Application of Genetic Algorithm for Random Forest Hyperparameter Optimization in Forecasting Indonesia's Monthly Oil and Gas Export Value by Destination Country

Maha Valne Datin Mahfujah Tambunan, Muhammad Syahputra Novelan

Abstract

Indonesia's oil and gas export value fell by 17.69 percent in 2025 and fluctuated sharply from month to month, so foreign exchange policy needs accurate forecasts. This study applies a Genetic Algorithm to optimize the hyperparameters of a Random Forest for forecasting Indonesia's monthly oil and gas export value by destination country. The data come from Statistics Indonesia (BPS) for January 2022 to July 2026 and cover 27 destination countries or regions. The features are the values of the previous one to three months, the three-month moving average, the month, and the country. Data from 2022 to 2025 were used for training and data from 2026 for testing, while the fitness function of the Genetic Algorithm was the RMSE on 2025 validation data. The Genetic Algorithm selected 50 trees, a maximum depth of 5, a minimum of 5 samples per leaf, and sqrt for max_features. On the test data, RMSE fell from 48,460.01 to 44,998.70 thousand US$ (7.14 percent) and WAPE fell from 45.09 to 42.70 percent compared with the default Random Forest. However, random search with the same evaluation budget produced an RMSE of 44,661.51, and a three-month moving average produced an RMSE of 40,708.76. This study concludes that hyperparameter optimization improves the Random Forest, but the advantage of the Genetic Algorithm over random search is not supported by this data.

Keywords: *genetic algorithm, random forest, hyperparameter optimization, time series forecasting, oil and gas exports*

Maha Valne Datin Mahfujah Tambunan
Magister of Information Technology, Universitas Pembangunan Panca Budi, Indonesia
e-mail: mhvlndtn04@gmail.com

Muhammad Syahputra Novelan
e-mail: putranovelan@dosen.pancabudi.ac.id
3rd International Conference Epicentrum of Economic Global Framework (ICEEGLOF)
Theme: Cross-Sector and Cross-Disciplinary Strategies for Achieving the SDGS
https://proceeding.pancabudi.ac.id/index.php/ICEEGLOF

Introduction

Oil and gas exports remain one of the components of Indonesia's international trade [1]. According to an official release from Statistics Indonesia (BPS) [2], Indonesia's oil and gas export value from January to December 2025 was US$13.07 billion, a decrease of 17.69 percent compared with 2024. In the same period, non-oil and gas exports rose 7.66 percent to US$269.84 billion. The monthly data used in this study also show that export values change from one month to the next. These fluctuations make it difficult to plan foreign exchange earnings and trade policy, especially because export values can be affected by changes in world oil prices and economic conditions [3].

These conditions call for a forecasting tool that can capture the fluctuating and not always linear patterns in monthly oil and gas export data. This is even more necessary because the data are split further across dozens of destination countries, each with its own value level and seasonal pattern [4]. Several earlier studies have used various methods to forecast Indonesia's export value. Yulisa et al. [5] used Long Short-Term Memory and Gated Recurrent Unit to forecast Indonesia's oil and gas export value. Junita and Kartikasari [6] used a Neural Network Autoregressive model. Random Forest (RF), introduced by Breiman [7], is a decision tree based ensemble method. In regression, this model combines the predictions of a number of trees by averaging them [8]. This study uses a single RF model trained on pooled data across countries, with the country identity encoded as one of the features.

The performance of classification algorithms can differ depending on the dataset and the problem being studied. Barutu et al. [9], including Muhammad Iqbal and Darmeli Nasution, compared C4.5 and Random Forest (RF) for predicting diarrheal disease. In that study, C4.5 reached an accuracy of 0.92, while RF reached 0.88. In contrast, Desvazulinda et al. [10], including Muhammad Syahputra Novelan, reported that RF was better in accuracy and prediction stability than C4.5 for classifying student eligibility. Aryza et al. [11] also applied RF to classify prospective recipients of an aid program. These results show that the choice and performance of an algorithm need to consider the context of the data used. Although the Genetic Algorithm and RF have been used in many studies, the relevant articles above still focus on classification. They show how RF is used in various classification problems, but none of them discusses optimizing RF hyperparameters with the Genetic Algorithm for time series forecasting. Therefore, this study applies the Genetic Algorithm to find an RF hyperparameter configuration that suits the dataset used, and uses it to forecast Indonesia's monthly oil and gas export value by destination country.

