Forecasting Import-Export Volume and Value Using Support Vector Machines and Random Forests

Hamda Yauma Istiqlala[b], Nughthoh Arfawi Kurdhi[a]

Department of Mathematics, Faculty of Mathematics and Natural Sciences, Sebelas Maret University, Indonesia

[a]Corresponding author: arfa@mipa.uns.ac.id

[b]hamdayauma@student.uns.ac.id

Abstract. National trade plays a crucial role in a country's economy by meeting the demand for goods and services. Imports and exports are vital components that significantly influence economic growth. This study aims to forecast and analyse the volume and value of imports and exports in the coming years, thereby helping the country to prepare for and enhance its competitiveness on the international stage. To achieve this, the study employs Support Vector Machine (SVM) and Random Forest methods, which have gained popularity for their high efficiency in handling classification and regression problems. These methods are particularly effective in producing accurate forecasts, especially when dealing with seasonal variations in data. The research also explores the potential of combining these methods for even more robust predictions. Additionally, the study considers external factors, such as global market trends and policy changes, which can influence trade dynamics. The findings of this research demonstrate the ability to predict future import and export volumes and values accurately. By comparing the error values in the form of Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) from the model, it is proven that the SVM model is highly effective for export-import forecasting. Lower error values indicate higher prediction accuracy, highlighting the model's reliability in capturing complex trade dynamics. The insights from this study can be instrumental for policymakers and businesses in strategising for future trade activities, allowing them to make informed decisions that align with anticipated market shifts.

Keywords: *Forecasting, Ekspor, Impor, SVM, Random Forest, MAPE, RMSE, MAE.*

INTRODUCTION

National trade is a crucial pillar in the global economy, enabling countries to meet the demand for goods and services(1). Two critical international trade activities, exports and imports, are fundamental components that indirectly influence economic growth(2). Both are essential, but accurate import-export forecasting plays a critical role in supporting government policymaking, preparing human resources, and establishing benchmarks for enhancing a country's competitiveness on the international stage(3). Therefore, precise forecasting of import-export volumes is of paramount importance. Imports represent an international business activity involving the entry of goods into a country's customs territory, in this case, Indonesia. These activities are undertaken by individuals or companies engaged in export-import transactions while adhering to applicable laws and regulations(4). Importing involves transporting goods and merchandise from one country to another, typically as part of international trade agreements. This process requires customs clearance in exporting and importing countries, ensuring the goods comply with all regulatory requirements. Imports play a vital role in international trade by fulfilling the domestic demand for goods that are either not produced locally or are insufficient to meet the population's needs. Moreover, importing goods in bulk can also be a strategic move for countries to access products at competitive prices, enhance the quality of life, and stimulate economic activity(5). As global supply chains become increasingly interconnected, the importance of efficient import processes and accurate forecasting grows, making it essential for countries to stay ahead in the competitive international market.

Exporting is selling goods and services from one country to another and serving as a primary source of income for many economies(6). By exporting, a country can expand the market for its domestic products, create jobs, and boost national income. Exports also contribute to economic diversification, enabling countries to reduce their reliance on domestic markets and mitigate financial risks(7). In today's globalised world, exporting has become increasingly important due to rising economic integration and strengthening international relationships. The export process involves not only economic factors but also political, social, and cultural considerations. Trade policies, international agreements, and government regulations significantly influence a country's ability to export goods and services. Moreover, shifts in global demand, emerging technologies, and innovations in production also play crucial roles in determining the competitiveness of export products(8). Understanding export potential is, therefore, essential for decision-makers who aim to formulate effective business strategies(9). A comprehensive grasp of the global market landscape allows countries to identify new opportunities, adapt to changes in consumer preferences, and leverage technological advancements(10). By fostering robust export strategies, nations can enhance their economic resilience, secure a stable income stream, and position themselves favourably in the international market.

Forecasting is both an art and a science involving predicting future events based on historical data and mathematical models(11). It encompasses using scientific methods—rooted in technology and mathematics—to project future trends(12). However, forecasting is not exclusively based on scientific or structured procedures; it can also involve intuition and informal group discussions. The concept of Support Vector Machines (SVM) was introduced by Vapnik et al. in 1996)(13). SVM is a versatile algorithm widely used for both classification and regression tasks. For regression, SVM operates on principles similar to classification but with some variations. It involves mapping input data into a high-dimensional space using kernel functions, which assist in finding optimal hyperplanes for separating data points(14). Several kernel functions can be employed for forecasting, including polynomial kernels, radial basis functions (RBFs), sigmoid functions, and linear functions. Each kernel function has its advantages depending on the nature of the data and the forecasting objectives(15). SVM is a supervised learning model investigating data to classify samples into distinct categories. Training algorithms for SVM build models that can accurately classify new datasets. SVM excels in linear and nonlinear classification and regression, demonstrating its robustness and adaptability in various machine-learning applications.

On the other hand, Random Forest offers significant advantages, including reduced overfitting and increased accuracy, by aggregating the predictions of multiple decision trees(16). This ensemble learning method handles complex data sets with numerous variables and intricate interactions. Applying machine learning methods, such as Random Forest, in forecasting import and export trends has gained traction among researchers worldwide(17). For example, a study by Zhang et al. (2020) in China demonstrates that the Random Forest model outperforms traditional linear models in predicting import trends. This advantage is particularly evident when dealing with large and complex datasets where simpler models do not easily capture interactions between variables (18). The effectiveness of the Random Forest method is attributed to its ability to average multiple decision trees, thereby reducing the risk of overfitting and enhancing predictive accuracy(19). However, achieving optimal performance with Random Forest requires careful parameter selection. Tuning parameters such as the number of trees, tree depth, and splitting criteria are crucial for maximising the model's accuracy. This research underscores the potential of Random Forest to improve forecasting precision across various contexts and highlights the importance of proper model configuration to leverage its full capabilities.

