Vol: 06 Issue: 06

International Journal of Data Science and Advanced Analytics (ISSN: 2563-4429)

# Predicting Crude Oil Price Using Time Series Statistical Modelling Techniques

Sunny Raj Pradeep Kumar Gupta[1] and Thomas Coombs[2]
[1]BASF, India, [2]eSystems Engineering Society, UK
[1]srg231985@gmail.com, [2]coombstom25@gmail.com

Corresponding author email: coombstom25@gmail.com

***Abstract*—**One of the important topics in Global Economy is crude oil prediction. Nonetheless, crude oil prediction is related to many factors that play a role in it and include geological factors, production, technical etc…Accurate prediction of crude oil is significant considering all the previous factors. However, these factors are volatile and vary under different conditions. Therefore, this study proposes using machine learning models for time series predictions of crude oil. More specifically, three types of convolutional neural networks were used being autoregressive integrated moving average (ARIMA), seasonal autoregressive moving average (SARIMA) and long short term memory (LSTM). The models were applied to three datasets of crude oil including Crude Oil Daily Price Data, Crude Oil Weekly Inventory Data and Crude Oil Monthly Production Data. Data analysis was conducted using Python. The outcomes of the predictions showed that LSTM had higher accuracy and prediction than ARIMA and SARIMA. The LSTM daily model achieved a MAPE score of as low as 0.087. Future research involve more reliable production data and include more reliable input parameters to improve accuracy.

***Keywords*—Crude oil, global economy, ARIMA, SARIMA, LSTM, prediction.**

## 1. Introduction

Crude oil is a naturally occurring substance found deep beneath the earth's surface [1-3]. It is a dark hydrocarbon found thousands of meters beneath the ground. It is an important building block for various industries, and its prices significantly impact the world economy [4-7]. Oil price forecasting is of great interest to scholars, researchers, and policymakers due to its significant effect on various economic sectors and markets [8-10]. Thus, it is imperative to study various crude oil price forecasting techniques.

Crude Oil prices have been fluctuating over time. In the recent Covid Crisis period in April 2020, the price of Crude Oil dropped as low as $20, and then skyrocketed as high as $120 in June 2022 [11].

Various machine learning models have been applied to predict crude oil prices [12-14]. Some reference models to predict crude oil prices include utilizing support vector regression (SVR) fuzzy time-series model utilised GASVR to predict and analyse crude oil prices. To explain the non-linearity associated with Crude Oil Prices Halleh Bostanchi utilised GARCH (Generalized autoregressive conditional heteroscedastic) and ARCH (Auto regressive conditional heteroscedastic) model. Though ARIMA models are widely used and trusted techniques for time series analysis [15-18].

### 1.1. Literature review

The prediction of crude oil can be incredibly complex and difficult to conduct for large scale refinery. A popular method to predict crude oil historically was linear models for time series forecasting in the past was the autoregressive integrated moving average (ARIMA) model [19-20]. Nonlinear patterns on the other hand, are difficult for ARIMA models to capture. More recently recurrent neural networks (RNN) have been evaluated by researchers using RNN-LSTM network [21]. These results obtained by the RNN and LSTM model were promising performing far superior to ARIMA/SARIMA models [22-25].

ARIMA models were developed in the 1970s and since have been used in a variety of applications predominantly in the fields of financial and economic forecasting. Several studies have looked at how well ARIMA models perform in various forecasting tasks to which the ARIMA was determined to be a trustworthy and efficient method for forecasting, especially when compared to conventional statistical and machine learning techniques [26].

Tidal Forecasting has also been utilized by the ARIMA model. By using two models, ARIMA and LSTM neural network were used in predicting tidal data. The ARIMA model produced forecast data when only linear components were considered and results accurate and few inaccuracies. The ARIMA prediction error sequence was trained using the LSTM neural network model and outputs nearly in agreement with the measured values [27].