Based on a literature search, studies that combine the Genetic Algorithm to optimize Random Forest hyperparameters for economic time series forecasting, especially Indonesia's monthly oil and gas export value by destination country, are still limited. This study offers novelty in three aspects. The first is applying the Genetic Algorithm to a time series forecasting regression problem, not only classification [12]. The second is building a single Random Forest model using panel data from various oil and gas export destination countries, with the destination country as one of the features so that the model can use patterns that may be similar across countries. The third is splitting the training and test data by time order, not randomly, so that the evaluation measures the model's ability to forecast periods it has never observed. In line with this, the study aims to design and test a forecasting model for Indonesia's monthly oil and gas export value by destination country using a Random Forest whose hyperparameters are optimized with the Genetic Algorithm, and then to compare its performance with a baseline Random Forest model that uses default hyperparameters. Academically, this study is expected to enrich the literature on applying the Genetic Algorithm to time series forecasting regression. Practically, the resulting model is expected to help policy makers, such as the Ministry of Trade or Statistics Indonesia, in projecting oil and gas export value by destination country and in preparing trade policies that respond better to trends identified early.

Literature Review

Previous studies on Indonesian export forecasting have used time series, neural network, and autoregressive approaches. Nurhasanah et al., Yulisa et al., and Junita and Kartikasari applied different forecasting methods to export or related energy data [4]–[6]. Soedanto et al. and Sadiah and Ginting also show that oil prices, international trade, and economic conditions are relevant to the analysis of Indonesian exports [1], [3]. These studies provide the economic and forecasting background for examining monthly oil and gas export values by destination country.

Random Forest is suitable for structured numerical data because it can model nonlinear relationships through an ensemble of decision trees. Its use for product demand prediction and crude oil price forecasting has been reported by Banjarnahor et al. and Rahkmawati et al. [8], [13]. Several studies have also examined hyperparameter optimization with Genetic Algorithms, including applications to Random Forest and ensemble models [12], [14]–[19]. However, most of these studies focus on classification or other prediction domains, while the application of Genetic Algorithm optimized Random Forest to monthly oil and gas export forecasting by destination country remains limited.

The evaluation of forecasting models should consider the scale of the target data and should not rely on a single metric [20]. Simple benchmarks such as naive and moving average models are also important for assessing whether a complex model provides additional value [21], [22]. Based on this literature, the present study applies a Genetic Algorithm to optimize Random Forest hyperparameters for time ordered panel data on Indonesian oil and gas exports. The optimized model is compared with a default Random Forest and simple reference models to assess its forecasting performance.

Research Methodology

3.1 Data Type and Source

This study uses secondary data in the form of Indonesia's monthly oil and gas export value by destination country, officially published by Statistics Indonesia (BPS) in monthly table format for 2022 to 2026. The original data are in a wide table, where one row represents one destination country or region and one column represents one month, in thousand US dollars. Each row in the BPS table falls into one of four categories: a main country or region row, a country breakdown row under a region, a regional subtotal row, and a national total oil and gas export row. For modeling at the destination country level, only the main country or region rows and the country breakdown rows are used, while the regional subtotal rows and the national total row are excluded to avoid double counting of export value during model training. After transforming the wide format into a panel time series format, where one row represents one combination of year, month, and country, 27 oil and gas export destination countries or regions were obtained that are consistently recorded throughout the observation period. The observations run from January 2022 to July 2026 because data for the following month had not yet been published by BPS when the research was carried out.

Figure 1 shows the total national monthly oil and gas export value from January 2022 to July 2026, which consists of 55 monthly observations. The highest value was recorded in August 2022 at 1,662,875.4 thousand US$, while the lowest occurred in May 2026 at 758,116.5 thousand US$. This movement, which goes up and down without a fixed pattern, shows that forecasting this data is indeed not easy.

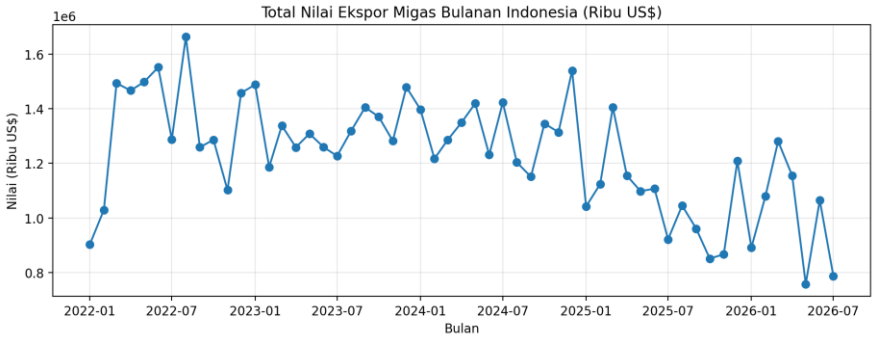


Figure 1. Total Monthly Oil and Gas Export Value of Indonesia (Thousand US$)