A technique known as the network search method is instrumental in identifying optimal parameters for models, which is crucial for accurately predicting unlabeled data, such as experimental datasets(20). This method is classified as a comprehensive approach because it first requires defining the type of prediction value being sought. It then evaluates and scores different parameter values to determine the most suitable one(21). The network search method is beneficial when the maximum desired value falls within a specified range defined by each independent variable's upper and lower bounds. For algorithms like Support Vector Machines (SVM) and Random Forest, employing network search techniques is essential for parameter optimisation (22). Network search systematically explores various parameter sets, allowing researchers to identify the combination that delivers the best model performance. This approach enables fine-tuning algorithms to achieve more accurate and reliable forecasting models. The ultimate goal is to enhance the model's predictive capabilities, leading to improved forecasting outcomes.

RESEARCH METHODS

This study analyses and forecasts the volume and value of imports and exports using relevant historical data. The dataset for this study comprises import and export records from one of Indonesia's ports, spanning the years 2017 to 2021. Key variables considered in the forecasting model include the volume and Free on-board (FOB) value of imports and exports. The research begins with collecting historical data, which will be scrutinised for missing values and normalised for analysis. The cleaned data will then be used to build the forecasting model, while a separate test dataset will be utilised to evaluate the model's performance(23). Forecasting time series data can be approached through various methods. Traditional forecasting techniques often struggle to capture complex patterns in the data(19).

In this study, we employ Support Vector Machines (SVM) and Random Forests, enhanced through Grid Search, to address these challenges(24). Adopting these machine learning methods reflects their growing popularity and efficiency in solving classification and regression problems. This study aims to assess the effectiveness of these advanced seasonal regression methods for time series analysis(25). We strive to develop hybrid models combining SVM and Random Forest to assist the government in making informed decisions that could stimulate economic growth. Once the model is trained, its performance will be evaluated using the test dataset, with evaluation metrics focusing on the Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) (26).

RESULTS AND DISCUSSION

Data Collection

The data presented in Table 1 and Table 2 provide a comprehensive view of import and export activities at one of Indonesia's ports from 2015 to 2021. Table 1 shows the historical figures for import and export volumes and their corresponding free-on-board (FOB) values. The volume data indicates the total quantity of goods handled through the port each year. At the same time, the FOB values represent the monetary worth of these goods at the point of shipment, excluding shipping and insurance costs. This historical data is crucial for identifying trends and patterns in trade activities over the years. For example, the volume of imports fluctuated from approximately 7.5 billion in 2015 to 19.9 billion in 2021, and the FOB values ranged from around 13.8 billion in 2015 to 20 billion in 2021, reflecting variations in trade value.

Table 2 shows the actual or target data for import and export volumes and FOB values. This table provides benchmarks for evaluating the forecasting models. The volume and FOB value figures presented here reflect the targets or observed outcomes that forecasting models aim to achieve. For instance, the volume of imports and exports in 2021 reached 24.7 billion and 12.7 billion, respectively, with FOB values of 24.8 billion and 23.7 billion. Comparing these figures with those in Table 1 allows for a critical assessment of forecasting accuracy. Table 1 is the foundation for developing and training forecasting models by offering historical data. In contrast, Table 2 provides the actual or target values necessary for evaluating the performance of these models. We can refine the models by analysing the differences between forecasted and actual values to enhance their accuracy and reliability in predicting future import and export volumes and values.

TABLE 1. Volume and Value of FOB Import and Export in 2015-2021

Tahun	Volume		FOB Value	
	Impor	Ekspor	Impor	Ekspor
2015	7.575.099.724	17.855.508.362	13.834.238.519	12.860.631.493
2016	12.860.631.493	18.745.023.600	13.593.078.453	13.225.617.968
2017	18.113.578.743	8.279.827.963	14.916.241.514	15.623.611.329
2018	16.329.264.807	8.917.386.074	16.229.591.088	14.795.634.023
2019	19.926.618.318	8.867.656.834	15.505.375.823	16.459.279.308
2020	17.598.146.661	9.335.942.648	15.637.718.463	14.292.945.130
2021	18.870.972.615	9.692.156.946	19.003.961.273	18.118.959.324

TABLE 2. Actual/Target Volume Data and FOB Value of Import Export

Year	Volume		FOB Value	
	Impose	Export	Impose	Export
2022	19.908.876.109	10.225.225.578	20.049.179.143	19.115.502.087
2023	21.003.864.295	10.787.612.985	21.151.883.996	20.166.854.702
2024	22.159.076.831	11.380.931.699	22.315.237.616	21.276.031.710
2025	23.377.826.057	12.006.882.942	23.542.575.685	22.446.213.454
2026	24.663.606.490	12.667.261.504	24.837.417.347	23.680.755.194

Forecasting Analysis Using SVM

The last five years of the entire dataset were selected as the validation pool. Given the seasonal component, SVM is automatically run, setting seasonality as t=5 (5-year calculation). The selected forecasting is then compared with the actual validation data. The comparison is shown in the figure.

Periode	Predicted Production impor (SVR)
1	12971751936
2	13969740106
3	14967728835
4	15965717284
5	16963705734
6	17961694183
7	18959682632
8	16464711589
9	17462699058
10	18460688487
11	19458676857
12	20456665306

(a)

Periode	Predicted Production expr (SVR)
1	10178847541
2	9011188271
3	9353334886
4	8959576024
5	8998916116
6	9340757042
7	9540966998
8	9582287228
9	9592918328
10	9610847938
11	9619589978
12	9619428711

(b)

Figure 1. Volume Forecasting Results

Periode	Predicted Production impor (SVR)
1	12998111584
2	14017110109
3	15487259355
4	15502994961
5	15678738508
6	16348879814
7	18167856338
8	15599061912
9	15986182633
10	17878861814
11	19687059892
12	24386002109

(a)

Periode	Predicted Production ekspor (SVR)
1	12843521326
2	14691673711
3	15175388888
4	15472685241
5	15578061683
6	16251696772
7	18181849157
8	15480857297
9	15881330731
10	16094192137
11	19647788163
12	24348113456

