International Journal of Applied Mathematics

Volume 38 No. 9s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

PREDICTING GAS OIL IMPORT QUANTITIES IN IRAQ USING MACHINE LEARNING TECHNIQUES

Ahmed Khalaf Zager Alsaedi[1], Ghazwan Hani Hussein[2], Fatimah Nazar Hamzah[3], Farah Alaa Mohammed[4], Ali Haitham Abdul amir[5]

[1] Physic department- Collage of science-Misan University, Iraq
ahmedkhalafzager@uomisan.edu.iq ,

[2] Computer Science department- Collage of Computer Science & IT- AlBarsh University, Iraq
ghazwan.hani@uobasrah.edu.iq ,

[3,4,5] Shatt Al-Arab University College, Basrah, Iraq, [3]fatima.nizar@sa-uc.edu.iq , [4] farah.alaa@sa-uc.edu.iq ,[5]ali.haitham.abdulamir@sa-uc.edu.iq

ABSTRACT

This study develops and evaluates predictive models for annual gas-oil import quantities in Iraq using historical production and consumption data. Implement and compare several supervised learning approaches Linear Regression, Polynomial Regression, Support Vector Regression, Random Forest, and XGBoost and propose a soft-voting ensemble (Voting Regressor) that aggregates these base learners. Models were trained with an 80/20 train/test split and evaluated using coefficient of determination (R^2), RMSE, and MAE. Results show marked differences in model behavior: Linear Regression performed poorly ($R^2\_train \approx 0.203$, $R^2\_test \approx 0.031$), confirming the relationship's nonlinearity; Polynomial and SVR achieved moderate explanatory power ($R^2\_test \approx 0.559$ and 0.582, respectively) but were sensitive to extreme observations; Random Forest yielded high accuracy with low errors ($R^2\_test \approx 0.992$, $RMSE\_test \approx 6.2\times10^7$); XGBoost delivered near-perfect fit ($R^2 \approx 1.000$, $RMSE\_test \approx 1.4\times10^6$), raising concerns of potential overfitting despite very small test errors. The proposed soft-voting ensemble combined model strengths to produce robust forecasts ($R^2\_test \approx 0.9987$, $RMSE\_test \approx 2.53\times10^7$, $MAE\_test \approx 1.86\times10^7$), yielding stable predictions and reduced variance. Conclude that ensemble methods, particularly soft voting of diverse regressors, substantially enhance short-term import forecasting accuracy under limited historical records.

Keywords: Predicting, gas oil, import, machine learning, Ensemble Learning, soft voting

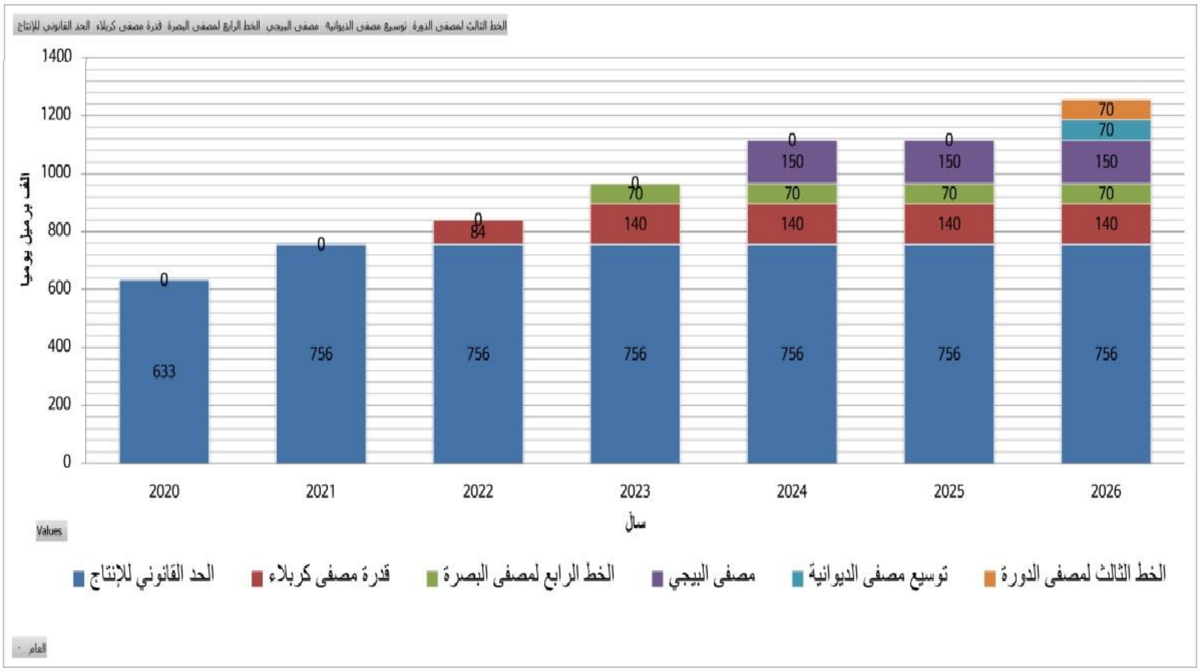
1. INTRODECTION

Iraq is a country rich in natural resources, and oil is one of those resources. It is a major source of energy. Sales of oil produced from Iraqi deposits constitute the primary and most important revenue stream for the Iraqi economy. Iraq has endured wars for more than two decades, preventing it from keeping pace with scientific and technological progress in the oil industry, reaching the level of developed countries, increasing its production, and improving the petroleum derivatives industry [1].

Received: August 12, 2025

International Journal of Applied Mathematics
Volume 38 No. 9s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Recently, Iraq has experienced events that have led to a decrease in its production of crude oil and petroleum derivatives. These events include wars and terrorist acts targeting the crude oil production sector, its transportation sector, and refining and oil industry facilities, causing a decrease in the availability of petroleum derivatives. Despite this, the government has worked to restore these facilities to contribute to meeting the largest portion of domestic consumption of petroleum derivatives by establishing new, advanced refineries to meet the necessary quantities needed for the growing domestic consumption, in addition to meeting a further portion through imports from other countries. Figure (1) illustrates the state's plan for the production line of petroleum derivatives for the period 2020-2026 [2].



المصدر:وزارة النفط العراقية، 3-12-2023 اوبيك وميس 2-12-2023

Fig. 1 shows the production line for petroleum derivatives.

Recent years have witnessed a significant increase in gasoil consumption in Iraq, due to improved living standards and the acquisition of a greater number of devices and equipment that rely on it as a fuel. This is in addition to the expansion of the activities of foreign oil companies within licensing rounds, which use it to operate their equipment. With domestic production declining in the face of this growing demand, the country has been forced to rely on gasoil imports to cover the shortfall in domestic consumption [3]. Therefore, a system has been proposed to predict future gasoil import quantities based on domestic production and consumption using machine learning techniques.

2. RELATED WORKS

Malyarenko-Olina et al. (2023) presented a systematic mathematical approach to forecasting the total volume of petroleum product consumption in Ukraine under the influence of economic variables and structural destruction resulting from the war. The study followed a "complex method" approach to energy forecasting in two stages: (first stage) forecasting the total volume of petroleum product consumption at the country level and economic activity groups, and (second stage) forecasting the structure of consumption by type using statistical analysis of historical data for the period 2015–2020. The study used annual statistical analysis, descriptive data tables, mathematical equation modeling, and methods for constructing the national energy balance. The results showed that gasoline consumption declined from approximately 2,360.8 to 1,767.7, while diesel fuel recorded higher levels (e.g., 4,770.9 → 5,173.9 during the period). The forecasts indicate a relative stability of total consumption until 2030, with an increasing share of diesel and liquefied petroleum gas (LPG) and a decreasing share of gasoline. The advantages of this research include its applied nature and mathematical precision, and its reliance on official data that inform national energy balance planning and the estimation of production and import needs. Its disadvantages, however, lie in the fact that the forecast relies on a relatively short historical period (2015–2020) and relative assumptions of post-shock economic stability. Furthermore, the results are sensitive to assumptions regarding structural coefficients and unit conversion, which may limit the accuracy of the forecast in contexts of extreme volatility or alternative scenarios for recovery and reconstruction [4].

Zhuang, Zeng, and Zheng (2025) addressed the complexity of oil and gas production forecasting due to heterogeneity of reservoir properties, diverse fluid properties, and multiple production technologies, which makes traditional methods limited in efficiency and generalizability. The study adopted a descriptive and comparative analytical approach, reviewing the development of forecasting techniques from empirical and regression models to machine learning and deep learning-based models such as neural networks (ANN, LSTM, CNN) and random forest and support vector algorithms (SVM, RF). The advantages and limitations of each approach were analyzed, highlighting their applications in conventional and unconventional oil fields. The results showed that deep learning models such as LSTM, CNN-LSTM, and GNN significantly improved the accuracy of production timing and well interaction prediction, outperforming traditional statistical methods in handling nonlinear and multi-source data. Some models achieved accuracy exceeding 97% in predicting well production over a long timescale. The advantages of this research include its comprehensive presentation of model development and its integration of physical and artificial intelligence perspectives. It also suggests future directions, such as integrating deep learning mechanisms with physical models and enhancing interpretability. Its drawbacks include its theoretical analytical nature, which lacks practical field applications. Furthermore, it does not provide extensive quantitative tests on real data to verify the performance of the proposed models [5].

