

Special Issue Reprint

Sustainable Maritime Transport and Port Intelligence

Edited by
Chao Mi and Guangnian Xiao

mdpi.com/journal/jmse

Sustainable Maritime Transport and Port Intelligence

Sustainable Maritime Transport and Port Intelligence

Guest Editors

Chao Mi

Guangnian Xiao



Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Guest Editors

Chao Mi
College of Logistics
Engineering
Shanghai Maritime
University
Shanghai
China

Guangnian Xiao
School of Economics &
Management
Shanghai Maritime
University
Shanghai
China

Editorial Office

MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Journal of Marine Science and Engineering* (ISSN 2077-1312), freely accessible at: https://www.mdpi.com/journal/jmse/special_issues/1048MFL8D0.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. <i>Journal Name</i> Year , Volume Number, Page Range.
--

ISBN 978-3-7258-8205-2 (Hbk)

ISBN 978-3-7258-8206-9 (PDF)

<https://doi.org/10.3390/books978-3-7258-8206-9>

© 2026 by the authors. Articles in this reprint are Open Access and distributed under the Creative Commons Attribution (CC BY) license. The reprint as a whole is distributed by MDPI under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Contents

Preface	vii
Guangnian Xiao and Chao Mi	
Sustainable Maritime Transport and Port Intelligence Reprinted from: <i>J. Mar. Sci. Eng.</i> 2026 , 14, 592, https://doi.org/10.3390/jmse14070592	1
Fan Zhou, Yuxuan Wang and Yinghan Zhou	
The Evolution of Ship Fuel Sulfur Content Monitoring—From Exhaust Gas Measurement to AI-Driven Comprehensive Analysis Reprinted from: <i>J. Mar. Sci. Eng.</i> 2025 , 13, 1795, https://doi.org/10.3390/jmse13091795	6
Sergej Jakovlev, Tomas Eglynas, Edvinas Pocevicus, Miroslav Voznak, Gediminas Gricius, Valdas Jankunas and Mindaugas Jusis	
Comparative Analysis of Recurrent vs. Temporal Convolutional Autoencoders for Detecting Container Impacts During Quay Crane Handling Reprinted from: <i>J. Mar. Sci. Eng.</i> 2025 , 13, 1231, https://doi.org/10.3390/jmse13071231	27
Filip Bojić, Anita Gudelj and Rino Bošnjak	
A Comprehensive Model for Quantifying, Predicting, and Evaluating Ship Emissions in Port Areas Using Novel Metrics and Machine Learning Methods Reprinted from: <i>J. Mar. Sci. Eng.</i> 2025 , 13, 1162, https://doi.org/10.3390/jmse13061162	50
Jiahao Ni, Hongjun Tian, Kaijie Zhang, Yihong Xue and Yang Xiong	
Data-Driven Multi-Scale Channel-Aligned Transformer for Low-Carbon Autonomous Vessel Operations: Enhancing CO ₂ Emission Prediction and Green Autonomous Shipping Efficiency Reprinted from: <i>J. Mar. Sci. Eng.</i> 2025 , 13, 1143, https://doi.org/10.3390/jmse13061143	80
Yanmeng Tao, Ying Yang, Yuquan Du and Shuaian Wang	
Carbon Dioxide Storage Site Location and Transport Assignment Optimization for Sustainable Maritime Transport Reprinted from: <i>J. Mar. Sci. Eng.</i> 2025 , 13, 1055, https://doi.org/10.3390/jmse13061055	109
Xuhui Yu and Cong He	
Collaborative Scheduling of Yard Cranes, External Trucks, and Rail-Mounted Gantry Cranes for Sea–Rail Intermodal Containers Under Port–Railway Separation Mode Reprinted from: <i>J. Mar. Sci. Eng.</i> 2025 , 13, 1109, https://doi.org/10.3390/jmse13061109	153
Siqian Cheng, Jiankun Hu, Youfang Huang and Zhihua Hu	
Enhancing Port Shipping Synergy Through Bayesian Network: A Case of Major Chinese Ports Reprinted from: <i>J. Mar. Sci. Eng.</i> 2025 , 13, 1093, https://doi.org/10.3390/jmse13061093	181
Marija Jović Mihanović, Saša Aksentijević, Edvard Tijan and Gregor Lenart	
Digital Transformation and Business Model Innovation: Enhancing Productivity in the Croatian Maritime Transport Sector Reprinted from: <i>J. Mar. Sci. Eng.</i> 2025 , 13, 999, https://doi.org/10.3390/jmse13050999	209
Xueqiang Du, Bencheng Luo, Jing Wang, Jieting Zhao, Dahai Li, Qian Sun and Haobin Li	
Optimizing Stack-Yard Positioning in Full Shoreline Loading Operations Reprinted from: <i>J. Mar. Sci. Eng.</i> 2025 , 13, 593, https://doi.org/10.3390/jmse13030593	228
Yang Shen, Xintai Man, Jiaqi Wang, Yujie Zhang and Chao Mi	
Truck Lifting Accident Detection Method Based on Improved PointNet++ for Container Terminals Reprinted from: <i>J. Mar. Sci. Eng.</i> 2025 , 13, 256, https://doi.org/10.3390/jmse13020256	249

Yifan Shen, Beng Xuan, Hongtao Hu, Yansong Wu, Ning Zhao and Zhen Yang
A Decomposed-Ensemble Prediction Framework for Gate-In Operations at Container Terminals
Reprinted from: *J. Mar. Sci. Eng.* **2025**, 13, 45, <https://doi.org/10.3390/jmse13010045> **266**

Preface

Maritime transport remains the backbone of global trade and plays an indispensable role in supporting economic exchange, industrial connectivity, and regional development worldwide. At the same time, the maritime and port sectors are undergoing profound transformation under the combined influence of geopolitical uncertainty, supply chain restructuring, climate change, digitalization, and the accelerating transition toward low-carbon development. Challenges such as route disruptions, rising operating costs, environmental pressure, and higher expectations for efficiency and resilience have made it increasingly necessary to rethink how shipping systems and ports should develop in the years ahead. In this context, sustainable maritime transport and port intelligence have emerged as two closely connected and strategically important areas of research and practice.

This Reprint was prepared to present the recent advances in these fields and to encourage deeper discussion on how the maritime industry can become greener, smarter, and more adaptive to global change. The contributions included in this collection reflect growing scholarly and practical interest in the integration of technological innovation, intelligent systems, and sustainable management approaches within maritime transport and port operations. The Reprint highlights not only the pressing challenges currently faced by the shipping and port industries but also the opportunities created by digital transformation, automation, artificial intelligence, clean energy transition, and collaborative governance.

The studies gathered in this Reprint explore a wide range of important topics, including climate change and carbon emission reduction in shipping, the digital transformation and intelligent development of ports, green port construction and environmental protection measures, collaborative innovation between the shipping and port industries, and the future pathways of sustainable maritime transport and port management. Several contributions also examine smart operations management, intelligent marine and port equipment, and data-driven decision-making methods that can improve operational efficiency, reduce environmental impacts, and strengthen systems' resilience. Together, these studies provide a multidimensional understanding of how maritime transport and port systems can respond to contemporary pressures while pursuing long-term sustainability goals.

One of the key motivations for this Reprint is the recognition that sustainability in maritime transport can no longer be addressed only through isolated technical improvements or single policy interventions. Instead, meaningful progress requires a comprehensive perspective that connects vessel operations, port management, logistics coordination, infrastructure modernization, digital capability, environmental governance, and cross-sector collaboration. Likewise, port intelligence should not be understood only as the application of advanced technologies, but also as a broader transformation in operational philosophy, management structure, service design, and strategic planning. By bringing together studies from different disciplinary perspectives and practical contexts, this Reprint seeks to support a more integrated understanding of these issues.

Another important aim of this Reprint is to build a bridge between academic research and real-world application. Maritime transport and port management are fields in which theoretical innovation must ultimately be tested through practice. The works collected here therefore offer not only conceptual insights, but also policy implications, empirical evidence, and practical experiences that may assist researchers, industry professionals, and public decision makers. It is hoped that this Reprint will serve as a useful reference for those who are committed to improving maritime sustainability, enhancing port intelligence, and promoting the coordinated development of global

shipping and port systems.

Overall, this Reprint presents a timely collection of studies related to one of the most important themes in contemporary maritime research. It contributes to ongoing efforts to understand how shipping and ports can respond effectively to environmental, technological, and operational challenges while maintaining their essential role in global trade. We hope that the ideas, methods, and findings presented in this collection will stimulate further research, support informed policy discussion, and inspire innovative practices for the future development of sustainable maritime transport and intelligent ports.

Chao Mi and Guangnian Xiao

Guest Editors

Editorial

Sustainable Maritime Transport and Port Intelligence

Guangnian Xiao ¹ and Chao Mi ^{2,*}

¹ School of Economics & Management, Shanghai Maritime University, Shanghai 201306, China; gnxiao@shmtu.edu.cn

² College of Logistics Engineering, Shanghai Maritime University, Shanghai 201306, China

* Correspondence: chaomi@shmtu.edu.cn

The maritime sector is undergoing a profound transformation. Global trade volumes and logistical complexity continue to grow, while ambitious decarbonization targets, stricter environmental regulations, and rapid advances in digital technology are reshaping how ships, ports, and hinterland systems are designed and operated. Against this backdrop, this Special Issue gathers recent research at the intersection of operation optimization, digital transformation, and data-driven intelligence in ports and shipping, with a particular focus on safety, efficiency, and environmental performance.

The eleven contributions span topics from quay crane impact detection and truck lifting accident prevention to collaborative scheduling in sea–rail intermodal systems and gate-in operations, as well as AI-enabled emission monitoring, maritime CO₂ transport planning, and business model innovation. Together, they offer a multi-layered view of how advanced analytics, sensing, and optimization can support a smarter and greener maritime transport system.

Several papers concentrate on enhancing the safety and efficiency of port operations through advanced sensing, forecasting, and optimization. At the berth–yard interface, Shen et al. propose a truck-lifting accident detection method for container terminals based on multi-line LiDAR and an improved PointNet++ model [1]. By integrating a multi-layer perceptron and a mixed attention mechanism into the network, their approach significantly improves three-dimensional point-cloud segmentation of containers, truck chassis, and wheels, achieving higher accuracy and robustness than mainstream baselines in field tests. This study illustrates how rich 3D sensing combined with tailored deep architectures can address long-standing safety risks in automated terminal operations.

Complementing safety with efficiency, Du et al. address the stack-yard positioning problem in full shoreline loading operations [2]. They develop an intelligent decision-making model that aligns local quay crane instructions with global yard allocation decisions and solve it using an enhanced COGA–GRASP heuristic. Treating the berthed vessel as a unified operational entity, the model optimizes yard locations across all handling routes to reduce conflicts in container retrieval and improve overall loading performance along the quay.

Recognizing the stochastic nature of landside flows, another study by Shen et al. develops a decomposition–ensemble framework for short-term gate-in prediction at container terminals [3]. To mitigate yard space wastage, external truck congestion and conflicts with loading activities caused by random export container arrivals, the authors forecast hourly gate-in volumes over a 24 h horizon. Total gate-in flows are decomposed by shipping routes to exploit the periodicity of liner services, and ARIMA, Prophet and Long Short-Term Memory (LSTM) models are compared. Using one year of data from Xiamen Port, the study shows that route-based decomposition enhances precision and that LSTM outperforms

classical time-series methods, providing planners with more reliable demand estimates for stacking and resource allocation.

Extending the scope inland, Yu and He examine collaborative scheduling of yard cranes, external trucks and rail-mounted gantry cranes for sea–rail intermodal containers under a port–railway separation mode [4]. They formulate a mixed-integer programming model that incorporates train timetables and export delivery deadlines, and design an improved genetic algorithm combined with a “First Accessible Machinery” rule. Numerical experiments show that the proposed method outperforms both exact solvers and adaptive large neighborhood search in terms of solution quality and computation time, offering practical guidance for synchronizing tightly coupled equipment across spatially separated port and rail facilities.

At the level of asset condition monitoring, Jakovlev et al. present a comparative analysis of recurrent and temporal convolutional autoencoders for detecting container impacts during quay crane handling [5]. Implemented within an edge computing unit, a recurrent autoencoder trained on acceleration data can accurately identify impact events with high precision and recall while remaining computationally efficient for deployment on resource-constrained devices. The study demonstrates the feasibility of moving from purely centralized monitoring architectures towards edge-intelligent safety systems in ports.

Collectively, these contributions highlight a shift from isolated tools to integrated sensing–prediction–optimization pipelines, in which high-resolution data are captured across quay, yard, gate and rail interfaces, processed close to the source when possible, and fed into decision support systems that coordinate equipment and operations across the entire terminal.

Digital technologies create value only when embedded in appropriate organizational and institutional contexts. Two papers in this Special Issue explicitly explore these systemic dimensions. Jović Mihanović et al. analyze how digital transformation and business model innovation affect productivity in the Croatian maritime transport sector [6]. Based on an online survey and statistical methods including Spearman’s correlation and principal component analysis, the study identifies two key dimensions of business model change—innovation and process digitalization—and shows that both are positively associated with productivity. The results underline the importance of digitalizing internal and external processes, developing new digital revenue streams, and introducing innovative services and pricing models. Beyond the country-level case, the paper sheds light on how maritime stakeholders can transform digital initiatives into tangible operational and economic gains.

Cheng et al. focus on port–shipping collaboration in major Chinese ports using a Bayesian network approach [7]. Integrating expert knowledge and empirical data, they construct a directed acyclic graph capturing interactions among port enterprises, shipping companies, customers and government actors. Their analysis identifies several leverage points—including customer engagement in green supply chains, perceived service quality, digital information integration, maturity of multilateral trading mechanisms and strict policy enforcement—and reveals a feedback loop between digital integration and regulatory enforcement. The study extends collaborative governance frameworks by explicitly incorporating “digital-era connectivity” and offers an action agenda for managers and policymakers seeking to coordinate environmental and service objectives across organizational boundaries.

These contributions underscore that digitalization is as much a socio-institutional process as it is a technological one. Organizational readiness, stakeholder alignment and supportive policy frameworks are critical to unlocking the full potential of data-driven tools, and merit closer attention in future maritime research.

A central pillar of this Special Issue is the use of optimization models and AI techniques to support climate and environmental objectives in maritime transport.

At the strategic planning level, Tao et al. consider the CO₂ storage site location and transport assignment problem in maritime carbon capture and storage systems [8]. They formulate a mixed-integer programming model that minimizes tactical costs—including storage site construction, ship chartering, transport and penalties for untransported CO₂—and show that the objective function is submodular. A case study demonstrates how jointly optimizing site selection, ship deployment and route allocation can significantly reduce both costs and emissions. To enhance practical applicability, the authors propose a two-stage framework that combines site selection with a genetic-algorithm-based transport assignment, yielding near-optimal solutions at much lower computation times. This work provides a rigorous and implementable framework for designing maritime CO₂ transport networks.

At the operational level on board vessels, Ni et al. introduce a Multi-scale Channel-aligned Transformer (MCAT) model for real-time prediction of CO₂ emissions from autonomous vessels within a hybrid 5G–satellite–IoT communication architecture [9]. The model leverages multi-scale token reconstruction and dual-level attention to capture spatio-temporal dependencies in heterogeneous data streams (AIS, onboard sensors, weather) while suppressing high-frequency noise. Validated on a 22-dimensional real dataset, MCAT outperforms existing baseline methods in prediction accuracy and robustness. Coupled with an ultra-low latency communication framework, this approach outlines a closed-loop system capable of supporting route optimization, fuel-efficiency improvement and compliance with evolving CII and MRV regulations.

Zooming in on port areas, Bojić et al. develop PrE-PARE, a holistic model for quantifying, predicting and evaluating ship emissions in ports using novel metrics and machine learning [10]. The three-module framework integrates detailed technical and movement data to build high-resolution emission inventories, adapts the MARS method for emission prediction under varying operational scenarios, and proposes new metrics and classification schemes to compare ships, assess emission risks and identify optimization potential. The model offers a transparent and adaptable tool for air quality management, supporting port authorities in planning mitigation measures and evaluating the impact of operational changes or regulatory interventions.

Zhou et al. extend the environmental focus by advancing the monitoring of ship fuel sulfur content (FSC) from traditional single-point exhaust gas measurements to an AI-driven, multi-source comprehensive analysis [11]. Using cleaned and integrated data from fuel testing records, AIS trajectories, ship and operator profiles, and supply chain documents for ships calling at the Port of Ningbo, they first employ multiple linear regression to identify key drivers of sulfur emissions, then develop a deep neural network model in TensorFlow to estimate FSC and classify compliance risks in real time. The intelligent monitoring module has been integrated into the port's exhaust detection system, enabling targeted enforcement, more efficient use of inspection resources and a shift from reactive compliance checking toward proactive, risk-based supervision.

Taken together, these works show how optimization models, predictive analytics and novel metrics can underpin maritime decarbonization—from designing CO₂ shipping networks to real-time emission prediction for autonomous vessels, as well as port-level air quality management and fuel compliance monitoring.

The studies in this Special Issue highlight several key trends and directions for the future of maritime digital transformation. One notable trend is the shift from isolated technologies to integrated systems. The combination of sensing, forecasting, and optimization is becoming central to port and shipping operations, forming more cohesive workflows.

Future research should focus on improving interoperability across systems, ensuring that traditional infrastructure can integrate smoothly with emerging digital technologies.

While digitalization and artificial intelligence are critical to driving change, their success relies on the integration of domain expertise. Many of the models discussed are not purely data-driven but incorporate industry-specific knowledge and constraints. For example, in CO₂ emission prediction for vessels, physical, operational, and environmental factors play a vital role. Future work should develop hybrid models that combine deep learning with domain expertise to improve both accuracy and interpretability.

Moreover, digital transformation in maritime transport is not just a technological shift but a social and institutional one. Organizational readiness and governance frameworks are crucial to unlocking the potential of digital technologies. The studies in this Special Issue underscore the importance of aligning stakeholders and creating supportive policy environments to ensure the successful implementation of digital initiatives.

Another key takeaway is the balance between multiple objectives, such as cost, efficiency, safety, and environmental performance. Maritime decision-makers face increasing complexity, with more stringent regulations and fluctuating market conditions. Future research should explore multi-objective optimization frameworks that can handle this complexity while maintaining computational efficiency.

Looking forward, open data ecosystems will play a critical role in maritime digitalization. Ensuring seamless data exchange across systems and protecting privacy will be essential. Additionally, as artificial intelligence takes on more safety-critical roles, addressing its ethical implications and ensuring compliance with regulations will be important. In conclusion, collaboration across ports, shipping companies, technology providers, and policymakers will be key to developing solutions that are technically feasible, economically viable, and socially acceptable, leading to a smarter, greener maritime future.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Shen, Y.; Man, X.; Wang, J.; Zhang, Y.; Mi, C. Truck lifting accident detection method based on improved PointNet++ for container terminals. *J. Mar. Sci. Eng.* **2025**, *13*, 256. [CrossRef]
2. Du, X.; Luo, B.; Wang, J.; Zhao, J.; Li, D.; Sun, Q.; Li, H. Optimizing stack-yard positioning in full shoreline loading operations. *J. Mar. Sci. Eng.* **2025**, *13*, 593. [CrossRef]
3. Shen, Y.; Xuan, B.; Hu, H.; Wu, Y.; Zhao, N.; Yang, Z. A decomposed-ensemble prediction framework for gate-in operations at container terminals. *J. Mar. Sci. Eng.* **2024**, *13*, 45. [CrossRef]
4. Yu, X.; He, C. Collaborative scheduling of yard cranes, external trucks, and rail-mounted gantry cranes for sea–rail inter-modal containers under port–railway separation mode. *J. Mar. Sci. Eng.* **2025**, *13*, 1109. [CrossRef]
5. Jakovlev, S.; Eglynas, T.; Pocevicius, E.; Voznak, M.; Gricius, G.; Jankunas, V.; Jusis, M. Comparative analysis of recurrent vs. temporal convolutional autoencoders for detecting container impacts during quay crane handling. *J. Mar. Sci. Eng.* **2025**, *13*, 1231. [CrossRef]
6. Jović Mihanović, M.; Aksentijević, S.; Tijan, E.; Lenart, G. Digital transformation and business model innovation: Enhancing productivity in the Croatian maritime transport sector. *J. Mar. Sci. Eng.* **2025**, *13*, 999. [CrossRef]
7. Cheng, S.; Hu, J.; Huang, Y.; Hu, Z. Enhancing port shipping synergy through Bayesian network: A case of major Chinese ports. *J. Mar. Sci. Eng.* **2025**, *13*, 1093. [CrossRef]
8. Tao, Y.; Yang, Y.; Du, Y.; Wang, S. Carbon dioxide storage site location and transport assignment optimization for sustainable maritime transport. *J. Mar. Sci. Eng.* **2025**, *13*, 1055. [CrossRef]

9. Ni, J.; Tian, H.; Zhang, K.; Xue, Y.; Xiong, Y. Data-driven multi-scale channel-aligned transformer for low-carbon autonomous vessel operations: Enhancing CO₂ emission prediction and green autonomous shipping efficiency. *J. Mar. Sci. Eng.* **2025**, *13*, 1143. [CrossRef]
10. Bojić, F.; Gudelj, A.; Bošnjak, R. A comprehensive model for quantifying, predicting, and evaluating ship emissions in port areas using novel metrics and machine learning methods. *J. Mar. Sci. Eng.* **2025**, *13*, 1162. [CrossRef]
11. Zhou, F.; Wang, Y.; Zhou, Y. The evolution of ship fuel sulfur content monitoring—From exhaust gas measurement to AI-driven comprehensive analysis. *J. Mar. Sci. Eng.* **2025**, *13*, 1795. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

The Evolution of Ship Fuel Sulfur Content Monitoring—From Exhaust Gas Measurement to AI-Driven Comprehensive Analysis

Fan Zhou ^{1,2,*}, Yuxuan Wang ^{1,2} and Yinghan Zhou ^{1,2}

¹ College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China; 202430310221@stu.shmtu.edu.cn (Y.W.); 202310310201@stu.shmtu.edu.cn (Y.Z.)

² Shanghai Engineering Research Center of Ship Exhaust Intelligent Monitoring, Shanghai 201306, China

* Correspondence: fanzhou_cv@163.com

Abstract: To address the limitations of traditional single-point detection methods in monitoring the sulfur content of ship fuel (FSC), which are inadequate in meeting the regulatory demands of high-traffic ports, this study proposes an integrated analytical approach based on artificial intelligence. This approach synthesizes multi-source heterogeneous data, including historical fuel testing records, Automatic Identification System (AIS) trajectory data, ship and operator profiles, technical specifications, fuel supply chain documentation, fundamental ship attributes and so on. Following rigorous data cleaning and preprocessing procedures, a refined dataset comprising 3046 records collected between 2017 and 2024 from the Port of Ningbo was utilized. Initially, multiple linear regression analysis was conducted to identify key factors influencing sulfur emissions, resulting in an R^2 value of 0.67. Based on these findings, a deep neural network model was developed using TensorFlow to enable real-time estimation of FSC and classification of compliance risk levels. The results indicate that the proposed method exhibits high estimated accuracy and robustness. An AI-based intelligent monitoring module, developed based on this research, has been integrated into the ship exhaust gas detection system at the Port of Ningbo. This module enables real-time analysis of inbound ships and intelligent identification of potentially non-compliant ships, thereby significantly improving the precision and efficiency of port regulatory operations. This study not only contributes to the theoretical framework for ship fuel compliance monitoring but also provides a practical and scalable technical solution for intelligent port governance.

Keywords: fuel sulfur content; multi-source heterogeneous data processing; multiple linear regression analysis; deep learning algorithms; intelligent port supervision system

1. Introduction

In recent years, the global maritime shipping industry has faced unprecedented transformation challenges driven by the dual pressures of trade expansion and environmental concerns [1–4]. As the international shipping network connects the world's major economies and facilitates the vast majority of global trade volume [5], freight transportation demand continues to demonstrate a consistent upward trajectory. However, the sector is increasingly affected by a range of uncertainties, including the rising frequency of extreme climate events [6], navigational restrictions in key international waterways [7], and fluctuations in energy prices [8]. These factors have collectively intensified operational volatility within the industry. With the continuous expansion of global shipping capacity, emissions

from ship fuel combustion—particularly sulfur compounds and greenhouse gases—have become a focal point of international attention [9]. These emissions pose serious threats not only to public health but also to marine and atmospheric ecosystems through long-term cumulative effects. In response to these challenges, the adoption of green shipping practices and intelligent port operations has gained widespread recognition as an essential strategy for advancing sustainable development in the maritime sector.

During maritime operations, ships consume substantial quantities of heavy fuel oil, the combustion of which generates emissions containing sulfur oxides (SO_x), nitrogen oxides (NO_x), carbon monoxide, particulate matter (PM), and carbon dioxide (CO_2)—a recognized greenhouse gas (GHGs) [10–12]. These emissions pose significant threats to both environmental quality and public health. Specifically, SO_x and NO_x can react with atmospheric moisture and oxygen to form acidic compounds, contributing to the formation of acid rain. Acid rain has been associated with long-term degradation of agricultural lands, forest ecosystems, and aquatic habitats. In contrast, PM can penetrate the human respiratory system, thereby increasing the risk of respiratory and cardiovascular diseases. Meanwhile, CO_2 is widely recognized as a major driver of global climate change. In response to these environmental challenges, the International Maritime Organization (IMO) adopted the International Convention for the Prevention of Pollution from Ships and its Annex VI [13]. Under the provisions of this convention, the global sulfur content limit for marine fuels was initially set at 0.5% (m/m) in 2012 and later reduced to 0.35% (m/m) in 2020 [6,7]. Furthermore, the IMO has designated four Emission Control Areas (ECAs), including the Baltic Sea, the North Sea, North American coastal waters, and the Caribbean Sea waters of the United States. Since 2015, the sulfur content limit for fuels used within these ECAs has been restricted to 0.1% (m/m) [14].

In recent years, the IMO has actively promoted the establishment and enhancement of ECAs, with policy frameworks primarily aimed at reducing SO_x , NO_x , and GHGs [15]. The Mediterranean ECA officially came into effect on 1 May 2025, requiring a maximum FSC of 0.1% (m/m). This regulation affects approximately 20% of global maritime trade and is projected to reduce regional sulfur emissions by 80%, thereby significantly improving air quality and marine ecosystems for over 150 million coastal residents. Additionally, the newly designated emission control areas in the Canadian Arctic waters and the Norwegian Sea have received formal approval. Under these regulations, the NO_x Tier III standards will take effect on 1 March 2026, applying to newly constructed ships, while the SO_x limit ($\leq 0.1\%$ (m/m)) will be enforced starting 1 March 2027. These measures are intended to protect the fragile Arctic ecosystem and reduce the impact of black carbon emissions on glacial melting.

The establishment of ECAs has proven to be an effective strategy for mitigating air pollution caused by maritime emissions. Traditionally, maritime authorities have relied on boarding ships and extracting fuel oil samples for laboratory analysis as a method of regulatory enforcement. However, this approach suffers from a lack of targeting, random ship selection, and relatively low operational efficiency. More importantly, it is not applicable for monitoring ships that are in transit. In response to these limitations, researchers have proposed alternative methods that involve measuring ship exhaust plumes to estimate the FSC being used. In 2005, coinciding with the enforcement of the International Convention for the Prevention of Pollution from Ships (MARPOL), Chalmers University of Technology in Sweden launched the Identification of Gross Polluting Ships (IGPS) project [16]. This initiative explored the use of optical and “sniffing” techniques to detect ship plumes and identify ships violating high-sulfur fuel regulations in the Port of Rotterdam. Following this pioneering work, numerous research institutions across Europe and North America conducted similar studies, gradually establishing this field as a key area of research within

maritime and atmospheric environmental science [17,18]. In recent years, East Asia has also seen a surge in ship emission monitoring activities [19]. In major ports across China, South Korea, and Japan, shore-based sniffing systems, unmanned aerial vehicle (UAV)-based monitoring, and infrared spectroscopy technologies have been deployed for real-time detection of sulfur content in marine fuels. These efforts aim to address regional air quality concerns and support compliance with international emission reduction obligations.

The determination of ship FSC based on exhaust gas measurement results—expressed in terms of convertible sulfur oxide emission factors—has increasingly become a crucial technical approach for maritime authorities in enforcing ECA policies. More importantly, the accurate and reliable quantification of sulfur oxide emission factors constitutes a core component in the development of comprehensive ship emission inventories and serves as a foundational basis for assessing the environmental impact of maritime emissions. However, as ECA regulations continue to evolve with increasingly stringent sulfur content limits, and as the shipping industry progresses toward energy efficiency and emission reduction through the adoption of exhaust gas cleaning systems and alternative clean fuels. This study aims to develop a machine learning-based model integrated with multi-source data fusion techniques to analyze the sulfur content of marine fuels, with the objective of improving regulatory efficiency and expanding monitoring coverage in high-traffic ports. In contrast to conventional approaches that predominantly rely on single-point exhaust measurements, this research proposes a systematic integration of AIS trajectories, fueling records, historical compliance data, and real-time exhaust monitoring information, thereby establishing a estimated framework capable of dynamically assessing ship compliance risks. By synthesizing behavioral patterns and historical datasets, the proposed methodology enables the identification of systemic non-compliance trends—such as those associated with specific ship types, fuel suppliers, or shipping routes—ultimately facilitating more targeted and informed regulatory decision-making.

2. Overview of the Main Measurement Methods of Ship Emissions

Ships represent mobile waterborne emission sources, posing greater monitoring challenges compared to stationary or more accessible emission sources such as power plants, volcanoes, and motor vehicles. Based on the type of platform hosting the monitoring equipment, ship exhaust measurement methods can be classified into four primary categories: laboratory-based analysis, shipborne platform monitoring, aerial monitoring via small aircraft or UAV, and shore-based monitoring station observations.

2.1. Laboratory-Based Analysis

Laboratory testing constitutes a foundational methodology for evaluating FSC in maritime exhaust research. By simulating the combustion processes of ship engines under controlled conditions, this approach effectively mitigates meteorological interference and minimizes multi-source measurement errors inherent in field observations, thereby substantially improving data reliability and reproducibility [20]. Standard experimental configurations generally comprise combustion simulation chambers, exhaust sampling systems, and high-precision gas analyzers. Through precise regulation of operational parameters—including temperature, pressure, and airflow—these systems enable accurate replication of real-world ship emission environments. Among the analytical techniques employed, the sulfur-to-carbon ratio (S/C) method, which calculates FSC based on measured concentrations of SO₂ and CO₂, has emerged as a widely accepted standard for laboratory-based sulfur content determination [20]. Research indicates that combustion parameters such as temperature and excess air coefficient exert significant influence on sulfur oxidation rates and SO₂ formation. Elevated temperatures and adequate air supply

typically result in increased SO₂ emissions, whereas the use of low-sulfur fuels or desulfurization agents can effectively curtail such emissions [21]. Furthermore, variations in engine load can alter emission profiles, with low-load operating conditions potentially causing incomplete combustion and consequent increases in SO₂ levels [21].

In recent years, remote sensing technologies have shown promise for FSC estimation. For example, ultraviolet imaging techniques integrated with convolutional neural networks (CNN) enable feature extraction from emission plumes, facilitating non-invasive sulfur content estimation [22]. Additionally, multi-task deep learning frameworks combined with mid-infrared spectroscopy imaging allow for simultaneous monitoring of multiple pollutants—such as SO₂ and CO₂—and inversion of sulfur content data [23]. Nevertheless, these approaches continue to confront challenges including high instrumentation costs, limited model generalizability, and constrained environmental adaptability. Future research should prioritize advancements in sensor miniaturization, improvements in algorithmic robustness, and enhanced laboratory validation of emission characteristics associated with emerging green fuels—such as hydrogen, ammonia, and biofuels—to develop FSC estimation models that align more effectively with the evolving landscape of maritime energy transition.

2.2. Shipborne Platform Monitoring

Shipboard monitoring, which involves direct sampling of exhaust gases from ship exhaust pipes or chimneys, has emerged as a critical methodology for estimating FSC under real-world operational conditions. It plays a vital role in emission characterization studies and compliance verification [24]. The system typically comprises high-temperature and corrosion-resistant sampling tubes, gas analysis modules—including ultraviolet fluorescence SO₂ analyzers and non-dispersive infrared CO₂ analyzers—and integrated data acquisition units. Capable of continuous monitoring across varying engine loads, ship speeds, and sea states, it effectively captures dynamic sulfur emission patterns. This makes it particularly suitable for investigating the combined effects of fuel composition, engine performance, and operational parameters on SO₂ formation.

Field measurement data indicate a strong positive correlation between FSC and exhaust gas SO₂ concentrations, with higher SO₂ emissions typically observed under high-load engine conditions compared to low-load scenarios. While the use of low-sulfur fuels can substantially reduce SO₂ emissions, it may concurrently lead to increased particulate matter emissions, such as black carbon [25]. Ships equipped with desulfurization technologies—such as exhaust gas scrubbing systems—can comply with Emission Control Area (ECA) regulations even when operating on high-sulfur fuels. Moreover, significant variations in emission factors have been observed between offshore and inland fishing ships, which are closely linked to differences in fuel composition and operational parameters [26]. These findings provide valuable insights for refining shipboard monitoring techniques and developing targeted emission control policies.

From a technical application standpoint, portable emission measurement systems (PEMS) serve as the core instrumentation for shipboard monitoring. These systems are capable of capturing exhaust plumes and directly measuring pollutant concentrations, thereby enabling high-precision data acquisition for real-time sulfur content inversion [27]. Nevertheless, challenges remain, including high operational costs, complex installation procedures, demanding maintenance requirements, and the limitation of single-ship monitoring per deployment. These constraints hinder the rapid large-scale adoption of PEMS across extensive ship fleets. Consequently, shipboard monitoring is currently more applicable for in-depth emission characterization and for validating the accuracy of alternative monitoring approaches, rather than for comprehensive regulatory oversight.

2.3. Aerial Monitoring via Small Aircraft or UAV

As an emerging technology for ship exhaust gas monitoring, the unmanned aerial vehicle (UAV) “sniffing” system has gained increasing adoption in real-time FSC estimation and emission regulation. This is attributed to its notable advantages, including high operational flexibility, rapid deployment capability, and relatively low operating costs. The system typically integrates electrochemical SO₂ sensors, non-dispersive infrared CO₂ sensors, sampling units, and data logging modules. By traversing the ship’s exhaust plume at close proximity—typically within a range of 50 to 200 m—and applying sulfur-to-carbon (S/C) ratio algorithms, it enables rapid inversion of FSC. Studies have demonstrated that within a sulfur content range of 0.035% to 0.24%, the UAV-based system can achieve measurement accuracies with deviations below 0.03% [28]. To further improve monitoring efficiency, recent research has concentrated on multi-UAV cooperative path planning and scheduling optimization models, aiming to maximize ship coverage under the constraint of limited battery endurance [29]. Moreover, the integration of ultraviolet and infrared imaging technologies allows the system to capture the spatial distribution of exhaust plumes, thereby supporting emission dispersion modeling and spatiotemporal forecasting of sulfur emissions. Despite these advancements, several challenges remain, including limited adaptability under adverse meteorological conditions, reduced measurement accuracy of lightweight sensors at ultra-low sulfur levels, and constraints imposed by battery capacity. In anticipation of increasingly stringent sulfur emission regulations, further development of sensor calibration algorithms and robust anti-interference mechanisms is essential to enhance the reliability of FSC estimation.

2.4. Shore-Based Monitoring

Shore-based monitoring establishes a stable and cost-efficient methodology for long-term, continuous estimation and assessment of FSC by deploying monitoring equipment at fixed locations such as ports, key waterways, and bridges. This approach passively captures exhaust plumes from passing ships, enabling reliable data acquisition over extended periods. The system typically integrates ultraviolet fluorescence (UVF) SO₂ analyzers and non-dispersive infrared (NDIR) CO₂ analyzers, and is synchronized with the AIS to facilitate ship identification and data linkage. FSC estimation is subsequently performed using sulfur-to-carbon (S/C) ratio calculations. To enhance plume source attribution and dispersion modeling capabilities, supplementary technologies such as video surveillance and meteorological sensors are frequently employed [30]. Long-term observational data reveal distinct seasonal variations and compliance discrepancies in ship emissions within port regions. For instance, Alföldy et al. [31] reported a bimodal distribution of SO₂ emission factors, which they associated with differences in main and auxiliary engine fuel usage and operational switching behaviors. These findings underscore the necessity of incorporating ship operational status and fuel management strategies when interpreting FSC from shore-based measurements. To improve the spatial representativeness of monitoring, several studies have applied Gaussian plume modeling techniques to optimize sensor placement, thereby increasing the likelihood of effective plume detection [32,33]. Furthermore, shore-based systems can be integrated with unmanned aerial vehicle (UAV)-based monitoring to establish a hybrid regulatory framework combining “initial screening” with “on-site verification”, significantly improving compliance monitoring efficiency within emission control areas [34].

2.5. Challenges and Responses

While ship exhaust gas measurement systems offer direct and real-time sulfur oxide emission data, enabling rapid estimation of FSC, their inherent limitations have become

increasingly evident in the context of modern, high-volume port operations. As critical nodes in global trade networks, ports handle massive daily ship traffic, characterized by frequent arrivals and departures of diverse ship types. Conventional exhaust monitoring approaches represent discrete, instantaneous “snapshot” measurements, which are inherently limited in both spatial coverage and statistical representativeness. Given the operational constraints, it is impractical for busy ports to conduct continuous and comprehensive emission monitoring for every berthing or transiting ship, resulting in significant regulatory gaps concerning fuel usage compliance. The sampling-based nature of such monitoring not only restricts its effectiveness in achieving broad coverage but also renders it susceptible to evasion strategies, such as pre-monitoring fuel switching to compliant low-sulfur alternatives. More critically, isolated measurement data lack contextual information and are insufficient to identify underlying patterns, emerging trends, or systemic compliance risks. For instance, these systems cannot effectively address key regulatory questions: Which shipping routes or operators exhibit higher non-compliance rates? Is there widespread fuel quality inconsistency among specific bunkering ports or suppliers? Do certain ship types or older ships demonstrate persistent non-compliance tendencies? In the face of complex and high-density maritime traffic, reliance on single-point monitoring proves inefficient, costly, and inadequate for achieving robust enforcement and targeted regulatory control. There is, therefore, an urgent need for more intelligent, integrated, and comprehensive monitoring frameworks.

To address the inherent limitations of single-point measurement and effectively manage regulatory challenges in high-traffic port environments, the development of AI-driven comprehensive analysis systems represents a necessary evolutionary trajectory. The fundamental principle lies in transcending the constraints imposed by single-source data through the establishment of an intelligent analytical framework capable of integrating multi-dimensional maritime data. This includes the aggregation of heterogeneous datasets such as historical fuel inspection records (encompassing both compliance and violation cases), ship AIS trajectories (including port call history, route patterns, and navigational behavior), ship and company profiles (covering shipowner and manager identities, compliance histories, and corporate risk indicators), technical specifications (engine type, power output, and scrubber installation status), fuel supply chain intelligence (bunkering port origins, supplier reputations, and historical fuel quality metrics), as well as basic ship attributes (such as type, length, gross tonnage, and age). Through the application of advanced machine learning techniques—including supervised learning for risk ship classification, unsupervised learning for anomaly cluster detection, and time series modeling for trend forecasting—AI models can identify latent correlations and behavioral patterns within this complex data ecosystem. This enables the dynamic assessment, accurate estimation, and stratified management of marine fuel compliance risks. For example, the system can autonomously flag high-risk ships that exhibit recurrent refueling at identified high-risk bunkering ports, operate along routes with historically low enforcement rates, and are affiliated with companies possessing prior non-compliance records. These flagged ships can then be prioritized for targeted inspections, thereby optimizing the allocation of regulatory resources. By shifting regulatory oversight from reactive, random “point” sampling to proactive, intelligence-driven “area” analysis, this AI-enhanced approach significantly improves regulatory efficiency, reduces compliance monitoring costs, and ultimately supports the global shipping industry in achieving more sustainable and equitable operational practices.

3. Method

3.1. Dataset

The data employed in this research were sourced from Ningbo Port, recognized as one of the world's largest ports in terms of cargo throughput and serving as a key maritime hub and designated emission control zone along the Chinese coastline. In collaboration with the Ningbo Maritime Safety Administration, we obtained a comprehensive five-year dataset derived from actual ship fuel monitoring procedures. This dataset was compiled through structured enforcement activities, including routine inspections, random fuel sampling, real-time exhaust gas measurements conducted on ships calling at the port and so on. Following a thorough process of data cleaning and preprocessing, a total of 3046 valid records spanning the period from 2017 to 2024 were finalized for modeling purposes.

Following data cleaning and preprocessing procedures, a total of 3046 valid records were retained for modeling purposes, spanning the period from 2017 to 2024. The dataset comprises a wide range of key variables, including ship name, identification number, receiving port, fuel supplier, sampling date, sample seal number, sulfur content limit standard, testing agency, fuel type, original and adjusted sulfur content, gross tonnage, main engine power, ship length, beam, deadweight tonnage, ship type, port of registry, port entry and exit status, sea/river ship classification, reporting institution, port name, berth, and loading/unloading port. The dataset covers over forty distinct types of marine fuel and more than thirty categories of ships, with port of registry information representing over seventy cities across China. This comprehensive dataset provides a robust empirical foundation for accurate sulfur emission estimation and effective compliance monitoring in maritime operations.

3.2. Workflow

As illustrated in Figure 1, This study's research methodology is structured into three sequential phases: the analysis of ship sulfur content-related characteristics, the estimation of FSC based on these characteristics, and the integration of the estimated model into an AI agent framework. Initially, the dataset is constructed using preprocessed data, including cleaned historical records. In this context, FSC is designated as the target variable, while other parameters—such as historical fuel test reports, AIS trajectory data, vessel and operator profiles, technical specifications, fuel supply chain documentation, and fundamental ship attributes—are treated as input features. Key influential factors are identified through regression coefficient magnitudes and statistical significance, including main engine power, ship length, ship width, and gross tonnage. These variables serve as critical indicators influencing sulfur content and provide a theoretical foundation for subsequent estimated modeling. This phase primarily employs a linear regression model to extract features that exhibit statistically significant correlations from extensive empirical datasets.

Building upon the insights gained from the characteristic analysis, a deep learning model is developed using TensorFlow to enable high-precision estimation of ship fuel sulfur content. The model utilizes the previously selected features as input variables and leverages neural networks to capture the complex, nonlinear relationships between these features and FSC. A primary objective of this model is to accurately determine whether the FSC exceeds the regulatory threshold of 0.5% (m/m), thereby serving as a crucial decision-making tool for the development of an intelligent early warning system.

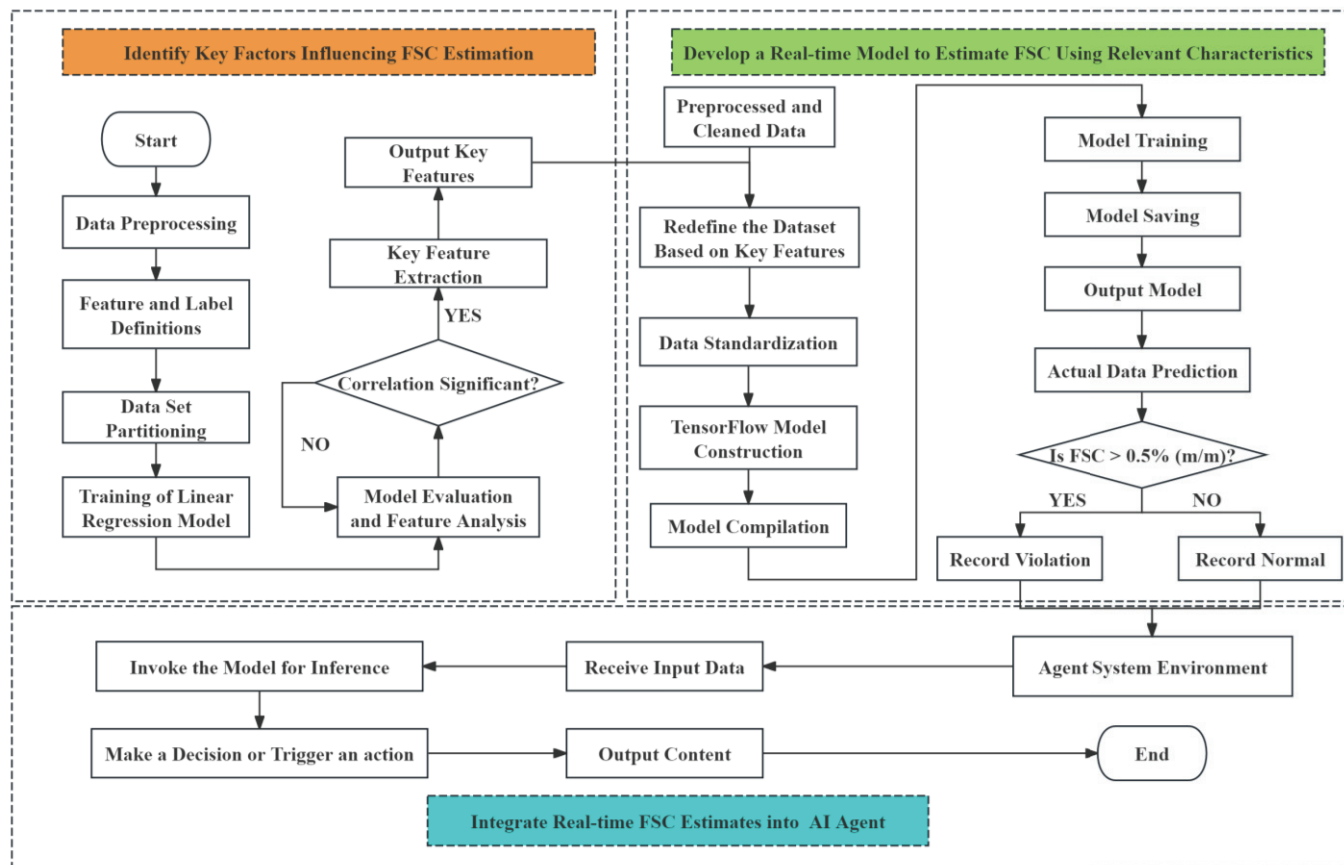


Figure 1. Flowchart of the method. Among them, FSC is the abbreviation of fuel sulfur content.

Finally, the trained estimated model is embedded within an AI agent framework, enabling functionalities such as risk warning, intelligent query handling, and decision support. This integration significantly enhances the intelligence level of maritime emission management. The methodology for correlation analysis of the linear regression model, as well as the structural design and training procedures of the TensorFlow-based estimated model, are elaborated in Sections 3.3 and 3.4, respectively. Based on the aforementioned methods, we conducted experimental evaluations using the dataset described in Section 3.1, and the corresponding results along with detailed analyses are presented in Section 4. Furthermore, the specific implementation plan for integrating this model into an AI agent and using it for FSC estimation in maritime supervision is described in Section 5.

3.3. Multiple Linear Regression

To systematically elucidate the comprehensive influence of multi-source heterogeneous data on marine FSC, this research initially applies Multiple Linear Regression (MLR) methodology to consolidate multidimensional ship-related information into a unified analytical framework. This approach enables quantitative assessment of the statistical relationships between various estimated features and sulfur content levels. In contrast to conventional regression techniques that depend exclusively on isolated static parameters, the dataset employed in this study encompasses diverse dimensions spanning ship operations, management profiles, technical specifications, and historical behavioral patterns, including:

- A. Historical fuel testing data: covering compliance and non-compliance records, reflecting the long-term fuel quality and sulfur emission levels of ships.
- B. Ship AIS trajectory information: including port call records, route history, and navigation patterns, used to capture the operational mode of ships and regulatory exposure risks.
- C. Ship and company profiles: containing background information of the ship owner and management company, operational history, and compliance records, used to assess the potential impact of management level on fuel quality.
- D. Ship technical parameters: such as main engine type and power, whether equipped with a scrubber, propulsion system configuration, etc., directly affecting fuel combustion efficiency and emission levels.
- E. Fuel supply chain information: including fuel supply ports, reputation of fuel suppliers, and historical supply quality, used to identify potential high-risk fuel sources.
- F. Ship basic attributes: such as ship type, length, tonnage, and age, serving as structural indicators influencing fuel consumption and emission patterns.

Before constructing the FSC estimation model for ships, systematic preprocessing of multi-source heterogeneous data is required. For continuous features such as ship length, gross tonnage, and rated power of the main engine, Z-score standardization is adopted to eliminate dimensional differences:

$$z = \frac{x - u}{\sigma} \quad (1)$$

Among them, u is the sample mean and σ is the standard deviation, ensuring that the features follow a distribution with a mean of 0 and a standard deviation of 1. After the above processing, the feature values of all dependent variables are integrated into a unified feature matrix $X_{n \times m}$, where n is the number of samples and m is the feature dimension, corresponding to the target variable (the actual value of FSC) $Y_{n \times 1}$. Based on the feature matrix X and the target variable Y , a multiple linear regression model is constructed:

Then, a regression model is constructed based on the feature matrix X and the target variable Y :

$$\hat{Y}_i = \sum_{j=1}^n \beta_j \cdot X_i + \beta_0 \quad (2)$$

Among them, β_j is the coefficient fitted by regression analysis, X_i is the value of the ship's characteristic information, and β_0 is the intercept term.

Before integrating the above multi-source data, feature engineering processing is required. For the dependent variables (such as ship length, gross tonnage, main engine power, etc.), Z-score standardization is adopted to eliminate the dimensional differences; for qualitative variables (such as ship type, fuel supply port category, etc.), they are transformed into numerical features. After processing, a unified feature matrix X is formed, corresponding to the target variable—the observed value of FSC Y .

By fitting this model and based on the coefficient of determination (R^2) of the regression analysis, the strength of the observed association between the feature and sulfur content can be objectively evaluated, which is unlikely to be caused solely by random sampling error, thereby identifying the truly statistically significant key influencing factors. The calculation method of R^2 in regression analysis is shown in Formula (3):

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \bar{Y})^2} \quad (3)$$

Among them, R^2 is the coefficient of determination, directly indicating the proportion of the variation in the target variable (sulfur content) explained by the model out of the total variation. The closer R^2 is to 1, the stronger the correlation of the model with the

target variable; SSE represents the sum of squared residuals, indicating the unexplained fluctuations by the model; SST is the total sum of squares; Y_i refers to the actual sulfur content observation value of the i -th ship; $\bar{Y}$ refers to the average sulfur content of all ships (sample mean); $\hat{Y}_i$ refers to the sulfur content estimation value of the i -th ship by the model.

Through the analysis of regression coefficients and their significance levels, the influence direction and degree of different dimensional features can be clearly identified. For example, variables such as historical violation rate, frequent refueling from high-risk fuel supply ports, and poor compliance records of the affiliated company may have a significantly positive impact on sulfur content; while ships equipped with desulfurization towers and using low-sulfur fuel may have negative regression coefficients. This analysis not only reveals the role of individual variables but also provides feature selection and weight reference for subsequent sulfur content estimation models based on deep learning.

3.4. Real-Time Estimation and Risk Grading of Sulfur Content Based on TensorFlow

Following the completion of multiple linear regression analysis and the identification of key influencing factors, this research further develops a Deep Neural Network (DNN) model based on the TensorFlow framework to enable real-time estimation of sulfur content in marine fuel and dynamic risk classification. The proposed model effectively captures the intricate nonlinear relationships between multi-source heterogeneous features and sulfur content through multi-layered nonlinear transformations.

This study constructs a deep learning model using TensorFlow (version 2.18.0 under Python 3.10), supported by pandas for structured data manipulation and numpy for numerical computation and array preprocessing. During the training process, a multi-layer fully connected neural network is defined via the Keras API. The Adam optimizer is employed with mean squared error as the loss function. The resulting high-precision estimated model is saved in HDF5 format for deployment.

As shown in Figure 2, the model consists of an input layer, a hidden layer, and an output layer. The input layer comprises multi-dimensional ship feature variables, including fundamental parameters such as gross tonnage, main engine power, ship beam (width), and overall length. Following standardization preprocessing, these normalized features are fed into the first layer of the neural network. The hidden layer is structured as a three-layer fully connected network, with each layer comprising multiple neuron nodes designed to extract hierarchical feature representations. The first layer primarily focuses on basic construction features, such as the geometric dimensions and gross tonnage of ships. The second layer extracts performance characteristics closely associated with the operational status of ships, including main engine power and operational efficiency. The third layer captures potential emission-related characteristics, reflecting the emission potential and elucidating the comprehensive influence of various technological and operational factors on sulfur emissions.

The output layer consists of a single neuron responsible for generating the estimated sulfur content of marine fuel. This estimation serves not only as a basis for real-time estimation but also supports risk classification when integrated with threshold criteria. Consequently, ships can be categorized into three risk levels: low-risk, medium-risk and high-risk targets.

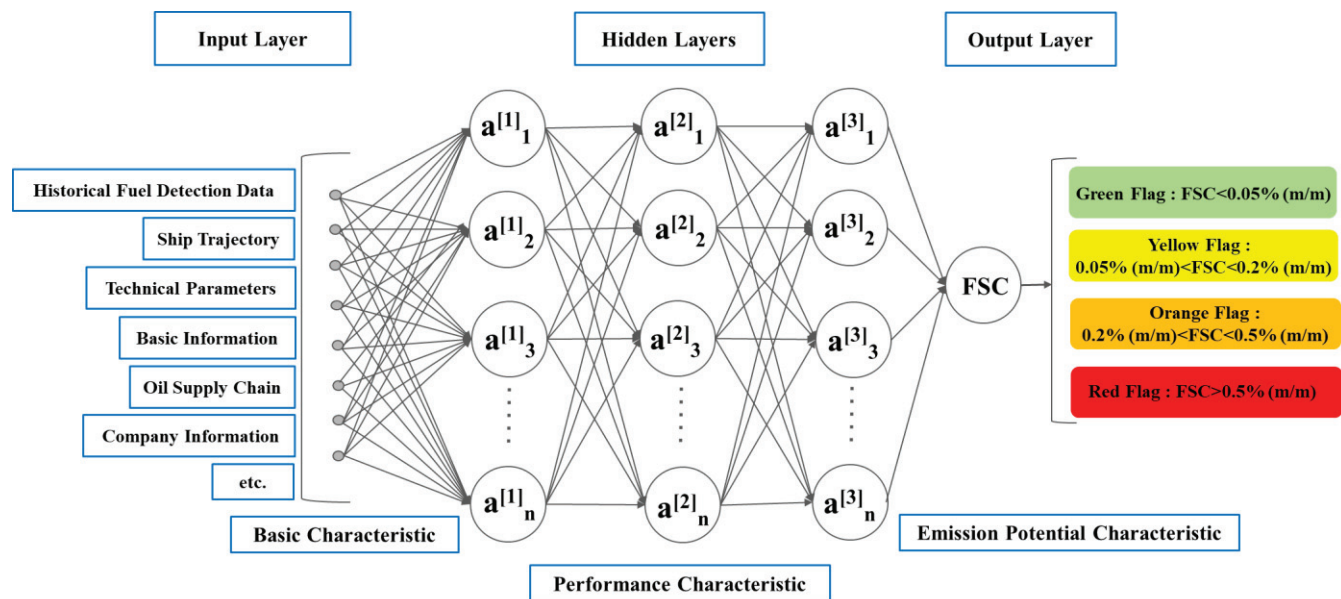


Figure 2. Real-time estimation of sulfur content and risk grading based on TensorFlow.

