network for the prediction of tomato prices.

- *Assistant Professor, Department of Computer Science, T.K. Govt Arts College, Vridhachalam, klmvmani@gmail.com*
- *Assistant Director, Tamil virtual Academy, Chennai, jeya_karthic@yahoo.com*

[6] undergo a comparison of the predictive performances of out- look hog price forecast using different alternative time series models. [7] investigated the chaos and nonlinearity in the future prices of crude oil. Based on various statistical and econometrical tests, the prices can be forecast. In addition, a comparison is made between linear and nonlinear-based models to forecast the oil prices. [8] presented a hybrid model for forecasting the monthly price of crude oil by the use of web mining, ANN and ARIMA concepts. It is defined that the nonlinear integration of the three approaches outperforms other methods. For instance, the rule base system of the text mining approach is based on the knowledge base presented human experts. Furthermore, neither the rules nor the knowledge base is accessible to the community. [9] presented SVM model to predict the monthly crude oil prices. This study proposes a new hybridization of seasonal-trend decomposition procedures based on loess (STDL) and optimal extreme learning machines (OELM) for short-term and mid-term forecasting of every day close price of CRUDE OIL index. The parameter tuning of ELM takes place by the use of grey wolf optimization algorithm (GWO) to further enhance the prediction performance of the ELM. The validation of the proposed work takes place by the use of two performance measures namely MASE and SMAPE. The experimental outcome showed the superiority of the proposed model over the compared methods. This paper is organized as follows, Section II discusses the proposed model, Section III provides the experimental analysis, and forthwith the paper is concluded in Section V.

2 THE PROPOSED STDL-OKELM MODEL

Fig. 1 defines the detailed working of the proposed STDL-OKELM method. First, the STDL, ELM, KELM and GWO algorithms are explained briefly. Afterwards, the proposed STDL-OKELM method gets formulated, and then respective steps are depicted in detail manner.

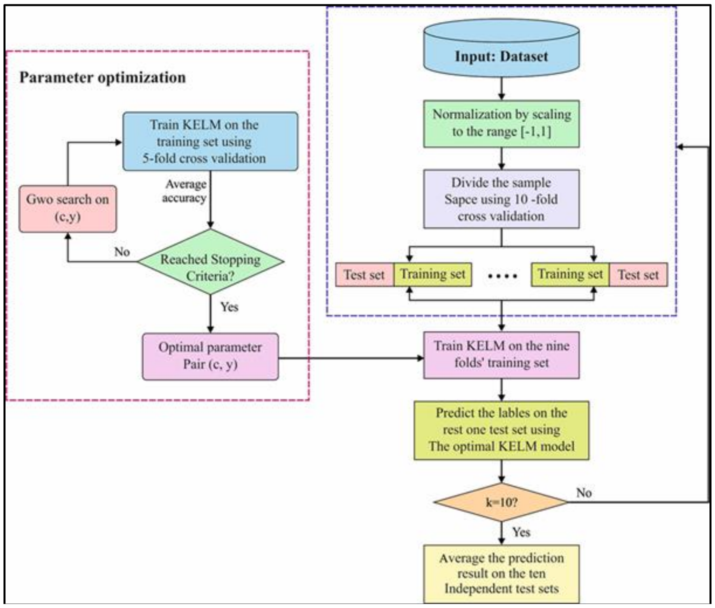

Fig. 1. Working process of GWO-KELM

2.1 Seasonal-trend decomposition procedure based on loess (STDL)

The STL method, initially presented by [10], it is a filtering process to decompose a time series into additive diversification elements. Really, it is varied from classical seasonal decomposition models like X-12-ARIMA and the ratio-to-moving-average method, STL offers rigid performance when dealing with the outliers in the investigated time series. Provided a closed prices of crude index series X_t, STDL will decompose X_t to the three additive elements of seasonal S_t, trend T_t, and remainder R_t elements: X_t= S_t+ T_t+ R_t. STDL is a repetitive produce contains a pair of recursive operative steps, inner and outer loops. Every passes via the inner loop performs seasonal smoothing revives the seasonal element, accompanied by a trend smoothing which update the trend element. After completion of the inner loop, robustness weight tends to be calculated in the outer loop, which is next employed to minimize the impact of the outliers to update the seasonal and trend elements in the succeeding inner loops.

Forecasting with decomposition

While decomposition is mostly helpful to study time series data, and searching historical changes over time, it can also be utilized in forecasting. Supposing an additive decay, the decayed time series could be written as

$$y_t = \hat{S}_t + \hat{A}_t \qquad (1)$$

where $\hat{A}_t = \hat{T}_t + \hat{R}_t$ is the seasonally adjusted element. Or, if a multiplicative decay has been utilized, we can write as,

$$y_t = \hat{S}_t \hat{A}_t \qquad (2)$$

where$\hat{A}_t = \hat{T}_t \hat{R}_t$.. To predict a decayed time series, we predict the seasonal element, $\hat{S}_t$, and the seasonally adjusted element $\hat{A}_t$, individually. It is generally assumed that the seasonal element is unchanging or changing very slow, so it is predict with easily getting the last year of the calculated element. In other words, seasonal naïve techniques are utilized to the seasonal element.