Chongyan Li and Fuzhong Wang (2025) conducted a study aimed at developing an accurate and easy-to-apply model for forecasting future natural gas demand in the Chinese logistics sector. A quantitative analytical approach based on a semi-hierarchical control mechanism with feedback was used to reduce error and improve forecast accuracy. Three independent forecasting methods were applied: gray forecasting, regression trend method, and Bass model, and then combined using the minimum sum of squared errors (LSSE) principle to arrive at a more stable composite model. The results showed that the error rate of the individual models

Volume 38 No. 9s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

ranged from 2.9% to 5.94%, while the composite model decreased to 2.9–3.1%, demonstrating a significant improvement in accuracy and stability. One of the most notable findings is that natural gas consumption in the logistics sector will continue to grow rapidly, driven by government policies and the drive for a low-carbon economy. The research's advantages include a novel methodological contribution by integrating predictive models within a semi-hierarchical control framework that is easy to apply, providing a tool for use in energy planning and environmental policies. Its drawbacks include its reliance on limited historical data (2010–2021) and the failure to incorporate modern artificial intelligence or machine learning techniques, despite their future importance. This makes the prediction results dependent on existing economic growth assumptions and energy policies [6].

Septri Damayanti, Siska Yosmar, and Nur Afandi (2023) conducted a study that addresses the problem of irregular fluctuations in the value of oil and gas imports in Indonesia, which impacts economic planning decisions and trade policies, given the lack of accurate forecasting studies. The study applies the Fuzzy Time Series Chen method, one of the modern methods for time series analysis, and evaluates its ability to predict the value of Indonesian oil and gas imports. The researchers adopted a quantitative analytical approach using secondary data from the Indonesian Central Statistics Agency (BPS) covering the period from January 2015 to July 2022 (91 months). The research tools included fuzzification and defuzzification stages, constructing fuzzy logical relationships (FLRs) and FLRGs, and then calculating the forecast accuracy using the Mean Absolute Relative Error (MAPE) index. The results showed that the MAPE value reached 19.969%, indicating good forecast accuracy. The forecast for the value of imports for August 2022 amounted to approximately USD 3,743,213 million. The research's advantages include providing a precise practical application of Chen's method in a realistic economic context, demonstrating the simplicity of the approach and its effectiveness in dealing with nonlinear and volatile data. It also contributed to expanding the use of fuzzy models in macroeconomics. Its disadvantages include its reliance on a single model without direct comparison with modern methods such as neural networks or hybrid models. Furthermore, the sample is limited to a relatively short period, which may limit the generalizability of the results across different periods or under changing economic conditions [7].

3. METHODOLOGY

3.1 DATASET

The dataset used in this research consists of annual measurements covering the period from 1985 to 2023, a total of thirty-four years, and includes three main variables: production, consumption, and import. These variables represent the primary indicators in analyzing the balance between supply and demand for gasoil in Iraq. All values in these variables are numerical and represent actual annual quantities extracted from official records. The data were collected from the official website of the State Oil Marketing Organization (SOMO) and from the receiving and delivery points of produced and imported diesel products affiliated with the Oil Products Distribution Company.

International Journal of Applied Mathematics
Volume 38 No. 9s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The data indicate a clear variation in production and consumption levels over the years. The first period was characterized by low or almost non-existent import values, while subsequent years witnessed a gradual increase in the volume of imports due to increased domestic demand and changing consumption patterns. This type of data is considered a time series due to its regular annual sequence, which makes it suitable for applying quantitative forecasting models using machine learning algorithms or time-series statistical models to estimate future import quantities based on the historical behavior of both production and consumption.

3.2 MACHINE LEARNING

Machine learning models are a modern method based on data analysis and exploring statistical patterns. They possess a high capacity to learn from historical data and discover hidden relationships between variables, making them an effective tool in forecasting processes. These models are characterized by their ability to process and analyze high-dimensional data to generate accurate and informed estimates. In this context, predictive machine learning algorithms are used to classify and analyze various data and identify future trends. They can be employed to forecast gas oil import quantities in Iraq based on historical production and consumption data. This study relies on the application of a set of traditional machine learning algorithms to build an efficient predictive model capable of accurately estimating future import quantities, supporting decision-makers in formulating economic and energy policies [8].

1- Random Forest (RF) Algorithm

This algorithm is one of the most efficient ensemble learning techniques for forecasting. It creates a large number of decision trees, each trained on random subsamples of data and attributes. The results are then combined to obtain the most accurate and stable final forecast [9]. In this research, the Random Forest algorithm is used to predict gas oil import quantities in Iraq based on historical production and consumption data, by analyzing patterns and nonlinear relationships between these variables. The algorithm's multi-decision aggregation nature reduces variance and improves model accuracy compared to using a single decision tree. It also has the advantage of being able to handle large and complex datasets without requiring prior assumptions about the distribution of variables. Furthermore, the Random Forest algorithm allows us to determine the importance of variables influencing import behavior, such as the impact of local production or annual consumption levels. Although the model is more stable and less prone to overfitting than individual models, it may exhibit relatively slow prediction speed when dealing with very large amounts of data, due to the large number of trees generated during the training process [10].

2- Linear Regression Algorithm

Linear regression is one of the simplest and most popular statistical methods used to analyze relationships between variables and predict future values. This model establishes a linear relationship between independent variables—such as production and consumption—and the dependent variable (imports), with the aim of determining the extent to which both production and consumption affect the volume of gas oil imports in Iraq. In this research, a linear regression model is used to estimate the direct quantitative relationship between supply and

Volume 38 No. 9s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

demand variables and formulate a predictive equation that enables calculating future import quantities based on historical trends in production and consumption. The model seeks to identify optimal regression coefficients (weights) that minimize the error between actual and predicted values, allowing for the construction of a simple yet effective predictive model that can support decision-makers in planning import and energy policies [11].

3- Support Vector Regression (SVR)

This algorithm is an extension of the Support Vector Machine (SVM) model, but is used for predicting continuous values rather than classification. It relies on the principle of finding a function that best fits the data while maintaining a small margin of error. In other words, it aims not only to minimize absolute error, but also to find a model that can generalize well to new data [12]. The algorithm transforms the data into a higher-dimensional space using kernel functions such as linear, polynomial, or RBF functions, and then searches for the optimal regression line or surface that separates the data within a certain margin without being significantly affected by outliers. SVR is characterized by its high ability to handle nonlinear and noisy data, in addition to its ability to generalize strongly even with small data sets. Its most prominent advantages include high accuracy and mathematical stability, while its disadvantages include high computational cost when dealing with large data sets and the difficulty of choosing optimal kernel parameters (such as C and γ). This algorithm is used in many applications, such as price forecasting [13]. In this study, it is used to predict the quantities of gas oil imports into Iraq by determining the extent of the impact of both production and consumption.

4- XGBoost (Extreme Gradient Boosting)

It is a standard machine learning model that follows the principles of ensemble learning by constructing successive decision trees that correct the errors of previous models, resulting in high accuracy and good execution speed for complex tasks. Methodologically, XGBoost relies on a gradient boosting framework that calculates the gradient of the loss function. Multiple experiments have shown that XGBoost achieves predictive superiority compared to similar algorithms. Its most notable advantages include its combination of high accuracy, relative ease of implementation, and the ability to analyze feature importance. Its disadvantages include its requirement for extensive hyperparameter tuning and the potential for overfitting if the regularization parameters are not well-tuned or if data with many outliers is used [14].

3.3 Ensemble Learning

Ensemble learning is an advanced approach in machine learning that combines the results of multiple predictive models rather than relying on a single model, with the goal of improving performance and reducing generalization error. The strategy is based on two main factors: first, diversifying the base learners so that each one provides a different perspective on the data, and second, combining their outputs through mechanisms such as voting, averaging, or stacking to achieve a "robust" model that is more accurate and stable than any single model[15] . For forecasting applications such as forecasting import quantities based on production and consumption ensemble learning is characterized by its ability to capture complex, nonlinear

Volume 38 No. 9s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

relationships between variables and increase stability across years compared to prediction by a single model.

In this study, a soft voting regression approach based on the Voting Regressor algorithm is employed to enhance the accuracy of gas oil import forecasting. The proposed method effectively reduces prediction variance and improves the model's generalization capability while minimizing the risk of overfitting a crucial advantage when working with limited time-series data. Through the soft voting mechanism, the outputs of multiple regression models are combined by averaging their predicted values, allowing the ensemble to leverage the strengths of each individual model. The overall prediction is then obtained from this averaged output [16]. as illustrated in Figure (2), which depicts the structural framework of the soft voting technique.