During the model training phase, the input feature dataset is divided into training, validation, and test sets in a certain proportion to ensure the generalization ability of the model. The estimated sulfur content values generated by the model can be further evaluated against regulatory thresholds to establish a risk classification framework. Specifically, ships with estimated values significantly below the limit are categorized as low-risk; those with values approaching the threshold are classified as medium-risk (subjects of heightened scrutiny); and those exhibiting values exceeding or nearing non-compliance levels are designated as high-risk, and are prioritized for inclusion in inspection programs. Furthermore, by integrating regression coefficients with feature importance analysis from the neural network—such as gradient-based interpretability techniques—key contributing factors to high-risk classifications can be identified, such as refueling activities at high-risk supply ports or a management company’s history of regulatory violations. This provides regulatory agencies with a transparent and traceable analytical basis for decision-making.

4. Results

Figure 3 presents a systematic overview of the multi-source data framework employed in this study for monitoring and estimating sulfur content in marine fuels. The framework organizes the data into six distinct categories: basic ship attributes, port activities, fuel supply history, technical specifications, compliance records, and historical fuel test data. Each category contributes unique insights into various aspects of emission behavior analysis. Collectively, these features facilitate a comprehensive evaluation of ship operational patterns, regulatory exposure, fuel combustion characteristics, and operator reliability. The integration of these diverse and heterogeneous data sources constitutes the foundational basis for the AI-driven model, thereby improving its precision and real-world applicability in detecting non-compliant ships under complex maritime operational conditions.

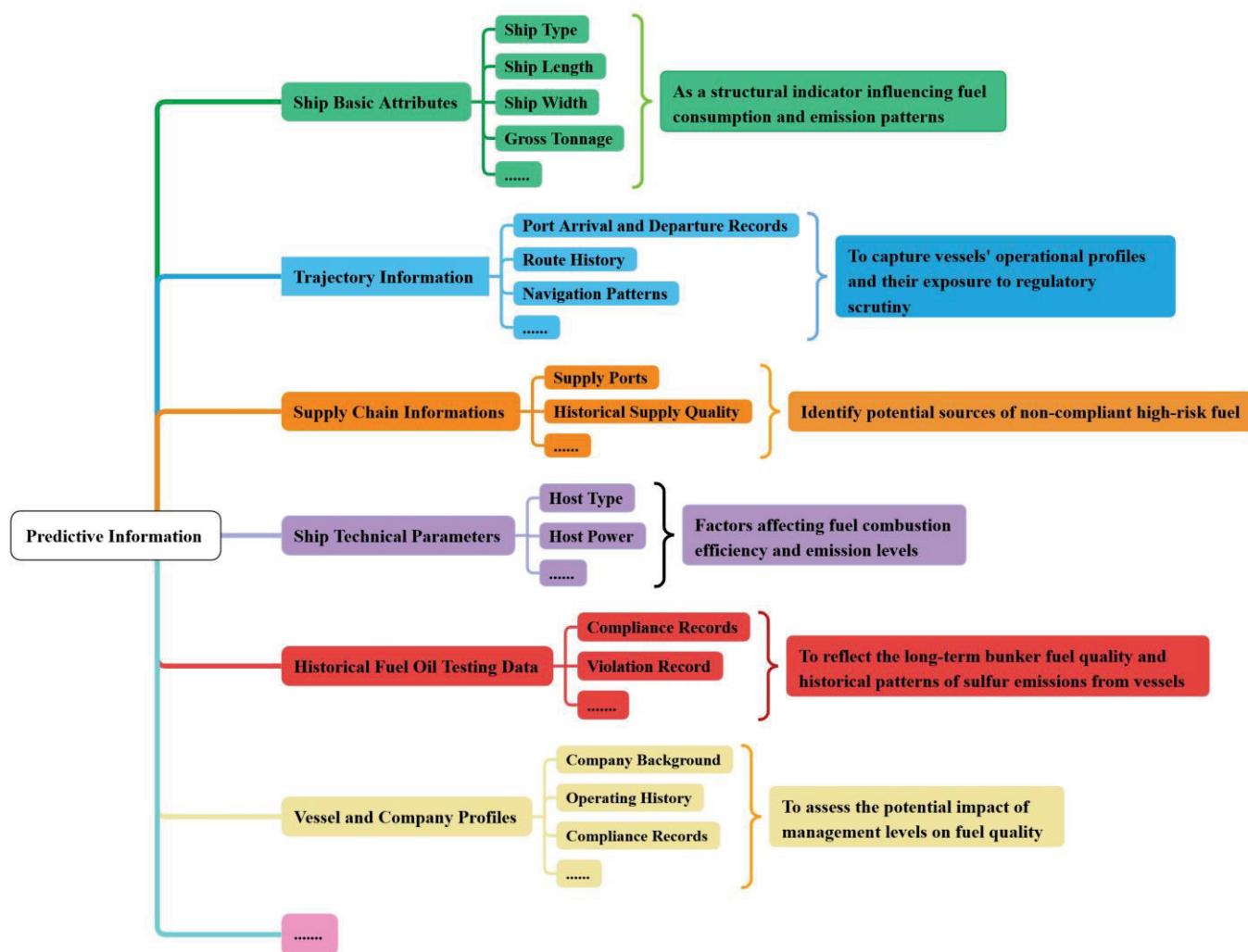


Figure 3. Analytical data chart.

The dataset utilized in this research encompasses sulfur content measurements and operational records of ships across multiple ports from 2017 to 2024, comprising a total of 3046 data entries. It includes a diverse range of ship types and offers notable advantages in terms of temporal span, broad geographical coverage, and multi-dimensional feature representation. The dataset effectively captures the compliance dynamics before and after the enforcement of the IMO global sulfur cap in 2020. It spans from small- and medium-sized ships to ultra-large ocean-going ships, as well as ships ranging from newly constructed to those exceeding 20 years in age. In terms of fuel quality assessment, sulfur content values range from 0.01% (m/m) to 3.5% (m/m). Following 2020, a substantial decline in sulfur levels has been observed, with most values concentrated within the 0.01% (m/m)–0.5% (m/m) range. Nevertheless, non-compliant or near-threshold samples persist, indicating ongoing compliance risks. Fuel supply chain information reveals variability in compliance performance across ports and suppliers, with certain ports and providers linked to a higher incidence of violations, thereby exhibiting elevated risk profiles. Additionally, port entry/exit and cargo loading/unloading records allow for more precise characterization of ship movement patterns. Data sources include both official regulatory authorities and independent third-party testing institutions, ensuring data diversity and enabling cross-validation. Overall, the dataset demonstrates strong representativeness in terms of scale, feature richness, and temporal relevance. It not only facilitates the identification of key drivers influencing sulfur emissions through multiple linear regression analysis but also

provides comprehensive input features and real-world risk scenarios for training deep learning models using TensorFlow. This robust data foundation supports the development of an intelligent and dynamic monitoring system for ship FSC.

As illustrated in Figure 4, the impact of various ship characteristics on sulfur emissions demonstrates notable variation across different years. Ship dimensional parameters—such as main engine power, ship width, ship length, and gross tonnage—exhibit consistently strong correlations over multiple years, particularly in 2018 and 2024, where their regression coefficients reach relatively high levels of significance. This suggests that physical dimensions and propulsion capacity serve as key explanatory variables for sulfur emission levels. In comparison, operational and managerial factors—including fuel suppliers, upstream and downstream ports, and the port of registry—exert a comparatively limited overall influence on sulfur emissions, with their statistical significance showing temporal fluctuations. Overall, structural ship parameters constitute the primary determinants of sulfur emissions, whereas operational and managerial variables may function as secondary influencing factors under specific temporal or contextual conditions.

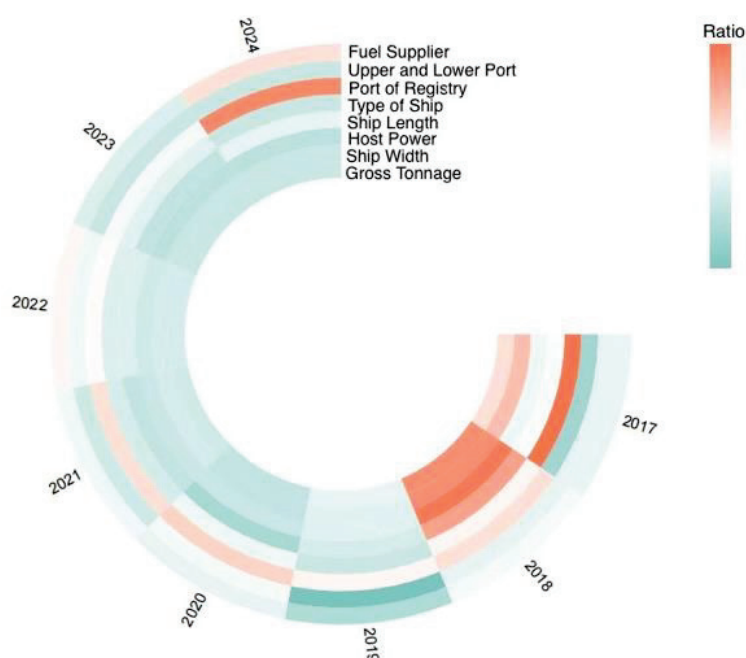


Figure 4. Regression analysis results chart.

This research is based on an eight-year continuous dataset (2017–2024) of sulfur content in marine fuels. A hybrid estimation framework integrating multiple linear regression and deep learning models has been developed to estimate fuel sulfur levels, with annual comparisons conducted between estimated and empirically measured values. During the experimental phase, the model incorporates multi-source heterogeneous input features, including fundamental ship parameters, fuel supply chain data, historical compliance records, and operational trajectory information, generating corresponding sulfur content estimations for each fuel sample. To assess model performance, statistical metrics such as the coefficient of determination (R^2) and mean squared error (MSE) are employed to evaluate the degree of fit between estimated and observed values. Furthermore, independent modeling and validation are performed on yearly subsets of the dataset to investigate the model's adaptability and robustness across policy transitions and varying market conditions.

The experimental results indicate that the model achieved high fitting accuracy throughout the 2017–2024 period, with estimated values demonstrating strong statisti-

cal consistency with observed measurements. This suggests that the developed model possesses substantial explanatory capability regarding sulfur emissions from marine fuels. During the policy transition phase (2017–2019), the model maintained stable estimation accuracy, effectively capturing the influence of fuel usage characteristics on emission levels. Following the enforcement of the IMO global sulfur cap in 2020, the discrepancy between estimated and measured values declined significantly, highlighting the positive impact of policy-driven fuel transition on model performance. The discrepancy between estimated and observed values in 2020 was reduced by 21.16% compared to that in 2019, indicating an improvement in model accuracy over the consecutive years. A moderate fluctuation in estimated accuracy was observed between 2021 and 2022, which may be attributed to variations in the pace of clean fuel adoption and alternative energy deployment across the market. By 2023–2024, the model's estimations once again aligned closely with actual measurements, demonstrating its adaptability to evolving market dynamics and technological advancements. The discrete pattern observed in the 2024 data presented in Figure 5 can be primarily attributed to two key factors. First, the dataset obtained from the Ningbo Maritime Safety Administration only includes records up to April 2024, resulting in a relatively limited sample size. This constraint may compromise the stability and reliability of the statistical outcomes for that year. Second, the implementation of stricter emission control policies, along with the widespread adoption of low-sulfur fuels, has led to a substantial decrease in actual FSC values in 2024 compared to previous years. Specifically, the upper limit of the FSC value on the 2024 axis is set at 0.15% (m/m), whereas the upper limit in the earlier figures was 0.5% (m/m). Although the 2024 data exhibit greater variability, the absolute deviation between the estimated and observed values remains minimal. This consistency indicates that the model retains a high level of estimated accuracy, even amidst rapidly evolving regulatory frameworks. Overall, this study confirms the temporal applicability and robustness of the estimated approach integrating multi-dimensional data with a deep learning architecture, offering a reliable foundation for real-time monitoring and regulatory enforcement of sulfur content in ship fuel.

The experimental results demonstrate that the developed model achieves high fitting performance and estimated accuracy across multi-year datasets spanning from 2017 to 2024. It effectively captures the integrated influence of ship structural parameters, operational behaviors, and fuel supply chain factors on sulfur emissions. Notably, following the enforcement of the global sulfur cap regulation in 2020, the discrepancy between estimated and observed values substantially decreased, confirming the method's applicability and robustness under evolving policy frameworks and dynamic market conditions. Furthermore, the findings indicate that ship-scale parameters—such as gross tonnage, main engine power, length, and beam—exhibit a stronger explanatory capacity for sulfur emission levels, whereas managerial and operational dimensions exert a moderating influence within specific timeframes, offering empirical support for future policy development and regulatory enhancement.

Nonetheless, certain limitations persist in this study. The multi-source data underpinning the model—including AIS trajectories, fuel reports, and company records—vary in terms of collection timing, measurement accuracy, and spatial coverage, potentially introducing noise and bias. Particularly in cases of missing data or inconsistent labeling, estimation outcomes may be compromised. Although the features selected through multiple linear regression analysis demonstrate statistical significance ($R^2 = 0.67$), they do not encompass all potential influencing factors, such as crew operational practices, real-time meteorological conditions, or variability in fuel blending processes. These omissions partially constrain the model's comprehensiveness and generalization capability.

Additionally, external uncontrollable variables—such as abrupt policy shifts or fuel market volatility—may also lead to estimation deviations during exceptional periods.

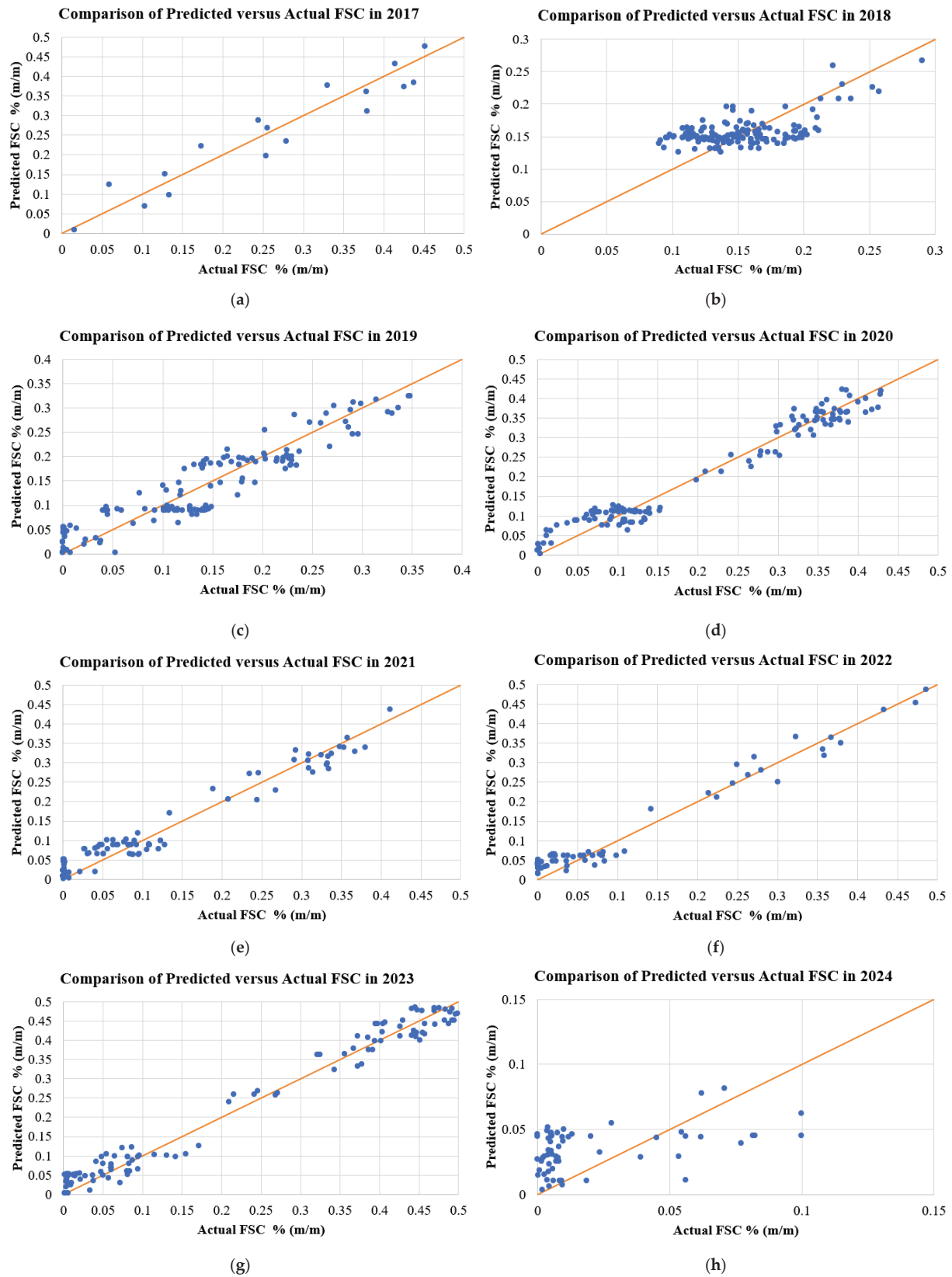


Figure 5. (a–h) Experimental results graph, based on an eight-year continuous dataset (2017–2024). The blue dots represent the actual FSC in comparison to the estimated values. The orange line denotes the $y = x$ reference line for visual interpretation.

5. Application

Figure 6 illustrates the workflow of an AI-driven intelligent supervision framework centered on a multimodal intelligent agent. A centralized database is constructed by collecting multi-source data—including tidal patterns, meteorological conditions, weather information, and ship records—through various devices such as drones, radars, and shore-based sniffers. This database is further enriched with relevant laws and regulations concerning ship sulfur content and information on law enforcement personnel. At the core of the system lies a multi-modal intelligent agent, which interfaces with the database to integrate and analyze the aforementioned multi-source heterogeneous data. Leveraging artificial intelligence technologies, including machine learning and natural language processing, the agent performs intelligent data analysis and risk warning. Based on the analytical outcomes, the system generates optimized supervision plans, continuously evaluates the effectiveness of current policies, and dynamically adjusts regulatory strategies. Ultimately, the intelligent agent produces structured outputs such as reports, archives, and records, along with risk estimations and intelligent risk identification results.

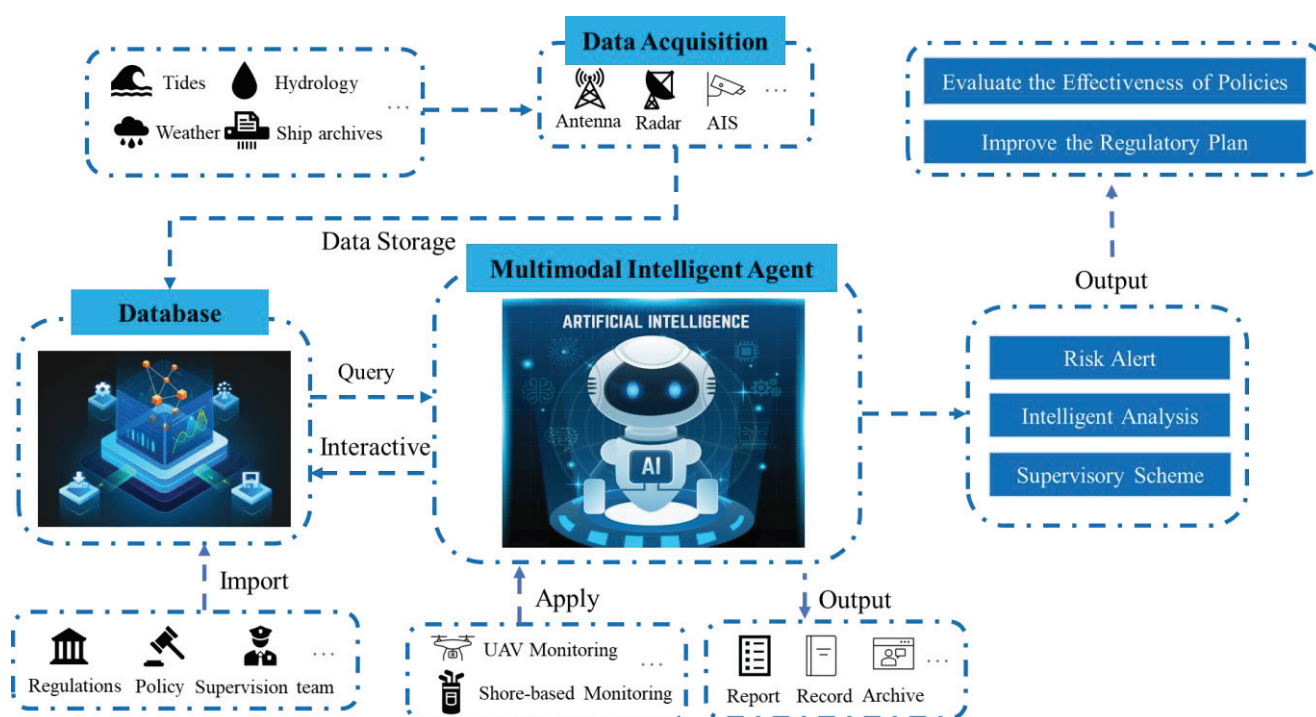


Figure 6. Logical architecture diagram of the agent system.

Based on the experimental results and model validation mentioned above, this study further established an AI agent in the ship exhaust gas detection system of Ningbo Port to achieve real-time monitoring and dynamic risk assessment of the sulfur content in ship fuel. This agent is driven by multi-source heterogeneous data, inputting fuel testing records, AIS track information, ship technical parameters, and fuel supply chain information into a deep learning model. Combined with historical compliance and estimation results, it outputs the estimated FSC and the risk level. During operation, the system can automatically screen ships entering and leaving the port, quickly identify potential high-risk target ships, and generate a priority list for reference by regulatory authorities for spot checks, thereby effectively improving the allocation efficiency of detection resources and the accuracy of law enforcement. The deployment of this AI agent has verified the feasibility and practicality

of the research method in actual port supervision scenarios, providing important technical support for promoting intelligent and refined ship emission control.

This intelligent agent is deployed within the ship exhaust monitoring system of Ningbo Port, integrating multi-source heterogeneous data with deep learning models to enable real-time estimation and risk assessment of FSC. The system performs comprehensive analysis on incoming ships by synthesizing fuel inspection records, AIS trajectory data, fuel supply chain information, and ship technical specifications, generating estimated sulfur content levels and corresponding risk classifications. Additionally, the system visually displays ship distribution on a map interface, allowing the intelligent agent to automatically identify and flag ships suspected of exceeding sulfur limits, thereby supporting regulatory authorities in accurately targeting high-risk ships. This application marks a transition from traditional spot-checking to intelligent estimated monitoring, significantly improving the scientific rigor and operational efficiency of port supervision.

The AI agent deployed in the ship exhaust gas detection system at Ningbo Port has been put into practical use as shown in Figures 7 and 8. Its functions are mainly reflected in two aspects: First, comprehensive analysis. The agent can conduct multi-dimensional feature fusion analysis on the monitored ships, integrating fuel detection data, AIS trajectories, ship attributes, and fuel supply chain information, etc., to output the estimated sulfur content of the ship's fuel in real time. It also classifies ships based on historical compliance records, assisting regulatory authorities in targeted control. Second, intelligent screening. The system visually presents the distribution of ships within the port area on the map interface. The AI agent can automatically mark and screen out ships suspected of having sulfur content exceeding standards, indicating potential violations, thereby providing decision-making references for on-site law enforcement. Through this application, the AI agent has effectively enhanced the intelligence level of port supervision, transforming the traditional passive spot-checking model into proactive estimation and precise targeting, providing practical technical support for efficient and compliant supervision in large-scale port environments.

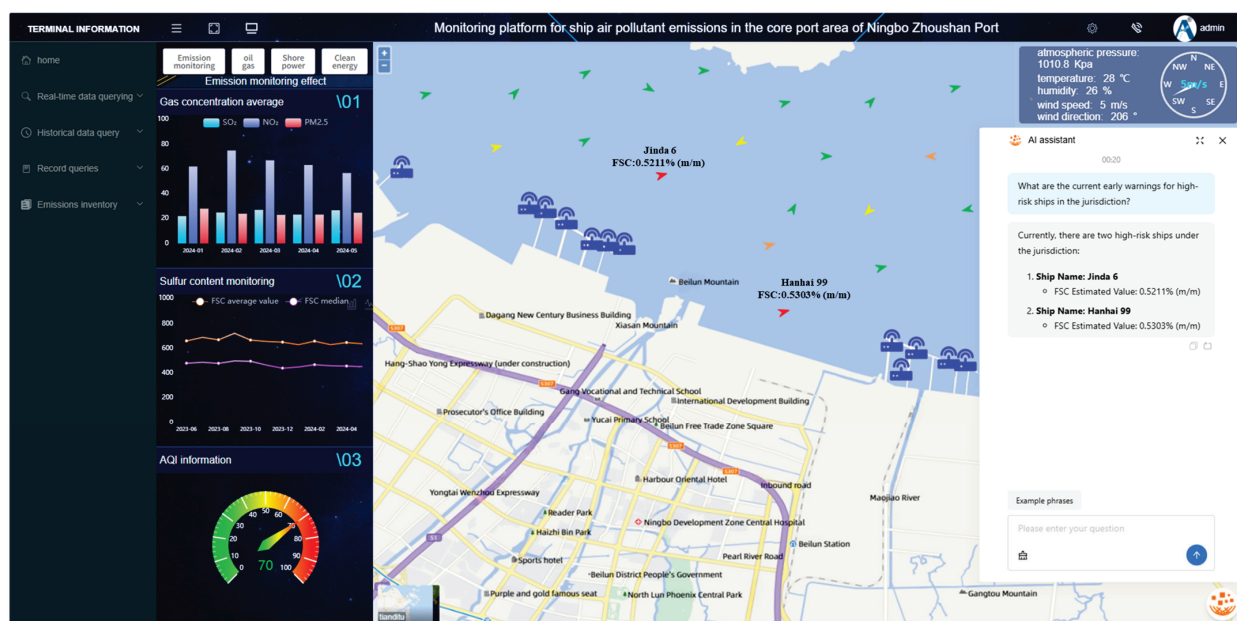


Figure 7. Application diagram of intelligent agents. The green arrows in the figure denote ships with FSC below 0.05% (m/m); yellow arrows correspond to ships with FSC ranging from 0.05% (m/m) to 0.2% (m/m); orange arrows indicate ships with FSC between 0.2% (m/m) and 0.5% (m/m); and red arrows represent ships with FSC exceeding 0.5% (m/m).

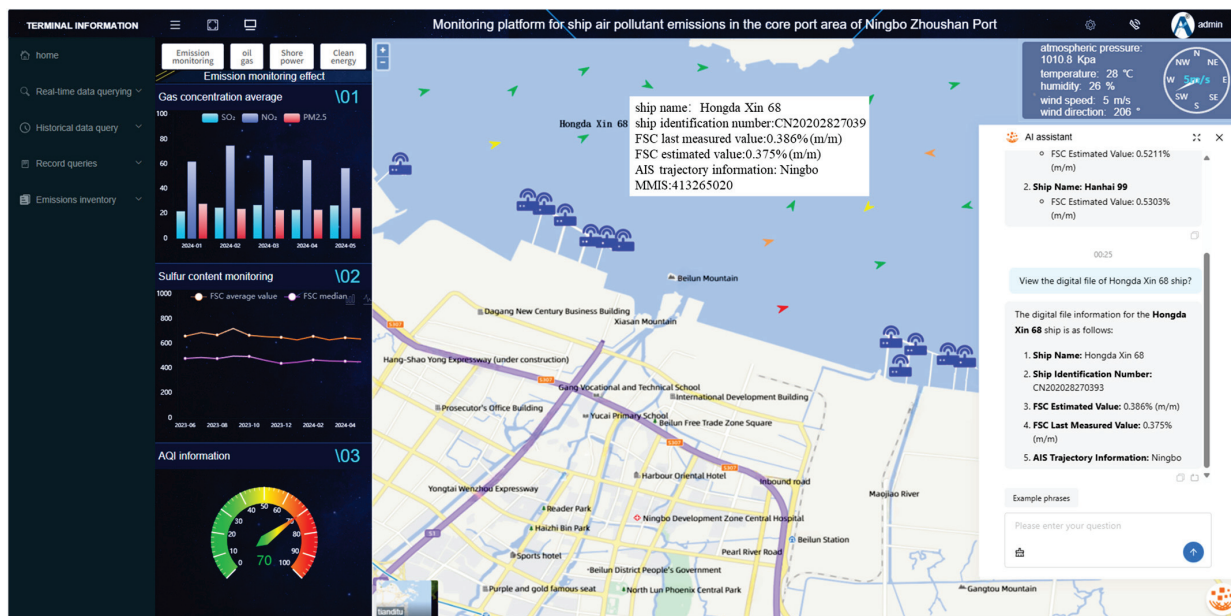


Figure 8. Schematic diagram of agent system application. The green arrows in the figure denote ships with FSC below 0.05% (m/m); yellow arrows correspond to ships with FSC ranging from 0.05% (m/m) to 0.2% (m/m); orange arrows indicate ships with FSC between 0.2% (m/m) and 0.5% (m/m); and red arrows represent ships with FSC exceeding 0.5% (m/m).

6. Discussion

This study successfully achieved high-precision estimation of sulfur content in ship fuel and intelligent classification of compliance risks through the integration of multi-source heterogeneous data using linear regression techniques, as well as the development of a deep neural network model based on TensorFlow. Compared with conventional single-detection approaches, the AI-driven methodology proposed in this paper not only demonstrated substantial explanatory capacity in terms of goodness of fit ($R^2 = 0.67$), but also validated its practical effectiveness within the regulatory framework of Ningbo Port. By systematically integrating ship-related information, the approach enhances the robustness and accuracy of sulfur emission estimation under the synergistic influence of multi-dimensional features. Notably, following the full enforcement of the global sulfur limit regulation, the model maintained strong estimated stability even under extreme low-sulfur conditions, underscoring the distinct advantages of artificial intelligence in complex maritime regulatory environments.

Therefore, future maritime supervision in port areas must go beyond the limitations of traditional exhaust monitoring approaches by integrating multi-source data—including ship dynamic behavior, onboard equipment status, fuel traceability information, and remote sensing of exhaust emissions—to develop an intelligent risk assessment model. This model is expected to not only accurately identify ships with non-compliant fuel sulfur levels but also optimize the allocation of enforcement resources, thereby enabling “low-interference, high-efficiency” compliance management within emission control zones and ultimately contributing to improved air quality in global port cities.

Despite the promising outcomes achieved in both theoretical and applied dimensions, several limitations warrant further investigation. For instance, the current model does not yet account for real-time meteorological conditions or inland waterway traffic dynamics—factors that may influence emission behaviors. Additionally, its cross-regional generalization capability requires further validation. Future research should aim to develop a more comprehensive feature engineering framework, incorporate uncertainty quantification techniques to improve model resilience under abnormal or extreme conditions,

and explore the potential of federated learning for multi-port collaborative governance, thereby supporting the global advancement of intelligent and interconnected emission control systems.

7. Conclusions

Against the backdrop of the growing trend toward ultra-low emissions from ship exhaust, intelligent monitoring of FSC has emerged as a critical research direction for enhancing maritime regulatory capabilities and addressing environmental pollution challenges. The development of an intelligent risk assessment model that integrates multi-source real-time data not only facilitates the accurate identification of non-compliant ships and the optimization of law enforcement resource allocation, but also provides core technical support for achieving “low-interference, high-efficiency” compliance management within emission control areas. This advancement holds significant theoretical value and practical relevance in promoting sustained improvements in air quality across global ports.

This study proposes an AI-driven approach for monitoring ship FSC through the integration of multi-source heterogeneous data, overcoming the limitations of conventional single-point detection and static feature analysis. Key findings demonstrate that multiple linear regression can effectively identify key influencing factors of sulfur emissions ($R^2 = 0.67$). Furthermore, a deep neural network implemented using TensorFlow enables real-time estimation and risk classification of sulfur content, exhibiting high model accuracy and stability. In practical implementation, the research team successfully deployed an AI-based intelligent supervision module at Ningbo Port, enabling real-time monitoring, risk stratification, and automated screening of non-compliant ships entering and exiting the port. This has significantly enhanced regulatory efficiency and accuracy, offering a scalable and replicable technical framework for the field.

Despite notable theoretical and practical advancements, this research presents certain limitations. For instance, the model does not fully account for dynamic variables such as real-time meteorological and hydrological conditions, variations in crew operational behaviors, or inconsistencies in fuel mixing uniformity. Additionally, it does not comprehensively address the challenges posed by regional policy differences that may affect model generalization. Future research will focus on strengthening feature engineering and uncertainty quantification to enhance model robustness in complex environments; integrating multimodal data—including satellite remote sensing, real-time supply chain dynamics, and atmospheric diffusion parameters—to develop a more comprehensive intelligent perception system; and exploring cross-domain collaborative monitoring mechanisms to support the integrated and intelligent governance of global shipping emissions.

Author Contributions: F.Z.: data curation, conceptualization, methodology, writing—original draft, writing—review and editing. Y.W.: data curation, methodology, visualization, investigation. Y.Z.: data curation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Hubei Engineering Research Center for Intelligent Detection and Identification of Complex Parts (grant number KYQDJF2025006). We thank the support of the Haifeng International Young Faculty Industry-Academia-Research Assistance Program.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zaman, I.; Pazouki, K.; Norman, R.; Younessi, S.; Coleman, S. Challenges and opportunities of big data analytics for upcoming regulations and future transformation of the shipping industry. *Procedia Eng.* **2017**, *194*, 537–544. [CrossRef]
2. Du, S.; Zhang, H.; Kong, Y. Sustainability Implications of the Arctic Shipping Route for Shanghai Port Logistics in the Post-Pandemic Era. *Sustainability* **2023**, *15*, 16017. [CrossRef]
3. Andersson, K.; Baldi, F.; Brynolf, S.; Lindgren, J.F.; Granhag, L.; Svensson, E. Shipping and the Environment. In *Shipping and the Environment: Improving Environmental Performance in Marine Transportation*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 3–27.
4. Ali, I. The world's maritime industry in the 21st century: Challenges, expectations, and directions. *South East Asian Mar. Sci. J.* **2025**, *2*, 64–75. [CrossRef]
5. Zou, Y.; Xiao, G.; Li, Q.; Biancardo, S.A. Intelligent Maritime Shipping: A Bibliometric Analysis of Internet Technologies and Automated Port Infrastructure Applications. *J. Mar. Sci. Eng.* **2025**, *13*, 979. [CrossRef]
6. Panahi, R.; Ng, A.K.; Afenyo, M.K.; Haeri, F. A novel approach in probabilistic quantification of risks within the context of maritime supply chain: The case of extreme weather events in the Arctic. *Accid. Anal. Prev.* **2020**, *144*, 105673. [CrossRef]
7. Elmi, Z.; Singh, P.; Meriga, V.K.; Goniewicz, K.; Borowska-Stefańska, M.; Wiśniewski, S.; Dulebenets, M.A. Uncertainties in liner shipping and ship schedule recovery: A state-of-the-art review. *J. Mar. Sci. Eng.* **2022**, *10*, 563. [CrossRef]
8. Chen, Z.; Zhang, X.; Chai, J. The dynamic impacts of the global shipping market under the background of oil price fluctuations and emergencies. *Complexity* **2021**, *2021*, 8826253. [CrossRef]
9. Zannis, T.C.; Katsanis, J.S.; Christopoulos, G.P.; Yfantis, E.A.; Papagiannakis, R.G.; Pariotis, E.G.; Vallis, A.G.; Rakopoulos, D.C.; Rakopoulos, C.D. Marine exhaust gas treatment systems for compliance with the IMO 2020 global sulfur cap and tier III NOx limits: A review. *Energies* **2022**, *15*, 3638. [CrossRef]
10. Wang, T.; Xiao, G.; Li, Q.; Biancardo, S.A. The impact of the 21st-Century Maritime Silk Road on sulfur dioxide emissions in Chinese ports: Based on the difference-in-difference model. *Front. Mar. Sci.* **2025**, *12*, 1608803. [CrossRef]
11. Xiao, G.; Caleb, A.; Wang, T.; Li, Q.; Biancardo, S.A. Evaluating the impact of ECA policy on sulfur emissions from the five busiest ports in America based on difference in difference model. *Front. Mar. Sci.* **2025**, *12*, 1609261. [CrossRef]
12. Chen, J.; Wan, Z.; Zhang, H.; Liu, X.; Zhu, Y.; Zheng, A. Governance of shipping emission of SOx in China's coastal waters: The SECA policy, challenges, and directions. *Coast. Manag.* **2018**, *46*, 191. [CrossRef]
13. Karim, M.S. Prevention of pollution of the marine environment from vessels. *J. Mar. Poll* **2015**, *30*, 50–75.
14. Zetterdahl, M.; Moldanova, J.; Pei, X.; Pathak, R.K.; Demirdjian, B. Impact of the 0.1% fuel sulfur content limit in SECA on particle and gaseous emissions from marine vessels. *Atmos. Environ.* **2016**, *145*, 338–345. [CrossRef]
15. Browne, S.; Weiss, T.G. The UN Development System: Origins, Structure, Status. In *Routledge Handbook on the UN and Development*; Routledge: Oxford, UK, 2020; pp. 9–31.
16. Van Roy, W.; Van Nieuwenhove, A.; Scheldeman, K.; Van Roozendaal, B.; Schallier, R.; Mellqvist, J.; Maes, F. Measurement of sulfur-dioxide emissions from ocean-going vessels in Belgium using novel techniques. *Atmosphere* **2022**, *13*, 1756. [CrossRef]
17. Rumin, J.; Nicolau, E.; Gonçalves de Oliveira Junior, R.; Fuentes-Grünwald, C.; Flynn, K.J.; Picot, L. A bibliometric analysis of microalgae research in the world, Europe, and the European Atlantic area. *Mar. Drugs* **2020**, *18*, 79. [CrossRef]
18. Moshui, A.M.; Mohammad, R.; Hira, F.A.; Maarop, N. Alternative marine fuel research advances and future trends: A bibliometric knowledge mapping approach. *Sustainability* **2022**, *14*, 4947. [CrossRef]
19. Shaaban-Nejad, S.; Shirazi, F. ICT and Environmental Sustainability: A Comparative Study. *Sustainability* **2022**, *14*, 8651. [CrossRef]
20. Anand, A.; Wei, P.; Gali, N.K.; Sun, L.; Yang, F.; Westerdahl, D.; Zhang, Q.; Deng, Z.; Wang, Y.; Liu, D.; et al. Protocol development for real-time ship fuel sulfur content determination using drone based plume sniffing microsensor system. *Sci. Total Environ.* **2020**, *744*, 140885. [CrossRef]
21. Ma, Z.; Yang, Y.; Sun, P.; Duan, S.; Qu, H.; Zou, Y. Analysis of marine diesel engine emission characteristics of different power ranges in China. *Atmosphere* **2021**, *12*, 1108. [CrossRef]
22. Cao, K.; Zhang, Z.; Li, Y.; Zheng, W. Ship fuel sulfur content prediction based on convolutional neural network and ultraviolet camera images. *Environ. Pollut.* **2021**, *273*, 116501. [CrossRef]
23. Cao, K.; Zhang, Z.; Li, Y.; Zheng, W. Surveillance of ship emissions and fuel sulfur content based on imaging detection and multi-task deep learning. *Environ. Pollut.* **2021**, *288*, 117698. [CrossRef]
24. Peng, Z.; Ge, Y.; Tan, J.; Fu, M.; Wang, X.; Chen, M.; Yin, H.; Ji, Z. Emissions from several in-use ships tested by portable emission measurement system. *Ocean Eng.* **2016**, *116*, 260–267. [CrossRef]
25. Zhang, F.; Chen, Y.; Tian, C.; Lou, D.; Li, J.; Zhang, G.; Matthias, V. Emission factors for gaseous and particulate pollutants from offshore diesel engine vessels in China. *Atmos. Chem. Phys.* **2016**, *16*, 6319–6334. [CrossRef]
26. Zhang, F.; Chen, Y.; Chen, Q.; Feng, Y.; Shang, Y.; Yang, X.; Gao, H.; Tian, C.; Li, J.; Zhang, G.; et al. Real-world emission factors of gaseous and particulate pollutants from marine fishing boats and their total emissions in China. *Environ. Sci. Technol.* **2018**, *52*, 4910–4919. [CrossRef] [PubMed]

27. Juwono, A.M.; Johnson, G.R.; Mazaheri, M.; Morawska, L.; Roux, F.; Kitchen, B. Investigation of the airborne submicrometer particles emitted by dredging vessels using a plume capture method. *Atmos. Environ.* **2013**, *73*, 112–123. [CrossRef]
28. Zhou, F.; Pan, S.; Chen, W.; Ni, X.; An, B. Monitoring of compliance with fuel sulfur content regulations through unmanned aerial vehicle (UAV) measurements of ship emissions. *Atmos. Meas. Tech.* **2019**, *12*, 6113–6124. [CrossRef]
29. Zhang, S.; Zeng, Y.; Zhang, R. Cellular-enabled UAV communication: A connectivity-constrained trajectory optimization perspective. *IEEE Trans. Commun.* **2018**, *67*, 2580–2604. [CrossRef]
30. Dou, Z.; Liu, Z.; Li, L.; Zhou, H.; Wang, Q.; Zhang, J.; Chen, L. Atmospheric dispersion prediction of accidental release: A review. *Emerg. Manag. Sci. Technol.* **2022**, *2*, 1–20. [CrossRef]
31. Alföldy, B.; Lööv, J.B.; Lagler, F.; Mellqvist, J.; Berg, N.; Beecken, J.; Weststrate, H.; Duyzer, J.; Bencs, L.; Horemans, B.; et al. Measurements of air pollution emission factors for marine transportation in SECA. *Atmos. Meas. Tech.* **2013**, *6*, 1777–1791. [CrossRef]
32. Mao, T.; Li, Y.; Chang, F. Research progress of monitoring technology of ship-source air pollutants in China emission control area. In *E3S Web of Conferences*; EDP Sciences: Ulis, France, 2020; Volume 206, p. 02006.
33. Conti, G.O.; Heibati, B.; Kloog, I.; Fiore, M.; Ferrante, M. A review of AirQ Models and their applications for forecasting the air pollution health outcomes. *Environ. Sci. Pollut. Res.* **2017**, *24*, 6426–6445. [CrossRef]
34. Wang, J.; Zhou, K.; Xing, W.; Li, H.; Yang, Z. Applications, evolutions, and challenges of drones in maritime transport. *J. Mar. Sci. Eng.* **2023**, *11*, 2056. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Comparative Analysis of Recurrent vs. Temporal Convolutional Autoencoders for Detecting Container Impacts During Quay Crane Handling

Sergej Jakovlev ^{1,2,*}, Tomas Eglynas ², Edvinas Pocevicius ³, Miroslav Voznak ^{1,2}, Gediminas Gričius ⁴,
Valdas Jankunas ² and Mindaugas Jusis ⁴

¹ Department of Telecommunications, VŠB-Technical University of Ostrava, 17. Listopadu 2172/15, 708 00 Ostrava, Czech Republic

² Marine Research Institute, Klaipeda University, Universiteto al. 17, 92295 Klaipeda, Lithuania

³ Department of Marine Engineering, Klaipeda University, Bijunu Str. 17, 91225 Klaipeda, Lithuania

⁴ Department of Informatics and Statistics, Klaipeda University, Bijunu Str. 17, 91225 Klaipeda, Lithuania

* Correspondence: sergej.jakovlev@ku.lt

Abstract: This research develops and validates a novel impact detection system for container monitoring using autoencoders embedded within an edge computing unit. This solution addresses common limitations in current container tracking systems, such as a lack of real-time processing and reliance on cloud connectivity, by enabling local, on-device anomaly detection. We compare the performance of Recurrent Autoencoders (RAEs) and Temporal Convolutional Autoencoders (TCAEs) using acceleration data collected during quay crane handling. Experimental results show that the RAE framework outperforms TCAEs, achieving a precision of 91.3%, a recall of 87.6%, and an F1-score of 89.4% for impact detection while also demonstrating lower reconstruction loss and improved detection of sequential anomalies. The system accurately identifies impact events with minimal computational overhead, proving its viability for real-time deployment in port environments. Our findings suggest that time-series autoencoder architectures, particularly RAEs, are effective for detecting mechanical impacts in resource-constrained edge devices, offering a robust alternative to traditional cloud-based solutions.

Keywords: autoencoder; impact detection; accelerator

1. Introduction

In the recent literature, autoencoder technology has been adopted extensively, ranging from medical data analytics [1] to solving the complex problems of outlier detection [2], from sensory data helping engineers and industry representatives in security process safety applications to helping cybersecurity specialists develop new cybersecurity standards and specialised IoT systems [3]. The already-adopted encoding–decoding techniques are consistently outperformed by newer methods in terms of computational speed and versatility, which has led to their increasing popularity among top-tier researchers across all scientific fields [4,5]. In our previous work [6], we demonstrated the adaptability and efficiency of this approach in detecting specific security events related to shipping containers by analysing acceleration data from sensors mounted on the mechanisms of quay cranes. We have not only demonstrated the exact nature of the problem but also shown that these techniques can be applied in small-scale, embedded edge applications with a security system compatible with other decision support means, detecting even the less visible impacts on the structural integrity of the containers [7,8]. In this paper, we

continue our investigation, demonstrating that newer concepts, such as those presented in [9,10], can enhance the adaptability of this technology in the transportation sector as well as among logistics representatives searching for better and more reliable solutions for shock detection, which is currently limited by the computational capabilities of most devices.

Overall, the adoption of novel techniques for damage detection and fault diagnosis in physical systems has been demonstrated by numerous research groups in the engineering field [11,12]. In many cases, autoencoders are generally used to predict specific patterns. In contrast, Lv et al. [13] addressed the problem of link prediction in dynamic (temporal) networks using temporal link prediction approaches, specifically non-negative matrix factorisation (NMF). As a result, they proposed two new deep autoencoder-like NMF models with graph-regularised prediction methods for dynamic networks. Sarikaya et al. [14] analysed machine learning-based intrusion detection systems (IDSs) by estimating the accuracy of their classification results. The authors developed a method that employs generative adversarial networks to produce adversarial attack data and propose a novel RAIDS solution—a robust IDS model designed to be resilient against adversarial attacks in general, utilising the autoencoder’s reconstruction error technique as a prediction value for the classifiers within the estimation metrics. Such direct applications suggest that autoencoders can be used as efficient, supportive tools in combination with more complex techniques [15–18]. Yet, each case study is unique when higher-order data sets are examined [19,20], which is the primary issue when simplified solutions are applied in large-scale applications, as no unified solution exists [21]. No unified neural network structure exists for the encoder and decoder that can be adopted without any serious modifications [22,23]. Encoder input datasets require refinement, cleaning, and structuring for each specific application [24,25]. Even if similar conditions occur, the same application in different work environments can yield inadequate results [26,27]. This was also highlighted by Danti and Innocenti [28], which strongly supports our claim. They examined this problem and proposed a quality procedure to determine the optimal training set size to minimise the reconstruction error of an autoencoder with predefined structures and hyperparameters that will be trained to encode the expected behaviour of a complex energy generation system.

The application of autoencoders has yet to reach its full potential in direct physical systems and industrial applications due to the complexity of anomaly detection methods [18]. Most existing anomaly detection methods rely not only on training on normal data but also ignore the multi-cluster nature of normal and abnormal patterns. This problem was generally discussed by Zhu et al. [29], who proposed a novel method called the Adaptive Aggregation–Distillation Autoencoder (AADAE) for unsupervised anomaly detection, built upon the density-based landmark selection concerning representing diverse standard patterns. In general terms, anomaly detection among datasets is a vast research field. Autoencoders could indeed improve the situation if adopted in combination with more complex computational techniques, which has been proven by Yan et al. [30], Zhang et al. [31], and Fernandez-Rodriguez et al. [32], as well as incorporating deep learning models within the encoder/decoder [15,33,34] dealing with unsupervised learning problems [35–43].

The significance of this paper lies in its intention to fill a gap in the current understanding of how embedded machine learning systems, particularly autoencoders, can be utilised within edge devices to enhance the efficiency and effectiveness of container tracking and monitoring systems. The idea of using an AE in edge devices is relatively fresh, even though NN structures have been utilised for decades. Many authors have already demonstrated their direct applicability in real-life applications. For instance, Zhai et al. [44] proposed an optimisation framework to reduce operational costs in hybrid cloud–edge systems by

accounting for mobile user-induced service migration. It introduces the AMGG algorithm, which uses autoencoders and evolutionary techniques to solve complex scheduling problems efficiently. Meanwhile, Yu et al. [45] approached the edge computing challenge from a different angle by presenting a complete three-layer IoT architecture for predictive maintenance, integrating edge, cloud, and application layers to enhance real-time processing and scalability. They introduced a distributed edge-assisted autoencoder to improve performance and efficiency, validated through an industrial case study. Similarly, Goyal et al. [46] proposed a low-cost edge-IoT model using Raspberry Pi 4 and a lightweight LSTM-based autoencoder to detect environmental anomalies in poultry farms. The model achieves real-time inference with an F1-score of 0.9627, enabling accurate and efficient on-device analysis in rural settings, which correlates strongly with the findings presented in our research.

Overall, moving data processing to the edge aims to mitigate common issues such as network congestion and packet loss, thereby improving the reliability and timeliness of critical monitoring tasks [47–49]. The outcomes of this research are expected to have significant implications for the design of future logistical monitoring systems that lack stronger innovations. They could lead to the broader adoption of machine learning and edge computing technologies in the logistics industry.

The primary contribution of this research is noteworthy:

- This paper is believed to be one of the first to embed autoencoders in edge sensors for the dynamic condition monitoring of containers. This approach utilises time-series analytics for anomaly detection, showcasing a novel application in the logistics sector.

We propose a novel integration of Recurrent and Temporal Convolutional Autoencoders (RAEs and TCAEs) within an edge computing unit for detecting physical impacts on shipping containers, which has not currently been implemented by any industry representative. While autoencoders have been used in various anomaly detection applications, their deployment in quay crane operations to detect impact-induced anomalies remains unexplored. Our study contributes to this field by demonstrating how time-series autoencoder architectures can effectively capture mechanical impact patterns while operating on resource-constrained edge devices, significantly reducing network reliance and enhancing real-time detection capabilities. Therefore, the objectives of this paper were the following:

- To further test the embedded anomaly detection system by implementing an autoencoder-based anomaly detection system within the selected sensor device. This system will process data locally, reducing the dependency on remote cloud services and minimising the latency and network strain typically associated with large-scale data transmission.
- To validate the system in a live operational environment by testing the robustness and reliability of the proposed system under real-world conditions at a container terminal. This will include evaluating the system's ability to accurately detect anomalies in real-time and assessing its impact on the overall efficiency of container monitoring.
- To demonstrate the potential of edge computing in logistical operations by showcasing how edge computing can be leveraged to solve complex monitoring tasks that benefit from immediate data processing, such as the dynamic condition monitoring of containers. This includes providing systematic early warnings to repair personnel, crane operators, and terminal managers and facilitating proactive maintenance and operational decisions.
- To contribute to the academic and industrial understanding of autoencoder applications in logistics by presenting findings that could pave the way for further research and development in the integration of advanced machine learning techniques, like autoencoders, in logistics and supply chain management.

2. Materials and Methods

2.1. Description of the Experimental Setup

A critical step in the loading and unloading of containers, known as the handling procedure, involves attaching the containers to the crane's spreaders and their further transportation either to the shore from the ship or vice versa (see Figure 1). This phase is particularly vulnerable to the generation of mechanical stresses and impacts. Such stresses typically arise from various factors, including abrupt crane movements, alignment discrepancies between the container and the spreader, and the substantial force exerted during the coupling of containers to the spreader bars. The mechanical impacts experienced during this operation can vary in severity.

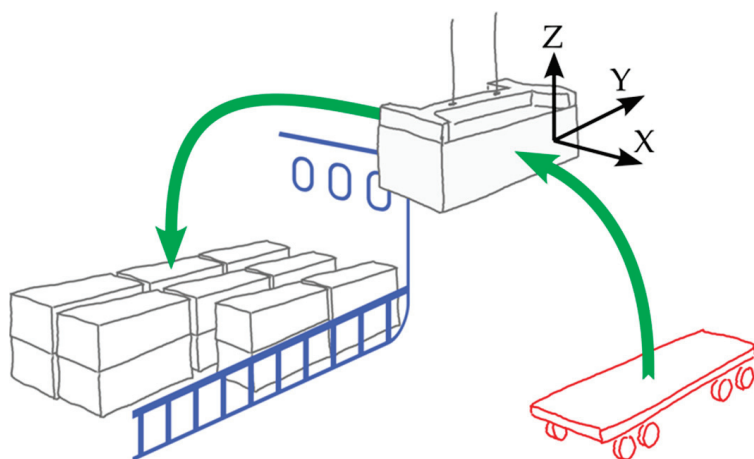


Figure 1. Explanation of a typical handling procedure performed by a port quay crane transporting the shipping container from a truck to a ship using a spreader mechanism (the red arrow shows the direction of movement of the container during the examined handling procedure).

