

RESEARCH ARTICLE

A New Approach for Forecasting Crude Oil Prices Based on Stochastic and Deterministic Influences of LMD Using ARIMA and LSTM Models

JAWARIA NASIR¹, MUHAMMAD AAMIR¹, (Senior Member, IEEE), ZAHOOR UL HAQ², SHEHZAD KHAN², MUHAMMAD YUSUF AMIN², AND MUHAMMAD NAEEM¹

¹Department of Statistics, Abdul Wali Khan University Mardan, Mardan 23200, Pakistan

²Pakhtunkhwa Economic Policy Research Institute (PEPRI), Abdul Wali Khan University Mardan, Mardan 23200, Pakistan

Corresponding author: Muhammad Aamir (aamirkhan@awkum.edu.pk)

ABSTRACT Crude oil is one of the non-renewable power sources and is the lifeblood of the contemporary industry. Every significant change in the price of crude oil (CO) will have an effect on how the global economy, including COVID-19, develops. This study developed a novel hybrid prediction technique that depends on local mean decomposition, Autoregressive Integrated Moving Average (ARIMA), and Long Short-term Memory (LSTM) models to increase crude oil price prediction accuracy. The original data is decomposed by local mean decomposition (LMD), and the decomposed components are reconstructed into stochastic and deterministic (SD) components by average mutual information to reduce the computation cost and enhance forecasting accuracy, predict each individual reconstructed component by ARIMA, and integrate the residuals with LSTM to capture the nonlinearity in residuals and help to find the final prediction result. The new hybrid model LMD-SD-ARIMA-LSTM has reduced the volatility and solved the issue of the overfitting problem of neural networks. The proposed hybrid technique is validated using publicly accessible data from the West Texas Intermediate (WTI), and forecast accuracy are compared using accuracy measures. The value of Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) for ARIMA, LSTM, LMD-ARIMA, LMD-SD-ARIMA, LMD-ARIMA-LSTM, LMD-SD-ARIMA-LSTM, and Naïve are 1.00, 1.539, 5.289, 0.873, 0.359, 0.106, 4.014 and 2.165, 1.832, 9.165, 1.359, 1.139, 1.124 and 3.821 respectively. From these results, it is concluded that the proposed model LMD-SD-ARIMA-LSTM has minimum values for MAE and MAPE which assured the superiority of the proposed model in One-step ahead forecasting. Moreover, forecasting performance is also compared up to five steps ahead. The findings demonstrate that the suggested approach is a helpful tool for predicting CO prices both in the short and long term. Furthermore, the current study reduces labor costs by combing the stationary and non-stationary Product Functions (PFs) into stochastic and deterministic components with improved accuracy. Meanwhile, the traditional econometric model can strengthen the prediction behavior of CO prices after decomposition and reconstruction, and the new hybrid forecasting method has better performance in medium and long-term forecasting of the CO price. Moreover, accurate predictions can provide reasonable advice for relevant departments to make correct decisions.

INDEX TERMS ARIMA, LMD, LSTM, stochastic and deterministic influences, crude oil prices.

I. INTRODUCTION

As the “lifeblood of the industry,” crude oil (CO) is the most significant strategic component raw element in contemporary

The associate editor coordinating the review of this manuscript and approving it for publication was Huaqing Li¹.

industrial society, the key to prosperity and national security, and the cornerstone of civilization [1]. It is connected to how the global economy is growing, and different economic data will be more significantly impacted by changes in oil prices. Predicting how the price of crude oil will fluctuate in the future is therefore quite important. Scholars have used a

range of research techniques to perform in-depth analyses and forecasts of global oil prices from the perspective of diversification. The artificial intelligence model and the econometric model, and the integrated forecasting model are roughly the three components of these research methodologies, according to the summary of the literature examined. For the models in econometrics The authors in [2] examined how well various ARIMA-GARCH (Generalized Autoregressive Conditional Heteroscedasticity) models performed in simulating and predicting the conditional mean and volatility of weekly crude oil prices. The researchers in [3] apply the findings of the ARIMA model compared with those of the Decomposition-based vector autoregressive model (DVAR), which is used to forecast the monthly price data for WTI crude oil. The authors in [4] examined the predictive power and impact of the Google index on CO price by incorporating it into the ARMA-GARCH and ARIMA models. By contrasting the suggested MMGARCH (Mixture Memory GARCH model) to other discrete unpredictable models. Klein and Walther [5] extend the literature on predictable and esteem predicting CO price returns. Stelios et al [6] compare the VAR model's capacity for forecasting to that of the Random Walk (RW) model and the AR model. The results are listed at the top of Table 1. Generally, the economic model assumes that the data are stable, regular, and linear. Under this assumption, the economic model can accurately predict the CO price. The international crude oil market, however, exhibits complex non-linear, and multidimensional characteristics of crude oil price movements. The intricate features concealed in crude oil may be too complicated for these conventional metering methods to detect. Support vector regression (SVR), Artificial Neural Network (ANN), Random Forest (RF), and other widely used non-linear models are applied to CO price forecasting, successfully fitting the non-linear CO price series as a result of the rapid growth of artificial intelligence. For example, the researchers in [7] utilized a neural network and genetic algorithm to predict the price of WTI CO. In the same way, the authors in [8] utilized a neural network to forecast the price term structure of crude oil futures. Fan et al [9] utilize an Imperialist Competitive Algorithm and Support Vector Regression (ICA-SVR) techniques to forecast the price of crude oil. Mostafa and El-Masry [10] projected CO price using gene expression programming (GEP); CO price is forecasted by Gao and Yalin [11] based on stream learning. The artificial intelligence model outperformed the conventional paradigm, in line with empirical research. A single artificial intelligence (AI) model cannot correctly represent the dynamic changes of complicated CO price time series responsible for the significant variations in the time series. However, an AI model can accurately anticipate non-linear and non-stationary sequences. However, the hybrid forecasting model overcomes the drawbacks of time series instability and nonlinearity and enhances the CO price prediction accuracy by combining a range of methodologies. In the past few years, integrated models for predicting the price of CO have

developed quickly. Tao et al [12] suggested a more effective EMD-SBM-FNN model that can capture the intricate dynamics of the price of crude oil. Zhang et al [13] introduced EEMD-PSO-LSSVM-GRACH a novel hybrid approach to forecast CO prices. Yu et al [14] used the EEMD-DCD-LSSVR model to predict the price of CO. The authors in [15] estimated the price of CO, by using bootstrap aggregation (bagging) and Stacked Denoising Auto Encoders (SDAE). In the same way, the authors in [16] use the EEMD-RVFL model to predict the price of CO. Moreover, the authors in [17] used the EEMD-EELM-ADD model as a unique decomposition-ensemble technique for predicting CO prices. Ding [18] created a hybrid model EEMD-ANN-ADD for predicting the price of CO. The authors in [19] use the DFN-AI model to predict the CO price. Similarly, the authors in [20] use the VMDICA-ARIMA hybrid model to predict the price of CO. In the same way, Zhang et al [21] suggested an algorithm for iterated combinations to predict the CO price. The authors in [22] combine RW and ARMA to predict the CO price. Similarly, Zheng et al [23] proposed EEMD and Dynamic Artificial Neural Network (DANN) to forecast the CO price. The authors in [24] showed load prediction, based on the long short-term memory (LSTM) model, based on Back Propagation Neural Network (BPNN) and Local Mean Decomposition (BPNN-LMD-LSTM). The design is based on a fixed-time consistency algorithm with random delay to predict the economic dispatch of microgrids. The authors in [25] proposed a landslide displacement prediction model, the local mean decomposition-bidirectional long short-term memory (LMD-BiLSTM), which depends on the time-frequency analysis method. The authors Heng Sun [26] utilizes method in three steps exhibits great potential applications in the RUL prediction of rotating machines. The authors in [27] LSTM, wavelet threshold denoising (WTD), and LMD have been integrated into a novel combined model called LMD-WTD-LSTM to estimate short-term gas consumption. In the same way, the authors in [28] introduced a new model which enhanced the accuracy of the predictions. The novel technique called variational mode decomposition (VMD) and used to predict the major factor time series utilizing its secondary factors. A new technique called multiscale forecasting model is introduced that produced an optimal forecast [29]. This model outperformed the compared model to forecast the complex time series data. In the same way, the authors in [30] decomposed the data into many features via VMD. Then the mutli-features are trained with machine learning classifiers. The authors in [31] forecasted the Daily PM2.5 and PM10 data employing a Robust LMD (RLMD) and moving window ensemble technique was done using linear and nonlinear modelling frameworks. The research mentioned above claims that the hybrid model mixes single models so that the benefits of each model balance out the drawbacks of the other models. As a result, the hybrid model is superior to the single model and offers us research suggestions. From the above discussions

the following research questions have been generated. How can the end-point impact be eliminated due to the complicated dynamic change of the crude oil price time series and additional information obtained on different frequencies of the crude oil price data itself? How the calculations for the hybrid model may be streamlined. How can CO price predictions be made more accurately? In the current study, we use Local Mean Decomposition (LMD) and an artificial intelligence model to forecast the price of CO: (1) Utilizing LMD to decompose the time series of CO price in an adaptive manner, removing the end-point effect, and further exploring the data's various frequencies. (2) This work uses average mutual information (AMI) to decrease the calculation amount while taking into account the growth of the hybrid model's calculation amount; (3) By separating the time series into random and stochastic variables, the econometric model is able to represent the volatility features of the crude oil price time series; (4) combining the outcomes of the prediction utilizing LSTM; (5) The experimental results demonstrate that the LMD-SD-ARIMA-LSTM suggested in this study outperforms a single model in terms of crude oil price prediction accuracy. They also demonstrate that the traditional econometric models can increase prediction accuracy through decomposition and aggregation. Researchers are still working on these problems. In comparison to previous studies, the novel hybrid model LMD-SD-ARIMA-LSTM has reduced the volatility and solved the issue of the overfitting problem of neural networks. The proposed hybrid technique is validated using publicly accessible data from the West Texas Intermediate (WTI), and forecast accuracy are compared using accuracy measures.

The organization of the study is as follows. Section I consists of an introduction and a literature review. Section II provides a brief description of the methods used in this study. In the same way, sections II-D and III consist of analysis and discussion along with a conclusion respectively.

II. METHODOLOGY

A. LOCAL MEAN DECOMPOSITION

Using adaptive time-frequency analysis, LMD is a technique for handling non-stationary signals [32]. Separating various envelope signals and purely frequency-modulated signals from the original signals is the foundation of the LMD approach. A physical significant product function (PF) component of the instantaneous frequency can be derived by multiplying the envelope signals with sole frequency modulated signals. The decomposition procedure for the initial signal $x(t)$ can be broken down into five steps:

- i Select all local extremum points n_i of the original signal $x(t)$ and calculate the mean m_i of adjacent extremum points n_i, n_{i+1} and envelope estimate α_i :

$$\alpha_i = \frac{|n_i - n_{i+1}|}{2} \tag{1}$$

The envelope estimate α_i and local means m_i are then used to smooth using the moving average to $m_{11}(t)$ and envelope estimate function $\alpha_{11}(t)$;

- ii Ignore the local mean function $m_{11}(t)$ in the original signal $x(t)$, that is:

$$h_{11}(t) = x(t) - m_{11}(t) \tag{2}$$

- iii Dividing it by $\alpha_{11}(t)$, $h_{11}(t)$ is the amplitude demodulated.

$$s_{11}(t) = h_{11}(t)/\alpha_{11}(t) \tag{3}$$

purely frequency modulated signal is Repeated iterations n times until $s_{1n}(t)$. Stopping iteration should be done when $\lim_{n \rightarrow \infty} \alpha_{1n}(t) = 1$.

