International Journal of Technology 12(5) 1048-1057 (2021)
Received August 2021 / Revised October 2021 / Accepted November 2021

International Journal of Technology

http://ijtech.eng.ui.ac.id

Ship Energy Efficiency Management Plan Development using Machine Learning: Case Study of CO_2 Emissions of Ship Activities at Container Port

Ivan Dewanda Dawangi[1], Muhammad Arif Budiyanto[1*]

[1]*Department of Mechanical Engineering, Faculty of Engineering, Universitas Indonesia, Kampus UI Depok, Depok 16424, Indonesia*

Abstract. Ship energy management has an effect both on cost efficiency and the environment due to the huge amount of CO_2 emission caused by ship activities. Meanwhile, research regarding efforts to save energy consumption in the container terminal area is scarce. This paper aims to estimate CO_2 emissions from ship activities in the container port. The influential variable of CO_2 emissions is in consideration to Ship Energy Efficiency Management Plan (SEEMP). The estimation of $C0_2$ emission starts from the ship activities when the ship approaches the port, which includes ship maneuvering and ship berthing. Ship's energy consumption and CO_2 emission were analyzed using random forest regression (RF) at the default setting, and then the effectiveness was verified using k-folds cross-validation. The analysis result showed there are five influential variables to reduce the CO_2 emission: (1) main engine power; (2) auxiliary engine power; (3) waiting time in a port basin; (4) maneuvering time; and (5) berthing time. Among those five variables, maneuvering, waiting in a port basin, and berthing have the same position at the top with the same amount of weight importance from the four attribute selection training results. The random forest model training and k-folds cross-validation confirmed that the model has 98.85% of accuracy. Finally, a fuel-efficient operation is discussed, and it can be concluded that by combining several voyage optimizations with a skilled operator and cold ironing when available, it is possible to reduce the CO_2 emission by 20%. The findings and proposed plan in this paper can become a reference to develop Ship Energy Efficiency Management Plan.

Keywords: CO_2 emission; Machine learning; Random forest regression; SEEMP; Ship maneuvering

1. Introduction

Ship energy management has an effect not only on cost efficiency but also on the environment due to the huge amount of CO_2 emission caused by ship activities. Reducing energy consumption has direct impacts on emissions, minimizes the environmental effect, and reduces operational costs. CO_2 emissions require reduction to improve port air quality, making emissions factor descriptions necessary (Budiyanto et al., 2019). Unfortunately, research regarding efforts to save energy consumption in the container terminal area is still rarely conducted (Budiyanto and Shinoda, 2020). While it requires development in energy management and operational planning in real time, most research does not discuss the relationship between working time, idle time, and energy consumption of each piece of

*Corresponding author's email: arif@eng.ui.ac.id; ivandewanda@rocketmail.com, Tel.: +62-8813872264
doi: 10.14716/ijtech.v12i5.5183

equipment in depth (Iris and Lam, 2019). The study about energy consumption in the container port was carried using using modality movement which is provide a consistent estimation for carbon emission (Huzaifi et al., 2020). Another aspect that contributed to the CO_2 emission is the layout of the container terminal tends to have different results between perpendicular and parallel layouts (Budiyanto et al., 2021).

Ship Energy Efficiency Management Plan (SEEMP) is a ship energy usage plan used as a milestone for energy efficiency development by ship owners and should reflect efforts to improve the ship's energy efficiency through four steps: planning, implementation, monitoring, evaluation, and improvement. Many companies have already applied environmental management systems under ISO 14001, which contain procedures for selecting the best measure for a particular vessel and then set goals for parameter measurement, control features, and relevant feedback (Witten et al., 2011). Noon data reports, which provide information on ship's fuel consumption, speed, and weather condition, are the traditional estimation method used for SEEMP development. Unfortunately, it is not fully utilized by shipping companies and is collected only for regulation due to limitations in other ships' information and the need to collect all the other ships' information to be systematically analyzed (Beşikçi et al., 2016).

