Bima Journal of Science and Technology, Vol. 8(4A) Dec, 2024 ISSN: 2536-6041

DOI: 10.56892/bima.v8i4.1148

A Review of Machine Learning Models Used in Forecasting of Petroleum Products Prices

Odion P.O.[1], Abdulrauf A.[2]*, Abodunrin O. B.[3] and Adam A.Y.[2]

[1]Department of Computer Science, Nigerian Defence Academy, Kaduna State, Nigeria
[2]Department of Computer Science, Federal Polytechnic Kaltungo, Gombe State, Nigeria
[3]Department of Computer Science, Federal School of Surveying Oyo, Oyo State, Nigeria
Corresponding Author email: abdulrashid.abdulrauf2022@nda.edu.n

ABSTRACT

Premium Motor Spirit (PMS) and other crude oil commodities play a pivotal role in shaping the nation's economy, influencing inflation, transportation costs, and overall economic stability. Accurate forecasting of these prices is essential for managing the inflation and ensuring economic stability. This study provides a systematic review of the application of machine learning models for predicting PMS and other crude oil commodity prices, focusing on literature from 2018 to 2024. The review identifies Support Vector Regression (SVR) and Long Short-Term Memory (LSTM) networks as the most reliable models for forecasting PMS prices. For crude oil and natural gas, models such as Artificial Neural Networks (ANN) and Random Forest are more effective. Additionally, hybrid models that combine ARIMA with machine learning techniques offer improved accuracy by capturing both short-term fluctuations and long-term trends. The findings underscore that no single model consistently outperforms others across all commodities, highlighting the need for tailored approaches based on the specific characteristics of each commodity and dataset.

Keywords: Crude Oil, Forecasting, Machine Learning models, Petroleum, Price and Review.

INTRODUCTION

Commodities have long played a fundamental role in shaping economies, serving as essential drivers of transactions across various sectors. In Nigeria, the cost of Premium Motor Spirit (PMS) and Natural Gas has a profound impact on the economy, affecting inflation, transportation costs, and overall economic stability (Abdullahi, 2024). The prices of PMS and natural gas are highly volatile, influenced by factors such as global oil price fluctuations, domestic refining capacity, and government policies like fuel subsidies (Abdullahi, 2024) . Studies have shown that fluctuations in PMS prices and other crude oil commodities have a direct effect on inflation in Nigeria, particularly in the short term, where price increases in PMS lead to immediate rises in the general cost of goods and services (Abdullahi, 2024; Bassey & Ekong, 2019).

Given this relationship, accurately forecasting PMS prices and other crude oil commodities is crucial for guiding policy decisions and economic planning, providing a critical tool for managing inflation and maintaining economic stability (Dimitriadou et al., 2018) . Machine learning techniques have proven to be powerful tools for forecasting crude oil prices due to their ability to effectively uncover relationships between commodity prices, historical data, and other variables (Agbo et al., 2022).

Machine learning enables the analysis of vast amounts of data that would be impossible for humans to process, allowing conclusions to be drawn from these analyses (Candanedo et al., 2018; Briones et al., 2016). These algorithms can identify behavioral patterns by analyzing the variables within datasets, helping to discover which factors are key drivers of changes in these patterns (Gonzalez-Briones et al., 2019) . Numerous machine learning models have been suggested for predicting crude oil prices (Zhao et al., 2017; Chen et al., 2017).

Bima Journal of Science and Technology, Vol. 8(4A) Dec, 2024 ISSN: 2536-6041

DOI: 10.56892/bima.v8i4.1148

Wang et al. (2021) highlighted that while certain methods perform well in specific scenarios, no single approach consistently outperforms the others. Each method comes with its own strengths and limitations. Due to the nonlinear, uncertain, and dynamic nature of energy price fluctuations, extracting meaningful patterns and creating a model that can reliably predict energy prices with a single forecasting technique remains a challenge.

In this regard, forecasting is crucial for predicting future trends in crude oil prices (Foroutan & Lahmiri, 2024). Advanced analytical models (Kou et al., 2021; Li et al., 2022; Lahmiri, 2023a), statistical methods (Lahmiri et al., 2022; Lahmiri, 2023b), and machine learning (Lahmiri et al., 2023) and deep learning algorithms (Amirifar et al., 2023; Amirshahi & Lahmiri, 2023a, 2023b) have been used to process large datasets, detect patterns, and make predictions valuable to short-term traders and long-term investors (Das et al., 2022; Jiang et al., 2022; Liang et al., 2023). Research shows that machine learning models often outperform traditional econometric methods in predictive accuracy (Dimitriadou et al., 2018; Herrera et al., 2019).

Support Vector Regression (SVR) and Convolutional Neural Networks (CNN) have demonstrated exceptional effectiveness in addressing the complexities of oil price fluctuations (Luo et al., 2019; Tissaoui et al., 2022). These techniques provide more reliable forecasts by uncovering hidden patterns that traditional statistical methods often fail to capture.

The context of Nigeria, characterized by its unique economic challenges and reliance on imported refined petroleum products, necessitates a tailored approach to price forecasting. Some crude oil commodities such as CNG are more affordable than PMS and the country's limited refining capacity means that fluctuations in global crude oil prices directly affect PMS prices domestically (Raifu & Afolabi, 2024; Patrick, 2021). Furthermore, the removal of fuel subsidies has been a contentious issue, with potential inflationary effects that complicate the economic landscape (Oladimeji & Umar, 2024) . Understanding these dynamics is crucial for developing effective forecasting models that can assist policymakers in mitigating adverse economic impacts.

Motivated by this, the study examines forecasting methodologies in the context of crude oil to improve the accuracy of price predictions. While recent advancements in deep learning models hold significant promise for time-series forecasting, their application in crude oil price prediction remains limited. This research aims to bridge this gap by addressing the following key questions.

- Which machine learning model provides the most reliable and accurate predictions for PMS and natural gas prices?
- In connection to the first question, is there a particular model that consistently outperforms others in predicting PMS, natural gas, and other crude oil products?
- Are hybrid models effective in forecasting PMS and natural gas prices?

