



Vaasan yliopisto
UNIVERSITY OF VAASA

OSUVA Open
Science

This is a self-archived – parallel published version of this article in the publication archive of the University of Vaasa. It might differ from the original.

Forecasting Crude Oil Prices using a Hybrid Model Combining Long Short-Term Memory Neural Networks and Markov Switching Model

Author(s): Shahbazbegan, Vahid; Hosseininesaz, Hamid; Shafie-Khah, Miadreza; Elmusrati, Mohammed

Title: Forecasting Crude Oil Prices using a Hybrid Model Combining Long Short-Term Memory Neural Networks and Markov Switching Model

Year: 2023

Version: Accepted Manuscript

Copyright ©2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Please cite the original version:

Shahbazbegan, V., Hosseininesaz, H., Shafie-Khah, M. & Elmusrati, M. (2023). Forecasting Crude Oil Prices using a Hybrid Model Combining Long Short-Term Memory Neural Networks and Markov Switching Model. In *2023 International Conference on Future Energy Solutions (FES)*, 1-6. IEEE. <https://doi.org/10.1109/FES57669.2023.10182444>

Forecasting Crude Oil Prices using a Hybrid Model Combining Long Short-Term Memory Neural Networks and Markov Switching Model

Vahid Shahbazbegian
School of Technology and Innovations
University of Vaasa
Vaasa, Finland
vahid.shahbazbegian@uwasa.fi

Hamid Hosseinesaz
School of Technology and Innovations
University of Vaasa
Vaasa, Finland
hamid.hosseinesaz@student.uwasa.fi

Mohammed Elmusrati
School of Technology and Innovations
University of Vaasa
Vaasa, Finland
mohammed.elmusrati@uwasa.fi

Miadreza Shafie-khah
School of Technology and Innovations
University of Vaasa
Vaasa, Finland
miadreza.shafiekhah@uwasa.fi

Abstract— Given the significant impact of crude oil prices on the global economy, accurately predicting their fluctuations is essential for effective decision-making in the energy sector. Therefore, this research aims to develop a hybrid model that can comprehensively capture the nonlinear and volatile characteristics of crude oil prices and provide accurate predictions. The proposed approach involves segmenting the time series into multiple sub-series, which capture the nonlinear and volatile characteristics of crude oil prices. The nonlinear sub-series is predicted using Long Short-Term Memory neural networks, while the volatile and fluctuating sub-series are forecasted using a Markov Switching model. The results of these predictions are combined using a linear combination to estimate the crude oil price time series. The proposed hybrid model provides a comprehensive understanding of the various factors that drive crude oil price fluctuations, making it a valuable tool for decision-making in the energy sector.

Keywords— crude oil, forecast, long short-term memory, Markov model, neural networks, time series.

I. INTRODUCTION

Crude oil has a crucial role in the world economy, so in 2021, providing almost 31% of the world's energy demand. With a share of over 10%, crude oil dominates as the most substantial and active commodity traded in commercial exchanges worldwide. It is purchased and sold among different players in the international oil market, including producing countries, oil companies, refineries, importing countries, and traders [1]. Apart from its role in the global economy, crude oil is also a significant factor in the economies of both exporting and importing countries. For exporting countries, a considerable portion of their revenue is generated from oil exports, while importing countries allocate a significant portion of their budget to import crude oil [2].

Economic researchers have been intrigued by the prospect of predicting oil prices for a considerable period of time [3]–[5]. However, due to the intricate and volatile nature of the oil market, accurately forecasting oil prices has proven to be challenging. Oil prices, like most goods, are determined by the forces of supply and demand. However, they are predominantly influenced by various other factors, such as past, present, and future events, climate, stock market performance, gross domestic product growth, political developments, as well as rumors and psychological

predictions [6]. These factors contribute to drastic changes in the oil market, which exhibit complex, dynamic, and irregular characteristics.

The study of time series forecasting and its associated applications has been extensively explored over the years [7], [8]. Time series refer to a collection of quantitative measurements that are recorded at regular time intervals. The primary objective of analyzing time series is to create a numerical model based on past data to predict future events. Nonetheless, in real-world scenarios, time series prediction can be extremely difficult due to inherent noise, instability, and irregularity [9].