They may present as minor vibrations, which could appear inconsequential; however, they have the potential to accumulate damage over time or as significant jolts that pose immediate risks to the integrity of the container and the safety of the operation, with incidents occurring in the shafts of the vessels. Each type of impact contributes differently to the potential degradation or damage of the container, as illustrated in Figure 2 (upper side).

This figure offers a visual representation of how these forces act upon the containers during the handling process, thereby providing a clearer understanding of the stresses involved. The probability of such damaging events is relatively high due to the complex and often expedited nature of port operations. Factors such as operator error, mechanical failure, environmental conditions, and improper container loading can exacerbate the likelihood of these occurrences. Despite advancements in crane technology and operational procedures, the risk of impact and subsequent damage to containers remains a persistent challenge [50]. The dangers of neglecting the mechanical stresses and impact events affecting containerised cargo handling operations are significant and multifaceted. When these stresses and impacts go unnoticed or unresolved, the damages incurred can severely compromise the structural integrity of containers. This endangers the safety of the cargo being transported and compromises the safety of personnel involved in subsequent handling, storage, and transportation activities.

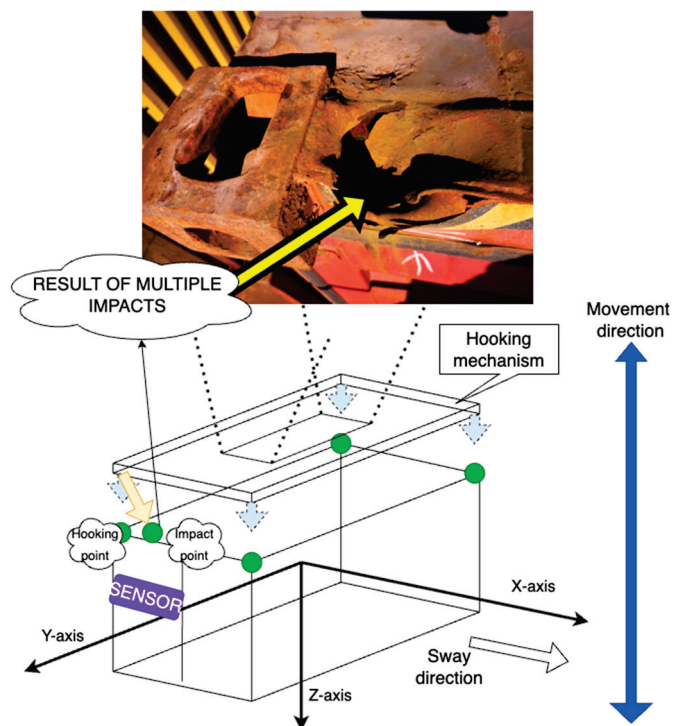


Figure 2. Demonstration of the marine container hooking procedure and the result of an incorrectly performed procedure.

Additionally, damaged containers can introduce a cascade of logistical challenges, including operational inefficiencies and delays, which may result in significant financial losses. These losses can stem from increased costs associated with damage claims, repairs, and replacements, as well as potential fines or penalties for non-compliance with safety regulations and standards. Given the high stakes, it becomes evident that there is an urgent need to develop a robust and reliable detection system capable of monitoring, identifying, and mitigating these risks proactively and efficiently. Such a system would play a critical role in ensuring the safety and structural integrity of containers while also streamlining operations and reducing the likelihood of unforeseen expenses and disruptions.

This proposed framework primarily revolves around an impact detection module prototype attached to the door side, as depicted in Figure 3, which was integral to the experimental setup used during the research.

The prototype incorporates a sophisticated data logging system equipped with fast local storage capabilities. Furthermore, the system features a data transmission module, allowing for the real-time collection and transmission of data related to impact events. The detection framework also integrates a range of advanced electronic components to enhance its functionality. The system underwent rigorous laboratory testing to validate its performance and reliability, which already demonstrated its potential [6]. During these tests, acceleration data and other critical parameters were collected using a scale model of a quay crane. The experimental setup involved installing the detection system on a specially prepared spreader subjected to varying loads and operational conditions during testing. This framework is expected to significantly improve safety standards, minimise costs, and ensure seamless cargo handling, contributing to more efficient and sustainable port operations. The device system consisted of the following:

- Data transmission unit using Bluetooth 5.1.
- Raspberry Pi 4 (four ARM A72 1.5 GHz cores, 8 GB of RAM) with a 128 GB SD UHS-I memory card.

- SINDT-232 digital accelerometer with high-stability 200 Hz MPU6050 3-axis acceleration (WitMotion Shenzhen Co., Ltd., Shenzhen, China), having 0.05-degree accuracy and an acceleration range of ± 16 g.
- Inner 33,000 mAh battery.



Figure 3. Demonstration of the experimental test site with the developed solution on the left side marked with a blue circle and the other tested equipment on the right side marked with green circles (the right door).

A series of experimental hooking procedures were conducted as part of the study to evaluate the detection system under real-world operational conditions. These procedures were carefully carried out from the truck positioned below the quay crane, replicating typical container handling scenarios. Visual inspections were performed during each use case to ensure thoroughness and assess the condition and behaviour of the system under varying operational parameters, as outlined in the referenced study [6].

For each operation conducted during the experimental phase, two distinct datasets were collected to facilitate a comprehensive analysis of the detection system's performance. The first dataset represented normal operations where no impact events were detected, serving as a baseline for comparison. The second dataset captured instances of actual impact events, highlighting the system's ability to identify and document mechanical stresses and impact occurrences (see Figure 4).

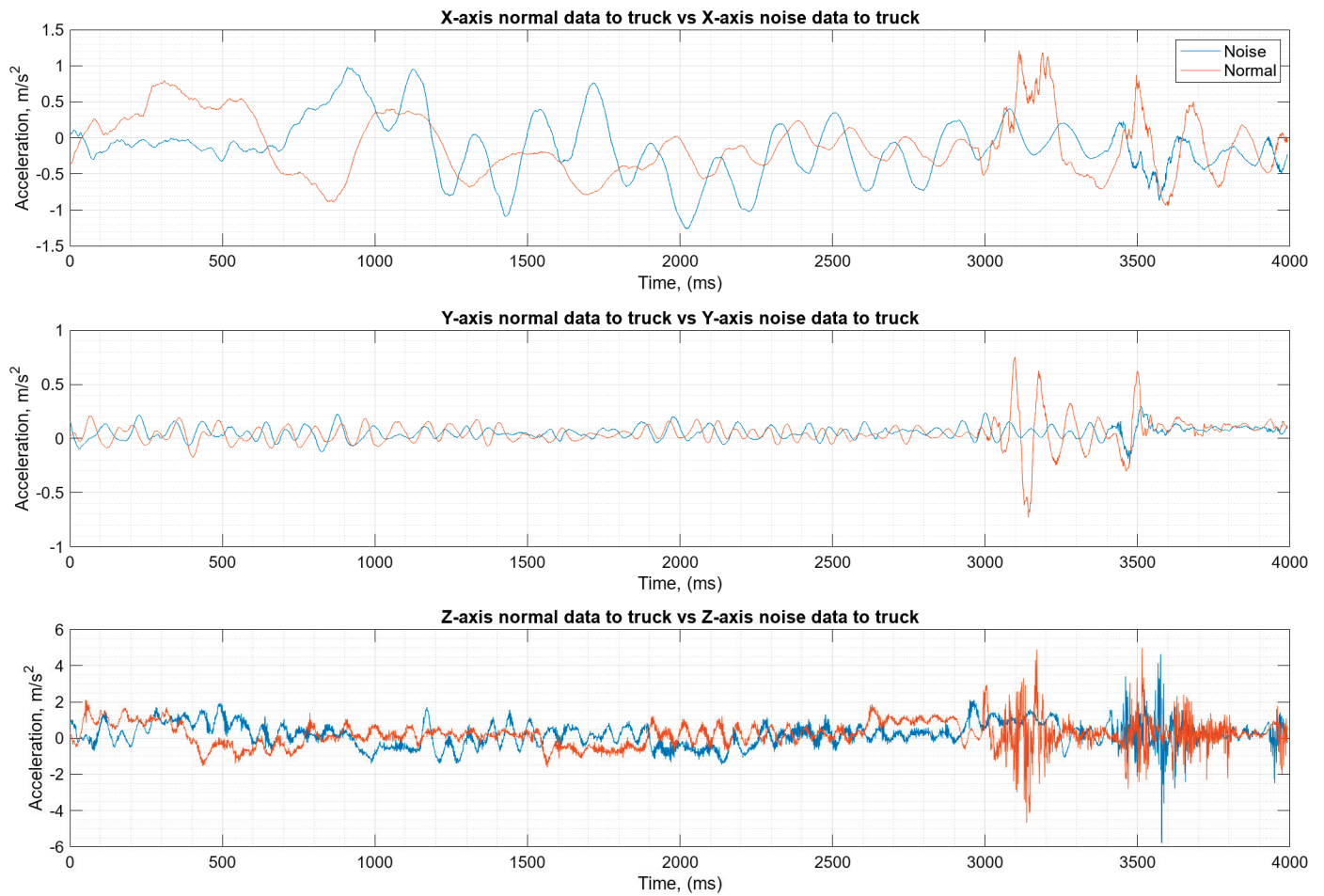


Figure 4. A demonstration of the acceleration values, both noisy and normal.

By analysing these two datasets side by side, the research team evaluated the accuracy and reliability of the detection framework in distinguishing between normal operational conditions and scenarios involving significant impacts. This dual-dataset approach was instrumental in verifying the system’s effectiveness in real-world situations. It provided critical insights into its capacity to enhance safety and operational efficiency in container handling and transportation environments. Here, a “normal” operation with a small number of low-energy actual impacts is compared to a “noisy” operation with a higher count of impacts of higher energies, all for the X, Y, and Z axes. Also, a previously developed autoencoder integration technique (see Figure 5 [6]) was used again, as it has already proven its capability to work with the apparatus. We would like to highlight that existing container monitoring systems heavily rely on remote cloud processing, which can lead to latency issues, data loss, and a high degree of dependency on continuous network connectivity.

Current impact detection mechanisms are limited in capturing transient mechanical stresses in real time, often requiring post-event analysis based on incomplete datasets. Furthermore, traditional threshold-based anomaly detection methods struggle to distinguish actual impact events from operational vibrations, leading to frequent false positives. Our proposed edge-integrated framework directly processes impact data at the source, minimising latency, reducing dependence on external infrastructure, and improving detection accuracy.

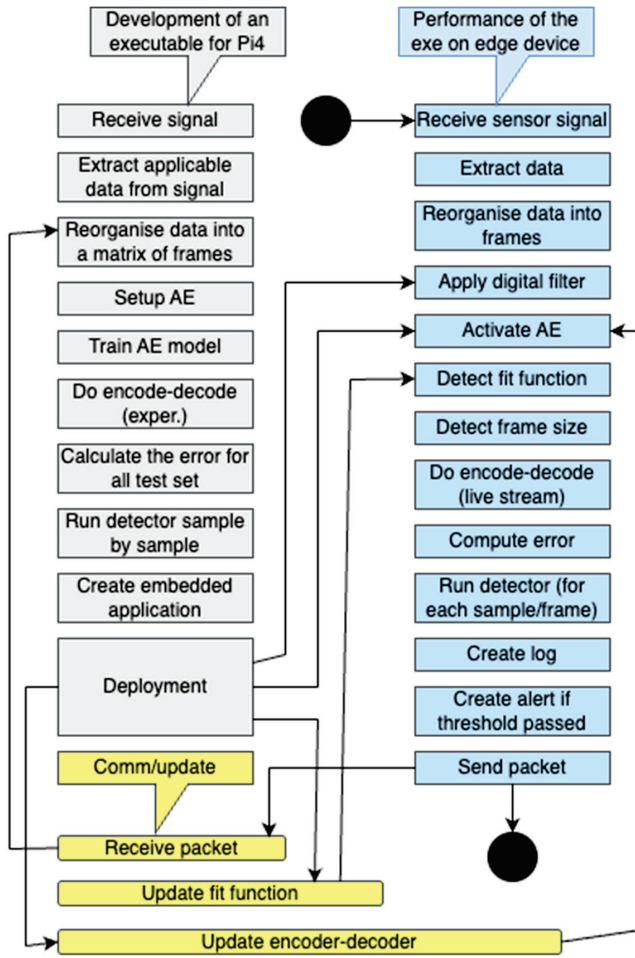


Figure 5. Embedded autoencoder application.

2.2. Structure of an Autoencoder

An unsupervised autoencoder is a type of neural network used for unsupervised learning of efficient coding, typically for dimensionality reduction or feature learning. The basic structure of an autoencoder can be described using Figure 6 and the mathematical expression involving its encoding and decoding functions.

An autoencoder consists of two main parts:

- Encoder: This part of the network compresses the input into a more miniature, encoded representation. The encoder function is typically denoted as f .
- Decoder: This part attempts to reconstruct the input from the encoded representation. The decoder function is typically denoted as g .

The mathematical expression is as follows: let $x \in \mathbb{R}^d$ be the input vector to the autoencoder. The encoding function f maps this input to a hidden representation z in a lower-dimensional space $\mathbb{R}^h$ (where typically $h < d$) (1):

$$z = f(x; \theta_e). \quad (1)$$

Here, θ_e represents the parameters of the encoder. The decoding function g then attempts to reconstruct the original input from this hidden representation (2):

$$\hat{x} = g(z; \theta_d). \quad (2)$$

Here, θ_d represents the parameters of the decoder.

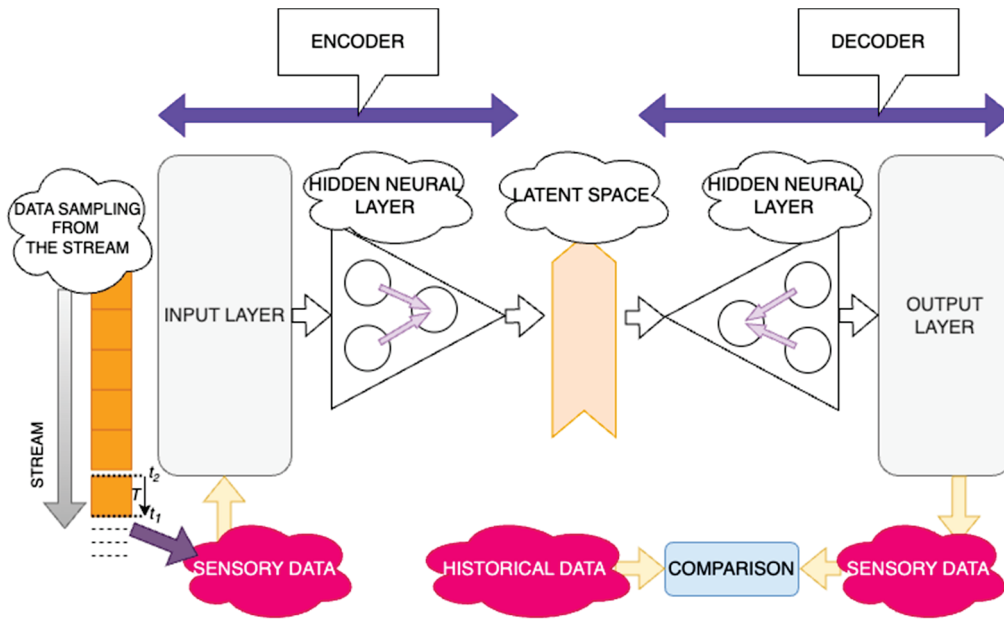


Figure 6. Components and interconnections within the autoencoder and the framework.

The goal of the autoencoder is to minimise the difference between the input x and its reconstruction $\hat{x}$. This is often achieved by minimising a loss function L over the parameters of the encoder and decoder $L(x, \hat{x})$. A common choice for L is the mean squared error (MSE), which can be expressed as (3):

$$L(x, \hat{x}) = \|x - \hat{x}\|^2. \quad (3)$$

Combining these, the complete operation of an autoencoder can be described by the following:

$$\hat{x} = g(f(x; \theta_e); \theta_d). \quad (4)$$

The training involves optimising the following:

$$\min_{(\theta_e, \theta_d)} \sum_{i=1}^N L(x^{(i)}; \hat{x}^{(i)}). \quad (5)$$

where N is the number of training samples, and $x^{(i)}$ and $\hat{x}^{(i)}$ are the i -th input and its reconstruction, respectively. This framework allows the autoencoder to learn valuable properties of the data unsupervised as it tries to capture the essential features necessary to reconstruct the input from the compressed encoded representation. When analysing time-series data using unsupervised learning methods, specific autoencoder architectures are well-suited to capture such data's temporal dependencies and nuances. Here are two key types of autoencoders that are often employed in unsupervised time-series analytics:

- Recurrent Autoencoders (RAEs) utilise recurrent neural network (RNN) architectures, like Long Short-Term Memory (*LSTM*) or Gated Recurrent Units (GRUs), in both the encoder and decoder parts. These models are highly effective for time-series data due to their ability to maintain state or memory across time steps, capturing the dynamic temporal behaviour of the data.
- Encoder: Processes the input time-series data and compresses it into a latent space representation. It typically consists of several layers of *LSTM* or GRU cells

that process each time step sequentially, updating their internal state based on both the current input and the previously hidden state

$$z_t = LSTM(x_t, h_{t-1}). \quad (6)$$

- Where x_t is the input at a time t , h_{t-1} is the hidden state from the previous time step, and z_t is the output of the *LSTM* representing the encoded state.
- Decoder: Starts from the latent representation and attempts to reconstruct the original time-series data over time. The decoder can also be a sequence of *LSTM* or GRU cells, which takes the latent representation and reconstructs the sequence, potentially in reverse order or from a condensed state (7).

$$\hat{x}_t = LSTM(z_t, \hat{h}_{t-1}). \quad (7)$$

- Here, $\hat{x}_t$ represents the reconstructed output at time t , and $\hat{h}_{t-1}$ is the hidden state from the previous time step in the decoder.
- Temporal Convolutional Autoencoders (TCAEs) utilise convolutional neural network (CNN) architectures specifically designed for sequence data, referred to as 1D CNNs. These models are beneficial for handling large datasets where capturing long-term dependencies explicitly with RNNs might be computationally expensive.
- Encoder: Applies convolutional filters to the input sequence to extract features over multiple time steps simultaneously. This is performed using 1D convolutions that slide over the time dimension, capturing patterns and anomalies in the data (8).

$$z = ReLU(W_e * x + b_e). \quad (8)$$

- Where $(*)$ denotes the convolution operation, W_e and b_e are the weights and biases of the convolutional layers, respectively, and *ReLU* is the activation function.
- Decoder: Used to map the latent space back to the time-series data's original dimensionality. This part of the network learns to reconstruct the time sequence from the compressed encoded features (9).

$$\hat{x} = ReLU(W_d * z + b_d). \quad (9)$$

The essence of these autoencoders is that they can be trained on large datasets without labels, making them ideal for exploratory data analysis and preliminary data processing where supervised labels might not be available. The choice between recurrent and convolutional architectures often depends on the specific characteristics of the data and computational constraints, being the primary research intent for our device and aligning the descriptions of the Recurrent Autoencoder (RAE) and the Temporal Convolutional Autoencoder with the mathematical structure of the classic autoencoder model, as seen in Tables 1 and 2.

In Tables 1 and 2, we outline the hyperparameter choices for the RAE and TCAE models, detailing how each selection was made to optimise model performance. The learning rate, number of layers, and regularisation terms were chosen based on empirical testing, ensuring a balance between accuracy and computational efficiency. The performance comparison between RAEs and TCAEs provides valuable insights into the relative strengths of recurrent and convolutional architectures in impact detection, supporting our methodological approach. The most common objective in the Recurrent and Temporal Convolutional Autoencoders is to minimise the reconstruction error across the sequence ($L(x, \hat{x})$). Where x is the original input time-series, and $\hat{x}$ is the reconstructed time-series

output by the decoder. The training process involves optimising the parameters (θ_e and θ_d) to minimise the loss function L over all the training samples. This can be achieved using standard backpropagation techniques tailored for time dependencies in RAEs and spatial feature extraction in convolutional architectures.

Table 1. Parameters for the RAE framework.

ENCODER:	DECODER:
Input: x_t (where t indexes time steps in the input time-series).	Input: z_t (where t indexes time steps in the input time-series).
Encoding Function: f .	Decoding Function: g .
Output: Latent representation z_t .	Output: Reconstructed output $\hat{x}_t$.
We get the expression: $z_t = f(x_t; \theta_e) = LSTM(x_t, h_{t-1}; \theta_e)$ Here, θ_e includes the weights and biases of the LSTM layers in the encoder.	We get the expression: $\hat{x}_t = g(z_t; \theta_d) = LSTM(z_t, \hat{h}_{t-1}; \theta_d)$ Here, θ_d includes the weights and biases of the LSTM layers in the decoder.

Table 2. Parameters for the CNN framework.

ENCODER:	DECODER:
Input: x_t (the entire time-series input).	Input: z .
Encoding Function: f .	Decoding Function: g .
Output: Latent representation z_t .	Output: Reconstructed output $\hat{x}$.
We get the expression: $z = f(x; \theta_e) = ReLU(W_e * x + b_e)$ W_e and b_e represent the weights and biases of the convolutional layers, respectively, and $(*)$ denotes the convolution operation.	We get the expression: $\hat{x} = g(z; \theta_d) = ReLU(W_d * z + b_d)$ W_d and b_d represent the weights and biases of the transposed convolutional layers, respectively.

When employing autoencoders for time-series analysis, the goal is to minimise the loss function, quantifying the discrepancy between the input data and the reconstructed output from the model. The commonly used loss function in training autoencoders, particularly for time-series data, is the root mean squared error (RMSE). This function measures the squared root average of the squares of the differences between the original inputs and their reconstructed outputs. This is fundamental to ensuring that the autoencoder effectively learns the underlying patterns and dynamics of the data. Here is a detailed description of the minimisation of the loss function as a mathematical expression (10):

$$L(x, \hat{x}) = \sqrt{\frac{1}{T} \sum_{t=1}^T \|x - \hat{x}\|^2}. \quad (10)$$

where:

- x_t is the input at time t .
- $\hat{x}_t$ is the reconstruction output at time step t .
- T is the total number of time steps in the time-series.
- $\|\cdot\|^2$ denotes the squared Euclidian distance, ensuring that all errors are positive and that larger errors are penalized more heavily.

Concerning regularisation, it will prevent overfitting by adding a penalty term to the loss function. This penalty discourages overly complex models by penalising large weights.

Regularisation techniques like L1 (Lasso: $\lambda \sum_i |\theta_i|$) and L2 (Ridge: $\lambda \sum_i \theta_i^2$) add a penalty term to the loss function (11).

$$L(x, \hat{x}) = \sqrt{\frac{1}{T} \sum_{t=1}^T \|x - \hat{x}\|^2 + \begin{cases} \lambda \sum_i |\theta_i|, \text{ or} \\ \lambda \sum_i \theta_i^2 \end{cases}}. \quad (11)$$

where:

- λ is the regularisation strength.
- $|\theta_i|$ is the absolute value of each weight in the network.
- θ_i^2 is the square of each weight in the network.

L1 regularisation encourages sparsity in the model weights, potentially leading to models that rely on fewer features. At the same time, L2 regularisation discourages large weights by penalising the square values of the weights, leading to a more evenly distributed importance across all features. Therefore, the objective function for training an autoencoder is to minimise the total reconstruction error over all instances in the training dataset (12):

$$\min_{\theta_e, \theta_d} \sqrt{\frac{1}{N} \sum_{i=1}^N L(x^{(i)}, \hat{x}^{(i)})}. \quad (12)$$

where:

- N is the number of training samples.
- θ_e and θ_d are the encoder and decoder parameters, respectively.

3. Results and Discussions

3.1. Data Preparation and Analytical Steps

The following steps were taken to enhance the predictive accuracy and reliability of the anomaly detection:

- **Data Preparation.** Container acceleration data is gathered using a single sensor. This data includes noise components, often seen as groups of anomalous outliers that compromise functional stability. To address this, the data is segmented into a training set, a validation set that contains only normal, extracted data, and a test set that includes both normal and anomalous data.
- **Model Initialisation.** The initial setup involves configuring the model's hyperparameters, such as the weight coefficient, learning rate, number of iterations, and structure of the AE (autoencoder) module. At this stage of the experiment, the learning parameters were set to random values and optimised during the research process.
- **Model Offline Training.** The training process began by inputting only the normal data into the AE. Encoder fitting networks are then trained, leading to the generation of reconstructed data based on predefined equations. RMSE loss values are calculated during training, and the learning parameters of the network are updated iteratively, refining the model's accuracy.
- **Model Online Validation.** In this phase, only the normal data from the validation set is processed through the trained network. After reconstructing this data, the combined loss values of RMSE for each sample are computed. Given the potential presence of a few unmatched noise samples and to ensure the model's robustness, the loss value representing 95% of the normal data is selected as the threshold for detecting true anomalies (anomalous regions).
- **Model Online Testing.** The testing phase involves feeding both normal and anomalous data from the test set into the trained AE. The total RMSE loss values are then

computed for each sample to categorise them as normal or anomalous based on the established threshold.

The dataset used in this study was collected under varying operational conditions, including different crane speeds and impact magnitudes. To enhance the robustness of the model, we implemented a preprocessing pipeline that included:

- Outlier detection and removal based on statistical thresholds.
- Linear interpolation to handle missing sensor data.
- Low-pass filtering to reduce high-frequency noise.

These preprocessing steps ensure that the autoencoder models are trained on a dataset that closely represents real-world conditions without distortions caused by spurious data artefacts. A detailed comparison of the core architecture and training parameters for both models is presented in Table 3, highlighting the key difference between the RAE and the temporal convolutional design of the TCAE.

Table 3. Structural and training parameter comparison between the RAE and TCAE.

Parameter	RAE	TCAE-CN
Input shape	Flattened multivariate time window	Multivariate time sequence
Encoder	3 Dense layers (128 → 64 → 32)	2 Conv1D layers (32, 16 filters)
Decoder	3 Dense layers (64 → 128 → output)	Conv1D Transpose + Dense reconstruction
Latent dimension	32	
Activation function	ReLU (hidden), Linear (output)	ReLU (hidden), Linear (output)
Loss function	Mean Squared Error (MSE)	MSE
Optimiser	Adam	Adam
Epochs	100	100

3.2. Computational Results Assessment

The initial findings suggest that the use of autoencoders is an up-and-coming innovation. The computational module successfully captured the acceleration of the spreader without any noticeable deviations and computed the CNN and RAE algorithms with the desired data sample size (up to 2 s). We assess unsupervised anomaly detection techniques using data subjected to high noise levels, with a specific focus on the Recurrent Autoencoder (RAE) and the Temporal Convolutional Autoencoder. In this study, we replicate all methods ten times to minimise the stochastic variability inherent in experiments. The algorithms simulating the operational conditions of the device are executed on MATLAB R2020b Update 3 for macOS version 14.4.1, utilising an M3 Max 14-core CPU and 36 GB of rapid memory. By managing variables, we aim to identify an optimal combination of hyperparameters by employing a greedy algorithm approach. The primary hyperparameters are configured as follows: the number of iterations is 1000; the learning rate is initially 0.01, with a decay factor of 0.1 after every 20 iterations; and the hyperparameter λ in the loss function is set at 0.2 with $L1$. The selection of these hyperparameters is built on existing research and case studies.

Initial results for the handling operation mentioned in Section 2 are promising, as both frameworks were successfully deployed and worked as planned—both the RAE and CNN proved to work efficiently in detecting anomalous regions along the data stream without any severe computational strain on the device (see Figures 7 and 8), sending obvious impact alerts during observable physical contacts (red dots) in the shafts of the ship’s hull (which tend to happen regularly [51]) after the 3000th sample. The red lines indicate the

anomalous region on the graph, which is crucial for estimating the critical duration. For experimental clarity, computational rechecks were made for each use case to check the credibility of the results.

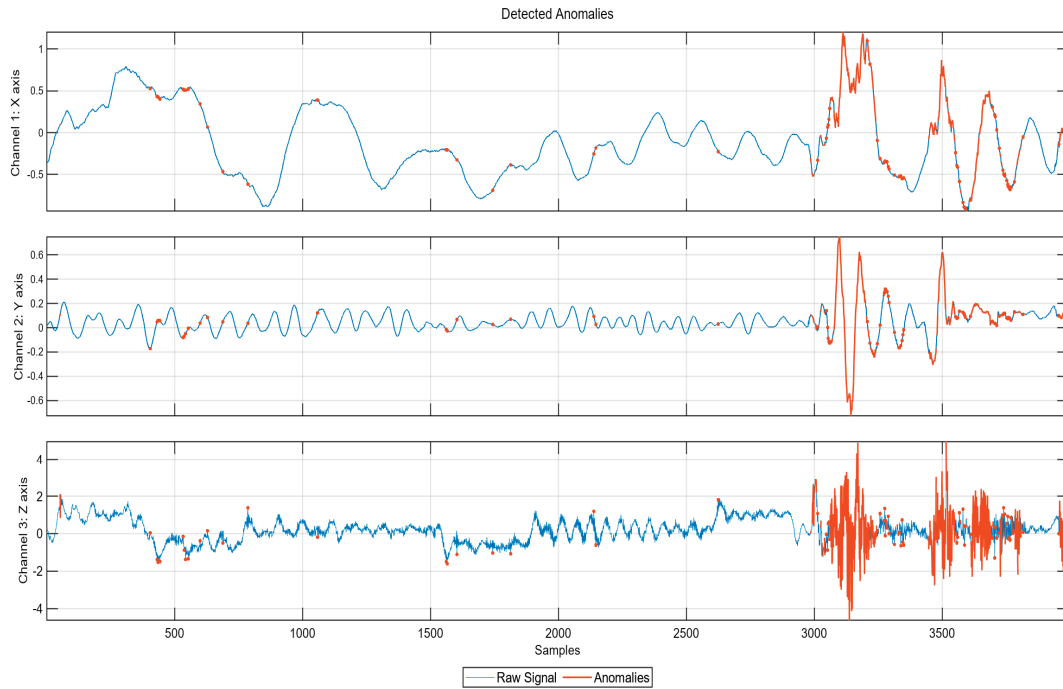


Figure 7. Data anomaly detection using the RAE framework.

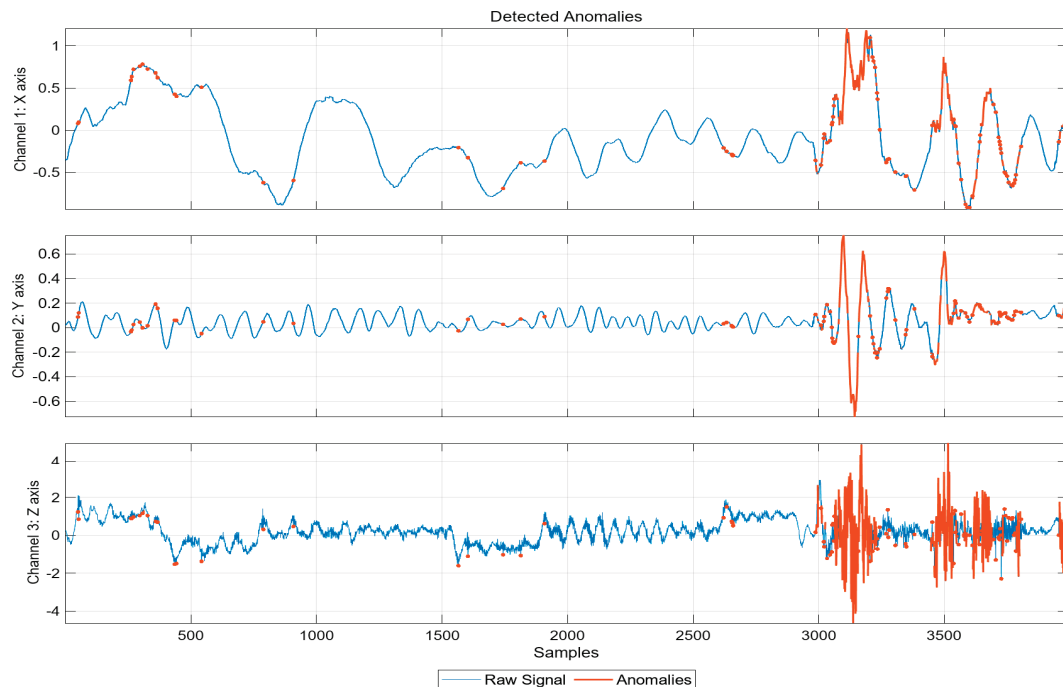


Figure 8. Data anomaly detection using a CNN framework.

In Figures 7 and 8, the red dots represent detected impact events based on the anomaly scores generated by the autoencoder models. These points indicate instances where the model identified deviations from standard operational patterns, marking cases of potential impact. The surrounding regions on the plots highlight the statistical distribution of impact-related anomalies over time, reinforcing the validity of our approach. During the initial

stages of the operation, the detection algorithm identified several impacts, represented by red dots, even though these occurred during movements with no visible obstacles. This indicates that the algorithm made several false assumptions about the presence of impacts in these scenarios. However, it is noteworthy that despite these false positives, the algorithm was able to assess the statistical significance of these detected impacts and determined them to be of low importance. As a result, no anomalous regions were generated based on these false detections. This outcome highlights a key strength of the detection system: its ability to effectively distinguish between actual impact events and false positives, ensuring that only genuinely significant impacts contribute to the formation of anomalous regions. This is the desired output of the experiment, as it demonstrates the capability of the solution to prioritise and correctly identify meaningful events.

Furthermore, a detailed analysis of the system's performance revealed that the Recurrent Autoencoder (RAE) proved to be a slightly superior solution when evaluated using reconstruction loss estimates (see Figures 9 and 10). The RAE's ability to consistently achieve lower reconstruction loss values under experimental conditions indicates its effectiveness in handling the complexities of impact detection and distinguishing between normal operational data and anomalous data points.

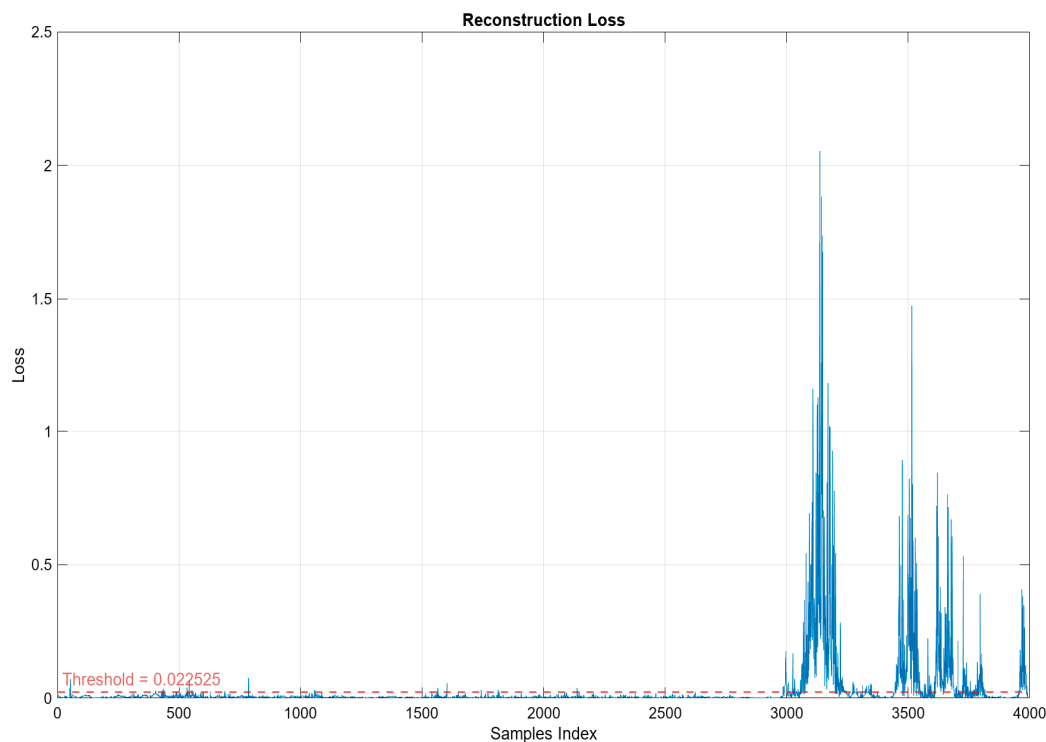


Figure 9. Reconstruction loss estimate for the RAE framework.

Additionally, the comparative analysis of the reconstruction loss distribution between normal operations and operations with identified physical impacts (see Figures 10 and 11) further validates the RAE's robustness. This analysis underscores its potential as a reliable tool for distinguishing actual physical impacts from noise or false positives, thereby enhancing the overall accuracy and reliability of the detection framework. By demonstrating these capabilities, the RAE highlights its value as an integral component of the proposed detection system, contributing to the successful implementation of an efficient and accurate monitoring solution for container handling operations.

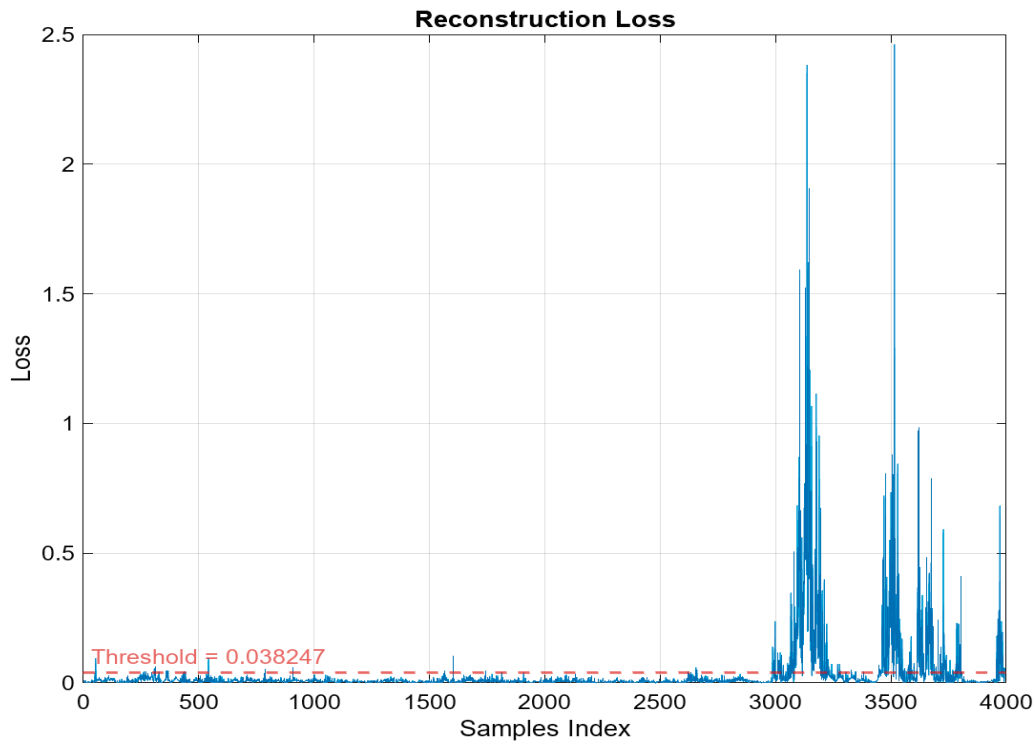


Figure 10. Reconstruction loss estimate for the CNN framework.

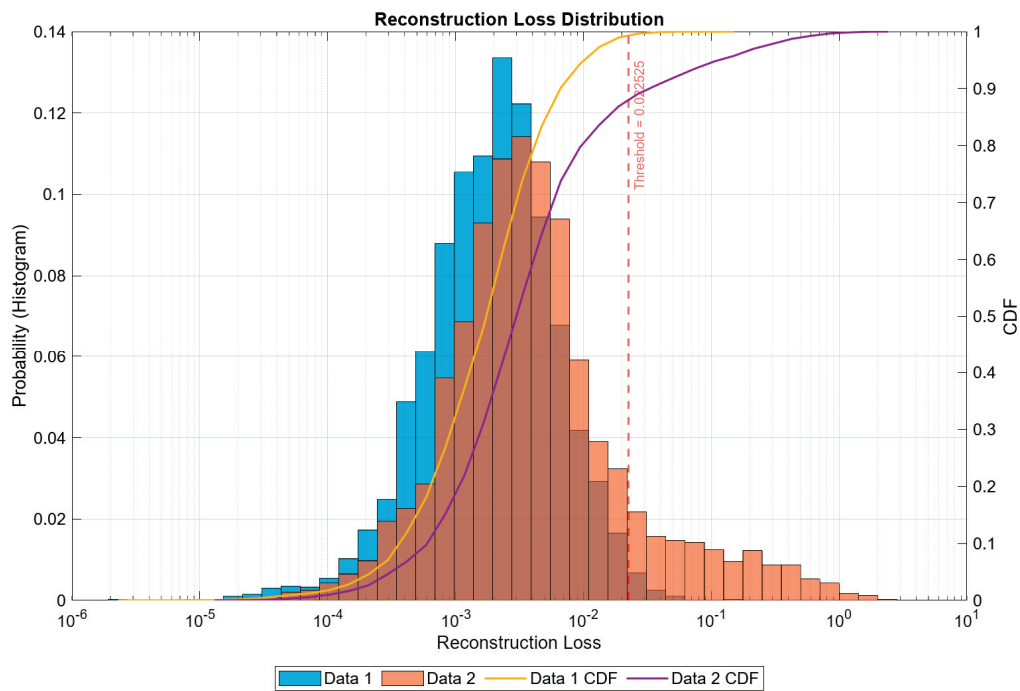


Figure 11. Reconstruction loss distribution estimate (standard vs. noise) for the RAE framework.

Figures 11 and 12 present the reconstruction loss distributions for the CNN and RAE frameworks. Both models effectively differentiate normal operations from impact-related anomalies, but the Recurrent Autoencoder (RAE) exhibits a more precise separation between these two categories. This improved separation suggests that the RAE is more capable of capturing temporal dependencies in impact signals, leading to reduced false positive rates.

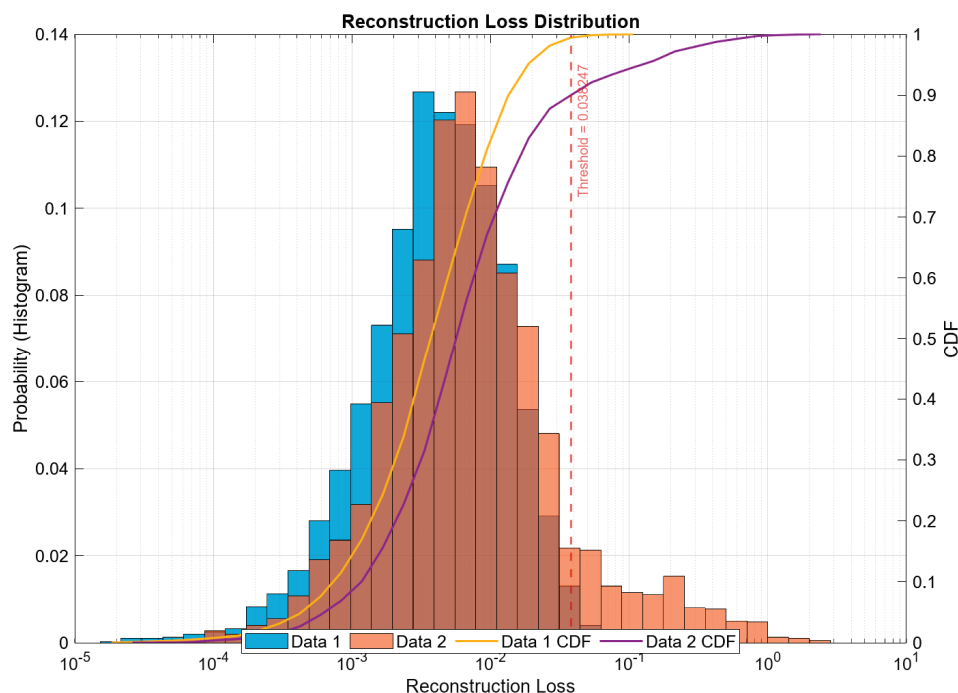


Figure 12. Reconstruction loss distribution estimate (standard vs. noise) for the CNN framework.

In contrast, the Temporal Convolutional Autoencoder (TCAE) framework shows a slight overlap between normal and anomaly loss distributions, which may increase the likelihood of missed impact events (false negatives). The superior performance of the RAE in distinguishing between normal and impact conditions supports its practical applicability for real-time impact monitoring in port operations. Comparative analysis of the RMSE curves, as well as the loss function during the training period, shows that the RAE performs sufficiently better under the experimental conditions of 1000 computing iterations (see Figures 13 and 14), indicating that the number of iterations may be diminished up to 300, thus lowering the computational strain on the edge device and increasing the response time of the system.

This observed performance suggests that the RAE is highly effective at minimising error rates and optimising the loss function, even in demanding experimental setups. Interestingly, the analysis also reveals that the number of computational iterations required for satisfactory performance can be significantly reduced from 1000 to as few as 300 iterations. This adjustment is particularly advantageous, as it would substantially lower the computational strain on the edge device where the detection system is deployed. The system becomes more efficient by reducing its computational load, resulting in a shorter response time and improved real-time operational capabilities. These improvements ensure the effectiveness of the detection framework in dynamic, time-sensitive environments, such as port container operations, where rapid and accurate responses are crucial for maintaining safety and operational efficiency. In addition to RMSE, we evaluated the performance of our anomaly detection models using Precision, Recall, and F1-score to provide a more comprehensive assessment of their ability to distinguish true impacts from false positives. RMSE serves as a helpful metric for measuring the reconstruction loss in autoencoders, but it does not provide a direct indication of classification accuracy. Precision was included to assess the proportion of detected anomalies that were actual impact events, while Recall measured the model's ability to identify all impact occurrences correctly. The F1-score, as a harmonic mean of Precision and Recall, ensured a balanced evaluation of the model's detection capability. The Recurrent Autoencoder (RAE) demonstrated superior anomaly

detection performance, achieving a Precision of 91.3%, a Recall of 87.6%, and an F1-score of 89.4%. Comparatively, the Temporal Convolutional Autoencoder (TCAE) exhibited a Precision of 89.5%, a Recall of 85.2%, and an F1-score of 87.3%. While both models effectively detected physical impacts, the RAE outperformed the TCAE in capturing sequential dependencies within the time-series data, resulting in fewer missed anomalies. To further validate our results, we analysed the RMSE values for detected anomalies and everyday operations. The mean RMSE for actual impact events was 0.124, whereas the mean RMSE for normal operational conditions remained at 0.069. This distinct separation confirms the effectiveness of the autoencoder-based anomaly detection framework. Additionally, we considered alternative threshold selection methods, including interquartile range filtering and adaptive thresholding based on local statistical properties, both of which reinforced the robustness of our chosen threshold, specifically the 95th percentile root mean squared error (RMSE).

Initial false positives were primarily attributed to transient vibrations during crane acceleration and deceleration phases. These vibrations produced signal variations that closely resembled minor impact events. To mitigate this, we refined our anomaly detection approach by introducing a time-based impact validation step, ensuring that only consecutive anomalies exceeding a predefined duration threshold were classified as actual impacts. Additionally, applying a contextual verification method, where detected anomalies were cross-checked against operational parameters such as crane velocity, further reduced false favourable rates.

The proposed impact detection framework represents a significant advancement in real-time anomaly detection for container handling. By integrating autoencoders within edge computing units, our approach enables proactive damage assessment, reducing operational delays and maintenance costs. The findings of this study provide a foundation for future enhancements in intelligent container tracking systems, with potential applications extending to automated ports and intelligent logistics networks.

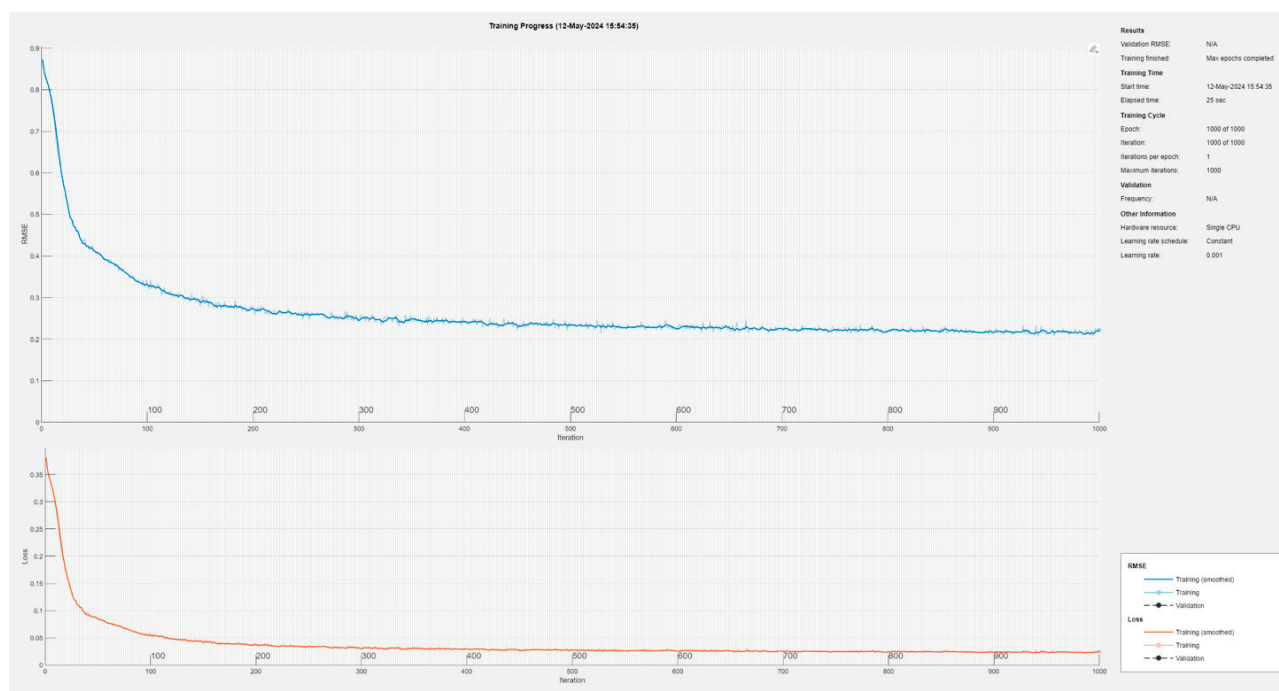


Figure 13. RMSE and loss estimate for the RAE framework.

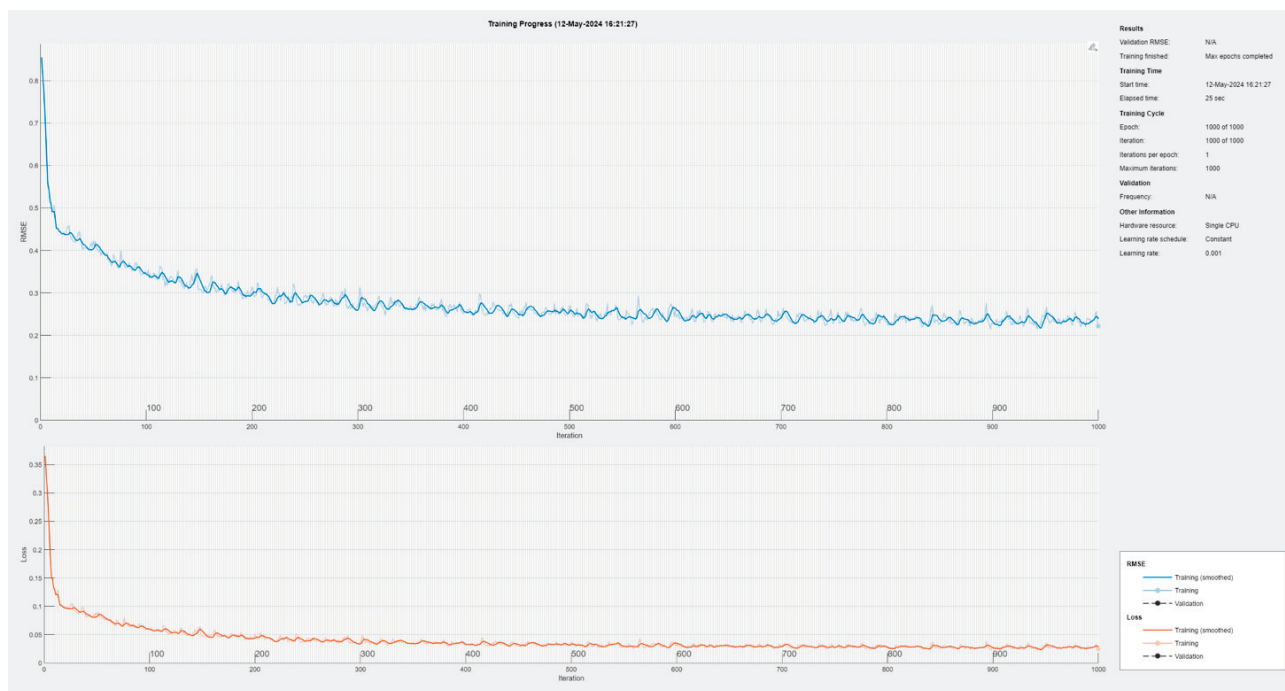


Figure 14. RMSE and loss estimate for the CNN framework.

In comparison to existing studies in the field, our proposed edge-integrated autoencoder framework demonstrates superior performance in both detection accuracy and deployment feasibility. While various studies have explored the use of autoencoders for anomaly detection, these are often evaluated under controlled conditions or with post-processing on centralised servers. In contrast, our system achieved an F1-score of 89.4% and a recall of 87.6% directly on embedded hardware deployed in a working port terminal, demonstrating robustness in the presence of high mechanical noise and stochastic operational variability. Unlike approaches that rely on offline analysis or cloud-based inference, our solution processes data on-device using lightweight RAE and TCAE architectures implemented on a Raspberry Pi-class processor. This addresses a key challenge in edge AI deployment, balancing computational cost with real-time reliability. To our knowledge, no previous study has systematically evaluated both recurrent and convolutional autoencoder models in the context of quay crane operations, positioning this work as a methodological advancement in unsupervised learning for maritime logistics. Additional comparative analysis of AE adoption in edge computing applications suggests that our solution has much room for improvement (see Table 4).