3.2 Reasearch Stages

This study was carried out in eight sequential stages. The first stage is collecting and cleaning the BPS data. The second stage is transforming the data from wide format to panel time series format. The third stage is feature engineering based on historical values. The fourth stage is splitting the training and test data based on time. The fifth stage is training the baseline Random Forest model with default hyperparameters [13]. The sixth stage is optimizing the Random Forest hyperparameters using the Genetic Algorithm [14]. The seventh stage is retraining the final Random Forest model using the optimized hyperparameters on all training data. The eighth stage is evaluating and comparing the performance of the baseline model and the optimized model, both in aggregate and by destination country. The whole process was done in the Python programming language, with the scikit-learn library for Random Forest modeling and a custom implementation of the Genetic Algorithm along with all of its genetic operators [16].

3.3 Feature Engineering

Random Forest is not a native time series model that automatically recognizes time order, so the monthly export series of each country first needs to be converted into tabular data. The features built consist of three lag values, one moving average, a month indicator, and the label encoding of the destination country. The lag feature at month t for a given country is defined as the actual export value k months earlier, as formulated in Equation (1).

$$lag_k(t) = y_{t-k}, \qquad k = 1, 2, 3 \tag{1}$$

The three-month moving average feature is computed from the average of the three actual values before month t, without including the value of month t itself so that no information leaks from the future, as formulated in Equation (2)

$$RollMean3(t) = \frac{1}{3}\sum_{k=1}^{3} y_{t-k} \tag{2}$$

Besides these two historical features, a month variable with values from 1 to 12 is added to capture seasonal patterns, along with a label-encoded variable of the destination country name so that a single model can learn the characteristics of each country in a panel setting. Rows that do not yet have a full three months of history, which generally occur at the start of each country's observation period, are removed from the modeling data because their feature values are incomplete.

3.4 Training and Test Data Split

The training and test data are separated chronologically in time order so that the test data represent the period after the training data. Next, the Random Forest hyperparameters are optimized using the Genetic Algorithm by adapting the approach of Togatorop et al [16], who used the Genetic Algorithm to search for Random Forest parameter combinations and compared its performance with a model using default parameters and other search methods. Data from 2022 to 2025 are used as training data, while the available 2026 data, which are

January to July, are used as test data that the model has never seen during training or the optimization process. For the fitness function of the Genetic Algorithm specifically, the training data are split again into GA training data from 2022 to 2024 and GA validation data from 2025, so the hyperparameter search never peeks at the test data at all. The number of observations in each part is shown in Table 1. Of the 27 destination countries, only 19 have test data because the other eight countries do not have complete export values for January to July 2026.

Table 1. Number of Observations in Each Data Subset (27 destination countries in total, 19 countries have test data)

Data Subset	Period	Number of Observations
GA training data	2022 - 2024	872
GA validation data	2025	186
Total training data	2022 - 2025	1.058
Test data	2026 (Jan - Jul)	115
Total	2022 - 2026 (Jul)	1.173

3.5 Random Forest Regressor Model

The Random Forest Regressor is used in this study to predict export value based on historical features, month, and destination country. Earlier studies have applied Random Forest to predict product demand [8] and to forecast Indonesian crude oil prices [13] For regression tasks, the final prediction is obtained by averaging the predictions of all trees, as formulated in Equation (3).

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} T_b(x) \tag{3}$$

where B is the total number of trees in the forest and $T_b(x)$ is the prediction of the b-th regression tree for feature vector x. The performance of Random Forest is strongly affected by five main hyperparameters that become the optimization targets of this study, namely the number of trees (n_estimators), the maximum tree depth (max_depth), the minimum number of samples required to split a node (min_samples_split), the minimum number of samples at a leaf node (min_samples_leaf), and the number of features considered at each node split (max_features). The search space for the values of these five hyperparameters is shown in Table 2.

Table 2. Random Forest Hyperparameter Search Space

Hyperparameter	Search Space Values
n_estimators	50, 75, 100, ..., 400 (multiples of 25)
max_depth	None, 3, 5, 7, 10, 15, 20
min_samples_split	2, 3, 4, 5, 6, 8, 10
min_samples_leaf	1, 2, 3, 4, 5
max_features	sqrt, log2, None

As a comparison, this study also trains a baseline Random Forest with the default scikit-learn hyperparameters without any tuning, to assess how much the Genetic Algorithm optimization benefits model performance. All Random Forest models in this study use a random_state of 42 so that the results can be reproduced.

3.6 Hyperparameter Optimization with the Genetic Algorithm

The Genetic Algorithm (GA) is a population-based evolutionary optimization method that represents candidate solutions as chromosomes and evaluates them with a fitness function. The search for solutions is done by forming new generations through selection,

crossover, and mutation operators [17]. In this study, one chromosome represents one combination of the five Random Forest hyperparameters in Table 2, with each gene being an index that points to one value in the search space of the corresponding hyperparameter.