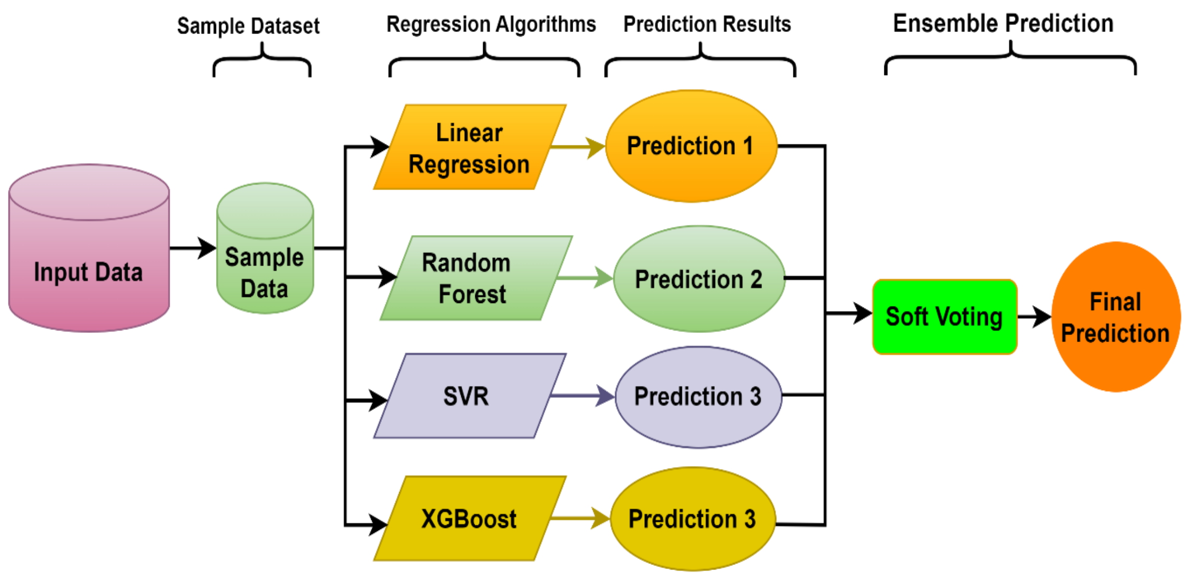


Fig. 2 structure of the soft voting technique.

3.4 Evaluation Metrics

1-Coefficient of Determination (R^2)

R^2 is one of the most important statistical measures used to evaluate the performance of regression models. It expresses the proportion of variance in the dependent variable that can be explained by the independent variables in the model. This measure ranges from zero to one, with $R^2 = 1$ indicating that the model fully explains the variance in the data, while $R^2 = 0$ indicates that the model cannot explain any of the variance. The value may also be negative if the model severely under-predicts, meaning that the model provides worse results than using the mean alone. This measure is widely used in economics, engineering, and time series analysis because it provides a simple interpretive indicator of how well the model fits the data (goodness of fit). However, relying on R^2 alone is not sufficient to evaluate a model; it must be compared with other measures such as RMSE and MAE to determine the accuracy of the predictions [17].

Received: August 12, 2025

Volume 38 No. 9s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

2-Root Mean Squared Error (RMSE)

This measure is one of the most widely used measures for evaluating the performance of forecasting and regression models. It expresses the average error between the actual values and the predicted values after squaring the differences and then taking the square root of the mean. This method gives greater weight to large errors than other measures, making it sensitive to large deviations in predictions. The lower the RMSE value, the more accurate the model is in predicting. This measure is widely used in statistical modeling, economics, and time series analysis applications because of its ability to measure the quality of predictions in the same units of measurement as the variable under study. However, RMSE does not indicate the direction of the error (overestimate or underestimate), so it is best used in conjunction with other measures such as MAE and R^2 to obtain a comprehensive assessment of model performance[18] .

3-Mean Absolute Error (MAE)

This measure is one of the basic statistical measures used to evaluate the performance of forecasting and regression models. It expresses the average absolute value of the difference between the actual and predicted values, regardless of the direction of the error (whether the model underestimates or overestimates). This measure is simple and easy to interpret, as it is presented in the same units as the variable under study, making it easy to understand and compare between models. MAE is a more stable measure than RMSE in cases with outliers, as it does not significantly amplify the impact of large errors. The lower the MAE value, the higher the model's predictive accuracy and performance. This measure is widely used in time series analysis, economics, and energy applications to measure the reliability of predictive models [19].

4. EXPERIMENTAL RESULTS

1-Random Forest

The current results obtained indicate that the Random Forest model achieved good predictive performance and represents a practical and reliable option for estimating import quantities based on production and consumption. The data was split 80% for training and 20% for testing. Model settings (n_estimators=200, max_depth=6, min_samples_split=4, min_samples_leaf=2, random_state=42) were used, yielding training metrics of R^2=0.9887, RMSE=82,077,663.60, and MAE=51,801,233.51, and testing metrics of R^2=0.9922, RMSE=61,971,460.24, and MAE=45,280,555.80.

The very high values of the coefficient of determination (R^2) for both the training and test sets indicate that the model explains a significant portion of the variance in the import quantity for both the training and test sets. The similarity between the R^2 values for the training and test sets suggests no apparent overfitting, meaning that the model generalizes well to data not seen during training. Figure (3) shows the Actual vs Predicted (Test Set) curve.

International Journal of Applied Mathematics
Volume 38 No. 9s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

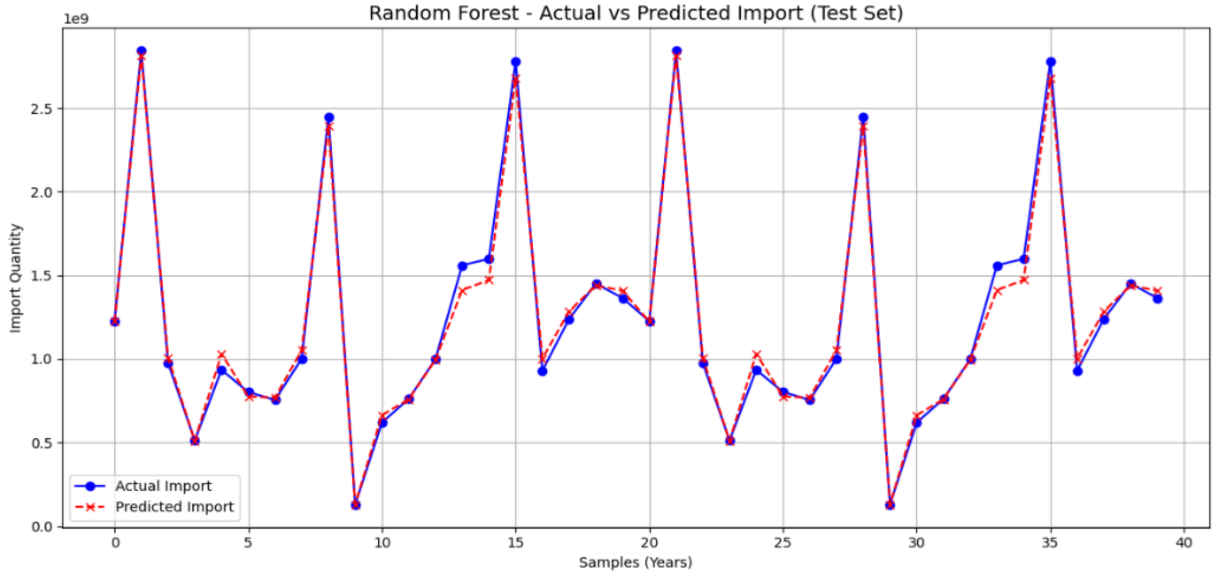


Fig. 3 Actual vs Predicted (Test Set) for Random Forest

RMSE and MAE measures These provide a practical measure of the magnitude of the error in units of the target variable (import quantity, measured in millions of liters per year). Given the general range of values in the graph (in billions), the RMSE and MAE represent relatively small errors relative to the annual import quantity, reinforcing the conclusion that the model is practically accurate for short-term forecasts.

Figure (4) shows that the distribution of errors is not homogeneous, as some samples show significantly high values, indicating the presence of anomalies or patterns that are not well represented in the current set of variables. This may be due to factors such as sudden changes in increased consumption or decreased production due to operational conditions or abnormal production conditions. This distribution, which is characterized by the spread of small and medium columns with the emergence of limited peaks, reflects good performance of the model in most samples.