- iv The corresponding envelope $\alpha_1(t)$ and the first component $PF_1(t)$ are obtained:

$$\begin{aligned} \alpha_1(t) &= \alpha_{11}(t)\alpha_{12}(t) \dots \alpha_{1n}(t) \\ &= \prod_{q=1}^n \alpha_{1q}(t) \end{aligned} \tag{4}$$

$$PF_1(t) = \alpha_1(t)s_{1n}(t) \tag{5}$$

- v The component $PF_1(t)$ is removed from $x(t)$, the new signal $u_1(t)$ is obtained and the process is repeated k times until the signal $u_k(t)$ is a constant the oscillations have stopped, too. Finally, the original signal $x(t)$ can be written as

$$x(t) = \sum_{p=1}^k PF_p(t) + u_k(t), t = 1 \dots n \tag{6}$$

B. ARIMA

Box and Jenkins first proposed the ARIMA model in the early 1970s. The following structure of the model is said to be the autoregressive integrated moving average model, which is defined as ARIMA (p,d,q): The dependent variable must be stationary (through the I-component), and the independent variables are taken as all lags of the dependent variable (the AR-component) and/or errors lags. In general, therefore, one might consider an ARIMA model to be a specific kind of regression model (the MA component).

In general, a model with ARMA looks like this:

$$Y_t = \alpha_0 + \sum_{i=1}^p \alpha_i Y_{t-i} + \sum_{j=1}^q \beta_j \varepsilon_{t-j} + \varepsilon_t \tag{7}$$

The AR coefficient and constant term are represented by p , whereas the MA coefficients are represented by q . The following steps are involved in modelling the ARIMA (p,d,q) model: First, the observation sequence's stationaries are tested. If the observed sequence is not stationary, a difference in times d must be used to convert the sequence into a stable time series. Second, after the difference, the stationary sequence is subjected to the white noise test when the observed sequence is stationary. The ARMA (p,q) model is fitted if the test result is a sequence that is not white noise. ACF and PACF can be used to determine p and q . The analysis is over if the test yields a white noise sequence. The fitted

TABLE 1. Descriptive statistics for the WTI crude oil prices.

min	median	max	MEAN	SD	SKEWNESS	KURTOSIS	ADF	J-B	PP
12.17	59.17	95.52	58.082	14.504	-0.530	3.399	-1.066 (0.928)	4.369 (0.497)	-1.941 (0.969)

Note: J-B is the Jarque-Bera Statistic; ADF is the augmented Dickey-Fuller Statistic; PP is the Phillips-Perron Statistic; * indicates rejections of the null hypothesis at the 1% significance level

ARMA (p,q) model’s residual sequence is next checked for white noise. The ARMA (p,q) model is re-fitted if the test result is a non-white noise sequence. If not, the analysis is over [33].

C. LSTM

It is a type of RNN with the capacity to take long-term dependability into account. Scientists The authors in [34] developed the LSTM in 1997. Because LSTMs can retain information over a longer amount of time and do not have long-term dependencies, they differ from other RNN techniques. Inside, LSTMs operate similarly to other RNN methods employing neural network gates and layers. They have a chain structure. The LSTM’s construction is designed to have a cell that runs the length of the device. Gates are used to control whether or not data may be transferred into the cell state. Additionally, there are parts known as gated cells that enable the storage of data from earlier LSTM outputs; this is the place where the memory-related aspects of LSTM come into play.

An advanced soft computing technique known as LSTM was developed from the Recurrent Neural Network (RNN). One of the numerous Artificial Neural Network (ANN) techniques, the RNN, was developed to address the ANN’s weakness in handling time correlation in the data sequence and enhances neurons in the networks with canonical connections to make it possible for RNN to create a sequence-to-sequence mapping between input and output data [24]. Unfortunately, the long-range dependencies are still a challenge for traditional RNN, which have difficulty learning the long-term temporal correlations due to expanding gradients or, conversely, vanishing gradients [25]. The authors in [34] used LSTM memory cells to get around this restriction. These cells use a three-gate mechanism made up of an input gate, an output gate, and a forget gate to store the temporal state of the networks [35]. Figure 1 shows an LSTM cell with all three of those gates as well as the cell state [36].

LSTM gates are simply used to limit the amount of information that can be transferred. They typically consist of a layer of a sigmoid neural network and an action of pointwise multiplication. While the forget gate is used to use selectively forget the information in the cell state, the input gate decides what new information will be stored in the current cell state. The output gate is then utilized to determine the value that we wish to output [22]. The forget gate is the initial component

of the LSTM cell. It can be expressed as follows and is used to regulate the magnitude to forget the previous cell’s concealed state:

$$f_t = (W_f h_{t-1} + U_f x_t + b_f), \tag{8}$$

where f_t denotes the value of the forgetting gate at the current cell, it ranges from 0 (completely forget) to 1 (completely keep) and W_f, U_f shows the network’s weights; b_f shows the value of the bias variable, h_{t-1} shows the prior hidden value, whereas x_t denotes the new input value at the current cell. The state of the cell is then updated using the input gate. This stage will involve two acts. First, we pass the previously hidden state value (h_{t-1}) to the input gate and the current input value (x_t) is shown in a sigmoid function in Eq. (9). The output of the input gate (i_t) determines how much extra information will be maintained in the current cell, where 0 denotes “totally disregard” and 1 denotes “completely keep”. Second, to aid with network regulation as stated in Eq, we also provide the previous hidden state (h_{t-1}) and the current input (x_t) into tanh function Eq. (10). Similarly, when it comes to the forget gate, there are some weights of the networks and bias values involved in this step, as denoted by W_i, W_c, U_i, U_c, b_i , and b_c .

$$i_t = (W_i h_{t-1} + U_i x_t), \tag{9}$$

$$C_t = \tanh(W_c h_{t-1} + U_c x_t + b_c). \tag{10}$$

The (current) cell state (C_t) may now be calculated to the information we currently have. The forget vector (f_t) will be pointwise multiplied by the preceding cell state (C_{t-1}). The output of the input gate (i_t), which has been pointwise multiplied with the cell candidate value (\tilde{C}), is then added to, as given in Eq. (11).

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}. \tag{11}$$

The output gate is used to determine the next hidden state in the last step (i.e., the value of the current hidden state, h). First, we use the sigmoid function as given in Eq to transmit the previously hidden state value (h_{t-1}) and current input value (x_t) into the function (12). Here, W_o, U_o and b_o are the corresponding networks weights and bias values for the output gate. The output of the tanh function is then pointwise multiplied with the sigmoid output from the output gate (O_t), as described in Eq. (12), using the newly discovered cell state (C_t) as input Eq. (13). The result of this final step is the value

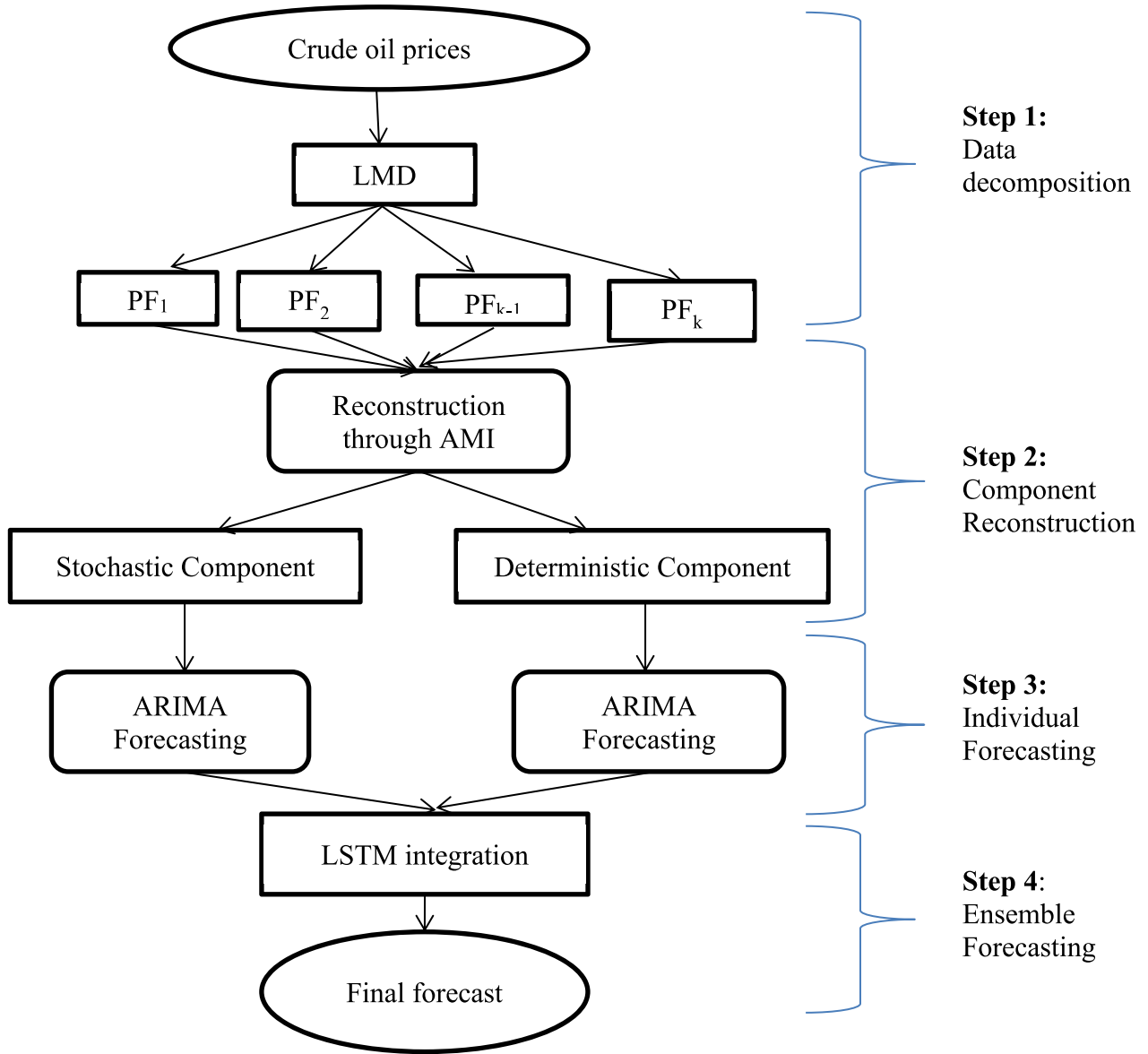


FIGURE 1. The flow chart of the proposed model.

of the current hidden state (h_t), which is sent along with the current cell state (c_t) to the following time step.

$$O_t = (W_o h_{t-1} + U_o x_t + b_o) \quad (12)$$

$$h_t = O_t \odot \tanh(C_t) \quad (13)$$

D. EVALUATION

In this study, metrics are assessed using three different prediction error criteria: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). The last one displays the level of inaccuracy in a percentage number, whereas the first two display the error level. As shown by Shahid et al in [37] and Hansun et al in [36] and [38], All three of those requirements can be

stated as:

$$MAE = \frac{1}{n} \sum_{t=1}^n |Y_t - F_t| \quad (14)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2} \quad (15)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|Y_t - F_t|}{Y_t} \quad (16)$$

where n shows the total number of values, the actual value is denoted by Y_t , and F_t shows the predicted value. Next, the Diebold Mariano statistic (DM) is used to compare the performance of the two forecasting errors [39], [40] and is defined as follows:

$$DM = \frac{\bar{k}}{\sqrt{Var(\bar{k})}} \quad (17)$$



FIGURE 2. Daily crude oil prices of the WTI market (Dollar per Barrel).

where $\bar{k} = \frac{1}{T} \sum_{t=1}^n k_t$, $k_t = (Y_t - F_{1,t})^2 - (Y_t - F_{2,t})^2$, $Var(\bar{k}) = \frac{1}{T}(z_0 + 2 \sum_{n=1}^k z_n)$ and $z_n = Cov(k_t, k_{t-n})$. $F_{1,t}$ signifies the first model predicted values, while $F_{2,t}$ representing the second model's predicted values.