2.2 Extreme Learning Machine (ELM)

ELM is primarily presented with [11], ELM is a unique learning technique to a single hidden- layer feed-forward neural network (SLFN). The flow diagram and single layer feed forward ELM is revealed in Fig. 2 and Fig. 3 respectively. Obviously various from the gradient-based learning technique utilized to usual ANN, the biases and input weights are contingently resolved, and consequently, the resultant weights are tuned utilizing easy matrix calculations in ELM, and that reduction training time in fine manner. In actual applications, the NNs are trained for finite training samples. In order to achieve best estimation, hidden layer concepts are used in the traditional neural applications. The learning speed of FFNNs is too slower than what we needed, and it has been a main blockage chaos in their applications to past decades. The gradient based method is the most common learning technique used in FFNN. But they are seems to be extremely slowly because of improper learning steps or can simply intersect to local minima.

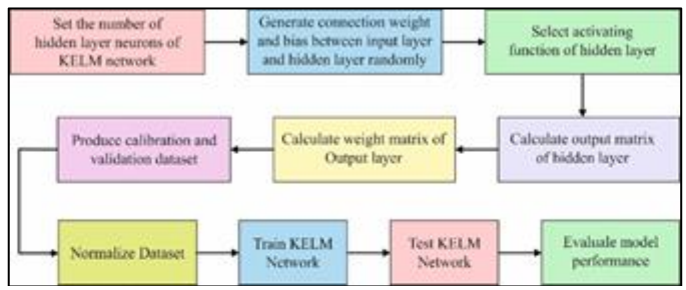

Fig. 2. Process Flow Diagram of KELM

This learning technique requires several iterative learning procedures for obtaining optimal action. The hidden layer concept and the non-linear activation functions make the learning algorithm very slow. Consequently, a model is introduced for making fast learning with reduced number of hidden layers. ELM as a rising technology has exposed its optimal and effective action in regression applications as well as in huge volume of dataset classification applications.

Provided a set of N different samples { (x j , t j) }$_j^N$ =1 with inputs X_j= [X_{j1} , X_{j2}, . . . ,X_{jn}] T ∈ _ $\vartheta^n$and outputs t_j= [t_{j1} , t_j2, . . . ,t_{jm}]T∈ _ _ ϑ^m, the ELM with $\tilde{N}$hidden neurons with activation function g (·) is mathematically modelled as

$$\sum_{j=1}^{L} \beta_j \phi(W_j X_i + b_j) \qquad i\epsilon[1,N] \qquad (3)$$

where W_j= [W_{j1}, [W_{j2} , . . . , [W_{jn}] T signifies the weight vector link the input neurons and the jth secret neurons, βj= [βj1 , βj2 , . . . , βjm] T signifies the weight vector of the link from the jth secret neurons to resultant neurons, and $b_j$specifies the threshold of the jth secret node.

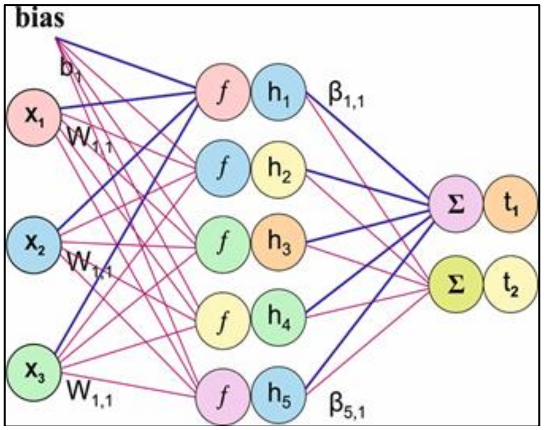

Fig. 3. *Single Layer Feed Forward ELM*

The relation between target input and output layers of network are defined by

$$Y_i = \sum_{j=1}^{L} \beta_j \, \phi\left(W_j X_i + b_j\right) = \; t_i \epsilon i i \epsilon [1, N] \qquad (4)$$

On the equation could be rewritten efficiently as H β= T , where β= [β1 , . . . , β˜ N] T , T = [t 1 , . . . , t N] T and H is the secret layer resultant matrix of ELM. The hidden neurons are transforming the input data into a different representation in two procedural ways. First, the data is predicted into the secret layer through the weights with biases of the input layer, and then implying it to the result of that a non-linear activation functions. Practically, ELMs are solved as common Neural Networks in their matrix form and the structure is shown in Fig. 4. The matrix form is represented here:

$$H = \begin{bmatrix} \emptyset(w_1 x_1 + b_1) & \cdots & \emptyset(w_L x_1 + b_L) \\ \vdots & \ddots & \vdots \\ \emptyset(w_1 x_N + b_1) & \cdots & \emptyset(w_L x_N + b_L) \end{bmatrix}. \qquad (5)$$

$$\beta = (\beta_1^T \cdots \beta_L^T)^T, \mathbb{T} = (y_1^T \cdots y_N^T)^T. \qquad (6)$$

And here is the important part of this method obtained. Given that T is the target we want to reach, a unique result of the system with least squared error could be found utilizing Moore-Penrose generalized inverse. Therefore, we can compute in one single operation the values of the weights of the secret layer that will result in the solution with the least error to forecast the destination T.