According to empirical studies on the oil economy, some scholars have discovered significant indications of regime-switching in the oil market. This implies that the drivers of oil prices determine two features of changes in oil prices. Firstly, changes in oil prices impact economic cycles, which suggests that different empirical models are required for different economic cycles. Secondly, while the empirical models remain the same, the influencing factors exhibit distinct characteristics under varying economic cycles, resulting in time-varying phenomena in the oil price [10].

Previously, Di Sanzo (2018) [11] proposed a time series model that combines long memory and Markov Switching (MS) dynamics to forecast oil price return volatility. The model is compared to various Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models and is found to outperform them in terms of forecasting accuracy. Aimer & Lusta (2021) [12] used Vector autoregression and Markov switching-Vector autoregression models to explore the relationships between oil prices, US dollar exchange rates, and stock markets. Their study finds that there are significant interrelations between the variables, and the relationships between them are non-linear. Additionally, the links between the variables change depending on the regime, with stronger links during high volatility periods. In their study, Hong et al. (2022) [13] examined how financial stress index impact oil price volatility using the Time-Varying Transition Probabilities-Markov Switching-GARCH model. Their findings indicated a significant correlation between the regimes of these two groups. Additionally, they found that the Time-Varying Transition Probabilities -Markov Switching-

GARCH model outperforms the symmetric model in predicting oil price volatility. Lastly, to validate the reliability of their model, they analyzed time series with various lags and found that their model is robust. Wang & Wang (2022) [14] enhanced the forecasting accuracy for both short-term and long-term horizons by merging the Mixed Data Sampling and Markov Switching-GARCH models. Their findings revealed that Markov-Switching-GARCH models typically outperform in the long-term horizon. The primary constraint of Markov Switching models lies in their linear assumption. Therefore, alternative approaches are necessary to accommodate non-linear patterns. Many studies have also utilized hybrid models to leverage the strengths of both linear and non-linear characteristics [15], [16]. Empirical evidence suggests that hybrid models outperform single models due to their capability to transform the data structure from linear to non-linear as time progresses [17], [18].

Among the studies, Vijh et al. (2020) [19] applies Artificial Neural Network (ANN) and Random Forest techniques to predict the next day's closing price for five companies across different sectors of operation. The models created using financial data perform well based on standard strategic indicators such as Root-mean-square deviation (RMSE) and Mean absolute percentage error (MAPE). Gupta & Nigam (2020) [20] proposes the use of ANN to predict crude oil prices, considering the different factors that affect it. The proposed ANN approach captures the unstable pattern of crude oil prices by finding the optimal lag and number of delay effects. By varying the lag over time, the proposed model is validated using the root mean square error and shows significantly better results compared to other models. Hu et al. (2020) [21] proposed a novel hybrid method to forecast copper price volatility by combining the GARCH model and deep neural network. The GARCH model provides statistical information about time-varying copper price volatility while the deep neural network uses domestic and international market factors to search for better nonlinear features. The method was found to significantly increase the predictive power of the neural network model.

Zhao et al. (2021) [6] proposed prediction models based on the use of Recurrent Neural Networks, Long Short-Term Memory, and Gated Recurrent Unit. The results suggest that Long Short-Term Memory and Gated Recurrent Unit models outperform the basic Recurrent Neural Networks model, and the addition of the attention mechanism layer helps improve the accuracy of the prediction model. Additionally, their study found that deeper neural networks did not necessarily improve results. Sohrabi et al (2022) [22] used an ANN with a whale optimization algorithm to predict West Texas Intermediate (WTI) petroleum prices. Five parameters, including gold price, coal price, natural gas price, Dollar-Euro exchange rate, and Dollar-Yuan exchange rate, are used as inputs. The ANN-whale optimization algorithm model improves the WTI price predicting accuracy up to 22% compared to the ANN. Highly correlated variables can significantly increase the accuracy of the forecast in predicting price or other variables.

Liu et al. (2012) [23] introduced a method for forecasting crude oil prices based on Empirical Mode Decomposition (EMD) and ANN. The EMD technique divides the oil price time series into several independent Intrinsic Mode Functions

(IMFs), which are then further divided into three categories: short-term fluctuations, extreme event impacts, and long-term trends. Using the results of this decomposition and synthesis, ANN is employed to model each of the IMFs. The research demonstrated that the proposed EMD-ANN technique outperforms ANN alone. Zhang et al. (2021) [24] developed a hybrid model to predict oil prices. They first applied ensemble empirical mode decomposition to extract the error terms, and then used particle swarm optimization and kernel extreme learning machine to forecast the oil price. To validate their findings, they used Brent and WTI oil prices. The results showed that their model significantly improved forecast accuracy. In [25], Lu et al. (2022) employed a Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN)-GA-SVR hybrid model for oil price prediction. Their study involved using the CEEMDAN technique to decompose the oil sector's time series into simpler ones, which were then predicted using the SVR model. The input parameters of the SVR model were optimized through the GA. The findings indicated that the proposed hybrid model outperformed the SVR model in predicting oil prices with higher accuracy.