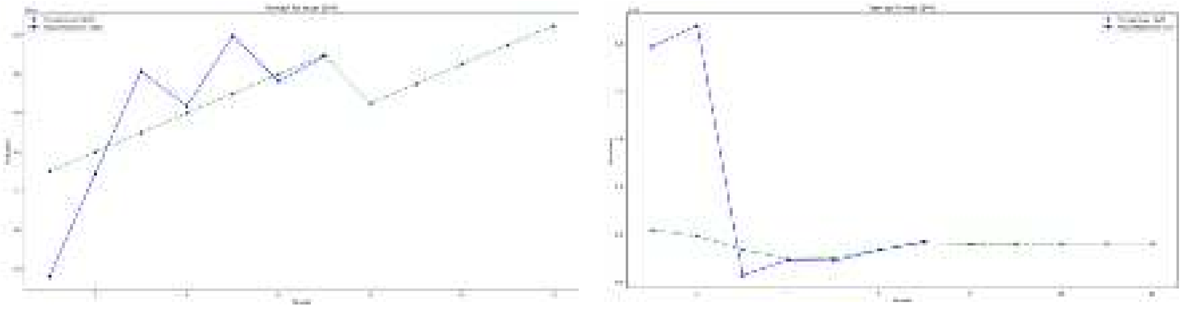
(b)

Figure 2. FOB Value Forecasting Results

The figure above presents the forecasting results using the Support Vector Machine (SVM) model for 2015 to 2026, focusing on the forecast for 2022 to 2026. The forecast predicts a continued increase in import volumes, which could signal a rise in domestic demand or changes in trade policies that encourage higher import levels. This upward trend suggests that ports and warehouses may need to adjust their capacities to effectively accommodate the growing import volumes. Conversely, the forecast for export volumes indicates stability with some improvement. This stability suggests consistent performance in the export sector and efficient export processes. This predictability in export volumes benefits ports as it supports more reliable planning and optimisation of export logistics.

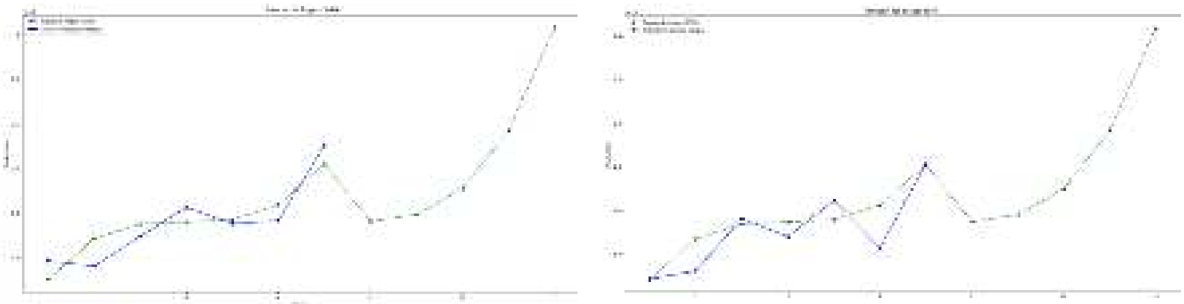
The forecasted increase in the FOB value of imports points to rising import costs, which could be attributed to higher international commodity prices, increased import volumes, or elevated tariffs. This trend is crucial for monitoring, as rising import costs can impact the domestic economy. In contrast, the FOB value of exports shows some fluctuations, reflecting the influence of international market prices and varying commodity types. These fluctuations may necessitate adjustments in foreign markets' pricing strategies or product offerings.

Overall, the forecast results reveal a trend of increasing volumes for both imports and exports, with the export sector remaining relatively stable with minor fluctuations in FOB value. This information is vital for strategic planning by ports and logistics providers, allowing them to enhance operational capacity and efficiency in response to these trends. Policymakers can use this forecasting analysis to assess the economic impact of import and export trade changes and make informed decisions to support financial stability and growth.



(a) (b)

Figure 3. Volume Forecasting Results Chart



(a) (b)

Figure 4. FOB Value Forecasting Results Chart

Based on Figures 3 and 4 display the forecast results for several bound variables. The subplots represent the forecast results for different periods. The subplot shows the general trend of the actual data (blue line) and the predicted data (green line). The comparison between these two lines provides an idea of how well forecasting models can capture historical patterns. Variations in data and predictions make it clear that some subplots show significant value fluctuations, which may reflect external factors affecting import-export activities.

Forecasting Analysis with Random Forest Regressor

The historical data collection, which includes import and export volume and value, undergoes a preprocessing phase. This phase involves optimising the data using advanced methods to enhance the accuracy of forecasting models. Specifically, the process employs a technique that builds many decision trees, combining their results to improve overall accuracy and reduce observational errors. Grid search techniques are also utilised to find the optimal parameter combinations that yield the best forecasting results.

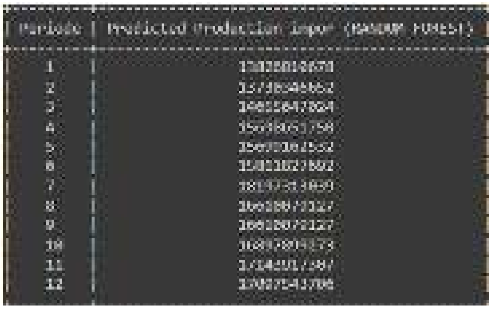
Periode	Predicted Production impor (RANDOM FOREST)
1	12758657219
2	12958559219
3	14519539881
4	16553088455
5	17937740420
6	18199967345
7	18199967345
8	15046441708
9	15046441708
10	18323215008
11	17484344100
12	18126269672

(a)

Periode	Predicted Production expr (RANDOM FOREST)
1	14801171681
2	14881171681
3	13202210022
4	10133150824
5	9267507838
6	3533210067
7	3533210067
8	12906468547
9	12900788587
10	11210762971
11	4833886339
12	4458040885