$$\mathrm{H}\beta = \mathrm{T}$$
$$\beta = \mathrm{H}^{\dagger} T$$

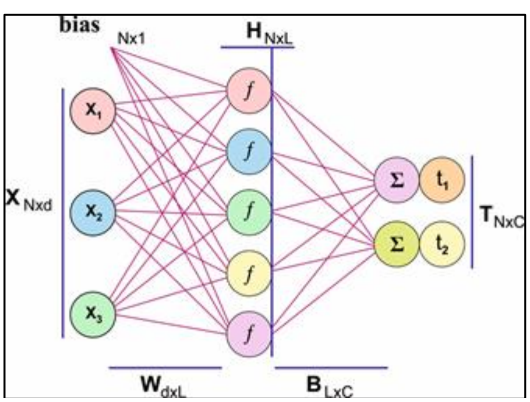

Fig. 4. *Structure of ELM*

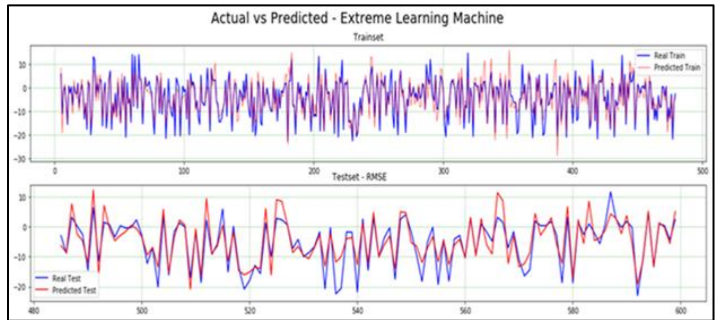

Fig. 5. *Model Resultant Shots of ELM*

This pseudo inverse in calculated using the Singular Value Decomposition. The resultant shots of the ELM are shown in Fig. 5.

2.3 Evolutionary KELM Technique

Recently, KELM is assumed to be an efficient learning model which has been emerged from ELM. It is applicable in exhibiting maximum number of real-time applications. In order to attain effective visibility in KELM, a recent work has been proposed the single secret layer feed forward NN which has been declared as normalized SLFNs in absence of tuning attributes from hidden layer. Hence, the simulation outcome of ELM in simplified SLFNs has been represented in Eq. (7).

$$f_L(u) = \sum_{p=1}^{L} \beta_p h_p(u) = h(u)\beta \qquad (7)$$

where $\beta = [\beta_1, \cdots, \beta_L]^Z$ denotes the vector of final result among secret layer of L nodes and outcome node, outcome vector of secret layer is equivalent to input u is mentioned as $h(u) = [h_1(u), \cdots, h_L(u)]$. Alternatively, the dataset could be matched among d-dimensional input space to the L-dimensional secret layer feature space H which is based on value of h(u), that refers the h(u) is feature mapping. Hence, it is termed as hypothesis through the aim of attaining lower training error from feed forward NN. Therefore, if norms are lower; then it helps to achieve optimal computation of networks. Thus, ELM attempts to reduce training fault and norm of the outcome weights at same time.

$$mn\|H\beta - Z\|^2 and \; \|\beta\| \qquad (8)$$

The H from former formation refers to hidden-layer output matrix

$$H = \begin{bmatrix} h(u_1) \\ h(u_2) \\ h(u_3) \end{bmatrix} = \begin{bmatrix} h_1(u_1) & \cdots & h_L(u_1) \\ \vdots & \ddots & \vdots \\ h_1(u_N) & \cdots & h_L(u_N) \end{bmatrix} \qquad (9)$$

The main purpose to decrease the norm of outcome weights $\|\beta\|$ tends to improve the distance of divided borders which belongs to various classes in ELM feature space. Furthermore, minimum amount of normal least square technique could be applied in ELM

$$\beta = H^{\dagger} Z \qquad (10)$$

where $H^{\dagger}$ implies the Moore-Penrose which is a normalized inverse of matrix. Here, the actual factor represents the users unaware of feature mapping, whereas kernel matrix is termed as kernel mapping operation in ELM would be considered by applying the given notation:

$$\Omega_{\mathrm{ELM}} = \mathrm{HH}^{\mathrm{Z}} : \Omega_{\mathrm{ELM}_{\mathrm{p,q}}} = h(u_p) \cdot h(u_q) = K(u_p, u_q) \qquad (11)$$

where h(u) represent the mapping function that maps data which is attained from input space in order to guarantee the linear division in secret layer feature space H. Therefore, the orthogonal presentation has been applied to estimate the MP generalized matrix's in reverse order as $H^{\dagger} = \mathrm{H}^{\mathrm{Z}}(\mathrm{HH}^{\mathrm{Z}})^{-1}$, as well as a positive constant C has been presented for the diagonal of HH^{Z}. The output notation of ELM could be defined as:

$$F(u) = h\beta = h(u)H^{\dagger}\left(\frac{I}{C} + HH^{\dagger}\right)^{-1} Z$$

$$= \begin{bmatrix} K(u, u_1) \\ \vdots \\ K(u, u_L) \end{bmatrix}^T \left(\frac{I}{C} + \Omega_{\mathrm{ELM}}\right)^{-1} Z \qquad (12)$$

There is no requirement for users to understand the details of hidden layer feature which match the coped kernel trick, whereas Radial Basis Function kernel (RBF) could be applicable. Here, RBF kernel function could be described as $K(u, u_p) = \exp\left(-\gamma \|u - u_p\|^2\right)$. According to the significant attributes which has been provided in RBF kernel is penalty parameter C as well as kernel parameter gamma γ.