The objective of this study is to create a hybrid approach that can predict crude oil prices. The proposed model leverages the CEEMDAN technique to partition the time series into multiple sub-series that capture the nonlinear and volatile nature of crude oil prices. The nonlinear sub-series are predicted using an ANN, while the volatile and fluctuating sub-series are forecasted using a MS model. Finally, the Brent crude oil price time series is estimated by combining the predictions through a linear combination of ANN and MS.

II. CRUDE OIL PRICE FORECASTING METHODOLOGY

The proposed hybrid model aims to decompose a time series of crude oil prices into multiple distinct series using CEEMDAN. Next, the LSTM algorithm is employed to predict the time series with nonlinear features. Additionally, a MS model is utilized to predict time series with volatile and fluctuating characteristics. The hybrid model then estimates the crude oil price time series by utilizing a linear combination of the aforementioned predictions. This research follows the below steps, which combine various methods to capture different facets of crude oil price fluctuations.

Step 1) Crude oil price time series is converted to IMF, depending on the level of decomposition, using the CEEMDAN. The price of crude oil is obtained by Eq (1):

$$X(t) = \sum_{i=1}^n x_i(t) \quad (1)$$

Step 2) Time series obtained from CEEMDAN can be considered as a component consisting of non-linear and fluctuating characteristics. The CEEMDAN is used to extract the nonlinear and volatile characteristics of crude oil prices. In Eq (2), the first component represents the fluctuating behavior and the second component depicts the nonlinear behavior.

$$\sum_{i=1}^n x_i(t) = \sum_{i=1}^m V_i(t) + \sum_{i=m+1}^n N_i(t) \quad (2)$$

Step 3) The volatile behavior of crude oil prices can be seen in the time series $\{IMF1, IMF2, \dots, IMF5\}$. MS model is used to model and predict the fluctuating component. The prediction value obtained from MS model is equal to $\hat{V}_i(t)$.

Step 4) The time series $\{IMF6, IMF7, \dots, IMF9\}$ have a non-linear feature, and ANN algorithm is used to model and predict them. The prediction value obtained from the ANN is equal to $\hat{N}_i(t)$.

Step 5) The values predicted by the ANN algorithm for the time series with non-linear properties are linearly combined with the values predicted by the MS model for the time series with fluctuating properties and forecast the price of crude oil. We get that it is defined as Eq (3).

$$\sum_{i=1}^n \hat{X}_i(t) = \sum_{i=1}^m \hat{V}_i(t) + \sum_{i=m+1}^n \hat{N}_i(t) \quad (3)$$

A. CEEMDAN

Huang et al. (1998) introduced EMD, a method that automatically breaks down signals into a group of IMFs. However, the technique suffered from modal aliasing and an endpoint effect, which restricted its applicability. To tackle this issue, Wu & Huang (2009) suggested a variant of EMD, which incorporates Gaussian white noise into EMD and cancels out the noise by averaging it multiple times. Though it partially resolves the modal aliasing problem, the decomposition efficiency remains poor, and modal aliasing and energy leakage persist in the low-frequency region. To overcome these limitations, Torres et al (2011) [26] proposed CEEMDAN, which involves adaptive noise reconstruction, resulting in extremely low error rates. CEEMDAN is effective at handling nonlinear and nonstationary signals since it adaptively adds Gaussian white noise at each stage of the decomposition process, leading to complete decomposition.

B. Artificial Neural Network-LSTM

Long Short-Term Memory Neural Networks (LSTM) neural networks were firstly introduced by Hochreiter and Schmidhuber in 1997 [27]. It is a type of recurrent neural network architecture that is designed to handle the problem of vanishing gradients and the difficulty of learning long-term dependencies in sequential data. LSTMs are capable of selectively remembering and forgetting information over time, which makes them particularly useful for processing sequential data, such as speech, natural language, time series data, and image captioning [28].