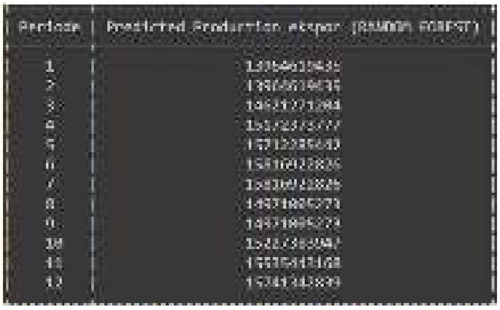
(b)

Figure 5. Volume Forecasting Results



Periode	Predicted Production impor (RANDOM FOREST)
1	13838010678
2	13738046662
3	14011847024
4	15698063758
5	15670162532
6	15011027892
7	18197313839
8	16610877127
9	16610877127
10	16897899373
11	17144017507
12	17087543706

(a)



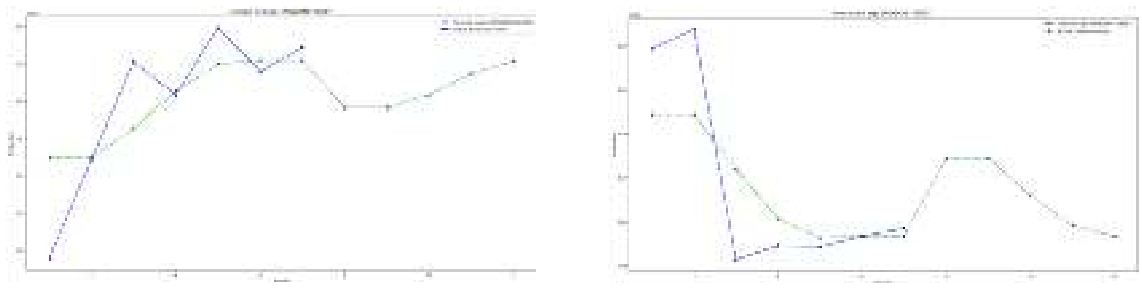
Periode	Predicted Production ekspor (RANDOM FOREST)
1	13764615435
2	13968619635
3	14621771284
4	15172373777
5	14713285683
6	15810222826
7	15810222826
8	14971885271
9	14971885271
10	15227383947
11	15535443168
12	15741342839

(b)

Figure 6. FOB Value Forecasting Results

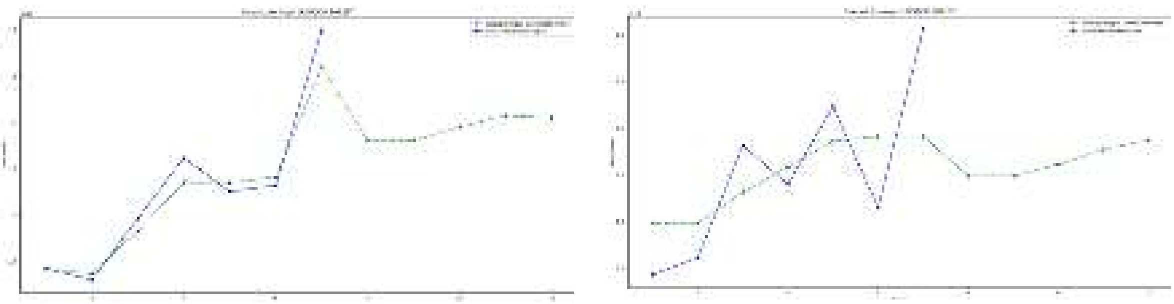
The forecast of import volumes for the period 1-10 exhibits fluctuations, which may be attributed to seasonal factors. This indicates that the Random Forest model effectively captures variations in historical data, reflecting its robustness in handling temporal changes. In contrast, the export volume predictions for the same period, as generated by the NUGA model, reveal different dynamics. Some periods show significant spikes or drops, suggesting that the NUGA model identifies patterns or anomalies that differ from the historical trends.

Evaluating forecasting models involves comparing their predictions with historical data. If the forecasted values closely align with past trends and demonstrate accuracy in capturing variations, it indicates that the models have been well-optimized. Effective optimisation ensures the models can predict future trends based on historical patterns and current data.



(a) (b)

Figure 7. Volume Forecasting Results Chart



(a) (b)

Figure 8. FOB Forecasting Results Chart

The graph reveals notable fluctuations in import volumes. The blue line, representing the actual data, displays sharp variations with distinct peaks and dips. In contrast, the predictions for import volumes show more stability, with the green line—indicating the model's forecasts—generally aligning with the blue line over most of the period. This demonstrates that while the Random Forest model effectively captures broader trends in import volumes, it may not fully account for all the observed fluctuations.

For the FOB (Free on Board) values, the green line shows a trend closer to the blue line than the import volume predictions. This indicates that the model is more successful in capturing overall trends in FOB data. However, there are still points where the predicted values deviate from the actual data, suggesting that the model might not fully capture some influencing factors affecting FOB values.

Overall, the graphical results suggest that the Random Forest model provides reasonably good predictions for import volumes and FOB values, but there is room for improvement. The fluctuations in the import volume graph highlight the need to explore additional variables or employ other techniques to enhance prediction accuracy. Similarly, while the FOB chart shows closer alignment between predictions and actual data, the existing deviations point to the complexity of the factors influencing FOB values, necessitating further evaluation and refinement of the model.

MAPE Calculation on the Support Vector Machine (SVM) Method

Calculating MAPE (Mean Absolute Percentage Error) is crucial for evaluating forecasting accuracy. By determining the smallest MAPE, we can identify the most accurate method for predicting Indonesia's port import and export volumes for 2022-2026. To calculate MAPE, we need the forecasted and actual values from previous data processing. The following table presents the results of the MAPE calculation for the SVM model's import volume forecasts.