2.4 GWO Algorithm

GWO (Grey Wolves Optimization) technique has been proposed. This GWO method is actually based on the behaviour of grey wolves and the hunting methodology of leadership hierarchy. Here, grey wolves are regarded as superior killers as they often live in a group of 7 to 15 members. According to the hunting technology, grey wolves are segmented into tetra types, namely into alpha (α), beta(β), delta(δ), and omega(ω) which is depicted in Fig. 6.

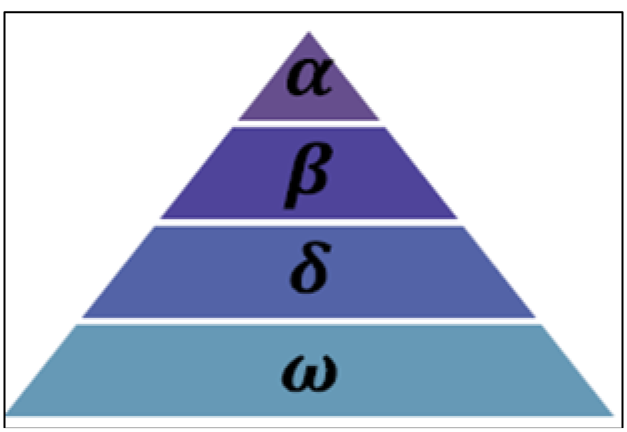

Fig. 6. *Hierarchy of grey wolves*

Initially, Alpha wolves are often declared as leaders of whole team since it is capable of deciding the resident place, chasing prey etc. Hence, these kinds of wolves are always leading as they orders neighbour wolves to obey their commands. Alpha wolves are most significant in generating novel solutions. Some of the wolves which are in the next level of alpha wolves have rights to make decisions in absence of alpha wolves. The main duty of other wolves is to follow the order and respond to alpha wolves accordingly. Then, delta wolves come under alpha a wolf that is known as subordinate wolves. The delta wolves come under the type of elders, sentinels, hunters, scouts, caretakers and so on. It follows the rules of alpha and beta wolves as well as controls the omega wolves or subordinates. Finally, omega is placed in last position which is considered as scapegoat. These types of wolf obey the rules arrived from hierarchy or dominant wolves. Although an omega does not play a crucial role, it helps neighbouring wolves to meet complex issues. Next, the hunting techniques of wolves are classified into 3 types as:

- Tracking and chasing
- Pursuing and encircling
- Attacking the prey

The main aim of GWO is to explore and exploit the prey or victim. As shown in Fig. 7, in the exploitation stage, identifying optimal solutions from local search space process takes place. The grey wolves encircles and attacks the prey has been applied in exploring maximum solutions of local search space. Subsequently, process of seeking for a prey is performed in exploration level where grey wolves finds prey by using global search space. When the prey is encircled, then wolves analyse the location of victim and encompass them. Therefore the location vector of victim is denoted by explore agents when it alters the location dependent on the obtained optimal result. Hence, the encircled prey could be expressed by:

$$\vec{D} = \left|\vec{C}.\overrightarrow{M_p}(k \propto) - \vec{M}(k)\right| \qquad (13)$$

$$\vec{M}(k+1) = \overrightarrow{M_p}(k) - \vec{A}.\vec{D} \qquad (14)$$

Where k implies the present iteration, $\vec{A}$ and $\vec{C}$ signifies coefficient vectors, location vector of the victim are signified by $\overrightarrow{M_p}(k)$, $\vec{M}$ indicated as position vector, || denotes the actual value of step by step propagation. Hence, vectors $\vec{A}$ and $\vec{C}$ are calculated as:

$$\vec{A} = 2\vec{a}.\vec{r} - \vec{a} \qquad (15)$$

$$\vec{C} = 2.\vec{r} \qquad (16)$$

In hunting stage, grey wolves have been directed with the help of α as well as some of the process has been provided to β and δ. In the primary phase, it is complex to obtain better solution due to the presence of maximum space while in hunting level, α is considered as optimal solution, β is declared as consecutive maximum solution whereas delta is last optimal candidate solution. Among other process, the attained solutions are upgraded for modification of lower grading solution. Therefore, the template to describe hunting principle could be provided as:

$$\vec{D}_\alpha = |\vec{C}_1 * \vec{M}_\alpha - \vec{M}|$$

$$\vec{D}_\beta = |\vec{C}_2 * \vec{M}_\beta - \vec{M}| \qquad (17)$$

$$\vec{D}_\delta = |\vec{C}_1 * \vec{M}_\delta - \vec{M}|$$

where $\vec{D}_\alpha$, $\vec{D}_\beta$, and $\vec{D}_\delta$ represents the manipulated distance vector among α, β and γ location for alternate wolves as well as $\vec{C}_1$, $\vec{C}_2$, and $\vec{C}_3$ are 3 coefficient vector which is applied to alter the distance vector. $\vec{M}$ is assumed to be placement of vector of another grey wolf (omega).