III. EMPIRICAL ANALYSIS

A. STATISTICAL DESCRIPTION OF DATA

West Texas Intermediate (WTI) data from the U.S. Energy Information Administration (EIA, <http://www.eia.doe.gov>) in this study, as depicted in Figure 1. The datasets for WTI crude oil from March 12, 2018, to February 14, 2022, having a total of size 999 samples were chosen. For forecasting accuracy, the data is divided into 75/25 training and testing parts respectively. The training part of a total of 749 samples comprised from March 12, 2018, to February 23, 2021, and the testing part of size 250 is from February 24, 2021, to February 14, 2022. A dramatic downward trend in the price of WTI can be seen in Figure 1 graphical analysis up until 11 January 2022. This is followed by another downward trend in the price of WTI up until 11 May 2020, which was seen up until 30 May 2020. It is simple to spot a most pronounced upward trend from May 12, 2020, to February 14, 2022. In contrast to other months, the seasonal plot in Figure 2 demonstrates that the annual variance is still minimal from January to May for year 2020 and from September to December for the year 2018. PP and ADF statistics are shown in Table 1 demonstrating

the sequence is not stationary. Jarque-Bera (J-B) statistics, however, indicate that the data is normal.

B. DATA DECOMPOSITION AND RECONSTRUCTION

This study proposed the reconstruction of PFs that are obtained from LMD. PFs from LMD is separated into two components, Deterministic and Stochastic. The stochastic and Deterministic PFs are modeled separately, and different models are selected for stochastic and deterministic components, ARIMA and LSTM models are then fitted for every stochastic and deterministic component all individual deterministic PFs are then combined for final forecasting. Figure 3 illustrates the suggested method's entire framework. WTI price data is decomposed by LMD. For the LMD method, The maximum number of generating function components is 20 and the maximum number of iterations is set to 30. The decomposition results demonstrate WTI CO price sequence is composed of three PFs and one residual.

The CO price sequence is divided into three PFs and a residual series, with frequency changes from high to low and a decreasing changing trend for the residual series, as shown in the aforementioned Figure 3. Then, considering the different influences of the decomposed PFs on the original series, the PFs are reconstructed according to the mutual information of each PF. AMI is also used to recreate the IMFs. AMI is only a visual assessment of the plots that are produced for each dataset's PFs, as shown in Figure 3, From the second

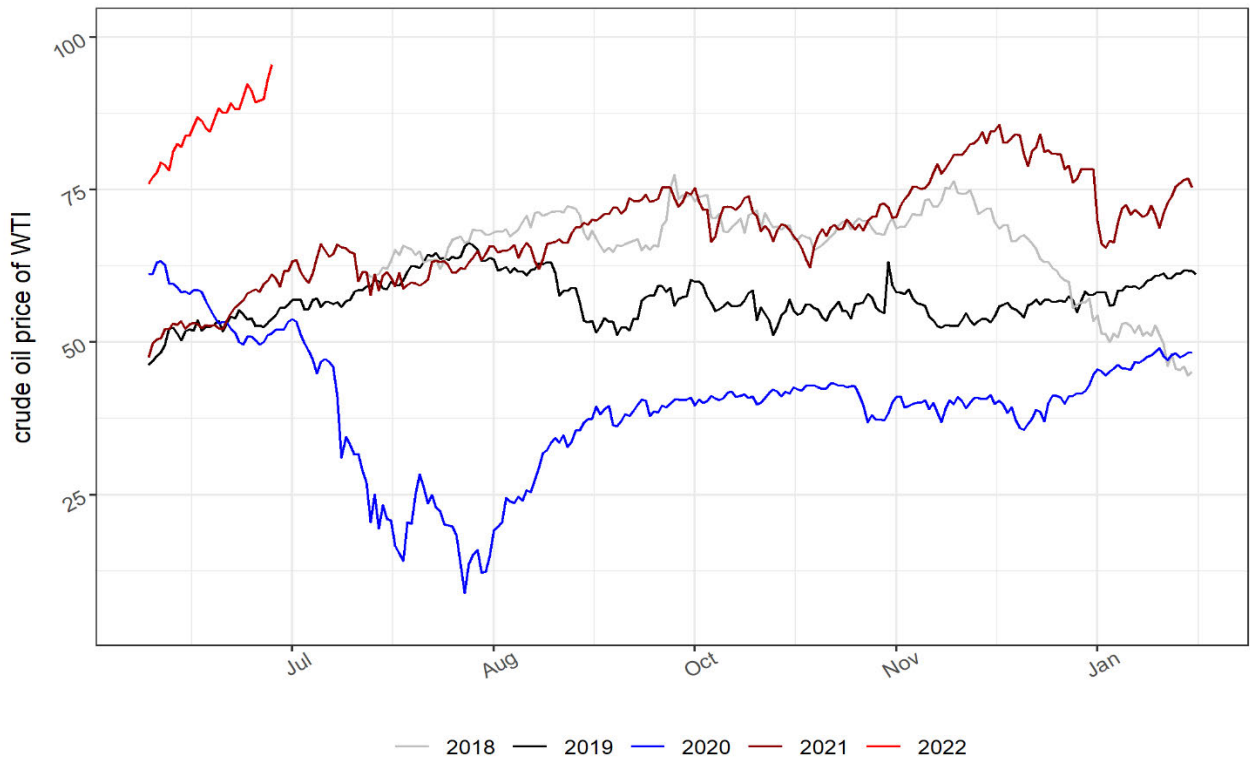


FIGURE 3. A seasonal plot of WTI from 2018-2022.

TABLE 2. ADF-test values for WTI, all PFs, and VPF.

	PF1	PF2	PF3	PF4	VPF	WTI
Adf-test	<0.0001	0.0003	0.1645	0.5007	0.7939	<0.0001

PF to the fourth PF, the AMI plots are seen to follow the same pattern, so the first four PFs are shown in Figure 4, As a result, the first PF is regarded as stochastic, while the rest are all deterministic. To create two components, add from second to fourth as deterministic and the first PF for stochastic. The Mutual information is calculated according to Eq. (9), by adding all deterministic series new series is formed called virtual product functions (VPF) is obtained and displayed in Figure 5. Stochastic and deterministic graphs are displayed in Figure 6.

Here, PF1 is considered to be a stochastic component, while PF2, PF3, and PF4 are treated as deterministic. Stochastic and deterministic (SD) components are treated separately and then combined to make SD components.

1) ARIMA MODEL

The basic assumption to apply the ARMA model is the stationarity of the time series model. Successive differences are taken to make time-series stationary [41], [42]. ADF test is used to determine whether the time series is stationary [43]. After obtaining a stationary time series appropriate model

for the time series by choosing the AR and MA words in the right order For selection, ACF and PACF plots of the best order are employed. The Adam algorithm is used in the Model for training and cross-validation. A 75:25 ratio of data is designated for training and testing, accordingly. Ljung-Box (LB) test is used for checking the model adequacy. Forecasting is done by fitting the ARIMA model to training and testing periods. By fitting data with the ARIMA model, we used R software. Figure 7 displays the residual of the WTI ARIMA fitted model with ACF lag and Residual plot.

The first and second plot of the above figure shows that residuals are uncorrelated, and data became stationary after taking lag1. Whereas the final LB statistics figure demonstrates that p-values are greater than 0.05 for all datasets [44]. A hypothesis shows no serial autocorrelation among the fitted residual is not significant. So, All of these techniques can offer the most accurate future projection. The forecasting accuracy of the models is shown below.

PF1 and PF2 test thinks that the residuals are distributed normally, and the series is stationary. However, PF3, PF4, and VPF need to be differenced. PF3 and PF4 have a difference

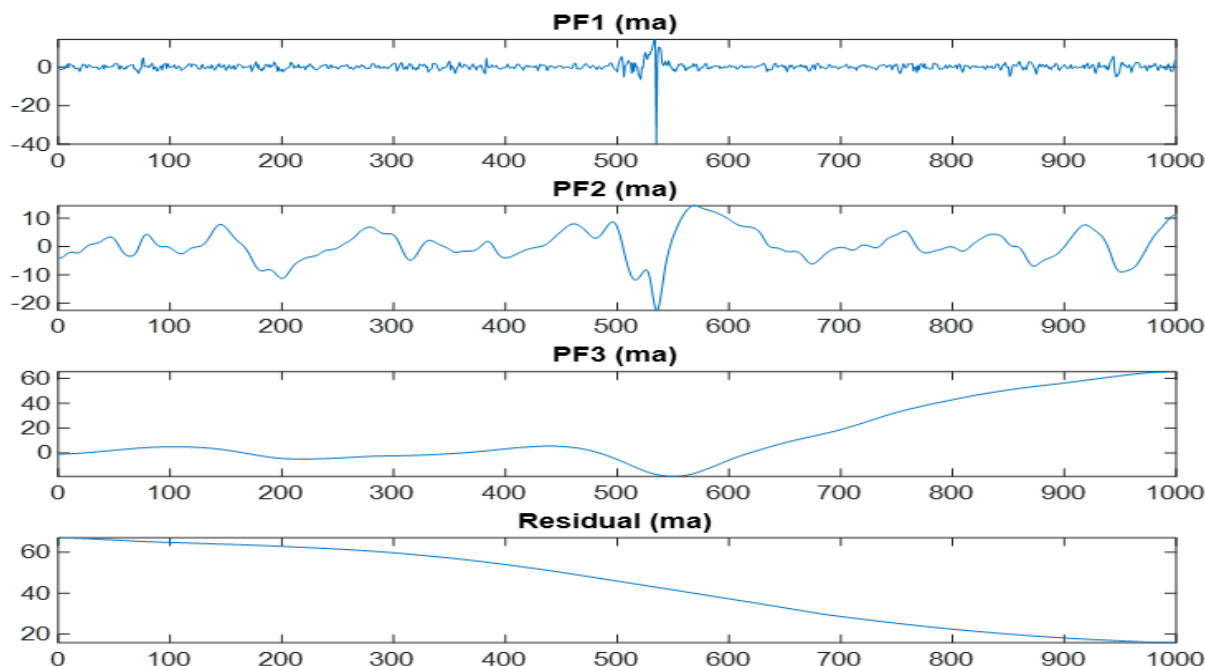


FIGURE 4. The PFs and residual plots for the WTI Crude oil price by LMD.