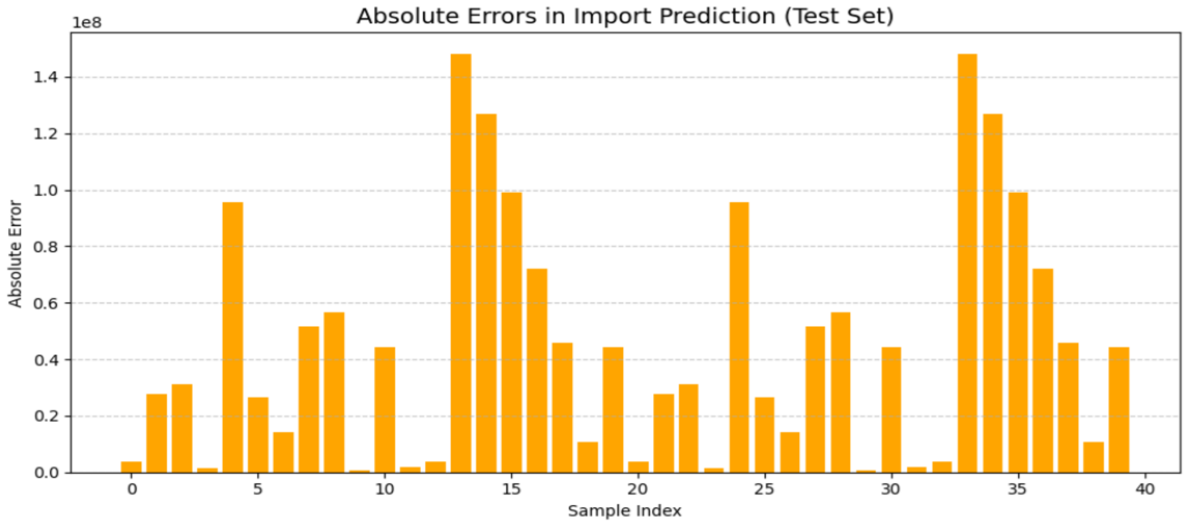


Fig. 4 Absolute Errors in import prediction

Volume 38 No. 9s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

2-Linear Regression Algorithm

The results of the linear regression model showed limited explanatory power in predicting gasoil import volumes. The coefficient of determination on the training data was ($R^2 = 0.2033$) and declined to ($R^2 = 0.0313$) on the test data. Both RMSE (approximately 6.9×10^8) and MAE (approximately 5×10^8) were high in both sets. The data was split 80% for training and 20% for testing. These results indicate that the model failed to capture the actual relationships between production, consumption, and import volume, reflecting the nonlinear nature of the relationship between the variables. The closeness of the error levels in the training and test data indicates that the problem is not caused by overfitting, but rather by underfitting or the absence of other influencing variables. Therefore, it can be said that the linear regression model is insufficient for practical use in forecasting gas oil import quantities due to its low explanatory power and high level of error in the predictions.

Figure (5) illustrates Actual vs. Predicted (Test Set) and shows that the variance and sharp peaks in the actual series are not captured by the model. Forecasts are more monotonous, smoother, and tend to represent a general trend or moving average rather than sharp spikes. This behavior suggests that the model captures general trends or the mean but fails to predict peaks and troughs. When the forecast is smooth compared to the volatile real data, we expect a large MAE/RMSE, as observed.

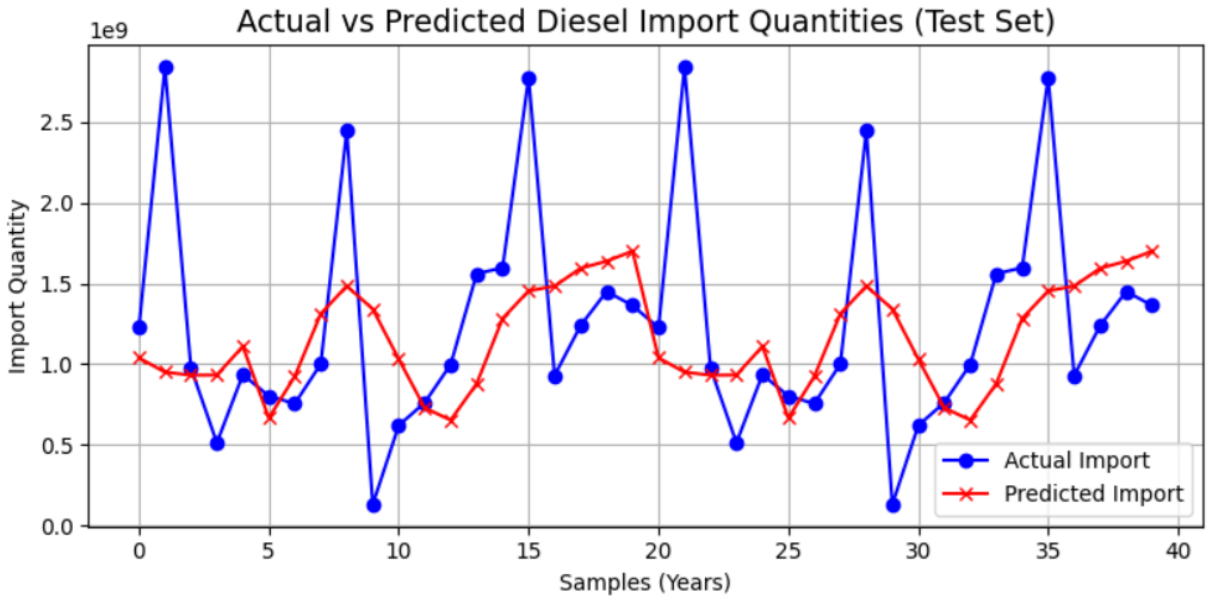


Fig. 5 illustrates Actual vs. Predicted (Test Set)

Figure (6) shows the absolute errors for each test set. There are very high peaks at certain sets. These peaks correspond to peaks where the model failed to adjust. The distribution of errors is not uniform; some years have small, acceptable errors, while others have enormous errors—indicating that a single, simple model does not provide consistent coverage of all cases over time. The presence of these peaks increases the RMSE further (since the RMSE gives greater weight to large errors), while the MAE is relatively less affected, but remains high because many errors are moderate to large.

Received: August 12, 2025

International Journal of Applied Mathematics
Volume 38 No. 9s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

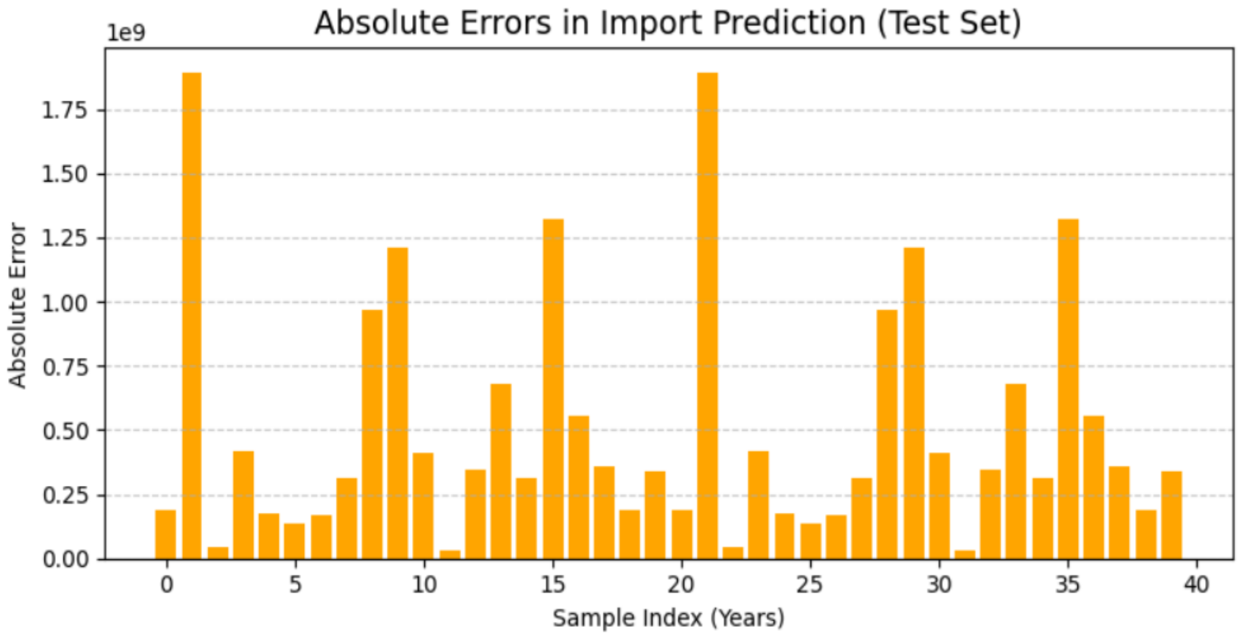


Fig. 6 Absolute Errors (Test Set)

3- Polynomial Regression Algorithm

The sample was divided into 80% training and 20% testing. The results of the 6-degree polynomial regression model demonstrated average performance in terms of explanatory power, with an R^2 value of 0.6742 for the training data and 0.5589 for the testing data, indicating that the model explains approximately 67% of the variance in the training data and 56% of the variance in the testing data. This small difference between the two groups reflects the model's ability to generalize to an acceptable degree, with a slight tendency to lose accuracy when predicting new data. The error metrics were RMSE $\approx 4.41\times10^8$ and MAE $\approx 3.01\times10^8$ in the training model, and RMSE $\approx 4.66\times10^8$ and MAE $\approx 3.25\times10^8$ in the testing model. These are relatively high values, given that import volumes are measured in hundreds of millions to billions. The higher RMSE value compared to MAE is expected, given the sensitivity of the former to large errors, indicating the presence of some outliers or exceptional cases that affected the model's accuracy. Overall, the results show that the model was able to represent the general trends and major fluctuations in the gas oil import data. However, it underestimated the intensity of peaks and failed to accurately represent some atypical cases. This results in an average performance in terms of practical accuracy, despite its reasonable ability to describe the overall behavior of the time series.