The architecture of an LSTM cell consists of three gates: an input gate, a forget gate, and an output gate. These gates regulate the flow of information through the cell and allow LSTMs to store and retrieve information from previous time steps. The input gate controls how much new information is added to the cell state, while the forget gate decides which information should be removed from the cell state. An LSTM neural network also consists of four main layers. The input layer receives input sequences and passes them on to the next layer. The LSTM layer contains the LSTM cells, which process the input sequences and learn to selectively store and retrieve information over time. The dense layer takes the

output of the LSTM layer and performs a nonlinear transformation on it to produce the final output. Finally, the output layer produces the final output of the network, such as a classification or regression prediction. An illustration of an LSTM neural network is indicated in Fig. 1.

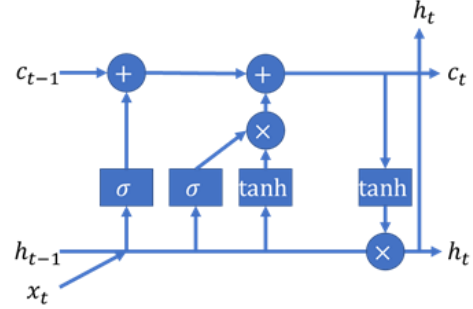


Fig. 1. Illustration of a LSTM neural network.

C. Markov switching

The MS model was first introduced by Quant in 1972 and Gould in 1973 and then developed by Hamilton in 1989 to derive business cycles. In general, in non-linear models, it is assumed that the behavior of the variable on which the modeling is done is different and changes in different situations. In the switching Markov model, parameters can change continuously from one time to another. If r_t is a vector of observed variables and S_t is an unobserved random change that can take the values 1, 2, 3..., S_t then can be expressed as a first-order Markov chain.

$$P_t(S_t = j/S_{t-1} = i, S_{t-2} = k, \dots) = P_t(S_t = j/S_{t-1} = i) = P_{ij} \quad (4)$$

P_{ij} is the transition probability. A Markov process is more stable if the probabilities are closer to 1. Considering a two-state process $S_t \in \{1, 2\}$, the probability of switching from state 1 to state 2 can be written as follows:

$$P_t(S_t = j/S_{t-1} = i, S_{t-2} = k, \dots) = P_t(S_t = j/S_{t-1} = i) = P_{ij} \quad (5)$$

$$P_{11} + P_{12} = 1 \text{ and } P_{21} + P_{22} = 1 \quad (6)$$

$$P_{11} = P_r[S_t = 1/S_{t-1} = 1] \text{ \& } P_{12} = P_r[S_t = 2/S_{t-1} = 1] \quad (7)$$

$$P_{22} = P_r[S_t = 2/S_{t-1} = 2] \text{ \& } P_{21} = P_r[S_t = 1/S_{t-1} = 2] \quad (8)$$

The MS model is estimated by adopting maximum likelihood techniques, in which the conditional density is defined as follows:

$$y_i = f(y_t/S_t = i, y_{t-1}; \theta) \text{ where } i = 1, 2, 3, \dots, k \quad (9)$$

And Marginal density is determined as follows:

$$f(y_t / \theta) = \sum_{i=1}^k f(y_t/S_t = i, y_{t-1}; \theta) P_r(S_t = i; \theta) \quad (10)$$

The likelihood function for estimating probabilities is obtained as follows:

$$P_r(S_t = 1, y_t; \theta) = \frac{f(y_t/S_t = 1, y_{t-1}; \theta) P_r(S_t = 1; \theta)}{f(y_t/y_{t-1}; \theta)} \quad (11)$$

One of the advantages of the MS method compared to other methods is the separation of the endogeneity of the observations of a variable and also emphasized the relationship between the observations of the variables, and in this respect, the MS method is completely different from the models based on structural failure and virtual variables.

III. CASE STUDY

In this research, Brent crude oil price data has been used for price modeling and forecasting. The data analyzed in this study includes 504 observations of Brent crude oil prices for the period 2021-2022 (<https://tradingeconomics.com>). The selected data are divided into two categories: training data and test data; 404 data are considered as training data and 100 data as test data.

IV. RESULT AND ANALYSIS

To check the performance of the presented hybrid algorithm, its results are compared with the single ANN algorithms and the single MS model:

In this paper, three criteria are utilized to analyze the results and compare the proposed methods. These criteria are the average absolute value of errors, the root mean square errors, and the average absolute value of the percentage of errors.