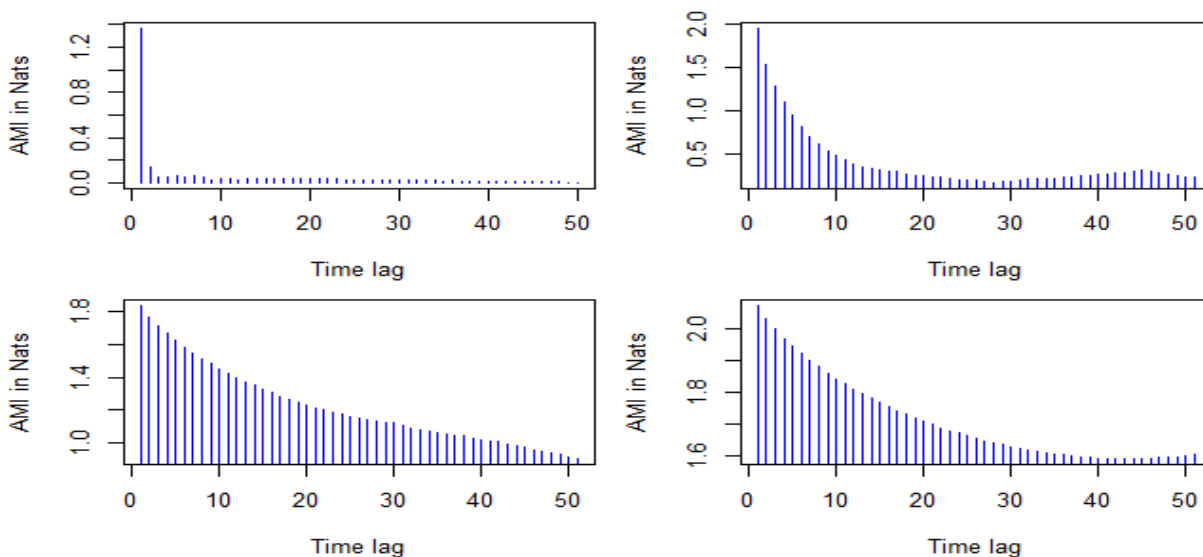


FIGURE 5. Average mutual information (AMI) Plots for PFs.

of order 2 and the difference order number of VPF is 1. Table 3 Box test results indicate that the p-value is greater than 0.05. This amount exceeds 0.05 as a significance level and hence the hypothesis that the residuals are white noise could not be rejected. Accept the supposition that there is no residual autocorrelation. ARIMA is successfully established. All four PFs are fitted for the LMD-ARIMA model whereas PF1 (stochastic component) and VPF(deterministic component) are added and fitted for the LMD-SD-ARIMA model.

2) STACK-LSTM

The residual of LMD-ARIMA and LD-SD-ARIMA is processed in a hybrid stack LSTM model for forecasting. In this technique output of the LSTM layer becomes the sequence of vectors to be used as an input to the subsequent LSTM layer. It has been seen in the Figure 8 input layers are used again in the second layer and multiple hidden layers are stacked one on top of another hidden layer. A three-dimensional input layer is required for the LSTM layer. RETURN_SEQUENCE=TRUE enables using the buried

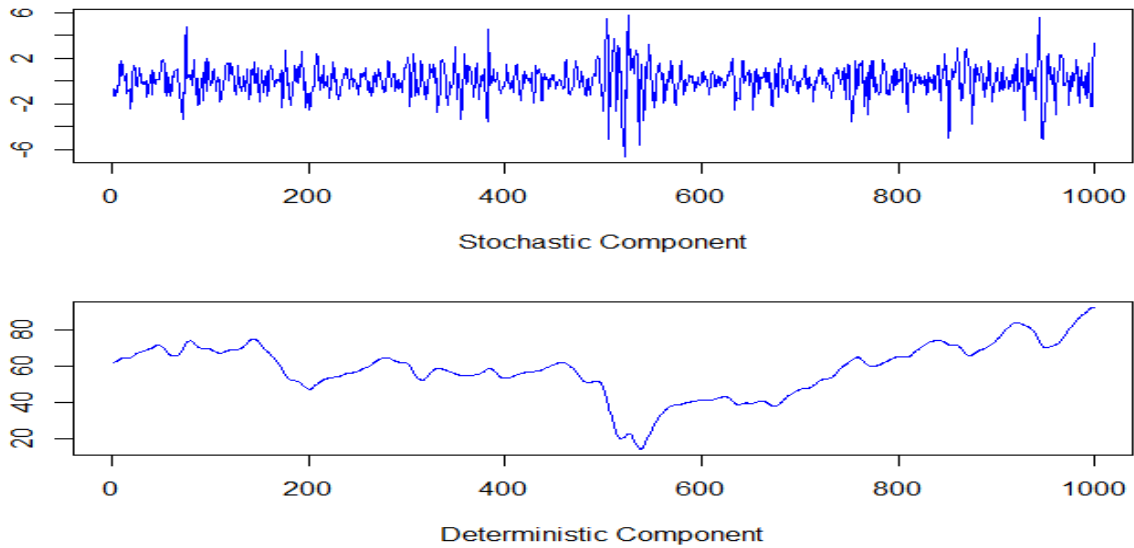


FIGURE 6. Stochastic and deterministic components plots.

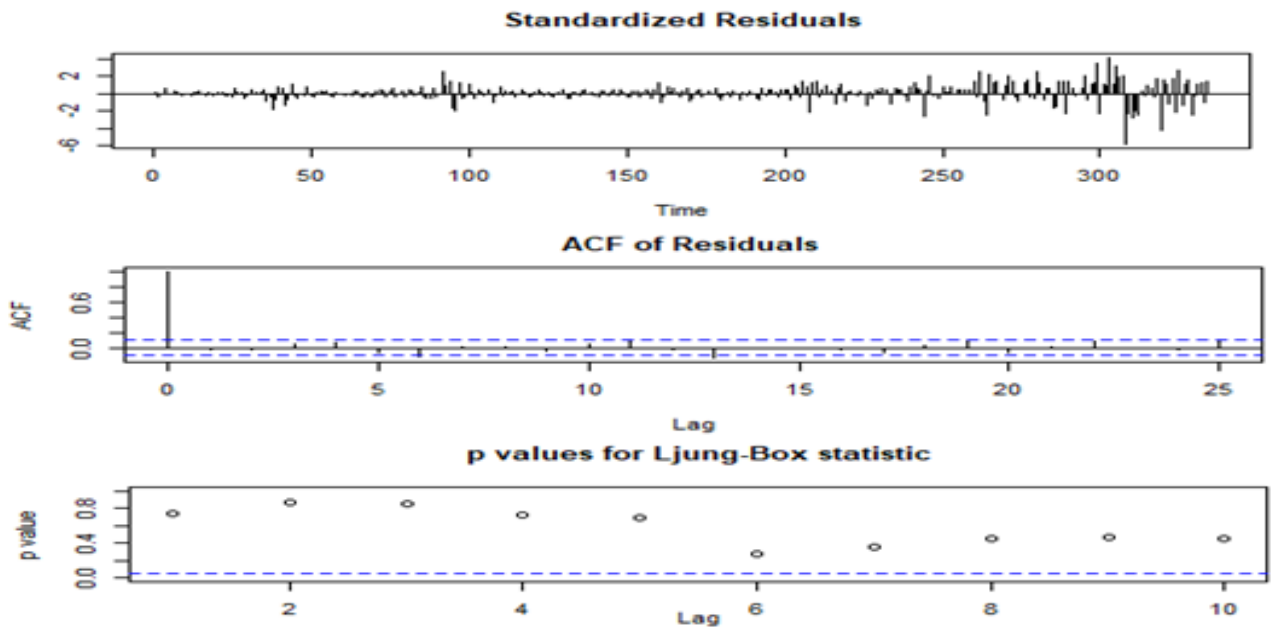


FIGURE 7. ARIMA Fitted model for WTI.

LSTM layer’s 3D results as the next layer’s input. It makes the model deeper and more accurate as a deep learning technique. M networks are made up of many LSTM hidden layers. A Deep Recurrent Neural Network is formed by several hidden layers (DRNN). Iterative weight updating utilizing training data is crucial for training the LSTM network. The stochastic gradient descent approach employs the Adaptive Movement Estimation (Adam) algorithm for weight updating [45]. For constructing and training the LSTM model Keras neural network API [46] uses Tensorflow and [47] developed Tensorflow library written in Python. For

numerical calculations, Tensorflow is employed as a machine learning framework. In contrast, Keras has a steep learning curve. Together, they deepen and improve the model’s accuracy.

In this paper, in our dataset, we represent each trading day of every month as a 1×20 input vector, where the number of features we would use in prediction is 20. The 3D vector forms the shape ($W \rightarrow l \rightarrow f$), where W represents the number of windows, l shows the length of a window, and f denotes the number of features, which is built in order to make output calculation easier. A maximum number of

TABLE 3. Accuracy measures of All PFs, VPF, and WTI crude oil prices.

ARIMA	AIC	AICc	BIC	RMSE	MAE	MAPE	Box-Ljung
PF1(2,0,1)	3043.03	3043.09	3067.56	1.18635	0.8632	226.6182	0.7526
PF2(5,0,0)	6020.73	6020.74	6025.64	0.006593	0.0046	0.5071	0.9472
PF3(4,2,0)	-11019.95	-11019.89	-10995.43	0.00238	0.0015	0.05751	0.0910
PF4(1,2,1)	-11545.94	-11545.91	-11531.22	0.00659	0.0046	0.5071	0.9472
VPF(4,1,2)	-7340.06	-7339.91	-7300.81	0.0067	0.0045	0.0104	0.0800
WTI(0,1,0)	3549.7	3549.71	3554.61	1.461871	1.0030	2.165	0.4469

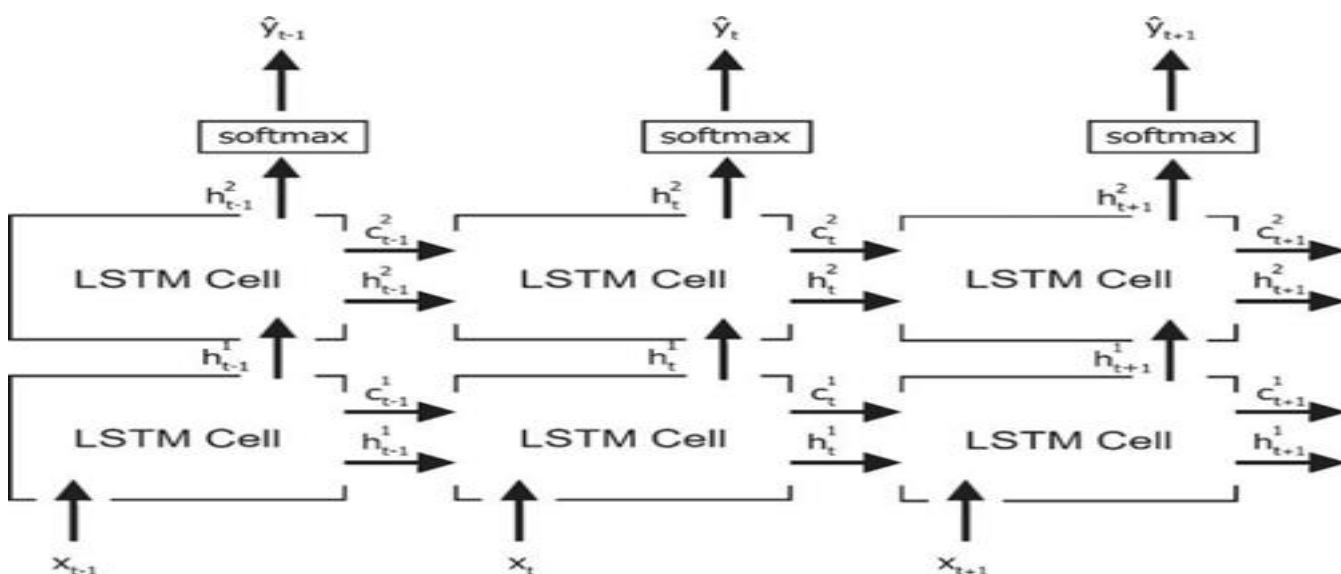


FIGURE 8. Stacked LSTM model.

neurons $(2*k+1)$ is selected to define the number of hidden layers, where k represents a number of inputs as the authors in [48] also show three layers make up the LSTM model: one layer followed by two hidden layers, and one dense layer. The first hidden layer’s output is connected to a second hidden layer, the third hidden layer, which is then coupled to a dense layer. Figure 9 depicts the relationship between each layer. After taking each hidden layer, for reducing the chance of overfitting dropouts are utilized. The authors in [49] define LSTM layers with trainable and non-trainable parameters as we generated in Table 4.