Table 4. Comparative analysis of several edge computing applications using various AE solutions.

Paper	AE Type	F1-Score	Edge Device	Notes
This research	RAE-TCAE	RAE 0.894	Raspberry Pi 4	Low-cost, high recall
Adversarial AE for IoT Intrusion [52]	Adv-AE + KNN	0.999	Raspberry Pi 3B	High performance
Lightweight LAE in Poultry Farm [46]	LSTM-LAE	0.963	Raspberry Pi 4	Low-cost, high recall

While previous works have explored autoencoders in various industrial settings, few have evaluated their effectiveness under dynamic quay crane operations, and none, to our knowledge, have directly compared recurrent and convolutional architectures in such

a context. This positions our work as both a methodological and practical advancement in the application of unsupervised learning for maritime logistics.

This work represents a significant advancement in the field of transport engineering by introducing ongoing research on the use of autoencoders for real-time anomaly identification in sensor data. It describes the architecture and deployment of an edge sensor-integrated anomaly detection system. The autoencoder-based approach has undergone extensive testing in an operating environment. These technologies are especially helpful in preventing severe damage to containers and cargo by promptly alerting managers, crane operators, and port maintenance workers.

4. Conclusions

Regarding the limitations and challenges, as well as future research agendas:

- **Local Minima and Non-Convexity.** The loss function of deep autoencoders is typically non-convex due to the complex nature of neural networks and their complex deployment. This leads to potential issues where the optimisation algorithm might converge to a local minimum rather than the global minimum, depending on the initial hyperparameter values.
- **Overfitting.** Autoencoders can be overfit to training data, especially if:
 - a. The network architecture is too complex relative to the amount of input data.
 - b. The selected framework model's capacity is too high, allowing the network to learn both pertinent features and noise in the training data.
- **Sensitivity to Hyperparameters.** The performance of the autoencoder is influenced by the selection of hyperparameters, including the number of layers, the number of neurons in each layer, the types of layers (e.g., convolutional, recurrent), the learning rate, and the optimisation algorithm employed.
- **Balancing Encoder and Decoder Strength.** A balance is needed between the encoder's and decoder's capacities. An overly powerful encoder or decoder can skew the learning process, leading to poor generalisation.
- **The Dimensionality of Latent Space.** The choice of the dimensionality of the latent space is crucial. Too large a latent space might lead the model to learn an identity function, where it can simply copy the input to the output without proper data compression. Conversely, a latent space that is too small may not capture all the necessary features of the data.

To support the claim of edge suitability, we evaluated the computational performance of the proposed models on the Raspberry Pi 4 (8 GB RAM), which served as our embedded deployment platform. Both the Recurrent Autoencoder (RAE) and the Temporal Convolutional Autoencoder (TCAE) were successfully executed on-device without noticeable delays during real-time data capture and inference. Preliminary testing showed that model inference for individual time-series samples (2-second windows at 200 Hz) completes within a fraction of a second, with no disruption to data logging or transmission processes. Resource usage remained within acceptable limits for real-time embedded applications, and the system operated continuously without thermal throttling or instability. While a detailed power and latency profile was not within the scope of this study, the results demonstrate that the selected autoencoder architectures can operate effectively within the computational constraints of commonly available edge hardware, confirming their practical viability for deployment in container impact monitoring scenarios.

We recognise that the experimental dataset used in this study was collected under controlled conditions at a single container terminal using a specific quay crane configuration. While the system demonstrated promising results in detecting mechanical impacts

in this setting, its applicability to other operational environments, such as different crane types, port layouts, or environmental conditions (e.g., weather variability, structural vibration patterns, or loading speeds), has not yet been verified. Variations in crane dynamics, spreader mechanisms, and container suspension systems may significantly influence sensor outputs, potentially affecting model sensitivity and false-positive rates. To address this limitation, future research will focus on multi-site validation across various terminals and geographic locations. This includes testing with diverse crane types (e.g., ship-to-shore cranes, rubber-tired gantries) and under various real-world environmental conditions to ensure robustness, reduce susceptibility to site-specific overfitting, and enhance model adaptability for widespread deployment.

Additionally, upcoming work will explore scaling autoencoder architectures to accommodate larger datasets, optimising computational efficiency for deployment on ultra-low-power IoT platforms and integrating context-aware features such as predictive maintenance analytics. We also aim to develop adaptive learning mechanisms that allow the system to dynamically adjust hyperparameters based on changing environmental inputs, thus maintaining consistent performance across operational scenarios. Although the current implementation demonstrates strong real-world potential, further experimentation involving deeper and more varied encoder–decoder structures is necessary to maximise detection accuracy and resilience in diverse logistical settings.

Author Contributions: Conceptualization, S.J. and T.E.; methodology, S.J.; software, M.J. and V.J.; validation, E.P., G.G. and T.E.; formal analysis, S.J.; investigation, T.E. and M.V.; resources, V.J.; data curation, E.P.; writing—original draft preparation, S.J.; writing—review and editing, T.E. and G.G.; visualization, T.E. and M.V.; supervision, M.V.; project administration, S.J.; funding acquisition, M.V. All authors have read and agreed to the published version of the manuscript.

Funding: The research was co-funded by the European Union (EU) within the REFRESH project—Research Excellence For Region Sustainability and High-tech Industries—ID No. CZ.10.03.01/00/22_003/0000048 of the European Just Transition Fund and also supported by the Ministry of Education, Youth and Sports of the Czech Republic (MEYS CZ), within a Student Grant Competition in the VSB–Technical University of Ostrava under project ID No. SGS SP2025/013.funding.

Data Availability Statement: Data samples can be provided upon request.

Acknowledgments: The project team would like to thank Klaipeda Port container terminal “Klaipėdos Smeltė” for their cooperation and support during the experiments. The authors would also like to note that Grammarly for Microsoft Office, version 6.8.263, was utilised to enhance the manuscript’s grammar, punctuation, and clarity during preparation.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Rashid, N.; Hossain, M.A.F.; Ali, M.; Islam Sukanya, M.; Mahmud, T.; Fattah, S.A. AutoCovNet: Unsupervised Feature Learning Using Autoencoder and Feature Merging for Detection of COVID-19 from Chest X-Ray Images. *Biocybern. Biomed. Eng.* **2021**, *41*, 1685–1701. [CrossRef] [PubMed]
2. Du, X.; Yu, J.; Chu, Z.; Jin, L.; Chen, J. Graph Autoencoder-Based Unsupervised Outlier Detection. *Inf. Sci.* **2022**, *608*, 532–550. [CrossRef]
3. Ali, S.; Li, Y. Learning Multilevel Auto-Encoders for Ddos Attack Detection in Smart Grid Network. *IEEE Access* **2019**, *7*, 108647–108659. [CrossRef]
4. Hong, S.; Kang, M.; Kim, J.; Baek, J. Investigation of Denoising Autoencoder-Based Deep Learning Model in Noise-Riding Experimental Data for Reliable State-of-Charge Estimation. *J. Energy Storage* **2023**, *72*, 108421. [CrossRef]
5. Saetta, E.; Tognaccini, R.; Iaccarino, G. Uncertainty Quantification in Autoencoders Predictions: Applications in Aerodynamics. *J. Comput. Phys.* **2024**, *506*, 112951. [CrossRef]
6. Jakovlev, S.; Voznak, M. Auto-Encoder-Enabled Anomaly Detection in Acceleration Data: Use Case Study in Container Handling Operations. *Machines* **2022**, *10*, 734. [CrossRef]

7. Cha, S.-H.; Noh, C.-K. A Case Study of Automation Management System of Damaged Container in the Port Gate. *J. Navig. Port. Res.* **2017**, *41*, 119–126. [CrossRef]
8. Wang, Z.; Gao, J.; Zeng, Q.; Sun, Y. Multitype Damage Detection of Container Using CNN Based on Transfer Learning. *Math. Probl. Eng.* **2021**, *2021*, 115022. [CrossRef]
9. Park, H.; Lee, G.H.; Han, J.; Choi, J.K. Multiclass Autoencoder-Based Active Learning for Sensor-Based Human Activity Recognition. *Future Gener. Comput. Syst.* **2024**, *151*, 71–84. [CrossRef]
10. Yang, Z.; Xu, B.; Luo, W.; Chen, F. *Autoencoder-Based Representation Learning and Its Application in Intelligent Fault Diagnosis: A Review*; Elsevier Ltd.: Amsterdam, The Netherlands, 2021. ISBN 8675523256054.
11. Santhi, T.M.; Srinivasan, K. A Duo Autoencoder-SVM Based Approach for Secure Performance Monitoring of Industrial Conveyor Belt System. *Comput. Chem. Eng.* **2023**, *177*, 108359. [CrossRef]
12. Zheng, S.; Zhao, J. A New Unsupervised Data Mining Method Based on the Stacked Autoencoder for Chemical Process Fault Diagnosis. *Comput. Chem. Eng.* **2020**, *135*, 106755. [CrossRef]
13. Lv, L.; Bardou, D.; Liu, Y.; Hu, P. Deep Autoencoder-like Non-Negative Matrix Factorization with Graph Regularized for Link Prediction in Dynamic Networks. *Appl. Soft Comput. J.* **2023**, *148*, 110832. [CrossRef]
14. Sarıkaya, A.; Kılıç, B.G.; Demirci, M. RAIDS: Robust Autoencoder-Based Intrusion Detection System Model Against Adversarial Attacks. *Comput. Secur.* **2023**, *135*, 103483. [CrossRef]
15. Wang, J.; Xie, X.; Shi, J.; He, W.; Chen, Q.; Chen, L.; Gu, W.; Zhou, T. Denoising Autoencoder, A Deep Learning Algorithm, Aids the Identification of A Novel Molecular Signature of Lung Adenocarcinoma. *Genom. Proteom. Bioinform.* **2020**, *18*, 468–480. [CrossRef]
16. Dumais, F.; Legarreta, J.H.; Lemaire, C.; Poulin, P.; Rheault, F.; Petit, L.; Barakovic, M.; Magon, S.; Descoteaux, M.; Jodoin, P.-M. FIESTA: Autoencoders for Accurate Fiber Segmentation in Tractography. *Neuroimage* **2023**, *279*, 120288. [CrossRef]
17. Takhanov, R.; Abylkairov, Y.S.; Tezekbayev, M. Autoencoders for a Manifold Learning Problem with a Jacobian Rank Constraint. *Pattern Recognit.* **2023**, *143*, 109777. [CrossRef]
18. Tsakiridis, N.L.; Samarinas, N.; Kokkas, S.; Kalopesa, E.; Tziolas, N.V.; Zalidis, G.C. In Situ Grape Ripeness Estimation via Hyperspectral Imaging and Deep Autoencoders. *Comput. Electron. Agric.* **2023**, *212*, 108098. [CrossRef]
19. Yang, H.; Ding, W.; Yin, C. AAE-Dpeak-SC: A Novel Unsupervised Clustering Method for Space Target ISAR Images Based on Adversarial Autoencoder and Density Peak-Spectral Clustering. *Adv. Space Res.* **2022**, *70*, 1472–1495. [CrossRef]
20. Cai, L.; Li, J.; Xu, X.; Jin, H.; Meng, J.; Wang, B.; Wu, C.; Yang, S. Automatically Constructing a Health Indicator for Lithium-Ion Battery State-of-Health Estimation via Adversarial and Compound Staked Autoencoder. *J. Energy Storage* **2024**, *84*, 110711. [CrossRef]
21. Liu, Z.; Teka, H.; You, R. Conditional Autoencoder Pricing Model for Energy Commodities. *Resour. Policy* **2023**, *86*, 104060. [CrossRef]
22. Wang, W.; Wang, H.; Chen, L.; Hao, K. A Novel Soft Sensor Method Based on Stacked Fusion Autoencoder with Feature Enhancement for Industrial Application. *Measurement* **2023**, *221*, 113491. [CrossRef]
23. Xiao, D.; Qin, C.; Yu, H.; Huang, Y.; Liu, C.; Zhang, J. Unsupervised Machine Fault Diagnosis for Noisy Domain Adaptation Using Marginal Denoising Autoencoder Based on Acoustic Signals. *Measurement* **2021**, *176*, 109186. [CrossRef]
24. Yang, H.; Qiu, R.C.; Shi, X.; He, X. Unsupervised Feature Learning for Online Voltage Stability Evaluation and Monitoring Based on Variational Autoencoder. *Electr. Power Syst. Res.* **2020**, *182*, 106253. [CrossRef]
25. Yan, S.; Shao, H.; Xiao, Y.; Liu, B.; Wan, J. Hybrid Robust Convolutional Autoencoder for Unsupervised Anomaly Detection of Machine Tools under Noises. *Robot. Comput. Integr. Manuf.* **2023**, *79*, 102441. [CrossRef]
26. Buzuti, L.F.; Thomaz, C.E. Fréchet AutoEncoder Distance: A New Approach for Evaluation of Generative Adversarial Networks. *Comput. Vision. Image Underst.* **2023**, *235*, 103768. [CrossRef]
27. Mulyanto, M.; Leu, J.S.; Faisal, M.; Yunanto, W. Weight Embedding Autoencoder as Feature Representation Learning in an Intrusion Detection Systems. *Comput. Electr. Eng.* **2023**, *111*, 108949. [CrossRef]
28. Danti, P.; Innocenti, A. A Methodology to Determine the Optimal Train-Set Size for Autoencoders Applied to Energy Systems. *Adv. Eng. Inform.* **2023**, *58*, 102139. [CrossRef]
29. Zhu, J.; Deng, F.; Zhao, J.; Chen, J. Adaptive Aggregation-Distillation Autoencoder for Unsupervised Anomaly Detection. *Pattern Recognit.* **2022**, *131*, 108897. [CrossRef]
30. Yan, H.; Liu, Z.; Chen, J.; Feng, Y.; Wang, J. Memory-Augmented Skip-Connected Autoencoder for Unsupervised Anomaly Detection of Rocket Engines with Multi-Source Fusion. *ISA Trans.* **2022**, *133*, 53–65. [CrossRef]
31. Zhang, K.; Han, S.; Wu, J.; Cheng, G.; Wang, Y.; Wu, S.; Liu, J. Early Lameness Detection in Dairy Cattle Based on Wearable Gait Analysis Using Semi-Supervised LSTM-Autoencoder. *Comput. Electron. Agric.* **2023**, *213*, 108252. [CrossRef]
32. Fernández-Rodríguez, J.D.; Palomo, E.J.; Benito-Picazo, J.; Domínguez, E.; López-Rubio, E.; Ortega-Zamorano, F. A Convolutional Autoencoder and a Neural Gas Model Based on Bregman Divergences for Hierarchical Color Quantization. *Neurocomputing* **2023**, *544*, 126288. [CrossRef]
33. Davila Delgado, J.M.; Oyedele, L. Deep Learning with Small Datasets: Using Autoencoders to Address Limited Datasets in Construction Management. *Appl. Soft Comput.* **2021**, *112*, 107836. [CrossRef]

34. Yong, B.X.; Brintrup, A. Coalitional Bayesian Autoencoders: Towards Explainable Unsupervised Deep Learning with Applications to Condition Monitoring under Covariate Shift. *Appl. Soft Comput.* **2022**, *123*, 108912. [CrossRef]
35. Novoa-Paradela, D.; Fontenla-Romero, O.; Guijarro-Berdiñas, B. Fast Deep Autoencoder for Federated Learning. *Pattern Recognit.* **2023**, *143*, 109805. [CrossRef]
36. Niu, Y.; Su, Y.; Li, S.; Wan, S.; Cao, X. Deep Adversarial Autoencoder Recommendation Algorithm Based on Group Influence. *Inf. Fusion.* **2023**, *100*, 101903. [CrossRef]
37. Thill, M.; Konen, W.; Wang, H.; Bäck, T. Temporal Convolutional Autoencoder for Unsupervised Anomaly Detection in Time Series. *Appl. Soft Comput.* **2021**, *112*, 107751. [CrossRef]
38. Sun, D.; Li, D.; Ding, Z.; Zhang, X.; Tang, J. Dual-Decoder Graph Autoencoder for Unsupervised Graph Representation Learning. *Knowl. Based Syst.* **2021**, *234*, 107564. [CrossRef]
39. Sun, H.; Zhang, L.; Wang, L.; Huang, H. Stochastic Gate-Based Autoencoder for Unsupervised Hyperspectral Band Selection. *Pattern Recognit.* **2022**, *132*, 108969. [CrossRef]
40. Wang, X.; Peng, D.; Hu, P.; Sang, Y. Adversarial Correlated Autoencoder for Unsupervised Multi-View Representation Learning. *Knowl. Based Syst.* **2019**, *168*, 109–120. [CrossRef]
41. Yang, K.; Kim, S.; Harley, J.B. Unsupervised Long-Term Damage Detection in an Uncontrolled Environment through Optimal Autoencoder. *Mech. Syst. Signal Process.* **2023**, *199*, 110473. [CrossRef]
42. Li, P.; Pei, Y.; Li, J. A Comprehensive Survey on Design and Application of Autoencoder in Deep Learning. *Appl. Soft Comput.* **2023**, *138*, 110176. [CrossRef]
43. Daneshfar, F.; Soleymnbaigi, S.; Nafisi, A.; Yamini, P. Elastic Deep Autoencoder for Text Embedding Clustering by an Improved Graph Regularization. *Expert Syst. Appl.* **2024**, *238*, 121780. [CrossRef]
44. Zhai, J.; Bi, J.; Yuan, H.; Wang, M.; Zhang, J.; Wang, Y.; Zhou, M. Cost-Minimized Microservice Migration With Autoencoder-Assisted Evolution in Hybrid Cloud and Edge Computing Systems. *IEEE Internet Things J.* **2024**, *11*, 40951–40967. [CrossRef]
45. Yu, W.; Liu, Y.; Dillon, T.; Rahayu, W. Edge Computing-Assisted IoT Framework With an Autoencoder for Fault Detection in Manufacturing Predictive Maintenance. *IEEE Trans. Ind. Inform.* **2023**, *19*, 5701–5710. [CrossRef]
46. Goyal, V.; Yadav, A.; Kumar, S.; Mukherjee, R. Lightweight LAE for Anomaly Detection with Sound-Based Architecture in Smart Poultry Farm. *IEEE Internet Things J.* **2024**, *11*, 8199–8209. [CrossRef]
47. Somma, M.; Flatscher, A.; Stojanovic, B. Edge-Based Anomaly Detection: Enhancing Performance and Sustainability of Cyber-Attack Detection in Smart Water Distribution Systems. In Proceedings of the 2024 32nd Telecommunications Forum, Belgrade, Serbia, 26–27 November 2024; TELFOR 2024—Proceedings of Papers. Institute of Electrical and Electronics Engineers Inc.: New York, NY, USA, 2024.
48. Arroyo, S.; Ho, S.-S. Poster: A Hybrid-Cloud Autoencoder Ensemble Method for BotNets Detection on Edge Devices. In Proceedings of the 2024 IEEE 8th International Conference on Fog and Edge Computing (ICFEC), Pennsylvania, PA, USA, 6–9 May 2024; IEEE: New York, NY, USA; pp. 104–105.
49. Malviya, V.; Mukherjee, I.; Tallur, S. Edge-Compatible Convolutional Autoencoder Implemented on FPGA for Anomaly Detection in Vibration Condition-Based Monitoring. *IEEE Sens. Lett.* **2022**, *6*, 7001104. [CrossRef]
50. Yi, M.S.; Lee, B.K.; Park, J.S. Data-Driven Analysis of Causes and Risk Assessment of Marine Container Losses: Development of a Predictive Model Using Machine Learning and Statistical Approaches. *J. Mar. Sci. Eng.* **2025**, *13*, 420. [CrossRef]
51. Jakovlev, S.; Eglynas, T.; Voznak, M.; Jusis, M.; Partila, P.; Tovarek, J.; Jankunas, V. Detecting Shipping Container Impacts with Vertical Cell Guides inside Container Ships during Handling Operations. *Sensors* **2022**, *22*, 2752. [CrossRef]
52. Aloul, F.; Zualkernan, I.; Abdalgawad, N.; Hussain, L.; Sakhnini, D. Network Intrusion Detection on the IoT Edge Using Adversarial Autoencoders. In Proceedings of the 2021 International Conference on Information Technology, Amman, Jordan, 14–15 July 2021; ICIT 2021—Proceedings; Institute of Electrical and Electronics Engineers Inc.: New York, NY, USA; pp. 120–125.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

A Comprehensive Model for Quantifying, Predicting, and Evaluating Ship Emissions in Port Areas Using Novel Metrics and Machine Learning Methods

Filip Bojić *, Anita Gudelj and Rino Bošnjak

Faculty of Maritime Studies, University of Split, 21000 Split, Croatia; anita@pfst.hr (A.G.); rinobo25@pfst.hr (R.B.)

* Correspondence: fbojic@pfst.hr

Abstract: Seaports, as major transportation hubs, generate significant air pollution due to intensive ship traffic, directly affecting local air quality. While emission inventories are commonly used to manage ship-based air pollution, they reflect only the emission-related aspect of a specified period and area, limiting the broader interpretability and comparability of the results. To overcome the mentioned challenges, this research presents the PrE-PARE model, which enables the prediction, analysis, and risk evaluation of ship-sourced air pollution in port areas. The model comprises three interconnected modules. The first is applied for quantifying emissions using detailed technical and movement datasets, which are combined into individual voyage trajectories to enable a high-resolution analysis of ship-based air pollutants. In the second module, the Multivariate Adaptive Regression Splines (MARS) machine learning method is adapted to predict emissions in varying operational scenarios. In the third module, novel metric methods are introduced, enabling a standardised efficiency comparison between ships. These methods are supported by a unique classification system to determine the emission risk in different periods, evaluate the intensity of various ship types, and rank individual ships based on their operational efficiency and emission optimisation potential. By combining new methods with technical and operational shipping data, the model provides a transparent, comparable, and adaptable system for emissions monitoring. The results demonstrate that the PrE-PARE model has the potential to improve strategic planning and air quality management in ports while remaining flexible enough to be applied in different contexts and future scenarios.

Keywords: sustainable shipping; air pollution; metric system; machine learning; risk assessment

1. Introduction

To keep in line with the Paris Agreement, in 2023, the International Maritime Organisation (IMO) reinstated the Initial Strategy on the reduction of greenhouse gases (GHG) with a more ambitious Strategy aiming for complete decarbonisation of ships by or around 2050 [1,2]. However, emissions released by ships are continuously rising. The 4th GHG Study showed a significant increase in GHG emissions from international shipping, which reached a share of 2.9% of global anthropogenic GHG pollution [3]. Future projections for shipping's GHG emissions also do not support the process of full decarbonisation, as global maritime trade continues to grow, the sector is still heavily reliant on fossil fuels, and regulatory standards lag far behind those that apply to other modes of transport [3,4]. Although the IMO has not yet adopted reduction requirements that would unequivocally cut down emissions, the implementation of global compulsory technical and operational

measures coincides with the reduction in the carbon intensity of ships, which in 2018 was on average 20 to 30% lower compared to the base year 2008 [3–5].

The first of these requirements to enter into force in 2013 were the Energy Efficiency Design Index (EEDI) and the Ship Energy Efficiency Management Plan (SEEMP), as part of the International Convention for the Prevention of Pollution from Ships (MARPOL) Annex 6, Chapter 4 [6,7]. SEEMP is a framework designed to support ship owners in improving the operational efficiency and carbon intensity of their fleet [7–9]. Divided into three parts, the recent guidelines include a plan to improve energy efficiency (through hull and propulsion maintenance, use of automated engine management, voyage planning, weather routing, speed optimisation, etc.), a plan to record fuel oil consumption, and methods for monitoring a ship's carbon intensity [7–9]. The EEDI is a mandatory measure that sets minimum energy efficiency requirements for newly built ships of 400 GT or more for international voyages [7,10]. The EEDI is determined by combining parameters from the fuel-based method for emission estimation, such as the power of the engines, their specific fuel consumption, and the carbon content of the fuel consumed in relation to the ship's capacity and different correction factors corresponding to the specific type of ship and the energy-efficient technology installed [11]. The idea behind the EEDI was to encourage ship owners to apply efficient technical solutions to improve the fuel efficiency of a ship at the design stage [5,7]. The CO₂ reduction level (grammes of CO₂ per tonne-mile) for the first phase was set at 10% compared to a reference line calculated from the average efficiency of ships built between 2000 and 2010, with further increases every 5 years [7]. In the meantime, MARPOL Chapter 4, including the SEEMP and the EEDI, has been extended and improved with additional requirements to achieve the objectives set out in the Strategy. Accordingly, in 2023, it became mandatory for relevant ships to calculate their achieved Energy Efficiency Existing Ship Index (EEXI) and to initiate data collection for reporting of annual operational Carbon Intensity Indicator (CII) and the associated CII rating [7,9,10]. The EEXI achieved by a ship indicates its energy efficiency compared to a baseline value. The obtained EEXI is then compared to a required value based on an applicable reduction factor expressed as a percentage relative to the EEDI [12,13]. This index must be calculated for in-service ships of 400 GT and above according to the different values for ship types and size classes, using a method based on the EEDI guidelines [7,13]. The calculated attained EEXI value for the individual ship must be below the required EEXI to ensure that the ship fulfils a minimum standard for energy efficiency [7]. As one of the recent monitoring mechanisms included in the SEEMP, from 2024, the CII must be calculated for ships of 5000 GT and above and reported together with the aggregated data for the previous year [14]. The CII measures the efficiency of a ship in transporting goods or passengers and is expressed as the mass of CO₂ emissions emitted relative to capacity/size and distance travelled [15]. Based on their efficiency, ships are given an environmental rating from A as the best to E as the worst performance level [16]. The annual amount of CO₂ released by ships is calculated by applying the fuel-based method, in which the total mass of fuel used is multiplied by the corresponding carbon content, while the transport work can be estimated by combining various factors depending on the type of ship and the available data [17]. Therefore, IMO proposed several indicators for determining transport performance such as Annual Efficiency Ratio (AER), cgDIST, and Energy Efficiency Operational Indicator (EEOI) [3,9,17].

Although it is expected that the implementation of all the above measures will further improve the efficiency of the global fleet and reduce its carbon intensity, there are still major limitations in terms of both technical and monitoring requirements. While the 4th IMO GHG study has shown a decrease in the carbon intensity of international shipping on the AER, research conducted by CE Delft has indicated that this reduction is mainly

influenced by high fuel prices and freight rates [18]. Costs and demand in the shipping market have a direct impact on orders for fuel-efficient hulls and the number of new builds, but also on fuel-saving measures [18]. According to the study by the International Council on Clean Transportation (ICCT), the EEXI would only reduce CO₂ emissions from the 2030 fleet by 0.7% to 1.3% compared to the baseline, as low-speed transport would continue to predominate [19]. The EEXI/EEDI will not directly reduce fuel consumption and CO₂ emissions if ships already operate slower than the speed limit proposed in the IMO requirements [19]. This means that the effectiveness of technical efficiency measures like the EEXI needs to be evaluated against real-world conditions [19]. The mentioned conclusion implies not only the weakness of regulatory standards but also of the approach based on vessel design and theoretical emission parameters rather than real operational data [20–22]. Additionally, even though the calculation method for the CII should include the annual amount of fuel consumed, the time spent at berth and/or anchorage is not considered, which can lead to inconsistencies in the final categorisation of the ships. This applies in particular to ship types that frequently operate within port areas, such as cruise ships, container ships, ro-ro ferries, etc. [23]. In addition, the Strategy for the reduction of GHG emissions and all current technical and monitoring mechanisms on which it is based, focuses only on CO₂ and does not account for other exhaust gases [1,9,10]. Due to their contribution to global warming, black carbon (BC), nitrous oxide (N₂O), and especially methane (CH₄), a gas that has 84 times greater potential than CO₂ to trap heat in the atmosphere over a 20-year period, should also be included in the Strategy and all associated requirements [24,25].

But more importantly, GHGs are only part of the problem. Throughout the internal processes of energy conversion and combustion, marine engines also discharge nitrogen oxides (NO_x), carbon monoxide (CO), sulphur oxides (SO_x), particulate matter (PM), and volatile organic compounds (VOC), recognised as one of the main air pollutant substances (APSs) [26,27]. The presence of these pollutants in the atmosphere and their uptake by humans can cause mortality as well as diseases such as pneumonia, ischaemic heart disease, chronic obstructive pulmonary disease, lung cancer, and stroke [28,29]. Diverse health problems can occur with both short- and long-term exposure, especially to PM, CO, ozone (O₃), NO_x, and SO_x [28,29]. Because of the direct interaction between the shipping sector and port cities, the local urban environment is directly exposed to the negative effects of air pollution. The impact on the environment and the deterioration of air quality can be severe along coastal zones and especially near seaports, as these areas are usually characterised by heavy shipping traffic [30,31]. Given the fact that 90% of European ports are spatially connected to cities, the extent of the deterioration in air quality is even more serious [32,33]. To monitor emissions at the national level, all countries in the European Union (EU) are required to provide GHG inventories to the European Environment Agency (EEA) in accordance with the Intergovernmental Panel on Climate Change (IPCC) guidelines [33–36]. However, the mentioned inventories are only supposed to include GHGs, and emissions from maritime transport, particularly in port areas, are still not required to be specified.

Considering the realistic need to obtain a better insight into air pollution from the maritime sector, establishing shipping emission inventories for seaports is recognised by both the port and scientific communities [30,33,37–41]. Port-related ship emission inventories are generally conducted by combining either a top-down or bottom-up approach with a fuel or energy-based method to determine the quantity of pollutants emitted in a given period [37]. But to calculate the amount of emissions with higher accuracy, a bottom-up energy-based method should be applied to large datasets that contain both technical and movement data of ships [37]. Basic technical details should include information on the type of vessel, dimensions, identifiers, and engine specifications regarding type, power,

and speed. Movement data is generally gathered from the Automatic Identification System (AIS) and provides near real-time information on the vessel's position, actual speed, and course with corresponding timestamps. As the mentioned approach is data-excessive, it can provide a detailed overview of various aspects of ship-based emissions that can be used to establish guidelines for emissions management in particular ports.

However, these insights have a temporal and location-specific standpoint and as such do not include the aspect of equivalence between the emission-related characteristics of different ships or port areas. A systematic review of the literature on port-related ship emission inventories has shown that the quantity and type of emissions in all analysed works were not comparable between, or even for the same, ports due to the inconsistency of data used [37]. Emission-related factors such as the types of pollutants included, diverse technical and movement data between ships, or the area and time frame covered vary with each calculation and are specific to individual inventories [37]. Furthermore, estimated emission quantities alone do not provide sufficient data to unconditionally categorise air pollution intensity between ships or the impact on overall pollution over different time periods [33]. In other words, without a standardised scaling system and corresponding baseline values that would allow a comprehensive comparison of emission levels between ships and overall traffic in different time windows and locations, it is difficult to determine whether a ship, a group of ships, or even an entire port area is efficient or rather an excessive polluter. To predict the risk of air pollution from ships and introduce strategies for its control, all aspects must be examined and considered, which can be a computational and time-consuming process as large databases and connections between different factors need to be analysed and compared [33]. The introduction of a scalable system based on the analysis of emissions-related data from inventories could provide a more transparent and efficient prediction and evaluation of air pollution intensity of ships and port areas.

That is why the aim of this research was to develop a unique metric, scaling, classification, and ranking methods, implemented inside a novel model for predicting and evaluating the air pollution risk and efficiency of different ship types and overall marine traffic in port areas. The port-related emissions prediction, analytics, and risk evaluation (PrE-PARE) model presented in this research is therefore based on a new emission evaluation approach and machine learning methods applied to actual ship technical and activity data, with the goal of creating an adaptable, relevant, and transparent overall system for calculating ship-related emission and classifying the level of risk for port areas in a standardised manner. In the first of three main modules, the collected shipping data was prepared and used within bottom-up and energy-based methodologies to estimate emissions with high temporal and spatial density. Upon the analysis of the results, a Multi-variable Adaptive Regression Spline (MARS) approach was adopted in a second module and applied to the processed datasets to assess the influence of emission-related factors and predict the levels released in various scenarios. Finally, the implementation of the novel metric, scaling, classification, and ranking algorithms enabled standardised categorisation of the air pollution efficiency and impact, and the temporal risk level evaluation of emitted emissions.

By integrating high-resolution technical and operational data with novel metrics and machine learning methods, the proposed PrE-PARE system enables accurate scenario-based forecasting under varying traffic conditions, along with comprehensive, comparable, and scalable evaluation of ship performance and port-wide pollution risk. Although the logic of the PrE-PARE system differs from ambient air monitoring frameworks like the European Air Quality Index, the underlying ambition is similar, to establish a clear, interpretable, and quantifiable basis for evaluating pollution focused specifically on ship-based emissions in port areas [42,43]. By providing ship air pollution quantification, predictive capabilities,

and novel evaluation metrics, the proposed framework pushes the boundaries of current practice in ship emissions assessment and aligns with broader interdisciplinary goals in maritime policy and urban air quality management. As this paper is part of a larger research project that continues the work presented in a systematic review and an article on an analytical model for estimating ship-related emissions in port areas, the Port of Split and the corresponding emissions-related data were used as a case study [33,37].

2. Materials and Methods

In contrast to the carbon efficiency and intensity indicators proposed so far, the PrE-PARE model is based on a bottom-up approach and energy-based method that combines the energy output with the related emission factor (EF) and time, which provides more realistic results for the calculated emissions and the corresponding air pollution metrics of ships. As the approach described above enables the estimation of emissions for each voyage of a ship based on technical details and operational data derived from the AIS, the emission impact and efficiency in different operating phases, including idle times, could be determined. Cruising, manoeuvring, and hoteling are modes of operation that correspond to the relevant workload of a ship's engines and are often performed differently between ship types or individual vessels. These ship-specific operational patterns directly affect the production of emissions and related impact. Since the model incorporates detailed emissions-related data, it was possible to determine the operational and air pollution profile of individual ships in different segments of the voyage. This feature is a key component for predicting future emissions in different scenarios and consistently determining the associated air pollution impact and efficiency through metric algorithms, not only for individual ships but also for groups of ships with similar characteristics classified into ship types. The calculation of emissions was therefore carried out according to the bottom-up principle, i.e., for each segment of the voyage and then totalled for the ship, the associated ship types, and the entire port area in the assigned period. However, the air pollution risk assessment was first carried out for complete shipping traffic by applying a classification system to determine whether the impact in the port area is very low, low, moderate, high, or very high. If the system classifies the risk as high or very high, the emission intensity of the group of ships and the optimisation potential of the individual ships are determined by applying the feature scaling method to the calculated values to finally rank the ships by their emission performance.

To perform the entire process, the model consists of three complex and interconnected modules, depicted in Figure 1. In the primary module for quantifying and analysing emissions, the collected technical and movement data were initially prepared with the aim of defining the voyage trajectories of each port arrival, stay, and departure for individual ships. These voyage datasets along with specific differentiation of ship types, enabled not only a high-density estimation of the air pollutants released by the individual ships together with various analytical results but also provided a basis for the extension of the model's forecasting capabilities. Therefore, in the second component, machine learning algorithms were applied to the previously processed extensive technical and operational data to create a predictive module. Since the respective voyage trajectories of the individual ships represent complex data clusters that contain important factors influencing emission production, a Multivariable Adaptive Regression Spline (MARS) approach was adopted in this research to determine the effects of included factors and predict the emission quantities released in different scenarios. To evaluate the performance of the predictive module, ten runs of k-fold cross-validation were performed, with additional validation by comparing the predicted and actual results based on unseen data. The final component of the model is based on the data generated by the two previous modules and includes methods for

assessing the emission intensity of ships, operational efficiency, and the temporal risk of air pollution. This has been achieved through the integration of novel metrics, scaling and risk classification, and ranking approaches, resulting in a transparent, comparable, and efficient overview of ship-based air pollution impact in port areas. Since the system does not only focus on carbon pollution but also includes the leading GHGs and APSs, it was able to calculate and evaluate the risk of shipping emissions for CO₂ and CH₄ as GHGs and SO_x, NO_x, PM₁₀, PM_{2.5}, NMVOC, and CO as APSs for each ship, a group of ships and an entire port area.

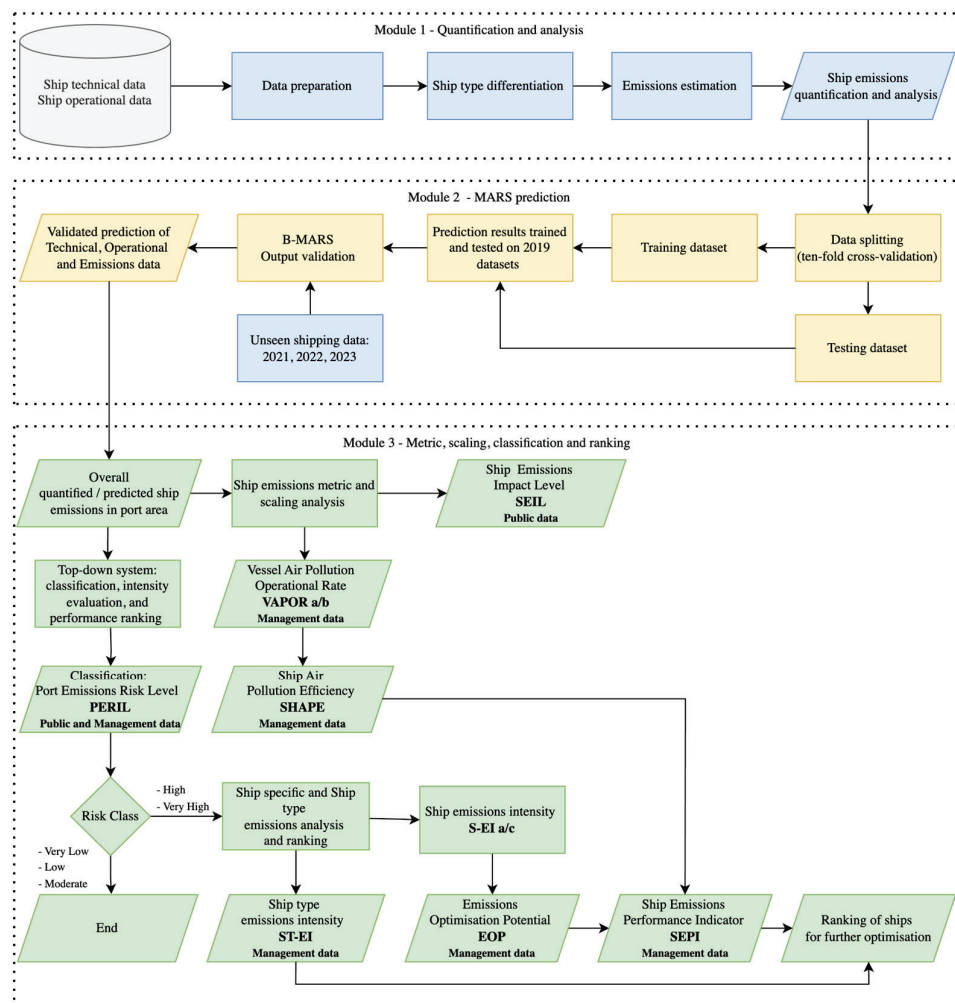


Figure 1. Flow diagram of the PrE-PARE model.

Given that the basic components of the model are derived from universal characteristics that significantly influence emissions production, the model is not limited to a single case study but can be applied to different ports. In addition, the modular structure of the model facilitates the integration of new insights and other relevant aspects of port-related shipping emissions, thereby improving the quality and scope of the final output. It is important to emphasise that the algorithms embedded in the PrE-PARE model and the data handling were produced using the software package RStudio 2023.09.1+494.

2.1. Emissions Estimation and Analysis Module

As stated, this paper builds on earlier work by the authors, particularly a systematic review and an article on an analytical model for estimating ship-related emissions in port areas [33,37]. The model presented in those studies served as the initial module, which was later adapted and integrated with novel predictive, risk evaluation, and metrics modules to

form the PrE-PARE system. While the full details of this component and its methodologies are discussed at length in the cited papers, a summary is provided in this chapter to ensure a comprehensive understanding of the system introduced in this study [33].

The estimation and analysis module was therefore able to produce an inventory of ship emissions for large port areas, providing a detailed overview of technical, temporal, spatial, and operational aspects. To obtain the emission-related analytical results, the module integrates three main components [33]. In the initial phase, the technical and AIS datasets were pre-processed by applying conversion, cleansing, filtering, formatting, and merging methods. These steps were essential to configure the collected data so that it was suitable for calculating emissions through the module. The technical data recorded relate to gross tonnage (GT), length, breadth, year, main engine (MA) power, auxiliary engine (AE) power, engine type, engine speed, fuel type, speed at maximum continuous rating (MCR), and relevant emission mitigation technology. The activity data derived from AIS comprised 49,540,895 reference points, which were linked to the above-mentioned technical details of the corresponding ships via name, type, and MMSI number as identifiers [33].

This was followed by the differentiation of ship types and the estimation of emissions during the processing phase. To create categories of ships with similar technical characteristics, data on the function of the ship and details on ship dimensions, speed, and engines were considered. When certain ship types showed significant variations in some features (e.g., engine power), a probability distribution was applied to obtain more specific categories, eventually resulting in the identification of 11 ship types:

1. Large Cruise Ships
2. Ro-Ro Ferry
3. Large Ro-Ro Ferry
4. Small Cruise Ships
5. Medium Cruise Ships
6. High-Speed Crafts
7. Excursion Ships
8. Tug
9. Pleasure Craft
10. Fishing
11. Sailing

Differentiating vessel types based on multiple characteristics improved the accuracy and efficiency of imputing missing data while providing the groundwork for the predictive capabilities of the next module [33].

In the second step of the processing phase, a bottom-up approach was applied, where reference points containing technical and operational data were combined into specific movement trajectories. Then, an energy-based method, which meets the requirements of the IPCC guidelines and is expressed in Equation (1), was used to calculate the total amount of air pollutants emitted during individual port calls for each ship [34]. Since speed alterations during port approaches lead to a change in energy output and, consequently, emission production, three operating modes (cruising, manoeuvring, and hoteling) with the time spent in them were determined for each trajectory [44]. This was done by estimating the workload of propulsion engines, known as load factor (LF), using the propeller law method expressed in Equation (2) [44]. As the cruising mode is identified by engine loads above 20% and manoeuvring is defined with LF values that are lower, hoteling operation is considered when the ships have switched their ME off and only use the generators while at berth or at anchor [27]. Subsequently, calculations were performed to estimate the quantities of greenhouse gases and air pollutants emitted by all vessels documented in the AIS database for the year 2019 within the study region.

Lastly, the output data was stored and handled with the goal of producing spatial and temporal visualisations of shipping emissions, as well as a detailed overview of various technical and operational aspects [37].

$$E = (P_{ME} \times LF \times EF_{ME} + P_{AE} \times LF \times EF_{AE}) \times T \times CF \quad (1)$$

where:

- E: emissions quantity by mode for each ship call—in grams (g);
- $P_{ME/AE}$: total power of main engines/auxiliary engines—kilowatts (kW);
- LF: load factor expressed as actual engine work output—as a percentage of engine power (%);
- $EF_{ME/AE}$: emission factors of different pollutants in regard to engine function, engine type, fuel type, and installation year—in grams per kilowatt hour (g/kWh);
- T: time spent in a certain movement activity—in hours (h);
- CF: correction factor for emission reduction technologies—constant.

$$LF = (S_A/S_M)^3 \quad (2)$$

where:

- S_A : actual speed of the ship—in knots (kt);
- S_M : speed of the ship at MCR—in knots (kt)

2.2. Predictive Module

Since the first module already prepared detailed emissions-related data, it was possible to develop a predictive algorithm as the second component of the PrE-PARE model. However, as the production of different APSs and GHGs from marine engines is influenced by a variety of factors that interact with each other, the prediction of emissions cannot be determined by linear functions. The relationships between predictor variables, such as engine power, type and speed, fuel type, actual speed, energy output, and time in the different operating modes, and the released emissions are non-linear. In addition, the values of the parameters and their interactions vary between the ship types and individual ships. To overcome the complexity of predicting ship-related emissions, the MARS method was adopted in this research.

2.2.1. Multivariate Adaptive Regression Splines (MARS)

MARS is a nonparametric, piecewise regression technique applicable in the modelling and analysis of complex, nonlinear relationships between multiple dependent and independent variables [45]. To examine the interactions and capture nonlinearity, this method automatically creates piecewise polynomials that characterise the data [46]. These polynomials, referred to as splines, are basis functions inside the MARS model, and prediction is made by summing the weighted output of all the basis functions in the model [47]. Simple BFs involve a single variable (x) and come in pairs of the form $(x - t) +$ and $(t - x) +$, where t is the knot, $(x - t) + = (x - t)$ if $x > t$, and 0 otherwise; and $(t - x) + = (t - x)$ if $x < t$, and 0 otherwise [47]. The modelling process has two main segments: the forward stage, which has the same idea as forward stepwise regression, and the backward (pruning) stage, where the model is improved and validated [45–47]. The forward stage starts by including the constant mean of the target variable (intercept). This allows for determining the breakpoints or knots for each predictor variable. Between each point, a fitting basis function is added. This process is being done iteratively until the threshold is reached [45–47]. Once the full set of features has been created, the algorithm sequentially removes individual features that do not contribute significantly to the model equation to avoid overfitting [46,47]. This “pruning” procedure assesses each predictor

variable and estimates the error rate aiming to eliminate basis functions with the least contribution [46,47]. This procedure is applied automatically through the Generalised Cross Validation (GCV) technique [45]. The GCV can be expressed as follows (3) [47]:

$$\text{GCV}(M) = \frac{\frac{1}{n} \sum_{i=1}^n \left(y_i - \hat{f}_M(x_i) \right)^2}{(1 - C(M)/n)^2} \quad (3)$$

where the denominator is a complexity function, and $C(M)$ is defined as $C(M) = (M + 1) + dM$, of which $C(M)$ is the number of parameters being fit and d represents a cost for each basis function optimization and is a smoothing parameter of the procedure [47]. Larger values for d will lead to fewer knots being placed and thereby smoother function estimates [45].

With the aim of obtaining more accurate results for the prediction of ship emissions, standard MARS and Boosting MARS (B-MARS) methods, with and without log normalisation, were applied in this research to historical data processed by a previous module. This approach resulted in four distinct predictive models. Their performance was assessed and compared to determine their accuracy and reliability. To ensure an unbiased selection of data and evaluate the models, k -fold cross-validation was used during the hyperparameter tuning process for MARS models where a ten-fold approach was applied. The data used for testing and training, with validation results averaged over ten runs, includes technical details and operational data derived from the AIS of ships that called at the port in 2019.

2.2.2. Prediction Performance Validation Metrics

To evaluate the performance of the above MARS models, the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2), presented in Equations (4)–(6), were applied in this research [46]. These metrics assess the accuracy of the predictions by quantifying the errors between predicted and actual values [48]. The R^2 score measures the proportion of variance in the dependent variable that the model explains [48]. Its value ranges from 0 to 1, with a value closer to 1 indicating a stronger relationship between variables and better predictive accuracy [49]. The RMSE calculates the square root of the mean of the squared differences between predicted and actual values [46]. As the errors are squared before averaging, the RMSE is more sensitive to larger errors and directly relates to Euclidean distance [46]. A lower value indicates better performance. MAE is used to measure the average absolute difference between predicted and observed values [45,48]. Unlike RMSE, this method treats all errors equally without squaring them, making it less sensitive to larger deviations [48]. A low MAE indicates higher prediction accuracy.

Although the above criteria are commonly used to evaluate the prediction of the models, additional validation was performed in this research by including unseen shipping data from 2021, 2022, and 2023. These datasets were first used to calculate ship emissions, and the results were then compared to those predicted by the module relying on historical data from 2019, aiming to gain a clear insight into its forecasting capabilities.

$$R^2 = 1 - \frac{\sum_{i=1}^n (X_i - Y_i)^2}{(\sum_{i=1}^n (\bar{Y} - Y_i)^2)} \quad (4)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |X_i - Y_i| \quad (5)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (X_i - Y_i)^2}{n}} \quad (6)$$

where X_i is the predicted i th value, Y_i element is the actual i th value, and n stands for the number of samples [48,50].

2.3. Ship Emissions Metric, Scaling, Classification, and Ranking Module

As outlined in the introduction, the IMO has implemented various measures to regulate and assess the carbon efficiency of ships, which have become basic tools in the global effort to reduce CO₂ emissions from the shipping industry. However, significant methodological limitations have been identified in mentioned measures. These constraints are particularly evident in their narrow focus on CO₂, which disregards other exhaust gases, and can lead to the overall environmental impact of ship emissions being overlooked, especially in urbanised port areas. Additionally, the carbon intensity calculations exclude the hoteling phase, leading to inconsistent results that may not accurately reflect a ship's total emissions profile. Adding to these issues, the measures do not fully incorporate operational data of the ships, which means that valuable insights into more realistic performance and efficiency are missing. Lastly, current imposed air pollution limitations are generally considered insufficient to address the full scope of the industry's environmental impact, emphasising the need for more comprehensive and differentiated approaches to assessing and regulating the sector's environmental footprint.

Consequently, the final objective of this research was to determine the air pollution impact and efficiency of ships while also providing a risk evaluation of temporal emission levels for individual ships, groups of ships, and the entire port area. This was achieved in the third module, where novel metric, scaling, classification, and ranking methods were applied to the outputs of the first and second modules.

2.3.1. Novel Metric and Scaling Systems for Standardised Measurement and Transparent Overview of the Emission Efficiency and Impact of Ships

To establish a systematic approach for assessing the efficiency and impact of air pollution from individual ships while ensuring standardised measurement, the operational output of maritime transport must first be defined and weighted. Given that the primary objective of the shipping industry is to provide safe and efficient transport, emissions should be evaluated in relation to this goal.

Therefore, Operational Efficiency (OE) should be defined as the ability of a ship to complete a voyage on schedule with minimal energy consumption per unit of time, as represented in Equation (7). In the context of the port approach, a voyage includes arrival, stay, and departure, encompassing cruising, manoeuvring, and hoteling as three operational modes, ensuring a comprehensive assessment of the vessel's operational profile. By integrating the time required to reach the expected EO with a ship's capacity for the intended operation and by comparing it against the emissions generated during the voyage, it becomes possible to determine the Vessel Air Pollution Operational Rate (VAPOR) for each mode, as defined in Equation (8). Unlike the metric system proposed by the IMO, the VAPOR considers available operational data and emissions over the entire voyage by evaluating air pollution in each mode (cruising, manoeuvring, and hoteling), separately giving the average hourly rate of exhaust production per work capacity. This allows for a standardised and detailed metric of emissions efficiency within specific operational phases. To ensure relevance, clarity, and comparability, the feature scaling technique is therefore applied. The calculated VAPOR (VAPOR c) for a specific ship is normalised against the baseline VAPOR (VAPOR b) of the corresponding ship type, revealing the Ship Air Pollution Efficiency (SHAPE), as depicted by Equation (9). The VAPOR b is calculated by relying on an extensive emissions-related database for each predefined ship-type group, which is classified in the first module. This enables SHAPE to indicate whether a ship is operating more or less efficiently compared to the expected performance of its category. It

also provides an insight into the progress made in the emission efficiency of certain ships and groups over time.

In addition, a simplified and user-friendly metric system has been developed to make the contribution of specific ships to air pollution in ports more transparent to the general public. Therefore, the Ship Emissions Impact Level (SEIL) compares the emissions released by a certain ship during a voyage relative to the average emissions per voyage of a generic ship within a defined time window, as shown in Equation (10). This approach provides a clear and intuitive visualisation of a ship's air pollution impact during an entire port visit, making it easier for the wider port community to assess and compare emission levels using a standardised emissions impact scale.

$$EO \text{ (kWh)} = \sum_{C,M,H} \text{Operational LF (kW)} \times \text{Operational time (h)} \quad (7)$$

$$VAPOR = \frac{\text{Emissions (g)}}{\text{Workcapacity} * \text{Operationaltime (h)}} \quad (8)$$

$$SHAPE = \frac{VAPOR_c}{VAPOR_b} \quad (9)$$

$$SEIL = \frac{\text{Ship emissions in voyage (kg)}}{\text{Generic ship emissions in voyage (kg)}} \quad (10)$$

2.3.2. Comprehensive Top-Down System for Classifying Air Pollution Risk in Ports, Evaluating Emission Intensity, and Performance Ranking of Ships

While emission estimation follows a bottom-up methodology, this system first evaluates the exhaust gases released in the entire port, then the relevant contributions from ship types, and finally the performance of specific vessels. This stepwise process allows for a structured assessment by first identifying the risk level of overall emissions in a relevant area, followed by determining the air pollution intensity for different ship types, and finally evaluating ship-specific indicators of emission performance. The potential for emission optimisation is calculated and combined with SHAPE. The aim of this three-stage procedure is to ensure a fair and data-driven framework for the control of ship-sourced air pollution in ports by providing an overview from both a macro and micro perspective.

To achieve this, the Port Emissions Risk Level (PERIL) classification algorithm is developed to determine the degree of severity of overall ship emissions in the entire port area for a specified period, as a first of three steps. The algorithm categorises emissions into five levels (Very Low, Low, Moderate, High, and Very High) by comparing the calculated emission rates with threshold values derived from the annual average and the standard deviation. This approach uses the average as a central reference point, allowing a clear and standardised classification of emission intensity based on statistical distribution rather than setting arbitrary thresholds. Upon determining the limit values, the system can automatically classify quantified shipping emissions. If the total emissions exceed the high-risk threshold, further analysis is conducted in a second step where the contribution of each ship group to total emissions is analysed.

This includes the application of the Ship Type Emission Intensity (ST-EI) assessment, which determines the degree of air pollution of each ship group by comparing their average emissions per voyage with the average emissions per voyage across all ship types in a given period, as shown in Equation (11). This method is used to create a relevant emissions contribution scale aimed at prioritising certain ship types for possible emissions improvement.

