

Discovering supply chain operation towards sustainability using machine learning and DES techniques: a case study in Vietnam seafood

243

Received 13 October 2023
Revised 7 March 2024
Accepted 7 July 2024

Luan Thanh Le

*Department of Business, Greenwich Alliance, FPT University, Hanoi, Vietnam, and
Trang Xuan-Thi-Thu*

Department of Business, Swinburne Alliance, FPT University, Hanoi, Vietnam

Abstract

Purpose – To achieve the Sustainable Development Goals (SDGs) in the era of Logistics 4.0, machine learning (ML) techniques and simulations have emerged as highly optimized tools. This study examines the operational dynamics of a supply chain (SC) in Vietnam as a case study utilizing an ML simulation approach.

Design/methodology/approach – A robust fuel consumption estimation model is constructed by leveraging multiple linear regression (MLR) and artificial neural network (ANN). Subsequently, the proposed model is seamlessly integrated into a cutting-edge SC simulation framework.

Findings – This paper provides valuable insights and actionable recommendations, empowering SC practitioners to optimize operational efficiencies and fostering an avenue for further scholarly investigations and advancements in this field.

Originality/value – This study introduces a novel approach assessing sustainable SC performance by utilizing both traditional regression and ML models to estimate transportation costs, which are then inputted into the discrete event simulation (DES) model.

Keywords Sustainable supply chain, Maritime shipping, Data driven, Discrete event simulation, Container ships

Paper type Research paper

1. Introduction

Sustainable maritime shipping is one of the agreements that is highly emphasized by the United Nations Convention on the Law of the Sea (UNCLOS) (Winther *et al.*, 2020). The shipping industry has experienced a significant increase in global transport emissions, leading to a pressing need to reduce fuel consumption in maritime transportation (Li and Tang, 2017). A fuel-efficient vessel not only helps reduce operational costs but also contributes to reducing greenhouse gas emissions, which are one of the major causes of global warming. Therefore, the development of an accurate fuel consumption estimation model for container ships has become crucial for the shipping industry.

Container ships are one of the most widely used vessels in the shipping industry, making a significant contribution to global trade. However, they are also among the largest fuel consumers in the industry. As a result, reducing fuel consumption in container shipping is crucial for cost savings and environmental sustainability. Accurate fuel consumption estimation models provide the basis for developing fuel-efficient shipping operations, leading



JEL Classification — C44, C45, C53, F47, O32

© Pacific Star Group Education Foundation. Licensed re-use rights only.

This research was funded by the research program by Foreign Trade University (Grant No: FTURP02-2020-16).

Declaration of interest: The authors report no conflict of interest.

Maritime Business Review
Vol. 9 No. 3, 2024
pp. 243-262
Emerald Publishing Limited
2397-3757

DOI 10.1108/MABR-10-2023-0074

to reduced fuel consumption and environmental impact. In addition, these models can help shipping companies comply with the International Maritime Organization (IMO) regulations, which aim to reduce greenhouse gas emissions from the shipping industry.

One of the IMO's primary objectives is to reduce greenhouse gas emissions from the shipping industry. The IMO adopted an ambitious strategy to reduce greenhouse gas emissions by at least 50% by 2050 compared to 2008 levels (Wilkinson, 2012). The strategy also aims to phase out greenhouse gas emissions entirely as soon as possible in this century. To achieve these objectives, the IMO has developed a range of measures, including the Energy Efficiency Design Index (EEDI), the Ship Energy Efficiency Management Plan (SEEMP) and the Data Collection System (DCS).

Accurate fuel consumption estimation models for container ships are essential for the shipping industry. These models help reduce fuel consumption and greenhouse gas emissions, leading to cost savings and environmental sustainability. The IMO has set ambitious objectives to reduce greenhouse gas emissions from the shipping industry, and fuel consumption estimation models play a crucial role in achieving these objectives. Therefore, it is imperative for shipping companies to develop and implement fuel-efficient shipping operations based on accurate fuel consumption estimation models to comply with the IMO's regulations and contribute to global efforts to combat climate change. The fundamental aim of this manuscript is to calibrate an accurate model for estimating fuel expenses associated with container ships. This model is then implemented on a simulation platform to facilitate a comprehensive assessment of the economic viability of the supply chain (SC).

The forthcoming sections are arranged in the following manner: [Section 2](#) amalgamates a compendium of pertinent research endeavors pertaining to the fuel consumption estimation paradigm for container vessels and the discrete event simulation (DES) methodology for SC. [Section 3](#) delineates the data collection, data preprocessing, estimation model design and simulation model design. [Section 4](#) expounds on the analysis results, estimation model validation and financial key performance indicators (KPIs) of the simulation model. Lastly, the concluding section offers a comprehensive summary of the study's findings.

2. State of the art

2.1 Data driven approach in maritime shipping and SC

The data-driven approach in ocean shipping and SC management has revolutionized the industry, enabling organizations to make informed decisions based on accurate and timely data (Liu *et al.*, 2023). It allows for better route optimization, improved fleet management, and enhanced SC visibility. Real-time data analysis enables proactive decision-making, minimizing disruptions and delays while optimizing resource allocation. Additionally, data-driven insights enable the identification of patterns and trends, facilitating predictive modeling and risk assessment. Tseng *et al.* (2022) conducted a comprehensive and data-driven analysis of the sustainable food SC, with a specific focus on comparing the unique attributes and considerations of Halal and non-Halal foods. A data-driven approach for the automatic design of ship routes between two ports was developed to demonstrate the efficacy of this data-driven approach in enhancing route design efficiency and accuracy, thereby offering significant potential for cost savings and operational improvements within the maritime industry (Wen *et al.*, 2020). Nguyen *et al.* (2022) deployed a knowledge mapping related to physical Internet (PI) and digital twin (DT) in supply chain management (SCM). The study collected 518 journal articles with ten key research lines to maximize PT/DT's application in SCM. The potential of this article included the improvement of operational efficiency, visibility and collaboration among stakeholders. This review encourages further studies to develop a new integration of a simplified model into SCM.

A review of 123 articles revealed that machine learning (ML) applications in SCM were still in a nascent stage (Ni *et al.*, 2020). The analysis demonstrated a lack of prolific authors to form a cohesive research community dedicated to ML applications in SCM. 10 ML algorithms were mostly employed in SCM. NN was considered the most prevalent algorithm, accounting for 54% of the total, followed by SVM at 21%. Han and Zhang (2021) calibrated an NN model for SC risk management. Piramuthu (2005) combined ML techniques to formulate a dynamically adaptable SC framework. A design of a robust demand prediction model was conducted using ARIMAX and NN models (Feizabadi, 2022). Abbasi *et al.* (2020) utilized several ML models to make decisions regarding the transshipment of blood units within a network of hospitals.

Another review of 93 articles focused on the application of various ML algorithms across different stages of the agricultural SC (Sharma *et al.*, 2020). The findings showed that the majority of ML models were implemented in the pre-production and production stages, such as crop yield prediction, predicting soil properties, irrigation management, weather prediction, drop protection, weed detection and livestock management. Some studies implemented ML for demand forecasting. Fewer studies focused on transportation, distribution and inventory management.

Furthermore, to harness and leverage the advantages of the data-driven approach, artificial neural network (ANN) has emerged as a powerful tool in various fields, including transportation and SCM (Sharma *et al.*, 2022). Their ability to learn from and adapt to complex datasets makes them valuable for optimizing operations, improving decision-making and enhancing efficiency. In the context of maritime transportation and SC, ANN applications have shown significant potential for addressing challenges and improving performance. Bodendorf *et al.* (2022) studied the use of ANN in improving cost estimation in manufacturing SC. ANN was used to propose a cost estimation model that could foresee the total costs of manufacturing products by researching several input factors, such as material and labor costs, production volume and product complexity. The dataset of actual costs was trained and tested by using a separate database to justify its accuracy. ANN-based cost estimation showed an accurate prediction and was aligned accordingly in comparison with its traditional method. An ANN-based cost model was concluded to be a valuable tool for cost-effective management. In the study for simplified SCM, cost-savings played an important role in enabling firms to make informed decisions about pricing, production volumes and resource allocation.