The structure of this paper is as follows: the "Literature Review" section presents an overview of relevant prior research. The "Methodology" section outlines the approaches used in this study. The "Results and Discussion" section provides an analysis of the reviewed publications, presents the findings, and discussion of results. Finally, the "Conclusion" section offers a summary of the paper.

REVIEW OF RELATED LITERATURES

Machine learning techniques have gained popularity for forecasting Premium Motor Spirit (PMS) prices due to the complexities of the oil market and the significant economic impact of fuel prices. Crude oil price volatility, driven by both domestic and global factors, requires adaptive predictive

DOI: 10.56892/bima.v8i4.1148

models. Recent literature emphasizes several machine learning techniques used to improve the accuracy of PMS and natural gas price forecasts.

One prominent approach is the use of Support Vector Regression (SVR), which has been shown to effectively model nonlinear relationships in time series data. Studies indicate that SVR, along with Kernel Ridge Regression (KRR), demonstrates superior performance in forecasting tasks compared to traditional econometric models (Hoque & Aljamaan, 2021) . This is particularly relevant in the case of PMS and Natural gas pricing, where the underlying factors influencing price fluctuations are often nonlinear and complex. The ability of SVR to handle such complexities makes it a valuable tool for predicting PMS and Natural gas prices in Nigeria, where market conditions can change rapidly due to both local and international influences (Agbo et al., 2022).

Su et al. (2022) proposed a hybrid forecasting model to improve the accuracy of predicting crude oil futures prices. The model was developed in three stages: initially, crude oil futures were predicted using SVM, ELM, and LSTM models. These predictions were then reconstructed using factor reconstruction (FR) and refined with XGBoost for a secondary prediction. Finally, residual sequences were forecasted using the GNB method, and the final prediction was derived by combining the residual and secondary predictions. When tested on OPEC's historical crude oil futures data, the model outperformed 16 general models and 4 recent models, demonstrating the effectiveness of combining multiple prediction techniques and residual sequence analysis.

Obulezi et al. (2023) compared three machine learning algorithms Artificial Neural Network (ANN), Boosted Regression Trees (BRT), and Support Vector Regression (SVR) to model transportation costs before and after policy changes. The study considered variables such as travel time, distance, average passengers per carriage, and fuel price. BRT outperformed the other models based on performance metrics. Time in minutes was found to be the most important factor in both scenarios, while the least important feature was average passengers per carriage before the policy and fuel price per litre after the policy.

Shahzad et al. (2024) explored the interactions between energy commodity futures, oil price futures, and carbon emission futures from a forecasting perspective, with an emphasis on environmental sustainability. Using daily data from January 2018 to October 2021, the study employed machine learning techniques, including multiple linear regression (MLR), artificial neural networks (ANN), support vector regression (SVR), and long short-term memory (LSTM). The results indicated that nonlinear models outperformed linear models in capturing the relationships between crude oil, heating oil, and carbon emission futures prices. Furthermore, the study found that carbon emission futures prices responded nonlinearly to significant fluctuations in oil and natural gas prices.

Wang et al. (2020) examined natural gas price prediction using three models: support vector regression (SVR), long-term short-term memory (LSTM), and an improved pattern sequence similarity search (IPSS). They proposed a new weighted hybrid model combining these techniques. Using U.S. daily natural gas spot prices, the model was trained with data and the results demonstrated that the hybrid model outperformed individual models, with IPSS showing the highest prediction accuracy.

Mohamed and Messaadia (2023) emphasize that artificial neural networks and support vector machines (SVMs) are among the most widely used artificial intelligence techniques for forecasting crude oil prices. These studies collectively underscore the

DOI: 10.56892/bima.v8i4.1148

increasing importance of advanced forecasting methods in improving the accuracy and reliability of oil price predictions. Additionally, the integration of various machine learning and deep learning approaches has been explored in several studies to enhance oil price forecasting. Such studies are as follows;

Lin et al. (2022) proposed an innovative hybrid model for predicting crude oil futures prices. This model combines BiLSTM, Attention mechanisms, and CNN to effectively capture both short-term and long-term dependencies within the time series data. By decomposing the crude oil futures price data using wavelet transform, the authors were able to predict the subsequences more accurately. However, the proposed model surpassed other comparative models across various evaluation metrics, demonstrating its effectiveness in predicting crude oil futures prices.

Fang et al. (2023) presented an enhanced slope-based method (ISBM) combined with empirical mode decomposition (EMD) and feed-forward neural network (FNN), referred to as the EMD-ISBM-FNN model, to decompose and forecast Brent crude oil prices. They used the ISBM-based EMD method to break down the time series into intrinsic mode functions (IMFs) and residuals, which were then fed into the FNN model for training. The study compared the forecasting performance of the EMD-ISBM-FNN model with benchmark models, including the EMD-FNN and FNN models, across two research frameworks and three strategies.

Fang et al. (2023) proposed a hybrid forecasting model, FinBERT-VMD-Att-BiGRU, to address the challenges of predicting crude oil prices due to the market's noisy and non-stationary nature. They applied FinBERT to extract sentiment information from news, used variational mode decomposition (VMD) to break down complex price series into simpler subseries, employed an attention mechanism to assign weights to input features, and utilized the BiGRU model for forecasting. Their experiments demonstrated that the sentiment-enhanced approach improved forecasting accuracy, and the weighting scheme in sentiment analysis further enhanced performance.

Liang et al. (2023) developed a novel Deep Reinforcement Learning (DRL) algorithm for multi-step ahead crude oil price forecasting across three major commodity exchanges. The algorithm introduced two key improvements: a dynamic action exploration mechanism based on stochastic processes aligned with commodity price fluctuations, and a dynamic update policy for network parameters grounded in approximate optimization theory. In their experiments, the algorithm's performance was compared with five state-of-the-art methods. The results showed that DRL's forecasting capability was highly effective and could be extended to other natural resources.

Sun et al. (2022) proposed a novel method for crude oil futures price forecasting by decomposing the original price series using the Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (ICEEMDAN) and reconstructing the sub-sequences with Permutation Entropy (PE) was used to decompose the data into high-frequency, low-frequency, and trend components. The Kernel Extreme Learning Machine (KELM), optimized by the Chaotic Sparrow Search Algorithm (CSSA), was employed to predict the low-frequency and trend components. Meanwhile, the high-frequency component was further decomposed using Empirical Mode Decomposition (EMD) and predicted using both the PE and CSSA-KELM models. The final forecasting value was obtained by nonlinearly integrating the results of all components with the CSSA-KELM model.