Figure (7) shows the actual import values versus the model's forecast values across time samples. We note that the model often underestimates high peaks and sometimes overestimates lows, or vice versa. When sharp actual peaks appear, the forecasts are less sharp and tend toward the average values, indicating that the model is able to capture the general trend but is inflexible enough to accurately represent unusual fluctuations.

Received: August 12, 2025

International Journal of Applied Mathematics

Volume 38 No. 9s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

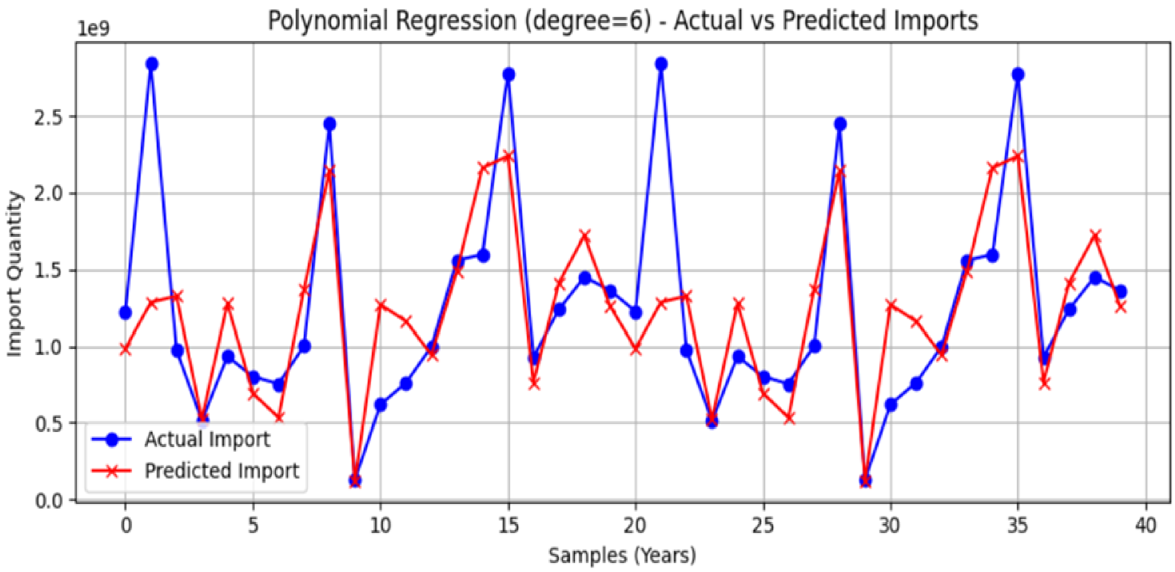


Fig .7 Prediction vs. Actual Values

Figure (8) shows the distribution of absolute errors for each time sample (each year). We notice that there is an inhomogeneous distribution of errors - most of the columns are small to medium, with some columns of very high height (peaks) emerging. This indicates that the model achieves acceptable performance in most years, but it clearly fails in a limited number of years that represent exceptional cases or abnormal points.

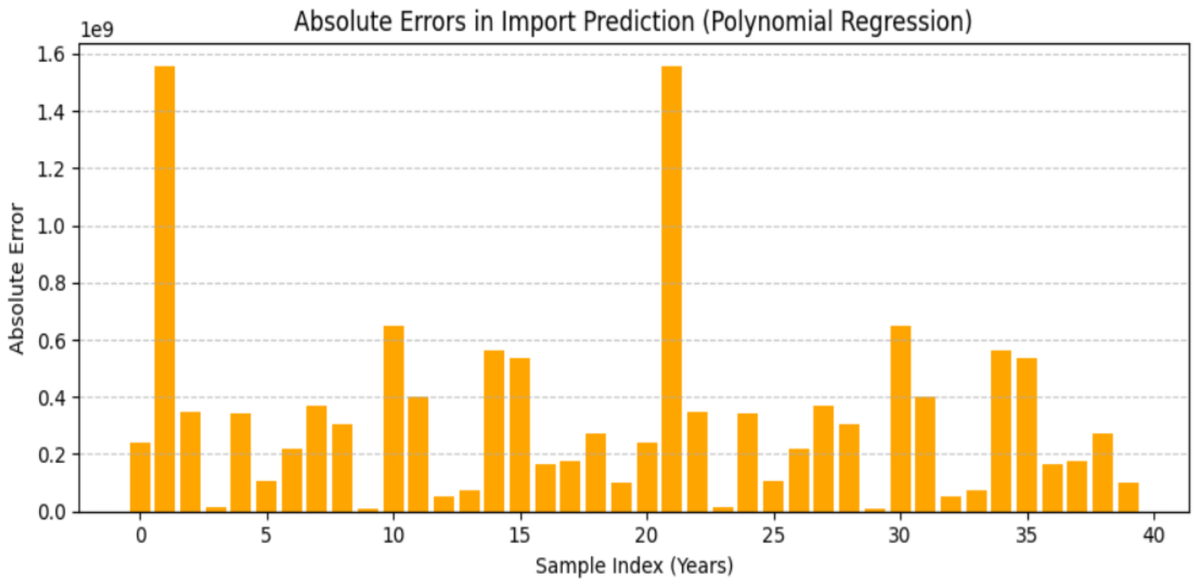


Fig. 8 Absolute Errors in Import Prediction

4-Support Vector Regression (SVR)

The results of the model, implemented using an RBF kernel and with settings C = 100, ε = 0.1, and γ = 'scale', showed moderate to good explanation ability for the relationship between the independent variables (production and consumption) and the dependent variable (imports). The R^2 value was 0.7255 for the training set and 0.5819 for the test set, meaning that the model

Volume 38 No. 9s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

explains approximately 72.6% of the variance in the training data and 58.2% of the variance in the test data. These values reflect the model's ability to capture general patterns in the data, but it does not cover all of the variance, especially when dealing with new data that it has not previously seen during training. As for the error metrics, they were high in absolute terms, with RMSE ≈ 4.05×10^8 and MAE ≈ 1.71×10^8 for the training data, compared to RMSE ≈ 4.55×10^8 and MAE ≈ 2.07×10^8 for the test data. The higher RMSE value compared to MAE indicates the presence of some large errors or outliers that have a greater impact on the mean square error, while the difference between the training and test metrics reflects a moderate loss of accuracy when generalizing—that is, a slight tendency toward higher performance on the training data compared to the test data.

Figure (9) shows the comparison between the predicted and actual values. We note that the predictions follow the general path of the time series to a reasonable degree. Many of the general peaks and humps in the predictions are close to the actual values, confirming that the model is able to capture the main trends and moderate fluctuations. However, significant differences are also observed at some extreme peaks. In these cases, the predicted values are less extreme than the actual values, or there is a slight time lag in the peak timing. This behavior indicates that the model does not fully capture exceptional events or sharp changes that may result from external factors not included in the variables.