$$\vec{M}_1 = \vec{M}_\alpha - \vec{A}_1 * \vec{D}_\alpha$$

$$\vec{M}_2 = \vec{M}_\beta - \vec{A}_2 * \vec{D}_\beta \qquad (18)$$

$$\vec{M}_3 = \vec{M}_\delta - \vec{A}_3 * \vec{D}_\delta$$

where $\vec{M}_1$ exposes the new location achieved under the application of α position $\vec{M}_\alpha$ as well as distance vector $\vec{D}_\alpha$, $\vec{M}_2$ denotes a novel location obtained using β position $\vec{M}_\beta$ and distance vector $\vec{D}_\beta$, $\vec{M}_3$ expresses the effective position vector that is computed with the help of delta position $\vec{M}_\delta$ and distance vector $\vec{D}_\delta$, and $\vec{A}_1$, $\vec{A}_2$, and $\vec{A}_3$ are the tertiary coefficient vectors that has been computed.

$$\vec{M}(k+1) = \frac{\sum_{i=1}^{n}\vec{X}_i}{n} \tag{19}$$

where $\vec{M}(k+1)$ represents the final position vector that is processed with the help of maximum value of every positions which is accomplished using α, β and δ (n=3).

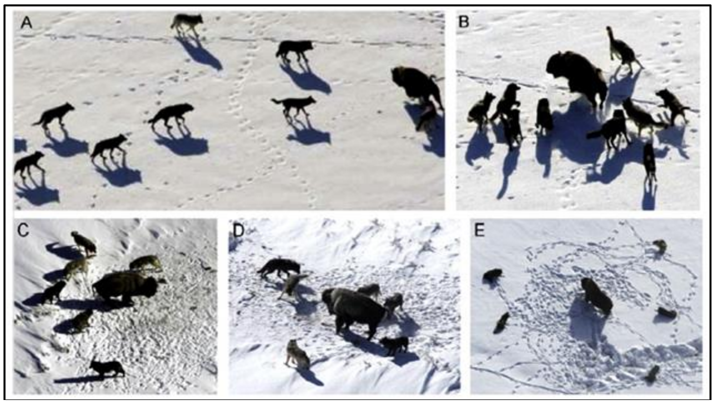

***Fig. 7**. Hunting behavior of grey wolves: (A) chasing, approaching, and tracking prey (B–D) pursuing, harassing, and encircling (E) stationary situation and attack*

2.5 GWO Based KELM

The feature of KELM could be managed by 2 significant attributes of C and γ. This research work proposes a novel method named as Grey Wolf Optimization Algorithm (GWO). The GWO helps in identifying two variables present in KELM. Therefore, the final outcome of KELM technique which is termed as GWO-KELM. The integration of GWO principle is applicable to compute the 2 important parameters of KELM. Here, GWO-KELM is comprised with two strategies along with parameter optimizing as well as classifying property estimation. While the procedure of attribute optimization is carried out, the interior 5-fold CV would be performed on the basis of maximum quantity of dataset. If the process of inner attributes optimization is paused then, optimum attribute pair could be induced to KELM predicting technique in order to detect the forecasting price from external loop which is based on application of outer 10-fold CV determination.

Here, the accuracy obtained from classification task has been considered while modelling the fitness function:

$$F = avgACC = \frac{\left(\sum_{p=1}^{K} testACC_p\right)}{K} \tag{20}$$

where $avgACC$ involved in function F refers the sample accuracy attained by KELM classification model through the 5-fold CV approach by interior parameter optimization technique. The pseudo code of GWO-KELM is provided in Algorithm 1.

Algorithm 1: The Pseudocode of GWO-KELM

Step 1. Obtain the training and testing dataset
Step 2. Initiate KELM train
Step 3. Assign KELM parameters
Step 4. Set MSE as a fitness function
Step 5. Initialize GWO population (P)
Step 6. Compute the fitness value of each candidate solution
Step 7. S=global best solution
Step 8. For it= 1 to max_ite, do
Step 9. For i = 1 to P do
Step 10. Update the position of the i^th wolf
Step 11. Evaluate the fitness of the i^th wolf
Step 12. Update personal best solution of the i^th wolf
Step 13. S=current global best solution
Step 14. End for
Step 15. End for
Step 16. End
Step 17. Obtain the optimal input weights and hidden biases of hidden layer neurons using S
Step 18. KELM Testing

2.6 The proposed hybrid STDL-OKELM method

In this section, the presented STDL-OKELM techniques are framed, and the relevant steps are implied in specify. Provided a crude oil closed price index series X t for t = 1 , 2 , . . . , n , a 3 step modelling process to the presented STDL-OKELM technique could be formulated for closed price of crude oil index forecasting (Fig. 8). As revealed in Fig. 8, the presented STDL- OKELM techniques are basically comprised of the subsequent 3 important procedural ways:

Procedure 1: Decomposition.