In addition to MAPE, two other evaluation metrics commonly used are MAE (Mean Absolute Error) and RMSE (Root Mean Square Error). MAE measures the average absolute difference between the predicted and actual values, providing a clear view of the forecast error magnitude. On the other hand, RMSE gives more weight to more significant errors, as it calculates the square root of the average squared error. By using these three metrics—MAPE, MAE, and RMSE—we can comprehensively assess the performance of the forecasting model and choose the most suitable method for analysing import and export volume data. The table below presents the error calculation results for the import volume forecast using the SVM model.

TABLE 3. Table of calculation results of Import Volume MAPE SVM Model

IMPOR					
TAHUN	**TARGET/ACTUAL**	**FORECAST**	**MAPE**	**MAE**	**RMSE**
2022	19.908.876.109	16.464.711.509	0,17	688.832.919,97	1,18623E+19
2023	21.003.864.295	17.462.699.959	0,17	708.232.867,16	1,25398E+19
2024	22.159.076.831	18.460.688.408	0,17	739.677.684,60	1,36781E+19
2025	23.377.826.057	19.458.676.857	0,17	783.829.839,95	1,53597E+19
2026	24.663.606.490	20.456.665.306	0,17	841.388.236,77	1,76984E+19
EROR			16,93%	3.761.961.548,45	3.771.956.419,30

TABLE 4. Table of SVM Model calculation results Export Volume MAPE

EKSPOR					
TAHUN	**TARGET/ACTUAL**	**FORECAST**	**MAPE**	**MAE**	**RMSE**
2022	10.225.225.578	9.582.287.228,00	0,06	642.938.350,03	4,1337E+17
2023	10.787.612.985	9.592.918.329,00	0,11	1.194.694.655,82	1,4273E+18
2024	11.380.931.699	9.610.047.938,00	0,16	1.770.883.760,99	3,13603E+18
2025	12.006.882.942	9.619.509.978,00	0,20	2.387.372.964,43	5,69955E+18
2026	12.667.261.504	9.619.428.734,00	0,24	3.047.832.770,26	9,28928E+18
EROR			15,37%	1.808.744.500,31	1.998.275.686,90

TABLE 5. Table of MAPE calculation results Import FOB Value SVM Model

IMPOR					
TAHUN	**TARGET/ACTUAL**	**FORECAST**	**MAPE**	**MAE**	**RMSE**
2022	20.049.179.143	15594961912	0,22	4.454.217.231,02	1,98401E+19
2023	21.151.883.996	15906102634	0,25	5.245.781.361,88	2,75182E+19
2024	22.315.237.616	17078863814	0,23	5.236.373.801,65	2,74196E+19
2025	23.542.575.685	19687659093	0,16	3.854.916.591,52	1,48604E+19
2026	24.837.417.347	24306902109	0,02	530.515.238,16	2,81446E+17
EROR			17,80%	3.864.360.844,85	4.240.747.862,67

TABLE 6. Table of MAPE calculation results FOB Export Value SVM Model

EKSPOR					
TAHUN	**TARGET/ACTUAL**	**FORECAST**	**MAPE**	**MAE**	**RMSE**
2022	19.115.502.087	15484857297	0,19	3.630.644.789,82	1,31816E+19
2023	20.166.854.702	15801330731	0,22	4.365.523.970,60	1,90578E+19
2024	21.276.031.710	16994192137	0,20	4.281.839.573,18	1,83342E+19
2025	22.446.213.454	19647700163	0,12	2.798.513.291,24	7,83168E+18
2026	23.680.755.194	24346113457	0,03	665.358.262,77	4,42702E+17
EROR			15,21%	3.148.375.977,52	3.430.682.425,33

Table 3 shows the MAPE, MAE, and RMSE results for forecasting import volumes using the SVM model. The MAPE values for import volumes range from 0.17 in 2022 to 0.17 in 2026, MAE values 688.832.919,97 in 2022 to 841.388.236,77 and RMSE values 1,18623E+19 in 2022 to 1,76984E+19, with an overall MAPE of 16,93%, MAE of 3.761.961.548,77 and RMSE of 3.771.956.419,3. This indicates that the model performs reasonably well but shows increased error in later years. The increase in error percentages over time suggests that while the model captures general trends, its accuracy diminishes as the forecast horizon extends. Thus, while the SVM model is effective for short-term forecasts, its predictive accuracy for more extended periods could be improved. Table 4 presents the MAPE, MAE, and RMSE calculation results for export volumes. The values fluctuate, with a minimum of 0.06 in 2022 and a maximum of 0,24 in 2026 to MAPE values, leading to an overall MAPE of 15.12, a minimum of 642.938.350,03 in 2022 and a maximum of 3.047.832.770,26 in 2026 to MAE values, leading to an overall MAE of 1.808.744.500,31 and a minimum of 1,4273E+18 in 2023 and a maximum of 9,28928E+18 in 2026 to RMSE values, leading to an overall RMSE of 1.998.275.686,90, . This variation reflects the model's differing accuracy across years, with significant deviations observed, particularly in the later forecast periods. The model's performance indicates that while it can forecast export volumes with moderate accuracy, there are challenges in maintaining consistent precision significantly further into the forecast period.

Table 5 details the error results for import FOB values. The MAPE, MAE, and RMSE values for FOB forecasts overall were 17,80%, 3.864.360.844, and 4.240.747.862. Table 6 provides the MAPE results for export FOB values. The MAPE, MAE, and RMSE values for FOB forecasts overall were 15,21%, 3.148.375.977, and 3.430.682.425. The significant deviations reflect the complexity and volatility in international markets, which may not be fully captured by the model.

Overall, the consistent results of MAPE, MAE, and RMSE for both volume and FOB values indicate that SVM could be an effective forecasting model for better accuracy. However, these findings highlight the need for further model comparisons or alternative forecasting techniques to achieve better accuracy and reliability in long-term predictions.