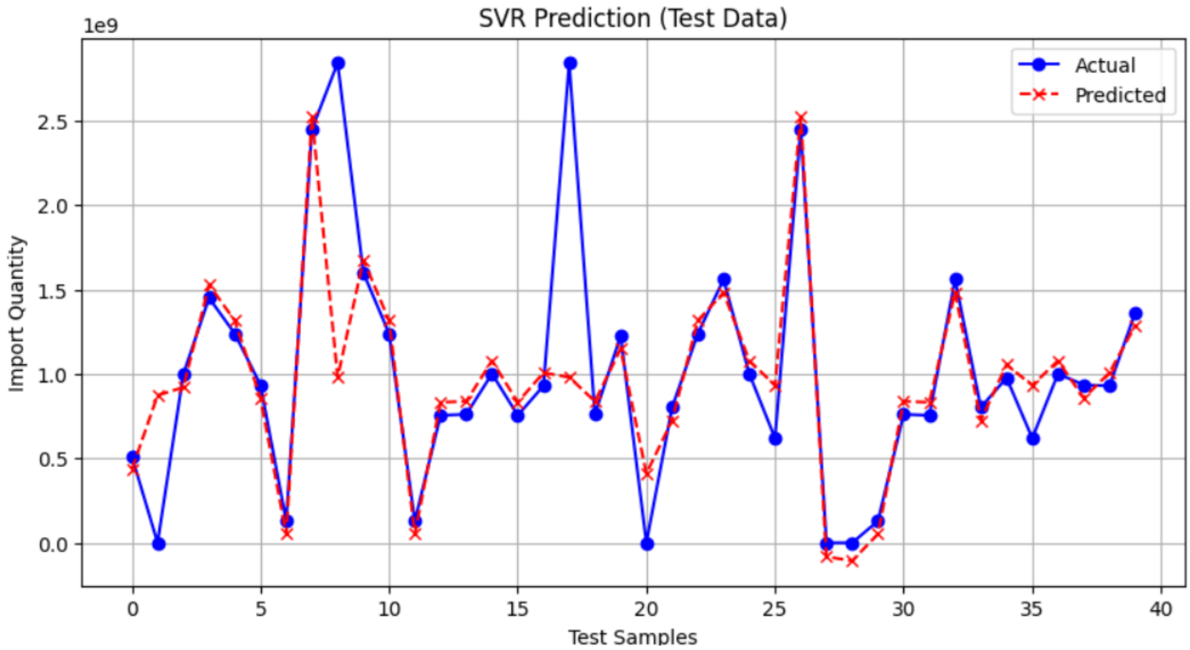


Fig. 9 Prediction vs. Actual Values

Figure (10) illustrates the absolute difference between the actual and predicted values for each sample from the test set. Most samples produce moderate to small errors, while some high peaks represent years with very large errors. The concentration of these peaks indicates the presence of outliers or periods that the model could not explain well—the same periods where a clear mismatch appeared in Figure (8). The high percentage of required errors contributes to

International Journal of Applied Mathematics

Volume 38 No. 9s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

a higher RMSE value, which confirms the need to address outliers or improve the model's sensitivity to these cases.

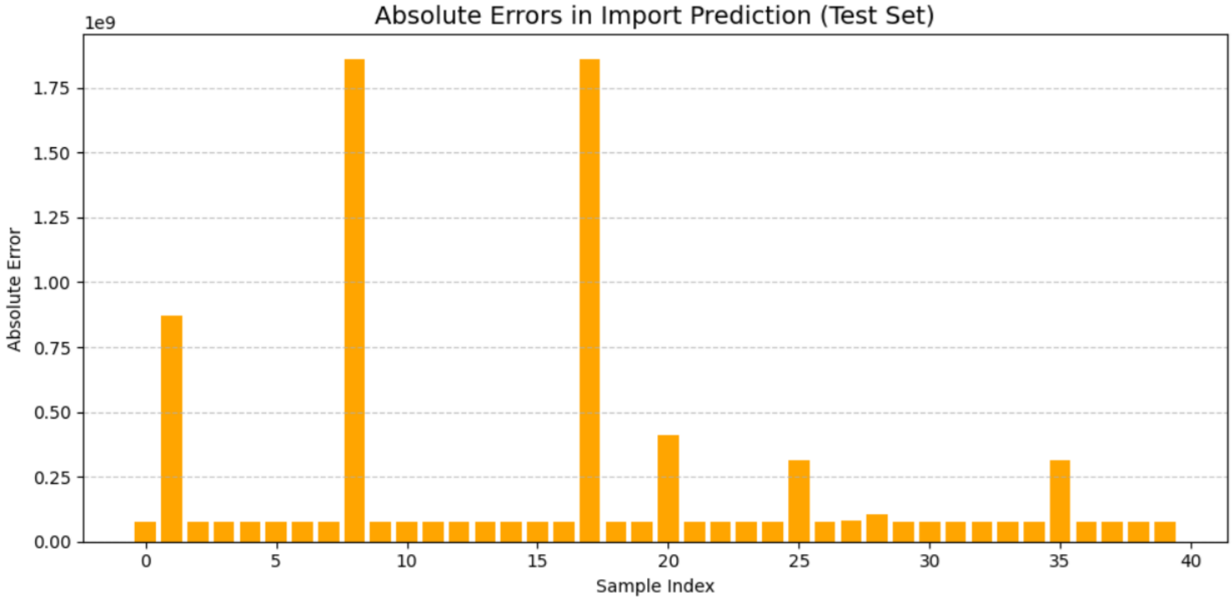


Fig. 10 Absolute Errors in Import Prediction

5-XGBoost (Extreme Gradient Boosting)

The evaluation results indicate that the XGBoost model performed very accurately in predicting gasoil import volumes on both the training and test data. The coefficient of determination (R^2) value was approximately 1.0000 in both parts, with very low values for both RMSE and MAE (RMSE ≈ 2.7×10^6 in training and 1.4×10^6 in testing, and MAE ≈ 1.6×10^6 and 1.07×10^6, respectively). These values indicate that the model was able to explain more than 99.9% of the variance in the data, with the differences between the actual and predicted values being almost negligible. This outstanding performance reflects the XGBoost algorithm's ability to handle nonlinear relationships and complex interactions between variables (production and consumption) when predicting the dependent variable (imports). This superiority is attributed to the nature of XGBoost as an ensemble algorithm that relies on building a large number of weak decision trees and gradually updating them to improve accuracy at each step via a gradient boosting mechanism.

Figure (11) shows a comparison between the actual and predicted values. We observe a near-perfect match between the two values across all samples, indicating that the model was able to track the behavior of the time series with very high accuracy, including the sharp peaks and declines in import quantities. This perfect consistency reflects the model's efficiency in capturing subtle temporal patterns and the complex relationships between production, consumption, and imports, and confirms that the model successfully predicts almost all short- and long-term changes without significant lag or deviation.

International Journal of Applied Mathematics
Volume 38 No. 9s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

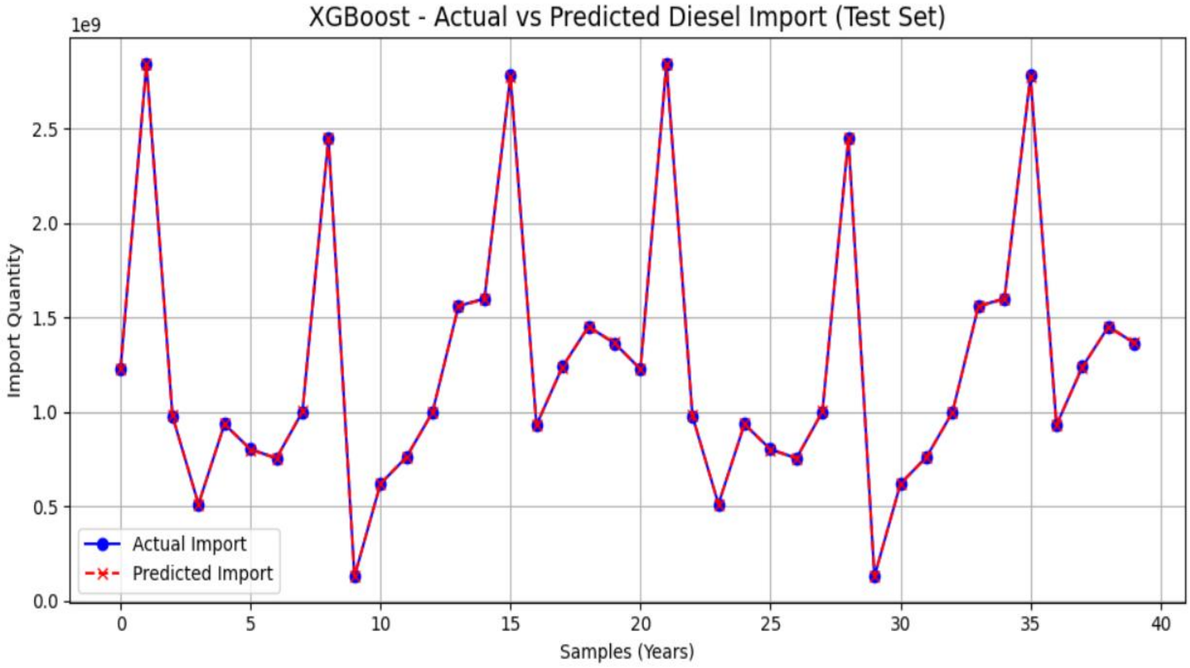


Fig .11 Actual vs Predicted Diesel Import Quantities

Figure (12) shows the absolute error values for each sample in the test set. We note that almost all of the columns are very small in height, not exceeding a few units in the 10^6 range. This is a very small error size compared to the actual import values, which average more than 10^9. This result indicates that the model did not exhibit significant deviation in any of the test years, and that the prediction errors are consistent and low over time.