The closed price index of crude oil series $X_t$is decayed into the seasonal S_t, trend T_t, and remainder $R_t$elements utilizing the STDL method as follows: X_t= S_t+ T_t+R_t, t = 1 , 2 , . . . , n . STDL is executed utilizing the Stl^2function of the library stats R package.

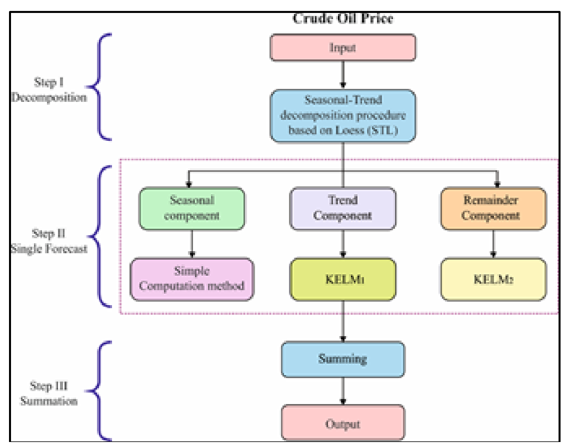

***Fig. 8.** The proposed STDL-OKELM model for crude oil price forecasting*

Procedure 2: Single prediction

OKELM is considered as a single predicting method for standalone method and prediction the taken out trend element T_t, t = 1 , 2 , . . . , n and remainder element R_t, t = 1 , 2 , . . . , n in a multi-step-ahead method utilizing the sequential scheme (forecast horizon H = 1, 2, and 3);this execution of OKELM and the model of multiple-step-ahead predicting are detailed as follows.To explain the working concept of multiple-step-

ahead predicting of OKELM utilized in this framework, the remainder element R_t, t = 1, 2, . . . ,n is resided as an example. Provided the remainder element R t ,t = 1 , 2 , . . . , n , as discussed in earlier section , an OKELM concepts is initially trained with minimized the squares of the in- sample single-step-ahead residuals. It depicts the model of multiple-step-ahead predicting with the achieved OKELM method. Besides utilizing the historical values of the remainder element (R_{t-d+1}, . . . ,R_{t-1}, R_t), the subsequent (single-step-ahead) value $\widehat{R}_{t+1}$is predicted. From the historical values of the remainder element (R_{t-d+2}, . . . ,R_{t-1}, R_t) and the previously predicted value $\widehat{R}_{t+1}$as the forecaster, the two-step-ahead forecasted value $\widehat{R}_{t+2}$ is received. As well, utilizing the historical values of the remainder element (R_{t-d+3}, . . . , R_{t-1}, R_t) and the earlier forecasted values $\widehat{R}_{t+1}$ and $\widehat{R}_{t+2}$ as the predictors, the 3-step- ahead forecasted value $\widehat{R}_{t+3}$is received. Therefore, the seasonal-naïve method considered to be a best method to forecasting the seasonal element. In the seasonal-naïve technique, the future value at period t + 3 is corresponding to the historical value at period t.
These methods could be represented as pursues:

$$\hat{S}_{t+3} = S_t, t = 1, 2, \ldots, n \tag{21}$$

where S t is the true value at period t of the seasonal element and ˆ S t+3 is the forecasted value at period t + 3 of the seasonal element.

Procedure 3: Summation
The forecast outcomes of the trend and remainder elements obtained by OKELM and the seasonal element using on seasonal-naïve technique in Procedure 2 are accumulated to get an aggregated outcome that are the last forecast closed price crude oil futures index.

3 PERFORMANCE VALIDATION
In this segment, the study design of the data representation, performance index criteria, and benchmark prediction methods are given in detail.

3.1 Data Representation
The tenure of three month closed index price crude oil is used as experimental datasets in this study because this crude oil plays a significant role in the everyday life of common people. The datasets are gathered from Multi Commodity Exchange India and are freely available from the official website [12] provided and regulated by Securities and Exchange Board of India (SEBI) since September 2015 [13]. The data from 07 Oct 2018 to 26 Dec 2018 are utilized as the holdout samples that are executed for evaluating the predicting action of the achieved methods. For evaluating the forecast capability of the presented STDL- ELM technique, we carry out one-, three-, and six-step-ahead predicting to examine the short and medium-term forecasting ability, correspondingly. It must be noted that an iterated scheme problem to executing the short and mid-term predicting. Additionally, in order to remove inputs in greater numeric ranges from dominating those in smaller numeric ranges, a linear transformation is adopted to adjust the actual original crude oil prices series scaled into the range of [0, 1] in this study.