Error Calculation on the Random Forest Method

In forecasting import and export volumes and FOB values, it is crucial to evaluate the accuracy of different models to determine the most reliable method for predicting future trends. This evaluation is typically conducted using the Mean Absolute Percentage Error (MAPE), which measures forecast accuracy by comparing the predicted values with the actual data. The following tables present the MAPE calculation results for the SVM and Random Forest models, focusing on import and export volumes and FOB values from 2022 to 2026. These results offer insights into the performance of each model and highlight areas for potential improvement.

In addition to MAPE, two other evaluation metrics commonly used in forecasting are Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). MAE measures the average absolute difference between the predicted and actual values, providing a simple measure of the prediction error without considering the direction of the error. RMSE, on the other hand, calculates the square root of the average squared error, making it more sensitive to more significant errors or outliers. By combining MAPE, MAE, and RMSE, we can obtain a more comprehensive view of the model's performance and determine which model provides the most accurate predictions for import, export, and FOB value trends.

TABLE 7. Table of MAPE calculation results of Import Volume of Random Forest Model

IMPOR					
TAHUN	TARGET/ACTUAL	FORECAST	MAPE	MAE	RMSE
2022	19.908.876.109	15.646.441.748	0,21	4.262.434.360,83	1,81683E+19
2023	21.003.864.295	15.646.441.748	0,26	5.357.422.546,81	2,8702E+19
2024	22.159.076.831	16.323.215.008	0,26	5.835.861.823,02	3,40573E+19
2025	23.377.826.057	17.484.344.390	0,25	5.893.481.666,73	3,47331E+19
2026	24.663.606.490	18.126.269.673	0,27	6.537.336.816,85	4,27368E+19
EROR			24,99%	5.577.307.442,85	5.628.454.584,62

TABLE 8. Table of MAPE Random Forest Model Export Volume calculation results

EKSPORT					
TAHUN	**TARGET/ACTUAL**	**FORECAST**	**MAPE**	**MAE**	**RMSE**
2022	10.225.225.578	12906980588	0,26	2.681.755.009,97	7,19181E+18
2023	10.787.612.985	12906980588	0,20	2.119.367.603,18	4,49172E+18
2024	11.380.931.699	11210982992	0,01	169.948.706,99	2,88826E+16
2025	12.006.882.942	9833886339	0,18	2.172.996.603,43	4,72191E+18
2026	12.667.261.504	9358869805	0,26	3.308.391.699,26	1,09455E+19
EROR			18,32%	2.090.491.924,57	2.340.076.127,32

TABLE 9. Table of MAPE calculation results Import FOB Value Random Forest Model

IMPORT					
TAHUN	**TARGET/ACTUAL**	**FORECAST**	**MAPE**	**MAE**	**RMSE**
2022	20.049.179.143	16.610.079.127	0,17	3.439.100.016,02	1,18274E+19
2023	21.151.883.996	16.610.079.127	0,21	4.541.804.868,88	2,0628E+19
2024	22.315.237.616	16.897.899.274	0,24	5.417.338.341,65	2,93476E+19
2025	23.542.575.685	17.143.917.307	0,27	6.398.658.377,52	4,09428E+19
2026	24.837.417.347	17.097.543.707	0,31	7.739.873.640,16	5,99056E+19
EROR			24,25%	5.507.355.048,85	5.703.532.731,44

TABLE 10. Table of MAPE calculation results FOB Export Value Random Forest Model

EKSPOR					
TAHUN	**TARGET/ACTUAL**	**FORECAST**	**MAPE**	**MAE**	**RMSE**
2022	19.115.502.087	14971005273	0,22	4.144.496.813,82	1,71769E+19
2023	20.166.854.702	14971005273	0,26	5.195.849.428,60	2,69969E+19
2024	21.276.031.710	15227383948	0,28	6.048.647.762,18	3,65861E+19
2025	22.446.213.454	15535443169	0,31	6.910.770.285,24	4,77587E+19
2026	23.680.755.194	15741342840	0,34	7.939.412.354,23	6,30343E+19
EROR			28,04%	6.047.835.328,81	6.189.553.446,59

Table 7 presents the error results for forecasting import volumes using the Random Forest model. The MAPE values range from 0,21 in 2022 to 0,27 in 2026, with an overall MAPE of 24,99%. The MAE values range from 4.262.434.360,83 in 2022 to 6.537.336.816,85 in 2026, with an overall MAE of 5.577.307.442,85. The RMSE values range from 1,81683E+19 in 2022 to 4,27368E+19 in 2026, with an overall RMSE of 5.628.454.584,62. This indicates that the Random Forest model shows a higher average error than the SVM. The increase in error percentages over time suggests that while the model is reasonably good at predicting short-term trends, its performance deteriorates for longer-term forecasts. The fluctuations in error rates highlight the need for additional tuning or incorporating more sophisticated features to improve accuracy. Table 8 shows the Error results for export volume forecasts using the Random Forest model. The MAPE values range from 0,26 in 2022 to 0,26 in 2026, with an overall MAPE of 18,32%. The MAE values range from 2.681.755.009,97 in 2022 to 3.308.391.699,26 in 2026, with an overall MAE of 2.090.491.924,57. The RMSE values range from 7,19181E+18 in 2022 to 1,09455E+19 in 2026, with an overall RMSE of 2.340.076.127,32. The higher error percentages compared to the SVM model suggest that the Random Forest model may not capture export volume trends as effectively, potentially due to its handling of seasonal or other complex factors influencing export volumes.

Table 9 details the error results for import FOB values the Random Forest model predicted. The MAPE values range from 0,17 in 2022 to 0,31 in 2026, with an overall MAPE of 24,25%. The MAE values range from 3.439.100.016,02 in 2022 to 7.739.873.640,16 in 2026, with an overall MAE of 5.507.355.048,85. The RMSE values range from 1,18274E+19 in 2022 to 5,99056E+19 in 2026, with an overall RMSE of 5.703.532.731,44. This higher error indicates that forecasting FOB values using the Random Forest model is more challenging, with considerable errors observed throughout the forecast period. Table 10 provides the Eror results for export FOB values using the Random Forest model. The MAPE values range from 0,22 in 2022 to 0,34 in 2026, with an overall MAPE of 28,04%. The MAE values range from 4.144.496.813,82 in 2022 to .939.412.354,23 in 2026, with an overall MAE of 6.047.835.328,81. The RMSE values range from 1,71769E+19 in 2022 to 6,30343E+19 in 2026, with an overall RMSE of 6.189.553.446,59. The higher error rates suggest that the Random Forest model may not be well-suited for capturing the nuances of FOB value fluctuations, possibly due to insufficient handling of market price variability or other external factors.