Another proposed model which was a mixed model of ARIMA and XGBoost. The model dataset was divided into approximation and error parts using a discrete wavelet transform [28]. This was then processed by the ARIMA model and further improved by XGBoost model (GSXGB). This is followed by the discrete wavelet transform. The DWT-ARIMA-GSXGB model offered superior approximation and generalisation capabilities for stock index opening price prediction. Moreover, the prediction capabilities were enhanced by using a single ARIMA model or XGBoost model [29]. Another time series forecasting method is SARIMA (Seasonal Autoregressive Integrated Moving Average). The SARIMA model combines moving average and autoregression. The ability to capture seasonal patterns in datasets and long-term trends are the key benefits of using the SARIMA model to forecast [30]. There have been numerous studies that have assessed the effectiveness of SARIMA models. An example of this is research conducted by Daryl and Aditya Winata which examined SARIMAs ability in forecasting Stock Market Prices. The results showed that the SARIMA model obtained a MAPE score of 36.05%, while the other two models had AIC scores of 580.165, 451.591, and 114.612. The system performed as expected, however, it was unable to correctly anticipate the Apple company's stock market value in real time [31].

In addition, a recent study by Jiaxin Liu and Zijian Zhao tried to predict dissolved gas concentration in power transformer oil. This study proposes a SARIMA model to predict the concentration of dissolved gases in transformer oil. It examines the relationship between the ambient temperature and gas concentration as well as the impact of the period parameter on the prediction effect. In comparison to the AR model and the LSTM model, the SARIMA model is determined to have the most accurate and consistent performance. When external factors are considered, accuracy increases [32].

Hochreiter and Schmidhuber created the Long Short-term Memory (LSTM) Recurrent Neural Network (RNN) architecture in 1997. As a result of the vanishing or exploding gradients problem in backpropagation, it is specifically created to address this issue. It also resolves RNNs difficulty in capturing long-term patterns in data. A comprehensive review of LSTM-based applications has been conducted by Norshakirah Aziz and Mohd Hafizul Afifi Abdullah to predict crude oil prices using RNN-LSTM Neural Network. The purpose of this study was to create an RNN-LSTM model utilising Tensor-Flow-Keras and Python to forecast changes in the price of crude oil based on historical data and other technical analysis indicators. The model performed well in predicting changes in the price of crude oil, and the addition of artificial intelligence techniques may open up new possibilities for the analysis of nonstationary data. By using an on-demand stream learning approach, the model mayalso be updated with fresh information on oil price [33-34].

Additionally, LSTM models have been applied to forecast stock prices with a feature fusion LSTM-CNN model by Taewook Kim and Ha Young Kim. The feature fusion Long Short-Term Memory-Convolutional Neural Network (LSTM-CNN) model for stock price prediction was addressed in this study. The model combines picture features from stock chart images with temporal data from stock time series. Based on data from the SPDR S&P 500 ETF, the performance of the LSTM and CNN model was compared to that of the single models (CNN and LSTM), and it was discovered to outperform the single models [35]. Additionally, it was discovered that a candlestick chart was the best stock chart illustration to utilise when predicting stock prices.

In one of the recent research studies conducted by Indrajeet Kumar and Bineet Kumar Tripathi LSTM network-assisted models were proposed for petroleum production forecasting. In this study an attention-based long short-term memory (A-LSTM) network is used in this article to present a technique for predicting petroleum production. A dynamic floating window with different window sizes is employed to prevent data leakage, and the A-LSTM network is quick and precise. The model was examined and contrasted with statistical, deep learning, hybrid, and machine learning approaches using actual production data. The A-LSTM network outperformed the other models, according to the results. In conclusion, LSTM models have been used for a variety of tasks and have proven to be quite successful at handling long-term dependencies and identifying patterns in data. LSTM models have been demonstrated to be an effective tool for a variety of tasks such as time series forecasting, natural language processing, computer vision, medical decision-making, and robotics, while additional study is required to further enhance their effectiveness [36].

To summarise the aforementioned models, the optimal time to employ ARIMA models is for seasonal data because they are good at capturing linear correlations in data. Like ARIMA models, SARIMA models also consider seasonal variations in the data. SARIMA models can be utilised for more complicated forecasting problems because they are more complex than ARIMA models. Long Short-Term Memory (LSTM) models are very helpful in Time series forecasting. These models can be applied to challenging forecasting jobs and are effective at capturing non-linear correlations in data. Complex patterns in data that simpler models would miss can be discovered by LSTM models. The complexity of the forecasting problem must be considered when comparing the performance of the ARIMA, SARIMA, and LSTM models. The ARIMA and SARIMA models are good enough for less complicated applications like forecasting weekly sales. However, LSTM models might perform better for more challenging tasks like stock price prediction. In addition, LSTM models can be used to detect non-linear relationships in data that simpler models would miss. Most of the machine learning models referenced during the Literature review built a univariate model based on price only. Only price has been considered as an input parameter to build these models. With this motivation this research focuses on multivariate nonlinear model building to include Price, Inventory and Production data.