DOI: 10.56892/bima.v8i4.1148

Foroutan & Lahmiri (2024) conducted a study to forecast crude oil and precious metals prices, focusing on West Texas Intermediate (WTI), Brent, gold, and silver markets. They implemented 16 deep and machine learning models, including LSTM, BiLSTM, GRU, BiGRU, CNN, and TCN variants, and compared their performance with baseline models like Random Forest, LightGBM, Support Vector Regression, and K-Nearest Neighbors. Their results showed that the TCN model performed best for WTI, Brent, and silver, achieving the lowest MAE values, while BiGRU excelled for gold.

Agbo et al. (2022) applied a visualization-based approach to forecast Premium Motor Spirit (PMS) prices across selected states in Nigeria, including Kogi, Lagos, Bauchi, Borno, Anambra, Rivers, and Abuja - FCT. The study utilized a dataset of monthly fuel prices collected by the Petroleum Products Pricing Regulatory Agency (PPPRA) and the National Bureau of Statistics (NBS) from January 2016 to December 2021. The analysis revealed continuous fluctuations in PMS prices across Nigeria's geopolitical regions, particularly in the North East. These fluctuations were influenced by disparities in fuel availability, regulatory enforcement, and price consistency during the study period.

Sofianos et al., (2024) focused on forecasting daily gasoline spot prices for New York and Los Angeles using a dataset that included gasoline prices and 128 other relevant variables. The study applied three tree-based machine learning algorithms: decision trees, random forest, and XGBoost, with a variable importance measure (VIM) technique to rank the most important explanatory variables. The optimal model, a random forest, achieved a mean absolute percent error (MAPE) of 3.23% and 3.78% in out-of-sample predictions. The first lag, AR (1), of gasoline was identified as the most critical variable in both markets, with energy-related variables dominating the top five.

Escribano & Wang (2021) proposed a new mixed random forest (RF) approach to model nonlinearities in forecasting gasoline prices. Their method identified explanatory variables with nonlinear impacts, threshold values, and the closest parametric approximations. The study applied the model to weekly gasoline prices in Spain from 2010 to 2019, which were cointegrated with international oil prices and exchange rates. The mixed RF approach successfully identified nonlinearities in the error correction term and the rate of change in oil prices.

Wang et al. (2021) developed a novel hybrid model for natural gas price prediction, combining CEEMDAN-SE and PSO-ALS-GRU to enhance accuracy. The model decomposes natural gas price series into sub-sequences using CEEMDAN-SE, predicts each sub-sequence with the PSO-ALS-GRU network, and then reconstructs the overall forecast. By optimizing GRU network hyperparameters with PSO-ALS, the model addresses nonlinear complexities and long-term dependencies.

Cen & Wang (2019) proposed a deep-learning crude oil price prediction using Long Short-Term Memory (LSTM) models. They applied transfer learning to extend training data, introducing a novel algorithm called data transfer with prior knowledge, which enhances data availability through three data types. Additionally, ensemble empirical mode decomposition was used to decompose time series into intrinsic mode functions, improving model training.

Li et al. (2021) proposed a combined model for forecasting monthly natural gas prices, integrating Deep Belief Network (DBN) with Variational Mode Decomposition (VMD) and Particle Swarm Optimization (PSO) as a hybrid approach.

Subsequent studies have also shown that Auto-Regressive Integrated Moving Average (ARIMA) was adopted to predict future trends in crude oil price forecasting

DOI: 10.56892/bima.v8i4.1148

(Shambulingappa, 2020) . Also, Yu et al., (2021) introduced a "decomposition-prediction-integration" approach using an STL-(ELM+ARIMA) model for forecasting fuel oil prices, demonstrating that this combination achieved the highest forecasting accuracy by effectively utilizing the strengths of both non-parametric and parametric models.

Alrweili & Fawzy (2022) proposed a hybrid ARIMA-ANN model for time series forecasting to capture both linear and nonlinear relationships in crude oil price data. Using monthly crude oil prices in Saudi Riyal from July 2001 to May 2021, they trained the model on 215 observations and tested it on 24. The hybrid model outperformed individual ARIMA and ANN models. Siddiqui (2019) developed an Autoregressive Neural Network (ARNN) model to enhance the forecasting of daily spot gas prices, tackling the nonlinear and nonstationary characteristics of gas price data.

Zhao et al. (2018) developed a three-step-ahead fuel cost prediction model using an ARIMA approach with historical data from Form EIA-923 and Texas natural gas fuel costs. By incorporating the spot price from a Texas gas hub, the model enhances prediction accuracy. The forecasted data was fit to a normal distribution, and Kullback-Leibler divergence was used to assess differences between real and predicted fuel cost distributions. Results indicate that the proposed forecasting algorithm is generally effective and merits further exploration.

Sokkalingam et al. (2021) examined the volatility of petroleum fuel prices in Malaysia, which have been regulated by the Automatic Price Mechanism (APM) since 2004. Despite the APM and the shift to a Managed Float System in 2016, fuel prices remain unstable due to factors such as international crude oil price fluctuations, foreign exchange volatility, and reduced subsidies. To address the challenges in predicting fuel prices, the authors employed the Autoregressive Integrated Moving Average (ARIMA) time series model for modelling and forecasting.

Agyare et al. (2024) performed a comparative analysis of gasoline and diesel price forecasting in Ghana using ARIMA and SARIMA models, based on monthly data from January 2016 to December 2023. They found ARIMA (0, 1, 2) to be the best model for gasoline and ARIMA (2, 1, 3) for diesel, with SARIMA models outperforming both for overall forecasting accuracy.

This study offers valuable contributions to the literature on forecasting prices in commodity markets. A summary of the reviewed studies on petroleum commodity price forecasting is presented in Table 1.