Hyperparameters that will increase accuracy and reduce the chance of overfitting the data are chosen [49]. The overfitting drop-out technique, which randomly selects cells within a layer based on the probability in such a way that their output is set to 0, is used to reduce overfitting as developed by Srivastava et al [50]. As a result, the ideal dropout is chosen at 30%, which gives the lowest MSE. Epoch was selected to 100 while performing this test, with one epoch equaling one iteration of all training data that was processed

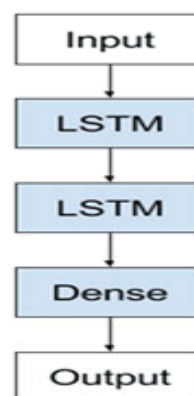


FIGURE 9. LSTM layers.

by the network in [51]. Each layer’s LSTM cells were programmed to have the following values: 41, 41, 64, 1, decay to 0.2, and window length to 22. We divided the training data into a batch size, with a batch size of 32, to facilitate the propagation of the training data across the

TABLE 4. LSTM layers and Parameters (Model: "sequential").

Layer (type)	Output Shape	Parameters
Lstm (LSTM)	(None, 20, 41)	7052
lstm_1 (LSTM)	(None, 20, 41)	13612
lstm_2 (LSTM)	(None, 41)	13612
Dense (Dense)	(None, 1)	42

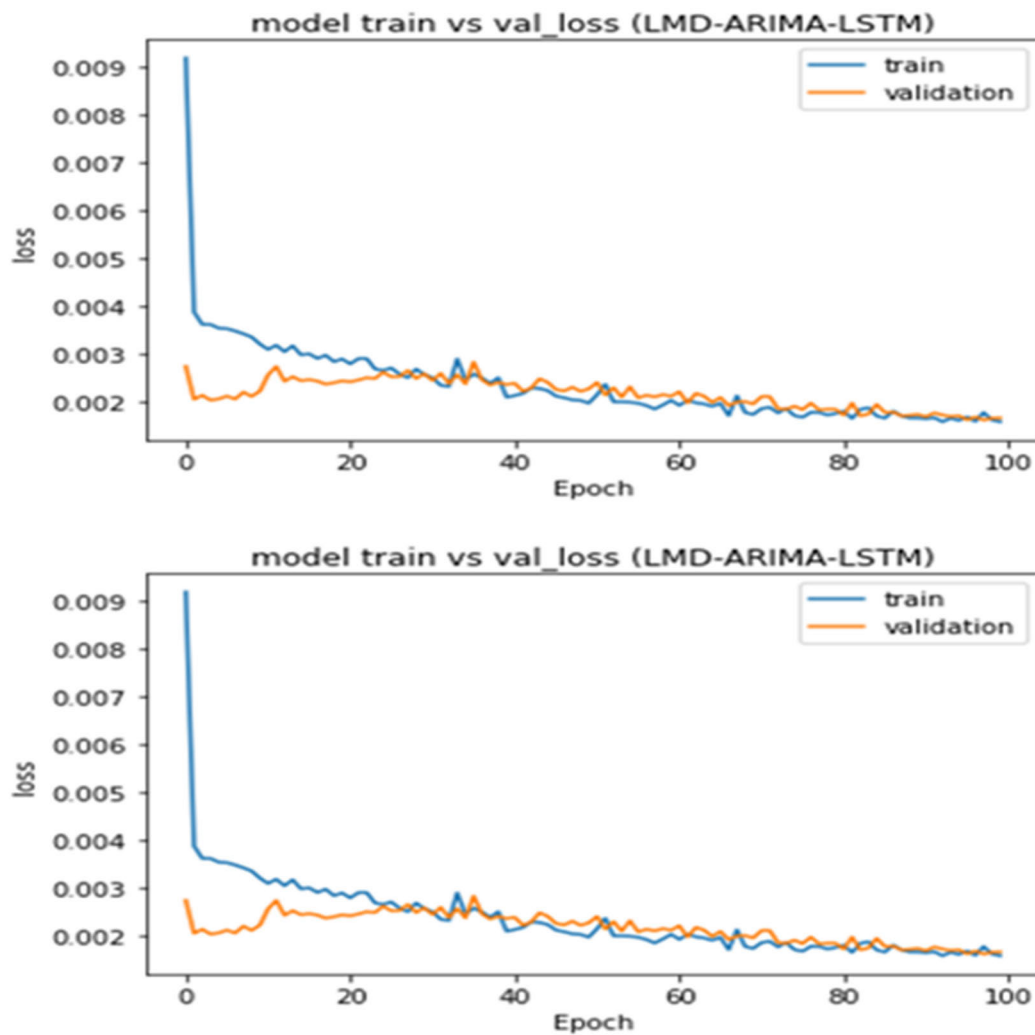


FIGURE 10. Validation plots for proposed models.

network. This indicates that the network is trained using the first 31 examples (0–31) from the training data, followed by the subsequent 31 samples (32–63). One epoch has been transmitted through the network once all samples have been propagated through it, hence the epoch continues until that happens [49].

In order to hasten learning, the loss function is used to calculate the difference between the desired output and the LSTM model output during the training. User-specified validation data, which we have defined to be 10% of the training data, is what we want as the output. We have taken the MSE as a loss function because it is frequently utilized for

TABLE 5. Prediction Accuracies of WTI.

Model	MAE	RMSE	MAPE
ARIMA	1.003	1.461	2.165
LSTM	1.539	1.793	1.832
LMD-ARIMA	5.289	7.314	9.165
LMD-SD-ARIMA	0.873	1.013	1.359
LMD-ARIMA - LSTM	0.359	0.503	1.139
Naïve	4.014	5.269	3.821
LMD-SD-ARIMA-LSTM	0.106	0.150	1.124

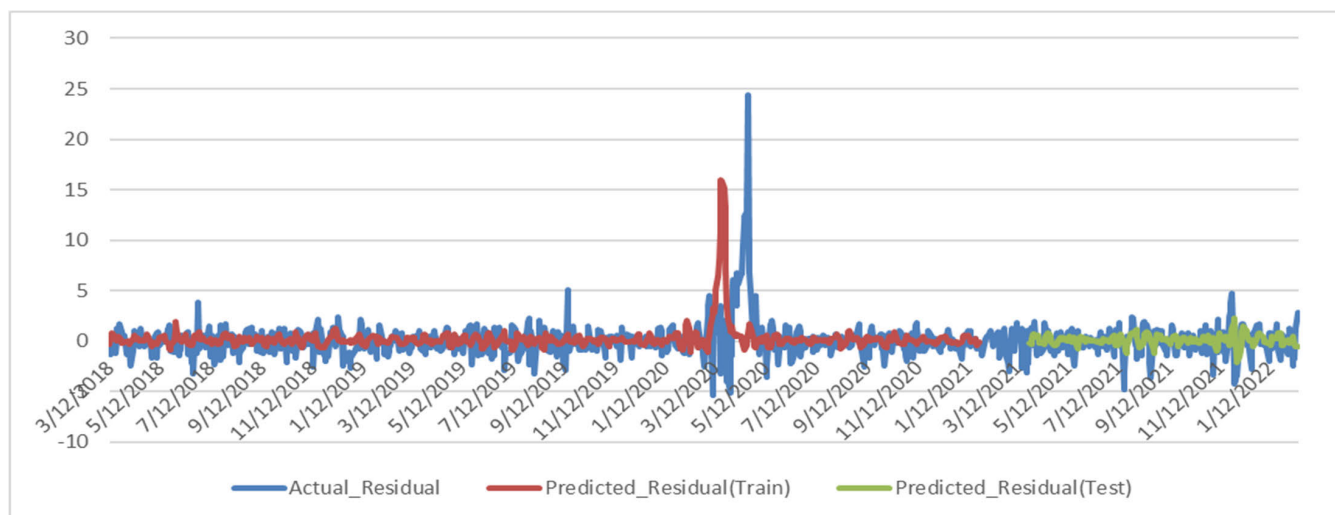


FIGURE 11. LSTM model fitting for LMD-SD-ARIMA residual model.

forecasting the time series [52]. We utilized the optimizer Adam to construct the LSTM model due to its superior results and quick convergence when other optimizers are compared, the authors in [49] gave advised us to utilize it as default. We take the decay as 0.3 while using optimizer Adam. With the dropout set to 30% and the optimizer Adam’s decay set to 0.3, Figure 10 depicts the validation loss and training of our LSTM model by using the best hyperparameter setting. Figure 10 shows that the model and data are most closely matched and that the MSE loss reduces with increasing epoch values.

IV. RESULTS AND DISCUSSIONS

The proposed new hybrid forecasting model LMD-SD-ARIMA-LSTM, which is compared with the ARIMA, LSTM, LMD-ARIMA, LMD-SD-ARIMA, LMD-ARIMA-LSTM methods for predicting the CO price of WTI. Root mean square error (RMSE), Mean absolute error (MAE), and

mean absolute percentage error (MAPE), all are used to calculate WTI data and to evaluate the efficiency of the model. Table 5 provides an overview of each model’s performance accuracy.

In comparison to other models, the hybrid LMD-SD-ARIMA-LSTM achieved the lowest RMSE of 0.150; as a result, this technique is useful for predicting the price of crude oil. Second, for the conventional econometric model, accuracy measurements of the individual model (i.e., ARIMA) and decomposition-ensemble models (i.e., LMD-ARIMA and LMD-ARIMA-LSTM) are compared, and the accuracy of the latter is lower than the former in forecasting. However, for the machine learning model, the three accuracy measures of the individual model (i.e. LSTM) and decomposition-ensemble models (i.e. LMD-ARIMA-LSTM and LMD-SD-ARIMA-LSTM) are compared, Decomposition-ensemble models’ accuracy metrics are lower than those of an individual model. It shows that

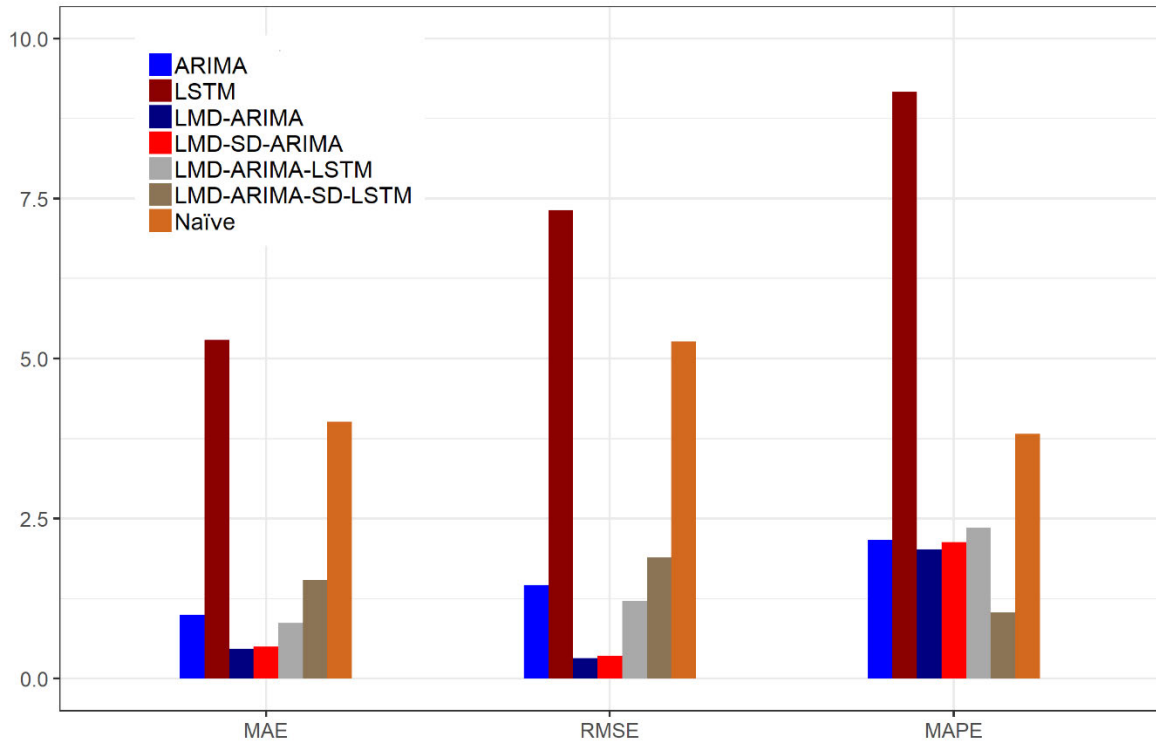


FIGURE 12. Accuracy measures of the hybrid model with other models.