## 2. **Materials and Methods**

Initially, data handling and merging were covered, followed by Exploratory Data Analysis (EDA), which includes scatter plots, box plots, histograms, bar plots, and 50DMA-200DMA plots. Stationarity tests (KPSS test and ADF test) were performed on the dataset, and then box-cox transformation and differencing were applied. ACF and PACF plots were made to get the q and p parameters for the ARIMA model.

Prior to the model-building process, the dataset was split into train and test datasets. Multiple time series forecasting models [37] (both linear and nonlinear) such as ARIMA Daily Price Data Model, SARIMA Daily Price Data Model, Univariate-Daily price data LSTM model, and Multivariate-Weekly Price vs Inventory data LSTM model was built.

### 2.1. **Datasets**

The three datasets consisting of Crude Oil Daily Price Data, Crude Oil Weekly Inventory Data and Crude Oil Monthly Production Data were imported into Python. The datasets were merged using pandas dataframe.

### 2.2. **Data selection**

All the data collected for this research work comes from the U.S. Energy Information Administration website. To support sound policymaking, effective markets, and public understanding of energy and its relationships with the economy and the environment, the U.S. Energy Information Administration (EIA), a key agency of the U.S. Federal Statistical System, collects, analyses, and disseminates data on energy. Data on coal, oil, gas, electricity, renewable energy, and nuclear energy are covered by EIA program. The U.S. Department of Energy includes EIA.

### 2.3. **Data pre-processing and transformation**

To make the datasets on Crude Oil Prices, Weekly Inventory data and Monthly Production data for Machine learning / deep learning algorithms, pre-processing is required. The handling of empty records, missing values and normalising of data values make up this pre-processing.

1. Normalisation / Standardisation: Before using any machine learning model, the primary notion is to normalise or standardise, i.e., set your features, variables, and columns of X. Each column of X is individually normalised by StandardScaler so that each column, feature, and variable have a value of mean= 0 and standard deviation=1.
2. Merge Price data with Inventory data: We have transformed Daily Price data to weekly since the time frame of Inventory data is weekly.
3. Merge Price data with Inventory data and Production data-: We have transformed Daily Price data to monthly, Inventory data to monthly since the time frame of Production data is monthly.
4. Missing value/ Blank value: Missing value/ Blank value if any is treated using the Forward fill method.
5. Stationarity tests: One of the important assumptions to building an Autoregressive model is to check the stationarity of the data set.
6. Non-Stationary to stationary Conversion-: To build an Auto-Regressive model convert non-stationary series to stationery.

### 2.4. **Exploratory data analysis**

The EDA analysis consisted of scatter plots between crude oil daily price and inventory data. This was followed by using a box-plot and interquartile range with the mean, quartiles, range and outliers summarised in the box-plot. 50-Day Moving Average and 200-Day Moving Average are plotted for price dataset. 50DMA and 200DMA is an excellent technical indicator for take buy/sell decisions in crude oil trading. When 50DMA is above 200DMA, the price is generally considered bullish and if 50DMA is below 200DMA then it is considered bearish.

### 2.5. **Model development**

First, three ARIMA models were built for the Daily Price Data, Weekly Price Data, and Month Price Data. The SARIMA model was then built to fit, define the index and make predictions, create a prediction data frame and plot the predictions and calculate errors. The LSTM model was them developed to for both univariate and multivariate models. Figure 1 shows the LSTM layer in detail:

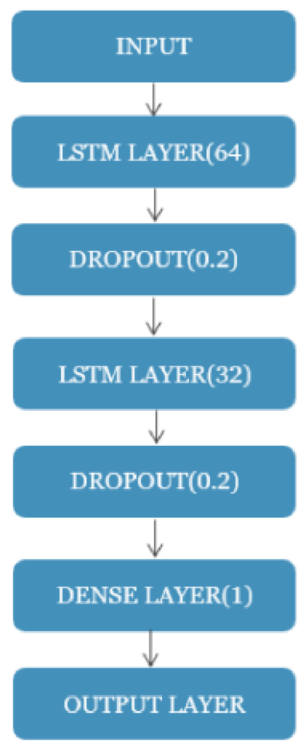


Figure 1. LSTM layers diagram.