Among many practices available, the fuel-efficient operation is chosen as the focus of SEEMP development in this paper as it is easier to be adapted. Many factors can be considered for fuel-efficient operation such as improved voyage planning, which can be achieved using the help of different software tools, weather routing for specific routes and trade areas, just in time, or good communication with the port in order to give maximum notice of berth availability and facilitate the use of optimum speed to maximize efficiency and minimize delay, speed optimization while coordinating with the availability of loading or discharge berths, and optimized shaft power using the automated engine management system to control speed. There are other types of factor as well that could contribute to SEEMP and needs further research, such as optimized ship handling (trim, ballast, rudder, hull, and propulsion system), fuel type, and even the ship's operational life service (MEPC, nd).

Machine learning is a method capable of analyzing the performance of existing systems, then determining the condition and improving the system performance to be more accurate. It could re-program computers when introduced to new information based on the initial learning strategies. The use of machine learning to develop green shipping industry has been done several times, including in determining the ideal ship fuel consumption for fuel efficiency and environmental preservation by conducting various training and tests on the ship's machinery data to create an estimated model containing predictions that can be compared with actual data to produce the best model for ship fuel consumption optimization (Singh and Dhiman, 2021).

Previously, machine learning was used to predict the energy consumption needed by ships using gradient boosting regression (GBR), random forest (RF), BP network (BP), linear regression (LR), and k-nearest neighbor regression (KNN) processes, but the discussion of the strategy regarding ship energy consumption reduction is limited to the efficiency of port facilities and the arrival time of ships (Uyanık et al., 2020). Calculations on CO_2 emissions and their distribution at ports are important to be investigated for SEEMP development (Peng et al., 2020). This research aims at helping to create decisions related to SEEMP development for ship owners and port authorities, which benefits green shipping activities, hence reducing CO_2 emission and increasing shipping efficiency by looking at not only the total emissions but also the amount of CO_2 emissions during the process of maneuvering, in the port basin, and berthing to make an analysis using machine learning models and a validation to replace the traditional estimation method, especially for container ports, which have not been widely researched.

2. Methods

Three port operational modes where a ship uses its engine and releases emission consist of port basin, maneuvering, and berthing (Bergovist and Monios, 2018). The data characteristics chosen from January 2019 to December 2020 of container port ship activity are recorded. This research then uses Rapid Miner Studio version 9.9, which was developed by RapidMiner company, to execute both data preprocessing, machine learning model training, and validation.

2.1. Ship's Data Analysis

The ship's data were acquired from previous research regarding the ship's CO_2 emission in various ports, using ship's AIS data provided by container port located in East Borneo, Indonesia which is important for the region specifically and the nation generally due to its capability of handling cargo growth and the facility limitation of the local port. The port's raw data consist of months, ship name, berthing and departure time, throughput, maneuvering time, in port basin time, berthing time, gross tonnage, main engine power, auxiliary engine power, emission during maneuvering, in a port basin, and berthing also with the total amount of emission each ship produced at the port. The data characteristics chosen as free variables to find the variable importance weight are as follows:

1) Main engine power
2) Auxiliary engine power
3) Time in the port basin
4) Time during maneuvering
5) Berthing time

The variables above were chosen to prevent multicollinearity between each variable, which could cause unreliability due to coefficients becoming significant or non-significant. Bottom-up is one of the methods able to calculate the ship's emission gas. Bottom-up is chosen due to higher spatial and temporal resolution from the ship's dynamic analysis (Tsagrisa and Pandis, 2021). The equation required to calculate a ship's emission is based on ICF International.

$$E_{total} = E_{ME} + E_{AE} \quad (1)$$
$$E_{ME} = TxP_{ME}xLF_{ME}xEF \quad (2)$$
$$E_{AE} = TxP_{AE}xLF_{AE}xEF \quad (3)$$

Total emission (E_{total}) of a ship consists of main engine emission (E_{ME}) and auxiliary engine emission (E_{AE}). The data needed in order to calculate emission from each engine are time (T) in hours, engine power (P), load factor (LF), and emission factor (EF). The CO_2 emission factor chosen is 683 g/kWh in the assumption that the fuel type used is medium-speed diesel (IPPC Tier 1). The ship load factors are divided into three during maneuvering, in the port basin, and berthing for each main engine and auxiliary engine (Weng et al., 2020).