The fitness value of each chromosome is determined by decoding the genes into a candidate Random Forest hyperparameter set, training the model on the training data, and then evaluating its predictions on the validation data using RMSE as the basis of optimization. The use of GA for Random Forest hyperparameter tuning was applied by Munshi et al. (2025) [18], while Essa et al. (2024) [19] used RMSE as the objective function to be minimized in GA-based optimization. Since the goal of this study is to minimize error, the fitness function is defined as the negative of the validation RMSE, as formulated in Equation (4).

$$fitness(\theta) = -RMSE_{val}(\theta) \tag{4}$$

where θ is the hyperparameter combination decoded from one chromosome. The smaller the validation RMSE, the larger the fitness value of the chromosome, because the value gets closer to zero. The evolution process runs with tournament selection of size k=3, which means choosing three chromosomes at random from the population and taking the one with the best fitness as a parent. The two selected parents then go through one-point crossover, which swaps the genes on one side of a random cut point between the two parents to produce two child chromosomes. Each gene in a child chromosome has a 20 percent mutation probability of being replaced with a new random value from the search space of the corresponding hyperparameter, so that population diversity is maintained and premature convergence can be avoided. The best chromosome in each generation is always carried over to the next generation through an elitism mechanism, so the best solution found so far is not lost because of crossover or mutation. The Genetic Algorithm parameters used are shown in Table 3. The random number generator for all genetic operators was given seed 42 and the search process was run once.

Table 3. Genetic Algorithm Parameters Used

Genetic Algorithm Parameter	Value
Population size (POP_SIZE)	12 chromosomes
Number of generations (N_GEN)	8 generations
Mutation probability	0.2 (20 percent)
Tournament size	3 chromosomes
Elitism mechanism	1 best chromosome per generation

The choice of the Genetic Algorithm (GA) for optimizing Random Forest hyperparameters is supported by Togatorop et al. [16], who reported that a combination of Random Forest and GA gave better accuracy than grid search in heart disease classification. That finding shows the potential of GA for hyperparameter optimization, but it does not prove that this method is always better on all kinds of data. Therefore, this study will test the performance of GA in forecasting Indonesia's monthly oil and gas export value.

3.7 Model Evaluation

The performance of the baseline Random Forest and the Random Forest optimized by the Genetic Algorithm is evaluated on the 2026 test data using RMSE, MAE, MAPE, and WAPE. Using several metrics gives a more complete picture of the evaluation, because the choice of metric needs to consider the characteristics of the data and should not depend on a single measure [20]. MAPE is undefined when the actual value is zero and can produce a very large percentage error when the actual value is close to zero. Therefore, WAPE is used as a complementary metric because it compares the sum of absolute errors with the sum of actual values and remains defined when some actual values are zero, as long as the sum of actual values in the test data is not entirely zero. The first three metrics are formulated in Equations (5), (6), and (7), respectively.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (5)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (6)$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (7)$$

Equation (7) is unstable on this data because several destination countries have very small or zero actual values in certain months, so even a small difference produces a very large percentage error. For this reason, the Weighted Absolute Percentage Error (WAPE) is added as a complementary metric, which is the total absolute error divided by the total actual value as in Equation (8). In this way, countries with large values get a proportional weight and a zero value in one row does not cause division by zero.

$$WAPE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{\sum_{i=1}^{n}|y_i|} \times 100\% \quad (8)$$

where n is the number of observations in the test data, y_i is the actual export value, and $\hat{y}_i$ is the value predicted by the model. For the MAPE calculation specifically, rows with an actual value of zero are excluded from the calculation to avoid division by zero, since some small oil and gas export destination countries sometimes record export values close to or even equal to zero in certain months. Besides the aggregate evaluation across all destination countries, these metrics are also computed separately for each of the 19 destination countries that have test data, because the range of export values between countries is very wide, from countries with values in the hundreds of millions of US dollars such as Singapore to small countries such as the Philippines or Myanmar, so an aggregate evaluation alone risks being biased toward countries with a large value scale.

3.8 Comparison Models

To assess the performance of the Random Forest (RF) optimized by the Genetic Algorithm, the model is compared with the baseline RF and three other comparison models. Using a simple model as a benchmark helps judge whether a complex model really gives a benefit. The study by Beck et al. [21] stresses the importance of comparing forecasting models with a naïve baseline. Yu et al. [22] also included naïve and moving average methods as comparison models in time series forecasting evaluation. The first comparison is the three-month moving average (MA3), which predicts the value of month t as the average of the last three actual values available, following Equation (2). The second comparison is the naïve method, which uses the actual value of the previous month to predict month t, following Equation (9).