## 3. **Results and Discussion**

This section evaluates the performance of the ARIMA, SARIMA and LSTM models chosen in this research. The models built are further evaluated using 3 metrics such as MSE (Mean Square Error), RMSE (Root Mean Square Error) and MAPE (Mean Absolute Percentage Error).

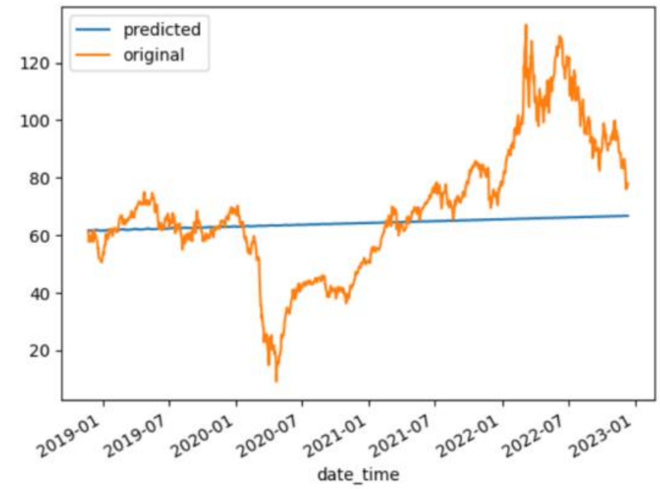


Figure 2. Predicted and test ARIMA model values

To assess the relative efficacy of the suggested models, each model's performance is contrasted with that of the baseline model. If any other comparable models were employed for the same task, the results are also compared to those models. Finally, conclusions are drawn from the findings, and the models' potential for improvement is noted.

### 2.6. ARIMA model

The ARIMA model Daily Price Data did not explain the non-linearity of the Crude Oil Price. The predicted and test values were not closely aligned and MAPE score was 0.25, MSE 505, and RMSE 22.49.

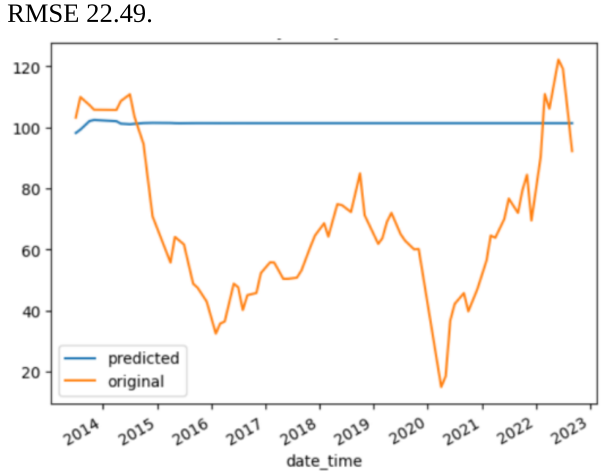


Figure 3. Predicted and test SARIMA model values

When the Weekly Price Data were input as the only parameter, again, the predicted and test values were not closely aligned. The MAPE values were 0.31, MSE 711, and RMSE 26.66. The final ARIMA model, Monthly Price Data was input as the only parameter which resulted in predicted and test value not closely aligned. The MAPE, MSE, and RMSE scores were 0.35, 1729, and 41.58, respectively.

### 2.7. SARIMA model

In this model, we used daily price as the only input parameter. The results look decent though it doesn't explain the non-linearity of the Crude Oil Price. A comparison between the predicted and test (original) values showed that they are not closely aligned. The test period was between January 2019 to December 2022. The MAPE score of this model was 0.25, MSE 520, and RMSE 22.81.

### 2.8. LSTM model

The first LSTM was univariate model with Daily Price Data as the input parameter, obtaining more significant results than ARIMA and SARIMA. The test period was between January 2019 to December 2022 and resulted in aligned predicted and test values.