Overall, the higher error values for ekspor impor forecasts forecasting aspects reveal that while the SVM model is effective to some extent, there are noticeable areas where it falls short compared to the Random Forest model. The higher error values for ekspor impor forecasts indicate a need for further refinement in the model or consideration of alternative forecasting methods to enhance accuracy and reliability in long-term predictions.

Comparison of MAPE SVM and Random Forest

Eror is a critical metric for evaluating the forecasting accuracy of different models. It quantifies prediction error as a percentage of the actual value, with smaller values indicating better accuracy. This comparison highlights the performance of the Support Vector Machine (SVM) and Random Forest models in forecasting import and export volumes and FOB (Free on Board) values.

TABLE 11. Eror Volume Impor

Metode Peramalan	MAPE	MAE	RMSE
SVM	16,93%	3.761.961.548,45	3.771.956.419,30
Random Forest	24,99%	5.577.307.442,85	5.628.454.584,62

TABLE 12. Eror Volume Ekspor

Metode Peramalan	MAPE	MAE	RMSE
SVM	15,37%	1.808.744.500,31	1.998.275.686,90
Random Forest	18,32%	2.090.491.924,57	2.340.076.127,32

TABLE 13. Eror FOB Impor

Metode Peramalan	MAPE	MAE	RMSE
SVM	17,80%	3.864.360.844,85	4.240.747.862,67
Random Forest	24,25%	5.507.355.048,85	5.703.532.731,44

TABLE 14. Eror FOB Ekspor

Metode Peramalan	MAPE	MAE	RMSE
SVM	15,21%	3.148.375.977,52	3.430.682.425,33
Random Forest	28,04%	6.047.835.328,81	6.189.553.446,59

Table 11 summarises the eror results for the import forecasts. For import volumes, the SVM model achieved a MAPE of 16,93%, MAE of 3.761.961.548,45 and RMSE 1.998.275.686,90, which is lower than the MAPE of 18,32%, MAE of 2.090.491.924,57 and RMSE 2.340.076.127,32 obtained using the Random Forest model. This suggests that the SVM model provides more accurate predictions for import volumes. In the performance analysis, MAPE, MAE, and RMSE values for the Support Vector Machine (SVM) model are lower than those for the Random Forest model. This indicates that the SVM model can make more accurate predictions with minor errors. The reduction in these values reflects the efficiency of the SVM model in handling the given data. Therefore, SVM demonstrates better performance in this context. In conclusion, the SVM model is recommended for applications requiring high accuracy. From the comparison, the SVM model consistently outperforms the Random Forest model in predicting import and export volumes, as evidenced by its lower MAPE values. This suggests the SVM model is generally better at capturing trends and making more accurate volume forecasts. Overall, the choice of model may depend on the specific forecasting requirements and the data characteristics. SVM is superior to Random Forest in situations where non-linear patterns dominate, the dataset is smaller, or there are many relevant features to predict. However, the performance of each model heavily depends on the specific dataset used, so it is always wise to compare their performance using metrics like MAPE, MAE, and RMSE before concluding. Further refinement and exploration of additional features or hybrid approaches could enhance the accuracy of both models.

CONCLUSION

This study concludes that the SVM (Support Vector Machine) model outperforms the Random Forest model for predicting the volume and value of imports and exports in the coming years. The SVM (Support Vector Machine) model demonstrates superior accuracy, as evidenced by its lower error values in forecasting import and export volumes and FOB values. This finding underscores the effectiveness of SVM in capturing complex patterns in the data compared to random Forest. The application of Param Grid techniques has enhanced our understanding of the influence of input features on model predictions, offering valuable insights into optimising model performance. By systematically exploring various parameter combinations, the Param Grid method has contributed significantly to improving the accuracy of forecasts, aligning with the growing recognition of the importance of precise model selection for reliable estimates.

Looking ahead, future research could benefit from several avenues of exploration. First, further refinement of the SVM model could be pursued to enhance its predictive capabilities, possibly by incorporating additional features or employing advanced tuning techniques. Second, expanding the analysis to include a broader range of import and export categories may provide a more comprehensive understanding of forecasting dynamics across different sectors. The third addition of variables is intended to compare error results to achieve higher accuracy. Lastly, applying the Param Grid method to other forecasting models could offer comparative insights and validate its effectiveness across diverse contexts. This study provides practical insights for policymakers, government agencies, and customs officials. By leveraging the strengths of the Random Forest model and exploring further refinements and applications, stakeholders can make more informed decisions and develop strategies to optimise trade operations and economic planning.

ACKNOWLEDGMENTS

We are grateful for the support and funding from Sebelas Maret University through the Institute for Research and Community Service, based on contract number 194.2/UN27.22/PT.01.03/2024.