Table 1: Summary of reviewed publications

Authors & Year	Title	Focus of Study	Methodology	Strengths	Weaknesses
Hoque & Aljamaan (2021)	Impact of hyperparameter tuning on machine learning models in stock price forecasting	Forecasting PMS and Natural gas prices in Nigeria	SVR & KRR	Effective for nonlinear relationships, superior performance compared to traditional econometric models.	Lack of comparison with other nonlinear methods
Su et al. (2022)	A new crude oil futures forecasting method based on	Forecasting crude oil futures prices	Hybrid model combining SVM, ELM, LSTM, FR,	Superior accuracy compared to other models, effective combination of	Complex model structure, may require significant computational

DOI: 10.56892/bima.v8i4.1148

	fusing quadratic forecasting with residual forecasting		XGBoost, and GNB	techniques.	resources.
Obulezi et al. (2023)	Machine learning models for predicting transportation costs inflated by fuel subsidy removal policy in Nigeria	Modeling transportation costs before and after policy changes	ANN, BRT, SVR	Effective for modeling transportation costs, BRT outperformed other models.	May require significant data preprocessing and tuning.
Shahzad et al. (2024)	Forecasting carbon emissions future prices using the machine learning methods	Forecasting energy commodity futures	MLR, ANN, SVR, LSTM	Nonlinear models outperformed linear ones, captured relationships between crude oil, heating oil, and carbon emission futures prices.	May not be suitable for short-term forecasting.
Wang et al. (2020)	Daily natural gas price forecasting by a weighted hybrid data-driven model	Forecasting natural gas prices	SVR, LSTM, IPSS	Hybrid model outperformed individual models, IPSS showed highest prediction accuracy.	May require significant computational resources for large datasets.
Mohamed & Messaadia (2023)	Artificial Intelligence Techniques for the Forecasting of Crude Oil Price: A Literature Review	Forecasting crude oil prices	ANN & SVM	Widely used for forecasting crude oil prices, effective for capturing nonlinear relationships.	May require significant data preprocessing and tuning.
Lin et al. (2022)	Forecasting crude oil futures prices using BiLSTM-Attention-CNN model with Wavelet transform	Forecasting crude oil futures prices	BiLSTM, Attention mechanisms, CNN	Effective for capturing both short-term and long-term dependencies, outperformed other models.	May require significant computational resources.
Fang et al. (2023)	A sentiment-enhanced hybrid model for crude oil price forecasting	Forecasting Brent crude oil prices	EMD-ISBM-FNN	Effective for decomposing and forecasting Brent crude oil prices.	May require domain expertise for feature engineering.
Fang et al. (2023)	Forecasting the crude oil prices with an EMD-ISBM-FNN model	Forecasting crude oil prices	FinBERT-VMD-Att-BiGRU	Effective for predicting crude oil prices with noisy and non-stationary data.	May require significant computational resources and data preprocessing.
Liang et al. (2023)	Crude oil price prediction using deep reinforcement learning	Multi-step ahead crude oil price forecasting	DRL	Effective for multi-step ahead crude oil price forecasting.	May require significant computational resources and domain expertise.
Sun et al. (2022)	A new secondary decomposition-reconstruction-	Forecasting crude oil futures prices	ICEEMDAN, PE, KELM, CSSA	Effective for decomposing and forecasting crude	May require significant computational

	ensemble approach for crude oil price forecasting			oil futures prices.	resources and domain expertise.
Foroutan & Lahmiri (2024)	Deep learning systems for forecasting the prices of crude oil and precious metals	Forecasting crude oil and precious metals prices	LSTM, BiLSTM, GRU, BiGRU, CNN, TCN	Effective for forecasting crude oil and precious metals prices.	May require significant computational resources and data preprocessing.
Agbo et al. (2022)	Visualization based approach for fuel (PMS) Price Forecast	Forecasting PMS prices in Nigeria	Visualization-based approach	Handles complexities in price fluctuations influenced by local/international factors in Nigeria.	Limited geographical scope and may not be suitable for quantitative forecasting.
Sofianos et al. (2024)	Forecasting East and West Coast Gasoline Prices with Tree-Based Machine Learning Algorithms	Forecasting daily gasoline spot prices	Decision trees, random forest, XGBoost	Effective for forecasting daily gasoline spot prices.	May require significant computational resources and data preprocessing.
Escribano & Wang (2021)	Mixed random forest, cointegration, and forecasting gasoline prices	Forecasting gasoline prices	Mixed random forest	Effective for modeling nonlinearities in forecasting gasoline prices.	May require domain expertise for feature engineering.
Wang et al. (2021)	Daily natural gas price forecasting by a weighted hybrid data-driven model	Forecasting natural gas prices	CEEMDAN-SE, PSO-ALS-GRU	Effective for natural gas price prediction.	May require significant computational resources and data preprocessing.
Cen & Wang (2019)	Crude oil price prediction model with long short term memory deep learning based on prior knowledge data transfer	Forecasting crude oil prices	LSTM with transfer learning and ensemble empirical mode decomposition	Effective for crude oil price prediction.	May require significant computational resources and data preprocessing.
Li et al. (2021)	A novel crude oil prices forecasting model based on secondary decomposition	Forecasting monthly natural gas prices	DBN, VMD, PSO	Effective for forecasting monthly natural gas prices.	May require significant computational resources and data preprocessing.
Shambulingappa (2020)	Crude oil price forecasting using machine learning	Forecasting crude oil prices	ARIMA	Effective for forecasting crude oil prices.	May not be suitable for capturing nonlinear relationships.
Yu et al. (2021)	Forecasting the price of fuel oil: a STL- (ELM+ ARIMA) combination approach	Forecasting fuel oil prices	STL-(ELM+ARIMA)	Effective for forecasting fuel oil prices.	May require significant computational resources and data preprocessing.
Alrweili & Fawzy	Forecasting crude oil prices using an	Forecasting crude oil prices	ARIMA-ANN	Effective for forecasting crude	May require significant

DOI: 10.56892/bima.v8i4.1148

(2022)	ARIMA-ANN hybrid model			oil prices.	computational resources and data preprocessing.
Siddiqui (2019)	Predicting natural gas spot prices using artificial neural network	Forecasting daily spot gas prices	ARNN	Effective for forecasting daily spot gas prices.	May require significant computational resources and data preprocessing.
Zhao et al. (2018)	Improvement to the prediction of fuel cost distributions using ARIMA model	Forecasting fuel cost	ARIMA	Effective for three-step-ahead fuel cost prediction.	May not be suitable for capturing nonlinear relationships.
Sokkalingam et al. (2021)	Forecasting petroleum fuel price in malaysia by arima model	Forecasting fuel prices in Malaysia	ARIMA	Effective for forecasting fuel prices in Malaysia.	May not be suitable for capturing nonlinear relationships.
Agyare et al. (2024)	Predicting Petrol and Diesel Prices in Ghana, A Comparison of ARIMA and SARIMA Models	Forecasting gasoline and diesel prices in Ghana	ARIMA & SARIMA	Effective for forecasting gasoline and diesel prices in Ghana.	May not be suitable for capturing nonlinear relationships.