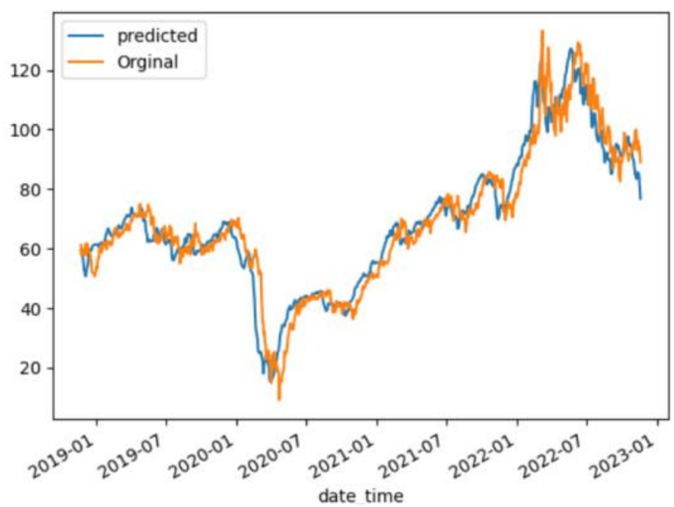


Figure 4. Daily Price Data for univariate LSTM model

The MAPE score was 0.151, MSE 207.2, and RMSE 14.39. When Monthly Price Data (test period between 2014 to 2022) were input into the LSTM univariate model, again the results of the predicted and test values were aligned providing in a better prediction of crude oil. The MAPE, MSE, and RMSE values were 0.207, 250.9, and 15.84, respectively.

Table 1. Univariate model metrics

| Evaluation Parameter | ARIMA Daily | ARIMA Weekly | ARIMA Monthly | SARIMA Daily | LSTM Daily | LSTM Weekly | LSTM Monthly |
|---|---|---|---|---|---|---|---|
| MSE | 505 | 711 | 1729 | 520 | 45.97 | 207.2 | 250.9 |
| RMSE | 22.49 | 26.66 | 41.58 | 22.81 | 6.78 | 14.39 | 15.84 |
| MAPE | 0.25 | 0.31 | 0.35 | 0.25 | 0.087 | 0.151 | 0.207 |

For the multivariate LSTM model, first, Weekly Price vs Inventory Data (test period between 2010 to 2022) were both input as parameters obtaining good results. The test and predicted values were aligned resulting in good crude oil price prediction with MAPE, MSE, and RMSE values of 0.159, 137.9, and 11.74, respectively. Finally, Monthly Price vs Inventory vs Production Data (test period between 2017 to 2022) were all input as parameters for the final multivariate LSTM model. This resulted in test and predicted values not completely aligned and would need further improvement before predicting crude oil price. The MAPE, MSE and RMSE values were 0.254, 170, and 13.06, respectively.

Table 2. Multivariate model metrics

| Evaluation parameter | LSTM Price vs Inventory | LSTM Price vs Inventory vs Production |
|---|---|---|
| **MSE** | 137.9 | 170 |
| **RMSE** | 11.74 | 13.06 |
| **MAPE** | 0.159 | 0.254 |

The results from the linear (ARIMA and SARIMA models) did not produce satisfactory results in predicting crude oil price. Moreover, the results of the lower timeframes perform better than those in higher timeframe models, with Daily Price Prediction for the LSTM model produced a MAPE score of 0.087. The major contributor to poor performing models is unreliable and uncorrelated production data. The availability of various input data in a lower timeframe could improve the performance of the model to a great extent.

## 4. Conclusions

In summary by performing a detailed comparison between ARIMA, SARIMA and LSTM models for crude oil prediction and examining the advantages and disadvantages of each model.

After comparing the results of all the models, it is evident that considering the nonlinear nature of Crude Oil Prices, Nonlinear models from the Neural Network family performed better in forecasting crude oil prices. This study provided insight into forecasting crude oil prices using ARIMA and SARIMA models, as well as LSTM models. LSTM models were found to be superior to the traditional models in terms of accuracy, with the daily model achieving a MAPE score of 0.087 and the weekly model achieving a MAPE score of 0.159. The multivariate model, while ambitious, did not perform as well due to unreliable production data, achieving a MAPE score of 0.254. Future research should further improve the multivariate models by using inventory data and production data on lower timeframes.

## References


[1] Simanzhenkov, V. and Idem, R., 2003. *Crude oil chemistry*. Crc Press.