2.2 ML approach in ship fuel consumption

Container ships are the backbone of the global shipping industry, and fuel consumption is a major cost component for shipping companies. To optimize fuel consumption and reduce operational costs, accurate fuel consumption models are essential. Two approaches have been proposed for predicting fuel consumption: top-down and bottom-up. In the top-down approach, fuel consumption is estimated based on the ship's overall performance, while in the bottom-up approach, fuel consumption is calculated by breaking down the ship's operations into smaller components. While both approaches have their advantages and disadvantages, recent research has shown that the bottom-up approach is more accurate and flexible. A simplified model was proposed for estimating the fuel consumption of container ships based on their voyage characteristics (Le *et al.*, 2020a, b). The models were developed using data from 100 to 143 container ships' voyage records and showed an average error rate of 10–20%. The advantage of this simplified model is that it requires fewer input data points and is computationally more efficient than traditional models. Dragović *et al.* (2018) estimated and analyzed ship exhaust emissions and externalities in two popular cruise destinations on the eastern coast of the Adriatic Sea using an activity-based approach. The results indicated an

increasing trend of air pollution in both ports over the past few years, and the study suggested that factors such as berth availability and accessibility can influence ship emissions in addition to the ship's operating characteristics.

In [section 2.1](#), we review studies applying ML to maritime shipping. Delving deeper into fuel consumption prediction models and related issues, ML continues to prove itself as an exceptionally robust method. [Yan et al. \(2021\)](#) examined 83 relevant studies spanning a 13-year period (from 2008 to 2021). They noted that a considerable proportion of over 40% of these studies employed ML models, particularly those based on black-box methodologies, for predicting fuel consumption.

Various ML models were constructed for a container ship, including ridge and LASSO regression, support vector regression (SVR), tree-based and boosting ([Uyanik et al., 2020](#)). [Li et al. \(2022\)](#) presented a technique to merge voyage report data with meteorological data and applied 11 commonly used ML models focusing on containerships ranging from 8,100 TEU to 14,000 TEU. [Yan et al. \(2020\)](#) proposed a dual-stage model to forecast and minimize fuel consumption for a dry bulk vessel. [Peng et al. \(2020\)](#) designed five ML models of gradient boosting (GRB), random forest (RF), backpropagation network (BP), linear regression (LR) and k-nearest neighbor (KNN) to estimate the energy consumption of ships in port at Jingtang Port in China. An application of long short-term memory (LSTM) NN to forecast real-time fuel consumption rates was implemented using variables such as water depth, water speed, wind speed and wind angle ([Yuan et al., 2021](#)). ANN models were designed to estimate fuel consumption for container ships in Korea ([Le et al., 2020a, b](#)). The results show that the ANN model provides accurate fuel consumption estimates with a high level of precision.

Among the studies employing ML models in the field of fuel consumption in maritime transportation mentioned above, none have demonstrated the extended application of these models within the SC. This involves utilizing simulation tools as a method to visualize the SC, aiding SC operators in emerging economies to observe the SC activities they are implementing. This is the primary objective of this research, aiming to lay the groundwork for proposing scenarios and addressing more complex problems in the future, making academic content more accessible to export businesses and thereby increasing the applicability of academic research.

2.3 Discrete event simulation

DES has become an increasingly popular tool for analyzing and optimizing SC systems due to its ability to model complex systems, simulate scenarios and analyze outcomes in a risk-free and cost-effective environment ([Legato and Mazza, 2023](#)). Over the years, several studies have employed DES to model and optimize SC systems, demonstrating its potential for enhancing SC efficiency. For instance, [Christopher and Peck \(2004\)](#) discussed the use of DES to model and evaluate SC risks, which could help SC managers develop contingency plans to mitigate risks. [Prinz et al. \(2019\)](#) demonstrated how DES could be used to model and evaluate the energy efficiency of forest chip supply systems. The study found that DES can effectively analyze complex SCs and identify opportunities to improve energy efficiency. [Moussa et al. \(2019\)](#) proposed an innovative approach to solving warehousing problems by combining ARIZ, lean management and DES. The study found that this approach could effectively identify and solve warehousing problems and improve overall performance. [Rabe et al. \(2020\)](#) explored how genetic algorithms and DES could be combined to speed up computational times in simheuristics. The study found that this approach could significantly reduce computational time while maintaining high levels of accuracy. [Amorim-Lopes et al. \(2021\)](#) demonstrated how DES could be used to improve picking performance in a large retail warehouse. The study found that combining probabilistic simulation, optimization and DES could effectively identify and solve performance problems and improve overall efficiency.

Oliveira *et al.* (2020) explored the use of DES to aid decision-making and mitigation in solid waste management. The study found that DES could be an effective tool for evaluating different waste management strategies and identifying opportunities to improve overall efficiency and sustainability. In the port logistics field, another study calibrated DES for two distinct queuing network models for quantitative evaluation of the two major logistic processes in a real container terminal for pure transshipment (Legato and Mazza, 2020).

Despite the growing body of literature on the use of DES in SCM, there is still a research gap in terms of its application to real-world situations. Many studies have focused on theoretical models and simulations, but there is a need for more empirical studies that test the effectiveness of DES in practical settings. Additionally, there is a need for more research on the integration of DES with other emerging technologies, such as blockchain and artificial intelligence, to further investigate SC operations for sustainability.

To bridge the research gap, this study provides a new approach evaluating sustainable SC performance by deploying the traditional regression and ML models of transportation cost as inputs to the DES model.

3. Conceptual framework and exploratory data analysis

3.1 Data collection

The data for this study were collected from multiple sources, with two main datasets identified: (1) the fuel consumption model data for container ships and (2) the SC simulation model data. The dataset (1) was obtained from a leading global shipping company with a branch in Vietnam, covering the period from October 2018 to December 2019. It consists of 8,197 voyages, including parameters such as vessel identification number, fuel consumption of container ships per voyage (in metric tons of fuel oil), departure port, destination port, departure time, arrival time, design speed (knot), average speed (knot), cargo weight (in metric tons) and draft (in metric tons).

The dataset comprises a diverse range of maritime transport routes, including Intra-Asia, Asia–North Europe, Asia–South America, Asia–North America, Asia–Mediterranean, Asia–Oceania, Intra-Europe, North America–South America and North Europe–North America (see Table 1). Additionally, the number of observations is adequate for running ML models, as per the review of similar studies such as the 724 observations (Uyanik *et al.*, 2020), 10,461 data rows (Bui-Duy and Vu-Thi-Minh, 2021), 645 rows (Yuksel *et al.*, 2023) and 242 entries (Yan *et al.*, 2020).

The dataset (2) was obtained from the General Department of Customs for the period from March to September 2022, specifically focusing on the catfish export volume through the Hai

Lane	Number of voyages
Intra-Asia	2,018
Asia–North Europe	1,098
Asia–South America	832
Asia–North America	1,784
Asia–Mediterranean	542
Asia–Oceania	329
Intra-Europe	539
North America–South America	416
North Europe–North America	639
<i>Total</i>	<i>8,197</i>

Source(s): Table by authors

Table 1.
Maritime shipping
lanes in shipping data

Phong, Quang Ninh and Ho Chi Minh ports. The dataset comprises 312 shipments. It includes variables such as the Bill of Lading numbers, customs declaration number, declaration time, number of containers, cargo weight (in metric tons), cargo value (in USD), transportation mode, name and address of exporting and importing companies, departure port, destination port and tax rates. The number of shipments is relatively acceptable, with an insignificant deviation compared to similar studies utilizing simulation platforms for SC, such as 443 shipments (Bui-Duy *et al.*, 2023).

3.2 Conceptual framework

The SC encompasses multiple stages with a plethora of parameters, including procurement, production, storage, inventory management, raw material management, transportation management, distribution management and last-mile delivery management. Within the scope of this study, we focus on leveraging the impact of ML models for predicting maritime transportation costs (main-leg transportation) on the financial and environmental metrics of the entire SC through DES techniques.

In maritime transportation, establishing a highly accurate model for estimating fuel consumption and consequently fuel cost is crucial. Previous studies have consistently identified speed as the most influential factor (following an exponential function) (Psaraftis and Kontovas, 2013; Wang and Peng, 2019), along with factors such as voyage distance (Meng *et al.*, 2016), vessel age (Zakerdoost and Ghassemi, 2019), displacement and draft (Bialystocki and Konovessis, 2016), load factor (Le *et al.*, 2020a, b), trim (Coraddu *et al.*, 2017) and weather conditions (Vettor and Guedes Soares, 2022). Due to the lack of data on technical specifications as well as weather conditions, and given that the main focus of this study is to explore SC dynamics rather than heavily emphasizing technical models, we employ the following variables: speed, vessel running time (derived from distance divided by speed), load factor, displacement and fuel purchase price for the maritime transportation cost model.