Table 1 Ship load factor

Load Factor	Main Engine	Auxiliary Engine
Maneuvering	0.2	0.5
In port basin	0.2	0.5
Berthing	0	0.4

The load factors percentage for the main engine and auxiliary engine chosen in this research for all situations is listed in Table 1 and based on the reference obtained from an interview with the port's sustainability manager and harbormaster (Styhre, et al., 2017).

2.2. Machine Learning Model and Validation

Research utilizing gradient boosting regression (GBR), random forest (RF), BP network (BP), linear regression (LR), and k-nearest neighbor regression (KNN) model found that the random forest model has the best performance compared to the other model. A random forest model is created in this research using RapidMiner Studio library with default setting used to train the models and estimate the importance factor of ship's emission. After the estimation is complete, k-folds cross-validation is needed in order to confirm the accuracy of the models created.

Random forest (RF) regression (Peng et al., 2020) is a parallel ensemble learning algorithm. Based on the decision tree-based learner, RF further introduces random attribute selection in the training process of the decision tree. Empirically, it works well for high-dimensional data. It is a compelling choice since it does not have a selection bias when creating a decision tree. The trees within are less correlated with each other compared to other tree-based algorithms (Breiman et al., 1984). Random forest is capable of measuring variable importance by examining non-linear relations between the object and explanatory variables (Mohana et al., 2021).

RF has merit compared to other machine learning methods due to it needing fewer tuning hyperparameters. Crucial hyperparameters in RF are mtry, which represents the number of variables at each split; ntree, representing the number of trees to grow; nodesize, representing the minimum size of terminal nodes; and sampsize, representing the sample size to draw. The ntree parameter is kept at the default setting of 500 trees considering past research (Mengxiao et al., 2021).

K-fold cross-validation is used to validate the accuracy of models mentioned before. It is usually advocated to measure the prediction error of a learner (Genuer et al., 2008). The dataset is split into k subsets before the validation experiment with the amount of k could begin with the value of k commonly set as 10. In each experiment, the union of k-1 subsets is split between training sets, while the rest subset is used as the test set. Finally, the mean value of all the experiments can be returned (Peng et al., 2020).

3. Results and Discussion

3.1. Random Forest Model Result and Variable Importance

This section presents and discusses the random forest training model result and the variable importance from RapidMiner Studio analysis; then, it discusses the possible plan for ship energy management in the port area. All ships using case study from container port from 2019 to 2020 are divided into two sets based on the total emission they emit at the port; those lower than the average total emission of 10,788 kg CO_2 are labeled as below average, while ships with total emission more than 10,788 are labeled as above average. The k-folds cross-validation model is also presented. The importance variable result will be chosen from the model result to be discussed to create an efficient energy management plan.

3.1.1. Variable importance of ship emission

The result of attribute selection model training to find the variable importance weight is shown, while the interface for the process section of RapidMiner Studio and training model can be seen in Figure 1. It was found that three data characteristics (maneuvering, in port basin, and berthing time) all have the same importance weight when measured by correlation, information gain, and information gain ratio, as shown in Figure 2, while both main engine and auxiliary engine power hold lower importance to the total emission category.