MATERIALS AND METHODS

In order to choose the related articles to review, we conducted a systematic literature review of relevant research in the field of machine learning performance in forecasting PMS and other crude oil commodities, published in the recent five years (2018–2024). The published papers involved journals, conference proceeding and technical reports.

Data Source

The search was conducted using Google Scholar, ResearchGate, Springer, and IEEE Xplore databases. The first approach utilized keywords such as premium motor spirit price, natural gas price, crude oil commodities price, and machine learning. The second approach employed the search string: “price forecasting using machine learning approaches” OR “premium motor spirit price forecasting” OR “natural gas price forecasting” OR “crude oil commodities price forecasting” OR “premium motor spirit, natural gas, and crude oil commodities price forecasting using machine learning techniques” to retrieve research papers.

Screening of Relevant Papers

After retrieving the papers from the databases, the next step was to screen them for relevance based on inclusion criteria. These criteria considered studies published between 2018 and 2024, research focused on applying machine learning approaches to forecast natural gas and crude oil prices, and articles specifically addressing the forecasting of premium motor spirit (PMS) prices using machine learning techniques.

Similarly, exclusion criteria were applied to eliminate publications published before 2018, articles without full-text availability, duplicates, papers not written in English, and studies that did not focus on the application of machine learning in forecasting the prices of crude oil commodities.

Keywording on the Basis of the Abstract

Abstracts are read to identify keywords and concepts that reflect the papers contribution. Using the specified keywords and search string, a total of 136 publications were retrieved from various electronic databases. After the screening, which was based on the title of papers, year of publication, relevancy,

DOI: 10.56892/bima.v8i4.1148

111 papers were discarded, leaving a total of 25 publication papers as shown in table 2 for further reviewing and analysis to obtain information.

RESULTS

The analysis of the reviewed publications on PMS and other crude oil commodities is shown in Table 2.

Table 2: Analysis of Reviewed Publications on PMS and other crude oil commodities.

Commodities	Number Of Publications	Reviewed Publications
Premium Motor Sprit	6	(Hoque & Aljamaan, 2021; Su et al., 2022; Obulezi et al., 2023; Shahzad et al., 2024; Agbo et al., 2022; Sokkalingam et al., 2021)
Compressed Natural Gas	1	(Hoque & Aljamaan, 2021)
Natural Gas	5	(Shahzad et al., 2024; Wang et al., 2020; Wang et al., 2021; Li et al. 2021; Zhao et al., 2018)
Crude Oil	9	(Su et al., 2022; Mohamed & Messaadia, 2023; Lin et al., 2022; Fang et al., 2023; Fang et al., 2023; Liang et al., 2023; Sun et al., 2022; Foroutan & Lahmiri, 2024; Cen & Wang, 2019; Alrweili & Fawzy, 2022)
Gasoline	3	(Sofianos et al. 2024; Escribano & Wang, 2021; Agyare et al., 2024)
Diesel	1	(Agyare et al., 2024)

The barchart in figure 1 shows the graphical representation of PMS and other crude oil commodities in reviewed publications.

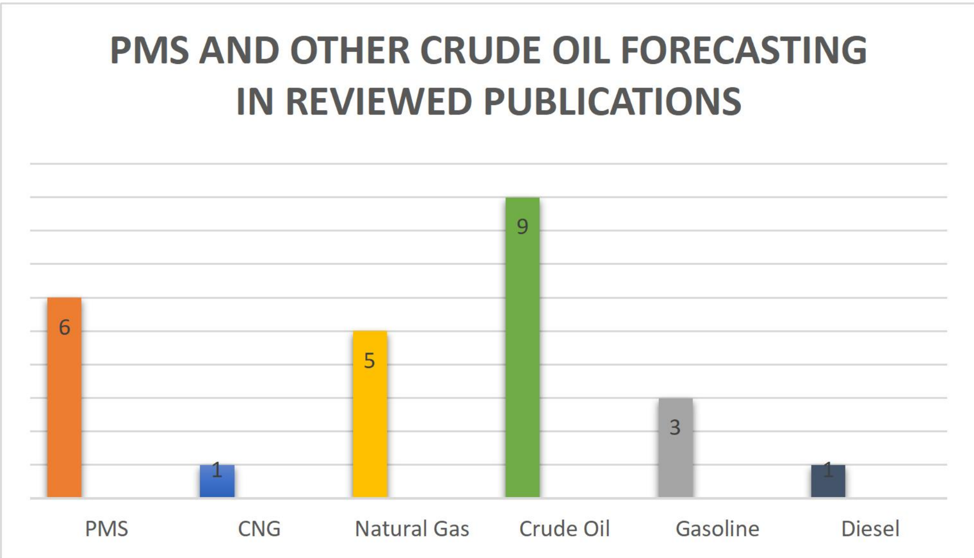

Figure 1: Graphical representation of PMS and other crude oil commodities in reviewed publications.

The Models Results by the Reviewed Publications

The figure 2 shows the results of the models used by the reviewed publications. It highlights the frequency of each method being applied in the reviewed publications.

- Support Vector Machines, Artificial Neural Networks, and ARIMA are the most frequently cited, with each being referenced 4 times.
- Long Short-Term Memory (LSTM) appears 3 times, while other methods like Convolutional Neural Networks and Attention Mechanisms are referenced twice.
- Many other algorithms are applied once, including Boosted Ridge Regression,

Bima Journal of Science and Technology, Vol. 8(4A) Dec, 2024 ISSN: 2536-6041

DOI: 10.56892/bima.v8i4.1148

Random Forest, and Decision Trees, among others.