While an increasing number of studies have calibrated ML for fuel consumption prediction, traditional regression methods remain applicable under certain circumstances. These include situations when there is limited data or insufficient data to train an effective ML model or when a simpler model structure is preferred to reduce computational complexity and enhance interpretability. Additionally, traditional regression approaches may be favored when there is a preference for model explainability over predictive accuracy.

3.3 Data preprocessing

Outliers can pose a serious problem for statistical analyses, as they can inflate error rates, obscure true effects and lead to inaccurate conclusions. Detecting and properly handling outliers is critical for ensuring the validity and reliability of research results (Wilcox, 2016). The identification and removal of outliers is essential for improving the accuracy and usefulness of data analysis. In this study, we utilize density-based spatial clustering of applications with noise (DBSCAN) as a density-based clustering algorithm, to detect outliers. One of the main advantages of DBSCAN for outlier detection is its ability to handle datasets with varying densities and irregular shapes. Unlike traditional methods, DBSCAN does not assume any particular distribution or shape of the data, making it robust to outliers that may not conform to typical patterns. Recent studies have demonstrated the effectiveness of DBSCAN in detecting outliers, such as identifying unusual traffic patterns in transportation networks (Gonzalez *et al.*, 2011). In implementing DBSCAN, two key hyperparameters that need to be appropriately set up include the minimum number of data points required to form a cluster (k) and the maximum distance between data points in the same cluster (ϵ). To identify the optimal value of k for DBSCAN clustering, we use the k-distance graph method. The k-distance graph is a plot of the distance to the k th nearest neighbor (kNN) for each point in

the dataset, sorted in descending order. We then identify the knee point in the k-distance graph, which is the point where the slope of the graph changes significantly. This knee point represents the optimal value of k . To identify the optimal value of ϵ for DBSCAN clustering, we use the elbow method. The elbow method involves plotting the kNN distances in ascending order and then computing the ratio of the differences between consecutive distances to the distances themselves. We then choose the elbow point, which is the point at which the ratio of differences starts to level off.

Along with using DBSCAN, we continue to use the Mahalanobis technique to detect outliers due to its advantage in detecting multivariate outliers (Hadi and Simonoff, 1993). Mahalanobis distance is defined as the distance between a multivariate point x and a reference point μ , normalized by the covariance matrix Σ of the dataset:

$$D^2 = (x - \mu)^T \Sigma^{-1} (x - \mu) \tag{1}$$

where T denotes the transpose, Σ^{-1} is the inverse of the covariance matrix, and D is the Mahalanobis distance. The Mahalanobis distance measures how many standard deviations away x is from the mean of the reference distribution, taking into account the correlations between variables. If D^2 exceeds a critical value, the corresponding observation may be considered an outlier.

3.4 Feature standardization

Standardizing the data involves transforming the original data values to a new scale to reduce the impact of differences in variable magnitudes (Cao *et al.*, 1999). It has several advantages that make it a suitable choice for many applications. One of the main advantages of standardization is that it centers the data at zero and scales it to have unit variance, which makes it easy to compare the relative magnitudes of different variables (Lichstein *et al.*, 2002). Standardization is often preferred in practice because it preserves the distributional shape of the data. Standardization scales the data to have a zero mean and unit variance using the formula:

$$x_{std} = (x - mean(x)) / sd(x) \tag{2}$$

where x is the original data value, $mean(x)$ is the mean of x and $sd(x)$ is the standard deviation of x .

Table 2 provides descriptive statistics for various variables related to the fuel consumption model of container ships. The wide range of fuel consumption and cargo weight suggests that the ship may be used for a variety of operations with different requirements. The low standard deviation of ship speed indicates that the ship may operate at a relatively consistent speed. The moderate range of sailing time values suggests that the ship can be used for different routes and distances. The low standard deviation indicates that the sailing time is relatively stable and consistent across different routes and distances. The

	Min	Max	Mean	SD	VIF
EC (MT)	1.8	2,562.7	70.7	95.5	
V (Knots)	5.2	24.4	15.0	2.7	1.05
T (Days)	0.03	36.0	9.63	1.7	1.06
C (MT)	2,840	86,040	30,960	19,444.2	2.78
CW (MT)	605	51,546	12,604	7,321.1	2.71
Disp (MT)	607	133,335	29,040	25,131.3	1.25

Source(s): Table by authors

Table 2.
Descriptive statistics

high standard deviation of ship capacity and cargo weight may suggest that the ship is used for a variety of cargoes, which can affect fuel consumption and sailing time. The moderate range of displacement values suggests that the ship has a medium to large size. The mean value of displacement indicates that the ship can carry medium to large volumes of cargo. The relatively high standard deviation indicates that the ship's displacement can vary significantly depending on the type and volume of cargo. Based on the analysis of the variance inflation factor (VIF), there is no evidence of significant multicollinearity among the independent variables in the regression model.

4. Model description

4.1 Energy cost model

4.1.1 *Traditional approach.* Firstly, in this study, the fuel consumption model per voyage (EC) for container ships is developed. Subsequently, the estimated energy consumption is multiplied by the fuel oil (FO) unit price at the loading port to establish the energy expense parameter (EE) for the DES model, as in Equation (3).

$$EE_i = EC_i \times FO_i \quad (3)$$

There exists a plethora of methods for estimating fuel consumption models for container vessels in the literature. In this study, we adopt a bottom-up approach to model development, which entails decomposing the system into its constituent parts, identifying key variables and estimating their effects on fuel consumption via MLR models. This approach confers several benefits, including improved precision of estimates, ease of interpretation and the potential for targeted optimization.

$$\text{Model LM1: } EC_1 = \alpha_1 TV^3 + \alpha_2 T + \varepsilon_1 \quad (4)$$

$$\text{Model LM2: } EC_2 = \alpha_1 TV^3 + \alpha_2 T + \alpha_3 Load + \varepsilon_2 \quad (5)$$

$$\text{Model LM3: } EC_3 = \alpha_1 TV^3 + \alpha_2 T + \alpha_3 Load + \alpha_4 Disp + \varepsilon_3 \quad (6)$$

where T is sailing time (hours); V is operational speed (knots); $Load$ presents the cargo load factor (%) defined by cargo weight (CW in ton) divided by ship's capacity (C in TEU); $Disp$ is displacement (ton); α_1 , α_2 , α_3 and α_4 are coefficients and ε_1 ; ε_2 and ε_3 are estimated errors.

[Le et al. \(2020\)](#) have proposed and rigorously tested the LM1 model, which serves as a foundational building block for subsequent models in the field of container ship fuel consumption estimation. Building on this seminal work, we enrich the LM2 candidate by incorporating an additional independent variable related to the cargo load factor, thereby improving the explanatory power of the model ([Le et al., 2020a, b](#)). In turn, the LM3 candidate extends the sophistication of the LM2 model by introducing displacement variables, thus capturing the effects of this key environmental factor on fuel consumption, as indicated by [Meng et al. \(2016\)](#). By advancing the theoretical and empirical foundations of container ship fuel consumption modeling in this manner, we contribute to the ongoing efforts to develop robust and accurate models for optimizing the fuel efficiency of container shipping operations.

Several studies have examined the correlation between displacement and fuel consumption in ships. For example, [Yang et al. \(2019\)](#) found that displacement was a significant factor in predicting fuel consumption. Similarly, it was found that fuel consumption increased proportionally with displacement and that this relationship was robust across different speed ranges ([Bocchetti et al., 2013](#)). By advancing the theoretical and empirical foundations of container ship fuel consumption modeling in this manner, we contribute to the ongoing efforts to develop robust and accurate models for optimizing the fuel efficiency of container shipping operations.