The average absolute value of errors provides a measure of the average deviation between the predicted values and the actual values. It is calculated by taking the absolute value of the difference between the predicted and actual values and then finding the average of these absolute differences. A smaller value for this criterion indicates better performance of the proposed methods in accurately predicting the values. The root means square errors measure the deviation between the predicted and actual values by taking the square root of the mean of the squared differences. It provides a measure of the magnitude of the errors, considering both the direction and the magnitude of the deviations. A smaller value for this criterion also indicates better performance of the proposed methods in accurately predicting the values.

The average absolute value of the percentage of errors measures the percentage deviation between the predicted and actual values, where the absolute difference is divided by the actual value and multiplied by 100. It provides a measure of the relative deviation between the predicted and actual values. A smaller value for this criterion indicates better performance of the proposed methods in predicting the values with higher accuracy.

Table 1 shows the performance of three different models used to predict values. The models are a Markov model, an ANN model, and a hybrid model. The performance is measured using the three criteria: mean square error, root mean square error, and average absolute value of the percentage of errors.

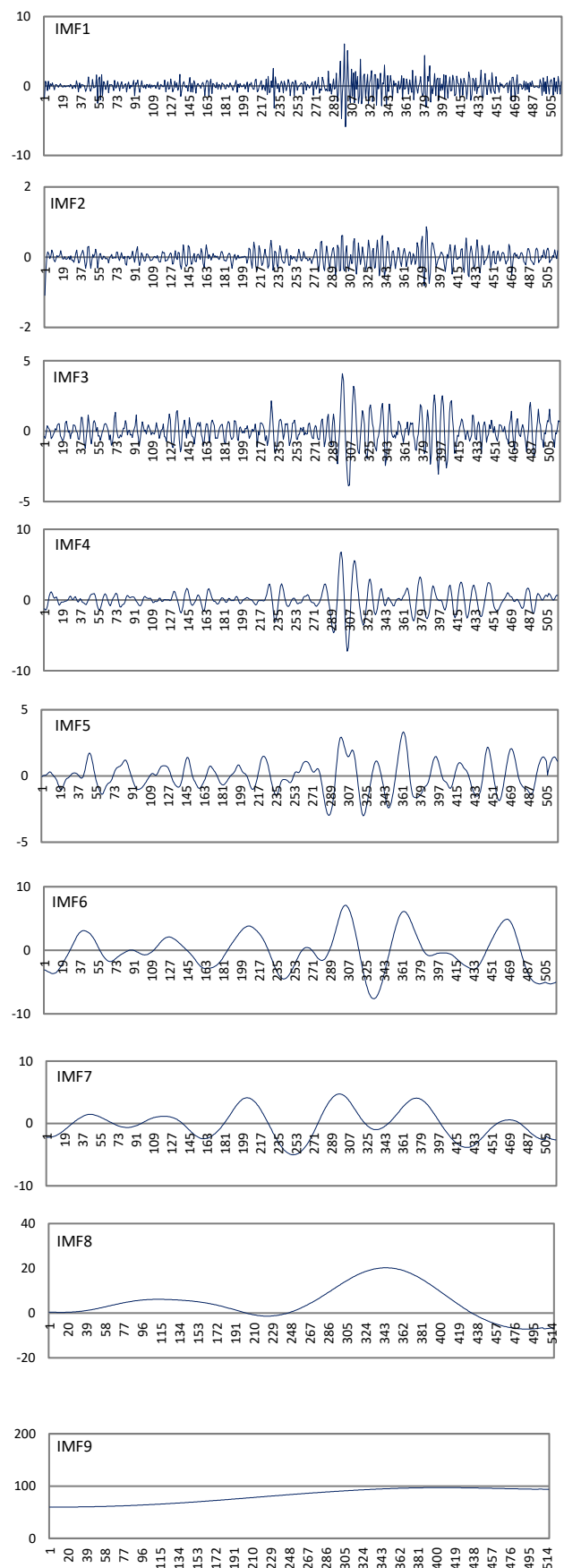


Fig. 2. Decomposition results of crude oil price series by the CEEMDAN method.