- Methods like PSO (Particle Swarm Optimization) and SARIMA are referenced multiple times as well.

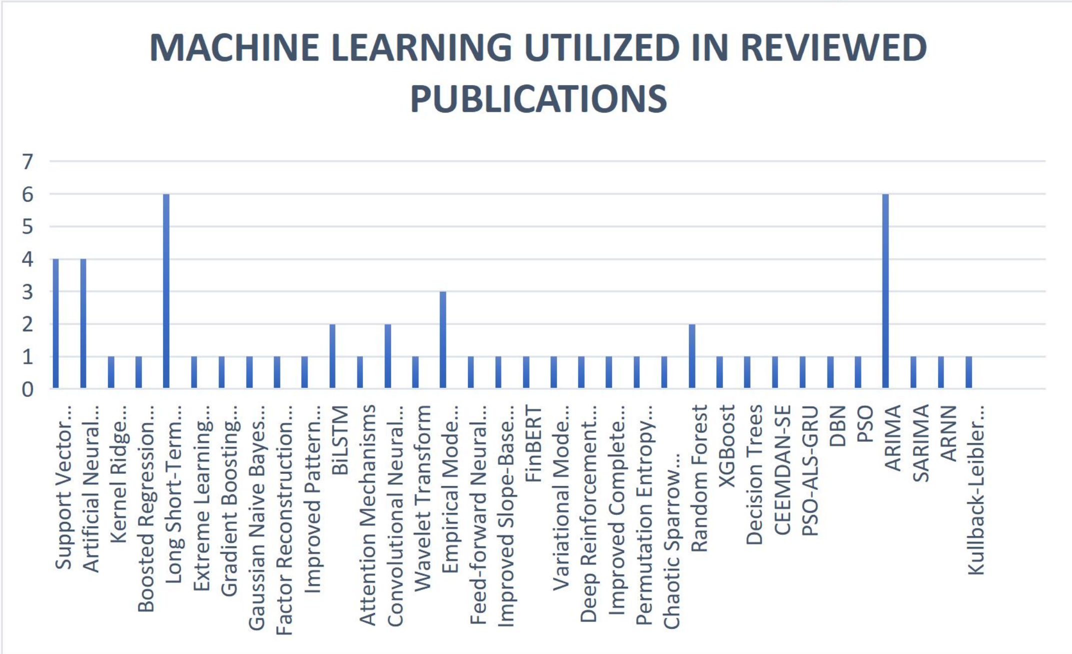

Figure 2: Visualization of machine learning techniques used in the reviewed publications.

DISCUSSION

The reviewed publications on machine learning models for predicting the prices of Premium Motor Spirit (PMS) and natural gas in Nigeria reveal diverse approaches. Given the volatility of crude oil prices, especially in Nigeria where the economy is heavily influenced by global oil prices, an accurate forecasting model is critical for managing inflation and ensuring economic stability. This study evaluated machine learning models applied to the forecasting of PMS and other crude oil commodities through a systematic review of publications from 2018 to 2024

One key observation is that machine learning techniques, particularly Support Vector Regression (SVR), Long Short-Term Memory (LSTM) networks, and Artificial Neural Networks (ANN), are frequently applied in this domain due to their ability to capture complex, nonlinear patterns in historical price data. These models have demonstrated better performance than traditional econometric approaches in handling the dynamic and often chaotic nature of energy prices. However, each model has its advantages depending on the commodity being forecasted, as shown in the review of publications. For instance, PMS and other crude oil prices fluctuations are subject to different influencing factors, and machine learning models that incorporate these unique factors provide more accurate forecasts.

The reviewed publication also answers the research questions, identifying Support Vector Regression (SVR) and Long Short-Term Memory (LSTM) as the most reliable models for predicting PMS and other crude oil commodity prices due to their ability to capture nonlinear relationships and long-term dependencies in time-series data. While no single model consistently outperforms others across all commodities, SVR and LSTM excel for PMS, whereas models like Artificial Neural Networks (ANN) and Random Forest are more effective for crude oil and natural gas.

DOI: 10.56892/bima.v8i4.1148

Hybrid models, such as ARIMA combined with machine learning techniques, offer enhanced accuracy by leveraging both statistical and machine learning strengths. Each model has unique properties: SVR and LSTM excel at handling complex data patterns, ANN is versatile but prone to overfitting, ARIMA is strong in short-term trend analysis, and Random Forest captures key features but may lack precision in time-series forecasting.

CONCLUSION

This study systematically reviews the use of machine learning models for predicting Premium Motor Spirit (PMS) and other crude oil commodity prices, emphasizing the critical role of accurate price predictions for economic stability in Nigeria. It identifies Support Vector Regression (SVR) and Long Short-Term Memory (LSTM) networks as the most reliable models for PMS due to their ability to capture nonlinear relationships and long-term dependencies. However, no single model consistently outperforms others across all commodities; for example, Artificial Neural Networks (ANN) and Random Forest models are more effective for crude oil and natural gas forecasting. Additionally, hybrid models combining traditional statistical methods like ARIMA with machine learning techniques enhance accuracy by addressing both short-term fluctuations and long-term trends. The study highlights the need to select models based on the specific characteristics of the commodity and dataset, contributing to improved price forecasting precision for informed policy making and economic planning.

REFERENCES

Abdullahi, I. (2024). Petroleum products pricing and price stability in nigeria: an ardl investigation. *Asian Journal of Advanced Research and Reports, 18*(5), 188-200. https://doi.org/https://doi.org/10.9734/ajarr/2024/v18i5647

Agbo, S. O., Ifeoluwa, Y. P., & Eric, A. S. (2022). Forecasting Premium Motor Spirit (PMS) and Energy Commodities Prices Using Machine Learning Techniques. A Review. *UMYU Scientifica, 1*(1), 194-203. https://doi.org/https://doi.org/10.56919/usci.1122.025

Agbo, S. O., Yemi-Peters, V. I., & Adewumi, S. E. (2022). Visualization based approach for fuel (PMS) Price Forecast. *FUW Trends in Science & Technology Journal, 7*(2), 1156-1160.