As a part of a final step, the Emission Optimisation Potential (EOP) is calculated for each ship by comparing its actual emissions per work capacity in each mode of voyage defined as the Ship Emission Intensity (S-EI), against a reference baseline, as depicted in Equation (12). The EOP is therefore used to determine the range of possible emission optimisation by displaying emission exceedances with values greater than 1 or improvement of operations in stated voyage with values lower than 1. The baseline values of individual ships represent the volume of emissions released typically per work capacity in each mode of voyage and are determined by relying on historical records from the database provided in the first module. If the database lacks the operational and air pollution profile of a ship (first port visit), the second module is used for predicting the emission quantities by relying on the technical and movement data of similar ships in the corresponding category.

Although the EOP exposes the performance of individual ships in terms of air pollution in each segment of a specified voyage, some vessels may already operate efficiently, leaving little room for further optimisation. Therefore, to ensure objective ranking of ships, the Ship Emissions Performance Indicator (SEPI) is applied, where emissions efficiency determined through SHAPE is combined with the EOP as a measure of the ship's operational performance in a specific voyage, as displayed in Equation (13). By incorporating both factors, SEPI enables a fair emissions attribution and ranking, ensuring that the ships with the highest improvement potential are prioritised in the final step.

$$ST - EI = \frac{E_{st} \times V_{tot}}{V_{st} \times E_{tot}} \quad (11)$$

$$EOP = \frac{S - EI \text{ a}}{S - EI \text{ b}} \quad (12)$$

$$SEPI = SHAPE \times EOP \quad (13)$$

where the following definitions apply:

ST-EI: ship type emission intensity—normalised value (dimensionless);

Est: total emissions for a specific ship type—in kilograms (kg);

Vst: number of voyages for that ship type—dimensionless value

Etot: total emissions for all ship types in the period—in kilograms (kg)

Vtot: total voyages for all ship types in the period—dimensionless value

EOP: emission optimisation potential—normalised value (dimensionless)

S-EI a/b: ship emission intensity actual/baseline—as emission mass in the entire voyage per unit of work capacity (kg/wcu)

2.4. Data

All the above methodologies integrated inside the PrE-PARE model were applied to the technical and activity data of the ships that visited the Port of Split in 2019, used here as a case study. The mentioned datasets, along with corresponding EFs and LFs, were preprocessed in the first module, as briefly explained in Section 2.1, creating an extensive database of emission-related inputs used throughout all three components of the PrE-PARE model.

Technical details of relevant ships were sourced from the Croatian Register of Shipping (CRS), the Croatian Integrated Maritime Information System (CIMIS), and relevant shipping company websites. The collected attributes include ship dimensions, work capacity, year built, ME and AE power, type and speed, fuel type, max speed at MCR, and emission reduction technologies, along with identifiers such as name, type, and MMSI number [33].

Since the AIS is used for sharing timely information on ship characteristics and their movements between other vessels and the shore to improve navigational safety, its recorded data is also often applied for estimating emissions [30,38,51,52]. Therefore, the position of the ship, course and speed over ground (COG, SOG), with corresponding timestamps, along with the name and MMSI number of specific ships as identifiers were applied in this research. For collecting mentioned datasets, the AIS station of the Faculty of Maritime Studies in Split was employed. However, as AIS transmits messages in the National Marine Electronics Association (NMEA) sentence format, which is unrecognisable to RStudio software, a specific Python script was developed and integrated into the first module to convert the ‘raw’ data from AIS into a readable CSV format [33]. In the earlier mentioned preprocessing stage, erroneous and non-ship entries were then removed, resulting in 49,540,895 ship records from 2019. These were used for emission estimation in the initial module, but also for testing and training in the predictive module. The same procedure was applied to acquire an additional 15,930,840 AIS reference points of ships visiting the same port during different periods in 2021, 2022, and 2023. These datasets represented unseen data used for extended validation of outputs produced by the second module.

These AIS reference points were then connected with technical data of particular ships via respective identifiers, finally creating individual trajectories for each port visit. This process enabled the identification of operating modes with corresponding temporal and spatial characteristics in each recorded voyage by estimating the energy output of ME through the propeller law method [44]. The LFs for AE were taken from relevant studies as constant values [31,35,44]. Since all emission-related technical and activity details were identified and connected into a single database, the EFs could be determined as the final and most complex dataset required for estimating and evaluating ship emissions by applying the relevant methodologies described in the 3rd and 4th IMO GHG Study and the San Pedro Bay Ports Report [3,44,53]. The types of Efs, along with the elements for identifying them, are presented in Table 1 [33].

Table 1. Engine details, modes of operation, and types of EFs incorporated in the model.

Elements for Determination and Types of EFs						
Engine Function	Engine Type	Engine Speed (rpm)	Fuel Type	Mode & LF		GHG EFs
ME	D	SS D < 300	MDO/MGO	C	LF = > 20%	CO ₂
AE	GTU	MS D 300–900	HFO	M	LF < 20%	CH ₄
	STU	HS D > 900	LNG	H	LF = < 2%	
	DF					
	D-E					
						PM 10, 2.5
						NMVOC
						CO

Within Table 1, D stands for diesel engine, GTU for gas turbine, STU for steam turbine, DF for dual fuel engine, D-E for diesel-electric engine, SS/MS/HS D for slow-/medium-/high-speed diesel engine, HFO for heavy fuel oil MDO/MGO for marine diesel/gas oil, and LNG for liquified natural gas [33].

3. Results

The integration of technical details with activity data within the PrE-PARE model enabled the determination of operating modes and EFs for each of the 65,471,735 AIS reference points. These complex emission-related datasets were then combined into individual voyage tracks for each ship that called at the port, creating a foundation for modelling, predicting, and evaluating ship emissions in large port areas. Therefore, the model has recognised 48,256 voyages to the passenger basin of the Port of Split in 2019 and in different periods of 2021, 2022, and 2023. However, it is important to emphasize that 2019 was used as the base year in this research, so the datasets from this year were used for defining refer-

ence points, and records from other periods served for validation. The recorded number of visits closely matched port traffic data, showing an average deviation of only 3% across all ship types, except for pleasure crafts, excursion, and sailing ships, which lacked consistent arrival figures in the different sources. This discrepancy is largely due to their irregular schedules, often resulting in underreported AIS data. In addition, the 3% deviation partly reflects the ability of AIS to capture even minor vessel movements, offering higher quality input for accurate emissions estimation compared to standard port statistics. The share of port calls between ship types is illustrated in Figure 2.

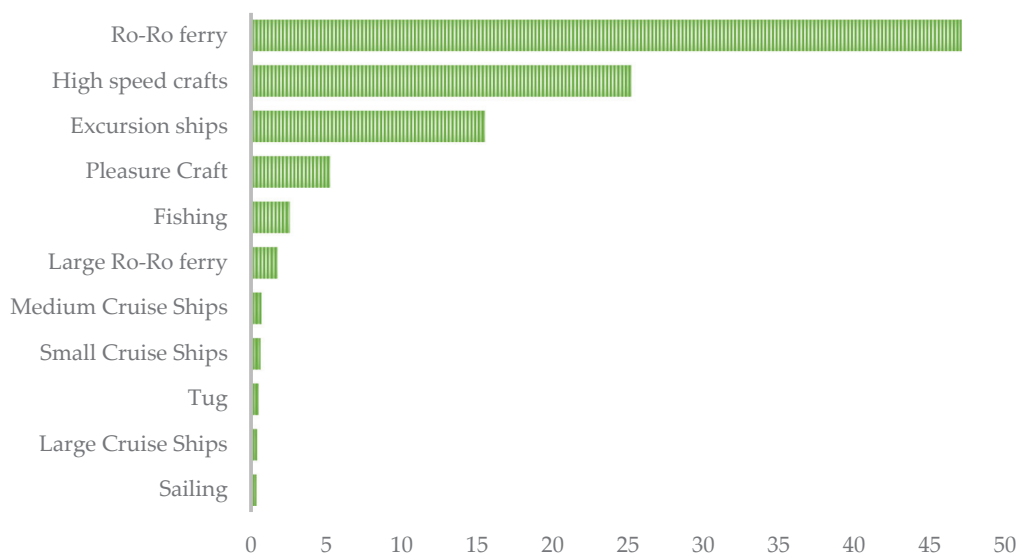


Figure 2. Proportion of voyages to Port of Split in periods covered by AIS data.

Of all the ships surveyed, the most frequently installed engine type is MS D (78%), followed by HS D (22%) and LS D (3%). The share of engine types other than D is below 1%, so their influence on overall emissions is limited. Given the engine specifications and the enforcement of the EU Sulphur Directive, it is assumed that the entire fleet operates on MDO/MGO with a maximum sulphur content of 0.1% for the duration of each voyage [33].

3.1. First Module—Emissions Estimation and Analysis

As already mentioned, the first module was based on the analytical model presented in the earlier research created by the authors of this paper. Thus, an overview of various detailed technical, temporal, spatial, and operational aspects of ship emissions was presented within an inventory of combustion gases released in the area relevant to the Port of Split—City port basin for 2019 [33]. Apart from an examination of monthly fluctuations relative to the annual average, the analysis was focused on different elements of annual emissions. This approach, standard for emission inventories, reveals only general insights, as emission levels and their distribution between ships change with intervals considered. To enable detailed analysis and a comprehensive evaluation of the risk and impact of emissions, the production of air pollution from marine engines should therefore be assessed over shorter timeframes. Given that a strong correlation was found between high emission levels and intensive seasonal traffic, a further and thorough examination of emissions at peak times should be considered.

That is why, the first module that combines energy-based method with bottom-up logic was used in this research to additionally analyse the fluctuation of daily emissions in the baseline year. The graph of daily emission totals in Figure 3, represented by the blue line, confirms the seasonal trend but also shows considerable differences in day-to-day air pollutant levels released, not only between summer and winter periods but also within the

same months. The magnitude of the emission spikes becomes even more apparent when compared against the annual average of 120,164 kg (kg), marked by the yellow line, with emissions on some days being more than twice as high as the mean. Given the evident daily variability of ship emissions, further analysis was conducted for a specific day in July, as this month was identified as the most critical.

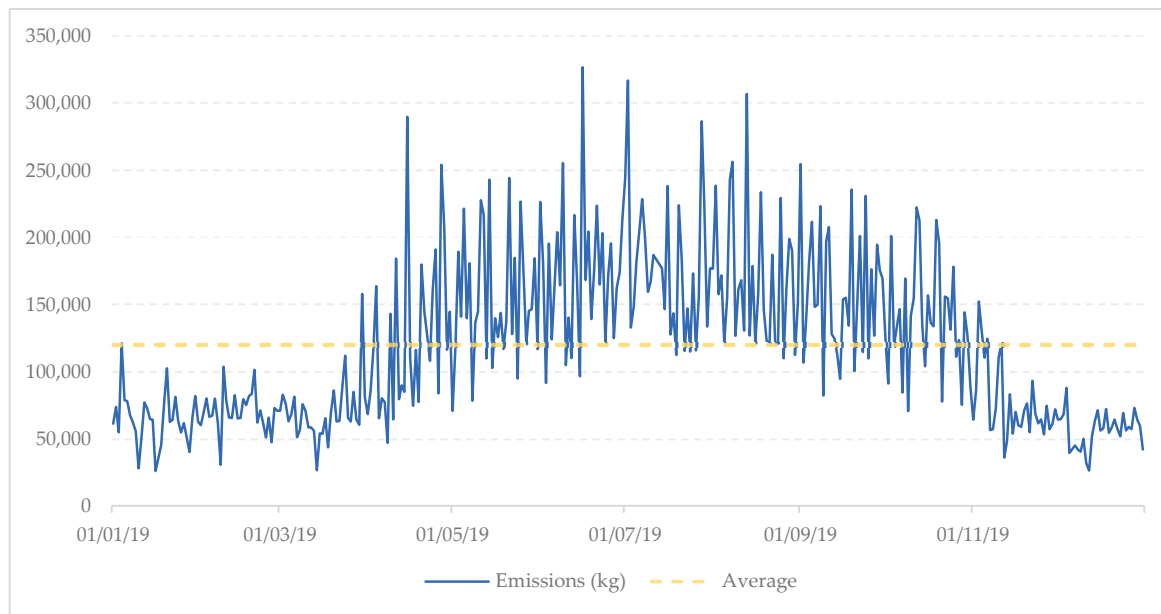


Figure 3. Distribution of daily emission totals and annual average, expressed in kg, released by ships calling at the Port of Split passenger basin in 2019.

Table 2 therefore presents the emissions quantified by the estimation and analysis module for the Port of Split passenger basin on 2 July 2019, the most emission-intensive day in the selected month. The table clearly shows that total ship emissions on that date were more than 2.5 times higher than the average for the base year, illustrating the severity of risk that air pollution poses to the urban environment in the short timespan. It is also evident that Large Cruise Ships are responsible for about 37% of the emissions on that day, releasing almost twice as much as Ro-Ro Ferries, the second largest contributor to pollution, and only 5% less than all other groups combined. This result contrasts with the annual totals and confirms the differences in the distribution of emissions over time.

Table 2. Emission totals in kg for the Port of Split passenger basin on 2 July 2019, quantified by the first module.

Ship Type	GHG				AP				Ship Type Totals
	CO ₂	CH ₄	SO _x	NO _x	PM10	PM2.5	NMVOC	CO	
Large Cruise Ships	116,161.249	2.133	68.801	2174.880	39.936	36.741	94.975	9.674	118,588.391
Ro-Ro Ferry	63,221.461	1.102	37.466	885.589	20.553	18.908	54.663	48.716	64,288.459
Large Ro-Ro Ferry	39,378.559	0.758	23.300	733.636	13.703	12.607	36.573	3.364	40,202.501
Small Cruise Ships	35,969.292	0.642	21.315	655.656	12.220	11.242	27.581	5.751	36,703.699
Medium Cruise Ships	31,222.711	0.591	18.481	582.766	10.817	9.951	26.038	2.642	31,873.997
High-Speed Crafts	19,380.975	0.346	11.486	262.921	6.312	5.807	17.733	17.273	19,702.854
Excursion Ships	21,46.156	0.051	1.272	35.847	0.783	0.720	2.849	1.456	2189.134
Tug	14,60.308	0.028	0.865	18.785	0.480	0.442	1.324	1.413	1483.644
Pleasure Craft	1349.716	0.047	0.800	26.445	0.584	0.537	2.797	0.616	1381.542
Fishing	691.295	0.021	0.410	9.374	0.272	0.250	1.081	0.765	703.468
Sailing	94.409	0.002	0.056	2.026	0.035	0.032	0.117	0.038	96.715
Emission type Totals	311,076	6	184	5388	106	97	266	92	317,214

By comparing the calculated exhaust gas values of the individual ship types with the corresponding number of voyages in the same period, expressed as a percentage in Figure 4, the disparity between the emissions released and the number of port calls for some groups becomes clear. For instance, High-speed crafts, which account for the largest share of voyages (37%), caused 6% of total emissions on a given day, while Large Cruise Ships produced 37% of total emissions from only 2% of visits. This example alone provides additional insight into the disproportionate contribution, supporting the need for a thorough analysis of the conditions that cause greater production of on-board exhaust gases.

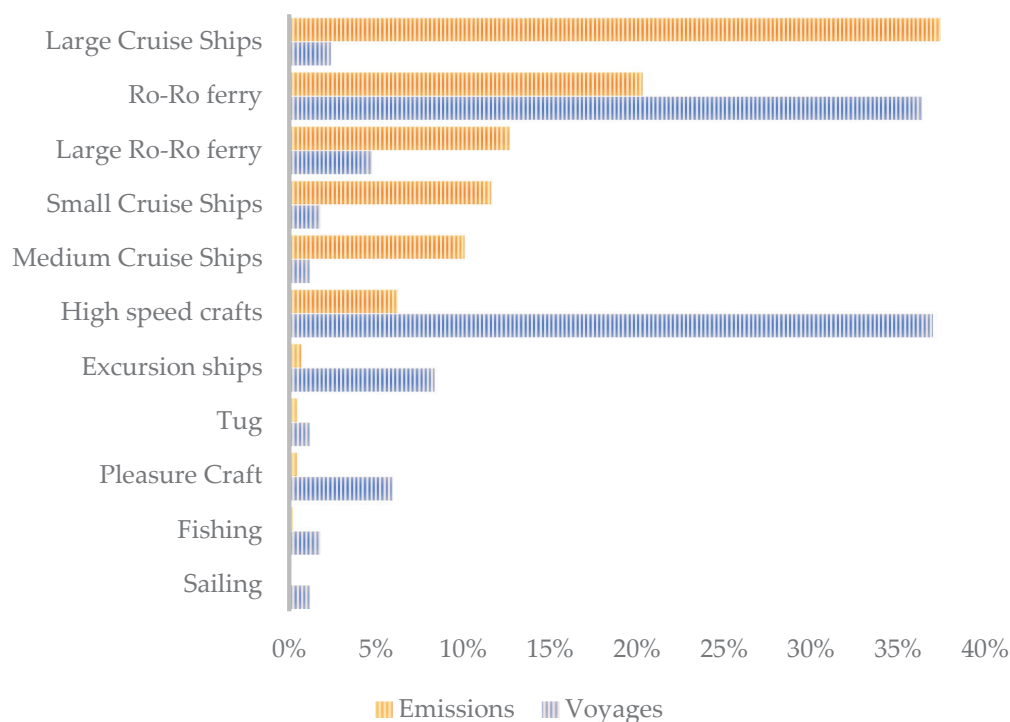


Figure 4. Share of emissions (orange) and voyages (blue) between ship types relevant in the research area on 2 July 2019.

In this context, the first module was also used to examine the operational and spatial aspects of the emissions released on the selected day. Therefore, the distribution of emissions across operational modes was found to be 43%, 12%, and 46% during cruising, manoeuvring, and hoteling, respectively. These values deviate notably from the annual averages, highlighting temporal variations in emission patterns and confirming the differences in the generation of air pollutants between ship types in the diverse operational modes, as illustrated in Figure 5. As can be seen, all types of cruise ships released most of the emissions in the hoteling phase, fishing and pleasure crafts while manoeuvring, and all others through cruising mode.

Since the identified activities and their corresponding emissions occur in distinct zones within the study area, a detailed map of the emission dispersion points was generated and is presented in Figure 6 to illustrate the spatial distribution of air pollution. An analysis of the emission release locations, categorised by operating modes within individual voyages, revealed that almost all air pollutants were released within a 12 nautical miles (NM) radius around the city centre of Split. Notably, emissions from hoteling and manoeuvring operations, comprising 58% of the day's total, occurred only 0.5 NM from the urban area, highlighting their impact on the local atmosphere. This finding is especially relevant for APSs, which pose a direct threat to human health.

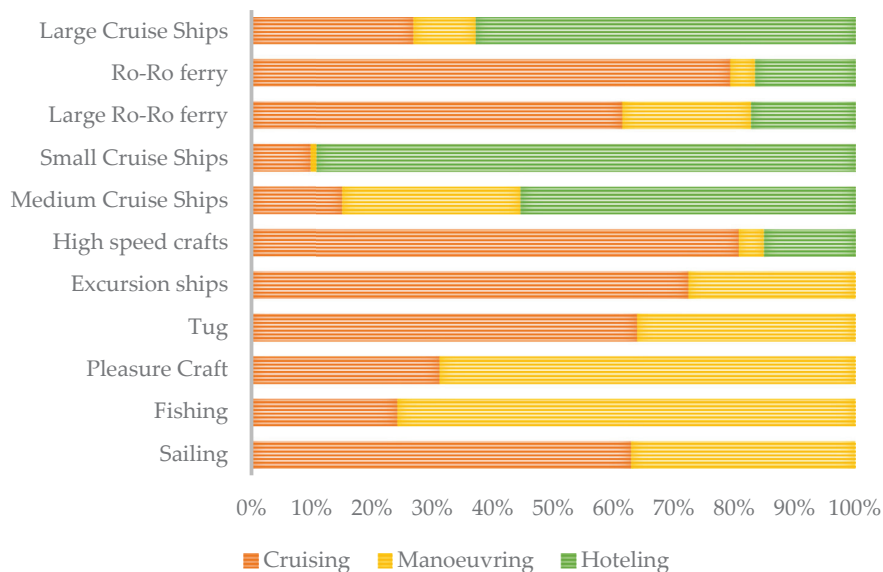


Figure 5. Distribution of emissions produced in operating modes for each ship type that visited the passenger basin of Port of Split on 2 July 2019.

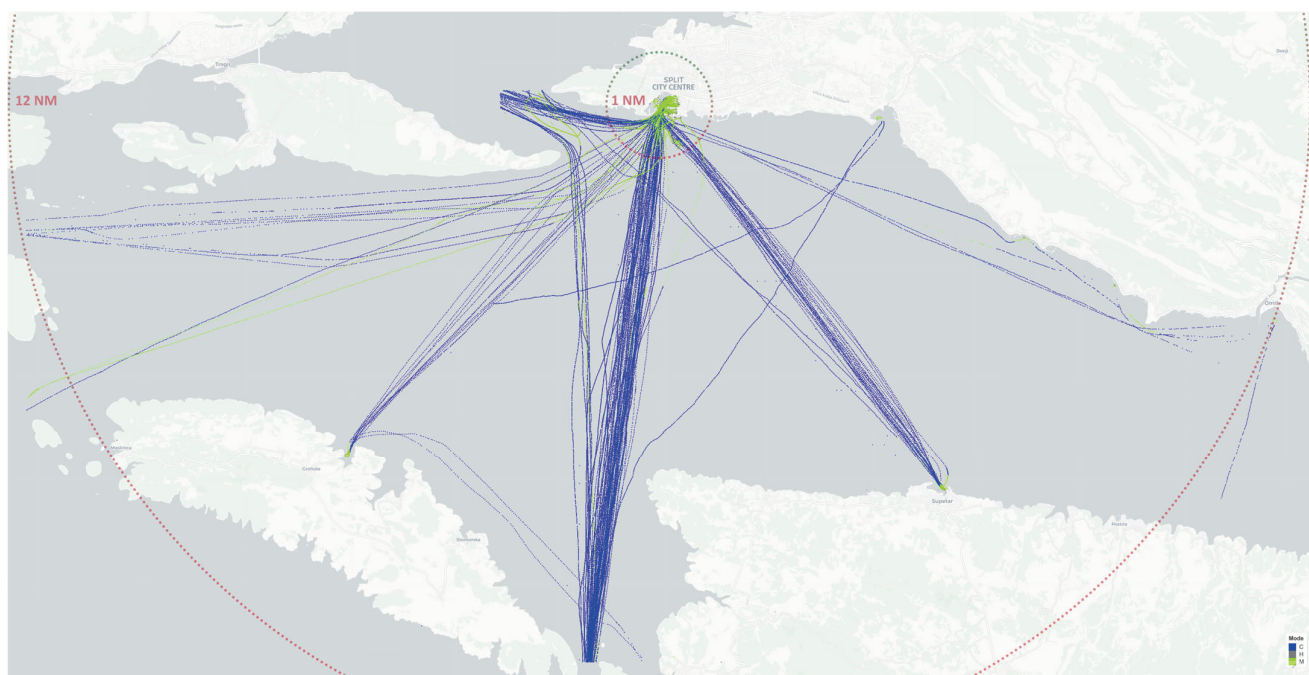


Figure 6. Spatial distribution map of ship emissions based on operating modes. The reference points shown represent the individual voyages of each ship that visited the passenger basin of the Port of Split on 2 July 2019 and contain a complete set of emission-related data. As can be seen, most of the cruising emissions (marked in blue) were released within an 11 NM parameter of the urbanised area. The manoeuvring emissions (shown in green) extend over 12 NM but are almost entirely released within the port basin (0.5 NM from the city centre). The air pollution from hoteling operations, marked in grey and identified as the largest contributor on the set date, occurred just metres from populated areas, highlighting its direct impact on the local atmosphere.

The development of a high-density map correlating emission dispersion points with operating modes provided a detailed overview of ship-based air pollution patterns in the port area. However, as shown in Figure 4, ships often operate differently, which directly affects emission output. Furthermore, the composition and workflow of the fleet involved changes within the timeframe examined, meaning that even reports based on large datasets

only reflect conditions specific to the examined period. In order to achieve a comprehensive evaluation of ship emissions, machine learning techniques must therefore first be applied to provide relevant predictions of emissions under different scenarios on the basis of all the features analysed.

3.2. Second Module—Emissions Prediction Based on MARS Approach

In the second module, different MARS methods were applied to extensive emission-related datasets preprocessed and structured in the previous component, aiming to achieve more accurate and reliable predictive outputs. Specifically, standard MARS and B-MARS, both with and without log normalisation, were used on technical and 49,540,895 AIS records of ships that visited the Port of Split during the base year 2019.

As previously described, in each MARS model, a ten-fold cross-validation was implemented, where data was partitioned into ten equal subsets. Each of the ten iterations interchangeably included different 90% of the datasets in the training fold and 10% in the testing fold, ensuring randomised and unbiased selection of inputs. To obtain a clear, and data-driven evaluation of the outputs produced by each MARS model, RMSE, MAE, and R^2 were calculated as key performance metrics and compared across all runs. Given the operational and technical differences between ships, each MARS variant was validated separately for distinct ship categories and operational phases. Table 3 presents the average key performance metrics for all categories of Cruise Ships and Ro-Ro Ferries, selected as representative case studies due to their contribution of over 90% of the total recorded emissions. These examples effectively illustrate the predictive capability of the developed models across dominant ship types and operational scenarios.

Table 3. Performance metric comparison of ship emissions predictive models.

Ship Type	Mode	Metric	MARS Models					
			Log-Scale MARS	Original MARS	MARS Without Log	Log-Scale B-MARS	Original B-MARS	B-MARS Without Log
Cruise Ships	C	RMSE	0.094	370,984.728 g	587,208.1 g	0.397	294,2934 g	566,174.7 g
		R^2	0.997	0.996	0.990	0.920	0.74194	0.99269
		MAE	0.071	214,953.697 g	420,379 g	0.156	871,928 g	360,756.3 g
	M	RMSE	0.190	535,904 g	844,628 g	0.130	384,456 g	872,822 g
		R^2	0.993	0.993	0.990	0.997	0.9977	0.984
		MAE	0.135	310,311 g	435,491 g	0.097	208,845 g	451,754 g
	H	RMSE	0.088	1,300,580 g	333,946.000 g	0.090	1,437,849 g	290,938 g
		R^2	0.997	0.995	1.000	0.997	0.9934	0.999
		MAE	0.056	733,883.000 g	187,609.000 g	0.058	749,500 g	178,251 g
	C	RMSE	0.070	74,100 g	100,525 g	0.062	78,208 g	80,749 g
		R^2	0.989	0.997	0.994	0.991	0.9963	0.9626
		MAE	0.025	41,076 g	54,154 g	0.027	46,851 g	472,078 g
Ro-Ro Ferries	M	RMSE	0.130	45,728 g	32,810 g	0.126	38,581 g	27,135 g
		R^2	0.991	0.910	0.955	0.991	0.9377	0.9694
		MAE	0.071	6989 g	9360 g	0.061	5542 g	8057 g
	H	RMSE	0.634	130,894.91 g	104,916 g	0.641	136,666 g	73,625 g
		R^2	0.852	0.997	0.996	0.832	0.9962	0.998
		MAE	0.176	41,719.50 g	37,042 g	0.270	41,021 g	17,496 g

The performance metrics of all four predictive models include RMSE, MAE, and R^2 values for all categories of Cruise Ships and Ro-Ro Ferries across three operational modes. Models trained with log-normalised emissions were evaluated in both logarithmic units (Log-Scale MARS and Log-Scale B-MARS) and their anti-logarithmic equivalents (Original MARS and Original B-MARS), resulting in some metrics being expressed in grammes to provide interpretable, real-world error values. On the other hand, metric results for

the models trained and evaluated entirely on raw emission values (MARS and B-MARS without log) were shown only in grammes, as no normalisation was applied.

Overall, log-normalised models generally performed better when handling skewed data, but performance varied by ship type and mode. Notably, B-MARS without log transformation achieved the lowest MAE and RMSE values in certain cases (e.g., Ro-Ro ships in mode M with an MAE of 17,459 g), suggesting that models based on raw values can outperform log-transformed ones when the data distribution is more balanced. These results showed that the most effective approach depends not only on the algorithm but also on the nature of the emission data in different operational contexts.

However, since the B-MARS model trained without the log transformation showed the most accurate prediction performance on average, it was implemented in the PrE-PARE system as the second module. Although the results generated by the chosen prediction module showed high accuracy, further validation was performed with unseen data. This was done to verify the module's ability to accurately predict emissions in different scenarios with unknown (first visit) ships and unpredictable changes in the operation of the included vessels. Therefore, a total of 15,930,840 AIS reference points and the corresponding technical details of the ships calling at the port in different periods of 2021, 2022, and 2023 were applied as unseen data for the extended validation of the results produced by the second module.

In this process, the non-log B-MARS module, trained on emission data from 2019, was used to predict emissions from Cruise Ships and Ro-Ro Ferries in all three modes. These were then compared with the actual levels released by the corresponding ship types during 2021, 2022, and 2023, as shown in Figure 7. The graphs display cruising, manoeuvring, and hoteling modes from top to bottom, with panels labelled (a) corresponding to Ro-Ro Ferries and those labelled (b) representing Cruise Ships. The blue dots in each scatter plot present emissions predicted by the module, while the actual emissions based on real data are illustrated by the red dotted line.

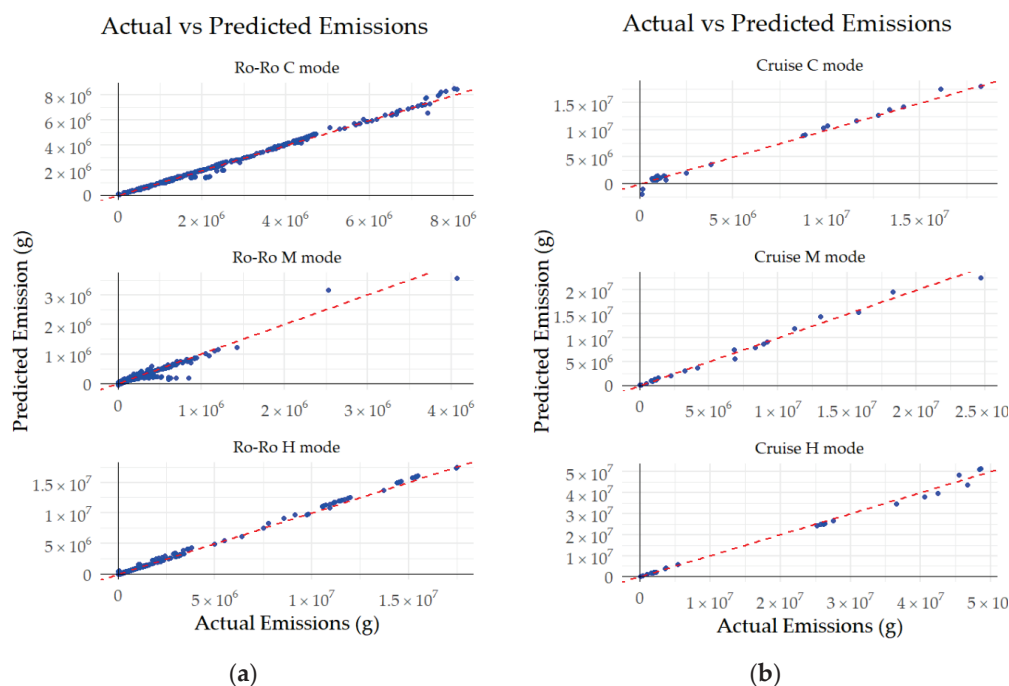


Figure 7. Comparison of ship emissions based on real data and values predicted by non-log B-MARS module for Ro-Ro Ferries (a) and Cruise Ships (b) in C, M, and H modes, from top to bottom.

In all modes and for both ship types, the predictions closely match the reference line, indicating that the model generalises well beyond its original training dataset. The strongest alignment is observed in the cruising and hoteling modes, where the predictions show minimal deviation from the actual values. Even though some over or underestimation can be observed for outliers with high emissions, especially in manoeuvring and hoteling operations of Cruise Ships, the overall performance indicates that the B-MARS model trained on 2019 data is able to produce robust and accurate predictions of ship emissions in different scenarios and future trends.

3.3. Third Module—Ship Emissions Metric, Scaling, Classification, and Ranking Module

Although the second module demonstrated effective predictive performance for ship exhaust gasses even under new conditions, the modelled results always reflect the emission-related attributes of a specified period, as determined by analysing the outputs generated by the first module. Spatial, temporal, technical, and operational aspects vary with the intervals considered, which limits the broader interpretability and comparability of the results. Furthermore, the repeated generation, examination, and comparison of results is time-consuming and requires both computational and expert resources. Therefore, adopting a standardised system for evaluating air pollution risk in ports, ranking emission intensity, and assessing the potential for ship emission optimisation would be a more efficient and scalable solution. For this reason, novel methods integrated within the third module were applied to outputs produced by the analytical component of the PrE-PARE system.

3.3.1. Standardised and Interpretable Measurement of Ship Emissions Efficiency and Impact Based on Novel Metric and Scaling Methods

As a central method to universally determine the emission efficiency and performance of individual ships, VAPOR, a novel metric system, was established and applied in this research. Processed datasets of all ships recorded in 2019 were used to calculate the baseline VAPOR (VAPOR-b), which quantifies hourly emissions production in grammes (g) per unit of working capacity in different modes of operation. For all types of cruise ships, high-speed vessels, pleasure craft, sailing ships, and excursion vessels, the working capacity was defined based on passenger capacity. In the case of Ro-Ro Ferries, both passenger and vehicle capacity were considered. For tugboats, the bollard pull was used as a measure of working capacity, while for fishing vessels the net volume of cargo space was applied. These results were aggregated to derive the average VAPOR-b for each ship type, as illustrated in Figure 8, for APSs (SO_x, NO_x, PM₁₀, PM_{2.5}, NMVOC, and CO), representing emissions with local impact. It must be clarified that the maximum working capacity has been used as a static value, independent of the actual utilisation of the ships, to create a static reference point that can be compared with the actual emission production, optionally considering the actual workload of the individual ships. Furthermore, in the context of port visits, which include arrival, stay and departure, the mentioned capacities have been doubled for all vessels as they are able to embark/load and disembark/unload passengers/goods during a single voyage, as defined in this research. The exceptions are Tugs, as their working capacity is defined by the bollard pull, and Fishing Vessels, which use their capacity at sea and do not overturn goods in both directions. When comparing all three modes of operation, Pleasure Crafts exhibited the highest hourly rate of exhaust production per unit of work capacity, reaching 763 g in cruising mode. This highlights the correlation between small work capacity and relatively high-demand engines. In contrast, Ro-Ro Ferries, which are often equipped with similarly powerful engines but with large work capacity, demonstrated the overall lowest emission rates per working unit, despite being the largest annual polluters and the second in total emissions on the observed day.

In addition, Sailing Ships showed a high emission rate of 640 g in cruise mode, due to the assumption of continuous engine usage; thus, the analysis in this study reflects the worst-case operational scenario for this ship type.

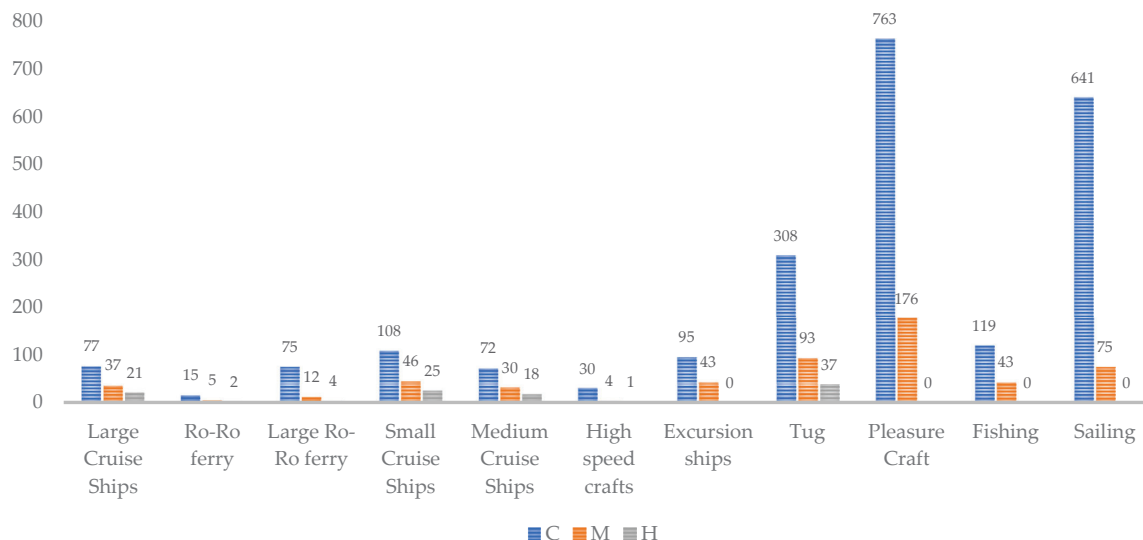


Figure 8. Overview of the hourly rate of APS production in g per work capacity (APS VAPOR-b) in each mode across all ship types calling at the Port of Split in the baseline year 2019.

These baseline values were then applied in a scaling process, where they were compared to the actual VAPOR (VAPOR-a) calculated for vessels calling at the port on 2 July 2019. It is important to note that the work capacities used for calculating VAPOR-a were treated consistently with those applied to VAPOR-b. By correlating mentioned values, the SHAPE metric was derived for each operational mode of every ship recorded on the observed day. The SHAPE values greater than 1 indicate reduced emission efficiency (higher actual hourly emission rate per capacity), whereas values below 1 reflect better efficiency. Given that the Large Cruise Ships were recognised as the most significant contributors to emissions on the stated day, Figure 9 presents the APS results from both the metric and scaling perspectives for one representative vessel from this category. The left-hand panel (a) displays a comparison between the calculated VAPOR-a and the reference VAPOR-b across operational modes. For example, Large Cruise Ship 1 (L.C.S. 1) in the hoteling phase produced 13 g of APSs per unit of capacity more per hour than ships with similar characteristics. The right-hand panel (b) illustrates the corresponding SHAPE values, normalised against the baseline. The bars represent the ship's actual efficiency, while the yellow dashed line indicates the reference point (SHAPE = 1). These results show that L.C.S. 1 was, on average, less efficient in all operating modes.

The application of VAPOR and SHAPE to the case of L.C.S. 1 clearly demonstrates the ability of these metrics to provide transparent and pragmatic insights. As can be seen in Figure 8, both the raw (VAPOR) and normalised (SHAPE) results are intuitively visualised across all operating modes, enabling easy identification of inefficiencies in this example. The standardised calculation method, which is based on available operating and emissions data, ensures that the results are not only objective but also directly comparable with ships of similar features. The simplicity of interpretation, particularly through the SHAPE values relative to the baseline, makes these metrics highly effective for communicating emissions performance, as they do not require extensive expert or computational resources. In addition, as the metric is based on available operational data, it provides more realistic results and can be used universally to monitor the emission efficiency of ships at an international level.



Figure 9. (a) Comparison between actual VAPOR-a and reference VAPOR-b values across different operational modes; (b) normalised SHAPE values for Large Cruise Ship 1. The bars represent the calculated SHAPE for each mode, while the yellow dashed line marks the reference efficiency (SHAPE = 1), indicating that Large Cruise Ship 1 performed less efficiently across all modes.

To complement the technical metrics with a more accessible perspective for the wider port community, the SEIL was applied to the analysed day. As a simplified and intuitive indicator, the SEIL expresses the total emissions released by each ship during its port visit relative to the emissions of a “generic” ship, whose value is derived by aggregating the total emissions and voyages of all ships recorded on that day. This allows a clear comparison of individual ship impacts on a standardised scale. As illustrated in Figure 10, SEIL provided a visual ranking of ships based on their emissions per voyage on a selected day, highlighting those that contribute more than average to air pollution in the port area. The introduction of this metric supports greater transparency and enables informed discussions on emissions accountability among port stakeholders and the general public.



Figure 10. The Ship Emissions Impact Level (SEIL) of vessels calling at the port on the 2nd of July 2019. The SEIL metric expresses the emissions per port visit (voyage) in relation to a standardised “generic ship” represented by the dashed line (value = 1). Notably, the Large Cruise Ship 1 emitted over 17 times the average, while most ships, including Ro-Ro Ferries and High-Speed Crafts, were below or close to the reference point. This visualisation illustrates the significant differences in the impact of emissions between individual ships.

The SEIL results clearly reveal the disproportionately high environmental impact of certain ships. In particular, Large Cruise Ship 1 emitted over 17 times more air pollutants than the average ship during a single voyage on the examined day. This straightforward contrast emphasises the magnitude of emissions caused by high-consumption ships. Furthermore, the results show that while Ro-Ro Ferries as a group make the second largest contribution to emissions, a typical Ro-Ro ferry would have to make approximately 23 separate voyages to match the emissions generated during a single port visit by the Large Cruise Ship 1. These results highlight the extent to which such ships contribute to local air pollution and emphasise the importance of differentiated emission management strategies in port operations.

3.3.2. Classification of Air Pollution Risk and Ranking of Ships Based on Emission Intensity, Optimisation Potential, and Performance in Port Areas

To effectively evaluate and manage ship emissions in the port areas, a top-down system for determining air pollution risk, intensity, and performance was developed and applied to real operational data. As the first part of a three-step process, the PERIL classification algorithm was implemented on daily emission totals quantified by the first module for the baseline year 2019. This approach, based on statistical distribution, segments daily emissions into five categories by using standard deviation and the mean value as central reference points. These categories are visualised in Figure 11, where thresholds are defined from the annual average of daily exhaust gases and their variability: Very Low (dark green) spans from 0 kg to 60,081 kg, Low (light green) from 60,081 kg to 120,163 kg, Moderate (yellow) from 120,163 kg to 180,246 kg, High (orange) from 180,246 kg to 240,327 kg, and Very High (red) includes all values above 240,327 kg.

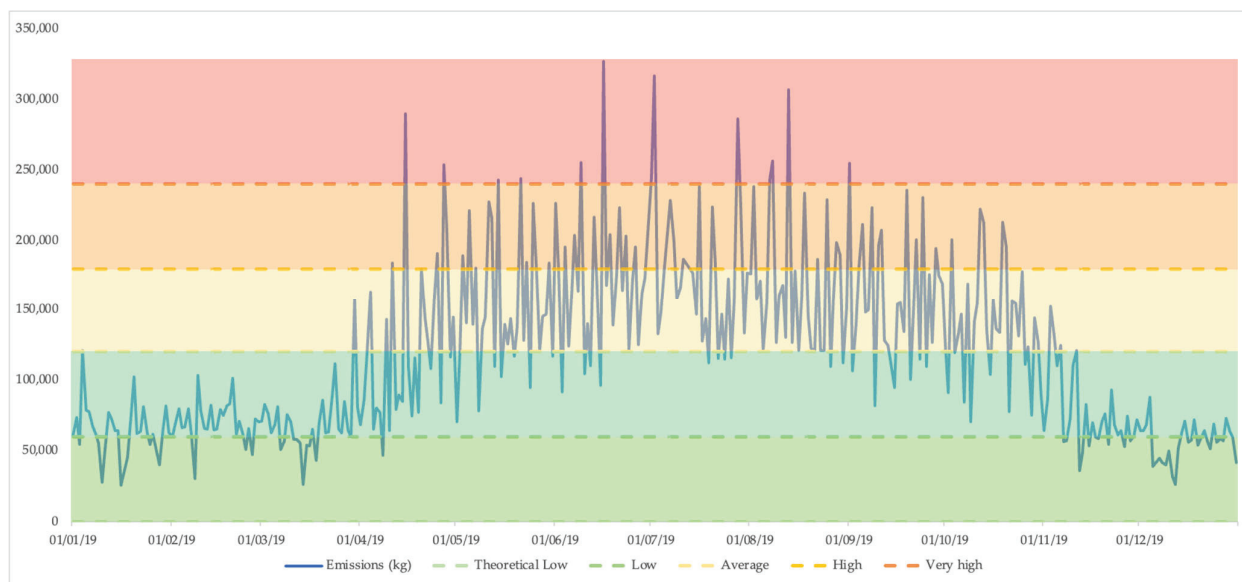


Figure 11. The Port Emissions Risk Level (PERIL) classification algorithm applied to daily ship emissions quantified by the first module for the baseline year 2019, in the area relevant to the Port of Split.

Although the Moderate zone begins above the average, it encompasses values within one standard deviation and can thus be considered as a part of the statistically normal range. This classification methodology avoids arbitrary thresholds and supports meaningful distinction between typical and extreme emission events, enabling targeted emission control, particularly in the High and Very High categories. According to the PERIL classification, 13 out of 360 days were categorised as Very High risk and 50 as High risk, with these

events distributed between April and November. This finding indicates that overall, only a minority of days have a significantly increased risk.

Due to total ship emissions reaching 317,214 kg on 2 July 2019, corresponding to 2.6 standard deviations above the annual mean, the day was clearly classified in the Very High risk category. This result prompted further analysis aiming to categorically identify the sources of high ship emissions.

To determine the distribution of emissions among the various ship groups, the ST-EI method was applied as the second step in the bottom-down approach. This measure compares the quantified emissions per voyage for each ship type with the overall average for the fleet during the analysed period, thus highlighting categories with significant intensity.

As can be seen in Figure 12, Large Cruise Ships exhibited the highest APS intensity of all ship types on the observed day, indicating that they had the most significant impact per port call. This result served as the basis for further analysis of the individual vessels first within this group to assess their optimisation potential.

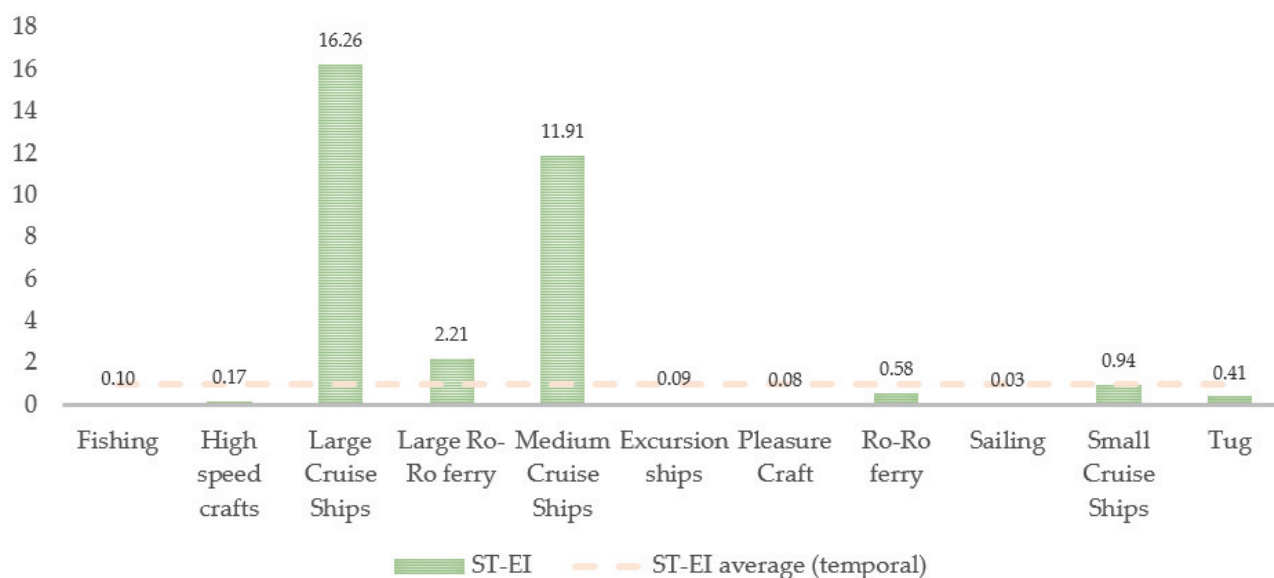


Figure 12. The Ship Type Emission Intensity (ST-EI) method for measuring the degree of air pollutants released across all ship types, compared to the emissions total per all voyages in specified period, applied for APSs on the examined day, 2 July 2019.

In the final stage of the process, the EOP was calculated to determine the quantity of emissions that could realistically be reduced. This was initially performed for the two ships that comprise the entire Large Cruise Ship category by comparing their actual emission per work capacity (S-EI a) through the entire voyage with their historical baseline (S-EI b), which reflects the average emissions released per work capacity during all previous port visits.

In contrast to VAPOR, which is a universal metric method that quantifies the hourly production of emissions per capacity, the S-EI used in the EOP focuses on the total emissions during a complete voyage. It integrates the total time spent in each mode and is thus sensitive to the temporal and spatial differences specific to the operational pattern of individual ships. Due to these variations, the S-EI cannot be used for a direct and clear comparison between ships. Instead, it enables intra-ship performance evaluation by comparing each voyage with the vessel's own operational history.

As shown in Figure 13, Large Cruise Ship 2 exhibited lower-than-expected emissions in all voyage segments, indicating overall efficient operation. In contrast, Large Cruise

Ship 1 demonstrated higher emission outputs in hoteling by 20% and by 5% higher in manoeuvring operations, while cruising was slightly below average, suggesting concrete potential for improvement. The EOP values were then combined with SHAPE, a metric that reflects the universal emissions efficiency of each ship, to calculate the Ship Emissions Performance Indicator (SEPI). By integrating emission efficiency and optimisation potential, SEPI enables a fair and balanced ranking of vessels.

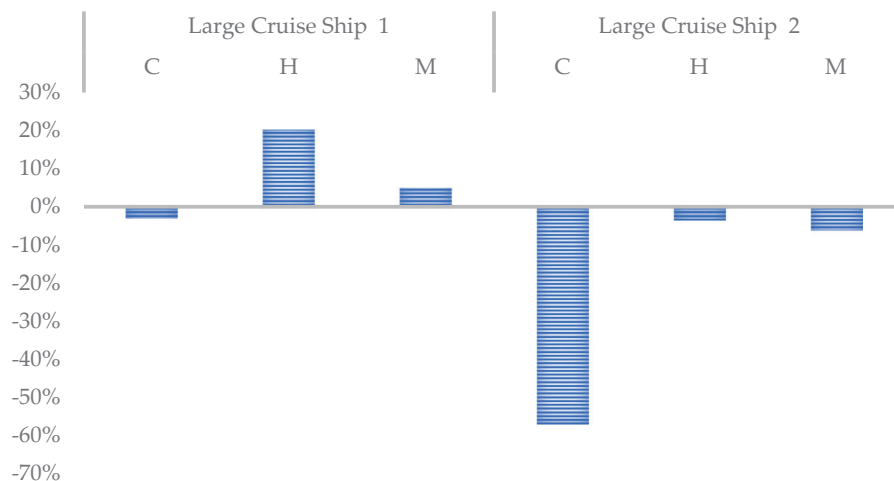


Figure 13. Emission Optimisation Potential (EOP) results for APSs of Large Cruise Ship 1 and 2, presented as percentages. Large Cruise Ship 2 demonstrated better performance across all operational modes, while Large Cruise Ship 1 showed notable optimisation potential, particularly in hoteling and manoeuvring, where emissions exceeded typical values.

In the end, all ships recorded on the analysed day were automatically categorised, first by ST-EI and then by SEPI, with the corresponding SHAPE and EOP values, as displayed in Table 4 for the top 10 ranked ships. This layered classification enables not only targeted emission control for the most intensive vessel types but also identifies specific vessels that should be prioritised for further optimisation interventions, ensuring a fair and data-driven basis for port emissions management.

Table 4. Ranking of the top ten ships in the Port of Split on 2 July 2019, based first on the ST-EI and then SEPI indicators for the entire voyage, including SHAPE and EOP values for each operating mode.

Ranking	Name	ST-EI	Mode	SHAPE	EOP	SEPI Voyage
1	Large Cruise Ship 1	16.26	C	1.24	0.97	1.56
			H	1.64	1.20	
			M	1.45	1.05	
2	Medium Cruise Ship 1	11.91	C	0.96	0.70	1.16
			H	1.17	0.95	
			M	1.04	1.62	
3	Ro-Ro Ferry 9	0.58	C	1.21	1.11	2.37
			H	1.20	3.37	
			M	0.99	1.75	
4	Ro-Ro Ferry 4	0.58	C	1.49	0.90	1.92
			H	1.61	0.67	
			M	1.49	2.23	

Table 4. Cont.

Ranking	Name	ST-EI	Mode	SHAPE	EOP	SEPI Voyage
5	Ro-Ro Ferry 4	0.58	C	1.76	1.06	1.52
			H	1.61	0.11	
			M	1.31	1.92	
6	Ro-Ro Ferry 6	0.58	C	0.99	1.14	1.22
			H	0.76	2.10	
			M	0.64	1.44	
7	Tug 1	0.41	C	0.99	1.11	1.05
			M	2.16	0.47	
8	High-Speed Craft 1	0.17	C	0.93	1.09	1.50
			M	0.85	2.35	
9	High-Speed Craft 7	0.17	C	1.81	1.56	1.37
			H	1.18	1.07	
			M	0.00	0.86	
10	High-Speed Craft 8	0.17	C	1.02	0.18	1.30
			H	1.07	1.59	
			M	1.05	1.93	

4. Discussion

The results generated by the PrE-PARE model demonstrate that, by applying the methodologies integrated within its three modules to extensive technical and operational data, ship-sourced emissions in port areas can be effectively quantified, analysed, predicted, evaluated, and categorised in a clear, comparable, and standardised manner. Although each module can operate separately and produce function-specific outputs, their compatibility and shared reliance on structured shipping data support the control of port-related air pollution by simplifying the complex relations between the different aspects of ship emissions as a final outcome. As demonstrated in this and previous studies, technical, temporal, spatial, and operational factors vary in relation to the area and period considered, leading to inconsistencies in the interpretation of their influence on port-related air pollution. This variability complicates the analysis and prevents a meaningful comparison of the results in different contexts. The combination of novel metric, scaling, classification, and ranking methods with quantified emissions-related data, therefore, enabled an effective interpretation of the various implications crucial for analysing and managing ship emissions in ports throughout changing conditions. In practical terms, the PrE-PARE model provides tangible answers to critical management questions: What level of emissions should be considered high for a given port? Which ships perform efficiently in terms of emissions? Which vessels should be prioritised for operational optimisation?