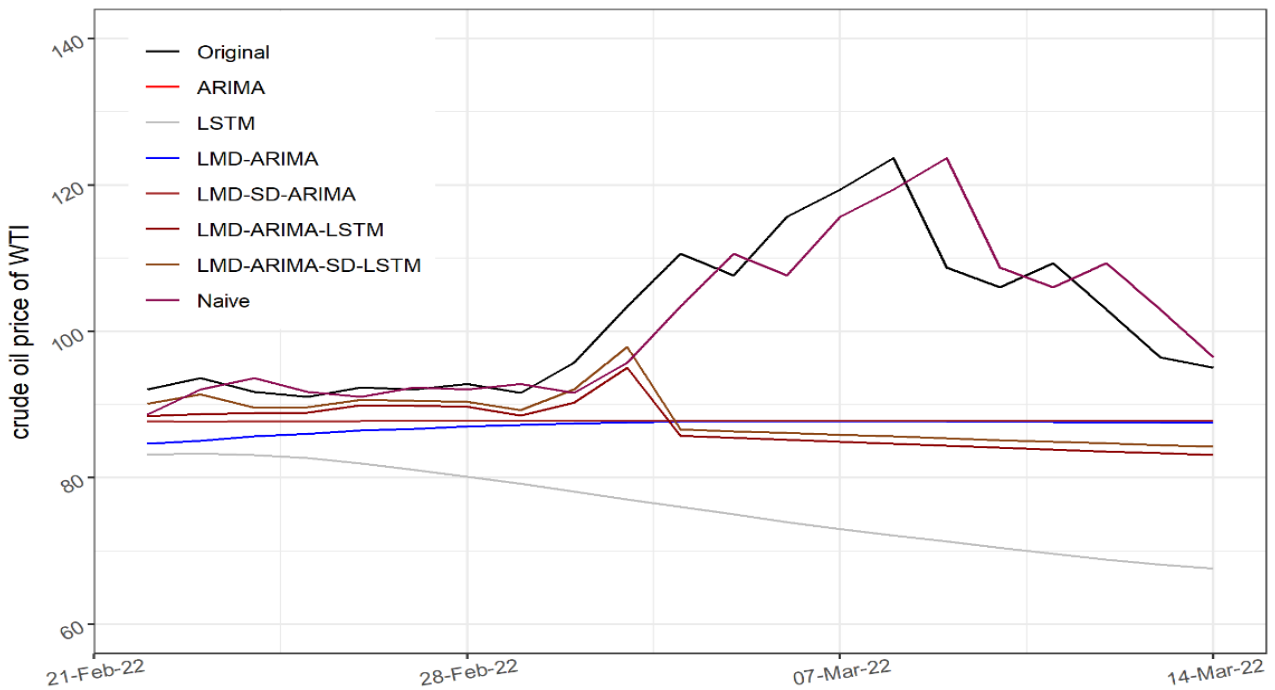


FIGURE 13. The plot of the forecasting results of all models.

after decomposition and reconstruction, the performance of the conventional econometric model for predicting oil prices

can be improved. LSTM model fitting of the LMD-ARIMA and LMD-SD-ARIMA residual model is shown in Figure 11.

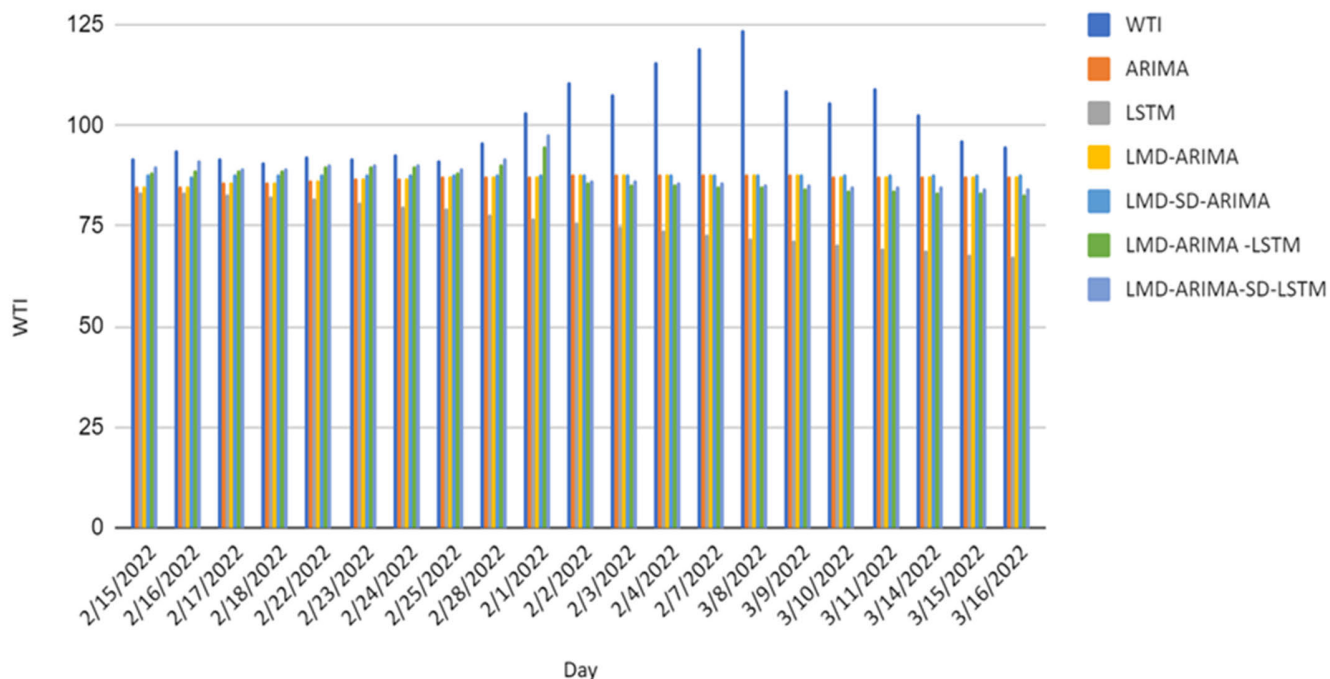


FIGURE 14. The plot of the forecasting results of all models.

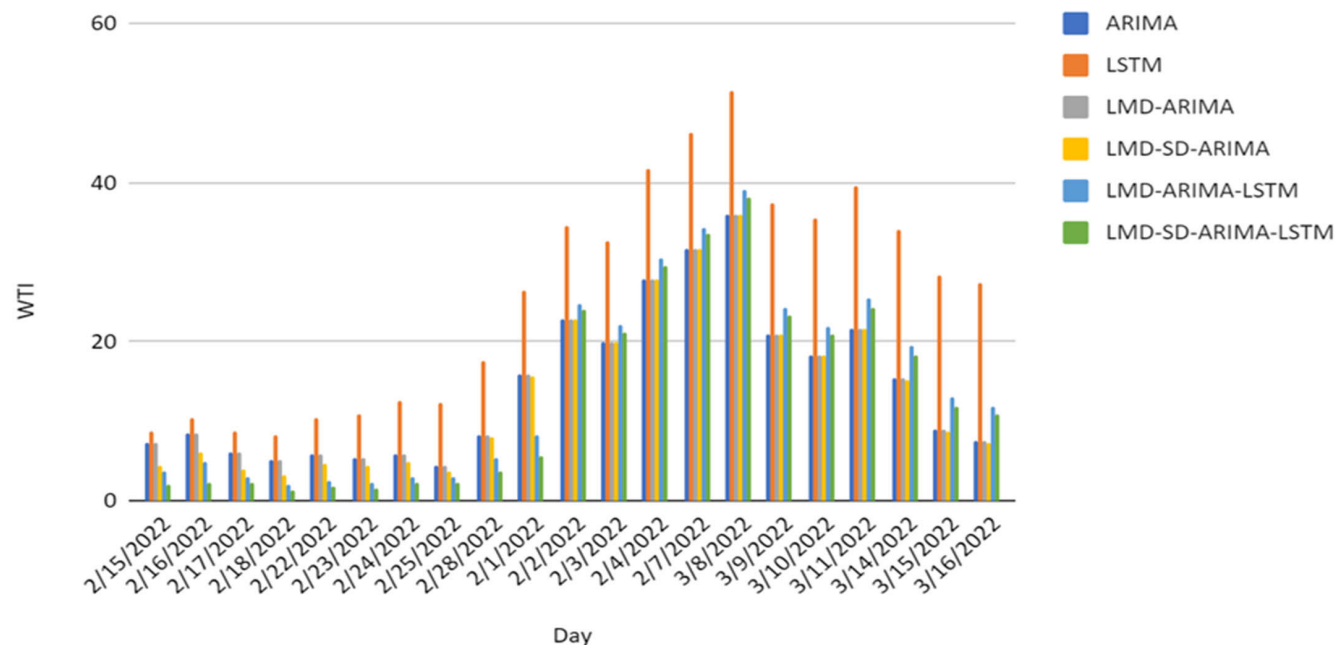


FIGURE 15. Graphical comparison of MAE WTI oil price forecast.

The predicting results of seven separate individual models for WTI crude oil price data are shown in Figure 12 as relative error histograms.

Twenty days of forecasting results for all seven models are shown in Figure 13 which are compared with 20 days of original WTI oil prices from 15th Feb 2022 to 15th March 2022.

Moreover, Table 6, shows that all the models are significantly different from each other.

Figures 14 and 15 compare the proposed model’s errors to those of the other models and provide a graphic comparison of the results of oil price forecasts using all seven models. The forecasting performance of LMD-SD-ARIMA-LSTM is the

TABLE 6. Diebold Mariano test for WTI Forecasting results using all selected models.