The third comparison is an RF with hyperparameters selected through random search, which means taking candidates at random from the search space in Table 2. The experiment evaluated 96 random search candidates, equal to 12 chromosomes over 8 generations in the GA. Each candidate was trained on the 2022 to 2024 training data and scored using RMSE on the 2025 validation data. The candidate with the lowest validation RMSE was then retrained on all 2022 to 2025 training data and evaluated on the January to July 2026 test data. The random search was run once with seed 7. Grid search is not used because the defined search space covers 11,025 combinations, far more than that evaluation budget.

Results

4.1 Hyperparameter Optimazation Results from Genetic Algorithm

The evolution ran for eight generations with a population of 12 chromosomes, so each generation tested 12 hyperparameter combinations and the whole process involved at most 96 model trainings. Figure 2 shows the best validation RMSE in each generation. The value dropped from 34,568.15 in the first generation to 33,814.49 in the seventh generation, then stayed there until the last generation. The decreases happened in the second, fourth, fifth, and seventh generations, while the mean RMSE of the whole population also went down from 35,356.00 to 34,297.85. This means the population as a whole moved toward better combinations, and it was not just one chromosome that happened to be good.

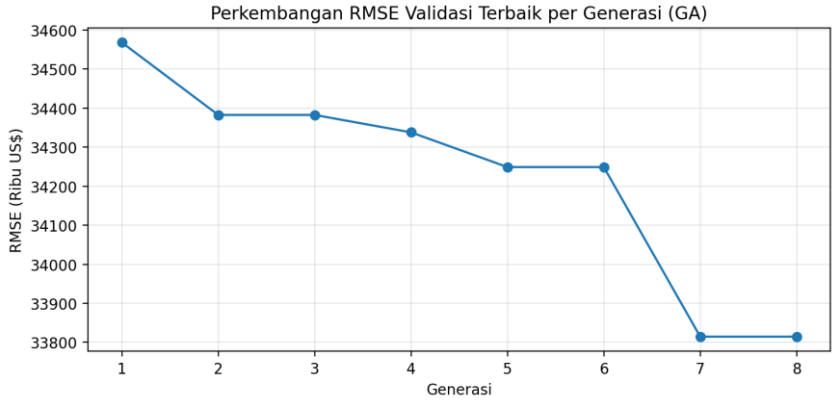


Figure 2. Progress of the Best Validation RMSE in Each Generation

The best hyperparameter combination found by the Genetic Algorithm is compared with the default scikit-learn settings in Table 4. The optimized combination gave a validation RMSE of 33,814.49 thousand US$, which is 7.94 percent lower than the validation RMSE of the default model at 36,730.65 thousand US$.

Table 4. Comparison of Default and Optimized Hyperparameters

Hyperparameter	Default setting	Genetic Algorithm result
n_estimators	100	50
max_depth	None (no limit)	5
min_samples_split	2	2
min_samples_leaf	1	5
max_features	All features	sqrt

The pattern that appears is quite clear. The Genetic Algorithm chose a smaller forest, shallow trees, leaves that contain at least five samples, and only a subset of features at each split. All of these point to a more controlled model that does not memorize the training data. This makes sense because the data used are relatively small, namely 872 observations at the search stage, while the export values fluctuate sharply.

4.2 Model Performance Comparion on Test Data

The test data consist of 115 observations from 19 destination countries from January to July 2026. Eight of the 27 countries are not in the test data because they do not have complete export values in that period. The evaluation uses RMSE, MAE, MAPE, and WAPE as formulated in Section 2.7, and compares all the comparison models in Section 2.8.

The aggregate MAPE on this data cannot be interpreted. Some countries have very small actual values, for example the United States with a mean of 5.52 thousand US$ per month, so even a small difference produces a percentage error of hundreds of thousands of percent. WAPE divides the total error by the total actual value so countries with large values get an appropriate weight. The results of all models are shown in Table 5.

Table 5. Model Performance on 2026 Test Data (RMSE and MAE in thousand US$)

Model	RMSE	MAE	WAPE (%)	MAPE (%)
Random Forest baseline	48,460.01	27,334.19	45.09	15,923.06
Random Forest + Genetic Algorithm	44,998.70	25,890.14	42.70	25,869.39
Random Forest + random search	44,661.51	25,864.10	42.66	25,496.55

3-month moving average (MA3)	40,708.76	22,904.10	37.78	10,075.40
Naive (previous month's value)	53,610.66	28,225.93	46.56	3,116.16

Compared with the baseline Random Forest, the Random Forest from the Genetic Algorithm reduced RMSE by 7.14 percent, MAE by 5.28 percent, and WAPE from 45.09 percent to 42.70 percent. The direction of this improvement is the same as that seen on the validation data, so the improvement did not happen only on the data used to search for hyperparameters. MAPE actually rose from 15,923.06 percent to 25,869.39 percent, but this number is dominated by countries with very small values so it is not used as a basis for conclusions.