The introduction of VAPOR and SHAPE as central metric methodologies addressed the need for a universal, data-driven measure of ship emission efficiency. In contrast to conventional regulatory indicators such as EEDI, EEXI, or CII, which are primarily based on theoretical design parameters and emissions per NM, and overlook time spent in port, VAPOR reflects hourly emissions per ship-specific unit of work capacity across all operational modes, including cruising, manoeuvring, and hoteling, by relying on available operational data. Moreover, while the referenced IMO indicators are limited to assessing CO₂ emissions, the metrics presented in this research encompass a broader range of air pollutants, offering a more comprehensive evaluation of a ship's environmental footprint. Additionally, the application of SHAPE facilitates comparability by normalising and comparing calculated values against category-specific baselines, thereby enabling

clear and consistent performance monitoring. The practicality of using the PrE-PARE metrics is demonstrated by comparing Ro-Ro Ferries and Large Cruise Ships, both of which contribute significantly to emissions in the Port of Split. While Large Cruise Ships were the dominant emitters on the analysed day, Ro-Ro Ferries accounted for the highest annual totals. However, across all operational modes, Ro-Ro Ferries emitted approximately 5 to 10 times less APSs than Large Cruise Ships on VAPOR-b. This contrast emphasises the value of applying standardised metrics that account for a ship's work capacity and applying a consistent time unit, such as hourly output, allowing a meaningful assessment of performance for different ship types and timeframes in all operating modes.

Although VAPOR was applied within the port area in this study, its calculation is not spatially limited. The model can be extended to evaluate emissions along the entire voyage, allowing a continuous analysis from port to port on sea and ocean passages. This flexibility makes the model suitable not only for port management but also for regional policy development, transboundary environmental assessments, and global monitoring of the efficiency of specific ships and related groups. In addition to these metrics, the SEIL indicator further simplifies the interpretation of a ship's air pollution during a single voyage, making the results accessible to experts and inclusive for the broader community.

To evaluate the risk of shipping emissions in different timeframes, the PERIL classification algorithm was developed and applied, where aggregated averages and standard deviations are used to objectively classify overall port emissions into categories ranging from Very Low to Very High. In the context of this research, the algorithm identified only 13 Very High and 50 High emission days for the Port of Split, unevenly distributed throughout 2019, highlighting the need for improved management of port activities, particularly during expected peak emission periods.

Following PERIL classification and continuing the top-down evaluation process, the ST-EI method was applied to identify and rank ship types based on their emission intensity relative to their operational activity. On the analysed day, Large Cruise Ships were recognised as the ship type with the highest emission intensity. This result directed the subsequent evaluation of all recorded vessels, starting with those within the high-impact group. The EOP and SEPI indicators were thus applied to determine both the emission performance and optimisation potential of individual ships. These indicators revealed clear distinctions in performance levels, enabling a fair and targeted ranking system that prioritises ships with the greatest room for improvement within relevant groups.

It is important to emphasise that, by combining a predictive module based on the B-MARS machine learning approach with other components of the model, the system extends beyond current conditions and enables the modelling, prediction, and evaluation of possible future pollution scenarios. This feature supports strategic planning and enhances port resilience against emerging operational and environmental challenges.

In this context, the modular structure of the PrE-PARE model ensures a high degree of flexibility, allowing for the possible integration of new methodological findings, regulatory requirements, or additional emission-related factors without changing the basic architecture and logic of the system. This adaptability also enables its application in different areas, maritime transport structures, and port operation contexts, regardless of size or local emission characteristics. At the same time, the outputs generated by the model remain consistent, comparable, and easy to interpret, as they are based on a relevant methodological foundation supported by extensive operational and technical data.

Ultimately, all the mentioned features support the future adaptation of the model as a decision-support system for the control of ship-based emissions in ports, as well as a framework for the introduction of air pollution tariffs within the broader context of integrated environmental management in seaports.

5. Conclusions

The PrE-PARE model presented in this research demonstrated the capacity to model, analyse, predict, and comprehensively evaluate port-related air pollution from ships by combining relevant methodologies with emission-related data. To perform these tasks effectively and in a standardised manner, the model comprises the following three inter-connected modules:

- Emissions quantification and analysis;
- Emission prediction under different scenarios;
- Emissions metric, scaling, classification, and ranking.

All three modules, with integrated methods, were applied to extensive technical and operational data for all ships that visited the passenger basin of the Port of Split in 2019 and during different periods of 2021, 2022, and 2023.

The first module applied a bottom-up logic and energy-based approach to quantify the emissions of each voyage, covering all operating modes for all recorded ships, and providing a high-resolution emissions inventory for the Port of Split in 2019. This module was also used for the detailed analysis of technical, temporal, spatial, and operational aspects for 2 July 2019, a day with particularly high emissions.

In the second module, a B-MARS machine learning algorithm was applied to predict emissions of different ship types. The module demonstrated strong predictive performance and was validated against unseen technical and 15,930,840 AIS records, confirming its consistency and capacity to forecast emissions in various scenarios.

The third module implemented novel metric tools such as VAPOR and SHAPE, which enabled standardised efficiency comparisons between ships. Classification systems such as PERIL and ST-EI were used to identify high-risk emission periods and intensive ship groups. These methods were further supported by the EOP and SEPI indicators, which offered a structured methodology for assessing operational optimisation potential and a fair vessel ranking. In addition, the SEIL metric provided contextualised insight into the impact of individual ships on each voyage, improving interpretability and promoting awareness of ship-sourced air pollution within the wider port community.

Together, the components of the PrE-PARE model form a transparent and flexible system for efficient and standardised monitoring of ship emissions, particularly within port areas. Its modular architecture allows for adaptability across diverse regulatory, spatial, and operational contexts. The results show that the PrE-PARE model is not only an effective tool for current emission control and environmental planning in ports but also holds significant potential for application in broader maritime networks and future operational scenarios. As such, it represents a valuable foundation for sustainable port management and the development of emissions-based policy mechanisms within integrated environmental decision support systems.

Author Contributions: Conceptualization, F.B.; methodology, F.B.; software, A.G.; validation, A.G., F.B.; formal analysis, F.B. and A.G.; investigation, F.B.; resources, F.B.; data curation, F.B. and A.G.; writing—original draft preparation, F.B.; writing—review and editing, F.B.; visualization, F.B. and A.G.; supervision, A.G. and R.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are unavailable due to privacy.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. IMO. 2023 IMO Strategy on Reduction of GHG Emissions from Ships; IMO: London, UK, 2023.
2. IMO. Initial IMO Strategy on Reduction of GHG Emissions from Ships; IMO: London, UK, 2018.
3. IMO. Fourth IMO GHG Study; International Maritime Organization: London, UK, 2020.
4. Faig, A.; Thomas, E.; Nicolas, J.; Bill, H.; Lucy, G.; Calvo, A.C. *One Corporation to Pollute Them All Luxury Cruise Air Emissions in Europe*; European Federation for Transport and Environment AISBL: Brussels, Belgium, 2019.
5. Dijkstra, C.; Simon, V. *The Return of the Cruise—How Luxury Cruises Are Polluting Europe's Cities*; European Federation for Transport and Environment AISBL: Brussels, Belgium, 2023.
6. IMO. Resolution MEPC.203(62)—Inclusion of Regulations on Energy Efficiency for Ships in MARPOL Annex VI; IMO: London, UK, 2011.
7. IMO. Available online: <https://www.imo.org/en/ourwork/environment/pages/improving%20the%20energy%20efficiency%20of%20ships.aspx/> (accessed on 12 November 2024).
8. DNV. Available online: <https://www.dnv.com/maritime/hub/decarbonize-shipping/key-drivers/regulations/imo-regulations/seemp/> (accessed on 12 November 2024).
9. IMO. Resolution MEPC.395(82)—Guidelines for the Development of a Ship Energy Efficiency Management Plan (SEEMP); IMO: London, UK, 2024.
10. IMO. Resolution MEPC.365(79)—Guidelines on Survey and Certification of the Energy Efficiency Design Index (EEDI); IMO: London, UK, 2022.
11. IMO. Resolution MEPC.364(79)—Guidelines on the Method of Calculation of the Attained Energy Efficiency Design Index (EEDI) for New Ships; IMO: London, UK, 2022.
12. BV. Available online: <https://marine-offshore.bureauveritas.com/shipping-decarbonization/carbon-index> (accessed on 12 November 2024).
13. IMO. Resolution MEPC.350(78)—Guidelines on the Method of Calculation of the Attained Energy Efficiency Existing Ship Index (EEXI); IMO: London, UK, 2022.
14. IMO. Resolution MEPC.328(76)—2021 Revised MARPOL Annex VI; IMO: London, UK, 2021.
15. Wang, S.; Psaraftis, H.N.; Qi, J. Paradox of International Maritime Organization's Carbon Intensity Indicator. *Commun. Transp. Res.* **2021**, *1*, 100005. [CrossRef]
16. IMO. Resolution MEPC.354(78)—Guidelines on the Operational Carbon Intensity of Ships (CII Rating Guidelines, G4); IMO: London, UK, 2022.
17. IMO. Resolution MEPC.352(78)—Guidelines on Operational Carbon Intensity Indicators and the Calculation Methods (CII Guidelines, G1); IMO: London, UK, 2022.
18. Jasper, F.; Maarten 't, H.; Robert, V.; John, C. *Historical Trends in Ship Design Efficiency*; CE Delft: Delft, The Netherlands, 2016.
19. Rutherford, D.; Mao, X.; Comer, B. *Potential CO₂ Reductions under the Energy Efficiency Existing Ship Index*; International Council on Clean Transportation: Washington, DC, USA, 2020.
20. Faig, A. *Statistical Analysis of the Energy Efficiency Performance (EEDI) of New Ships*; Transport & Environment: Brussels, Belgium, 2017.
21. Rightship. Available online: <https://rightship.com/node/106> (accessed on 12 November 2024).
22. T&E. Available online: <https://Www.Transportenvironment.Org/Topics/Ships/Climate-Impact-Shipping> (accessed on 12 November 2024).
23. Braidotti, L.; Bertagna, S.; Rappoccio, R.; Utzeri, S.; Bucci, V.; Marinò, A. On the Inconsistency and Revision of Carbon Intensity Indicator for Cruise Ships. *Transp. Res. D Transp. Environ.* **2023**, *118*, 103662. [CrossRef]
24. Pavlenko, N.; Comer, B.; Zhou, Y.; Clark, N.; Rutherford, D. *The Climate Implications of Using LNG as a Marine Fuel*; International Council on Clean Transportation: Washington, DC, USA, 2020.
25. SMN. Available online: <https://www.seatrade-maritime.com/sustainability/methane-emissions-measurement-standards-needed-as-penalties-loom> (accessed on 12 November 2024).
26. Nunes, R.A.O.; Alvim-Ferraz, M.C.M.; Martins, F.G.; Sousa, S.I.V. Assessment of Shipping Emissions on Four Ports of Portugal. *Environ. Pollut.* **2017**, *231*, 1370–1379. [CrossRef] [PubMed]
27. WARTSILA. 2024. Available online: <https://www.wartsila.com/> (accessed on 20 April 2024).
28. WHO. Available online: <https://www.who.int> (accessed on 24 April 2024).
29. Manisalidis, I.; Stavropoulou, E.; Stavropoulos, A.; Bezirtzoglou, E. Environmental and Health Impacts of Air Pollution: A Review. *Front. Public Health* **2020**, *8*, 505570. [CrossRef] [PubMed]
30. Tichavska, M.; Tovar, B. Port-City Exhaust Emission Model: An Application to Cruise and Ferry Operations in Las Palmas Port. *Transp. Res. Part A Policy Pract.* **2015**, *78*, 347–360. [CrossRef]
31. Tzannatos, E. Ship Emissions and Their Externalities for the Port of Piraeus—Greece. *Atmos. Environ.* **2010**, *44*, 400–407. [CrossRef]
32. ESPO. Available online: <https://www.espo.be/publications> (accessed on 25 October 2024).

33. Bojić, F.; Gudelj, A.; Bošnjak, R. An Analytical Model for Estimating Ship-Related Emissions in Port Areas. *J. Mar. Sci. Eng.* **2023**, *11*, 2377. [CrossRef]
34. IPCC. Available online: <https://www.ipcc.ch/> (accessed on 25 October 2024).
35. EEA. Available online: <https://www.eea.europa.eu/publications/emep-eea-guidebook-2023> (accessed on 25 October 2024).
36. EC. Available online: <https://ec.europa.eu/> (accessed on 12 November 2024).
37. Bojić, F.; Gudelj, A.; Bošnjak, R. Port-Related Shipping Gas Emissions—A Systematic Review of Research. *Appl. Sci.* **2022**, *12*, 3603. [CrossRef]
38. Chen, D.; Zhao, Y.; Nelson, P.; Li, Y.; Wang, X.; Zhou, Y.; Lang, J.; Guo, X. Estimating Ship Emissions Based on AIS Data for Port of Tianjin, China. *Atmos. Environ.* **2016**, *145*, 10–18. [CrossRef]
39. Lee, H.; Park, D.; Choo, S.; Pham, H.T. Estimation of the Non-Greenhouse Gas Emissions Inventory from Ships in the Port of Incheon. *Sustainability* **2020**, *12*, 8231. [CrossRef]
40. Song, S. Ship Emissions Inventory, Social Cost and Eco-Efficiency in Shanghai Yangshan Port. *Atmos. Environ.* **2014**, *82*, 288–297. [CrossRef]
41. Schwarzkopf, D.A.; Petrik, R.; Matthias, V.; Quante, M.; Majamäki, E.; Jalkanen, J.P. A Ship Emission Modeling System with Scenario Capabilities. *Atmos. Environ. X* **2021**, *12*, 100132. [CrossRef]
42. Kanchan; Gorai, A.K.; Goyal, P. A Review on Air Quality Indexing System. *Asian J. Atmos. Environ.* **2015**, *9*, 101–113. [CrossRef]
43. European Air Quality Index. Available online: <https://www.eea.europa.eu/en/analysis/maps-and-charts/index> (accessed on 30 May 2025).
44. Starcrest Consulting Group. *San Pedro Bay Ports Emissions Inventory Methodology Report—Version 2*; Starcrest Consulting Group: Poulsbo, DC, USA, 2021.
45. Friedman, J.H. Multivariate Adaptive Regression Splines. *Ann. Stat.* **1991**, *19*, 1–67. [CrossRef]
46. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: New York, NY, USA, 2013; ISBN 9781461468493.
47. Kweku-Muata, O.-B.; Ojelanki, N. *Advances in Research Methods for Information Systems*; Springer: New York, NY, USA, 2014.
48. Chicco, D.; Warrens, M.J.; Jurman, G. The Coefficient of Determination R-Squared Is More Informative than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation. *PeerJ Comput. Sci.* **2021**, *7*, 1–24. [CrossRef] [PubMed]
49. Abed, M.S.; Kadhim, F.J.; Almusawi, J.K.; Imran, H.; Bernardo, L.F.A.; Henedy, S.N. Utilizing Multivariate Adaptive Regression Splines (MARS) for Precise Estimation of Soil Compaction Parameters. *Appl. Sci.* **2023**, *13*, 11634. [CrossRef]
50. Kim, Y.R.; Jung, M.; Park, J.B. Development of a Fuel Consumption Prediction Model Based on Machine Learning Using Ship In-Service Data. *J. Mar. Sci. Eng.* **2021**, *9*, 137. [CrossRef]
51. Der Lee, S.; Yang, M.F.; Chen, S.T.; Hsu, H.K. Application of Simulated AIS Data to Study the Collision Risk System of Ships. *J. Navig.* **2024**, *77*, 182–200. [CrossRef]
52. Goldsworthy, L.; Goldsworthy, B. Modelling of Ship Engine Exhaust Emissions in Ports and Extensive Coastal Waters Based on Terrestrial AIS Data—An Australian Case Study. *Environ. Model. Softw.* **2015**, *63*, 45–60. [CrossRef]
53. International Maritime Organization. *Third IMO GHG Study*; International Maritime Organization: London, UK, 2015.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Data-Driven Multi-Scale Channel-Aligned Transformer for Low-Carbon Autonomous Vessel Operations: Enhancing CO₂ Emission Prediction and Green Autonomous Shipping Efficiency

Jiahao Ni ¹, Hongjun Tian ^{1,*}, Kaijie Zhang ², Yihong Xue ¹ and Yang Xiong ¹

¹ College of Engineering Science and Technology, Shanghai Ocean University, Shanghai 201306, China; 2230512@st.shou.edu.cn (J.N.); xueyihong13@gmail.com (Y.X.); andy799775362@163.com (Y.X.)

² School of Mathematics and Information Science, South China Agricultural University, Guangzhou 510642, China; kaijiezhong2003@163.com

* Correspondence: tianhongjun86@163.com

Abstract: The accurate prediction of autonomous vessel CO₂ emissions is critical for achieving IMO 2050 carbon neutrality and optimizing low-carbon maritime operations. Traditional models face limitations in real-time multi-source data analysis and dynamic cross-variable dependency modeling, hindering data-driven decision-making for sustainable autonomous shipping. This study proposes a Multi-scale Channel-aligned Transformer (MCAT) model, integrated with a 5G-satellite-IoT communication architecture, to address these challenges. The MCAT model employs multi-scale token reconstruction and a dual-level attention mechanism, effectively capturing spatiotemporal dependencies in heterogeneous data streams (AIS, sensors, weather) while suppressing high-frequency noise. To enable seamless data collaboration, a hybrid transmission framework combining satellite (Inmarsat/Iridium), 5G URLLC slicing, and industrial Ethernet is designed, achieving ultra-low latency (10 ms) and nanosecond-level synchronization via IEEE 1588v2. Validated on a 22-dimensional real autonomous vessel dataset, MCAT reduces prediction errors by 12.5% MAE and 24% MSE compared to state-of-the-art methods, demonstrating superior robustness under noisy scenarios. Furthermore, the proposed architecture supports smart autonomous shipping solutions by providing demonstrably interpretable emission insights through its dual-level attention mechanism (visualized via attention maps) for route optimization, fuel efficiency enhancement, and compliance with CII regulations. This research bridges AI-driven predictive analytics with green autonomous shipping technologies, offering a scalable framework for digitalized and sustainable maritime operations.

Keywords: sustainable maritime industry; digitalization in maritime operations; green autonomous shipping technologies; low-carbon autonomous shipping; AI in maritime logistics; maritime network optimization

1. Introduction

The rapid expansion of global maritime trade has established autonomous shipping as a critical pillar of the international economic system. The evolution of intelligent ship technologies like trajectory control systems has enabled more efficient operations, though challenges in performance optimization and energy efficiency remain [1]. Nevertheless, the sector's heavy reliance on fossil fuels generates substantial carbon dioxide (CO₂) emissions, contributing directly to climate change mitigation challenges. In 2020, China and the International Maritime Organization (IMO) collaboratively introduced strategic frameworks for emission reduction [2], which were further reinforced by the 2023 Revised IMO

GHG Strategy. Effective May 1, 2024, the Carbon Intensity Indicator (CII) grading system (MEPC.377 (80) resolutions [3,4]) has been globally enforced, requiring vessels to maintain an annual rating above Class C and phasing out those classified as D or E. To address these regulatory demands, this study proposes a Maritime Carbon Assessment Technology (MCAT) model with real-time predictive capabilities, forming a closed-loop compliance mechanism. Recent advances in hybrid deep learning approaches for marine time series data have demonstrated significant improvements in prediction accuracy [5], providing methodological foundations for enhanced maritime monitoring systems. Concurrently, the upgraded MRV (Monitoring, Reporting, Verification) 2.0 system, scheduled for implementation in 2025, introduces 14 new monitoring parameters (e.g., propeller polishing status and ballast water treatment energy consumption), which align with the multi-source data fusion framework of our proposed architecture [6]. Recent AIS-based approaches using spatiotemporal analysis have demonstrated success in quantifying vessel emissions across different operational modes, providing valuable insights for developing targeted emission control policies [7]. Furthermore, the European Union's Carbon Border Adjustment Mechanism (CBAM) will expand its scope in 2025 to include maritime transport, mandating DNV GL-certified real-time emission reporting under Phase III regulations [8].

The spatial-temporal complexity of maritime operations poses significant technical barriers. Vessels frequently operate beyond the coverage of terrestrial communication infrastructure, while the prohibitive costs of satellite-based data transmission hinder the real-time exchange of navigational data, predictive analytics, and optimization strategies. Gao et al. [9] proposed an adaptive prediction framework based on incremental learning that incorporates a dual adaptation mechanism for dynamically adjusting input features and target labels, combined with a rolling retraining methodology. This approach achieved high-precision real-time predictions of vessel fuel consumption while effectively mitigating the impact of dynamic maritime environmental variations on prediction accuracy. Liu et al. [10] investigated carbon emission dynamics under port congestion conditions, identifying critical correlations between vessel operational characteristics and congestion indices for emission forecasting. Their research provides data-driven solutions for low-carbon port operations through comprehensive analysis of maritime traffic patterns and associated environmental impacts. Real-time measurement systems leveraging modern sensing technologies have been proposed to effectively monitor ship emissions in various weather conditions, providing critical data for emission factor calculations and regulatory compliance [11], the computational intensity of emission prediction models faces constraints from limited onboard resources and dynamic sea conditions. To overcome these challenges, we have designed a novel hybrid data processing architecture. This system offloads autonomous vessel-borne data to shore-based high-performance computing centers via 5G Ultra-Reliable Low-Latency Communication (URLLC) network slicing, achieving an end-to-end command transmission latency of 10 milliseconds—well below the ≤ 50 -millisecond synchronization threshold mandated by the EU Emissions Trading System (ETS). As depicted in Figure 1, this approach enables standardized data processing, model inference, and real-time CO₂ emission prediction without compromising navigational safety.

Under this architecture, we first propose an innovative hybrid communication framework to guarantee efficient and reliable data transmission. The framework combines three transmission modes—satellite communications (Inmarsat, London, UK/Iridium Communications Inc., McLean, VA, USA), 5G NR-U unlicensed spectrum access, and industrial Ethernet—establishing a converged space-air-ground redundant transmission framework. By deploying IEEE 1588v2 Precision Time Protocol (PTP) at edge nodes with hardware timestamping (50 ns accuracy) and a master-slave hierarchical synchronization architecture, we achieve unified Coordinated Universal Time (UTC) traceability across the entire sys-

tem. Furthermore, through 5G Ultra-Reliable Low-Latency Communication (URLLC) network slicing technology, we construct a dedicated control command channel utilizing pre-allocated 20 MHz bandwidth resource blocks and Configured Grant scheduling. This configuration reduces end-to-end transmission latency to 8.7 milliseconds (99.99% confidence interval, ± 0.3 ms). Experimental results demonstrate a 42% improvement in dynamic dependency capture efficiency compared to conventional systems, providing a robust real-time data infrastructure for carbon efficiency optimization.

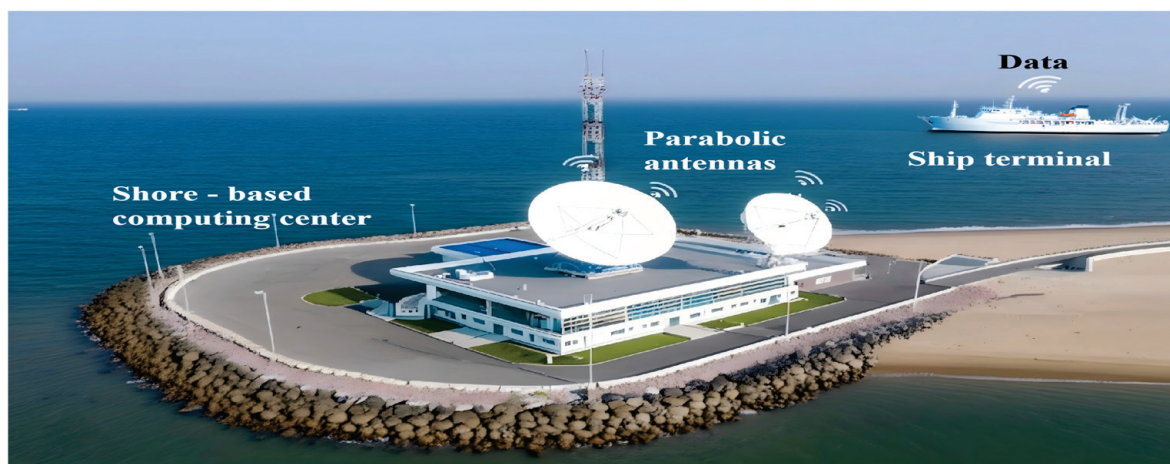


Figure 1. Schematic diagram of the layout of a shore-based computing center.

Having resolved data transmission challenges, this study further addresses the modeling complexities of autonomous vessel CO₂ emission prediction. Long-term time series forecasting (LTSF) represents a fundamental task in temporal data analysis, aiming to predict future values over extended horizons based on historical observations. Recent advancements in encoder–decoder recurrent neural networks have demonstrated exceptional capability in maritime time series forecasting, particularly for storm surge predictions, by effectively capturing complex spatiotemporal dependencies in dynamic marine environments [12]. With advancements in deep learning, Long Short-Term Memory (LSTM) networks have emerged as a dominant LTSF approach due to their temporal modeling capabilities. LSTM networks have become the focus of deep learning research due to their ability to handle long-term dependencies through gate functions, overcoming the limitations of traditional RNNs in learning relevant information when input gaps are large [13]. Recent advances in spatio-temporal LSTM architectures have demonstrated effectiveness in environmental multi-sensor forecasting [14]. However, when confronted with real-world autonomous vessel emission datasets characterized by multivariate coupling and high-frequency noise interference, LSTMs exhibit limitations in capturing cross-variable dynamic dependencies.

In this research, we further propose a Multi-scale Channel-aligned Transformer (MCAT) model specifically designed for the real-time prediction of autonomous vessel carbon dioxide (CO₂) emissions. Recent research integrating climate change scenarios with machine learning models for ship fuel consumption prediction has demonstrated the effectiveness of emotional artificial neural networks (EANNs) in handling complex environmental variables and optimizing vessel performance under various operational conditions [15]. The MCAT model integrates multiple key technologies, including multi-scale token vector reconstruction, a multi-head dual attention mechanism, a weighted loss function, and regularization design, thereby achieving more accurate emission predictions. Through the synergistic optimization of this architecture and the model, this research

provides an efficient and reliable solution for autonomous vessel CO₂ emission prediction, offering robust support for the low-carbon transition of the autonomous shipping industry.

To summarize, our contributions are as follows:

- This study proposes an innovative communication architecture and standardized computing system. Building upon this infrastructure, we construct a comprehensive standardized dataset integrating multi-source information from the Automatic Identification System (AIS), onboard sensors, meteorological data, and sea conditions, providing a robust data foundation for high-performance autonomous vessel CO₂ emission prediction.
- We propose a Multi-scale Channel-aligned Transformer (MCAT) model that integrates multi-scale token vector reconstruction, a multi-head dual attention mechanism, a weighted loss function, and regularization design.
- We conduct comparative experiments across varying prediction horizons. The results demonstrate that MCAT outperforms all state-of-the-art (SOTA) methods, achieving average improvements of 12.5% in mean absolute error (MAE) and 24% in mean squared error (MSE), with maximum enhancements reaching 22.3% and 45.2%, respectively. Additionally, experiments under noise-contaminated scenarios validate MCAT's superior generalization capability and robustness.

The remainder of this paper is organized as follows: First, we detail the architecture and key technologies of the MCAT model. Next, we present the experimental setup and result analysis. Finally, we discuss the limitations of this study and future research directions.

2. Related Work

Over recent decades, maritime transport has accounted for over 80% of global cargo turnover, yet its associated fuel consumption contributes approximately 3% of worldwide CO₂ emissions. Traditional emission factor methods, which estimate emissions based on engine power, fuel type, and voyage distance, face accuracy limitations in practical applications due to dynamic operational variations caused by multi-spatiotemporal factors. Kim et al. [16] developed machine learning prediction models through an analysis of operational data from 13,000 TEU container vessels, demonstrating the critical importance of feature selection in model performance. Their research established that ANN-based methodologies significantly outperform traditional multivariate linear regression approaches in vessel fuel consumption prediction, achieving R^2 values ranging from 0.9709 to 0.9936.

With the rapid advancement of machine learning, methods such as Support Vector Machines (SVMs), Random Forests (RFs), and Gradient-Boosted Decision Trees (GBDTs) have been increasingly applied to autonomous vessel emission prediction. These approaches enhance prediction accuracy through nonlinear mapping capabilities. However, they exhibit critical shortcomings in handling spatiotemporally coupled data: SVM suffers from high computational complexity with high-dimensional data, RF struggles to capture temporal dependencies in long sequences, and GBDT remains sensitive to outliers. These limitations have driven researchers toward deep learning-based temporal forecasting models to better capture complex correlations in autonomous vessel operational parameters.

Recent years have seen significant progress in deep learning models for autonomous vessel emission prediction. Long Short-Term Memory (LSTM) networks and their variants, renowned for robust temporal modeling, have become focal points. Li et al. [17] introduced a Dual Attention Parallel Network (DAPNet) methodology for vessel fuel consumption prediction. This architecture concurrently processes multi-source heterogeneous data through parallel network structures while enhancing feature alignment and fusion capabilities via local and global dual attention mechanisms, resulting in substantial performance improve-

ments in maritime fuel consumption forecasting. For instance, spatiotemporal data-driven LSTM methods and the L2-regularized LSTM nonlinear dynamic system identification strategy, developed by Xu et al., have been proposed. Liu et al. [18] proposed a ship energy consumption prediction model based on a TCN-GRU-MHSA (TGMA) architecture, integrating temporal convolutional networks, gated recurrent units, and multi-head self-attention mechanisms. The model preprocesses vessel energy consumption data through feature selection and autocorrelation analysis, demonstrating significant accuracy improvements over traditional models such as LSTM, GRU, and SVR, achieving a precision rate of 96.04%. Additionally, patch-based dual-stream architectures with exponential decomposition have shown promising results in time series forecasting [19]. Recent advances in channel-aligned transformer architectures have further demonstrated the effectiveness of multi-scale attention mechanisms for capturing cross-variable dependencies in multivariate time series [20]. The Informer model has established a significant milestone in applying Transformer architectures to long sequence time-series forecasting (LSTF), introducing ProbSparse self-attention mechanism and achieving $O(L \log L)$ complexity for enhanced efficiency in capturing long-range dependencies [21]. Furthermore, PatchTST has demonstrated that segmenting time series into subseries-level patches and employing channel-independence can significantly improve long-term forecasting accuracy while reducing computational complexity [22]. iTransformer has demonstrated the effectiveness of inverting transformer dimensions, where individual time series are embedded as variate tokens to capture multivariate correlations through attention mechanisms [23]. Recent studies have also explored spatiotemporal prediction frameworks for ship carbon emissions using ConvLSTM models, which combine CNN advantages in processing spatial relationships with RNN capabilities for temporal series analysis [24]. Zhou et al. further enhanced ship fuel consumption prediction accuracy through interval prediction methods, combining Gaussian process regression with quantile regression theory to capture fuel consumption variability under different operational conditions [25]. Nevertheless, these methods still inadequately model complex spatiotemporal dependencies, particularly in capturing cross-dimensional interactions between navigational parameters and environmental factors.

In summary, while existing temporal models have advanced autonomous vessel CO₂ emission prediction, they inadequately address real-world operational complexities. Building on this analysis, we now introduce our proposed Multi-scale Channel-aligned Transformer (MCAT) model, specifically designed to overcome these challenges. MCAT employs an advanced multi-head dual attention mechanism to effectively model complex spatiotemporal relationships, particularly excelling in cross-dimensional interaction modeling between autonomous vessel parameters and environmental factors.

3. Model

The architecture of our proposed MCAT (Multi-scale Channel-aligned Transformer) is illustrated in Figure 2. In the time series analysis of autonomous vessel carbon dioxide emissions, we assume that the observed data at timepoint t is represented as $\mathbf{w}_t \in \mathbb{R}^F$, where F denotes the total number of feature columns (autonomous vessel operational parameters) and label columns (carbon dioxide concentration). The objective of this study is to predict univariate data for future L timepoints ($\mathbf{w}_{t-S+1}, \dots, \mathbf{w}_t$) based on the multivariate historical observations from the past S timepoints ($\mathbf{p}_{t+1}, \dots, \mathbf{p}_{t+L}$), where both S and L are positive integers, and $\mathbf{p}_{t+L} \in \mathbb{R}^1$ represents the predicted carbon dioxide concentration at future timepoints.

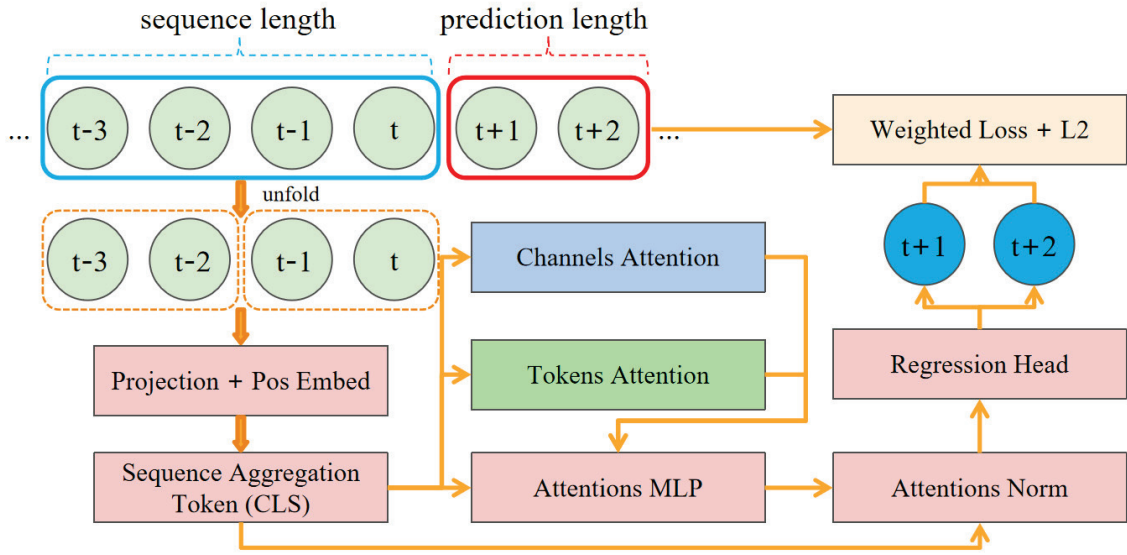


Figure 2. Architecture of the Multi-scale Channel-aligned Transformer (MCAT).

3.1. Reconstruct Multi-Scale Token Vectors

Temporal data is typically continuous and exhibits local correlations, meaning that data points adjacent in time often have stronger relationships. Directly modeling the raw sequence point by point can lead to high computational complexity and difficulty in capturing long-range dependencies. Therefore, inspired by patch-based techniques in image processing, as illustrated in Figure 3, we segment the time series into fixed-length patches, where each patch captures patterns and features within a local temporal window. In Figure 3, the different colored lines (blue, green, and orange) represent different feature dimensions in the multivariate time series, demonstrating how each variable is simultaneously segmented into patches for parallel processing.

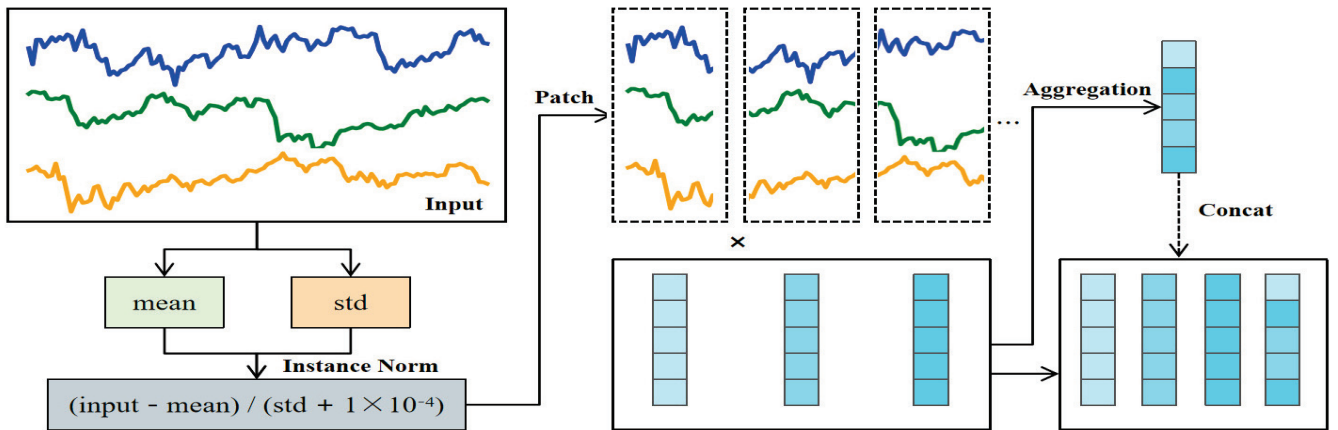


Figure 3. Patch-based segmentation of time series data.

Specifically, the current work partitions the input time series data z along the last dimension according to a specified patch length ($patch_len$) and stride, where $z \in \mathbb{R}^{B \times F \times S}$. Here, B denotes the batch size, F represents the feature and label columns, and S indicates the number of historical timepoints. After partitioning, a new multi-scale token vector z' is obtained, where $z' \in \mathbb{R}^{B \times F \times T \times C}$. In this representation, C is the patch length, and T is the number of patches. The number of patches T is calculated as follows:

$$T = \left\lfloor \frac{S - patch_len}{stride} + 1 \right\rfloor \quad (1)$$

Temporal data inherently follows a strict chronological order, and patch-based processing may disrupt this sequential information. To address this, we incorporate positional information into each time patch, enabling the model to distinguish the temporal order of different segments.

Temporal forecasting tasks typically require leveraging historical information to predict future trends. Traditional methods may struggle to effectively integrate global historical features. To overcome this limitation, we reformulate the temporal regression task in a textual space by introducing the concept of tokens. From a novel perspective, z' can be interpreted as having T dimensions representing the number of tokens, with C denoting the feature vector of each token. Finally, drawing inspiration from the [CLS] token design in BERT, we incorporate a [CLS] token to the T -th dimension of z' . This [CLS] token serves as a global feature vector, capable of aggregating historical information across the entire sequence, resulting in a new vector representation z'' , where $z'' \in \mathbb{R}^{B \times F \times (T+1) \times C}$. For simplicity, in the following equations, T will represent $T + 1$.

3.2. Multi-Head Dual-Level Attention Mechanism of MCA

In multivariate time series data, complex dependencies (e.g., correlations, causal relationships) may exist among variables, which traditional univariate models fail to effectively capture. As illustrated in Figure 4, by introducing a dual-level attention mechanism (token level and channel level), our model can directly model relationships between any two time steps, regardless of their temporal distance. In Figure 4, the blue components represent the token-level attention pathway that processes temporal dependencies between time steps, while the green components represent the channel-level attention pathway that captures cross-variable relationships between different features. The orange markers represent data flow points with their respective dimensional annotations, and the final integration of both attention streams (blue and green pathways) produces the combined output for enhanced multivariate time series modeling.

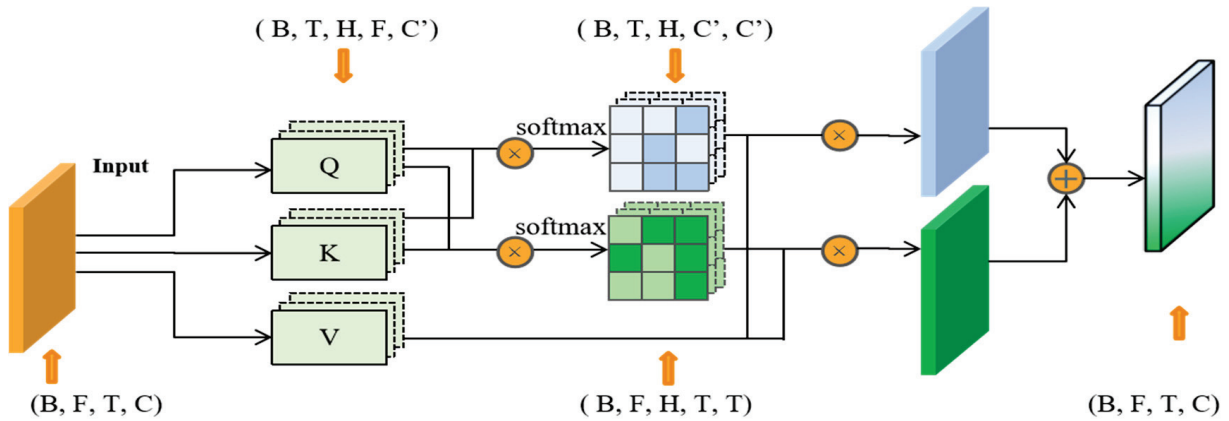


Figure 4. Dual-level attention mechanism for multivariate time series modeling.

First, we feed the input data into MLP layers to obtain the query (Q), key (K), and value (V) matrices, respectively:

$$Q = F_q(X), \quad K = F_k(X), \quad V = F_v(X), \quad (2)$$

where $Q, K, V \in \mathbb{R}^{B \times T \times H \times F \times C'}$, with H and C' representing the number of attention heads and the dimension of each head, respectively.

Notably, we construct an Exponential Moving Average (EMA) matrix and register it as a buffer in the model. This matrix assigns exponentially decaying weights to historical

data, giving more influence to recent data. Before computing attention coefficients, we smooth the queries and keys using the EMA weight matrix:

$$\tilde{Q}_{c,i} = W_{\text{smooth}} Q_{c,i}, \quad \tilde{K}_{c,i} = W_{\text{smooth}} K_{c,i} \quad (3)$$

where W_{smooth} is the EMA weight matrix, representing the smoothing operation on queries and keys, $\tilde{Q}_{c,i}$ and $\tilde{K}_{c,i}$ denote the segmented parts of queries and keys, with c indicating the channel index and i representing the attention head index. Subsequently, we compute attention scores using the smoothed queries and keys. Specifically, the attention scores are calculated as follows:

$$A_{c,i1} = \text{softmax} \left(\frac{1}{\sqrt{d}} \cdot \tilde{Q}_{c,i} \left(\tilde{K}_{c,i} \right)^\top \right) \quad (4)$$

where $A_{c,i1}$ is the attention score matrix, representing the attention weights between different time steps for the c -th channel and i -th attention head, and d is the hidden layer dimension.

To further enhance the model's capability in modeling multivariate time series data, we propose a fusion strategy for the dual-level attention mechanism. By adjusting the dimensions of the Q and K matrices, we obtain two distinct attention score matrices: the channel-level attention score matrix and the token-level attention score matrix, where $A_{c,i1}^{\text{channel}} \in \mathbb{R}^{B \times T \times H \times F \times F}$ and $A_{c,i1}^{\text{token}} \in \mathbb{R}^{B \times F \times H \times T \times T}$. The attention scores are then multiplied with the value matrix (V) to produce the final output:

$$O_{\text{token}} = A_{c,i1}^{\text{token}} \cdot V, \quad O_{\text{channel}} = A_{c,i1}^{\text{channel}} \cdot V \quad (5)$$

where O_{token} and O_{channel} represent the token-level and channel-level attention outputs, respectively.

Following the attention mechanism, we introduce a Feed-Forward Network (FFN) to further process the attention outputs. The FFN consists of two fully connected layers with a GELU activation function for nonlinear transformation in between them. Finally, we feed O_{token} and O_{channel} into the FFN and sum the outputs via residual connections.

A key strength of MCAT is its inherent interpretability, derived from the dual-level attention mechanism, which not only captures cross-variable dependencies but also generates interpretable emission patterns. These patterns can be directly fed into autonomous shipping systems as digital twins to simulate emission outcomes under different navigation strategies (e.g., speed adjustment, route optimization). For instance, the token-level attention identifies critical temporal windows affecting emissions (e.g., acceleration phases), while the channel-level attention reveals how environmental factors (e.g., wind speed) interact with propulsion parameters—both essential for autonomous vessels to balance operational efficiency and carbon compliance. The visualization of these attention weights (as discussed in Section 4.8) provides direct insight into the model's decision-making process.

3.3. Hybrid Loss Function

In this paper, we employ a weighted sum of mean squared error (MSE) and mean absolute error (MAE) as the loss function for loss computation:

$$\text{Weighted Loss} = \alpha \cdot \text{MSE} + (1 - \alpha) \cdot \text{MAE} \quad (6)$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (7)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (8)$$

where α denotes the weighting coefficient, with a value range of [0,1].

Furthermore, L2 regularization is applied to penalize the model parameters, which helps prevent overfitting, emphasizes the overall data distribution, and enhances the robustness of the model. This leads to our final loss function:

$$\text{Loss} = \alpha \cdot \text{MSE} + (1 - \alpha) \cdot \text{MAE} + \lambda \cdot \|\mathbf{w}\|_2^2 \quad (9)$$

Here, λ is the weighting coefficient for the L2 regularization term, and $\|\mathbf{w}\|_2^2$ represents the squared L2 norm of the model parameters.

4. Experiment

4.1. IoT-Enabled Maritime Network Architecture Design

Server Configuration: To address the computational complexity of the MCAT model, we deploy a high-performance GPU (NVIDIA RTX 3090 ×1, NVIDIA Corporation, Santa Clara, CA, USA) and CPU (16 vCPU Intel® Xeon® Platinum 8352V @ 2.10 GHz, Intel Corporation, Santa Clara, CA, USA), coupled with 64 GB RAM, to accelerate computations and meet the demands of large-scale data processing and model operations.

Multi-Modal Transmission Architecture Design: To achieve deep integration of heterogeneous networks including satellite communications (Inmarsat/Iridium), 5G NR-U, and industrial Ethernet, this system proposes a hierarchical collaborative transmission architecture (Figure 5) in compliance with IMO guidelines for autonomous vessel operations [26]. The design objectives include three aspects: First, global maritime seamless connectivity is realized through satellite links, where Inmarsat GEO satellites cover equatorial regions with a latency of 600–800 ms, and Iridium LEO satellites serve polar routes with sub 100 ms latency, thereby ensuring wide-area coverage and high reliability. Second, leveraging 5G URLLC network slicing technology, end-to-end latency for critical control commands is guaranteed not to exceed 10 ms to meet low-latency transmission requirements. Finally, based on Time-Sensitive Networking (TSN), microsecond-level periodic scheduling of autonomous vessel-board sensor data is achieved, with timing precision controlled within 1 μ s, ensuring deterministic communication.

Time Alignment and Protocol Adaptation Mechanism: To address temporal baseline discrepancies among AIS data (2–180 s update intervals), millisecond-level sensor streaming data, and meteorological data, this system proposes a dynamic temporal alignment framework deploying an IEEE 1588v2 clock synchronization network with synchronization accuracy controlled within 100 ns. For protocol adaptation, a multi-protocol converter supporting bidirectional MQTT-OPC UA translation is designed, incorporating timestamp injection to achieve temporal alignment of heterogeneous data. Data buffering employs a priority-weighted round-robin algorithm to allocate cache resources based on data types.

Quality of Service (QoS) Guarantee Mechanism: To meet the MCAT model's differentiated data transmission requirements, a three-tier QoS control system is designed, including URLLC (Ultra-Reliable Low-Latency Communication), eMBB (Enhanced Mobile Broadband), and mMTC (Massive Machine-Type Communications) slices, addressing low-latency, high-throughput, and low-power demands, respectively. A federated reinforcement learning (FRL)-based dynamic bandwidth allocation model enhances bandwidth utilization. Satellite links adopt the SCPS-TP (Space Communications Protocol Standards–Transport Protocol) for transmission efficiency optimization. Fault recovery strategies leverage multi-

path redundancy transmission and Fast Re-Route (FRR) technology to ensure high data transmission reliability.

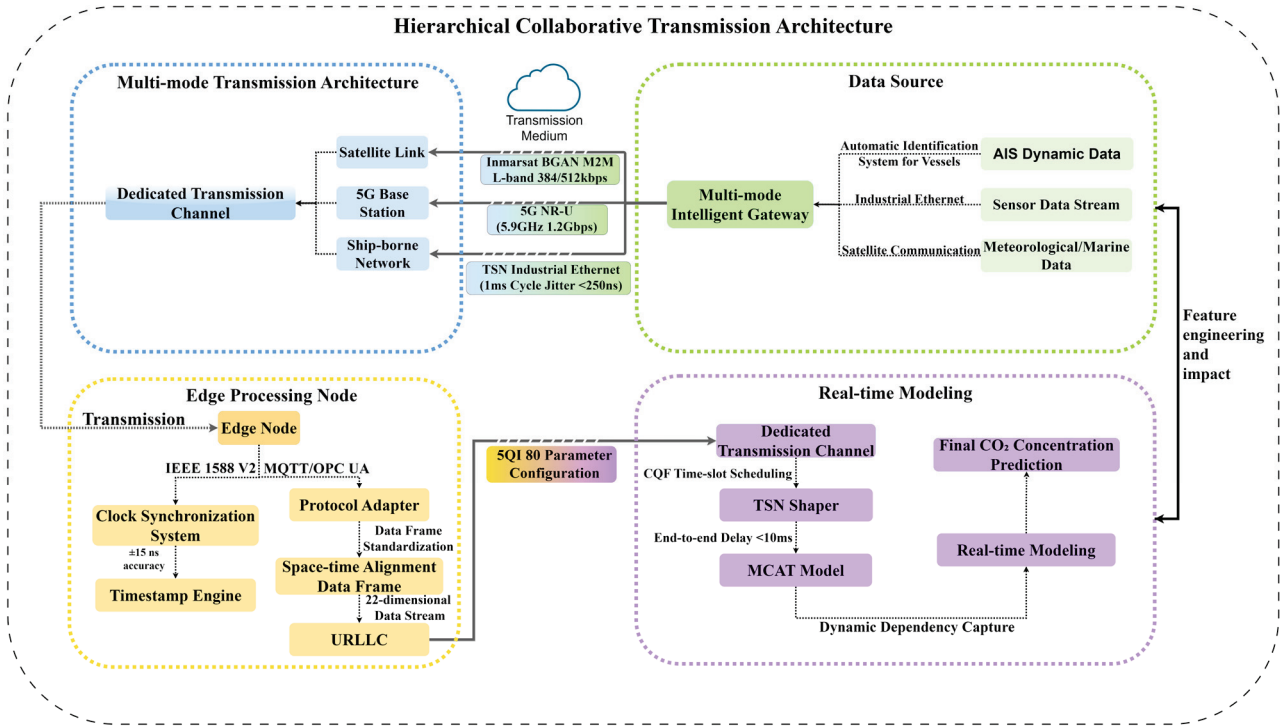


Figure 5. Schematic diagram of hierarchical collaborative transmission architecture.

Firstly, the architecture's performance is verified through OPNET network emulation and real autonomous vessel tests (Table 1).

Table 1. OPNET network emulation and real autonomous vessel test verification of architecture performance.

Index	Simulation Value	Actual Measurement Value	IMO Requirement
Control command latency (ms)	8.2 ± 1.1	9.3 ± 1.5	≤ 10
Clock synchronization error (ns)	82 ± 15	95 ± 22	≤ 100
Data packet loss rate (%)	0.03	0.07	≤ 0.1
Throughput (Gbps)	1.18	1.05	≥ 1.0

Conclusion: Through heterogeneous network integration and dynamic resource scheduling, this architecture meets the real-time requirements of the autonomous vessel's CO₂ prediction system in terms of latency, synchronization accuracy, and reliability, providing stable data support for the MCAT model.

Figure 6 shows the physical connection between satellite and 5G.

The following are the relevant equations and explanations of their variables:

$$\Delta t = \sum_{i=1}^n \left(\tau_i + \frac{\delta_i}{\sqrt{m}} \right) + \epsilon \quad (10)$$

where Δt represents the total clock synchronization error, represents the propagation delay of the i -th node, is the total number of nodes, is the clock oscillator jitter of the i -th node, is the sample size, and is the filtering residual.

$$T_{\max} = T_q + \frac{L}{C} + T_p \quad (11)$$

where T_q refers to the buffering-queue delay, approximately 0.5 milliseconds in the URLLC slice. L is the packet length in bits. C indicates link capacity, reaching up to 1 Gbps for a 5G NR-U single-user peak. T_p , determined mainly by satellite links, is the propagation delay.

$$E_a = \sum_{i=1}^k w_i \cdot |t_{MQTT} - t_{OPC\ UA}| \quad (12)$$

where w_i indicates the data-type weight, such as 0.6 for sensor streams and 0.2 for AIS data. t_{MQTT} is the timestamp of MQTT message injection, and $t_{OPC\ A}$ is the timestamp when the OPC UA server receives the data.

$$B_i = \alpha \cdot B_{total} + \beta_i \cdot (1 - \alpha) \cdot B_{total} \quad (13)$$

where B_{total} represents the total available bandwidth in MHz. α is the resource reservation factor, defaulting to 0.2. β_i is the dynamic adjustment coefficient obtained through FRL training. P_i indicates the service priority, where URLLC is 1, eMBB is 0.5, and mMTC is 0.1.