Models	DM statistic	p-values
ARIMA with LSTM	-5.707	1.14E-08
ARIMA with LMD-ARIMA	0	3.34E+05
ARIMA with LMD-SD-ARIMA	4.454	8.40E-06
ARIMA with LMD-ARIMA-LSTM	-2.308	2.00E-02
ARIMA with LMD-ARIMA-SD-LSTM	-1.163	2.44E-01
ARIMA with Naïve	3.635	2.77E-04
LSTM with LMD-ARIMA	5.707	1.14E-08
LSTM with LMD-SD-ARIMA	5.889	3.87E-09
LSTM with LMD-ARIMA-LSTM	6.442	1.18E-10
LSTM with LMD-ARIMA-SD-LSTM	6.434	1.24E-10
LSTM with Naïve	5.071	3.95E-07
LMD-ARIMA with LMD-SD-ARIMA	4.454	8.41E-06
LMD-ARIMA with LMD-ARIMA-LSTM	-2.308	2.09E-02
LMD-ARIMA with LMD-ARIMA-SD-LSTM	-1.163	2.45E-01
LMD-ARIMA with Naïve	3.635	2.77E-04
LMD-SD-ARIMA with LMD-ARIMA-LSTM	-2.866	4.15E-03
LMD-SD-ARIMA with LMD-ARIMA-SD-LSTM	-1.757	7.89E-02
LMD-SD-ARIMA with Naïve	3.467	5.26E-04
LMD-ARIMA-LSTM with LMD-ARIMA-SD-LSTM	6.069	1.28E-09
LMD-ARIMA-LSTM with Naïve	3.517	4.35E-04
LMD-ARIMA-SD-LSTM with Naïve	3.353	7.98E-04

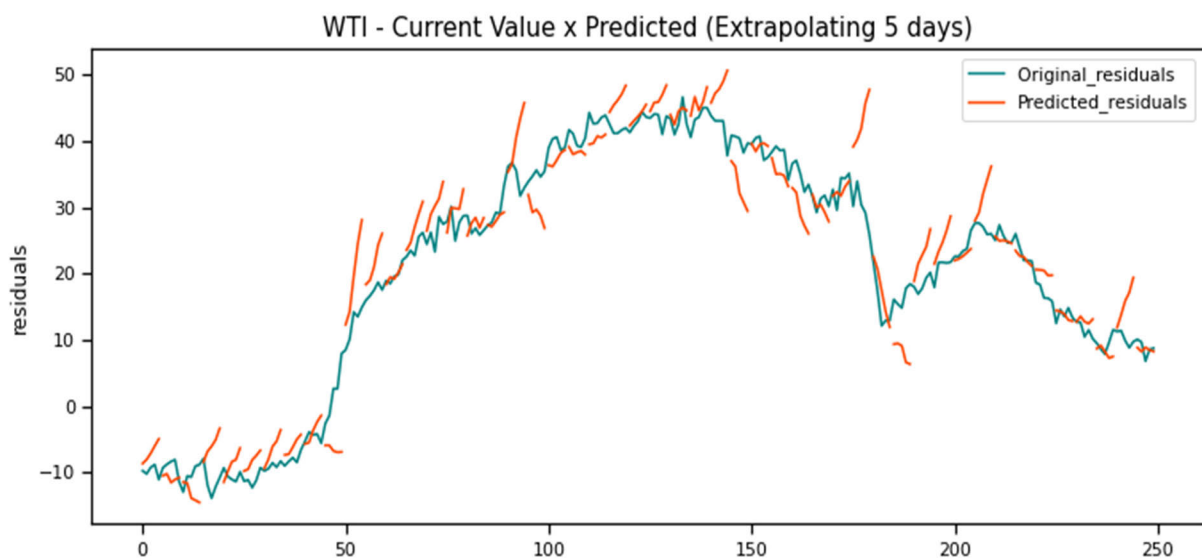


FIGURE 16. Five days ahead forecasting plot.

best, as shown by the three figures. Additionally, compared to the short- and medium-term forecast, the LMD-SD-ARIMA-LSTM technique performs better in the long-term forecast.

For improved presentation, actual data is now compared with the function that divides up each five-day forecast. For the testing data set, the original residuals and predicted

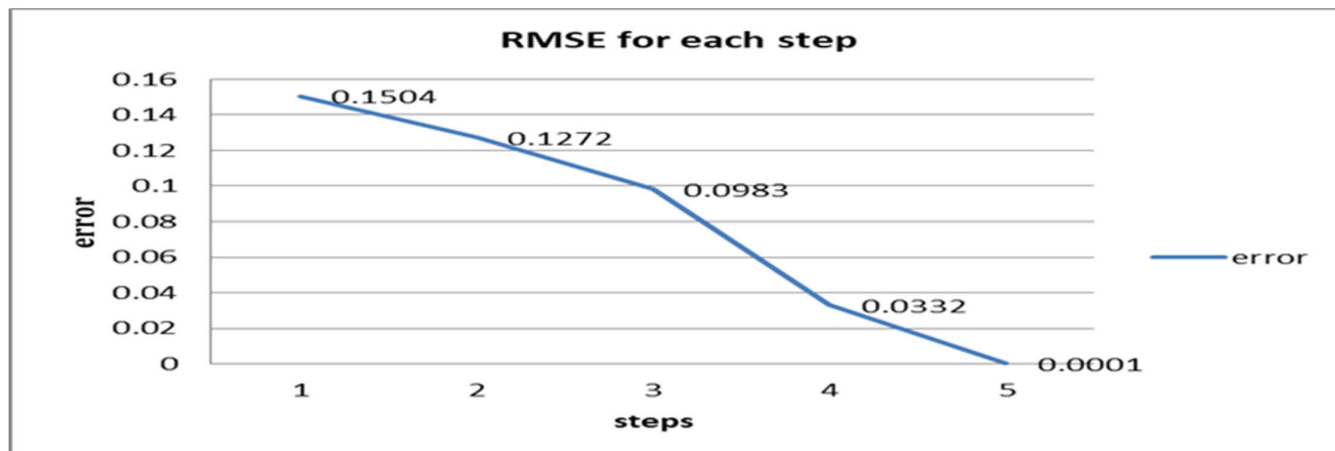


FIGURE 17. Accuracy measures at each step.

values five days ahead are compared and presented in Figure 16.

The average error for each step is analyzed separately. Usually, the tendency of RMSE decreases according to the extrapolated periods due to uncertainties. As predicted, at each step the tendency of RMSE decreases. As in 1st step, we got RMSE is 0.1504, for the second step it decreases to 0.1272 and it continues to decrease till it reaches 0.0001 for five-step ahead.

The RMSE for all five steps is presented in Figure 17.

V. CONCLUSION

Crude oil is one of the non-renewable power sources and is the lifeblood of contemporary industry. The stability of the global economic market benefits from accurate crude oil price forecasting. The LMD-SD-ARIMA-LSTM hybrid prediction approach, which is discussed in this study, is based on the LMD, ARIMA, and LSTM methodologies. The proposed hybrid technique is validated using the WTI CO prices. This study decrease the efforts by collecting the stationary and non-stationary IMFs into stochastic and deterministic components with improved accuracy. The investigation demonstrates, in comparison to the other five approaches, the novel hybrid method significantly increases the prediction accuracy of the CO price. Additionally, the results demonstrate that the conventional econometric model can enhance oil price prediction accuracy following decomposition and reconstruction. Moreover, the new hybrid forecasting system performs better when predicting the price of CO over the medium and long term. Meanwhile, accurate predictions can provide reasonable advice for relevant departments in order to make correct decisions.

A. LIMITATIONS OF THE STUDY

In this study, we only used the univariate time series data.

B. FUTURE RECOMMENDATIONS

The current study can be extended using LSTM based on EEMD and other decompositions methods. Moreover, we can extend the current study into the bivariate and multivariate data.

In the future, some other traditional econometric forecasting models and other machine learning methods will be explored and studied. The factors influencing crude oil prices will also be taken into account, and it will be further investigated to see if the novel hybrid forecasting approach is appropriate for multi-variate forecasting of crude oil prices.

REFERENCES

- [1] G. Wu and Y.-J. Zhang, "Does China factor matter? An econometric analysis of international crude oil prices," *Energy Policy*, vol. 72, pp. 78–86, Sep. 2014.
- [2] H. Mohammadi and L. Su, "International evidence on crude oil price dynamics: Applications of ARIMA-GARCH models," *Energy Econ.*, vol. 32, no. 5, pp. 1001–1008, Sep. 2010.
- [3] H. Xie, X. Zhang, and S. Wang, "The more the better: Forecasting oil price with decomposition-based vector autoregressive model," *Int. J. Energy Statist.*, vol. 1, no. 1, pp. 45–53, Mar. 2013.
- [4] T. Yao and Y.-J. Zhang, "Forecasting crude oil prices with the Google index," *Energy Proc.*, vol. 105, pp. 3772–3776, May 2017.
- [5] T. Klein and T. Walther, "Oil price volatility forecast with mixture memory GARCH," *Energy Econ.*, vol. 58, pp. 46–58, Aug. 2016.
- [6] S. Bekiros, R. Gupta, and A. Paccagnini, "Oil price forecastability and economic uncertainty," *Econ. Lett.*, vol. 132, pp. 125–128, Jul. 2015.
- [7] H. Chiroma, S. Abdulkareem, and T. Herawan, "Evolutionary neural network model for west Texas intermediate crude oil price prediction," *Appl. Energy*, vol. 142, pp. 266–273, Mar. 2015.
- [8] J. Baruník and B. Malinská, "Forecasting the term structure of crude oil futures prices with neural networks," *Appl. Energy*, vol. 164, pp. 366–379, Feb. 2016.
- [9] L. Fan, S. Pan, Z. Li, and H. Li, "An ICA-based support vector regression scheme for forecasting crude oil prices," *Technol. Forecasting Social Change*, vol. 112, pp. 245–253, Nov. 2016.
- [10] M. M. Mostafa and A. A. El-Masry, "Oil price forecasting using gene expression programming and artificial neural networks," *Econ. Model.*, vol. 54, no. 1, pp. 40–53, 2016.
- [11] S. Gao and Y. Lei, "A new approach for crude oil price prediction based on stream learning," *Geosci. Frontiers*, vol. 8, no. 1, pp. 183–187, Jan. 2017.
- [12] T. Xiong, Y. Bao, and Z. Hu, "Beyond one-step-ahead forecasting: Evaluation of alternative multi-step-ahead forecasting models for crude oil prices," *Energy Econ.*, vol. 40, pp. 405–415, Nov. 2013.