REFERENCE

1. Chalid DA, Cokrodiharjo VR. Peramalan Harga Saham Menggunakan Metode Support Vector Machine (SVM). Jurnal Pasar Modal dan Bisnis. 2021 Feb 25;3(1):61–74.
2. Pramana MW, Purnamasari I, Prangga S, Universitas), Kuaro MJ, Kelua G, et al. Peramalan Data Ekspor Nonmigas Provinsi Kalimantan Timur menggunakan Metode Weighted Fuzzy Time Series Lee [Internet]. Vol. 14. 2021. Available from: www.unipasby.ac.id
3. Chalid DA, Cokrodiharjo VR. Peramalan Harga Saham Menggunakan Metode Support Vector Machine (SVM). Jurnal Pasar Modal dan Bisnis. 2021 Feb 25;3(1):61–74.
4. Ahwan A. Forecasting of Export and Import Value of Industrial in Indonesia based on Monte Carlo Algorithm. Vol. 176, International Journal of Computer Applications. 2020.
5. Hauptmann A, Schmerer HJ, Schwanebeck B. Plant-level adjustments to imports and exports at the extensive margin. China Economic Review. 2024 Oct 1;87.
6. Dave E, Leonardo A, Jeanice M, Hanafiah N. Forecasting Indonesia Exports using a Hybrid Model ARIMA-LSTM. In: Procedia Computer Science. Elsevier B.V.; 2021. p. 480–7.
7. Hauptmann A, Schmerer HJ, Schwanebeck B. Plant-level adjustments to imports and exports at the extensive margin. China Economic Review. 2024 Oct 1;87.
8. Purwoko CFF, Sediono S, Saifudin T, Mardianto MFF. Prediksi Harga Ekspor Non Migas di Indonesia Berdasarkan Metode Estimator Deret Fourier dan Support Vector Regression. Inferensi. 2023 Mar 29;6(1):45.
9. Yolanda Agatha S, Nurkhamid M, Tipe Madya Pabean Juanda K, Keuangan Negara STAN P, Sektor B V, Selatan T. HIT RATE PEMERIKSAAN BARANG BAWAAN PENUMPANG: PENDEKATAN DATA MINING. Jurnal Perspektif Bea dan Cukai. 2022;6(1):2022.
10. Mei J, He D, Harley R, Habetler T, Qu G. A random forest method for real-time price forecasting in New York electricity market. In: IEEE Power and Energy Society General Meeting. IEEE Computer Society; 2014.
11. Rum IA, Tukker A, de Koning A, Yusuf AA. Impact assessment of the EU import ban on Indonesian palm oil: Using environmental extended multi-scale MRIO. Science of the Total Environment. 2022 Dec 20;853.
12. Petropoulos F, Apiletti D, Assimakopoulos V, Babai MZ, Barrow DK, Ben Taieb S, et al. Forecasting: theory and practice. Vol. 38, International Journal of Forecasting. Elsevier B.V.; 2022. p. 705–871.
13. Ameri R, Hsu CC, Band SS, Zamani M, Shu CM, Khorsandroo S. Forecasting PM 2.5 concentration based on integrating of CEEMDAN decomposition method with SVM and LSTM. Ecotoxicol Environ Saf. 2023 Nov 1;266.

14. Joachims T. Making large scale SVM learning practical. 1999; Available from: http://www-ai.informatik.uni-dortmund.de/ls8-reports.html
15. Sinulingga MA, Asih D, Maruddani I, Hoyyi A, Statistika D, Sains F, et al. VECTOR AUTOREGRESSIVE STABILITY CONDITION CHECK UNTUK PEMODELAN DAN PREDIKSI SUMBER PENERIMAAN PABEAN BELAWAN. JURNAL GAUSSIAN [Internet]. 2020;9(2):193–203. Available from: https://ejournal3.undip.ac.id/index.php/gaussian/
16. Gatera A, Kuradusenge M, Bajpai G, Mikeka C, Shrivastava S. Comparison of random forest and support vector machine regression models for forecasting road accidents. Sci Afr. 2023 Sep 1;21.
17. Gatera A, Kuradusenge M, Bajpai G, Mikeka C, Shrivastava S. Comparison of random forest and support vector machine regression models for forecasting road accidents. Sci Afr. 2023 Sep 1;21.
18. Ren Y, Zhang L, Suganthan PN. Ensemble Classification and Regression-Recent Developments, Applications and Future Directions [Review Article]. Vol. 11, IEEE Computational Intelligence Magazine. Institute of Electrical and Electronics Engineers Inc.; 2016. p. 41–53.
19. RAMADHAN MM, SITANGGANG IS, NASUTION FR, GHIFARI A. Parameter Tuning in Random Forest Based on Grid Search Method for Gender Classification Based on Voice Frequency. DEStech Transactions on Computer Science and Engineering. 2017 Oct 10;(cece).
20. Semarang. PROCEEDING SEMINAR NASIONAL SCIENCE AND ENGINEERING 2016 “ Quantum Technology for Better Future” Eco-friendly Manufacturing, Design and Construction Methods. 2016.
21. FORECASTING KEBERANGKATAN PENUMPANGDOMESTIKDIBANDARA NGURAH RAI BALI PERIODE 2014 – 2018DENGAN METODE ARIMA.
22. Soewignjo S, Septia Sari NWW, Mediani AP, Kamil MAZ, Amelia D, Chamidah N. CHINESE YUAN EXCHANGE RATE AGAINST THE INDONESIAN RUPIAH PREDICTION USING SUPPORT VECTOR REGRESSION. BAREKENG: Jurnal Ilmu Matematika dan Terapan [Internet]. 2024 Aug 2;18(3):1683–94. Available from: https://ojs3.unpatti.ac.id/index.php/barekeng/article/view/12217
23. Verleysen Michel. 12th European Symposium on Artificial Neural Networks, ESANN 2004 : Bruges, Belgium, April 28-30, 2004. D-side; 2004. 575 p.
24. Kecman V. Support Vector Machines – An Introduction. In 2005. p. 1–47.
25. Fitri E, Riana D. ANALISA PERBANDINGAN MODEL PREDICTION DALAM PREDIKSI HARGA SAHAM MENGGUNAKAN METODE LINEAR REGRESSION, RANDOM FOREST REGRESSION DAN MULTILAYER PERCEPTRON. METHOMIKA Jurnal Manajemen Informatika dan Komputerisasi Akuntansi [Internet]. 2022 Apr 30;6(1):69–78. Available from: https://ejurnal.methodist.ac.id/index.php/methomika/article/view/926
26. Rute M, Listrik Heterogen K. Asumsi dan Notasi HEVRP PERNYATAAN MASALAH.