4.1.2 Design ML architecture. For ML models, we utilize the ANN algorithm for the forecasting model. ANN is a powerful computational tool inspired by the structure and functioning of biological NNs (Bal Beşikçi *et al.*, 2016). The configuration of an ANN model involves several key parameters and settings that profoundly impact its performance. Firstly, we employ a feedforward architecture, specifically a multilayer perceptron (MLP), which enables the model to process information in a unidirectional flow. The model architecture comprises two hidden layers, each consisting of 200 hidden nodes strategically positioned to extract intricate patterns and relationships from the input data. To enhance the network's learning capabilities, we employ the Rectified Linear Unit (ReLU) activation function as in Equation (7), known for its ability to capture nonlinearities and facilitate efficient gradient-based optimization.

$$\text{ReLU: } f(x) = \max(0, x) \quad (7)$$

To ensure robustness and evaluate the model's generalization abilities, we employed a ten-fold cross-validation technique. This approach divides the dataset into ten subsets, with each subset taking turns as the validation set while the remaining nine subsets serve as the training data. By repeating this process ten times, we obtain a comprehensive assessment of the model's performance across different data partitions. Three ANN models, namely ANN1, ANN2 and ANN3, are determined based on input layers, where ANN1, ANN2 and ANN3 have the number of input nodes equivalent to the number of independent variables as in models LM1, LM2 and LM3, respectively. The output layer of the ANN model is an EC variable.

4.2 Simulation model

In this section, we adopt the energy expense model into the DES architecture for the catfish SC as a case study. AnyLogistix platform, including multiple parameters, is utilized to simulate the real-world SC. AnyLogistix has the ability to model complex and heterogeneous SCs, taking into account multiple levels of detail such as transportation, inventory management and demand patterns (Ivanov, 2019). Its fitness for this purpose lies in the fact that it uses advanced optimization algorithms to model and simulate SC operations, allowing for the evaluation of different scenarios and the identification of optimal solutions. Additionally, AnyLogistix offers a user-friendly interface and supports integration with other tools and software systems, making it a versatile and powerful tool for SC analysis and optimization.

The customer locations are obtained from the General Department of Vietnam Customs and converted to longitude and latitude using a geographic information system (GIS) tool. Customer locations are distributed in the USA, EU and China, and each location has a different demand pattern for catfish based on historical data in 2019. The number of sites in the SC is also modeled, including farms, processing plants and sea ports. An inventory policy was established to ensure sufficient safety stock levels at each site while minimizing excess inventory and associated holding costs. Transportation between sites is modeled using a combination of modes, including container trucks and container ships. Sourcing policies from farms to processing plants to departure ports and from arrival ports to customer locations are set up based on the closest criterion, while sourcing policies between sea ports are set based on a split-by-ratio function.

5. Analysis results

5.1 Model validation

Model validation is an essential step in the development of any simulation model. One way to evaluate the accuracy of a simulation model is through statistical measures such as Mean Absolute Percentage Error (MAPE), Root Mean Squared Percentage Error (RMSPE) and

Mean Percentage Error (MPE). These statistical measures are used to compare the actual and simulated data.

MAPE is a widely used measure of forecasting error that expresses the difference between the actual and predicted values as a percentage of the actual values. It is defined as the average of the absolute percentage errors, where the absolute percentage error is the absolute difference between the actual and predicted values divided by the actual value. MAPE is a relative error measure, which means that it is unaffected by the scale of the data.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \times 100\% \quad (8)$$

where A_t is the actual value at time t , F_t is the estimated value at time t and n is the total number of observations.

RMSPE is another measure of forecasting error that is similar to MAPE but takes into account the squared errors rather than the absolute errors. RMSPE is the square root of the mean squared percentage error, where the squared percentage error is the square of the difference between the actual and predicted values divided by the actual value. RMSPE is useful because it penalizes large errors more than small errors, making it a good measure of the overall accuracy of a forecasting model.

$$RMSPE = \sqrt{\frac{1}{n} \sum_{t=1}^n \left(\frac{A_t - F_t}{A_t} \right)^2} \times 100\% \quad (9)$$

MPE is a measure of the bias of a forecasting model, which is defined as the difference between the actual and predicted values. It is expressed as a percentage of the actual value, and positive values indicate that the model is over-predicting, while negative values indicate that the model is under-predicting. MPE is useful for detecting systematic errors in the model, such as a consistent over- or under-prediction of demand.

$$MPE = \frac{1}{n} \sum_{t=1}^n \left(\frac{A_t - F_t}{A_t} \right) \times 100\% \quad (10)$$

5.2 Model selection

Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are used in statistical modeling to determine the best-fit model among a set of candidate models. They are measures of the goodness of fit of a model to the data while also taking into account the complexity of the model.

AIC and BIC both balance the fit of the model to the data against the complexity of the model. The difference between AIC and BIC is the penalty term for model complexity. AIC has a lower penalty term, whereas BIC has a higher penalty term. As a result, BIC tends to select simpler models than AIC.

Both AIC and BIC are based on likelihood functions, which estimate the probability of observing the data given the model. AIC and BIC are defined as

$$AIC = -2 \log(L) + 2k \quad (11)$$

$$BIC = -2 \log(L) + \log(n) \times k \quad (12)$$

where L is the maximum value of the likelihood function for the model, k is the number of parameters in the model and n is the number of observations in the data set. The AIC and BIC values are minimized to select the best-fit model.

Table 3 presents the results of three models for predicting outcome variables. All three models have p -values less than 0.01 ($p < 0.01$), which indicates a statistically significant relationship between the independent variables and the dependent variable at the 99% confidence interval. LM1 has an RMSPE of 24.3%, an MPE of 5.02% and an MAPE of 20.3%. Meanwhile, LM2 performs better than LM1, with RMSPE of 23.6%, MPE of -3.48% and MAPE of 14.4%. The best MLR model is LM3, which has the lowest RMSPE of 22.1%, MPE of -2.98% and MAPE of 13.2%.

The empirical results demonstrate the superiority of all ANN models over their respective MLR counterparts. Specifically, ANN1 exhibits a noteworthy reduction in the RMSPE, MPE and MAPE when compared to LM1. Similarly, ANN2 surpasses LM2 and ANN3 demonstrates a remarkable improvement in all aforementioned indices compared to LM3. These findings underscore the enhanced predictive performance and efficacy of ANN models in capturing the complex relationships within the dataset. Among the various models considered, it is evident that ANN3 outperforms the rest in terms of cross-validation metrics, specifically RMSPE, MPE and MAPE. The results demonstrate the remarkable accuracy achieved by ANN3, with a notable RMSPE of 16.1%, an MPE of 1.15%, and impressive MAPE of 7.03%. In particular, the MAPE improvement offered by ANN3 is striking, surpassing the poorest-performing model, LM1, by a substantial margin of over 13%. Furthermore, it is noteworthy that even when employing the same set of input variables, ANN3 exhibits superior accuracy than LM3, with an additional gain of over 6%. This highlights the robustness and efficacy of the ANN3 model in capturing intricate patterns and nuances within the dataset, resulting in enhanced predictive performance.

In addition to the performance metrics, the table also reports other important information, such as the adjusted R^2 , AIC and BIC values. For all predicted models, the adjusted R^2 value is high, indicating a good fit to the data. Furthermore, the AIC and BIC values were consistent across all models, with LM1 having the highest values and ANN3 having the lowest values. Therefore, the results suggest that ANN3 is the best-performing model, with the lowest RMSPE, MPE and MAPE values as well as the lowest AIC and BIC values.

These findings underscore the significant advantages offered by the ANN approach, showcasing its ability to uncover hidden relationships and optimize the model's fit to the data. By leveraging the power of ANNs, ANN3 has demonstrated its capability to provide more accurate and reliable predictions, contributing to improved decision-making processes and bolstering the overall quality of the analytical framework.