- [13] J.-L. Zhang, Y.-J. Zhang, and L. Zhang, "A novel hybrid method for crude oil price forecasting," *Energy Econ.*, vol. 49, pp. 649–659, May 2015.
- [14] L. Yu, Z. Wang, and L. Tang, "A decomposition–ensemble model with data-characteristic-driven reconstruction for crude oil price forecasting," *Appl. Energy*, vol. 156, pp. 251–267, Oct. 2015.
- [15] Y. Zhao, J. Li, and L. Yu, "A deep learning ensemble approach for crude oil price forecasting," *Energy Econ.*, vol. 66, pp. 9–16, Aug. 2017.
- [16] L. Tang, Y. Wu, and L. Yu, "A non-iterative decomposition-ensemble learning paradigm using RVFL network for crude oil price forecasting," *Appl. Soft Comput.*, vol. 70, pp. 1097–1108, Sep. 2018.
- [17] L. Yu, W. Dai, and L. Tang, "A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting," *Eng. Appl. Artif. Intell.*, vol. 47, pp. 110–121, Jan. 2016.
- [18] Y. Ding, "A novel decompose-ensemble methodology with AIC-ANN approach for crude oil forecasting," *Energy*, vol. 154, pp. 328–336, Jul. 2018.
- [19] M. Wang, L. Zhao, R. Du, C. Wang, L. Chen, L. Tian, and H. E. Stanley, "A novel hybrid method of forecasting crude oil prices using complex network science and artificial intelligence algorithms," *Appl. Energy*, vol. 220, pp. 480–495, Jun. 2018.
- [20] E. Jianwei, Y. Bao, and J. Ye, "Crude oil price analysis and forecasting based on variational mode decomposition and independent component analysis," *Phys. A, Stat. Mech. Appl.*, vol. 484, pp. 412–427, Oct. 2017.
- [21] Y. Zhang, F. Ma, B. Shi, and D. Huang, "Forecasting the prices of crude oil: An iterated combination approach," *Energy Econ.*, vol. 70, pp. 472–483, Feb. 2018.
- [22] Y. Chen, C. Zhang, K. He, and A. Zheng, "Multi-step-ahead crude oil price forecasting using a hybrid grey wave model," *Phys. A, Stat. Mech. Appl.*, vol. 501, pp. 98–110, Jul. 2018.
- [23] J. W. Zheng, S. X. Li, and Y. Kun, "A new hybrid model for forecasting crude oil price and the techniques in the model," *Adv. Mater. Res.*, vol. 974, pp. 310–317, Jun. 2014.
- [24] F. Wang, Z. Xuan, Z. Zhen, K. Li, T. Wang, and M. Shi, "A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework," *Energy Convers. Manage.*, vol. 212, May 2020, Art. no. 112766.
- [25] G. Sun, C. Jiang, X. Wang, and X. Yang, "Short-term building load forecast based on a data-mining feature selection and LSTM-RNN method," *IEEE Trans. Electr. Electron. Eng.*, vol. 15, no. 7, pp. 1002–1010, Jul. 2020.
- [26] H. Sun, M. Xia, Y. Hu, S. Lu, Y. Liu, and Q. Wang, "A new sorting feature-based temporal convolutional network for remaining useful life prediction of rotating machinery," *Comput. Electr. Eng.*, vol. 95, Oct. 2021, Art. no. 107413.
- [27] S. Peng, R. Chen, B. Yu, M. Xiang, X. Lin, and E. Liu, "Daily natural gas load forecasting based on the combination of long short term memory, local mean decomposition, and wavelet threshold denoising algorithm," *J. Natural Gas Sci. Eng.*, vol. 95, Nov. 2021, Art. no. 104175.
- [28] Y.-N. Lu, Y.-L. Bai, L.-H. Tang, W.-D. Wan, and Y.-J. Ma, "Secondary factor induced wind speed time-series prediction using self-adaptive interval type-2 fuzzy sets with error correction," *Energy Rep.*, vol. 7, pp. 7030–7047, Nov. 2021.
- [29] R. Li, Y. Hu, J. Heng, and X. Chen, "A novel multiscale forecasting model for crude oil price time series," *Technol. Forecasting Social Change*, vol. 173, Dec. 2021, Art. no. 121181.
- [30] H. He, M. Sun, X. Li, and I. A. Mensah, "A novel crude oil price trend prediction method: Machine learning classification algorithm based on multi-modal data features," *Energy*, vol. 244, Apr. 2022, Art. no. 122706.
- [31] Z. Wang, H. Chen, J. Zhu, and Z. Ding, "Daily PM 2.5 and PM 10 forecasting using linear and nonlinear modeling framework based on robust local mean decomposition and moving window ensemble strategy," *Appl. Soft Comput.*, vol. 114, Jan. 2022, Art. no. 108110.
- [32] J. S. Smith, "The local mean decomposition and its application to EEG perception data," *J. Roy. Soc. Interface*, vol. 2, no. 5, pp. 443–454, 2005.
- [33] L. S. Dar, M. Aamir, Z. Khan, M. Bilal, N. Boonsatit, and A. Jirawattana-panit, "Forecasting crude oil prices volatility by reconstructing EEMD components using ARIMA and FFNN models," *Frontiers Energy Res.*, vol. 10, pp. 1–20, Oct. 2022.
- [34] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [35] M. Phi, "Illustrated guide to LSTM's and GRU's: A step by step explanation," *Towards Data Sci.*, vol. 9, pp. 1–17, Sep. 2018.
- [36] S. Hansun, A. Wicaksana, and M. B. Kristanda, "Prediction of Jakarta City air quality index: Modified double exponential smoothing approaches," *Int. J. Innov. Comput., Inf. Control*, vol. 17, no. 4, pp. 1363–1371, 2021.
- [37] F. Shahid, A. Zameer, and M. Muneeb, "Predictions for COVID-19 with deep learning models of LSTM, GRU and bi-LSTM," *Chaos, Solitons Fractals*, vol. 140, Nov. 2020, Art. no. 110212.
- [38] S. Hansun and M. B. Kristanda, "Forecasting foreign tourist arrivals to Bali: Hybrid double exponential smoothing approach," *Int. J. Eng. Res. Technol.*, vol. 12, no. 11, pp. 1864–1868, 2019.
- [39] W. Gao, M. Aamir, A. B. Shabri, R. Dewan, and A. Aslam, "Forecasting crude oil price using Kalman filter based on the reconstruction of modes of decomposition ensemble model," *IEEE Access*, vol. 7, pp. 149908–149925, 2019.
- [40] P. Xu, M. Aamir, A. Shabri, M. Ishaq, A. Aslam, and L. Li, "A new approach for reconstruction of IMF's of decomposition and ensemble model for forecasting crude oil prices," *Math. Problems Eng.*, vol. 2020, pp. 1–23, Oct. 2020.
- [41] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*. Hoboken, NJ, USA: Wiley, 2015.
- [42] M. Aamir and A. Shabri, "Modelling and forecasting monthly crude oil price of Pakistan: A comparative study of ARIMA, GARCH and ARIMA Kalman model," in *Proc. AIP Conf.*, 2016, Art. no. 060015.
- [43] G. Elliot, T. J. Rothenberg, and J. H. Stock, "Efficient tests for an autoregressive unit root," *Econometrica*, vol. 64, pp. 813–836, Dec. 1996.
- [44] M. Aamir and A. Shabri, "Improving crude oil price forecasting accuracy via decomposition and ensemble model by reconstructing the stochastic and deterministic influences," *Adv. Sci. Lett.*, vol. 24, no. 6, pp. 4337–4342, 2018.
- [45] Y. Liu, L. Guan, C. Hou, H. Han, Z. Liu, Y. Sun, and M. Zheng, "Wind power short-term prediction based on LSTM and discrete wavelet transform," *Appl. Sci.*, vol. 9, no. 2, p. 1108, 2019.
- [46] Y.-T. Tsai, Y.-R. Zeng, and Y.-S. Chang, "Air pollution forecasting using RNN with LSTM," in *Proc. IEEE 16th Intl Conf Dependable, Autonomic Secure Comput., 16th Intl Conf Pervasive Intell. Comput., 4th Intl Conf Big Data Intell. Comput. Cyber Sci. Technol. Congress(DASC/PiCom/DataCom/CyberSciTech)*, Aug. 2018, pp. 1074–1079.
- [47] D. Smilkov, N. Thorat, Y. Assogba, C. Nicholson, N. Kreeger, P. Yu, S. Cai, E. Nielsen, D. Soegel, S. Bileschi, and M. Terry, "TensorFlow.js: Machine learning for the web and beyond," in *Proc. Mach. Learn. Syst.*, vol. 1, 2019, pp. 309–321.
- [48] K. G. Sheela and S. N. Deepa, "Review on methods to fix number of hidden neurons in neural networks," *Math. Problems Eng.*, vol. 2013, pp. 1–11, May 2013.
- [49] N. Reimers and I. Gurevych, "Reporting score distributions makes a difference: Performance study of LSTM-networks for sequence tagging," 2017, *arXiv:1707.09861*.
- [50] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 56, pp. 1929–1958, 2014.
- [51] I. Kaastra and M. Boyd, "Designing a neural network for forecasting financial and economic time series," *Neurocomputing*, vol. 10, no. 3, pp. 215–236, 1996.
- [52] S. Makridakis and M. Hibon, "Exponential smoothing: The effect of initial values and loss functions on post-sample forecasting accuracy," *Int. J. Forecasting*, vol. 7, no. 3, pp. 317–330, Nov. 1991.



JAWARIA NASIR was born in Rawalpindi, Pakistan, in 1985. She received the M.Phil. degree from Quaid-i-Azam University, Islamabad, in 2010. She is currently pursuing the Ph.D. degree with Abdul Wali Khan University Mardan. She was worked as a Lecturer with the Federal Urdu University of Arts, Science, and Technology. Her research interests include time series decomposition and machine learning methods.



MUHAMMAD AAMIR (Senior Member, IEEE) received the bachelor's degree (Hons.) in statistics from the University of Peshawar, Pakistan, in 2005, the master's degree in statistics from Quaid-i-Azam University, Islamabad, Pakistan, in 2013, and the Ph.D. degree in mathematics from Universiti Teknologi Malaysia, Malaysia, in 2018. In 2006, he joined the Department of Statistics, Islamia College University Peshawar, Pakistan, as a Lecturer. In 2009, he moved to Abdul Wali Khan University Mardan, Pakistan, where he was a Lecturer and became an Assistant Professor, in 2013. His current research interests include sampling, machine learning for mathematical problem diagnostics, econometrics, and time series. He also serves as an Active Reviewer for various international journals, including IEEE ACCESS, *PLOS ONE*, *Resources Policy*, *Energy*, *Journal of Cluster Science*, and *Heliyon*.



ZAHOOOR UL HAQ is currently pursuing the Ph.D. degree in agricultural economics and business with the University of Guelph (UoG), Canada. He is also working as a Professor in economics and the Vice Chancellor with Abdul Wali Khan University Mardan (AWKUM). Before joining AWKUM, he worked with the Ontario Ministry of Agriculture, Food and Rural Affairs, Canada, as an Economist, and United Arab Emirates University. In addition, he completed two postdoctoral studies,

one from UoG, and Gatton School of Business and Economics, University of Kentucky, USA. He received numerous awards. He also received the Fulbright Pre-Doctoral Scholarship, Ontario Graduate Scholarship, the Agricultural Institute of Canada Foundation's Award "Douglas MacRorie Memorial Scholarship," the Toronto Milk Producer's Award, and the HEC Best Teachers Award, in 2016. His research interests include agricultural economics and business, microeconomics, econometrics, and financial econometrics. His research portfolio comprises more than 70 articles published in leading journals. His research also got the Best Research Paper Award at the 41st International Business Research Conference, in April 2017, with the Imperial College London.



SHEHZAD KHAN received the Ph.D. degree in finance from Universiti Teknologi Malaysia. He is currently working as an Assistant Professor in finance with Abdul Wali Khan University Mardan (AWKUM), Pakistan, where he is also the Deputy Director (Research) with the Pakhtunkhwa Economic Policy Research Institute (PEPRI). He is also the Managing Editor of the *Journal of Applied Economics and Business Studies*. His research interests include but not limited to initial public offerings, capital structure, ESG, corporate governance, CSR, financial econometrics, and behavioral finance.



MUHAMMAD YUSUF AMIN received the Ph.D. degree in finance from Eastern Mediterranean University, North Cyprus. He is currently working as an Assistant Professor (Finance) with the Institute of Business Studies and Leadership, Abdul Wali Khan University Mardan. In addition to this, he is the Managing Editor of *Journal of Business and Tourism*. His research interests include but is not limited to capital structure, dividend policy, public finance, financial market, corporate finance, investment, and econometrics.



MUHAMMAD NAEEM received the bachelor's and master's degrees in statistics from Islamia College University Peshawar, Pakistan, in 2010 and 2014, respectively. He is currently pursuing the Ph.D. degree with the Department of Statistics, Abdul Wali Khan University Mardan, Pakistan. His research interests include time series modeling, linear models, machine learning, applied statistics, and computational statistics.

...