Agyare, S., Odoi, B., & Wiah, E. N. (2024). Predicting Petrol and Diesel Prices in Ghana, A Comparison of ARIMA and SARIMA Models. *Asian Journal of Economics, Business and Accounting, 24*(5), 594-608. https://doi.org/https://doi.org/10.9734/ajeba/2024/v24i51333

Alrweili, H., & Fawzy, H. (2022). Forecasting crude oil prices using an ARIMA-ANN hybrid model. *Journal of Statistics Applications & Probability, 11*(3), 845-855. https://doi.org/http://dx.doi.org/10.18576/jsap/110308

Amirifar, T., Lahmiri, S., & Zanjani, M. K. (2023). An NLP-deep learning approach for product rating prediction based on online reviews and product features. *IEEE Transactions on Computational Social Systems, 2023*, 1-13. https://doi.org/https://doi.org/10.1109/TCSS.2023.3290558

Amirshahi, B., & Lahmiri, S. (2023a). Hybrid deep learning and GARCH-family models for forecasting volatility of cryptocurrencies. *Machine Learning with Applications, 12*, 1-14.

Amirshahi, B., & Lahmiri, S. (2023b). Investigating the effectiveness of Twitter sentiment in cryptocurrency close price prediction by using deep learning. *Expert Systems, 2023,* 1-23.

DOI: 10.56892/bima.v8i4.1148

https://doi.org/https://doi.org/10.1111/exsy.13428

Briones, A. G., Chamoso, P., & Barriuso, A. (2016). Review of the main security problems with multi-agent systems used in e-commerce applications. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, 5*(3), 55-61.

Candanedo, I. S., Nieves, E. H., González, S. R., Martín, M. T., & Briones, A. G. (2018). Machine learning predictive model for industry 4.0. In Knowledge Management in Organizations. *In International Conference on Knowledge Management in Organizations.* Žilina, Slovakia: Springer International Publishing. https://doi.org/https://doi.org/10.1007/978-3-319-95204-8_42

Cen, Z., & Wang, J. (2019). Crude oil price prediction model with long short term memory deep learning based on prior knowledge data transfer. *Energy, 169*, 160-171.

Chen, Y., He, K., & Tso, G. K. (2017). Forecasting crude oil prices: a deep learning based model. *Procedia computer science, 122*, 300-307.

Das, S., Nayak, J., Kamesh Rao, B., Vakula, K., & Ranjan Routray, A. (2022). Gold price forecasting using machine learning techniques: Review of a decade. *Computational Intelligence in Pattern Recognition: Proceedings of CIPR, 2021*, 679-695.

Dimitriadou, A., Gogas, P., Papadimitriou, T., & Plakandaras, V. (2018). Oil market efficiency under a machine learning perspective. *Forecasting, 1*(1), 157-168. https://doi.org/https://doi.org/10.3390/forecast1010011

Escribano, Á., & Wang, D. (2021). Mixed random forest, cointegration, and forecasting gasoline prices. *International Journal of Forecasting, 37*(4), 1442-1462. https://doi.org/https://doi.org/10.1016/j.ijforecast.2020.12.008

- Fang, T., Zheng, C., & Wang, D. (2023). Forecasting the crude oil prices with an EMD-ISBM-FNN model. *Energy, 263*(6), 125407. https://doi.org/https://doi.org/10.1016/j.energy.2022.125407

Fang, Y., Wang, W., Wu, P., & Zhao, Y. (2023). A sentiment-enhanced hybrid model for crude oil price forecasting. *Expert Systems with Applications, 215(4)*, 119329. https://doi.org/https://doi.org/10.1016/j.eswa.2022.119329

Foroutan, P., & Lahmiri, S. (2024). Deep learning systems for forecasting the prices of crude oil and precious metals. *Financial Innovation, 10*(1), 1-40. https://doi.org/https://doi.org/10.1186/s40854-024-00637-z

Gonzalez-Briones, A., Hernandez, G., Corchado, J. M., Omatu, S., & Mohamad, M. S. (2019). Machine learning models for electricity consumption forecasting: a review. *In 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS).* Riyadh, Saudi Arabia: IEEE. https://doi.org/https://doi.org/10.1109/CAIS.2019.8769508

Herrera, G., Constantino, M., Tabak, B., Pistori, H., Su, J., & Naranpanawa, A. (2019). Long-term forecast of energy commodities price using machine learning. *Energy, 179*, 214-221. https://doi.org/https://doi.org/10.1016/j.energy.2019.04.077

Hoque, K., & Aljamaan, H. (2021). Impact of hyperparameter tuning on machine learning models in stock price forecasting. *IEEE Access, 9*, 163815-163830. https://doi.org/https://doi.org/10.1109/access.2021.3134138

Jiang, H., Hu, W., Xiao, L., & Dong, Y. (2022). A decomposition ensemble

DOI: 10.56892/bima.v8i4.1148

based deep learning approach for crude oil price forecasting. *Resources Policy, 78*, 102855.

Kou, G., Olgu Akdeniz, Ö., Dinçer, H., & Yüksel, S. (2021). Fintech investments in European banks: a hybrid IT2 fuzzy multidimensional decision-making approach. *Financial innovation, 7*(1), 39.

Lahmiri, S. (2023a). Multifractals and multiscale entropy patterns in energy markets under the effect of the COVID-19 pandemic. *Decision Analytics Journal, 7*, 100247.

Lahmiri, S. (2023b). A comparative study of statistical machine learning methods for condition monitoring of electric drive trains in supply chains. *Supply Chain Analytics, 2*, 100011.

Lahmiri, S., Bekiros, S., & Avdoulas, C. (2023). A comparative assessment of machine learning methods for predicting housing prices using Bayesian optimization. *Decision Analytics Journal, 6*, 100166.

Lahmiri, S., Bekiros, S., & Bezzina, F. (2022). Complexity analysis and forecasting of variations in cryptocurrency trading volume with support vector regression tuned by Bayesian optimization under different kernels: An empirical comparison from a large dataset. *Expert Systems with Applications, 209*, 118349.

Li, G., Yin, S., & Yang, H. (2022). A novel crude oil prices forecasting model based on secondary decomposition. *Energy, 257*, 124684.