5.3 SC performance

Figure 1 represents the cost and revenue breakdown for a business operation, with six different estimated models being used to predict the associated costs. The costs and revenue are categorized into various components, such as facility cost, inbound processing cost, initial

Model	RMSPE (%)	MPE (%)	MAPE (%)	Adjusted R^2	AIC	BIC
LM1	24.3	5.02	20.3	0.9522 (0.000***)	57,763.9	57,784.7
LM2	23.6	-3.48	14.4	0.9557 (0.000***)	57,611.3	57,639.0
LM3	22.1	-2.98	13.2	0.9648 (0.000***)	57,478.4	57,512.9
ANN1	20.6	2.04	11.1	0.9813 (0.000***)	56,732.1	56,815.5
ANN2	17.9	-1.73	8.39	0.9845 (0.000***)	56,454.9	56,543.7
ANN3	16.1	1.15	7.03	0.9877 (0.000***)	55,810.2	56,011.6

Note(s): In parentheses: p -value at 99% confidence interval (***)

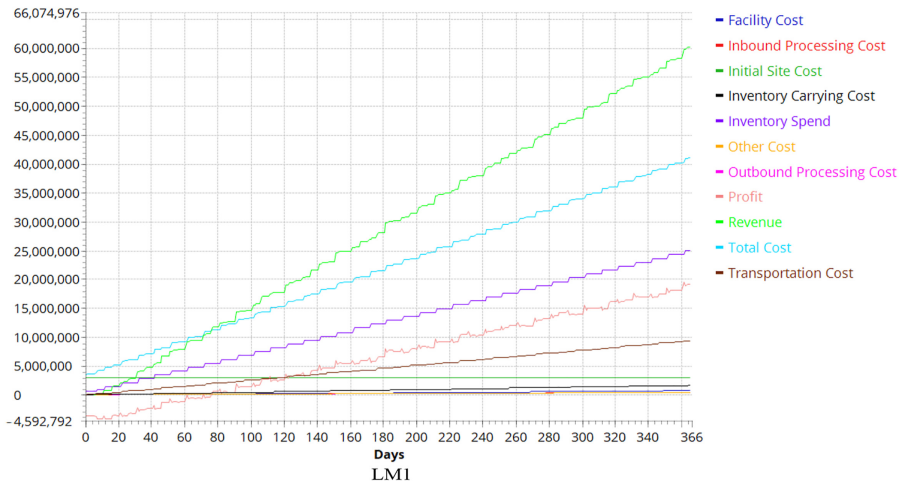
Source(s): Table by authors

Table 3.
Model validation and
evaluation

site cost, inventory carrying cost, inventory spend, other costs, outbound processing cost, profit, revenue and transportation cost.

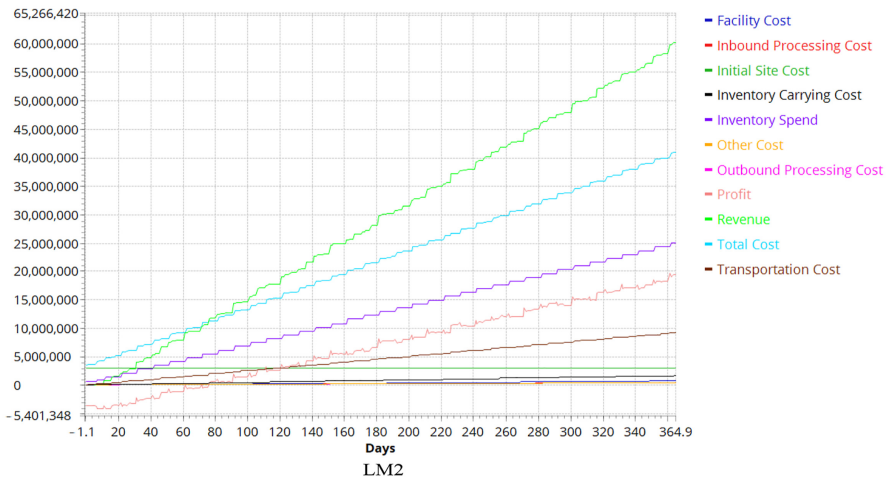
Comparing the costs of the estimated models, it can be seen that there are some variations in the costs predicted by each model. For instance, the simulation results reveal variations in inbound processing costs among the models. LM1 has the highest cost at USD 465,234, followed by LM2 at USD 445,816 and LM3 at USD 435,816. The ANN models show lower costs, with ANN3 having the lowest at USD 429,091. The analysis of outbound processing costs indicates similar trends. LM1 exhibits the highest cost at USD 454,456, while LM2 and

Profit and Loss Statement



(a)

Profit and Loss Statement

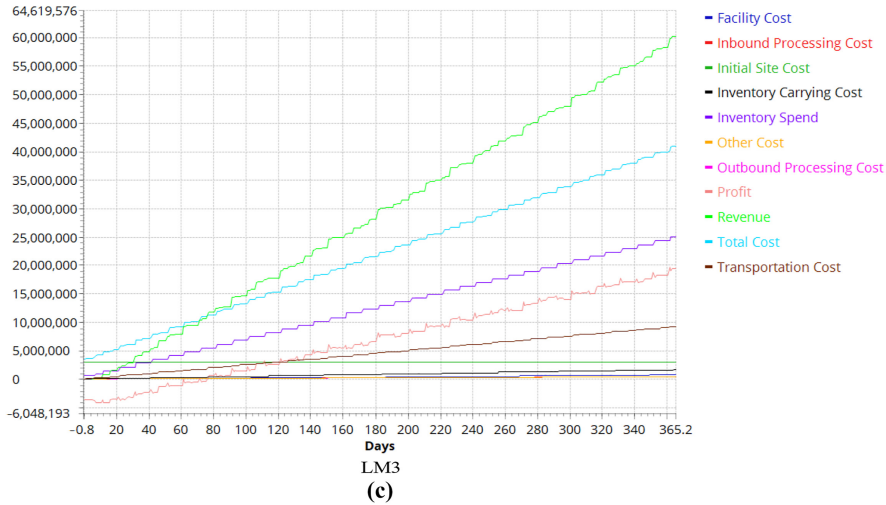


(b)

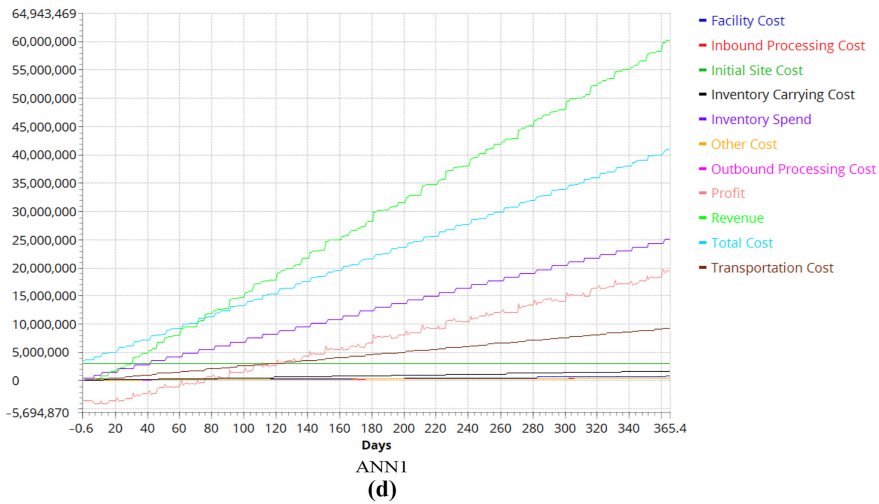
Figure 1.
Supply chain
performance by
estimation model

(continued)

Profit and Loss Statement

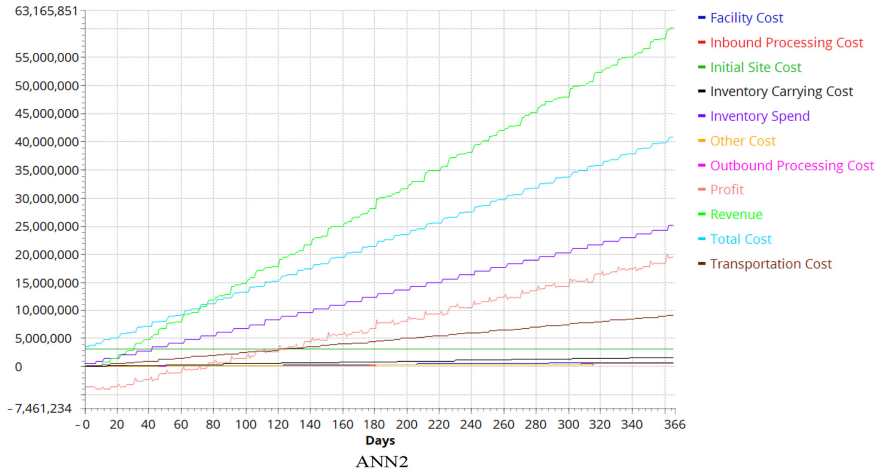


Profit and Loss Statement



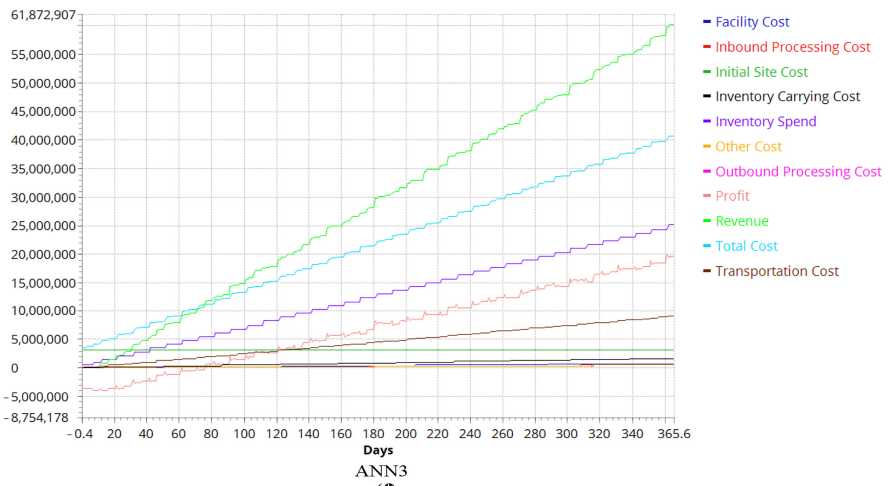
(continued) **Figure 1.**

Profit and Loss Statement



(e)

Profit and Loss Statement



(f)

Figure 1.