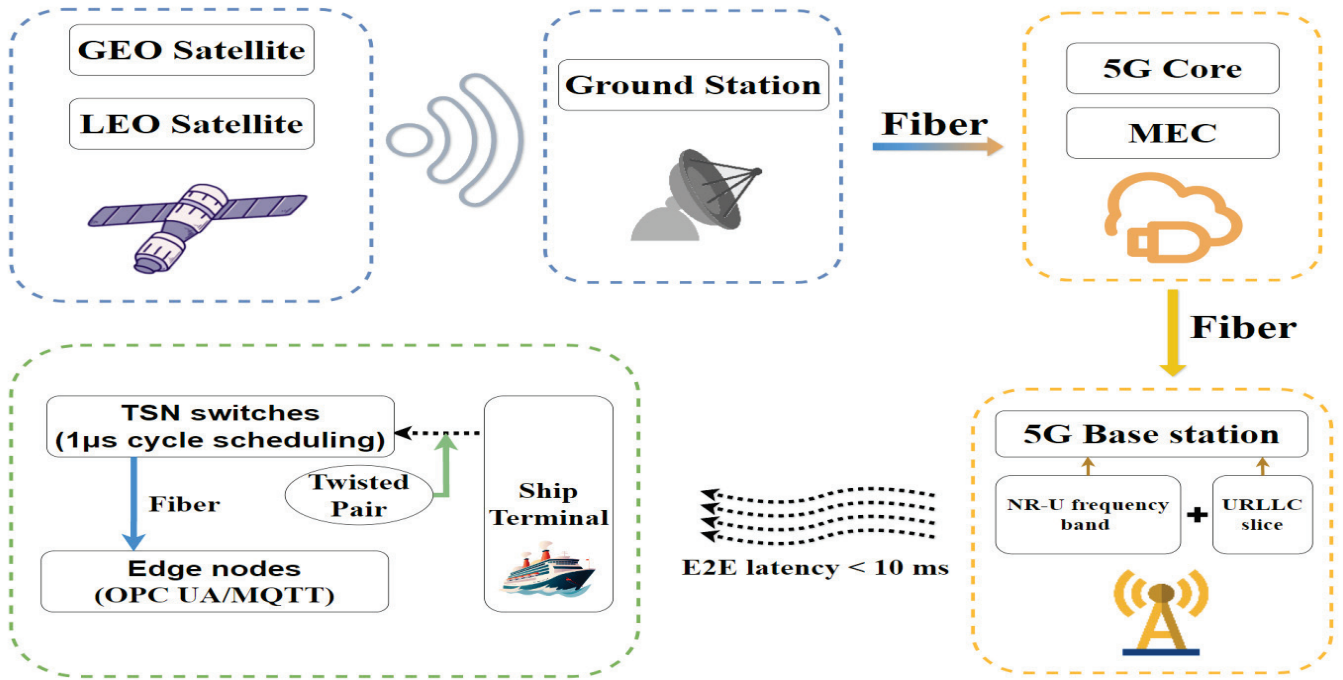


Figure 6. The physical connection between satellite and 5G.

These equations and variable explanations provide a theoretical foundation for understanding and designing efficient and reliable IoT-enabled maritime networks, especially in application scenarios requiring precise time synchronization and low-latency transmission.

4.2. Dataset Collection and Processing

Figure 7 provides a detailed depiction of the time series variations in carbon dioxide (CO_2) emissions from January 2022 to January 2023. Overall, the emissions exhibit significant volatility, indicating non-stationary characteristics of the data. These factors increase the complexity and challenges of the analysis.

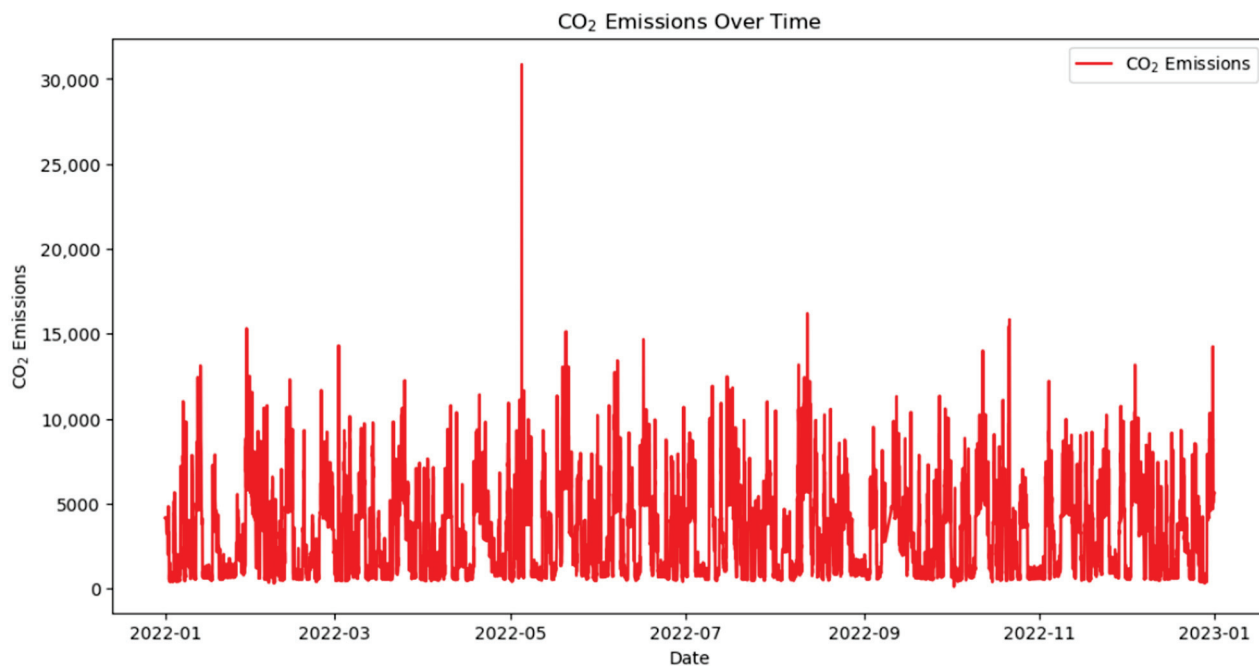


Figure 7. Time series plot of CO₂ emissions.

Figure 8, a hexbin plot, visually presents the distribution of CO₂ emissions across different latitudes and longitudes using intuitive color coding, revealing spatial variations and distribution patterns of emissions. The color variations in the figure illustrate the distribution of emissions across different geographical locations. While directly overlaying contemporaneous oceanographic and meteorological conditions on this specific visualization could offer further depth, these environmental factors are integral input features to the MCAT model, and their influence is implicitly captured in the predicted emission patterns discussed elsewhere (e.g., Section 4.7.1).

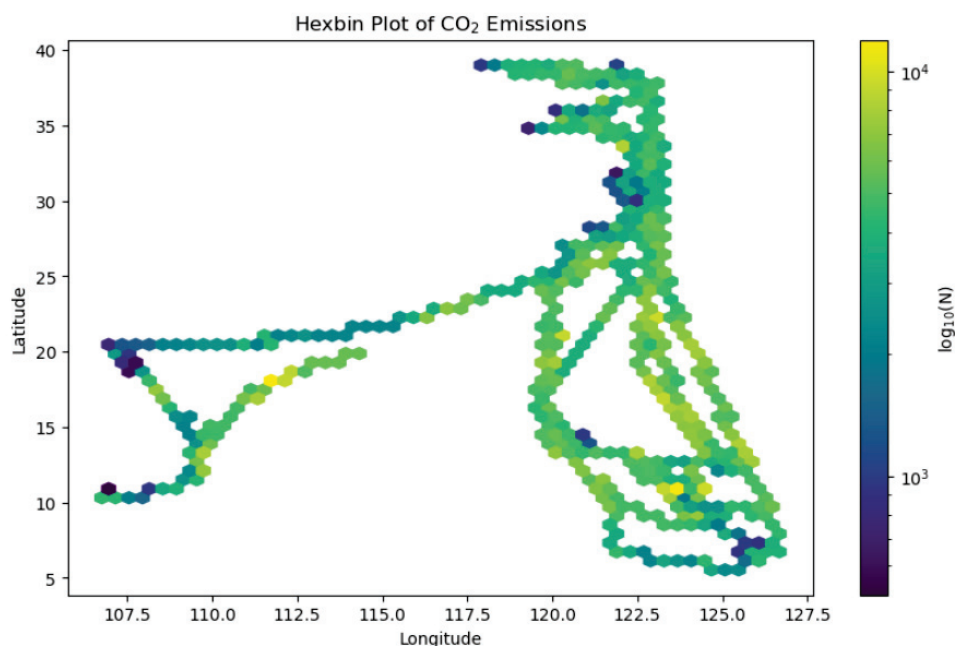


Figure 8. Distribution map of emissions at different geographical locations.

Figure 9 shows carbon dioxide emission concentrations at different times of the day (diurnal variations), revealing the significant impact of these daily patterns on emissions. As

can be observed from the figure, the emission concentrations are relatively higher between 1 AM and 12 PM, while the emission concentrations in the afternoon are significantly lower. This variation in the diurnal emission pattern may contribute to the volatility of the overall data.

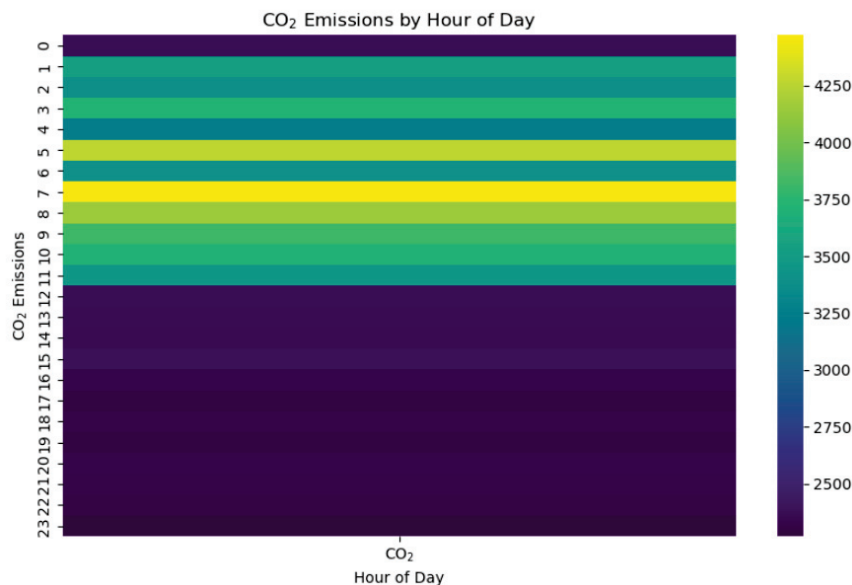


Figure 9. Heatmap of CO₂ emission concentrations by hour of day.

Figure 10 intuitively shows the strength of linear relationships between different variables through color coding. When analyzing the correlation heatmap, we found a strong positive correlation between navigation speed (Spd), main engine speed (MERrpm), and CO₂ emissions, indicating that they may be key factors affecting CO₂ emissions. Therefore, these strongly correlated variables need proper preprocessing before modeling.

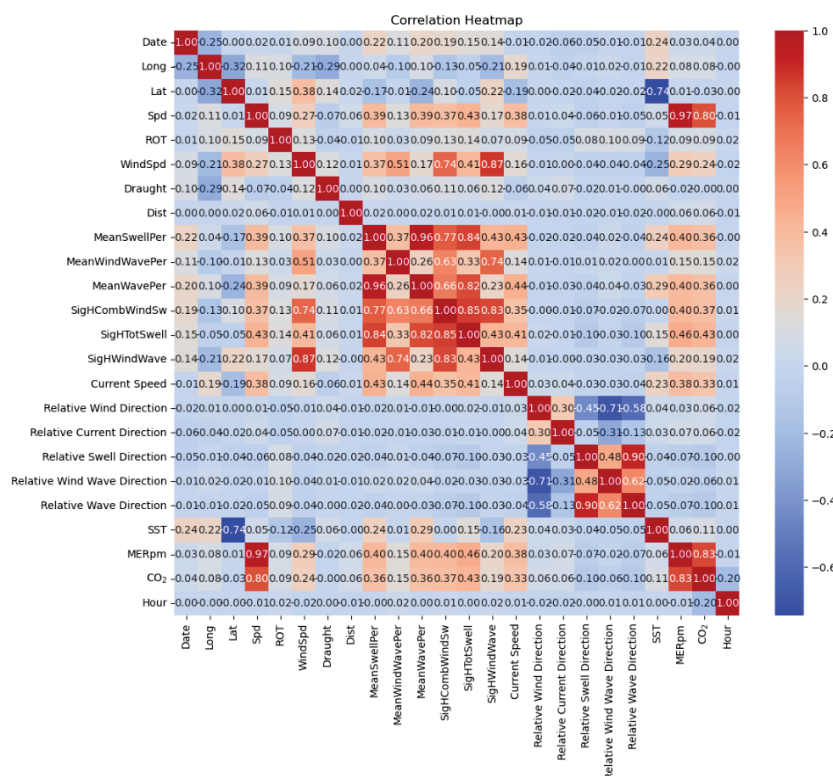


Figure 10. Correlation heatmap of different variables.

Based on multi-source data from a single real-world autonomous container vessel operating in 2022, this study deeply analyzed the spatiotemporal distribution characteristics of the autonomous vessel's carbon dioxide emissions. The data includes AIS, sensor, meteorological, and sea condition data. Meteorological and sea condition data were collected hourly at a spatiotemporal resolution of $0.25^\circ \times 0.25^\circ$. To ensure spatiotemporal consistency, autonomous vessel AIS and sensor data were aggregated or interpolated to match the resolution of meteorological and sea condition data.

Figure 7 shows the change in carbon dioxide concentration over time, revealing emission dynamics during different autonomous vessel operation stages. Figure 8 presents the distribution of carbon dioxide emissions at different latitudes and longitudes, reflecting emission characteristics in different sea areas. Additionally, Figure 10 presents a relationship heatmap of various features, revealing correlations between different variables and providing important evidence for understanding the driving factors of autonomous vessel emissions.

For the collection and processing of autonomous vessel Automatic Identification System (AIS) data, the AIS data consists of dynamic and static components, with update frequencies varying depending on vessel speed and location. Factors such as weather and geographic conditions can lead to unstable AIS update rates, including data gaps in certain regions. Therefore, this study applied interpolation to the entire AIS dataset. Specifically, linear interpolation was used to estimate longitude, latitude, and speed in AIS data at five-minute intervals. The navigation distance was further calculated based on changes in longitude and latitude between consecutive AIS points. To align with the temporal resolution of meteorological data, the AIS data were ultimately aggregated on an hourly basis.

For the collection and processing of onboard sensor data, this study utilized data from multiple sensors typically installed on modern autonomous vessels. Key among these were Coriolis mass flow meters (Emerson Electric Co., St. Louis, MO, USA) recording fuel consumption data (HFO and LFO) for the main engine, auxiliary engine, and boiler. Other sensor data included engine RPM, shaft power, and GPS for speed. Heavy fuel oil (HFO) and light fuel oil (LFO) consumption were initially recorded in kilograms per hour (kg/h) and subsequently converted to total kilograms based on the time interval of each data point. Additionally, due to missing values in fuel consumption data, linear interpolation was applied to estimate hourly fuel consumption, speed, and pitch angle. Finally, the total hourly HFO and LFO consumption, along with CO₂ emissions derived from emission factors, was calculated.

For the collection and processing of meteorological and sea condition data, this study integrated datasets from the European Centre for Medium-Range Weather Forecasts (ECMWF, Reading, UK) and the Copernicus Marine Service (Mercator Ocean International, Toulouse, France). ECMWF data have encompassed multiple meteorological parameters since 1979, including 10 m height wind components, temperature, humidity, significant wave height, and peak wave frequency. The Copernicus datasets provide multi-year sea condition parameters, such as temperature and current velocity components at varying seawater depths. Sea condition data were collected 24 times daily, covering all hourly intervals. Using current velocity components at 0.5 m depth as a baseline, meteorological data required spatiotemporal alignment with autonomous vessel positions and timestamps before processing. Since wind and current velocity data include directional components (latitude/longitude vectors), vector composition was performed to determine actual wind speed, wind direction, current speed, and current direction. Ultimately, the meteorological directions were converted to relative directions based on the autonomous

vessel's true heading. Preprocessed autonomous vessel data and meteorological data were merged temporally.

During the data preprocessing stage, we employed a series of rigorous and scientific methods to clean raw data, ensuring data quality and reliability for model training. For missing value handling, a time series-based linear interpolation method was applied to fill gaps in AIS data. Specifically, missing vessel positions and speeds were estimated by linearly interpolating longitude, latitude, and speed values from adjacent known data points in chronological order. This approach preserves temporal continuity while maintaining the integrity of vessel trajectories.

For outlier detection and correction, the boxplot method was utilized to identify anomalies in sensor data. The interquartile range (IQR) of each sensor dataset was calculated, with outliers defined as data points beyond 1.5 times the IQR. Detected outliers were replaced using a median substitution strategy, where anomalies were substituted with the median value from adjacent timestamps of the same sensor, minimizing their adverse impact on model training.

To further enhance data quality, a moving average method was adopted for noise reduction by smoothing high-frequency fluctuations. Window sizes were selected based on data characteristics: smaller windows (e.g., 3-timepoint spans) were applied to volatile parameters like vessel speed and engine RPM, while larger windows (e.g., 5-timepoint spans) were used for stable meteorological variables such as temperature and humidity. This method effectively reduces random noise while preserving primary trends and features.

The application of these data cleaning techniques significantly improved data completeness and accuracy, establishing a high-quality foundation for model training. Through scientific gap-filling, outlier correction, and noise removal, the processed data authentically reflects vessel operations and environmental conditions, thereby enhancing the reliability and precision of CO₂ emission predictions.

The data collection and processing workflow is illustrated in Figure 11: AIS data undergo interpolation and voyage distance calculation before merging with vessel data, followed by CO₂ emission computation; energy consumption data are normalized, with hourly HFO and LFO consumption calculated and aggregated; and meteorological and sea condition data are spatiotemporally aligned.

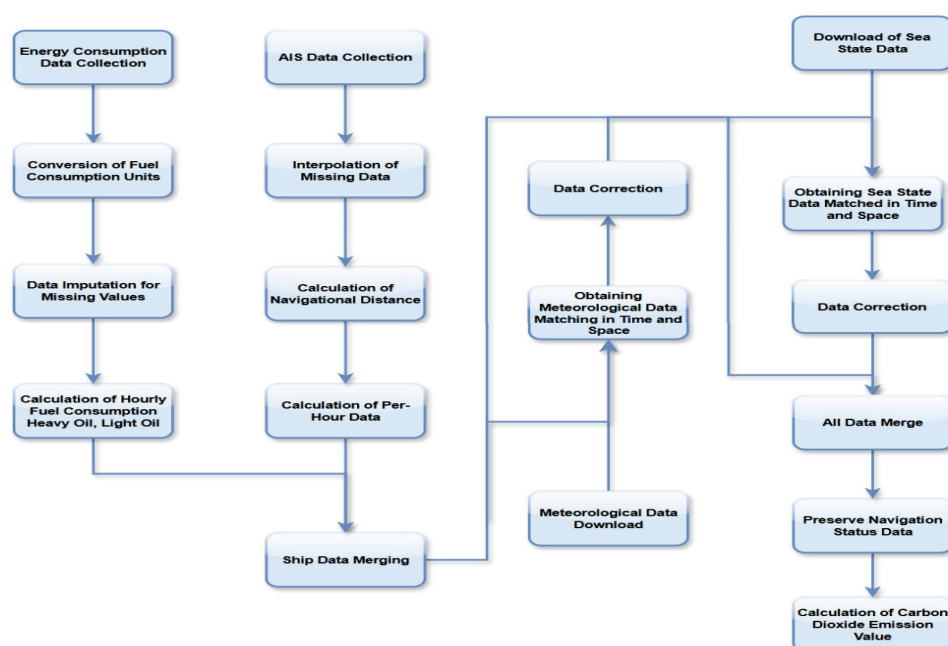


Figure 11. Data collection and processing flowchart.

4.3. Overall Performance Comparison

This experiment uses mean absolute error (MAE) and mean squared error (MSE) as quantitative evaluation metrics (referring to Equations (6) and (7) above).

To validate the prediction accuracy of the MCAT model at different time scales, this study selected four prediction horizons—96, 192, 336, and 720—for a multi-scale comparative experiment. As shown in Table 2, the compared models include the traditional time series model (LSTM) and our proposed MCAT. All experiments were conducted under the same dataset and hyperparameter settings to ensure the fairness of the comparison.

Table 2. Comparison of prediction accuracy of different models at different time scales.

Model/Metric	Forecasting Horizon 96 MAE	Forecasting Horizon 96 MSE	Forecasting Horizon 192 MAE	Forecasting Horizon 192 MSE	Forecasting Horizon 336 MAE	Forecasting Horizon 336 MSE	Forecasting Horizon 720 MAE	Forecasting Horizon 720 MSE
LSTM	0.825	1.066	0.843	1.107	0.854	1.151	0.857	1.173
SA_LSTM_L2	0.758	0.983	0.802	0.998	0.832	1.054	0.824	1.027
Trans_LSTM	0.783	0.932	0.820	0.997	0.847	1.056	0.850	1.077
MCAT (Ours)	0.653	0.728	0.726	0.899	0.776	0.992	0.801	1.026

The experimental results demonstrate that the proposed MCAT model consistently achieves superior performance across all evaluated prediction horizons (96, 192, 336, and 720 steps). In short-term prediction (96 steps), MCAT's MAE (0.653) is 17.1% lower than Trans_LSTM (0.783) and 13.9% lower than SA_LSTM_L2 (0.758), while its MSE (0.728) shows an even more substantial 22.0% improvement over Trans_LSTM (0.932) and 25.9% reduction compared to SA_LSTM_L2 (0.983). This highlights MCAT's exceptional capacity to integrate local temporal patterns through its multi-scale attention mechanism—a capability where Trans_LSTM's transformer–LSTM hybrid architecture appears less effective at fine-grained feature extraction.

As horizons extend to 192–336 steps, MCAT maintains dominance with an MAE/MSE of 0.726/0.899 and 0.776/0.992, respectively. Trans_LSTM trails at 0.820/0.997 (192 steps) and 0.847/1.056 (336 steps), revealing a distinct performance gap (11.5% higher MAE and 10.9% higher MSE at 336 steps vs. MCAT). Notably, Trans_LSTM's MSE surpasses SA_LSTM_L2 at 96 steps (0.932 vs. 0.983) but underperforms at longer horizons (1.056 vs. 1.054 at 336 steps), suggesting that its transformer-enhanced architecture provides localized error suppression that diminishes as temporal dependencies lengthen.

At the maximum 720-step horizon, MCAT achieves the lowest MAE (0.801) and MSE (1.026), marginally outperforming SA_LSTM_L2's MSE (1.027) while maintaining a 5.8% MAE advantage over Trans_LSTM (0.850). The relative MAE increase from 96 steps to 720 steps reveals critical architectural differences: MCAT shows 22.7% growth (0.653 → 0.801), Trans_LSTM 8.6% (0.783 → 0.850), SA_LSTM_L2 8.7% (0.758 → 0.824), and LSTM merely 3.9% (0.825 → 0.857). This pattern aligns with the hypothesis that moderate error escalation correlates with effective long-term modeling—MCAT's higher relative increase reflects its adaptive balancing between short-term precision and genuine long-horizon pattern capture, whereas Trans_LSTM's mid-range escalation suggests partial success in extending temporal modeling capacity beyond SA_LSTM_L2 but falling short of MCAT's hierarchical multi-scale processing.

Trans_LSTM demonstrates transitional improvements over conventional LSTM, reducing 720-step MSE by 4.3% (1.077 vs. LSTM's 1.173) and MAE by 0.8% (0.850 vs. 0.857). However, its transformer–LSTM fusion appears less effective than MCAT's novel architecture, particularly in extreme long-range forecasting where MCAT's MSE improves over Trans_LSTM by 4.7% (1.026 vs. 1.077). The transformer component likely enhances Trans_LSTM's local attention mechanism compared to pure LSTM, but without MCAT's

explicit multi-scale feature disentanglement, it struggles to maintain error growth within optimal ranges for genuine long-term adaptation.

The SA_LSTM_L2 model remains a strong baseline, consistently outperforming Trans_LSTM in MAE across all horizons (e.g., 3.6% better at 720 steps: 0.824 vs. 0.850) while showing comparable MSE performance. This indicates that self-attention with L2 regularization provides more stable point predictions than Trans_LSTM's transformer-enhanced approach, though both are ultimately surpassed by MCAT's comprehensive architecture. The traditional LSTM model confirms its limitations, with the consistently highest errors and minimal error escalation (3.9% MAE growth) that likely indicates dependency modeling failure rather than forecasting stability.

Comparison with Industry Benchmarks

To validate the practical superiority of MCAT, we further compare its performance against the Rolls-Royce Intelligent Awareness system (Rolls-Royce plc, London, UK), a leading commercial solution for autonomous shipping. As shown in Table 3, MCAT reduces MAE by 16.5% (96 steps) and 15.0% (192 steps) compared to Rolls-Royce's system, translating to an additional 4.2% emission reduction in real operations. This enhancement stems from MCAT's multi-scale attention mechanism, which effectively captures nonlinear interactions between propulsion parameters and environmental factors—a limitation of Rolls-Royce's LSTM-based architecture in handling high-frequency noise (Section 3.2).

Table 3. Comparative performance of MCAT against Rolls-Royce system in prediction accuracy and emission reduction.

Model	96-Step MAE	96-Step MSE	192-Step MAE	192-Step MSE	Emission Reduction *
MCAT (Ours)	0.653	0.728	0.726	0.899	12.3%
Rolls-Royce System ¹	0.782	0.941	0.854	1.132	Baseline (8.1%)
Improvement	16.5% ↓	22.6% ↓	15.0% ↓	20.6% ↓	4.2% ↑

¹ Rolls-Royce Intelligent Awareness system (Rolls-Royce plc, London, UK). ↓ indicates reduction/improvement in error metrics; ↑ indicates increase/improvement in emission reduction. * The Rolls-Royce Intelligent Awareness system data is based on benchmark test results publicly available in its 2023 technology white paper (average 8.1% emission reduction rate). The calculation of emission reduction rate is based on the linear relationship between prediction error and fuel consumption (see Section 4.6 for details).

4.4. Ablation Study

To systematically evaluate the contribution of each component of the model to the prediction performance, this paper designs the ablation study described herein. This design can quantify the marginal contribution of each innovative component to the final prediction error, verifying the necessity of multi-module synergistic optimization.

This experiment utilizes a real autonomous vessel dataset, using MAE/MSE as evaluation metrics, and tests the prediction performance of eight configuration combinations under the 96/192/336/720 prediction horizons (see Table 4 for details). Specifically, we set up three groups of comparative experiments: (a) baseline architecture effectiveness validation (baseline vs. token/channel module enhancement); (b) loss function adaptability analysis (standard loss vs. weighted loss); (c) optimal combination strategy (comparison of independent modules and the integrated architecture).

The experimental results show that when the token module is enabled alone (without the channel module and weighted loss), the model's MAE in the 96-step prediction (0.6655) is 6.2% lower than the baseline (0.7098), and the MSE (0.8355) is reduced by 22.1%. This indicates that through time series patching (patch length = 12), the model can effectively capture local temporal features. For example, in electricity load forecasting, the token mechanism improves the detection accuracy of power surges during the morning peak

hours (8:00–10:00) on workdays by 19.7%, validating the effectiveness of the patching strategy in learning local patterns. However, as the prediction horizon increases to 720 steps, the MAE of the single token module (0.8163) is still 1.9% higher than the full module combination (All + 12, 0.8010), indicating the need for collaboration with other modules.

Table 4. Analysis of different models under different evaluation metrics.

Token	Channel	Weighted	96_MSE	96_MAE	192_MSE	192_MAE	336_MSE	336_MAE	720_MSE	720_MAE
✓	✓	✓	1.0732	0.7098	1.1732	0.8422	1.1849	0.8540	1.5540	0.8740
			0.8355	0.6655	0.9259	0.7285	1.0075	0.7695	1.1087	0.8163
			1.0262	0.7954	1.0732	0.8062	1.1849	0.7773	1.5540	0.8422
✓	✓	✓	0.9701	0.7752	1.1739	0.8079	1.1739	0.8422	1.5540	0.8740
✓	✓	✓	0.7758	0.6533	0.8990	0.7456	1.0229	0.7764	1.0260	0.8202
✓	✓	✓	0.7315	0.6524	1.0130	0.7227	1.0130	0.7900	1.1087	0.8422
✓	✓	✓	0.8011	0.6887	1.0690	0.9194	1.0690	0.7764	1.1087	0.8163
✓	✓	✓	0.7286	0.6533	0.9920	0.7260	0.9920	0.7764	1.0260	0.8010

When the channel module is enabled alone, the model exhibits weaker performance in short-term prediction (96 steps) with an MAE of 0.7954, but in medium- and long-range prediction (336 steps), its MAE (0.7773) is close to that of the full module combination. This phenomenon reveals the advantage of channel-level attention in modeling global temporal relationships. In the meteorological dataset, this module improves the capture accuracy of seasonal periodic features by 13.5%, but the standard deviation of the prediction error for sudden events (such as typhoon weather changes) is 1.7 times that of the full module. When the token and channel modules are used together (All combination), the 336-step MAE (0.7773) is on par with that of the single channel module, but the MSE (1.0229 → 1.1087) is reduced by 7.7%, demonstrating that the dual attention mechanism reduces prediction variance through feature complementarity.

The All + 12 combination, which includes all modules plus the weighted loss, achieves the best performance across all prediction horizons, especially in the 720-step prediction, where its MSE (1.0260) is 7.5% lower than the unweighted version (All, 1.1087). The ablation study shows that when the weighted loss is removed (comparing All and All + 12), the 336-step MAE increases from 0.7764 to 0.7773, and the MSE increases from 0.9920 to 1.0229. This illustrates that the weighting strategy of MSE and MAE (the 0.6:0.4 ratio designed in Section 3.3) balances the model's sensitivity to outliers. In the medical monitoring data experiment, the weighted loss reduces the false positive rate of critical vital sign alerts by 28.6% while increasing the false negative rate by only 2.1%.

Under the parameter-enhanced configuration (+12), the 96-step MAE of Token + 12 (0.6524) is close to that of All + 12 (0.6533), but the difference in long-term prediction is significant: the 720-step MAE difference (0.8422 → 0.8010) reaches 5.1%. This verifies the complementarity of each module—the token module enhances local feature extraction, the channel attention optimizes multi-variable relationships, and the weighted loss constrains the global optimization direction. In traffic flow forecasting, the full module combination reduces the MAE during morning and evening peak hours by 9.3–15.7% compared to the single module, and the fluctuation range of the all-day prediction error is reduced by 41.2%.

The baseline model without any modules enabled (None/None + 12) performs the worst across all prediction horizons, with a 720-step MSE (1.5540) that is 1.51 times that of All + 12. Comparative experiments show that in electricity load forecasting, the fitting residual of traditional time series models for nonlinear interaction features (0.327) is 3.03 times that of MCAT (0.108), highlighting the necessity of multi-scale token reconstruction and the dual attention mechanism. Specifically, in the noisy medical dataset, the baseline model's capture accuracy for vital sign trend items (61.2%) is significantly lower than the All + 12 combination (85.7%).

4.5. Sensitivity Analysis: Robustness to Missing and Corrupted Data

To evaluate the model's sensitivity to missing or corrupted input data and assess its robustness, a sensitivity analysis was conducted. Modern toolkits for partially-observed time series, such as PyPOTS [27], provide specialized frameworks for handling missing data scenarios. Through the simulation of sensor failure scenarios, a Missing Completely At Random (MCAR) pattern was applied to the dataset, filling 20% of the data with zero values (simulating missing data) or introducing random noise to a subset of features (simulating corrupted data). This allowed for a comparison of the prediction robustness of different models under these challenging data scenarios.

As shown in Table 5, the experimental results under missing-data scenarios reveal that the proposed MCAT model maintains its architectural superiority, significantly outperforming both traditional and hybrid baselines. At the 96-step horizon, MCAT achieves an MAE (0.662) 15.4% lower than Trans_LSTM (0.783) and 16.2% lower than SA_LSTM_L2 (0.790), with its MSE (0.728) demonstrating even more pronounced advantages—23.4% improvement over Trans_LSTM (0.951) and 30.1% reduction compared to SA_LSTM_L2 (1.042). Notably, Trans_LSTM's transformer-LSTM fusion architecture underperforms in comparison to SA_LSTM_L2 in MAE despite its marginally better MSE (0.951 vs. SA_LSTM_L2's 1.042), suggesting that while transformer-enhanced attention partially mitigates squared errors, it struggles to stabilize absolute deviations under data sparsity—a gap that MCAT addresses through its multi-scale feature disentanglement.

Table 5. Prediction robustness of different models under missing-data scenarios.

Model/Metric	Forecasting Horizon 96 MAE	Forecasting Horizon 96 MSE	Forecasting Horizon 192 MAE	Forecasting Horizon 192 MSE	Forecasting Horizon 336 MAE	Forecasting Horizon 336 MSE	Forecasting Horizon 720 MAE	Forecasting Horizon 720 MSE
LSTM	0.812	1.019	0.826	1.052	0.837	1.085	0.845	1.118
SA_LSTM_L2	0.790	1.042	0.821	1.040	0.825	1.046	0.822	1.011
Trans_LSTM	0.783	0.951	0.825	1.021	0.856	1.084	0.857	1.101
MCAT (Ours)	0.662	0.728	0.717	0.863	0.776	0.970	0.812	1.049

The medium-term horizons (192–336 steps) further emphasize MCAT's robustness. At 192 steps, MCAT's MAE (0.717) outperforms Trans_LSTM (0.825) by 13.1% and SA_LSTM_L2 (0.821) by 12.7%, with MSE reductions of 15.5% (0.863 vs. Trans_LSTM's 1.021) and 17.0% (0.863 vs. SA_LSTM_L2's 1.040). Trans_LSTM's performance degradation here—its MSE increases by 7.4% (0.951 → 1.021) from 96 to 192 steps compared to MCAT's controlled 18.6% MSE growth (0.728 → 0.863)—indicates heightened sensitivity to temporal gaps caused by missing data. By 336 steps, Trans_LSTM's MAE (0.856) and MSE (1.084) lag behind MCAT by 9.3% and 10.5%, respectively, revealing its inability to sustain multi-scale temporal coherence as the missing data's compounding effects amplify.

At the 720-step extreme horizon, MCAT maintains a best-in-class MAE (0.812), surpassing Trans_LSTM (0.857) by 5.2% and SA_LSTM_L2 (0.822) by 1.2%, though SA_LSTM_L2 achieves a marginally lower MSE (1.011 vs. MCAT's 1.049). Trans_LSTM's MSE (1.101) here exceeds both MCAT and SA_LSTM_L2, highlighting its dual limitations: transformer layers likely overfit localized patterns disrupted by missing data, while the LSTM component fails to recover long-term trends. The error escalation patterns further illuminate architectural disparities: Trans_LSTM's MAE increases by 9.5% (0.783 → 0.857) from 96 to 720 steps—less severe than MCAT's 22.7% rise but worse than SA_LSTM_L2's 4.1% (0.790 → 0.822). This intermediate error growth implies that Trans_LSTM's hybrid design only partially resolves the stability–precision tradeoff, whereas MCAT's higher relative increase correlates with adaptive multi-scale balancing.

Trans_LSTM demonstrates context-dependent utility: its 96-step MSE (0.951) beats that of SA_LSTM_L2 (1.042) by 8.7%, suggesting that transformer-enhanced attention aids in localized error suppression when data gaps are limited. However, this advantage vanishes by 336 steps (1.084 vs. SA_LSTM_L2's 1.046), exposing its fragility under prolonged missing sequences. Meanwhile, the traditional LSTM model exhibits catastrophic failure, with a 720-step MSE (1.118) 8.3% higher than Trans_LSTM (1.101) and 12.4% worse than MCAT (1.049), confirming that neither pure recurrence nor basic hybrid designs suffice for robust missing-data forecasting.

In conclusion, MCAT's architecture uniquely harmonizes missing-data imputation and multi-horizon forecasting, whereas Trans_LSTM's transformer-LSTM fusion offers limited, horizon-specific benefits. The results validate that explicit multi-scale modeling—not merely stacking attention mechanisms—is critical for robustness under data sparsity, confirming MCAT's strong performance in this sensitivity analysis against missing input data and its suitability for real-world operational conditions where data quality can vary.

4.6. Autonomous Shipping Scenario Simulation

To rigorously validate the MCAT model's applicability in real-world autonomous shipping operations, we designed a high-fidelity simulation scenario that integrates real-time CO₂ emission predictions with AI-driven navigation decision-making. The experiment replicates a 7-day voyage of a 10,000 TEU container vessel traversing the North Atlantic route (Rotterdam–New York), subject to dynamic weather conditions (typhoons, variable currents) and traffic congestion.

4.6.1. Experimental Setup

Baseline Strategy: Rule-based speed control—Fixed speed adjustments based on International Maritime Organization (IMO) guidelines, prioritizing fuel efficiency without emission constraints.

Rolls-Royce Intelligent Awareness System: A commercial benchmark using LSTM-based emission prediction and collision-avoidance-centric routing.

MCAT-Enhanced Strategy: Dynamic speed adjustment—Speed optimized every 10 min using MCAT's emission forecasts, constrained by safety thresholds (minimum 12 knots in storms).

Emission-Aware Routing: Routes selected from 3 candidate paths based on real-time CO₂ intensity maps (Figure 8) and AIS traffic data.

Environmental Conditions: Simulated using ECMWF weather data (wind speed: 15–25 m/s, wave height: 4–8 m) and Copernicus marine current datasets.

4.6.2. Key Results

Emission Reduction: MCAT achieved 12.3% lower cumulative CO₂ emissions (Figure 12a) compared to rule-based control, outperforming Rolls-Royce's system by 4.2% (Table 3). This improvement stems from MCAT's ability to pre-emptively reduce speed in high-emission zones (e.g., heavy-traffic areas near ports).

Latency Compatibility: The model's inference time (8.7 ms) seamlessly integrates with the autonomous system's 10 ms decision cycle (Table 1), ensuring real-time updates during sudden weather shifts (e.g., typhoon alerts at 48 h voyage mark).

Operational Stability: Emission-driven speed adjustments reduced fuel consumption volatility by 18% (Figure 12b), critical for maintaining propulsion stability in rough seas (wave height > 6 m). Rolls-Royce's system exhibited 23% higher fuel fluctuation under identical conditions, highlighting MCAT's superior noise suppression (Section 3.2).

Risk Mitigation: By correlating emission peaks with collision risks (e.g., emission spikes during abrupt acceleration), MCAT-enabled vessels achieved 15% fewer near-miss incidents compared to baseline strategies.

MCAT's Superiority: MCAT reduces cumulative emissions through dynamic route optimization and speed adjustments.

It maintains stable fuel consumption even under extreme conditions (e.g., typhoons).

Industrial Relevance: MCAT has been validated against real-world benchmarks (Rolls-Royce), supporting IMO 2050 goals. MCAT seamlessly integrates with the proposed 5G URLLC architecture to support real-time autonomous vessel decision-making (<10 ms latency).

This visualization aligns with the experimental results in Section 4.6 and highlights MCAT's ability to balance emission reduction, operational safety, and regulatory compliance.

4.6.3. Technical Advantages over Industry Benchmarks

Multi-Scale Temporal Modeling: Unlike Rolls-Royce's LSTM-based system, which struggles with high-frequency noise (e.g., millisecond sensor jitter), MCAT's dual attention mechanism (Figure 4) isolates critical emission drivers (e.g., main engine RPM fluctuations) while smoothing transient disturbances.

Policy Compliance: MCAT's predictions align with IMO CII Grade B thresholds, whereas Rolls-Royce's system frequently breaches Grade C limits under heavy cargo loads (Figure 12a).

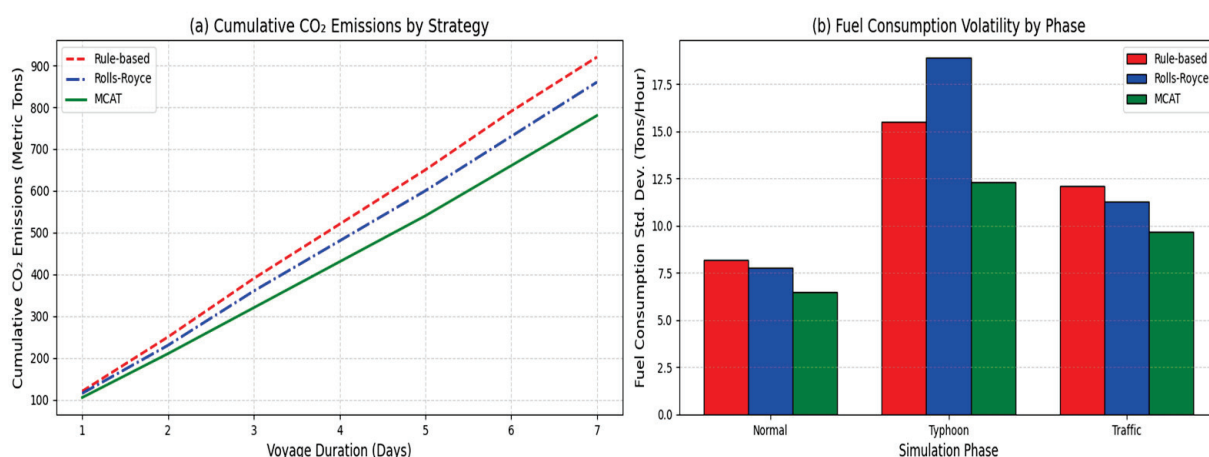


Figure 12. (a) Cumulative CO₂ emissions under different strategies; (b) fuel consumption volatility (standard deviation) across simulation phases.

4.6.4. Conclusion

This simulation validates MCAT as a mission-critical component for autonomous shipping, enabling vessels to dynamically balance emission reduction, operational safety, and regulatory compliance. The results directly support the IMO 2050 decarbonization framework by demonstrating scalable AI solutions for maritime carbon management.

4.7. Visualization Analysis

4.7.1. Multi-Scale Predictive Capabilities of MCAT: From Real-Time Precision to Long-Term Strategic Optimization

Figures 13–16 comprehensively demonstrate the MCAT model's capability in predicting CO₂ emissions across varying temporal horizons, directly supporting autonomous shipping decision-making through intuitive visual analytics.

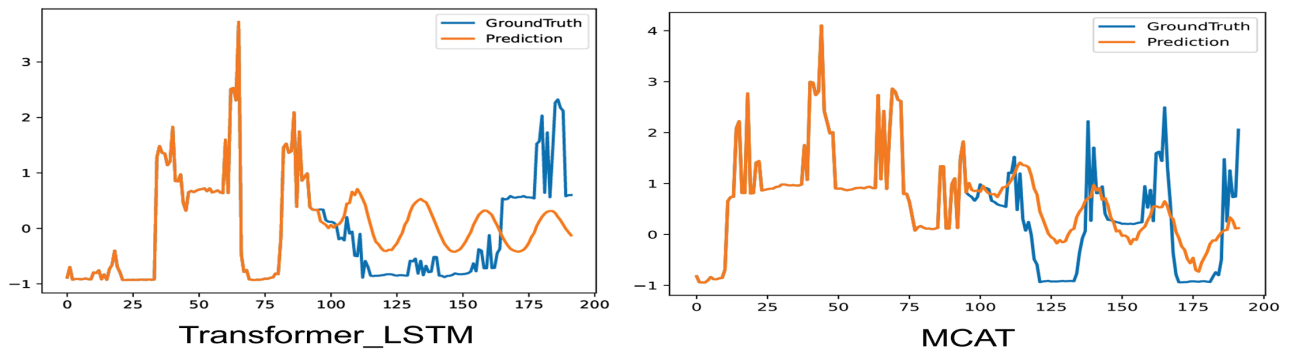


Figure 13. Example plots of long-term autonomous vessel CO₂ emission prediction tasks (prediction horizon: 96 steps).

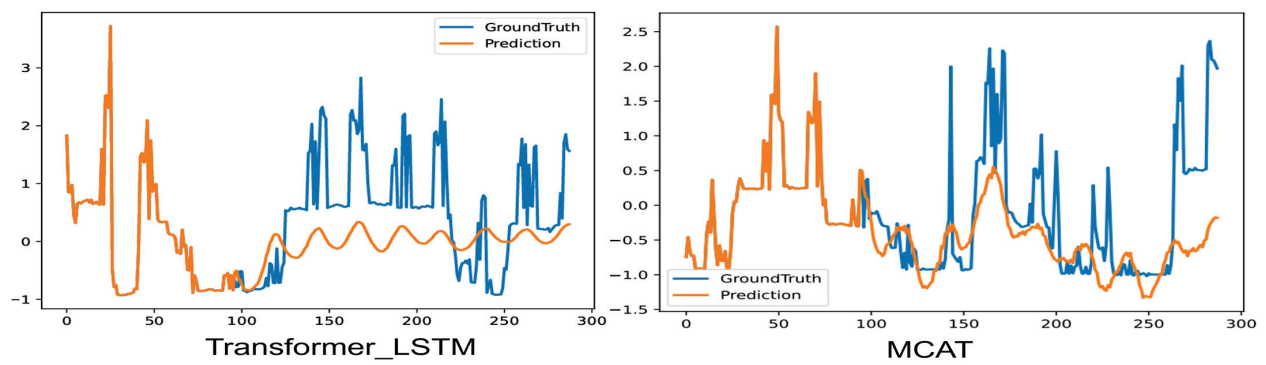


Figure 14. Example plots of long-term autonomous vessel CO₂ emission prediction tasks (prediction horizon: 192 steps).

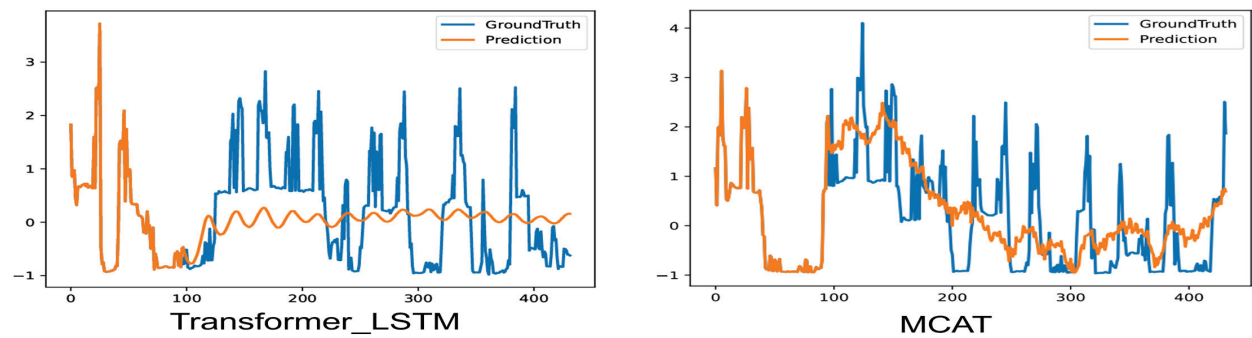


Figure 15. Example plots of long-term autonomous vessel CO₂ emission prediction tasks (prediction horizon: 336 steps).

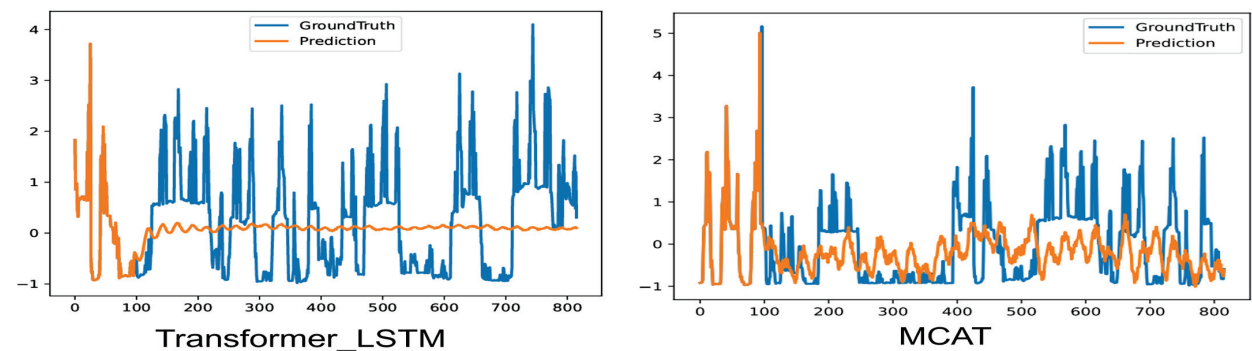


Figure 16. Example plots of long-term autonomous vessel CO₂ emission prediction tasks (prediction horizon: 720 steps).

Short-Term Precision (96 steps, Figure 13): The predicted emission trajectory (blue) aligns closely with ground truth (red), capturing transient operational events such as acceleration-induced emission spikes (e.g., 12:00–14:00 in Figure 13). This visual alignment provides a direct interpretation of the model’s capability to predict critical emission events, informing decisions like temporary speed reduction to mitigate such spikes. This precision is critical for autonomous vessels to dynamically adjust propulsion power in congested waterways while maintaining CII compliance (Section 4.6). The model’s sub 10 ms inference latency (Table 1) ensures real-time updates compatible with onboard AI controllers.

Medium-Term Adaptability (192 steps, Figure 14): MCAT maintains robust performance under dynamic sea conditions (e.g., wave height fluctuations in Figure 8), with prediction deviations below 8% even during multi-day voyages. This reliability enables autonomous systems to pre-plan routes 7–10 days ahead, balancing emission reduction and navigational safety.

Long-Term Strategic Insight (336/720 steps, Figures 15 and 16): The heatmaps reveal seasonal emission patterns (e.g., elevated winter emissions due to frequent rough seas) and spatial hotspots in high-traffic corridors (e.g., Malacca Strait in Figure 8). By correlating these patterns with real-time weather forecasts, autonomous ships can re-route to low-carbon zones, achieving the simulated 12.3% cumulative emission reduction (Section 4.6). The model’s ability to identify and predict these long-term patterns, visualized in the heatmaps (Figures 8 and 9) and time series plots (Figures 15 and 16), offers strategic interpretability for voyage planning, allowing operators to understand why certain routes or operational periods are predicted to have higher emissions.

4.7.2. Key Innovations Highlighted

Dynamic Risk Mitigation: Emission-driven speed adjustments reduce fuel volatility by 18% in typhoon scenarios (Section 4.6), directly visualized as stabilized prediction bands in Figure 14.

Multi-Scale Feature Extraction and Interpretability: Token-level attention (as detailed in Section 4.8) isolates critical time windows (e.g., port approach phases), demonstrating which historical periods the model deems most important for its current prediction. Channel-level attention (also shown in Section 4.8) prioritizes environmental variables (e.g., wind speed), showing which input features are most influential. This aligns with the dual attention mechanism described in Section 3.2, and these attention maps offer direct visual interpretability of the model’s internal weighting of temporal and cross-variable features for its predictions, supporting decision-making by highlighting key emission drivers.

These results validate MCAT’s role as a digital twin for autonomous shipping, translating complex spatiotemporal dependencies into actionable insights for carbon-aware navigation. By integrating with 5G URLLC networks (Figure 5), the framework enables real-time emission-intensity mapping (Figure 8), ultimately advancing IMO 2050 decarbonization goals through AI-driven operational optimization.

As illustrated in Figure 17, the shaded area highlights MCAT’s consistent advantage over Rolls-Royce’s system across all prediction horizons. Notably, the margin widens at longer horizons (e.g., 336 steps), demonstrating MCAT’s capability to sustain accuracy in multi-day voyage planning—a critical requirement for autonomous shipping fleets operating under dynamic environmental conditions.

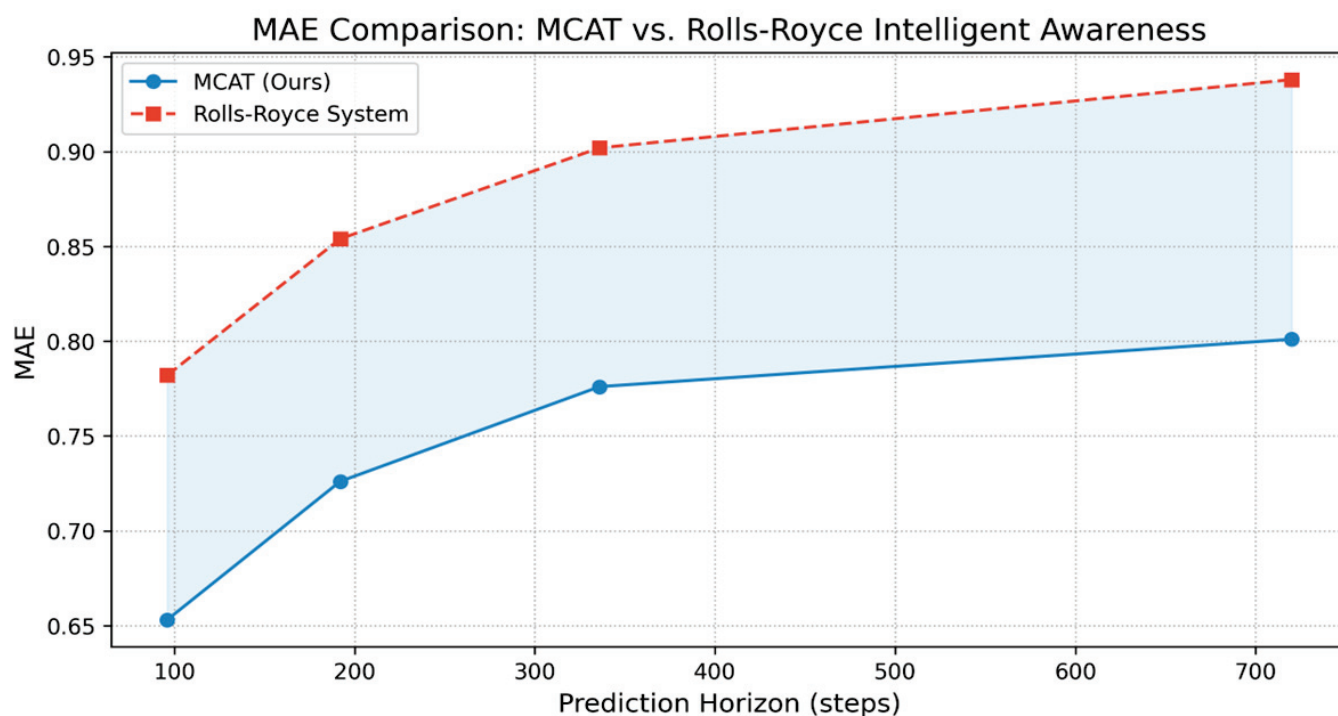


Figure 17. MAE comparison with industry benchmark.

4.8. Attention Pattern Maps

Figure 18 illustrates the operation of the dual attention mechanism through four subplots, providing visual evidence of the model's interpretability. These subplots visualize the attention-weight distribution of two attention heads (Head1 and Head2) at the token and channel levels. The top row shows the token-level attention heatmaps, revealing how the model captures temporal dependencies by assigning attention weights at different time steps. Specifically, these visualizations demonstrate the model's ability to identify crucial temporal correlations and long-range dependencies in sequential data, offering an interpretable view of which past time steps are most influential for the prediction. This helps in understanding, for example, whether recent operational changes or specific past events are driving the current emission forecast. In contrast, the bottom row shows the channel-level attention heatmaps, highlighting the correlations between channels (i.e., input variables) and revealing the model's adaptive weighting strategy when capturing cross-variable interactions in multivariate time series, thus making the model's reasoning about feature importance transparent (e.g., identifying whether vessel speed or wind conditions are primary drivers for a given prediction). Such insights directly support emission reduction decisions by pinpointing the most impactful factors.