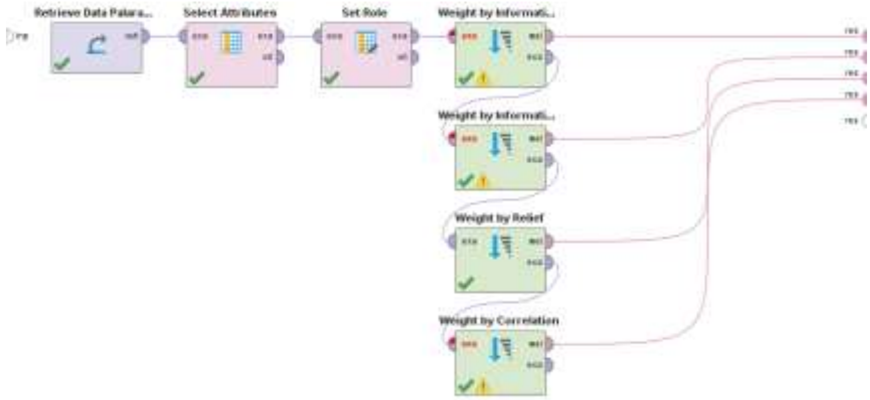


Figure 1 Attribute selection process

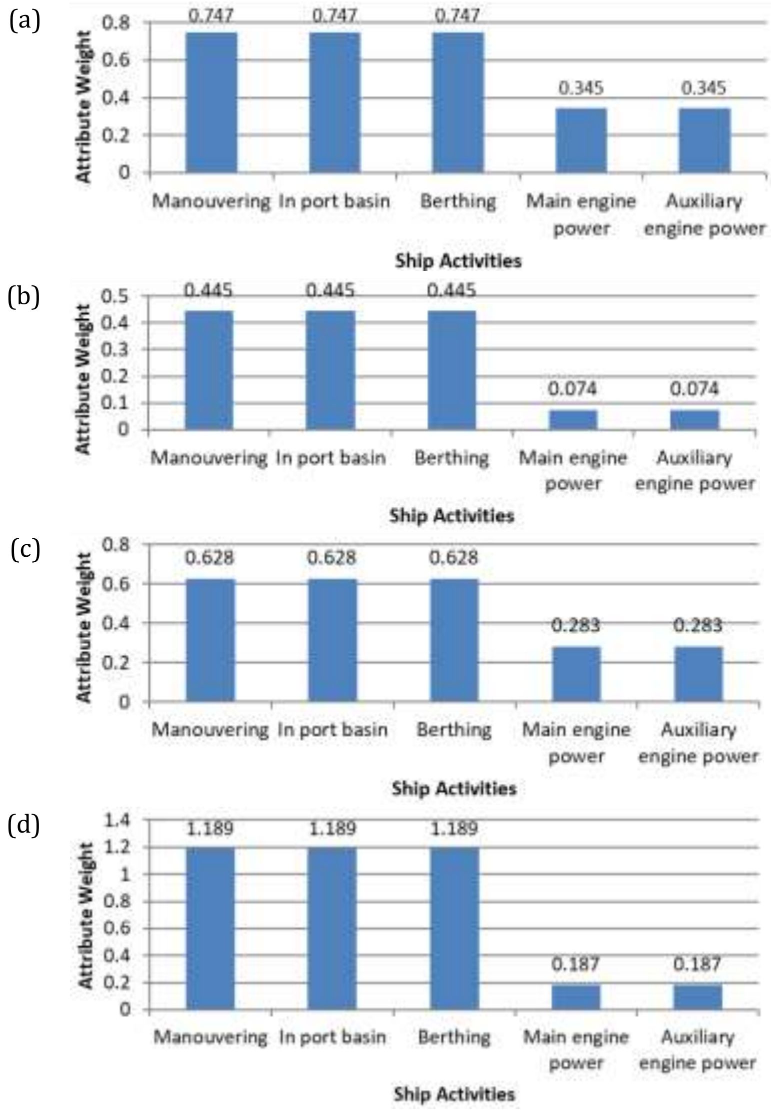


Figure 2 Attribute weight based on: (a) correlation; (b) information gain; (c) information gain ratio; and (d) relief

It can be inferred from above results that the amount of time spent during each operations holds more importance than the engine power itself. The reason for this might lie in the switching of engine load input being the same for every ship; it becomes the constraint to each engine while operation time does not have such constraint. Time spent for maneuvering and waiting in the port basin is affected by the port congestion and weather condition at the time, while berthing time is affected by not only the type and amount of cargo each ship carries but also by the port's work capabilities in cargo handling. An energy management plan that focuses on appropriate timing and switching method of engine load factor between each process is the key to the fuel-efficient operation.

3.1.2. Random forest model and validation

This paper uses a random forest regression model to train the ships' data with 500 trees and a maximal depth of 10. A confidence vote is chosen to find the accuracy of the data; the model process inputted in the RapidMiner Studio can be seen in Figure 3. The dataset is prepared by selecting the attributes chosen to be trained; the emission category is then set for the role label. The data is then split into 70% for training and 30% for the example set before being trained with random forest regression model and applied for classification performance.