[2] Gisser, M. and Goodwin, T.H., 1986. Crude oil and the macroeconomy: Tests of some popular notions: Note. *Journal of Money, Credit and Banking*, *18*(1), pp.95-103.

[3] Tsui, K.K., 2011. More oil, less democracy: Evidence from worldwide crude oil discoveries. *The Economic Journal*, *121*(551), pp.89-115.

[4] Rafiq, S., Salim, R. and Bloch, H., 2009. Impact of crude oil price volatility on economic activities: An empirical investigation in the Thai economy. *Resources policy*, *34*(3), pp.121-132.

[5] Sharma, A., Singh, G., Sharma, M. and Gupta, P., 2012. impact of Crude oil Price on Indian economy. *International Journal of Social Sciences & Interdisciplinary Research*, *1*(4), pp.95-99.

[6] Zhang, X., Wu, Q. and Zhang, J., 2010, October. Crude oil price forecasting using fuzzy time series. In *2010 Third International Symposium on Knowledge Acquisition and Modeling* (pp. 213-216). IEEE.

[7] Talreja, R., Coombs, T., Assi, S., Ismail, N.A., Jayabalan, M., Liatsis, P., Mahyoub, M., Al-Hamid, A. and Kolivand, H., 2022, December. Well Log Data Preparation and Effective Utilization of Drilling Parameters Using Data Science Based Approaches. In *The International Conference on Data Science and Emerging Technologies* (pp. 388-402). Singapore: Springer Nature Singapore.

[8] Pai, P.F. and Lin, C.S., 2005. A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega*, *33*(6), pp.497-505.

[9] Ariyo, A.A., Adewumi, A.O. and Ayo, C.K., 2014, March. Stock price prediction using the ARIMA model. In *2014 UKSim-AMSS 16th international conference on computer modelling and simulation* (pp. 106-112). IEEE.

[10] Kirubakaran, S.S., Selvadurai, J.J., Ashok, V., Nivin, J.A., Aron, A.D. and Bijolin, E., 2023, April. A Brief Assessment on Stock Market Forecast using Long Short Term Memory (LSTM) Algorithm. In *2023 7th International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 1473-1478). IEEE.

[11] Nasir, J., Aamir, M., Haq, Z.U., Khan, S., Amin, M.Y. and Naeem, M., 2023. A new approach for forecasting crude oil prices based on stochastic and deterministic influences of LMD Using ARIMA and LSTM Models. *IEEE Access*, *11*, pp.14322-14339.

[12] Alvarez-Ramirez, J., Soriano, A., Cisneros, M. and Suarez, R., 2003. Symmetry/anti-symmetry phase transitions in crude oil markets. *Physica A: Statistical Mechanics and its Applications*, *322*, pp.583-596.

[13] Malki, A., Atlam, E.S., Hassanien, A.E., Ewis, A., Dagnew, G. and Gad, I., 2022. SARIMA model-based forecasting required number of COVID-19 vaccines globally and empirical analysis of peoples' view towards the vaccines. *Alexandria Engineering Journal*, *61*(12), pp.12091-12110.

[14] Kumar, I., Tripathi, B.K. and Singh, A., 2023. Attention-based LSTM network-assisted time series forecasting models for petroleum production. *Engineering Applications of Artificial Intelligence*, *123*, p.106440.

[15] Babu, S.B., Suneetha, A., Babu, G.C., Kumar, Y.J.N. and Karuna, G., 2018. Medical disease prediction using grey wolf optimization and auto encoder based recurrent neural network. *Periodicals of Engineering and Natural Sciences*, *6*(1), pp.229-240.

[16] Kanchymalay, K., Sallehuddin, R., Salim, N. and Hashim, U.R., 2017, May. Time series based forecasting for crude palm oil price utilizing neural network algorithms. In *2017 6th ICT International Student Project Conference (ICT-ISPC)* (pp. 1-4). IEEE.

[17] Wu, G. and Zhang, Y.J., 2014. Does China factor matter? An econometric analysis of international crude oil prices. *Energy Policy*, *72*, pp.78-86.

[18] Wati, M., Haviluddin, A.M., Septiarini, A. and Hatta, H.R., 2023. Autoregressive Integrated Moving Average (ARIMA) Model for Forecasting Indonsesian Crude Oil Price. *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI)*, *9*(3), pp.720-730.