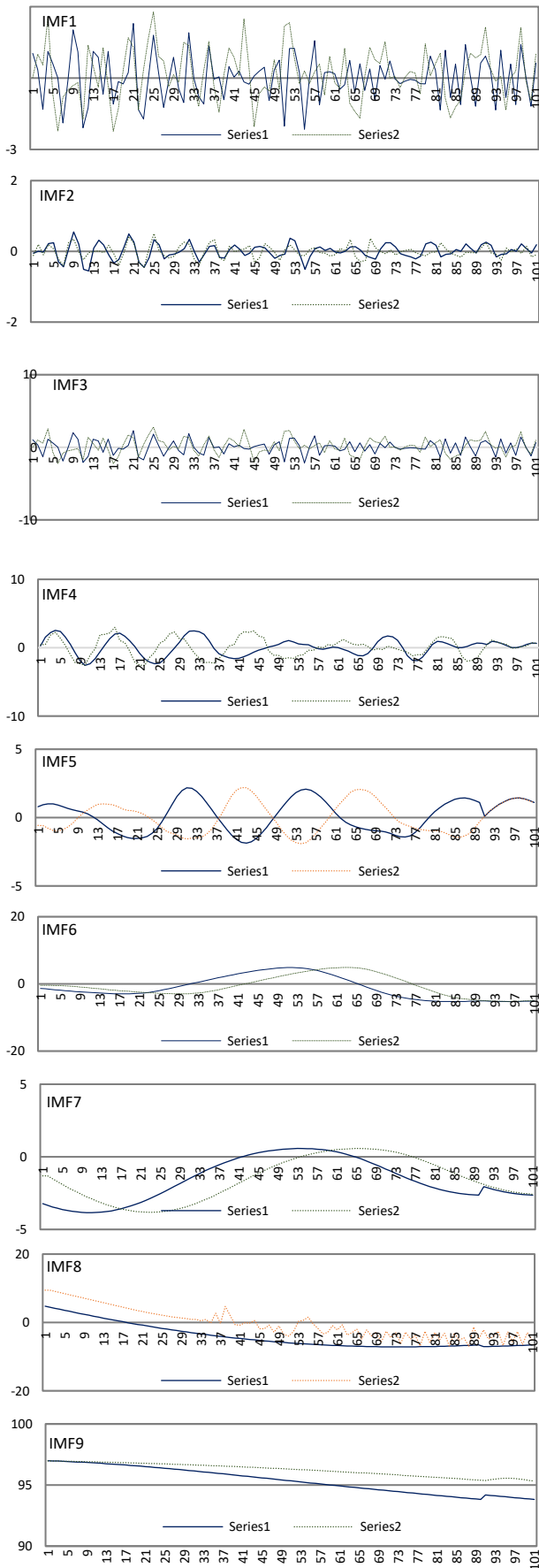


Fig. 3. Prediction of crude oil price series using the proposed approach (series 1 and series 2 indicate actual and predicted value, respectively).

For the Markov model, the mean square error is 2.99%, the root mean square error is 4.17%, and the average absolute value of the percentage of errors is 0.03377%. As explained earlier, a smaller value for each criterion indicates better performance.

For the ANN model, the mean square error is 2.60%, the root mean square error is 4.35%, and the average absolute value of the percentage of errors is 0.023%. Compared to the Markov model, the ANN model has a smaller mean square error and a smaller average absolute value of the percentage of errors. However, the root mean square error is slightly larger for the ANN model, indicating a slightly higher magnitude of errors. For the hybrid model, the mean square error is 3.02%, the root mean square error is 3.65, and the average absolute value of the percentage of errors is 0.034. The hybrid model has a larger mean square error and a larger average absolute value of the percentage of errors compared to the Markov model. However, the root mean square error is smaller for the hybrid model, indicating a lower magnitude of errors.

Overall, Table I suggests that the ANN model performs the best in predicting the values, as it has the smallest mean square error and average absolute value of the percentage of errors. However, it is important to consider all three criteria and the specific application when choosing the best model. The hybrid model may still be a good choice for some applications, as it has a lower magnitude of errors compared to the Markov model.