Source(s): Figure by authors

LM3 have slightly lower costs at USD 441,971 and USD 448,432, respectively. Among the ANN models, ANN3 shows the lowest outbound processing cost at USD 437,673. The evaluation of inventory spend showcases consistent costs across the different models. The traditional models (LM1, LM2 and LM3) demonstrate similar inventory spending figures, ranging from USD 25,094,437 to USD 25,055,480. Similarly, the ANN models (ANN1, ANN2 and ANN3) also exhibit comparable inventory spend, with values ranging from USD 25,064,164 to USD 25,006,129. These findings suggest that inventory spend is not significantly influenced by the choice of estimated models.

The analysis reveals notable variations in transportation costs among the different logistics models. LM1 exhibited the highest transportation cost at USD 9,325,341, followed by LM2 at USD 9,181,794 and LM3 at USD 9,170,655. Comparatively, the ANN models show slightly lower costs, with ANN3 having the lowest transportation cost of USD 9,163,036. However, there are certain costs that remain constant across all three models, such as the initial site cost, facility cost, inventory carrying cost, other cost, profit, revenue and total cost. These costs are likely to be relatively fixed and not significantly affected by the estimated model being used.

Therefore, for the fuel cost estimation model, despite the significantly higher accuracy of ANN models compared to traditional MLR models (up to nearly 13% higher), when running DES, the discrepancy in transportation costs between ANN3 and ML1 is only about 2%. This may be because, during DES execution, the system utilizes surplus resources to ensure that the supply chain operates smoothly and efficiently. Furthermore, although the difference is only 2% in percentage terms, the absolute value amounts to approximately USD 162,305. If the model were applied to the entire seafood supply chain of Vietnam, with a volume of USD 9.3 bn, this difference would amount to 22,870,250-USD. This represents a significant discrepancy between the two forecasting models. Therefore, the utilization of ANN models will assist SC operators in making decisions with less deviation.

6. Implications

6.1 Theoretical implications

This study provides several important theoretical implications, as follows: Firstly, the authors have successfully constructed robust ANN models for predicting the fuel costs of container ships. These models achieve higher forecasting accuracy than traditional regression methods. This accomplishment is significant in studies concerning fuel costs and emissions in maritime operations. Secondly, the study highlights the potential application of this model in SC research. SC scholars can fully leverage the model outputs to illustrate SC simulations by examining SC variations as fuel costs and constituent factors change. Through robust SC simulation tools such as AnyLogistix, AnyLogic or ExtendSim, researchers can observe SC operations and variations under different scenarios, providing strategic recommendations for SC operators. Finally, in line with the explosion of ML and DL algorithms, highly accurate forecasting models can be combined with robust simulation models, contributing to innovation in integrating multiple advanced methods in research.

6.2 Practical implications

This study also provides some practical values. Firstly, as previously mentioned, estimating fuel consumption and consequently fuel costs holds significant importance not only for scholars but also for ship owners and operators in fleet management practices. Secondly, having such an accurate model can assist environmental policymakers in measuring emissions from transportation activities as a basis for carbon tax policies, as it is not feasible to install monitoring stations to measure these emissions at sea. Thirdly, for the seafood SC in Vietnam and many emerging economies, small-scale businesses are primarily focused on aquaculture and selling to trading enterprises, which then export to foreign partners. Consequently, they often lack the tools to observe the entire SC, making it difficult to explore different scenarios to improve the current SC. Through this research, observing SC activities becomes more accessible, and scenarios aimed at enhancing various aspects of the SC (such as maritime transportation costs) can be envisioned more easily, avoiding the expenses and risks associated with real-world validation. Finally, the approach of this research marks an incredibly crucial platform for improving traditional SCs toward economic sustainability and

7. Concluding remarks

In summary, this paper has successfully achieved its initial research objective, which is to provide seafood SC operators in Vietnam with a detailed and accurate view of their chain operations through the combined use of two modern and robust techniques: the ANN algorithm and the DES platform. Vietnam's seafood SC operators, accustomed to conventional production and transactional practices based on experience, may now benefit from scientifically evaluated insights from advanced DES-ML simulation model analysis. Specifically, the ANN algorithm is employed to construct a fuel cost prediction model, serving as a key input parameter system in the main stage of the entire SC. Undeniably, the fuel cost estimation predicated on the ANN3 model generates outcomes that are superior in terms of precision and accuracy vis-à-vis the other models considered. In addition, the validation results indicate that the inclusion of the cargo load factor variable contributes significantly to improving the estimation model's accuracy. This is demonstrated by the marked increase (about 6%) in accuracy realized in the transition from the LM1 to the LM2 or (about 4%) from the ANN1 to the ANN2. While the displacement variable does not exert as strong an influence on estimation accuracy as the cargo load factor, it still plays a role in mitigating estimation errors, resulting in a reduction of approximately 1%. Notably, the ANN3 model emerged as the top performer in terms of estimation accuracy. It is also noteworthy that the MPE, MAPE and RMSPE validation errors are quite similar to those found in the estimation models advanced by [Le et al. \(2020a, b\)](#). These findings underscore the feasibility and reliability of the estimation model, which can be a valuable tool in predicting ship fuel consumption and fostering sustainable shipping practices.

The simulation model gives a breakdown of the costs and revenue for an SC operation and provides estimates using three different models. The analysis can provide insights into the costs and revenue associated with the operation as well as highlight potential areas for improvement in the SC. One key aspect of the operation is the transportation cost, which is variable across the three models. This cost may be influenced by factors such as fuel prices, sailing times and sailing speed. Understanding the factors that influence transportation costs is important for optimizing the SC operation. For instance, if delivery times can be optimized, transportation costs may be reduced. Another important aspect of the operation is inventory management. The inventory spend is also variable across the three models, suggesting that the estimated model used to predict inventory requirements may not be accurate. Implementing better inventory management strategies, such as just-in-time inventory management, may help reduce inventory spend and improve overall efficiency in SC operations.

The limitation of this study lies in its lack of detailed exploration into specific strategies or solutions to enhance SC operations, as it only reaches the level of constructing a visual model to facilitate SC operators' easy understanding of their operations. However, visualizing the SC based on the combination of two robust and modern methods, such as ML and DES, can help the authors develop future research directions to address complex challenges. One typical problem could be investigating the impact of slow steaming on the entire SC to determine the optimal speed that meets both the financial and environmental KPIs of the chain. Another potential problem could involve carbon taxation on maritime transportation routes to Europe and America. Thus, how such tax policies would affect Vietnam's seafood SC and what solutions SC operators need to prepare to maintain economic profitability and enhance the competitiveness of the chain are questions worth exploring.