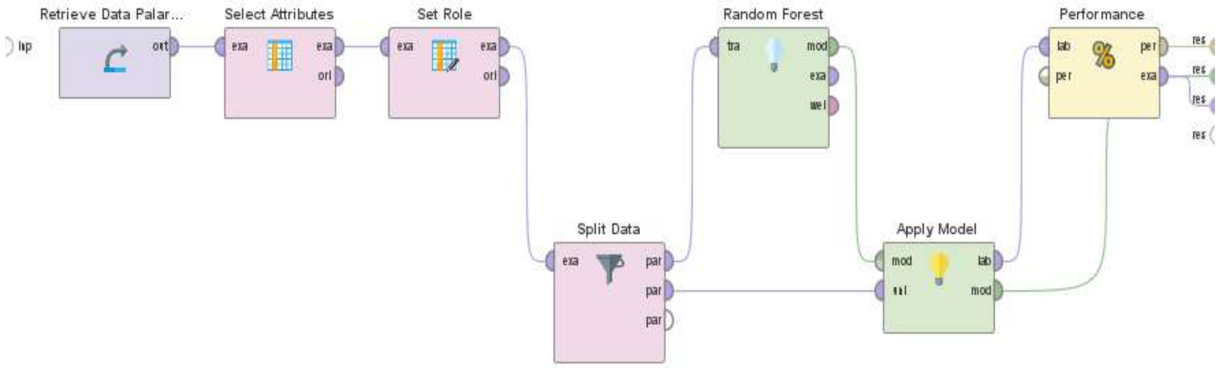


Figure 3 The random forest model process

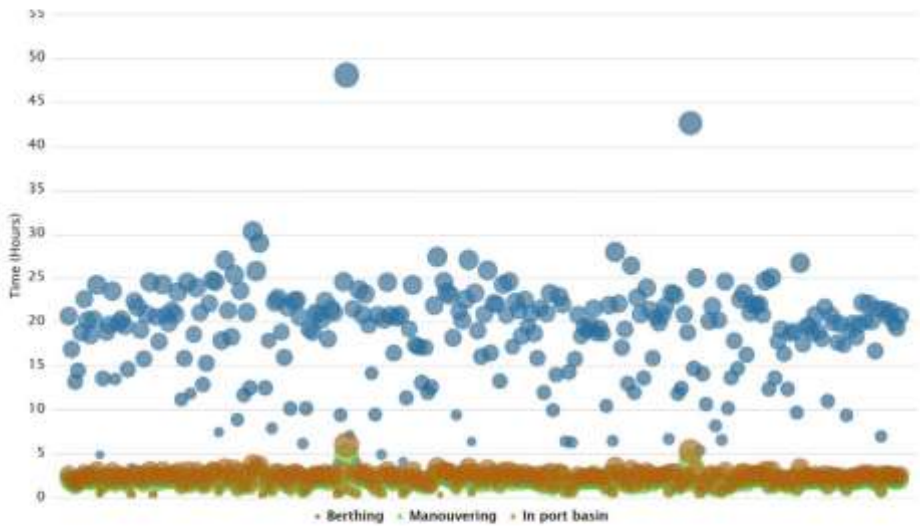
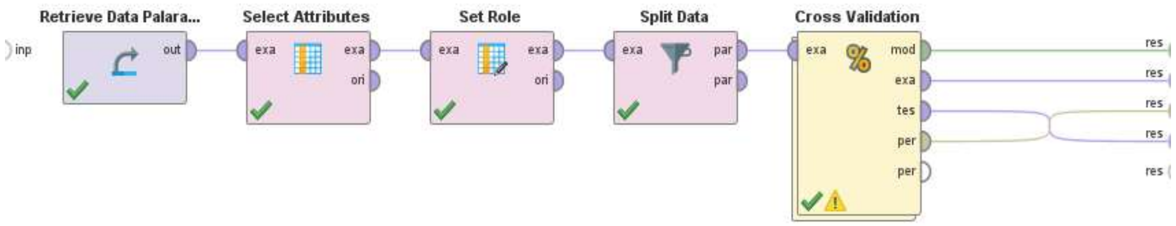
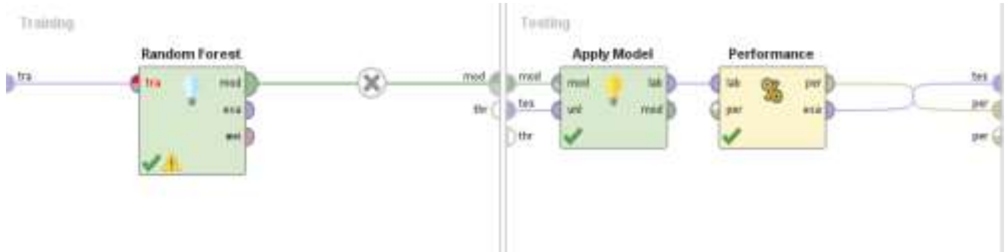