There are two other findings that need to be reported as they are. First, random search with the same budget gave an RMSE of 44,661.51, which differs by only 0.75 percent from the Genetic Algorithm and is actually slightly lower. On the validation data, random search found a best RMSE of 34,248.97, which equals the Genetic Algorithm's value in the fifth generation. So this study finds no evidence that the Genetic Algorithm is better than random search. Second, the very simple three-month moving average gave an RMSE of 40,708.76, about 9.5 percent lower than the Random Forest from the Genetic Algorithm.

4.3 Evaluation by Destination Country

Because the value scale differs widely between countries, model performance is also viewed by country. Table 6 shows the eight countries with the highest mean exports in the test period. Overall, the Random Forest from the Genetic Algorithm has a lower RMSE than the baseline in 12 of the 19 countries. The largest improvements are in countries with large values, namely Singapore down 12.2 percent (from 105,043 to 92,199) and Malaysia down 11.3 percent (from 48,184 to 42,742). For China the improvement is only 3.0 percent, while for Japan the RMSE actually rose from 75,565 to 79,412.

Table 6. Performance by Country on Test Data (values in thousand US$)

Country	Mean actual	RMSE GA	MAE GA	MAPE GA (%)	RMSE baseline
Singapore	372,309	92,199	73,742	18.7	105,043
Malaysia	174,957	42,742	36,646	19.9	48,184
China	133,068	40,466	37,076	29.5	41,717
Other Asia	115,544	58,364	51,289	64.7	59,316
Japan	99,194	79,412	68,244	4,125.3	75,565
Thailand	45,934	83,465	66,868	7,063.6	93,739
Philippines	32,063	63,883	50,401	13,238.1	70,268
Netherlands	12,750	8,673	8,020	2,459.5	7,638

MAPE can still be read only for the three largest countries, namely Singapore, Malaysia, and China, with errors of 18.7 percent to 29.5 percent. For countries with small or very unstable values such as Japan, Thailand, and the Philippines, MAPE reaches thousands of percent. In addition, the naive model turns out to have a lower RMSE than the Random Forest from the Genetic Algorithm in 12 of the 19 countries. Figure 3 shows the actual values and the predictions of the Random Forest from the Genetic Algorithm for Singapore, China, Malaysia, and Japan.



Figure 3. Actual Values and Predictions of Oil and Gas Exports for Four Destination Countries, January to July 2026

Figure 3 shows a pattern that is consistent with the error numbers. The predictions move smoothly around the value level of the last few months and do not capture sharp spikes, for example exports to Singapore in June, to Malaysia in April, and to Japan in March. For China, the predictions are above the actual values in all seven test months, while exports to Japan in May were close to zero and were not anticipated by the model.

4.4 Feature Contribution to Predictions

Table 7 shows the feature importance in the final model. The four features based on historical values contribute 97.34 percent of the total importance, with the three-month moving average being the highest. The country feature contributes only 2.30 percent and the month feature only 0.35 percent.

Table 7. Feature Importance in the Optimized Random Forest

Feature	Importance (%)
roll_mean_3	38.82
lag3	23.33
lag1	22.98
lag2	12.21
country_enc	2.30
month	0.35

Discussion

The results show that the RF with hyperparameters optimized by the Genetic Algorithm produced lower error than the baseline RF on both the validation and test data. The study by Togatorop et al. [16] also applied the Genetic Algorithm to optimize RF hyperparameters, although in a heart disease classification case. The difference in problem and data characteristics means the results of the two studies cannot be compared directly.

The advantage of the Genetic Algorithm over other search methods is not yet proven in this study. Random search with the same evaluation budget gave a slightly lower test RMSE than RF-GA. Because each method was run only once, that difference is not enough to conclude that one search method is better in general. This study does not compare GA with grid search or Bayesian optimization, so it draws no conclusion about those two methods.

The most important finding to note is that the Random Forest lost to the three-month moving average. The roll_mean_3 feature is already the most important feature of the model, so the model basically learned to follow the value level of the last few months, and then added a correction that turned out to be not always right. This pattern can be seen in the monthly totals across all countries. In May 2026, the actual value was 758,048 thousand US$ while the prediction reached 1,271,177, or 67.7 percent higher. In July the prediction was 40.7 percent above the actual value and in January it was 22.8 percent above. The model seems slow to respond to sudden drops. The month feature, which barely contributes, also indicates that seasonal patterns cannot yet be learned from data that cover only four years.