References

- Abbasi, B., Babaei, T., Hosseinfard, Z., Smith-Miles, K. and Dehghani, M. (2020), "Predicting solutions of large-scale optimization problems via machine learning: a case study in blood supply chain management", *Computers and Operations Research*, Vol. 119, 104941, doi: [10.1016/j.cor.2020.104941](https://doi.org/10.1016/j.cor.2020.104941).
- Amorim-Lopes, M., Guimarães, L., Alves, J. and Almada-Lobo, B. (2021), "Improving picking performance at a large retailer warehouse by combining probabilistic simulation, optimization, and discrete-event simulation", *International Transactions in Operational Research*, Vol. 28 No. 2, pp. 687-715, doi: [10.1111/itor.12852](https://doi.org/10.1111/itor.12852).
- Bal Beşikçi, E., Arslan, O., Turan, O. and Ölçer, A.I. (2016), "An artificial neural network based decision support system for energy efficient ship operations", *Computers and Operations Research*, Vol. 66, pp. 393-401, doi: [10.1016/j.cor.2015.04.004](https://doi.org/10.1016/j.cor.2015.04.004).
- Bialystocki, N. and Konovessis, D. (2016), "On the estimation of ship's fuel consumption and speed curve: a statistical approach", *Journal of Ocean Engineering and Science*, Vol. 1 No. 2, pp. 157-166, doi: [10.1016/j.joes.2016.02.001](https://doi.org/10.1016/j.joes.2016.02.001).
- Bocchetti, D., Lepore, A., Palumbo, B. and Vitiello, L. (2013), "A statistical control of the ship fuel consumption", RINA, Royal Institution of Naval Architects – Design and Operation of Passenger Ships 2013, doi: [10.3940/rina.pass.2013.01](https://doi.org/10.3940/rina.pass.2013.01).
- Bodendorf, F., Merkl, P. and Franke, J. (2022), "Artificial neural networks for intelligent cost estimation—a contribution to strategic cost management in the manufacturing supply chain", *International Journal of Production Research*, Vol. 60 No. 21, pp. 6637-6658, doi: [10.1080/00207543.2021.1998697](https://doi.org/10.1080/00207543.2021.1998697).
- Bui-Duy, L. and Vu-Thi-Minh, N. (2021), "Utilization of a deep learning-based fuel consumption model in choosing a liner shipping route for container ships in Asia", *Asian Journal of Shipping and Logistics*, Vol. 37 No. 1, pp. 1-11, doi: [10.1016/j.ajsl.2020.04.003](https://doi.org/10.1016/j.ajsl.2020.04.003).
- Bui-Duy, L., Le, L.T., Vu-Thi-Minh, N., Hoang-Huong, G., Bui-Thi-Thanh, N. and Nguyen, P.N. (2023), "Economic and environmental analysis of biomass pellet supply chain using simulation-based approach", *Asia Pacific Management Review*, Vol. 28 No. 4, pp. 470-486, doi: [10.1016/j.apmr.2023.02.002](https://doi.org/10.1016/j.apmr.2023.02.002).
- Cao, Y., Williams, D.D. and Williams, N.E. (1999), "Data transformation and standardization in the multivariate analysis of river water quality", *Ecological Applications*, Vol. 9 No. 2, pp. 669-677, doi: [10.1890/1051-0761\(1999\)009\[0669:DTASIT\]2.0.CO;2](https://doi.org/10.1890/1051-0761(1999)009[0669:DTASIT]2.0.CO;2).
- Christopher, M. and Peck, H. (2004), "Building the resilient supply chain", *The International Journal of Logistics Management*, Vol. 15 No. 2, pp. 1-14, doi: [10.1108/09574090410700275](https://doi.org/10.1108/09574090410700275).
- Coraddu, A., Oneto, L., Baldi, F. and Anguita, D. (2017), "Vessels fuel consumption forecast and trim optimisation: a data analytics perspective", *Ocean Engineering*, Vol. 130, pp. 351-370, doi: [10.1016/j.oceaneng.2016.11.058](https://doi.org/10.1016/j.oceaneng.2016.11.058).
- Dragović, B., Tzannatos, E., Tselentis, V., Meštrović, R. and Škurić, M. (2018), "Ship emissions and their externalities in cruise ports", *Transportation Research Part D: Transport and Environment*, Vol. 61, pp. 289-300, doi: [10.1016/j.trd.2015.11.007](https://doi.org/10.1016/j.trd.2015.11.007).
- Feizabadi, J. (2022), "Machine learning demand forecasting and supply chain performance", *International Journal of Logistics Research and Applications*, Vol. 25 No. 2, pp. 119-142, doi: [10.1080/13675567.2020.1803246](https://doi.org/10.1080/13675567.2020.1803246).
- Gonzalez, H., Han, J., Ouyang, Y. and Seith, S. (2011), "Multidimensional data mining of traffic anomalies on large-scale road networks", *Transportation Research Record*, Vol. 2215 No. 1, pp. 75-84, doi: [10.3141/2215-08](https://doi.org/10.3141/2215-08).
- Hadi, A.S. and Simonoff, J.S. (1993), "Procedures for the identification of multiple outliers in linear models", *Journal of the American Statistical Association*, Vol. 88 No. 424, pp. 1264-1272, doi: [10.1080/01621459.1993.10476407](https://doi.org/10.1080/01621459.1993.10476407).
- Han, C. and Zhang, Q. (2021), "Optimization of supply chain efficiency management based on machine learning and neural network", *Neural Computing and Applications*, Vol. 33 No. 5, pp. 1419-1433, doi: [10.1007/s00521-020-05023-1](https://doi.org/10.1007/s00521-020-05023-1).

- Ivanov, D. (2019), *Supply Chain Simulation and Optimization with any Logistix*, 4th ed., Berlin School of Economics and Law, Berlin.
- Le, L.T., Lee, G., Park, K.S. and Kim, H. (2020a), "Neural network-based fuel consumption estimation for container ships in Korea", *Maritime Policy and Management*, Vol. 47 No. 5, pp. 615-632, doi: [10.1080/03088839.2020.1729437](https://doi.org/10.1080/03088839.2020.1729437).
- Le, L.T., Lee, G., Kim, H. and Woo, S.H. (2020b), "Voyage-based statistical fuel consumption models of ocean-going container ships in Korea", *Maritime Policy and Management*, Vol. 47 No. 3, pp. 304-331, doi: [10.1080/03088839.2019.1684591](https://doi.org/10.1080/03088839.2019.1684591).
- Legato, P. and Mazza, R.M. (2020), "Queueing analysis for operations modeling in port logistics", *Maritime Business Review*, Vol. 5 No. 1, pp. 67-83, doi: [10.1108/MABR-09-2019-0035](https://doi.org/10.1108/MABR-09-2019-0035).
- Legato, P. and Mazza, R.M. (2023), "Queueing networks for supporting container storage and retrieval", *Maritime Business Review*, Vol. 8 No. 4, pp. 301-317, doi: [10.1108/MABR-01-2023-0009](https://doi.org/10.1108/MABR-01-2023-0009).
- Li, X.Y. and Tang, B.J. (2017), "Incorporating the transport sector into carbon emission trading scheme: an overview and outlook", *Natural Hazards*, Vol. 88 No. 2, pp. 683-698, doi: [10.1007/s11069-017-2886-3](https://doi.org/10.1007/s11069-017-2886-3).
- Li, X., Du, Y., Chen, Y., Nguyen, S., Zhang, W., Schonborn, A. and Sun, Z. (2022), "Data fusion and machine learning for ship fuel efficiency modeling: part I – voyage report data and meteorological data", *Communications in Transportation Research*, Vol. 2, 100074, doi: [10.1016/j.commr.2022.100074](https://doi.org/10.1016/j.commr.2022.100074).
- Lichstein, J.W., Simons, T.R., Shriner, S.A. and Franzreb, K.E. (2002), "Spatial autocorrelation and autoregressive models in ecology", *Ecological Monographs*, Vol. 72 No. 3, pp. 445-463, doi: [10.1890/0012-9615\(2002\)072\[0445:SAAMI\]2.0.CO;2](https://doi.org/10.1890/0012-9615(2002)072[0445:SAAMI]2.0.CO;2).
- Liu, L., Shibasaki, R., Zhang, Y., Kosuge, N., Zhang, M. and Hu, Y. (2023), "Data-driven framework for extracting global maritime shipping networks by machine learning", *Ocean Engineering*, Vol. 269, 113494, doi: [10.1016/j.oceaneng.2022.113494](https://doi.org/10.1016/j.oceaneng.2022.113494).
- Meng, Q., Du, Y. and Wang, Y. (2016), "Shipping log data based container ship fuel efficiency modeling", *Transportation Research Part B: Methodological*, Vol. 83, pp. 207-229, doi: [10.1016/j.trb.2015.11.007](https://doi.org/10.1016/j.trb.2015.11.007).
- Moussa, B.F.Z., De Guio, R., Dubois, S., Rasovska, I. and Benmoussa, R. (2019), "Study of an innovative method based on complementarity between ARIZ, lean management and discrete event simulation for solving warehousing problems", *Computers and Industrial Engineering*, Vol. 132, pp. 124-140, doi: [10.1016/j.cie.2019.04.024](https://doi.org/10.1016/j.cie.2019.04.024).
- Nguyen, T., Duong, Q.H., Nguyen, T.V., Zhu, Y. and Zhou, L. (2022), "Knowledge mapping of digital twin and physical internet in supply chain management: a systematic literature review", *International Journal of Production Economics*, Vol. 244, 108381, doi: [10.1016/j.ijpe.2021.108381](https://doi.org/10.1016/j.ijpe.2021.108381).
- Ni, D., Xiao, Z. and Lim, M.K. (2020), "A systematic review of the research trends of machine learning in supply chain management", *International Journal of Machine Learning and Cybernetics*, Vol. 11 No. 7, pp. 1463-1482, doi: [10.1007/s13042-019-01050-0](https://doi.org/10.1007/s13042-019-01050-0).
- Oliveira, D.R.L., Fagundes, L.D., da Silva Lima, R. and Montaña, M. (2020), "Discrete event simulation to aid decision-making and mitigation in solid waste management", *Mitigation and Adaptation Strategies for Global Change*, Vol. 25 No. 1, pp. 67-85, doi: [10.1007/s11027-019-09859-4](https://doi.org/10.1007/s11027-019-09859-4).
- Peng, Y., Liu, H., Li, X., Huang, J. and Wang, W. (2020), "Machine learning method for energy consumption prediction of ships in port considering green ports", *Journal of Cleaner Production*, Vol. 264, 121564, doi: [10.1016/j.jclepro.2020.121564](https://doi.org/10.1016/j.jclepro.2020.121564).
- Piramuthu, S. (2005), "Machine learning for dynamic multi-product supply chain formation", *Expert Systems with Applications*, Vol. 29 No. 4, pp. 985-990, doi: [10.1016/j.eswa.2005.07.004](https://doi.org/10.1016/j.eswa.2005.07.004).
- Prinz, R., Väättäin, K., Laitila, J., Sikanen, L. and Asikainen, A. (2019), "Analysis of energy efficiency of forest chip supply systems using discrete-event simulation", *Applied Energy*, Vol. 235, pp. 1369-1380, doi: [10.1016/j.apenergy.2018.11.053](https://doi.org/10.1016/j.apenergy.2018.11.053).