Figure 4 The ship total emission size and the amount of time spent during each operation

The random forest model training results show that the berthing process takes the longest time, which creates more CO_2 emission, as can be seen in Figure 4, with an average amount of time 18.41 hours, while the average waiting time in a port basin only took 2.3 hours and 1.84 hours for average maneuvering time. However, due to the main engine being inactive during the berthing process, the differences in average emissions created are not large, with 6.119 kgCO_2 from berthing, 2.50 kgCO_2 from waiting time, and 2.1556 kgCO_2 from maneuvering. The berthing process for container ships takes more time depending on the amount of cargo it carries, the port facilities, and work efficiency. The availability and capability of the ship's equipment for berthing, such as deck crane or even good mooring equipment, can help quicken the process along with the port's cargo handling capability.



(a)



(b)

Figure 5 The process of: (a) cross-validation; and (b) random forest model

The k-fold cross-validation use 10 folds with the process model in RapidMiner Studio, as seen in Figure 5, showed that the model has 98.85% accuracy of prediction. With the absolute error amount of 0.024 and squared error of 0.008, It can be inferred that the random forest model is proven to have a good validation, which means that the same type of dataset and model setting can be applied to predict the CO_2 emission of any year in container port.

3.2. Energy Management Plan Discussion

Based on the acquired variable importance, the appropriate SEEMP for fuel-efficient operation can be discussed. Since maneuvering in port basin and berthing time have the same importance weight, a fuel-efficient operation plan is required for each process. In order to satisfy the required Energy Efficiency Design Index (EEDI) reduction rates of 15–20% emission reduction for container ship (MARPOL, nd), some options are available to increase energy efficiency without having to do the major conversion for the ships.

3.2.1. Voyage optimization

Voyage optimization is a concept concerning the reduction of resistance under defined constraints by optimizing the ship routing, speed, scheduling, ballast, and trim. It uses safe handling, fuel consumption, seaworthiness, time spent at ports, origin to destination route and distance, and also service time written in contracts as the main constraints (Xing et al., 2020).

Combining weather routing and speed optimization to maneuver the port could help reduce the time needed hence reducing the CO_2 emission. Unlike slow steaming, speed optimization will adjust the speed needed to arrive at berth for just-in-time arrival, while weather routing creates the most efficient route instead of just reducing the speed all the way to berth, which does not seem very desirable where time hold higher importance. This will not only reduce CO_2 emission but also prevent waiting time. Communication with the port in giving the schedule and information regarding when and which berth is available will be a huge key to this plan.

Operating resistance optimization can also be done by adding sensors and using various decision support systems to reduce calm water resistance. The system could adjust the engine and ballasts to fit the required draft, trim, and steering conditions for fuel-efficient maneuvering. Voyage optimization has varying CO_2 emission reduction potential up to 10% (Xing et al., 2020); installing sensors and decision support system-based software has a promising potential to provide a strategic approach and is recommended to help reduce maneuvering time (Beşikçi et al., 2016).

3.2.2. Cold ironing

Cold ironing is the practice of providing electrical source from shore to the ship while at the berthing process. The ship could turn off the diesel generators during the berthing process, which cost the most time at the port; this would reduce the CO_2 emission by the ship itself greatly. Since cold ironing can be done with the use of low-carbon or zero-carbon energy sources such as wind and solar energy, it could even potentially reduce ship CO_2 emission at port up to 40% (Sciberas et al., 2016). The shipowner is greatly advised to comply with port authorities to be allowed for cold ironing at berth whenever available since berthing has the longest time compared to maneuvering and waiting in the port basin, significantly reducing both CO_2 emission and fuel cost.