In terms of use, this model is reasonably suitable as a support tool for the main export destinations such as Singapore, Malaysia, and China, whose percentage errors are still around 19 percent to 30 percent. For small-value countries, the model cannot be relied on and simple methods are often better. Thus, the contribution of this study is better stated as evidence that hyperparameter optimization improves Random Forest in forecasting oil and gas exports by country, as well as a reminder that machine learning models should always be tested against simple comparison models.

Several limitations need to be acknowledged. The test data are only 115 observations from seven months and 19 countries. The Genetic Algorithm was run once with one seed value, and the random search comparison was also run once, so the small difference between them cannot be considered meaningful. The validation uses only one year, the model does not yet include external variables such as world oil prices or exchange rates, and grid search has not been compared. Future research can repeat the experiment with several seeds, use multi-fold time series validation, add external variables, and complete the comparison with grid search.

Conclusion

This study applied the Genetic Algorithm to optimize the hyperparameters of a Random Forest in forecasting Indonesia's monthly oil and gas export value for 27 destination countries or regions, using Statistics Indonesia data from January 2022 to July 2026. The Genetic Algorithm found a combination of 50 trees, a maximum depth of 5, min_samples_split of 2, min_samples_leaf of 5, and max_features of sqrt. The validation RMSE fell from 34,568.15 in the first generation to 33,814.49 thousand US$, or 7.94 percent lower than the default settings.

On the January to July 2026 test data, the optimized Random Forest reduced RMSE from 48,460.01 to 44,998.70 thousand US$ (7.14 percent), MAE from 27,334.19 to 25,890.14 thousand US$ (5.28 percent), and WAPE from 45.09 percent to 42.70 percent. RMSE was lower than the baseline in 12 of the 19 countries, with the largest improvements for Singapore and Malaysia. Thus, hyperparameter optimization is shown to improve Random Forest performance in this case. The aggregate MAPE cannot be used as a basis for assessment because the actual values of several countries are close to zero.

The advantage of the Genetic Algorithm as a search method is not yet proven. Random search with the same budget of 96 evaluations gave an RMSE of 44,661.51 thousand US$, which differs by only 0.75 percent from the Genetic Algorithm. The three-month moving

average even gave an RMSE of 40,708.76 thousand US$, about 9.5 percent lower than the optimized Random Forest. The four features based on historical values contribute 97.34 percent of feature importance, while the month and country features barely play a role. The model is reasonably suitable for the main destinations such as Singapore, Malaysia, and China with MAPE of 18.7 percent to 29.5 percent, but it cannot be relied on for small-value countries.

This conclusion is limited by several things. The Genetic Algorithm and random search were each run only once with one seed, the validation uses only one year, the test data are only 115 observations, and grid search has not been compared. Future research is advised to repeat the experiment with several seeds and multi-fold time series validation, compare it with grid search and Bayesian optimization, add external variables such as world oil prices and exchange rates, and test a modeling approach that separates countries by the scale of their export values.

References

[1] F. U. Sadiah and A. L. Ginting, "Pengaruh Perdagangan Internasional Ekspor Migas dan Non Migas terhadap Perdagangan di Indonesia," vol. 06, no. 1, pp. 1–15, 2024.

[2] R. Wahyu, M. Dwi, R. W. Saputra, and M. D. Kartikasari, "Pendekatan Metode Prophet dalam Peramalan Nilai Ekspor Jawa Tengah untuk Perencanaan Ekonomi Daerah Pendekatan Metode Prophet dalam Peramalan Nilai Ekspor Jawa Tengah untuk Perencanaan Ekonomi Daerah," vol. 14, no. 2, pp. 309–320, 2026.

[3] E. Soedanto, P. D. Utami, A. Salsabillah, and B. Refangga, "DAMPAK FLUKTUASI HARGA MINYAK DUNIA TERHADAP EKONOMI DI INDONESIA," vol. 2, no. 1, pp. 231–242, 2025.

[4] E. Nurhasanah, Y. Sukmawaty, P. S. Statistika, and U. L. Mangkurat, "Peramalan Ekspor Migas di Indonesia Menggunakan Pendekatan," vol. 7, no. 1, pp. 87–97, 2024, doi: 10.13057/ijas.v7i1.84934.

[5] P. N. Yulisa, M. Al Haris, P. R. Arum, and U. M. Semarang, "Peramalan Nilai Ekspor Migas di Indonesia dengan Model Long Short Term Memory (LSTM) dan Gated Recurrent Unit (GRU)," vol. 16, no. 1, pp. 328–341, 2023.

[6] T. P. Junita and M. D. Kartikasari, "APPLICATION OF THE NEURAL NETWORK AUTOREGRESSIVE (NNAR) METHOD FOR FORECASTING THE VALUE OF," vol. 18, no. 1, pp. 341–348, 2024.