***Table 1** Description of MCX CRUDEOIL DEC2018 data for last 90 days*

Parameter (for 45 days span)	*Price*
High	*5316*
Low	*3346*
Mean	*4217.4444*
Standard deviation	*565.1206*
Median	*4104*

3.2 Measurement criteria
In the measure of MASE and SMAPE are utilized as forecast accuracy measures:

$$SMAPE = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{Xt - \widehat{Xt}}{(|Xt| + |\widehat{Xt}|)/2}\right| \tag{22}$$

$$MASE = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{Xt - \widehat{Xt}}{\frac{1}{N-1}\sum_{i=2}^{N}|Xi - Xi - 1|}\right| \tag{23}$$

where $X_t$is the observation at period t, $\widehat{Xt}$ is the predict of X_t, and T and N are the numbers of observations in the evaluation sample and hold-out sample (in this case, N = 90 and T = 50), correspondingly. Fig. 9 shows the closing prices of crude oil commodity in Multi Commodity Exchange (MCX) market. The datasets are available freely from archives which are listed in their own website. As discussed in this approach the time series closed price data were taken for the three months from October 2018 to December 2018. Those data listed as start price, top price and closed price for every day. Start price reflects the previous day closed price. The top price reveals the surge in the price behaviour in a day. Closed price designated as closing time price of the particular commodity, here it denotes the crude oil. The fluctuation seen in the price series dataset shows chaotic characteristics like political influence, economic recession and so on that are factors affect the crude oil prices.
Z

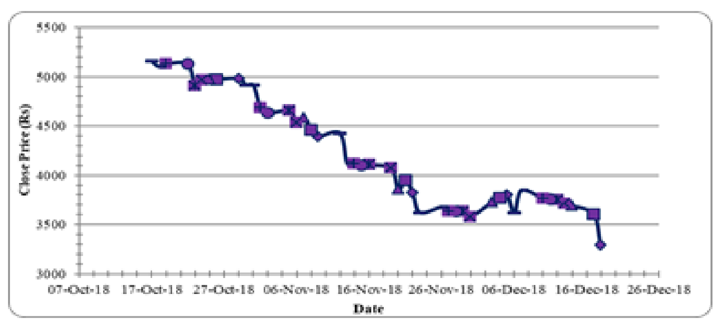

***Fig. 9**. Closing prices of MCX CRUDEOIL from Oct-2018 to Dec2018 (3 months)*

***Table 2** The feasible number of hidden nodes in STDL-OKELM and ELM Model*

No of Hidden Nodes	*Commodity: Crude Oil*
STDL-OKELM for trend element	6
STDL-OKELM for reminder element	6
Only ELM model	5

***Table 3** The final number of input, hidden, and output nodes in OKELM.*

Nodes Description	*Commodity: Crude Oil*
The number of input nodes	*10*
The number of hidden nodes	*6*
The number of output nodes	*1*

The above table consist of nodes description details such as number of input nodes, number of secret nodes and the final output nodes that are driven in ELM classification that derives single layer feed forward technique.

3.3 Results Analysis

Utilizing the research based mentioned beyond, the forecasted experiments for the crude oil price series are conducted. Aforesaid, the forecasted actions of every verified methods are evaluated utilizing the above mentioned accuracy measures.

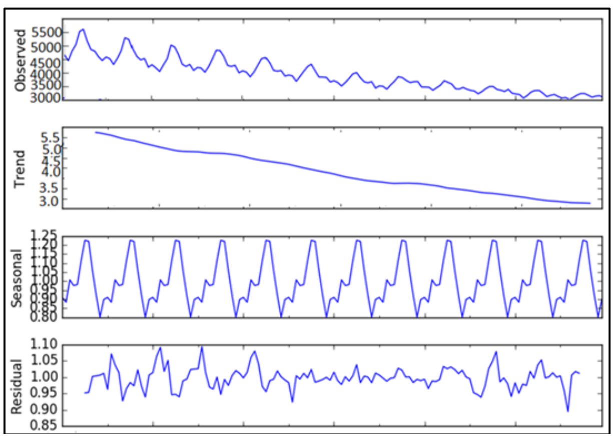

***Fig. 10.** The decomposition results of closed price index crude oil series by STDL-OKELM*

The decay results of the crude oil closed price index utilizing STDL is shown in Fig. 10. As discussed in proposed method, can be seen that crude oil closed prices index are decomposed into seasonal, trend, and remainder (residual) elements utilizing the STDL method. Entire seasonal elements of the crude oil closed index prices shows a three-month cycle. Furthermore, a stable growth and surge trend can be seen in the trend element observed in fourth quarter of 2018 calendar year. After the decomposition, the OKELM linked with the sequential scheme is utilized to predict the pull out trend and remainder elements, as the seasonal-naïve techniques are utilized for predicting the pull out seasonal element. At the end, the forecast outcomes of the seasonal, trend, and remainder elements are aggregated to create an accumulated outcome. The predicting executes of the three examined methods (i.e., STDL-OKELM, ELM, and SARIMA) across the 3 forecast horizons (i.e., H = 1, 2, and 3) to SMAPE and MASE are revealed in Table 4.