TABLE I. COMPARISON OF PERFORMANCE OF DIFFERENT METHODS

Index	Approach		
	<i>Markov</i>	<i>ANN</i>	<i>Hybrid</i>
MEA (%)	2.9941	3.6094	3.0287
RMSE (%)	4.1783	4.3500	3.6521
MAPE (%)	0.0333	0.0232	0.0342

V. CONCLUSION

This research developed a hybrid model that could comprehensively capture the nonlinear and volatile characteristics of crude oil prices and provide accurate predictions, given the significant impact of crude oil prices on the global economy. The proposed approach involved segmenting the time series into multiple sub-series and predicting the nonlinear sub-series using Long Short-Term Memory neural networks, while forecasting the volatile and fluctuating sub-series using a MS model. The results of these predictions were combined using a linear combination to estimate the crude oil price time series. The proposed hybrid model provided a comprehensive understanding of the various factors that drive crude oil price fluctuations, making it a valuable tool for decision-making in the energy sector.

In comparison, this study compared the performance of three different models, a Markov model, an ANN model, and a hybrid model, in predicting values to identify the best model for a specific application. The methodology involved

measuring the performance of each model using three criteria, mean square error, root mean square error, and average absolute value of the percentage of errors. The most important results showed that the neural network model outperformed the other two models with the smallest mean square error and average absolute value of the percentage of errors. However, the hybrid model had a lower magnitude of errors compared to the Markov model, indicating that it may still be a good choice for some applications. Overall, this study provided useful information for researchers and practitioners in selecting an appropriate model for predicting values in their specific applications.