[19] Shumway, R.H., Stoffer, D.S., Shumway, R.H. and Stoffer, D.S., 2017. ARIMA models. *Time series analysis and its applications: with R examples*, pp.75-163.

[20] Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S. and Ciccozzi, M., 2020. Application of the ARIMA model on the COVID-2019 epidemic dataset. *Data in brief*, *29*, p.105340.

[21] Ariyo, A.A., Adewumi, A.O. and Ayo, C.K., 2014, March. Stock price prediction using the ARIMA model. In *2014 UKSim-AMSS 16th*

*international conference on computer modelling and simulation* (pp. 106-112). IEEE.

[22] Jadhav, V., Chinnappa, R.B. and Gaddi, G.M., 2017. Application of ARIMA model for forecasting agricultural prices.

[23] Tse, R.Y., 1997. An application of the ARIMA model to real-estate prices in Hong Kong. *Journal of Property Finance*, *8*(2), pp.152-163.

[24] Guha, B. and Bandyopadhyay, G., 2016. Gold price forecasting using ARIMA model. *Journal of Advanced Management Science*, *4*(2).

[25] Juberias, G., Yunta, R., Moreno, J.G. and Mendivil, C., 1999, April. A new ARIMA model for hourly load forecasting. In *1999 IEEE Transmission and Distribution Conference (Cat. No. 99CH36333)* (Vol. 1, pp. 314-319). IEEE.

[26] Mondal, P., Shit, L. and Goswami, S., 2014. Study of effectiveness of time series modeling (ARIMA) in forecasting stock prices. *International Journal of Computer Science, Engineering and Applications*, *4*(2), p.13.

[27] Wang, Y. and Guo, Y., 2020. Forecasting method of stock market volatility in time series data based on mixed model of ARIMA and XGBoost. *China Communications*, *17*(3), pp.205-221.

[28] Alim, M., Ye, G.H., Guan, P., Huang, D.S., Zhou, B.S. and Wu, W., 2020. Comparison of ARIMA model and XGBoost model for prediction of human brucellosis in mainland China: a time-series study. *BMJ open*, *10*(12), p.e039676.

[29] Jayaraman, V., Parthasarathy, S., Lakshminarayanan, A.R. and Singh, H.K., 2021, September. Predicting the quantity of municipal solid waste using XGBoost model. In *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 148-152). IEEE.

[30] Poola, K. and Sekhar, P.H., 2021. Prediction of rainfall by using extreme gradient boost (XG boost) in Vishakapattanam area, Andhra Pradesh. *International Journal of Statistics and Applied Mathematics*.

[31] Kumar, D.S., Thiruvarangan, B.C., Vishnu, A., Devi, A.S. and Kavitha, D., 2022, March. Analysis and prediction of stock price using hybridization of sarima and xgboost. In *2022 International Conference on Communication, Computing and Internet of Things (IC3IoT)* (pp. 1-4). IEEE.

[32] Zha, W., Liu, Y., Wan, Y., Luo, R., Li, D., Yang, S. and Xu, Y., 2022. Forecasting monthly gas field production based on the CNN-LSTM model. *Energy*, *260*, p.124889.

[33] Shahid, F., Zameer, A. and Muneeb, M., 2021. A novel genetic LSTM model for wind power forecast. *Energy*, *223*, p.120069.

[34] Livieris, I.E., Pintelas, E. and Pintelas, P., 2020. A CNN–LSTM model for gold price time-series forecasting. *Neural computing and applications*, *32*, pp.17351-17360.

[35] Lin, Y., Yan, Y., Xu, J., Liao, Y. and Ma, F., 2021. Forecasting stock index price using the CEEMDAN-LSTM model. *The North American Journal of Economics and Finance*, *57*, p.101421.

[36] Wöllmer, M., Kaiser, M., Eyben, F., Schuller, B. and Rigoll, G., 2013. LSTM-modeling of continuous emotions in an audiovisual affect recognition framework. *Image and Vision Computing*, *31*(2), pp.153-163.

[37] Alkhammash, E.H., Kamel, A.F., Al-Fattah, S.M. and Elshewey, A.M., 2022. Optimized multivariate adaptive regression splines for predicting crude oil demand in Saudi arabia. *Discrete Dynamics in Nature and Society*, *2022*(1), p.8412895.