[7] N. Asyifa and M. S. Novelan, "Performance Analysis of the K-Means Algorithm in Classifying Organic and Inorganic Waste Types," pp. 6681–6690, 2025.

[8] L. Banjarnahor, M. I. Sarif, and M. S. Novelan, "Analisis Machine Learning untuk Prediksi Permintaan Produk pada Percetakan PT . Next Sihar Dagulon Menggunakan Algoritma Random Forest," vol. 5, no. 2, pp. 1523–1530, 2026.

[9] S. Barutu, M. Iqbal, and D. Nasution, "Comparative Analysis of the C4 . 5 and Random Forest Algorithms for the Prediction of Diarrheal Disease," vol. 4, no. 07, pp. 625–635, 2025, doi: 10.58471/esaprom.v4i07.

[10] E. Desvazulinda, M. Iqba, and M. S. Novelan, "Analysis of the Decision Tree (C4 . 5) and Random Forest Algorithms to Determine Student Eligibility for Final Project Assignments Based on Academic Requirements," vol. 7, pp. 1667–1675, 2026, doi: 10.30865/json.v7i4.9947.

[11] S. Aryza, M. Furqon, Z. Lubis, and A. R. Nasution, "Explainable Data-Driven Machine Learning for Identifying MBG Program Beneficiaries in Medan City," vol. 7, no. 1, pp. 1109–1118, 2026, doi: 10.30596/jcositte.v7i1.29560.

[12] M. K. Suryadi, R. Herteno, S. W. Saputro, M. R. Faisal, and R. A. Nugroho, "Comparative Study of Various Hyperparameter Tuning on Random Forest Classification with SMOTE and Feature Selection Using Genetic Algorithm in Software Defect Prediction," vol. 6, no. 2, pp. 137–147, 2024.

[13] Y. Rahkmawati, S. Annisa, H. Hafid, and E. Safitri, "Comparison of ARIMA , Random Forest , and Hybrid ARIMA-Random Forest Models in Forecasting Indonesian Crude Oil Prices," vol. 11, no. 1, pp. 593–612, 2026.

[14] B. P. Kaur, H. Singh, R. Hans, S. K. Sharma, C. Sharma, and M. M. Hassan, *A Genetic algorithm aided hyper parameter optimization based ensemble model for respiratory disease prediction with Explainable AI*. 2024. doi: 10.1371/journal.pone.0308015.

[15] P. Prima, A. R. I. O. Adriansyah, and A. N. U. R. Usyaid, "Prediksi Retensi Mahasiswa Menggunakan Algoritma Random Forest dengan Optimasi Algoritma Genetika," vol. 11, no. 1, pp. 115–126, 2026.

[16] P. R. Togatorop, M. Sianturi, D. Simamora, and D. Silaen, "Optimizing Random Forest using Genetic Algorithm for Heart Disease Classification," vol. 13, no. 1, pp. 60–71, 2022.

[17] M. Binjubier, M. Arfian, M. Othman, S. Kasim, and H. Amnur, "Optimizing Genetic Algorithm by Implementation of An Enhanced Selection Operator," vol. 8, no. November, pp. 1643–1650, 2024.

[18] T. A. Munshi, K. Popi, L. N. Jahan, M. F. Howladar, and M. Hashan, "Applied Computing and Geosciences Stacking modeling with genetic algorithm-based hyperparameter tuning for uniaxial compressive strength prediction," *Appl. Comput. Geosci.*, vol. 27, no. February, p. 100276, 2025, doi: 10.1016/j.acags.2025.100276.

[19] M. G. Essa, C. M. Elias, and C. M. Elias, "Utilization of genetic algorithm in tuning the hyper-parameters of hybrid NN-based side-slip angle estimators," *Neural Comput. Appl.*, vol. 36, no. 30, pp. 19055–19074, 2024, doi: 10.1007/s00521-024-10115-3.

[20] H. Khoshvaght, R. Ramyad, A. Razmjou, and M. Khiadani, "Journal of Environmental Chemical Engineering A critical review on selecting performance evaluation metrics for supervised machine learning models in wastewater quality prediction," *J. Environ. Chem. Eng.*, vol. 13, no. 6, p. 119675, 2025, doi: 10.1016/j.jece.2025.119675.

[21] N. Beck, J. Dovern, and S. Vogl, "Mind the naive forecast ! a rigorous evaluation of forecasting models for time series with low predictability," 2025, doi: 10.1007/s10489-025-06268-w.

[22] Y. Yu, J. Peeperkorn, J. De Smedt, and J. De Weerdt, *A benchmarking study on process model forecasting : univariate vs . multivariate approaches*. 2025.