REFERENCES

- [1] N. E. S. Derani and P. Naidu, "The Impact of Utilizing Social Media as a Communication Platform during a Crisis within the Oil Industry," *Procedia Economics and Finance*, vol. 35, pp. 650–658, Jan. 2016, doi: 10.1016/S2212-5671(16)00080-0.
- [2] N. Sehgal and K. K. Pandey, "Artificial intelligence methods for oil price forecasting: a review and evaluation," *Energy Systems*, vol. 6, no. 4, pp. 479–506, Nov. 2015, doi: 10.1007/S12667-015-0151-Y/TABLES/24.
- [3] A. Safari and M. Davallou, "Oil price forecasting using a hybrid model," *Energy*, vol. 148, pp. 49–58, Apr. 2018, doi: 10.1016/J.ENERGY.2018.01.007.
- [4] J. Beckmann, R. L. Czudaj, and V. Arora, "The relationship between oil prices and exchange rates: Revisiting theory and evidence," *Energy Econ*, vol. 88, p. 104772, May 2020, doi: 10.1016/J.ENERGY.2020.104772.
- [5] L. Han, Q. Lv, and L. Yin, "Can investor attention predict oil prices?," *Energy Econ*, vol. 66, pp. 547–558, Aug. 2017, doi: 10.1016/J.ENERGY.2017.04.018.
- [6] L. Zhao, L. Cheng, Y. Wan, H. Zhang, and Z. Zhang, "A VAR-SVM model for crude oil price forecasting," *International Journal of Global Energy Issues*, vol. 38, no. 1–3, pp. 126–144, 2015, doi: 10.1504/IJGEI.2015.069485.
- [7] Z. Wu and N. E. Huang, "ENSEMBLE EMPIRICAL MODE DECOMPOSITION: A NOISE-ASSISTED DATA ANALYSIS METHOD," <https://doi.org/10.1142/S1793536909000047>, vol. 1, no. 1, pp. 1–41, Nov. 2011, doi: 10.1142/S1793536909000047.
- [8] M. I. Haque and A. Rahman Shaik, "Predicting Crude Oil Prices During a Pandemic: A Comparison of Arima and Garch Models," 2021, doi: 10.14254/1800-5845/2021.17-1.15.
- [9] G. A. Christodoulakis and S. E. Satchell, "Hashing GARCH: a reassessment of volatility forecasting performance," *Forecasting Volatility in the Financial Markets*, pp. 227–247, Jan. 2007, doi: 10.1016/B978-075066942-9.50012-1.
- [10] X. Gong, K. Guan, L. Chen, T. Liu, and C. Fu, "What drives oil prices? — A Markov switching VAR approach," *Resources Policy*, vol. 74, p. 102316, Dec. 2021, doi: 10.1016/J.RESOURPOL.2021.102316.
- [11] S. Di Sanzo, "A Markov switching long memory model of crude oil price return volatility," *Energy Econ*, vol. 74, pp. 351–359, Aug. 2018, doi: 10.1016/J.ENERGY.2018.06.015.
- [12] N. M. Aimer and A. A. Lusta, "Exchange rates and oil price under uncertainty and regime switching: A Markov-switching VAR approach," *Economic Journal of Emerging Markets*, vol. 13, no. 2, pp. 200–215, Oct. 2021, doi: 10.20885/EJEM.VOL13.ISS2.ART9.
- [13] Y. Hong, L. Wang, C. Liang, and M. Umar, "Impact of financial instability on international crude oil volatility: New sight from a regime-switching framework," *Resources Policy*, vol. 77, p. 102667, Aug. 2022, doi: 10.1016/J.RESOURPOL.2022.102667.
- [14] J. Wang and J. Wang, "Forecasting energy market indices with recurrent neural networks: Case study of crude oil price fluctuations," *Energy*, vol. 102, pp. 365–374, May 2016, doi: 10.1016/J.ENERGY.2016.02.098.
- [15] T. Mingming and Z. Jinliang, "A multiple adaptive wavelet recurrent neural network model to analyze crude oil prices," *J Econ Bus*, vol. 64, no. 4, pp. 275–286, Jul. 2012, doi: 10.1016/J.JECONBUS.2012.03.002.
- [16] Y. J. Zhang and J. L. Zhang, "Volatility forecasting of crude oil market: A new hybrid method," *J Forecast*, vol. 37, no. 8, pp. 781–789, Dec. 2018, doi: 10.1002/FOR.2502.
- [17] J. Chai, L. M. Xing, X. Y. Zhou, Z. G. Zhang, and J. X. Li, "Forecasting the WTI crude oil price by a hybrid-refined method," *Energy Econ*, vol. 71, pp. 114–127, Mar. 2018, doi: 10.1016/J.ENERGY.2018.02.004.
- [18] L. Lin, Y. Jiang, H. Xiao, and Z. Zhou, "Crude oil price forecasting based on a novel hybrid long memory GARCH-M and wavelet analysis model," *Physica A: Statistical Mechanics and its Applications*, vol. 543, p. 123532, Apr. 2020, doi: 10.1016/J.PHYSA.2019.123532.
- [19] M. Vijh, D. Chandola, V. A. Tikkiwal, and A. Kumar, "Stock Closing Price Prediction using Machine Learning Techniques," *Procedia Comput Sci*, vol. 167, pp. 599–606, Jan. 2020, doi: 10.1016/J.PROCS.2020.03.326.
- [20] N. Gupta and S. Nigam, "Crude Oil Price Prediction using Artificial Neural Network," *Procedia Comput Sci*, vol. 170, pp. 642–647, Jan. 2020, doi: 10.1016/J.PROCS.2020.03.136.
- [21] Y. Hu, J. Ni, and L. Wen, "A hybrid deep learning approach by integrating LSTM-ANN networks with GARCH model for copper price volatility prediction," *Physica A: Statistical Mechanics and its Applications*, vol. 557, p. 124907, Nov. 2020, doi: 10.1016/J.PHYSA.2020.124907.
- [22] P. Sohrabi, H. Dehghani, and R. Rafie, "Forecasting of WTI crude oil using combined ANN-Whale optimization algorithm," <https://doi.org/10.1080/15567249.2022.2083728>, vol. 17, no. 1, 2022, doi: 10.1080/15567249.2022.2083728.
- [23] H. Liu, C. Chen, H. Q. Tian, and Y. F. Li, "A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks," *Renew Energy*, vol. 48, pp. 545–556, Dec. 2012, doi: 10.1016/J.RENENE.2012.06.012.
- [24] T. Zhang, Z. Tang, J. Wu, X. Du, and K. Chen, "Multi-step-ahead crude oil price forecasting based on two-layer decomposition technique and extreme learning machine optimized by the particle swarm optimization algorithm," *Energy*, vol. 229, p. 120797, Aug. 2021, doi: 10.1016/J.ENERGY.2021.120797.
- [25] Y. Lu, J. Luo, Y. Cui, Z. He, and F. Xia, "Improved CEEMDAN, GA, and SVR Model for Oil Price Forecasting," *J Environ Public Health*, vol. 2022, 2022, doi: 10.1155/2022/3741370.
- [26] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pp. 4144–4147, 2011, doi: 10.1109/ICASSP.2011.5947265.
- [27] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Networks*, vol. 18, no. 5–6, pp. 602–610, Jul. 2005, doi: 10.1016/J.NEUNET.2005.06.042.
- [28] J. Martens and I. U. Ca, "Learning Recurrent Neural Networks with Hessian-Free Optimization Ilya Sutskever," 2011.