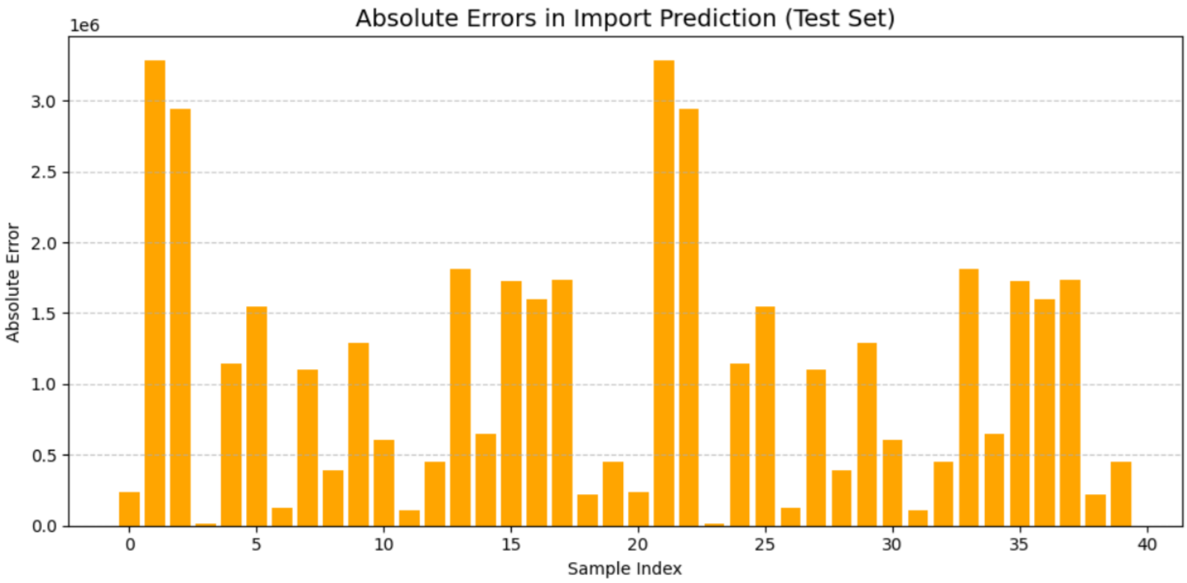


Fig. 12 Absolute Errors in Import Prediction

6-Soft Voting Ensemble Model

The results of the Soft Voting Ensemble learning model demonstrated outstanding performance in predicting gas oil import volumes. This is attributed to the model's ability to combine the

Volume 38 No. 9s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

strengths of several algorithms XGBoost, Polynomial Regression, Random Forest, and Supporting Edge Regression (SVR) allowing it to simultaneously incorporate linear and nonlinear patterns. On the training data, the model achieved an $R^2 = 0.9981$, RMSE = 33.7×10^6, and MAE = 21.4×10^6, while the test results achieved an $R^2 = 0.9987$, RMSE = 25.3×10^6, and MAE = 18.6×10^6. These values indicate the model's very high ability to explain variance in the data by averaging the predictions generated by each algorithm, which reduces variance in results and improves the stability and accuracy of the predictions. The extremely low error rate and high R^2 values indicate that the model was able to represent the relationships between production, consumption, and imports with near-perfect accuracy. The small difference between training and testing performance reflects that the model has excellent generalization ability and does not suffer from overfitting, enhancing the reliability of its results when applied to future data.

Figure (13) shows the relationship between actual and predicted values. The predicted values curve matches the actual values curve, confirming the model's high accuracy in tracking general trends and seasonal fluctuations in import data. This visual consistency indicates that the model was able to accurately represent temporal patterns and accommodate minor changes in the data without significant deviations.

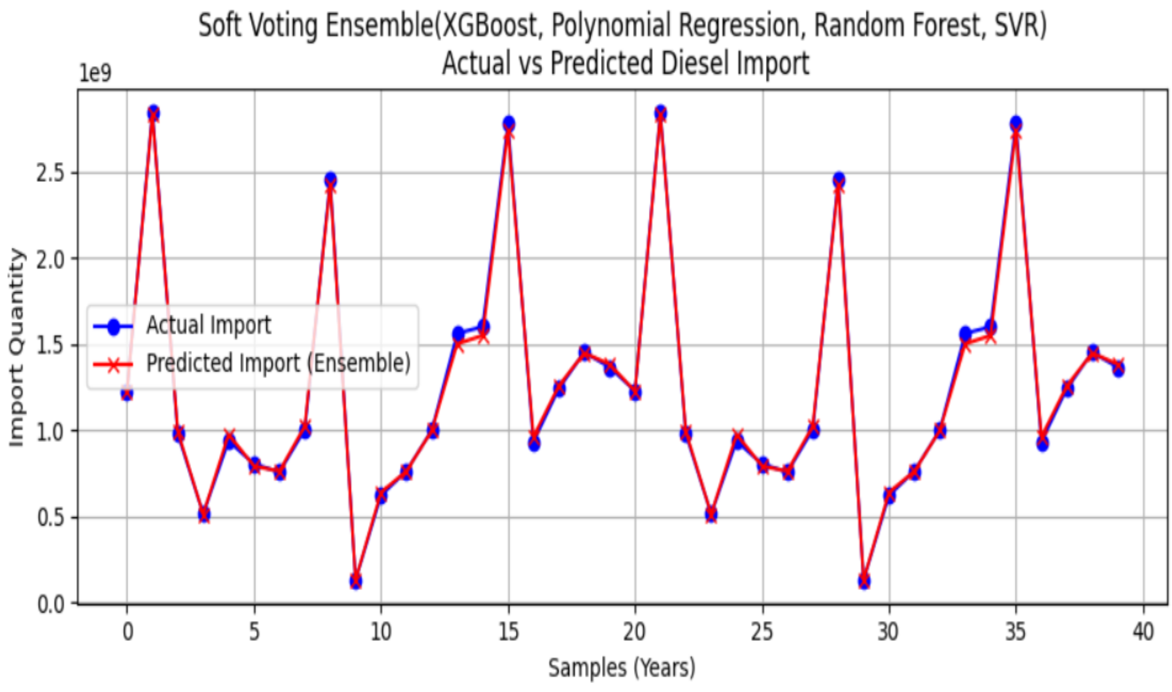


Fig .13 Actual vs Predicted Diesel Import Quantities

Figure (14) shows the absolute errors for forecasting gas oil imports. The error levels in most samples are very low, with slight increases in some years that may have experienced unexpected changes in import behavior. However, the values of these errors remain limited compared to the overall data range, confirming that the forecasts are stable and realistic.

Received: August 12, 2025

International Journal of Applied Mathematics
Volume 38 No. 9s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

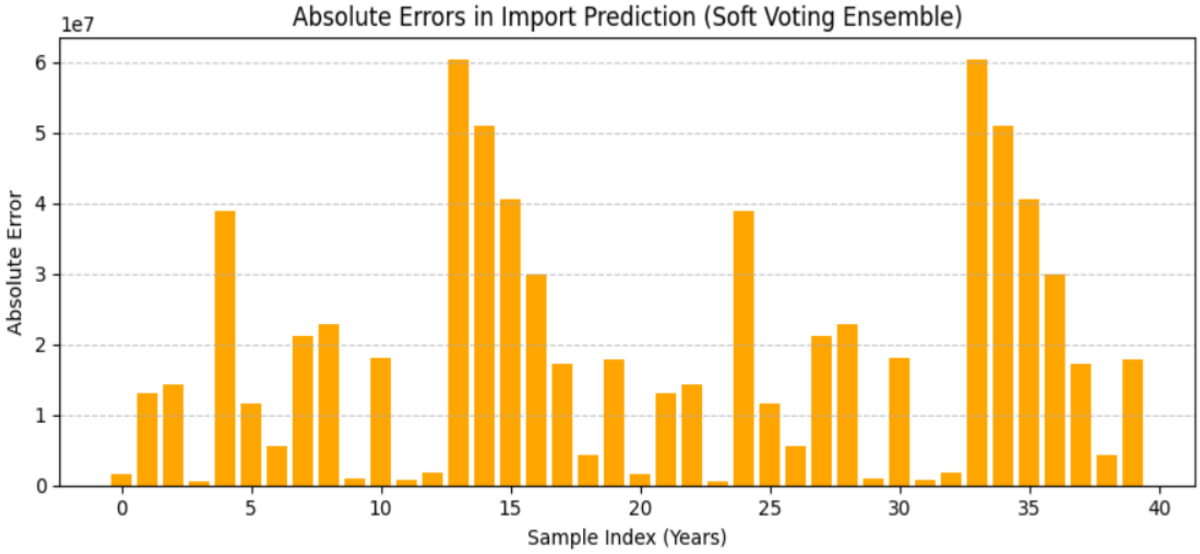


Fig. 14 Absolute Errors in Import Prediction

5. Discussion of results

In Table (1), the results of the individual models indicate a clear discrepancy in the explanatory power and prediction accuracy of the different algorithms. The results of the linear regression model showed very poor performance, indicating its failure to represent the true relationship between production, consumption, and import quantities due to its linear simplicity and inability to handle complex nonlinear relationships in the data. The polynomial regression model significantly improved performance thanks to its nonlinear nature, which enabled it to capture general trends. However, the relatively high RMSE and MAE values indicate the persistence of large errors in some samples, especially during periods with high peaks. The SVR model performed well, with a slight decrease in accuracy on the test, reflecting its average ability to generalize and capture nonlinear patterns. In contrast, the Random Forest model distinguished itself with very strong performance and low RMSE and MAE values, reflecting its high efficiency in handling nonlinear relationships and interactions between variables, while maintaining excellent generalization ability, making it the best individual model among the tested models. The XGBoost model delivered optimal results ($R^2 = 100$), indicating overfitting caused by excessive data retention at the expense of generalization, despite a significant reduction in error.