Li, T., Kou, G., Peng, Y., & Philip, S. Y. (2021). An integrated cluster detection, optimization, and interpretation approach for financial data. *IEEE transactions on cybernetics, 52*(12), 13848-13861.

Liang, X., Luo, P., Li, X., Wang, X., & Shu, L. (2023). Crude oil price prediction using deep reinforcement learning. *Resources Policy, 81*, 103363.

Lin, Y., Chen, K., Zhang, X., Tan, B., & Lu, Q. (2022). Forecasting crude oil futures prices using BiLSTM-Attention-CNN model with Wavelet transform. *Applied Soft Computing, 130*(2), 109723. https://doi.org/http://dx.doi.org/10.1016/j.asoc.2022.109723

Luo, Z., Cai, X., Tanaka, K., Takiguchi, T., Kinkyo, T., & Hamori, S. (2019). Can we forecast daily oil futures prices? experimental evidence from convolutional neural networks. *Journal of Risk and Financial Management, 12*(1), 9. https://doi.org/https://doi.org/10.3390/jrfm12010009

Mohamed, N. A., & Messaadia, M. (2023). Artificial Intelligence Techniques for the Forecasting of Crude Oil Price: A Literature Review. *In 2023 International Conference On Cyber Management And Engineering (CyMaEn)*. Bangkok, Thailand: IEEE. https://doi.org/https://doi.org/10.1109/CyMaEn57228.2023.10050945

Obulezi, O. J., Chinedu, E. Q., Oramulu, D. O., Etaga, H. O., Onyeizu, N. M., & Ejike, C. O. (2023). Machine learning models for predicting transportation costs inflated by fuel subsidy removal policy in Nigeria. *International Research Journal of Modernization in Engineering Technology and Science, 5*(9), 1053-1070. https://doi.org/https://www.doi.org/10.56726/IRJMETS44687

Oladimeji, L. A., & Umar, M. (2024). Fuel subsidy removal and welfare of the working class in nigeria: policy options for the state. *Journal of Social Sciences and Management Studies, 3*(2), 1-7. https://doi.org/https://doi.org/10.56556/jssms.v3i2.709

Patrick, O. N. (2021). Determinants of domestic energy prices in nigeria (1980-2020). *Resources and*

DOI: 10.56892/bima.v8i4.1148

Environmental Economics, 3(2), 305-311. https://doi.org/https://doi.org/10.25082/ree.2021.02.006

Raifu, I. A., & Afolabi, J. A. (2024). Simulating the inflationary effects of fuel subsidy removal in nigeria: evidence from a novel approach. *Energy RESEARCH LETTERS, 5*. https://doi.org/https://doi.org/10.46557/001c.94368

Shahzad, U., Sengupta, T., Rao, A., & Cui, L. (2024). Forecasting carbon emissions future prices using the machine learning methods. *Annals of Operations Research, 17*(6), 1296. https://doi.org/https://doi.org/10.3390/en17061296

Shambulingappa, H. S. (2020). Crude oil price forecasting using machine learning. *International Journal of Advanced Scientific Innovation, 1*(1), 1-11.

Siddiqui, A. W. (2019). Predicting natural gas spot prices using artificial neural network. *In 2019 IEEE 2nd International Conference on Computer Applications & Information Security (ICCAIS).* Riyadh, Saudi Arabia: 1-6. https://doi.org/10.1109/CAIS.2019.8769586

Sofianos, E., Zaganidis, E., Papadimitriou, T., & Gogas, P. (2024). Forecasting East and West Coast Gasoline Prices with Tree-Based Machine Learning Algorithms. *Energies, 17*(6), 1296. https://doi.org/https://doi.org/10.3390/en17061296

Sokkalingam, R., Sarpong-Streetor, R. M., Othman, M., Daud, H., & Owusu, D. A. (2021). Forecasting petroleum fuel price in malaysia by arima model. *In Proceedings of the 6th International Conference on Fundamental and Applied Sciences.* Springer, Singapore. https://doi.org/https://doi.org/10.1007/978-981-16-4513-6_58

Su, M., Liu, H., Yu, C., & Duan, Z. (2022). A new crude oil futures forecasting method based on fusing quadratic forecasting with residual forecasting. *Digital Signal Processing, 130*, 103691. https://doi.org/https://doi.org/10.1016/j.dsp.2022.103691

Sun, J., Zhao, P., & Sun, S. (2022). A new secondary decomposition-reconstruction-ensemble approach for crude oil price forecasting. *Resources Policy, 77*, 102762. https://doi.org/https://doi.org/10.1016/j.resourpol.2022.102762

Tissaoui, K., Zaghdoudi, T., Hakimi, A., & Nsaibi, M. (2022). Do gas price and uncertainty indices forecast crude oil prices? fresh evidence through xgboost modeling. *Computational Economics, 62*(2), 663-687. https://doi.org/https://doi.org/10.1007/s10614-022-10305-y

Wang, J., Cao, J., Yuan, S., & Cheng, M. (2021). Short-term forecasting of natural gas prices by using a novel hybrid method based on a combination of the CEEMDAN-SE-and the PSO-ALS-optimized GRU network. *Energy, 233*. https://doi.org/https://doi.org/10.1016/j.energy.2021.121082

Wang, J., Lei, C., & Guo, M. (2020). Daily natural gas price forecasting by a weighted hybrid data-driven model. *Journal of Petroleum Science and Engineering, 192*, 1-11. https://doi.org/https://doi.org/10.1016/j.petrol.2020.107240

Yu, F., Liu, Y., & Zhang, C. (2021). Forecasting the price of fuel oil: a STL-(ELM+ ARIMA) combination approach. *In Journal of Physics: Conference Series, 1903*(1), 1-6. https://doi.org/10.1088/1742-6596/1903/1/012048

Zhao, Y., Li, J., & Yu, L. (2017). A deep learning ensemble approach for crude

DOI: 10.56892/bima.v8i4.1148

oil price forecasting. *Energy Economics, 66*, 9-16.

Zhao, Z., Fu, C., Wang, C., & Miller, C. J. (2018). Improvement to the prediction of fuel cost distributions using ARIMA model. *In 2018 IEEE power & energy society general meeting.* Portland, USA: 1-5. https://doi.org/10.1109/PESGM.2018.8585984