***Table 4** Prediction accuracy measures of different models Commodity: CRUDE OIL*

Accuracy Measure	*Prediction Horizon (h)*	*MODELS*		
		STDL-OKELM	*ELM*	*SARIMA*
SMAPE	*1*	*3.187*	*4.112*	*5.328*
	2	*3.101*	*4.127*	*5.789*
	3	*3.621*	*3.814*	*5.815*
MASE	*1*	*0.821*	*0.874*	*1.296*
	2	*0.827*	*0.882*	*1.352*
	3	*0.923*	*0.940*	*1.398*

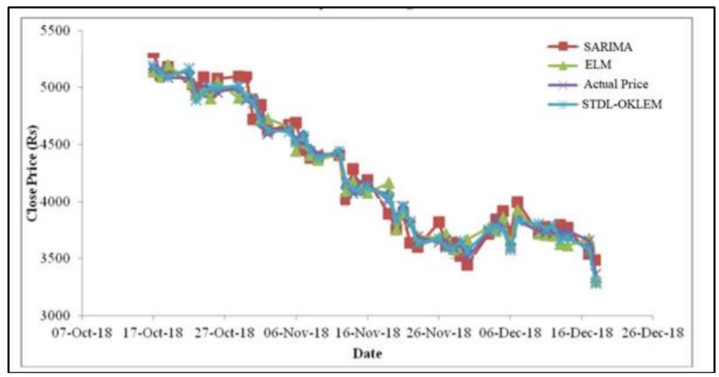

***Fig. 11**. Actual and predicted crude oil futures index closed prices using the ELM, SARIMA and proposed STDL_OKELM forecasting model.*

From Table 4 and Fig. 11, it is clear that the presented STDL-OKELM technique is the optimal one for volatile crude oil closed price index predicting across every forecast horizons (i.e., one-, two-, and three-step-ahead) for the crude oil closed index price series. In 2 methods except STDL-OKELM cannot model seasonality directly. So, the prior data processing, namely decomposition and seasonal adjustment, is essential and critical for creating an optimal predictor that is executed as the presented STDL-OKELM technique in this concept. In this approach, the SARIMA is behaved as the worst predicting concept for the crude oil price series, despite the forecast horizons and accuracy measures considered. On account of the above analysis of the experimental outcomes acquire in this study, one significant thing can be concluded. The presented STDL-OKELM technique is usually the optimal executing method, rather than other models listed in this approach, for crude oil price predicting with respect to the SAMPE and MASE criteria.

4 CONCLUSION

As the cause of economical characteristics and political instances, the price of crude oil seems to be volatile. The future closed price indexed data of crude oil which listed in MCX commodity market, India in Energy category are collected and they are used for analysis. This approach proposes a new hybrid method combining STDL and OKELM for seasonal forecasting of one-, two-, and three months crude oil prices in the MCX Commodity market, India. The experimental outcomes that are acquired shows that the proposed STDL-OKELM method brought out the optimal forecast execution relative to the mentioned competitors i.e. ELM and SARIMA for short to mid-term month based forecasting. Therefore, the proposed STDL-OKELM technique can be utilized as an effective tool for predicting crude oil prices with highly seasonal. Besides crude oil prices, the proposed STDL-OKELM method potentially executed to other complex seasonal predicting roles in the commodity market, namely copper and gold price predicting. Besides, these are considered to be most influential economic growth impact elements for every country.

5 REFERENCES

[1] A. Dieng , Alternative forecasting techniques for vegetable prices in Senegal, Revue sénégalais de recherchesagricoles et agroallementalress 1 (2008) 5–10 .

[2] G. Martin-Rodriguez , J. Caceres-Hernandez , Canary tomato export prices: com- parison and relationships between daily seasonal patterns, Span. J. Agric. Res. 11 (2013) 882–893 .

[3] G. Nasira , N. Hemageetha , Vegetable price prediction using data mining classification technique. Pattern Recognition, in: 2012 International Conference on: IEEE Informatics and Medical Engineering (PRIME), 2012, pp. 99–102 .

[4] P. Melin , A. Mancilla , M. Lopez , W. Trujillo , J. Cota , S. Gonzalez , Modular neural networks with fuzzy integration applied for time series forecasting, in: Analysis and Design of Intelligent Systems Using Soft Computing Techniques, Springer, 2007, pp. 217–225 .

[5] G. Nasira , N. Hemageetha , Vegetable price prediction using data mining classification technique. Pattern Recognition, in: 2012 International Conferen

[6] E.V. Colino , S.H. Irwin , P. Garcia , Improving the accuracy of outlook price forecasts, Agric. Econ. 42 (2011) 357–371 .

[7] S. Moshiri, and F. Foroutan, “Forecasting nonlinearcrude oil futures prices,” The Energy Journal vol. 27, pp. 81-95, 2005.

[8] S. Wang, L. Yu, and K. Lai. “Crude oil price forecasting with TEI@I methodology.” Journal of Systems Science and Complexity, vol. 18(2),145-166, 2005.

[9] Xie, W., Yu, L., Xu, S. and 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.

[10] Cleveland, R. B., Cleveland, W. S., McRae, J. E., &Terpenning, I. J. (1990). STL: A seasonal-trend decomposition procedure based on loess. Journal of Official Statistics, 6(1), 3–33. http://bit.ly/stl1990

[11] G. Huang, Q. Zhu, C. Siew, Extreme learning machine: theory and applications, Neurocomputing 70 (2006) 489–501.

[12] Dataset Archieve: http://www.mcxindia.com/products/energy/crude-oil

[13] V. Veeramanikandan, M. Jeyakarthic, Forecasting of Commodity Future Index using a Hybrid Regression Model based on Support Vector Machine and Grey Wolf Optimization Algorithm, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9, Issue-2, PP. No. 2856-2862December 2019.