Comparing the performance of the Soft Voting model with the best individual model, the Random Forest, we note that both models achieved very high results. However, the Soft Voting model slightly outperformed in terms of the balance between training and testing, with further reductions in RMSE and MAE values, reflecting improved prediction accuracy and performance stability. This superiority is attributed to the nature of the ensemble learning technique, which relies on combining multiple different models into a single system, reducing the bias of each individual model and improving the final model's ability to generalize. Therefore, it can be said that the Soft Voting model represents the best overall model for

Volume 38 No. 9s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

predicting gas oil import quantities compared to the performance and efficiency of individual models.

Table 1 Results of evaluation metrics for all models

Models	Model Evaluation Metrics					
	R^2 Train	R^2 Test	RMSE Train	RMSE Test	MAE Train	MAE Test
Random Forest	98.87	99.22	$8*10^7$	$6*10^7$	$5*10^7$	$4*10^7$
Linear Regression	20.33	3.13	$6.9*10^8$	$6.9*10^8$	$5*10^8$	$5*10^8$
Polynomial Regression	67.42	55.89	$4.41*10^8$	$4.66*10^8$	$3.01*10^8$	$3.25*10^8$
XGBoost	100	100	$2.7*10^6$	$1.4*10^6$	$1.6*10^6$	$1.07*10^6$
SVR	72.55	58.19	$4.05*10^8$	$4.55*10^8$	$1.71*10^8$	$2.07*10^8$
Soft Voting	99.81	99.87	$33.7*10^6$	$25.3*10^6$	$21.4*10^6$	$18.6*10^6$

6. CONCLUSIONS

Based on the experimental results of the implemented machine learning models, it can be concluded that advanced regression algorithms such as Random Forest and XGBoost demonstrated superior performance in forecasting gas oil import quantities using production and consumption data. The findings indicate that the relationship between variables is inherently nonlinear, which makes traditional models such as Linear Regression less effective in capturing the true behavior of the data. In contrast, nonlinear models such as SVR and Polynomial Regression provided better representations of complex patterns but showed sensitivity to outliers. The proposed ensemble learning model based on the Voting Regressor algorithm proved highly effective in enhancing prediction accuracy and reducing variance by combining the strengths of the four base models (Linear Regression, Random Forest, SVR, and XGBoost). This approach achieved an optimal balance between accuracy, stability, and generalization capability, making it a suitable method for short-term forecasting, particularly under limited temporal data conditions.

In conclusion, the integration of machine learning algorithms within an ensemble framework represents an effective and reliable approach for forecasting gas oil import quantities, with the potential for extending the proposed methodology to other petroleum products or related industrial sectors.

DATA AVAILABILITY STATEMENT

All data, models and code generated or used during the study appear in the submitted article.

Received: August 12, 2025

DECLARATION OF CONFLICTING INTERESTS

The author(s) declare(s) that there is no conflict of interest

REFERENCES

[1] A. A. Wali, "Analysis of Iraqi Oil Production Since 1970 with Formulation of perceptions about Future Scenarios of Oil Revenue Management." [Online]. Available: https://www.researchgate.net/publication/338950735

[2] D. Bamber, F. Saleh, and R. Panditharathna, "THE IRAQI OIL INDUSTRY: HISTORY AND OVERVIEW," *European Journal of Political Science Studies*, vol. 6, no. 1, Feb. 2023, doi: 10.46827/ejpss.v6i1.1424.

[3] A. Y. Fahad and H. Y. Shaalan, "Investment Trends in the Iraqi Oil Sector (Upstream – Oil Refining – Gas Treatment - Transportation - Distribution) (2003 – 2020)," *Journal of Petroleum Research and Studies*, vol. 12, no. 2, pp. 123–147, Jun. 2022, doi: 10.52716/jprs.v12i2.662.

[4] O. Y. Maliarenko and N. Y. Maistrenko, "A Methodical Approach to Forecasting the Consumption of Petroleum Products by Their Main Types," *Energotehnologii i Resursosberezenie*, vol. 74, no. 1, pp. 14–24, Mar. 2023, doi: 10.33070/etars.1.2023.02.

[5] Y. Zhuang, H. Zeng, and X. Cheng, "The Development, Application and Prospect of Oil and Gas Production Forecasting Technologys," *International Journal of Natural Resources and Environmental Studies*, vol. 6, no. 3, pp. 12–22, Aug. 2025, doi: 10.62051/ijnres.v6n3.02.

[6] C. Li and F. Wang, "Forecasting of natural gas consumption in China's logistics industry based on semi-hierarchical control," *PLoS One*, vol. 20, no. 6 June, Jun. 2025, doi: 10.1371/journal.pone.0325788.

[7] S. Damayanti, S. Yosmar, and N. Afandi, "IMPLEMENTATION OF FUZZY TIME SERIES CHEN FOR FORECASTING INDONESIAN OIL AND GAS IMPORTS VALUE," *Barekeng*, vol. 17, no. 2, pp. 685–694, Jun. 2023, doi: 10.30598/barekengvol17iss2pp0685-0694.

[8] S. Huang *et al.*, "A systematic review of mathematical and machine learning models of Avian Influenza," Dec. 01, 2025, *Elsevier B.V.* doi: 10.1016/j.onehlt.2025.101203.

[9] H. A. Salman, A. Kalakech, and A. Steiti, "Random Forest Algorithm Overview," *Babylonian Journal of Machine Learning*, vol. 2024, pp. 69–79, Jun. 2024, doi: 10.58496/bjml/2024/007.

[10] G. H. Hussien and Z. N. Nemer, "A Comprehensive Review of Advances in Tongue Image Classification Techniques for Diabetes Identification," *Iraqi Journal of Intelligent Computing and Informatics (IJICI)*, vol. 4, no. 1, Mar. 2025, doi: 10.52940/ijici.v4i1.90.

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[11] K. Qu, "Research on linear regression algorithm," *MATEC Web of Conferences*, vol. 395, p. 01046, 2024, doi: 10.1051/matecconf/202439501046.

[12] B. Bargam, A. Boudhar, C. Kinnard, H. Bouamri, K. Nifa, and A. Chehbouni, "Evaluation of the support vector regression (SVR) and the random forest (RF) models accuracy for streamflow prediction under a data-scarce basin in Morocco," *Discover Applied Sciences*, vol. 6, no. 6, Jun. 2024, doi: 10.1007/s42452-024-05994-z.

[13] M. Safari, A. H. Rabiee, and J. Joudaki, "Developing a Support Vector Regression (SVR) Model for Prediction of Main and Lateral Bending Angles in Laser Tube Bending Process," *Materials*, vol. 16, no. 8, Apr. 2023, doi: 10.3390/ma16083251.

[14] D. Boldini, F. Grisoni, D. Kuhn, L. Friedrich, and S. A. Sieber, "Practical guidelines for the use of gradient boosting for molecular property prediction," *J Cheminform*, vol. 15, no. 1, Dec. 2023, doi: 10.1186/s13321-023-00743-7.

[15] S. S. Rezk and K. S. Selim, "Metaheuristic-based ensemble learning: an extensive review of methods and applications," Oct. 01, 2024, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s00521-024-10203-4.

[16] G. H. Hussein and Z. N. Nemer, "CLASSIFICATION OF DIABETIC TONGUE IMAGES USING DEEP LEARNING," *NED University Journal of Research*, vol. XXII, no. 3, pp. 155–176, Sep. 2025, doi: 10.35453/nedjr-ascn-2025-0025.r1.

[17] A. Torres-Hernandez and A. Torres-Hernandez, "Proposal for the Application of Fractional Operators in Polynomial Regression Models to Enhance the Determination Coefficient R^2 on Unseen Data Proposal for the Application of Fractional Operators in Polynomial Regression Models to Enhance the Determination Coefficient R 2 on Unseen Data", doi: 10.48550/arXiv.2503.11749.

[18] T. O. Hodson, "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not," Jul. 19, 2022, *Copernicus GmbH*. doi: 10.5194/gmd-15-5481-2022.

[19] J. Qi, J. Du, S. M. Siniscalchi, X. Ma, and C.-H. Lee, "On Mean Absolute Error for Deep Neural Network Based Vector-to-Vector Regression," Aug. 2020, doi: 10.1109/LSP.2020.3016837.