3.2.3. Operator skill

The human factor is the non-technical factor to fuel-efficient operation. Appropriate skills and knowledge are needed by the ship crews and operators, not only to properly control the ship in every situation but also to be able to fully utilize the systems installed for fuel-efficient operation measures. Up-to-date ship operation training is important for both existing and future operators whenever any new system is to be implemented on the fleet. The shipowners also need to be able to make the right decision for ship activity and time spent at the port. It is possible to reduce CO_2 emissions by increasing energy awareness and operation capability up to 10% (Jensen et al, 2018).

3.2.4. Type of auxiliary engine

On the technical operation aspect, the trend of the used dual fuel for main engine auxiliary engine increase by the year. The use of LNG Fuel for marine diesel would increase power and efficiency (Budiyanto et al., 2020). In some cases the retrofit of the diesel engine to the dual-fuel is needed, the ISO tank for LNG is the key for the fuel management system (Pamitran et al., 2019).

4. Conclusions

This paper uses the machine learning method to predict the variable importance of CO_2 emission with RapidMiner Studio as the main tool to create a data model and analysis. A random forest model is created using five data characteristics as free variables. After running the model, variable importance is then selected and analyzed along with the effects of the selected variables on the ship's CO_2 emission. Finally, the SEEMP is developed based on the discussed result. Several options to minimize the CO_2 emission during maneuvering, in the port basin, and berthing are discussed, using the results of the training model. While cold ironing has a huge CO_2 emission reduction port, not all ports can provide it, so a combination between voyage optimization methods and trained operators is crucial for achieving the required EEDI reduction rates. Further research to minimize the CO_2 emission by analyzing the port's facilities, such as cold ironing feasibility and work efficiency improvement, need to be done, as cooperation with the port holds the crucial part during berthing processes, such as cold ironing availability and information exchange, to achieve just-in-time arrival. The method proposed by this paper can be used to develop fuel-efficient operations at ports for SEEMP development, especially container ports in developing countries such as Indonesia, where environmental health is not the top priority.

Acknowledgements

This study is supported by funding for Basic Research from the Ministry of Research and Technology/National Research and Innovation Agency of the Republic of Indonesia for Fiscal Year 2021. Number: NKB-035/UN2.RST/HKP.05.00/2021.

References

Bergovist, R., Monios, J., 2018. *Green Ports: Inland and Seaside Sustainable Transportation Strategies.* Elsevier

Breiman, L., Friedman, J., Olsen, R., Stone, C., 1984. *Classification and Regression Trees.* Wadworth

Budiyanto, M.A., Shinoda, T., 2020. Energy Efficiency on the Reefer Container Storage Yard: An Analysis of Thermal Performance of Installation Roof Shade. *Energy Reports,* Volume 6(Supplement 2), pp. 686–692

Budiyanto, M.A., Huzaifi, M.H., Sirait, S.J., 2019. Estimating of CO_2 Emission in a Cotainer Port based on Modality Movement in the Terminal Area. *International Journal of Technology,* Volume 10(8), pp. 1618–1625

Budiyanto, M.A., Huzaifi, M.H., Sirait, S.J., Prayoga, P.H.N., 2021. Evaluation of CO_2 Emissions and Energy use with Different Container Terminal Layouts. *Scientific Reports*, Volume 11, pp. 1–14

Budiyanto, M.A., Pamitran, A.S., Wibowo, H.T., Murtado, F.N., 2020. Study on the Performance Analysis of Dual Fuel Engines on the Medium Speed Diesel Engine. *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences*, Volume 68(1), pp. 163–174

Beşikçi, E.B., Arslan, O., Turan, O., Ölçerc, A.L., 2016. An Artificial Neural Network Based Decision Support System for Energy Efficient Ship Operations. *Computers & Operations Research,* Volume 66, pp. 393–401

Genuer, R., Poggi, J., Tuleau, C., 2008. Random Forests: Some Methodological Insights. *Rapport de recherche,* pp. 1–35