- Psarafitis, H.N. and Kontovas, C.A. (2013), "Speed models for energy-efficient maritime transportation: a taxonomy and survey", *Transportation Research Part C: Emerging Technologies*, Vol. 26, pp. 331-351, doi: [10.1016/j.trc.2012.09.012](https://doi.org/10.1016/j.trc.2012.09.012).
- Rabe, M., Deininger, M. and Juan, A.A. (2020), "Speeding up computational times in simheuristics combining genetic algorithms with discrete-Event simulation", *Simulation Modelling Practice and Theory*, Vol. 103, 102089, doi: [10.1016/j.simpat.2020.102089](https://doi.org/10.1016/j.simpat.2020.102089).
- Sharma, R., Kamble, S.S., Gunasekaran, A., Kumar, V. and Kumar, A. (2020), "A systematic literature review on machine learning applications for sustainable agriculture supply chain performance", *Computers and Operations Research*, Vol. 119, 104926, doi: [10.1016/j.cor.2020.104926](https://doi.org/10.1016/j.cor.2020.104926).
- Sharma, R., Shishodia, A., Gunasekaran, A., Min, H. and Munim, Z.H. (2022), "The role of artificial intelligence in supply chain management: mapping the territory", *International Journal of Production Research*, Vol. 60 No. 24, pp. 1-24, doi: [10.1080/00207543.2022.2029611](https://doi.org/10.1080/00207543.2022.2029611).
- Tseng, M.L., Ha, H.M., Tran, T.P.T., Bui, T.D., Lim, M.K., Lin, C.W. and Ali, M.H. (2022), "Data-driven on sustainable food supply chain: a comparison on Halal and non-Halal food system", *Journal of Industrial and Production Engineering*, Vol. 39 No. 6, pp. 430-457, doi: [10.1080/21681015.2022.2040622](https://doi.org/10.1080/21681015.2022.2040622).
- Uyanik, T., Karatuğ, Ç. and Arslanoğlu, Y. (2020), "Machine learning approach to ship fuel consumption: a case of container vessel", *Transportation Research Part D: Transport and Environment*, Vol. 84, 102389, doi: [10.1016/j.trd.2020.102389](https://doi.org/10.1016/j.trd.2020.102389).
- Vettor, R. and Guedes Soares, C. (2022), "Reflecting the uncertainties of ensemble weather forecasts on the predictions of ship fuel consumption", *Ocean Engineering*, Vol. 250, 111009, doi: [10.1016/j.oceaneng.2022.111009](https://doi.org/10.1016/j.oceaneng.2022.111009).
- Wang, S. and Peng, C. (2019), "Model and analysis of the effect of China's potential domestic emission control area with 0.1% sulphur limit", *Maritime Business Review*, Vol. 4 No. 3, pp. 298-309, doi: [10.1108/MABR-03-2019-0012](https://doi.org/10.1108/MABR-03-2019-0012).
- Wen, Y., Sui, Z., Zhou, C., Xiao, C., Chen, Q., Dong, S.H. and Zhang, Y. (2020), "Automatic ship route design between two ports: a data-driven method", *Applied Ocean Research*, Vol. 96, 102049, doi: [10.1016/j.apor.2019.102049](https://doi.org/10.1016/j.apor.2019.102049).
- Wilcox, R.R. (2016), *Introduction to Robust Estimation and Hypothesis Testing*, 4th ed, Academic Press, Burlington.
- Wilkinson, J.M. (2012), "Reducing greenhouse gas emissions from livestock", *Livestock*, Vol. 17 No. 5, pp. 25-27, doi: [10.1111/j.2044-3870.2012.00138.x](https://doi.org/10.1111/j.2044-3870.2012.00138.x).
- Winther, J.G., Dai, M., Rist, T., Hoel, A.H., Li, Y., Trice, A., Morrissey, K., Juinio-Meñez, M.A., Fernandes, L., Unger, S., Scarano, F.R., Halpin, P. and Whitehouse, S. (2020), "Integrated ocean management for a sustainable ocean economy", *Nature Ecology and Evolution*, Vol. 4 No. 11, pp. 1451-1458, doi: [10.1038/s41559-020-1259-6](https://doi.org/10.1038/s41559-020-1259-6).
- Yan, R., Wang, S. and Du, Y. (2020), "Development of a two-stage ship fuel consumption prediction and reduction model for a dry bulk ship", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 138, 101930, doi: [10.1016/j.tre.2020.101930](https://doi.org/10.1016/j.tre.2020.101930).
- Yan, R., Wang, S. and Psarafitis, H.N. (2021), "Data analytics for fuel consumption management in maritime transportation: status and perspectives", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 155, 102489, doi: [10.1016/j.tre.2021.102489](https://doi.org/10.1016/j.tre.2021.102489).
- Yang, L., Chen, G., Rytter, N.G.M., Zhao, J. and Yang, D. (2019), "A genetic algorithm-based grey-box model for ship fuel consumption prediction towards sustainable shipping", *Annals of Operations Research*, Vol. 2029, doi: [10.1007/s10479-019-03183-5](https://doi.org/10.1007/s10479-019-03183-5).
- Yuan, Z., Liu, J., Zhang, Q., Liu, Y., Yuan, Y. and Li, Z. (2021), "Prediction and optimisation of fuel consumption for inland ships considering real-time status and environmental factors", *Ocean Engineering*, Vol. 221, 108530, doi: [10.1016/j.oceaneng.2020.108530](https://doi.org/10.1016/j.oceaneng.2020.108530).

Yuksel, O., Bayraktar, M. and Sokukcu, M. (2023), "Comparative study of machine learning techniques to predict fuel consumption of a marine diesel engine", *Ocean Engineering*, Vol. 286 No. 1, 115505, doi: [10.1016/j.oceaneng.2023.115505](https://doi.org/10.1016/j.oceaneng.2023.115505).

Zakerdoost, H. and Ghassemi, H. (2019), "A multi-level optimization technique based on fuel consumption and energy index in early-stage ship design", *Structural and Multidisciplinary Optimization*, Vol. 59 No. 5, pp. 1417-1438, doi: [10.1007/s00158-018-2136-7](https://doi.org/10.1007/s00158-018-2136-7).

Corresponding author

Luan Thanh Le can be contacted at: luanlt13@fe.edu.vn

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgroupublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com