Huzaifi, M.H., Budiyanto, M.A., Sirait, S.J., 2020. Study on the Carbon Emission Evaluation in a Container Port based on Energy Consumption Data. *Evergreen*, Volume 7(1), pp. 97–103

Iris, C., Lam, J.S.L., 2019. A Review of Energy Efficiency in Ports: Operational Strategies, Technologies and Energy Management Systems. *Renewable and Sustainable Energy Reviews,* Volume 112, pp. 170–182

MARPOL, nd. Annex VI Chapter 4 Regulatio 21.6: Review of Phases and Reduction Rates

Mengxiao, L., Hu, S., Ge, Y., Heuvelink, G.B.M., Ren, Z., Huang, X., 2021. Using Multiple Linear Regression and Random Forests to Identify Spatial Poverty Determinants in Rural China. *Spatial Statistics,* Volume 42, https://doi.org/10.1016/j.spasta.2020.100461

MEPC, nd. 70/18/Add.1 Annex 10, 2016 Guidelines for the Development of a Ship Energy Efficiency Management Plan (SEEMP)

Pamitran, A.S., Budiyanto, M.A., Maynardi, R.D.Y., 2019. Analysis of ISO-tank Wall Physical Exergy Characteristic – Case Study of LNG Boil-Off Rate from Retrofitted Dual Fuel Engine Conversion. *Evergreen*, Volume 6(2), pp. 134–142

Peng, Y., Liu, H., Li, X., Huang, J., Wang, W., 2020. Machine Learning Method for Energy Consumption Prediction of Ships in Port Considering Green Ports. *Journal of Cleaner Productions,* Volume 264, https://doi.org/10.1016/j.jclepro.2020.121564

Mohana, R.M., Reddy, C.K.K., Anisha, P.R., Murthy, B.V.R., 2021. Random Forest Algorithms for the Classification of Tree-based Ensemble. *Materials Today: Proceedings*, https://doi.org/10.1016/j.matpr.2021.01.788

Sciberas, E.A., Zahawi, B., Atkinson, D.J., Juandó, A., Sarasquete, A., 2016. Cold Ironing and Onshore Generation for Airborne Emission Reductions in Ports. *In:* Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment, pp. 67–82

Jensen, S., Lützen, M., Mikkelsen, L.L., Rasmussen, H.B., Pedersen, P.V., Schambya, P., 2018. Energy-Efficient Operational Training in a Ship Bridge Simulator. *Journal of Cleaner Production,* Volume 171, pp. 175–183

Singh, J., Dhiman, G., 2021. A Survey on Machine-Learning Approaches: Theory and Their Concepts. *Materials Today: Proceedings*, https://doi.org/10.1016/j.matpr.2021.05.335

Styhre, L., Winnes, H., Black, J., Lee, J., Le-Griffin, H., 2017. Greenhouse Gas Emissions from Ships in Ports – Case Studies in Four Continents. *Transportation Research Part D: Transport and Environment*, Volume 54, pp. 212–224

Tsagrisa, M., Pandis, N., 2021. Multicollinearity. *American Journal of Orthodontics and Dentofacial Orthopedics,* Volume 159(5), pp. 695–696

Uyanık, T., Karatuğ, Ç., Arslanoğlu, Y., 2020. Machine Learning Approach to Ship Fuel Consumption: A Case of Container Vessel. *Transportation Research Part D: Transport and Environment*, Volume 84, https://doi.org/10.1016/j.trd.2020.102389

Weng, J., Shi, K., Gan, X., Li, G., Huang, X., 2020. Ship Emission Estimation with High Spatial-Temporal Resolution in the Yangtze River Estuary using AIS Data. *Journal of Cleaner Production,* Volume 248, https://doi.org/10.1016/j.jclepro.2019.119297

Witten, I.H., Frank, E., Hall, M.A., 2011. *Data Mining: Practical Machine Learning Tools and Techniques.* Elsevier Inc

Xing, H., Spence, S., Chen, H., 2020. A Comprehensive Review on Countermeasures for CO_2 Rmissions from Ships. *Renewable and Sustainable Energy Reviews,* Volume 134, https://doi.org/10.1016/j.rser.2020.110222