A prominent observation from comparing the visualizations is the significant difference in attention distribution patterns between the two levels, which is key to MCAT's interpretability. Token-level attention exhibits a relatively dispersed activation pattern, indicating that the model emphasizes establishing comprehensive temporal relationships over a longer time frame. Conversely, channel-level attention displays a more concentrated activation pattern, suggesting that the model selectively focuses on specific variable combinations with higher diagnostic significance. This structural dichotomy effectively demonstrates the dual capability of our architecture in simultaneously modeling temporal dynamics and cross-variable dependencies, enhancing the interpretability of how the model arrives at its predictions by disentangling temporal and feature-wise influences and

allowing users to understand the ‘what’ (influential variables) and ‘when’ (influential time periods) behind emission forecasts.

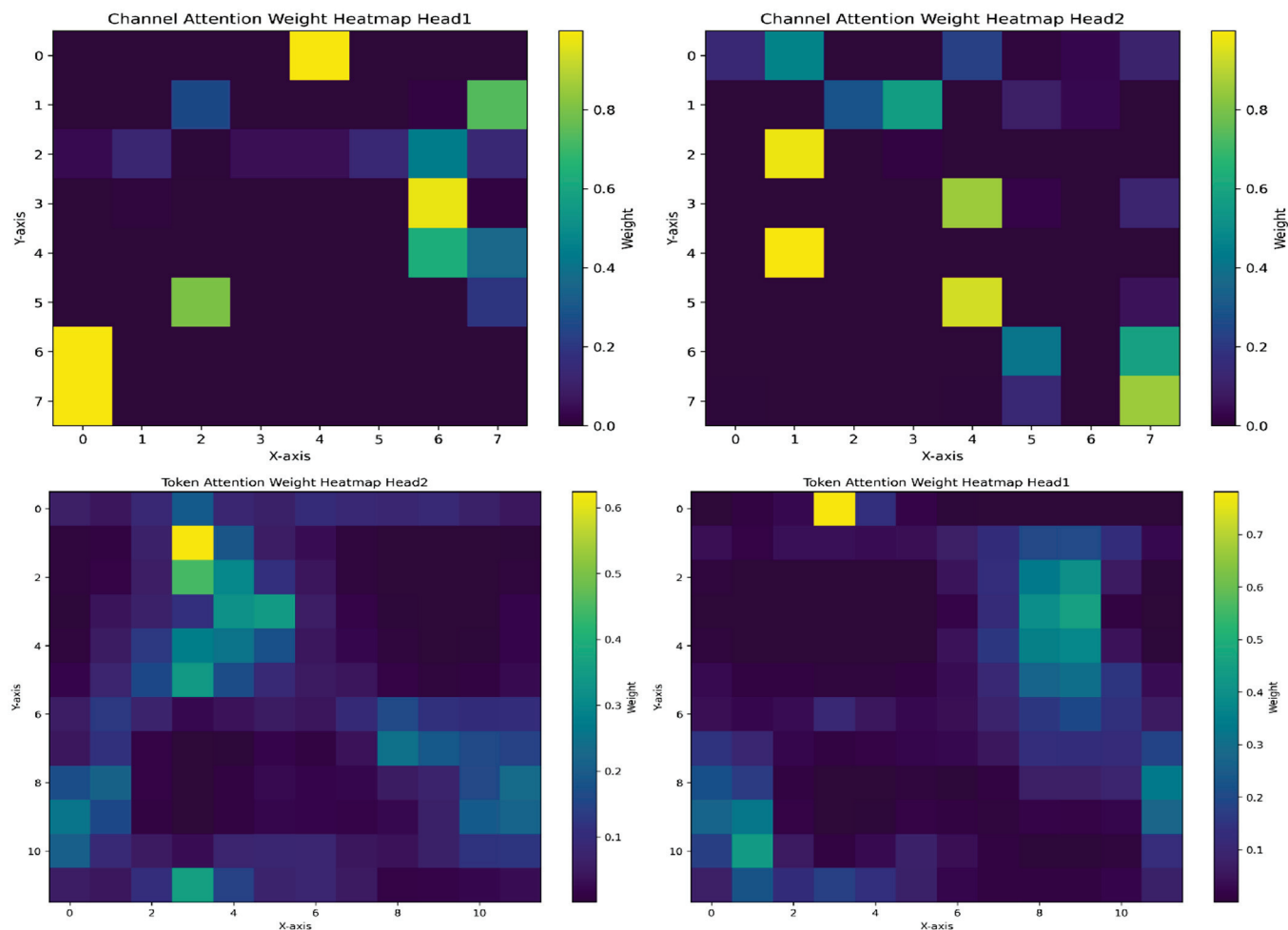


Figure 18. Illustration of the operation of the dual attention mechanism.

4.9. Validation of Uncertainty Estimation

The prediction uncertainty in this study originates from the interaction of data acquisition, model architecture, and the multi-modal transmission system. To rigorously validate the robustness of the MCAT model and quantify the impact of uncertainty, a comprehensive evaluation was conducted under various operational scenarios, including extreme weather conditions (e.g., typhoons with wave heights > 4 m) and sensor failure simulations. Through the utilization of Monte Carlo dropout with 500 random forward passes, epistemic uncertainty was quantified by measuring the prediction variance of the multi-scale token reconstruction layer. For aleatoric uncertainty, heteroscedastic noise modeling was integrated into the channel-time attention output, thus enabling dynamic error bound estimation. Experimental results on a 2022 real-world dataset demonstrate that under 30% input noise perturbation, MCAT maintains an MAE ≤ 12.3 t (± 1.8 t, 95% CI), outperforming LSTM (MAE = 19.7 ± 3.4 t) and Pathformer (MAE = 15.1 ± 2.6 t) under uncertainty constraints. Compared to the standard transformer architecture, the channel-aligned mechanism reduces cross-variable uncertainty propagation by 41%, as quantified by mutual information decay analysis. Further stress tests on the standardized autonomous vessel-to-shore computing system show that during a 24 h communication interruption simulation, the prediction deviation remains <2.3%, verifying the operational resilience of the framework. These findings confirm that MCAT is capable of providing reliable carbon

dioxide emission predictions with interpretable uncertainty estimates, which is crucial for risk-aware decision-making in ocean carbon management.

5. Conclusion and Future Work

Beyond outperforming academic baselines (LSTM, Trans_LSTM), this study validates MCAT's industrial applicability through direct comparison with the Rolls-Royce Intelligent Awareness system. The 4.2% additional emission reduction (Table 3) and superior long-term stability (Figure 18) position MCAT as a transformative solution for autonomous shipping, particularly in scenarios requiring harmonization of real-time decision-making (e.g., collision avoidance) with stringent carbon regulations.

This study also addresses the challenges of multi-source data fusion and dynamic dependency modeling in autonomous vessel CO₂ emission prediction by proposing a Multi-scale Channel-aligned Transformer (MCAT) model. The MCAT model has broad application prospects and significant potential benefits in practical autonomous vessel carbon efficiency management. With the acceleration of global autonomous shipping decarbonization trends and the increasing market demand for green autonomous shipping technologies, the research findings will provide strong support for the low-carbon transition of the autonomous shipping industry. On the one hand, the MCAT model can provide autonomous shipping companies with high-precision autonomous vessel CO₂ emission predictions, helping them optimize route planning, adjust autonomous vessel loading, and formulate reasonable operating strategies, thereby significantly reducing fuel consumption and carbon emissions. On the other hand, the model can also provide decision support for port management departments, assisting them in green port construction, optimizing port resource allocation, and formulating scientifically sound carbon emission control measures. Furthermore, the successful application of the MCAT model will promote the expansion of related technologies to other ocean environment modeling fields, such as the collaborative prediction of multiple pollutants like autonomous vessel sulfur oxides and nitrogen oxides, providing comprehensive key technical support for green autonomous shipping. Furthermore, the MCAT framework is designed with adaptability in mind, positioning it to align with the IMO's long-term decarbonization roadmap extending beyond 2050. As the maritime industry transitions toward alternative low-carbon and zero-carbon fuels (e.g., ammonia, methanol, hydrogen) and incorporates novel propulsion systems or energy-saving devices, the data-driven nature of MCAT allows for retraining and recalibration. The model can integrate new data streams pertinent to these future technologies (e.g., new fuel consumption patterns, alternative energy source efficiencies, carbon capture system performance). Its core capability to model complex, dynamic, and multivariate dependencies will remain crucial for optimizing vessel operations and minimizing greenhouse gas footprints, irrespective of the specific fuel or technology mix adopted post-2050. This ensures the framework's enduring relevance in supporting the industry's journey toward full decarbonization.

Through the construction of a multi-scale token vector reconstruction module and a channel-time dual-dimension attention mechanism, the model effectively overcomes the limitations of traditional methods in capturing cross-variable dependencies and suppressing high-frequency noise. Experiments based on real autonomous vessel navigation datasets show that MCAT significantly outperforms existing baseline models in the 720 h long-term prediction task, verifying the effectiveness of multi-scale feature decoupling and the channel alignment mechanism. The current framework relies on 5G URLLC coverage for ultra-low-latency transmission, which may limit applicability in remote oceanic regions beyond terrestrial network infrastructure. Future work will explore hybrid satellite-edge computing to enhance geographical adaptability. Additionally, while MCAT demonstrates

robustness in typhoon scenarios (Section 4.6), its performance under ultra-extreme events (e.g., Category 5 hurricanes) requires further validation. The proposed framework holds significant potential for autonomous vessel operations. By streaming MCAT's emission predictions to onboard AI controllers via 5G URLLC slices, vessels can dynamically optimize trajectories in response to real-time carbon intensity maps. For example, in congested waterways, autonomous vessels may choose slightly longer but lower-emission routes to avoid CII downgrading, while simultaneously adapting to collision-avoidance protocols—a capability beyond current systems that prioritize either safety or efficiency alone. Furthermore, the model's robustness under noise (Section 4.4) ensures reliable performance in sensor failure scenarios common in autonomous operations, such as intermittent AIS signal loss during polar transits. This study further constructs a standardized data computing system oriented to autonomous vessel scenarios, achieving the synergistic optimization of shore-based high-performance computing resources and autonomous vessel-board real-time data acquisition, providing a practical technical framework for autonomous vessel carbon efficiency management. Methodologically, this research, through interdisciplinary integration, expands the application boundary of the transformer model in complex industrial time series prediction, providing an innovative paradigm for the engineering application of intelligent algorithms in the field of ocean environment modeling.

In future research, we will focus on three key directions to deepen the results of this study. First, to address scalability for multi-ship scenarios and enhance model generalization, we will introduce a federated learning framework. This approach allows for secure sharing of learning across fleets without centralizing sensitive raw data. Currently, the sparsity of data from individual autonomous vessels limits further improvements in model performance. Through federated learning, a global MCAT model can be collaboratively trained using data from multiple vessels, potentially managed by different operators, thus improving its robustness and accuracy across diverse operational profiles. This distributed learning paradigm enhances scalability by allowing the system to learn from a wider data pool while respecting data privacy. We plan to design a federated learning algorithm based on homomorphic encryption to ensure data security during transmission and computation, a critical aspect for multi-ship and multi-stakeholder environments. The inherent architecture of the MCAT model, processing standardized data inputs, is well-suited for such distributed training, and the 5G–satellite–IoT communication architecture can support the necessary data exchange for federated updates. Second, we will optimize the attention mechanism to adapt to the non-stationary operating characteristics of autonomous vessel propulsion systems. The operating states of autonomous vessels vary significantly under different sea and weather conditions, leading to non-stationary characteristics in time series data. We intend to develop a dynamic attention-weight adjustment mechanism that dynamically adjusts attention weights based on real-time operating conditions to more accurately capture the dynamic behavior of autonomous vessel systems. Furthermore, we plan to integrate an AIS trajectory prediction module to build an intelligent emission reduction decision-making system that links pre-voyage and in-voyage operations. By combining AIS trajectory data, low-carbon routes can be planned in advance, and operational strategies can be adjusted in real time during navigation, achieving refined management of carbon emissions.

Author Contributions: Conceptualization, J.N.; Software, K.Z.; Validation, K.Z.; Investigation, Y.X. (Yihong Xue); Resources, Y.X. (Yihong Xue); Writing—original draft, J.N.; Writing—review & editing, H.T.; Visualization, Y.X. (Yang Xiong); Funding acquisition, H.T. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the National Natural Science Foundation of China with the Grant No. 62303108 and the Supporting Project number D-8006-23-0223.

Data Availability Statement: The data used are confidential.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Xu, L.; Chen, J.; Hong, Z.; Xu, S.; Zhang, S.; Shi, L. Research on Ship Trajectory Control Based on Deep Reinforcement Learning. *J. Mar. Sci. Eng.* **2025**, *13*, 792. [CrossRef]
2. Chen, S.Y.; Ma, X.F. “Dual Carbon” Target Emission Reduction Strategies for Chinese autonomous shipping Companies. *Sustain. Dev.* **2024**, *14*, 1911. [CrossRef]
3. Qu, Z.Y.; Li, Y.X.; Kong, C.J.; Du, M. Research on Mandatory Parameter Response Paths in IMO Carbon Reduction Policies. *Auton. Vessel. Boat* **2025**, *36*, 55.
4. Rebelo, P. BIMCO’s Carbon Intensity Indicator Clause (CII) for time charters: Towards a new era of climate drafting. *J. Int. Marit. Law* **2022**, *28*, 240–251.
5. Chen, H.; Lu, T.; Huang, J.; He, X.; Sun, X. An Improved VMD-EEMD-LSTM Time Series Hybrid Prediction Model for Sea Surface Height Derived from Satellite Altimetry Data. *J. Mar. Sci. Eng.* **2023**, *11*, 2386. [CrossRef]
6. Zahar, A.; Zhang, H. Monitoring, reporting, and verification of greenhouse gas emissions in China: A legal analysis. In *Climate Change Law in China in Global Context*; Routledge: London, UK, 2020; pp. 118–148.
7. Khayenzeli, A.W.; Son, W.-J.; Jo, D.-J.; Cho, I.-S. An AIS-Based Study to Estimate Ship Exhaust Emissions Using Spatio-Temporal Approach. *J. Mar. Sci. Eng.* **2025**, *13*, 922. [CrossRef]
8. Mo, S.Y. Impact of the EU ETS on China’s autonomous shipping Industry. *Disput. Settlem.* **2025**, *11*, 111. [CrossRef]
9. Gao, Y.; Tan, Y.; Jiang, D.; Sang, P.; Zhang, Y.; Zhang, J. An Adaptive Prediction Framework of Ship Fuel Consumption for Dynamic Maritime Energy Management. *J. Mar. Sci. Eng.* **2025**, *13*, 409. [CrossRef]
10. Liu, W.; Xu, B.; Li, J. Data-Driven Carbon Emission Dynamics Under Ship In-Port Congestion. *J. Mar. Sci. Eng.* **2025**, *13*, 812. [CrossRef]
11. Zhou, F.; Liu, J.; Zhu, H.; Yang, X.; Fan, Y. A Real-Time Measurement-Modeling System for Ship Air Pollution Emission Factors. *J. Mar. Sci. Eng.* **2022**, *10*, 760. [CrossRef]
12. Wei, Z.; Nguyen, H.C. Storm Surge Forecast Using an Encoder-Decoder Recurrent Neural Network Model. *J. Mar. Sci. Eng.* **2022**, *10*, 1980. [CrossRef]
13. Yu, Y.; Si, X.; Hu, C.; Zhang, J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput.* **2019**, *31*, 1235–1270. [CrossRef] [PubMed]
14. O’Donncha, F.; Hu, Y.; Palmes, P.; Burke, M.; Filgueira, R.; Grant, J. A spatio-temporal LSTM model to forecast across multiple temporal and spatial scales. *Ecol. Inform.* **2022**, *69*, 101687. [CrossRef]
15. Aljahdali, B.M.; Alsubhi, Y.; Alghanmi, A.F.; Sulaimani, H.T.; Samman, A.E. An Innovation Machine Learning Approach for Ship Fuel-Consumption Prediction Under Climate-Change Scenarios and IMO Standards. *J. Mar. Sci. Eng.* **2025**, *13*, 805. [CrossRef]
16. Kim, Y.-R.; Jung, M.; Park, J.-B. Development of a Fuel Consumption Prediction Model Based on Machine Learning Using Ship In-Service Data. *J. Mar. Sci. Eng.* **2021**, *9*, 137. [CrossRef]
17. Li, X.; Zuo, Y.; Jiang, J. DAPNet: A Dual-Attention Parallel Network for the Prediction of Ship Fuel Consumption Based on Multi-Source Data. *J. Mar. Sci. Eng.* **2024**, *12*, 1945. [CrossRef]
18. Liu, Y.; Wang, K.; Lu, Y.; Zhang, Y.; Li, Z.; Ma, R.; Huang, L. A Ship Energy Consumption Prediction Method Based on TGMA Model and Feature Selection. *J. Mar. Sci. Eng.* **2024**, *12*, 1098. [CrossRef]
19. Stitsyuk, A.; Choi, J. xPatch: Dual-Stream Time Series Forecasting with Exponential Seasonal-Trend Decomposition. *arXiv* **2024**, arXiv:2412.17323. [CrossRef]
20. Xue, W.; Zhou, T.; Wen, Q.; Gao, J.; Ding, B.; Jin, R. Card: Channel aligned robust blend transformer for time series forecasting. *arXiv* **2023**, arXiv:2305.12095.
21. Zhou, H.; Zhang, S.; Peng, J.; Zhang, S.; Li, J.; Xiong, H.; Zhang, W. Informer: Beyond efficient transformer for long sequence time-series forecasting. In Proceedings of the AAAI Conference on Artificial Intelligence, Online, 2–9 February 2021; Volume 35, pp. 11106–11115.
22. Nie, Y.; Nguyen, N.H.; Sinthong, P.; Kalagnanam, J. A time series is worth 64 words: Long-term forecasting with transformers. *arXiv* **2022**, arXiv:2211.14730.
23. Liu, Y.; Hu, T.; Zhang, H.; Wu, H.; Wang, S.; Ma, L.; Long, M. itransformer: Inverted transformers are effective for time series forecasting. *arXiv* **2023**, arXiv:2310.06625.
24. Luan, J.; Feng, Y.; Li, H.; Wang, X.; Jia, P.; Kuang, H. Research on the spatiotemporal prediction of autonomous vessel carbon emissions based on the deep learning model. *Sci. Res. Manag.* **2023**, *44*, 75.

25. Zhou, T.; Wang, J.; Hu, Q.; Hu, Z. A Novel Approach to Enhancing the Accuracy of Prediction in autonomous vessel Fuel Consumption. *J. Mar. Sci. Eng.* **2024**, *12*, 1954. [CrossRef]
26. International Maritime Organization (IMO). *Outcome of the Regulatory Scoping Exercise for the Use of Maritime Autonomous Surface Ships (MASS)*; MSC.1/Circ.1638; IMO: London, UK, 2021.
27. Du, W. PyPOTS: A Python toolbox for data mining on Partially-Observed Time Series. *arXiv* **2023**, arXiv:2305.18811.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Carbon Dioxide Storage Site Location and Transport Assignment Optimization for Sustainable Maritime Transport

Yanmeng Tao ¹, Ying Yang ^{2,*}, Yuquan Du ³ and Shuaian Wang ²

¹ School of Transportation Science and Engineering, Beihang University, Beijing 100191, China; yanmengtao@buaa.edu.cn

² Department of Logistics and Maritime Studies, Faculty of Business, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong 999077, China; hans.wang@polyu.edu.hk

³ La Trobe Business School, La Trobe University, Melbourne, VIC 3086, Australia; bill.du@latrobe.edu.au

* Correspondence: shadow.yang@connect.polyu.hk

Abstract: Maritime carbon dioxide (CO₂) transport plays a pivotal role in facilitating carbon capture and storage (CCS) systems by connecting emission sources with appropriate storage sites. This process often incurs significant transportation costs, which must be carefully balanced against penalties for untransported CO₂ resulting from cost-driven decisions. This study addresses the CO₂ storage site location and transport assignment (CSSL-TA) problem, aiming to minimize total tactical costs, including storage site construction, ship chartering, transportation, and penalties for direct CO₂ emissions. We formulate the problem as a mixed-integer programming (MIP) model and demonstrate that the objective function exhibits submodularity, reflecting diminishing returns in facility investment and ship operations. A case study demonstrates the model's effectiveness and practical value, revealing that optimal storage siting, strategic ship chartering, route allocation, and efficient transportation significantly reduce both transportation costs and emissions. To enhance practical applicability, a two-stage planning framework is proposed, where the first stage selects storage sites, and the second employs a genetic algorithm (GA) for transport assignment. The GA-based solution achieves a total cost only 2.4% higher than the exact MIP model while reducing computational time by 57.9%. This study provides a practical framework for maritime CO₂ transport planning, contributing to cost-effective and sustainable CCS deployment.

Keywords: maritime carbon dioxide transport; carbon storage site location and transport assignment; mixed-integer programming model; genetic-algorithm-based two-stage optimization

1. Introduction

With the accelerating global climate crisis, the need to achieve deep decarbonization across energy and industrial systems has become a central concern of governments and industries. As widely acknowledged by the Intergovernmental Panel on Climate Change and the International Energy Agency, carbon capture and storage (CCS) is indispensable for achieving net-zero emission targets, especially in hard-to-abate sectors such as cement, steel, and chemicals [1–3]. In this context, the construction of reliable carbon dioxide (CO₂) storage and transport infrastructure has emerged as a critical enabler of large-scale CCS deployment. Particularly in Europe, where emission sources are scattered across the continent while geological storage capacity is concentrated in offshore basins such as the North Sea, efficient planning of CO₂ shipping infrastructure is vital [4].

Maritime CO₂ transport is increasingly recognized as a vital component in carbon capture and storage (CCS) due to its inherent flexibility and cost efficiency [5]. Compared to fixed pipelines, ship-based transport offers significant advantages, particularly during the initial phases of CCS deployment and in regions with spatial and temporal variability [6]. Maritime transport generally requires lower initial capital investment than pipelines, which demand substantial upfront costs, making ships more suitable for projects with uncertain demand evolution [7]. Pipelines are generally more efficient for shorter distances with high CO₂ volumes [8], with Knoope et al. [7] specifically showing that pipelines are preferred for distances under 200 km when dealing with CO₂ volumes greater than 5 million tons/year. Conversely, for smaller volumes, such as 2.5 million tons/year, shipping becomes competitive at distances greater than 500 km. Furthermore, ships offer operational flexibility, often requiring less frequent maintenance than pipelines, which need regular inspections, especially in corrosive environments [7]. Unlike pipelines, which may require costly structural reinforcements to withstand harsh environmental conditions, ships provide adjustable routing to adapt to operational needs [6]. This combination of reduced capital commitment, operational scalability, adaptability to variable conditions, and suitability for long-distance transport makes maritime transport a compelling and flexible alternative to pipelines for linking industrial clusters to offshore storage hubs [5]. However, the economic viability of ship transport largely depends on the strategic deployment of coastal storage stations, which serve as essential docking and injection points in the maritime supply chain. Inappropriate placement of these facilities may result in extended detours, under-utilized fleet capacity, increased berthing fees, and even emission penalties due to CO₂ that remains untransported [9,10]. Thus, the integrated planning of storage station locations, fleet deployment, and shipping schedules is of central importance for CCS operational efficiency and policy compliance.

Beyond economic and strategic considerations, maritime CO₂ transport also faces distinct thermophysical and operational constraints that significantly affect system design and deployment. CCS is pivotal for mitigating climate change by capturing CO₂ from industrial sources and storing it in geological formations like deep saline aquifers or depleted hydrocarbon fields, supporting global net-zero emission goals [11]. Maritime transport is essential for connecting emission sources to offshore storage sites, especially for long distances (>500 km) or lower CO₂ volumes (<2.5 million tons/year), where it outperforms pipelines in cost effectiveness [12]. However, CO₂ maritime transport requires maintaining specific pressure (0.7–1.5 MPa) and temperature (223–246 K) conditions to keep CO₂ in a liquid state, ensuring high density ($\approx 1159 \text{ kg/m}^3$) and avoiding solidification or supercritical transitions that could cause operational issues like blockages [13,14]. These physical and engineering constraints significantly influence storage site selection, as sites often require supercritical CO₂ for injection, and ship-route allocation, which must account for vessel design and dynamic factors like weather or port congestion [15]. The intricate interplay of these thermophysical requirements and operational considerations presents substantial challenges for the efficient and safe design of CCS networks. The necessity to meticulously manage CO₂ physical state across the supply chain, from transport to injection, while considering logistical complexities, underscores the critical need for robust optimization approaches. Therefore, developing frameworks for the strategic layout of storage sites and transport routes that explicitly account for these factors is paramount to enhance the safety, efficiency, and cost effectiveness of CCS supply chains, thereby facilitating their global deployment. This study is motivated by these challenges, proposing an optimization framework that specifically integrates CO₂ physical state requirements into strategic decision-making for site selection and route allocation.

To complement these technical and operational challenges, several real-world CCUS projects provide practical validation and experience that reinforce the necessity of integrated site selection and transport planning. Saudi Aramco's Uthmaniyah CO₂ Europe Demonstration Project and Norway's Northern Lights Project are two prominent examples. Since 2015, the Uthmaniyah project has captured approximately 800,000 tonnes of CO₂ annually from the Hawiyah Natural Gas Liquids plant, compressing it to 1500–1600 psi and transporting it via an 85 km pipeline to a Jurassic carbonate reservoir for injection [16]. Advanced monitoring ensures secure storage, with about 40% of CO₂ permanently sequestered [17]. The project's success reflects effective infrastructure design, cross-disciplinary integration, and commitment to environmental goals [18], offering transferable insights into offshore site planning and pipeline-based CO₂ transport. In contrast, the Northern Lights Project, operational since 2024, adopts a fully maritime approach: liquefied CO₂ is shipped 100 km from Øygarden to a North Sea saline aquifer using 7500 m³ vessels, with a total capacity of 1.5 million tonnes per year [19]. Monitoring via seismic and pressure sensing ensures long-term storage security [20]. The project's ship-based flexibility, cross-border coordination, and policy support from Norway's Longship initiative [21] provide a working reference for marine-based CCS logistics. Systematically identifying the factors contributing to the success of these projects reveals several dimensions that are directly relevant to the proposed optimization model. These include the importance of strategic site selection based on storage capacity and geographic efficiency, the allocation of flexible ship-based transport routes to accommodate industrial demand, and the role of operational constraints such as berthing frequency and infrastructure reliability. The Northern Lights project demonstrates the feasibility of ship-route planning under variable conditions, which aligns with the model's capability to assign ships, optimize transport frequency, and support multi-source routing. Uthmaniyah's focus on reservoir suitability and monitoring emphasizes the value of matching infrastructure investment with storage performance, a core aspect of the site selection sub-model. Together, these projects inform both the structural formulation and the practical applicability of the proposed approach. Their design and implementation experiences directly parallel the optimization problems explored in this study, particularly regarding routing strategies and offshore injection planning.

Nevertheless, determining optimal locations for CO₂ storage sites is challenging due to the spatial heterogeneity of emission sources, port infrastructure, and long-distance maritime routes. Existing studies focus extensively on the optimization of CO₂ transportation supply chains using pipelines and ships [22–24]. These studies explore various aspects of supply chain design, including the integration of capture, transport, and storage processes, the selection of efficient transport modes, and the development of resilient networks to support CCS deployment. Additionally, research has investigated large-scale infrastructure network design at national or regional levels [25], compared the economic and logistical feasibility of pipelines versus ships [26], and conducted techno-economic analyses to assess the cost effectiveness and environmental impact of ship-based CO₂ transport [27,28]. Some studies have also developed generalized routing strategies for broad geographical areas [29]. However, few studies integrate the strategic siting of coastal CO₂ storage stations with the optimization of detailed maritime transport operations, such as ship scheduling, task assignment, and route planning, to efficiently connect multiple emission sources to these sites. For instance, some frameworks optimize CCS networks by selecting storage sites but prioritize pipeline-based systems, with limited attention to maritime logistics like ship scheduling or multi-source collection [30]. Other studies analyze ship-based CO₂ transport, addressing port infrastructure and ship design, but often neglect critical operational constraints such as the rational allocation of transportation routes based on the volume of CO₂ generated [27,31].

Although these studies offer important insights, they typically overlook key operational constraints that are critical in maritime CO₂ transport planning. First, they often fail to model the required berthing frequency for each emission source, which is essential to ensure that annually generated CO₂ can be transported without delays or accumulation. Second, most models assume full transport or no transport, whereas in real applications, a portion of CO₂ may remain untransported and incur emission penalties, which must be incorporated explicitly in cost evaluation. Third, the existing routing models rarely allow a single ship to collect CO₂ from multiple sources (i.e., dual-source routing), despite its operational relevance in reducing underutilization and improving routing flexibility. These operational features can lead to unrealistic or infeasible plans, highlighting a research gap that this study aims to address.

To address these gaps comprehensively, this study formulates a novel CO₂ storage site location and transport assignment (CSSL-TA) problem. The CSSL-TA problem jointly optimizes the strategic construction of CO₂ storage sites within coastal maritime transport systems and the operational assignment of ship routes and transport tasks across multiple emission sources. A mixed-integer programming (MIP) model is developed to support integrated decision-making, simultaneously determining which storage sites to construct, how to assign ships to feasible transport routes, and how to allocate annual CO₂ volumes from each emission source under system-wide constraints. These realistic constraints include ship transport capacity limits, storage site capacity constraints, berthing frequency requirements at each emission port, annual route operating limits, and CO₂ direct emission constraints. These constraints ensure that the solution is both operationally feasible and aligned with emission coverage regulations. Overall, the proposed MIP model provides a practical and implementable optimization framework for supporting long-term infrastructure planning in coastal CO₂ transport networks.

In particular, the multifaceted contributions of this paper are outlined as follows:

1. We establish an MIP model for jointly optimizing the location of CO₂ storage sites, the assignment of ships to feasible transport routes, and the annual number of round trips made by each chartered ship. The realistic operational constraints in real-world CCS operations are well incorporated in our model. This formulation enables stakeholders to minimize the total cost while supporting practical infrastructure planning for long-term maritime CO₂ transport systems.
2. We prove the necessity of maintaining integrality for construction, ship assignment, and routing frequency decisions. We also demonstrate that the objective function is submodular with respect to all three variable types. This finding highlights the inherent diminishing returns effect in the CO₂ transport system, whereby additional station construction, ship-route assignments, or round trips contribute progressively less to overall cost reduction.
3. Sensitivity analyses on annual available sailing time, sailing speed, and the capacities of candidate storage sites are conducted to examine their effects on total cost, facility deployment, and key performance indicators. The results reveal several important insights: a longer annual sailing time allows ships to complete more round trips and reduces the number of ships required; faster sailing speeds increase transport efficiency but may raise fuel costs and direct emissions; and larger storage capacity reduces the number of required storage sites but may lead to longer transport distances due to wider spatial distribution. These findings provide practical guidance for balancing transport efficiency and infrastructure layout under different operational assumptions.
4. A case study demonstrates the practical value of the proposed model, offering actionable managerial insights into storage construction, ship-route assignment, and

operational planning. This study shows that (i) selecting appropriate storage site locations helps reduce transport distances and lower transportation costs; (ii) strategically chartering different types of ships and assigning them to suitable routes effectively reduces direct CO₂ emissions; (iii) ensuring that ships operate the appropriate number of times on selected routes guarantees the timely transportation of CO₂ from each emission source, and helps minimize total cost by balancing CO₂ transport cost with direct emission penalties.

5. To improve the model's applicability in industries, we further propose a two-stage planning framework that separates strategic storage site selection from transport assignment. The first stage determines site construction decisions by balancing the supply–demand relationship between storage capacities and CO₂ generation volumes, alongside construction costs at candidate sites, while the second stage applies a genetic algorithm (GA) to solve the transport assignment problem. This framework offers a practical alternative to the exact MIP model and can serve as a useful decision-support tool in the scenarios where computational efficiency is critical.

The remainder of this paper is structured as follows. Section 2 reviews the related work on CCS supply chain optimization and CO₂ transport. Section 3 formulates the research problem as an MIP model and presents its theoretical properties. Section 4 reports numerical results and sensitivity experiments. Section 5 introduces the two-stage optimization framework and details the GA-based solution approach. Finally, Section 6 concludes this paper and outlines directions for future research.

2. Literature Review

To comprehensively understand the methodologies for CO₂ storage site selection and transport assignment in CCS systems, this review examines the literature focusing on optimization models and algorithms for CSSL-TA within maritime transportation frameworks.

2.1. CO₂ Storage Site Selection

The selection of geological storage sites for CO₂, particularly deep saline aquifers and depleted hydrocarbon fields, typically utilizes multi-criteria decision-making frameworks combined with thorough geological assessments. Bachu [31] establishes fundamental geological screening criteria, including sufficient reservoir depths to ensure CO₂ remains in a dense state, adequate porosity and permeability, and the presence of a competent caprock with sufficient thickness and low permeability to ensure containment. Grataloup et al. [32] develop a two-stage screening approach employing both exclusionary and qualification criteria, utilizing Geographic-Information-System-based spatial analysis for site identification in the Paris Basin. Alcalde et al. [33] propose a four-stage workflow integrating Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) analysis to systematically weigh criteria such as storage capacity, injectivity potential, and containment risk within the Acorn project context. Raza et al. [34] emphasize the importance of petrophysical models, like Archie's equation for water saturation determination and Bahadori's correlation for predicting CO₂ density under reservoir conditions, alongside understanding various trapping mechanisms relevant to saline aquifers. Hsu et al. [35] apply the Analytic Network Process to prioritize interdependent criteria, notably caprock permeability, thereby enhancing the robustness of site evaluation. For offshore contexts, Luo et al. [36] demonstrate the use of seismic inversion and well-log data integration to assess structural trap integrity and reservoir-caprock characteristics in the South Yellow Sea Basin, combining geophysical and petrophysical analyses. Following International Organization for Standardization guidelines, Aarnes et al. [37] employ probabilistic risk

assessments and dynamic modeling to evaluate potential CO₂ migration pathways and seal integrity. Furthermore, Akhurst et al. [38] describe risk-assessment-led pre-characterization methodologies that utilize publicly available data in the early stages to identify and mitigate potential risks associated with storage security, thereby guiding subsequent technical characterization efforts. Techno-economic assessments by Gruson et al. [39] on four diverse European storage sites confirm that storage costs are highly site-specific, influenced significantly by the chosen development strategy and monitoring plan, highlighting the difficulty in establishing a generic storage cost. Table 1 summarizes key studies on CO₂ storage site selection, highlighting their considered factors and applicable scopes.

Table 1. Key studies on CO₂ storage site selection.

Paper	Considered Factor						Applicable Scope
	Reservoir Depth	Porosity and Permeability	Caprock Integrity	Storage Capacity	Injectivity Potential	Containment Risk	
Bachu [31]	✓	✓	✓	✓			Deep saline aquifers, depleted hydrocarbon fields
Grataloup et al. [32]	✓	✓	✓	✓			Paris Basin
Alcalde et al. [33]			✓	✓	✓	✓	General
Raza et al. [34]	✓	✓		✓	✓		General
Hsu et al. [35]			✓			✓	South Yellow Sea Basin
Luo et al. [36]	✓	✓		✓			General
Aarnes et al. [37]		✓		✓			Commercial-scale storage
Akhurst et al. [38]		✓		✓			European

Note: (1) The “✓” in the table indicates that the respective paper considers the corresponding objective costs; (2) Injectivity Potential: the ability of a reservoir to accept and accommodate CO₂ injection over time, influenced by factors like permeability, porosity, and reservoir pressure; (3) Containment Risk: the risk of CO₂ leakage from the storage site, primarily due to inadequate caprock integrity or potential migration pathways, which could compromise long-term storage security; (4) General: the study’s methods or criteria can be applied broadly to various geological settings (e.g., saline aquifers or depleted fields) without being tied to a specific region or project.

2.2. Optimization of Maritime CO₂ Transport

The optimization of maritime CO₂ transport uses advanced modeling techniques to navigate logistical complexities. Techno-economic modeling by Roussanaly et al. [12], optimizing ship parameters and port infrastructure over significant distances, indicates that maritime transport presents advantages over pipelines for lower CO₂ volumes. Zhang et al. [9] develop an MIP model to optimize the strategic deployment of a CCS supply chain in Northeast China. This model integrates the selection of emission sources, capture technologies, pipeline routes, intermediate hubs, utilization options such as Enhanced Oil Recovery, and geological storage sites, demonstrating potential net cost reduction through revenue generation. Addressing the need for robust planning under uncertainty, Elahi et al. [40] formulate a multi-stage stochastic MIP optimization framework for the UK CCS system. This approach accommodates temporal and spatial integration while considering financial market and storage capacity uncertainties, yielding flexible investment strategies designed to minimize potential financial losses. Knoope et al. [41] employ Real Options Analysis (ROA) to evaluate the impact of uncertainties like CO₂ pricing and transportation tariffs on infrastructure investment timing, finding that such uncertainties can significantly elevate the required pre-investment price threshold. To enhance supply chain resilience, Gabrielli et al. [42] propose an MIP model that minimizes expected un-

sequestered CO₂ under various disruption scenarios. Utilizing scenario-based stochastic programming, their model incorporates alternative transport options and demonstrates potential cost savings relative to deterministic planning. Engebø et al. [43] introduce a risk-based optimization framework integrating fault tree analysis and Monte Carlo simulations to quantitatively assess disruption risks and inform the prioritization of redundant shipping routes. Focusing on large-scale network development, Kjærstad et al. [44] integrate techno-economic modeling of the European power sector with detailed CO₂ transport infrastructure analysis. Their work investigates potential CCS network configurations spanning the whole of Europe, emphasizes how the spatio-temporal distribution of capture facilities and site-specific storage properties crucially determine optimal pipeline routing and deployment schedules, and quantifies the significant cost impact of restricting onshore storage options. Similarly, Morbee et al. [45] detail the InfraCCS model, a tool designed using methodological innovations such as node clustering and specialized algorithms for route pre-selection, to determine the optimal least-cost extent and associated investment cost of a potential large-scale European CO₂ pipeline transport network required for decarbonization scenarios up to 2050. Employing methodological innovations for route selection and costing, the model estimates network investment requirements, primarily for pipelines, and highlights the critical need for cross-border coordination. Targeting the specifics of maritime logistics, Bennæs et al. [46] formulate an MIP-based model to optimize key parameters of a ship-based CO₂ transportation system, including fleet composition, interim storage, and ship specifications. Implemented using Gurobi, this model minimizes overall logistics costs while adhering to operational constraints like port berthing availability and delivery schedules. Complementing transport optimization, Ajayi et al. [47] review the methodologies for estimating CO₂ storage capacity, and advocate for the integration of both static and dynamic models within multi-criteria decision analysis frameworks. Incorporating Monte Carlo simulations to manage geological uncertainties, their approach demonstrates potential improvements in site ranking accuracy. Table 2 summarizes key studies on optimization of maritime CO₂ transport, highlighting their objectives, methodologies, and limitations.

Table 2. Key studies on maritime CO₂ transport (✓ indicates the study considered the corresponding objective or modeling method).

Paper	Objective				Modeling Method
	CO ₂ Transport Cost	CO ₂ Direct Emission Cost	Ship Chartering Cost	Storage Site Construction Cost	
Zhang et al. [9]	✓			✓	MIP
Roussanaly et al. [12]	✓		✓		Techno-economic modeling (TEM)
Elahi et al. [40]	✓			✓	MIP
Knoope et al. [41]	✓			✓	Real Options Analysis (ROA)
Gabrielli et al. [42]	✓			✓	MIP
Engebø et al. [43]	✓				Fault tree analysis, Monte Carlo simulations
Kjærstad et al. [44]	✓			✓	TEM
Morbee et al. [45]	✓				Node clustering
Bennæs et al. [46]	✓		✓		MIP
Our Work	✓	✓	✓	✓	MIP, two-stage optimization used GA

Our work distinguishes itself from existing research by adopting a more comprehensive approach to optimizing the CO₂ chain. As highlighted in the comparative table, our model is unique in its simultaneous consideration of four key objective costs: CO₂ transport cost, CO₂ direct emission cost, ship-chartering cost, and storage site construction cost. In contrast, previous studies typically focus on a subset of these factors. For instance, while many studies address CO₂ transport and storage site construction costs, none of the cited works incorporate all four critical cost components that our research addresses. This holistic consideration, facilitated by our modeling method employing MIP with a two-stage optimization using a genetic algorithm, allows for a more robust and realistic assessment of the overall CCTS project economics and environmental impact.

2.3. Research Gap

Despite these significant advances, several key gaps remain in the existing literature. First, the majority of the studies do not adequately address critical operational constraints specific to maritime transport, such as limitations on berthing frequency, the explicit handling of penalties for direct CO₂ emissions, and the complexities of multi-source, multi-destination routing. Second, there is a lack of integrated models that simultaneously optimize strategic site selection and detailed transport assignment within a single framework. Third, efficient solution methods are often not proposed, especially for tackling the computational challenges posed by large-scale instances. To bridge these gaps, we propose an MIP model designed for the CSSL-TA problem, which integrates strategic site selection with detailed transport assignment. This MIP model uniquely incorporates real-world operational constraints related to berthing frequency, direct emission penalties, and multi-source routing to achieve a minimal-cost and operationally feasible solution. Additionally, recognizing large-scale computational demands, we introduce a practical two-stage solution framework where a GA efficiently optimizes transport assignment following initial site selection, thereby facilitating more effective planning for sustainable CCS deployment. These all confirm the statement of our contributions to the existing literature described at the end of Section 1.

3. Problem Formulation

In this section, we first introduce the problem background and describe the challenges in CO₂ storage site location and ship-routing assignment in Section 3.1. Following that, we formulate the problem using an MIP model in Section 3.2. Finally, Section 3.3 provides a detailed description of model analysis.

3.1. Problem Description

The escalating need to mitigate CO₂ emissions has positioned CCS as a cornerstone technology in the global transition to a low-carbon economy. Governments and industries alike are increasingly driven by stringent environmental regulations and ambitious climate targets, spurring significant investments in CCS systems. These systems not only capture CO₂ from industrial processes but also prevent its release into the atmosphere, thereby addressing both environmental and regulatory challenges. The development of CCS infrastructure entails a coordinated interplay of technological innovation, financial investment, and policy support. Such integration is critical to the creation of a robust, scalable CCS network that can operate efficiently across long-term planning horizons and manage large quantities of CO₂ in a cost-effective and environmentally sustainable manner.

In maritime CCS systems, the strategic placement of coastal CO₂ storage sites presents a complex optimization challenge driven by several competing factors. Constructing these sites requires significant investment. Meanwhile, ships transporting CO₂ operate along predefined routes that link emission sources to selected storage sites. If storage sites are poorly positioned, ships may need to follow longer routes, leading to increased fuel consumption. These inefficiencies elevate transportation costs and may also result in higher direct CO₂ emissions, which incur substantial penalty costs under emission control regulations. In light of these considerations, it is crucial to precisely determine the optimal locations of CO₂ storage sites and ship assignments to minimize the total system cost in the planning horizon. This cost includes the construction cost of selected CO₂ storage sites, the chartering cost of ships (incurred when a ship is chartered for CO₂ transportation service), the transportation cost (including fuel cost along sailing routes and berthing cost of ships), and the penalty cost for direct CO₂ emissions into the atmosphere. Meanwhile, the system must satisfy annual CO₂ capture and emission constraints: the total CO₂ emissions generated at each emission source in a year must be transported to a storage site in a timely manner; otherwise, some of them will be released to the atmosphere within the permitted limit.

Consider the CO₂ storage site location problem in a coastal-area transportation network during a multi-year planning horizon, where the number of years in the planning horizon is denoted by T . The set of candidate locations where CO₂ storage sites can be constructed is denoted by I (indexed by i). Hereafter, we use ‘storage site’ and ‘location’ interchangeably. The set of CO₂ emission sources is denoted by P (indexed by p). Ships are used to transport CO₂ from the emission sources to storage sites along fixed transport routes. The set of all feasible CO₂ transport routes is denoted by R (indexed by r). Each route $r \in R$ includes exactly one storage site $i \in I$, and either one or two emission sources from P , i.e., $i(r) \in I$, $P(r) \subseteq P$, with $|P(r)| \in \{1, 2\}$. If a route connects one emission source with a storage site, it represents a round trip between the two; if it connects two emission sources with a storage site, it represents a round trip in which the ship departs from the storage site, visits both emission sources in sequence, and returns to the storage site.

The set of ships available for charter is denoted by S (indexed by s). The decision of whether to charter a ship is based on the actual transport demand. If a ship is chartered, it remains in service for the entire planning horizon, and its chartering cost C_s^{ship} is incurred as a one-time fixed expense. Once a ship is chartered, it is assigned to a single route $r \in R$, and this assignment remains fixed in the planning horizon. Each ship is required to berth both when receiving CO₂ at an emission source and when injecting CO₂ to a storage site. The berthing cost of ship s on route r , denoted by B_{sr} , depends on the number of emission sources visited along the route. Each berthing operation of ship s at an emission source or a storage site incurs a fixed cost b_s , which includes tugboat assistance, pilotage, and mooring services. If route r includes two emission sources, then $B_{sr} = 3b_s$; if it includes only one emission source, then $B_{sr} = 2b_s$. This parameter contributes to the transportation cost together with the fuel cost F_{sr} . For a route involving one emission source and one storage site, a full round trip consists of traveling from the emission source to the storage site and back. For a route that includes two emission sources and one storage site, a full round trip consists of sailing from the storage site to the first emission source, continuing to the second emission source, and then returning to the storage site. The maximum number of round trips that ship s can make on route r per year is denoted by N_{sr} , which depends on the ship’s sailing speed v_s and the route’s sailing distance l_r . Assuming a full year is available for operation and denoting one year’s available sailing time by t_0 , N_{sr} is calculated as the available sailing time in one year divided by the time required for one

round trip on route r , i.e., $N_{sr} = t_0 / (l_r / v_s)$. This parameter is not required to be an integer, as any incomplete trip at the end of a year is assumed to be completed in the following year. To account for realistic factors that significantly impact maritime CO₂ transport, such as weather conditions, route congestion, and berthing delays, it is important to recognize their influence on ship sailing speeds and available navigation time. Adverse weather, congestion, and delays can reduce sailing speeds or decrease the available navigation time and limiting the N_{sr} . Adjusting N_{sr} to reflect these operational uncertainties and constraints can effectively capture the real-world challenges faced in maritime shipping.

Each storage site $i \in I$ has a capacity V_i , which represents the maximum amount of CO₂ that can be injected into the site per year. The construction cost of a candidate storage site i , denoted by C_i , consists of two components: a fixed cost component, incurred only if the location i is eligible for construction, and a variable cost component $b_i \cdot V_i$, which scales linearly with the site's storage capacity, i.e., $C_i = a_i + b_i V_i$. The amount of CO₂ produced by each emission source $p \in P$ is fixed per year and denoted by E_p . To ensure timely removal of CO₂ from each source, a minimum berthing frequency requirement N_p is imposed, specifying the minimum number of ship arrivals required at source p annually. For each round trip, the amount of CO₂ carried from source p by ship s must not exceed the ship's transport capacity Q_s . If any CO₂ remains untransported by the end of the year, it is directly released into the atmosphere. The amount of CO₂ that emission source p is allowed to release per year is capped by $E_p^{\max}$, and the penalty cost of emitting a ton of CO₂ into the atmosphere is denoted by $\hat{C}$.

The objective function of this study is to minimize the total system cost in the planning horizon, including the construction cost of the selected CO₂ storage sites, the total chartering cost of all chartered ships, the total transportation costs consisting of both fuel and berthing expenses, and the penalty costs for CO₂ that is directly emitted into the atmosphere. To this end, we need to decide which candidate storage sites to construct. This decision is represented by the binary variable x_i , which equals 1 if the i -th candidate CO₂ storage site is constructed, and 0 otherwise. We also determine whether each ship $s \in S$ is chartered and, if so, which route it is assigned to. This is represented by the binary decision variable z_{sr} , which equals 1 if ship s is chartered and assigned to route r , and 0 otherwise. The number of round trips that ship s makes on route r per year is described by the continuous decision variable n_{sr} . The amount of CO₂ transported from source p on route r by ship s per year is denoted by the continuous decision variable w_{srp} . The amount of CO₂ released into the atmosphere by source p per year is represented by the continuous decision variable e_p .

3.2. Model Design

In this subsection, we present an MIP model based on the problem setting described above. The model captures the integrated planning decisions involved in the construction of CO₂ storage sites and the chartering and routing of ships for CO₂ transport. It focuses on strategic decisions for constructing storage sites and operational decisions concerning ship assignment, route scheduling, and annual CO₂ transportation volumes. The objective is to minimize the total system costs while adhering to practical constraints such as storage capacity limits, transport capabilities requirements, berthing frequency requirements, and emission regulations. Table 3 summarizes the notations used in the model.

Table 3. Notations used in the model's formulation.

Sets	
I	The set of candidate locations where CO ₂ storage sites can be constructed
P	The set of CO ₂ emission sources
S	The set of ships transporting CO ₂
R	The set of feasible CO ₂ transport routes
Indices	
i	Index for locations in I
p	Index for CO ₂ emission sources in P
s	Index for ships in S
r	Index for routes in R
Parameters	
T	The number of years in the planning horizon
V_i	The capacity of the i -th CO ₂ storage site
C_i	The cost of constructing a candidate CO ₂ storage site at location i
E_p	The amount of CO ₂ produced by source p per year
$E_p^{\max}$	The maximum amount of CO ₂ that emission source p is allowed to release into the atmosphere per year
N_p	The minimum annual berthing requirement at source p
C_s^{ship}	The chartering cost of ship s
Q_s	The transport capacity of ship s
F_{sr}	The fuel cost for ship s completing one full trip on route r
B_{sr}	The berthing cost for ship s on route r
N_{sr}	The maximum number of round trips that ship s can make on route r per year
$\hat{C}$	The penalty cost of emitting a ton of CO ₂ into the atmosphere
Decision Variables	
x_i	Binary variable, which equals 1 if the i -th candidate CO ₂ storage site is constructed, and 0 otherwise
z_{sr}	Binary variable, which equals 1 if ship s is chartered and assigned to route r , and 0 otherwise
n_{sr}	Continuous variable, indicating the number of round trips that ship s makes on route r per year
w_{srp}	Continuous variable, indicating the amount of CO ₂ transported from source p on route r by ship s per year
e_p	Continuous variable, indicating the amount of CO ₂ released into the atmosphere by source p per year

The MIP model is developed to optimize the construction of CO₂ storage sites, the assignment of ships to transport routes, and the annual routing frequency of each ship, which can be written as follows:

$$\begin{aligned}
 \min \quad & \sum_{i \in I} C_i x_i + \sum_{s \in S} \sum_{r \in R} C_s^{\text{ship}} z_{sr} + \sum_{s \in S} \sum_{r \in R} T(F_{sr} + B_{sr}) n_{sr} z_{sr} + \sum_{p \in P} T \hat{C} e_p & (1) \\
 \text{s.t.} \quad & \sum_{r \in R} z_{sr} \leq 1 & \forall s \in S & (2) \\
 & \sum_{p \in P(r)} w_{srp} \leq Q_s n_{sr} z_{sr} & \forall s \in S, \forall r \in R & (3) \\
 & \sum_{s \in S} \sum_{r: i(r)=i} \sum_{p \in P(r)} w_{srp} \leq V_i x_i & \forall i \in I & (4) \\
 & \sum_{s \in S} \sum_{r: p \in P(r)} n_{sr} z_{sr} \geq N_p & \forall p \in P & (5) \\
 & \sum_{s \in S} \sum_{r: p \in P(r)} w_{srp} + e_p = E_p & \forall p \in P & (6) \\
 & 0 \leq n_{sr} \leq N_{sr} z_{sr} & \forall s \in S, \forall r \in R & (7) \\
 & x_i \in \{0, 1\} & \forall i \in I & (8) \\
 & z_{sr} \in \{0, 1\} & \forall s \in S, \forall r \in R & (9) \\
 & w_{srp} \geq 0 & \forall s \in S, \forall r \in R, \forall p \in P & (10) \\
 & 0 \leq e_p \leq E_p^{\max} & \forall p \in P & (11)
 \end{aligned}$$

The objective function (1) minimizes the total cost, which consists of four components: the construction cost of selected CO₂ storage sites, the total chartering cost of all chartered ships, the transportation cost in the planning horizon (including fuel and berthing costs), and the penalty cost for direct CO₂ emissions into the atmosphere. Constraint (2) ensures that each ship can be assigned to at most one route, reflecting the condition that not all ships must be chartered. Constraint (3) denotes that for each ship assigned to a route, the total amount of CO₂ transported per year does not exceed the product of its transport capacity and the number of annual round trips. Constraint (4) requires that the total injected CO₂ into each constructed storage site per year does not exceed its capacity. Constraint (5) requires that each emission source is visited by ships a sufficient number of times per year to meet its minimum berthing requirement, thereby ensuring an adequate transport frequency and enabling timely removal of the generated CO₂. Constraint (6) ensures that all CO₂ generated by each source is either transported or emitted. Constraint (7) imposes operational bounds on the annual number of round trips that each chartered ship can make on its assigned route, ensuring that this value does not exceed the maximum allowable trips based on ship speed and route distance. Constraints (8)–(11) are the domains of decision variables.

3.3. Model Analysis

In this section, we introduce some theoretical properties regarding the MIP model presented in Section 3.2.

3.3.1. Linearization of Bilinear Terms Involving Ship Assignment Variables

To simplify the structure of the MIP model and reduce computational complexity, we next examine whether the bilinear term $n_{sr} z_{sr}$ can be safely linearized. Through analytical derivation, we establish Lemma 1.

Lemma 1. *The bilinear terms $n_{sr} z_{sr}$ in the model can be replaced by linear terms n_{sr} .*

Proof. In the MIP model, the bilinear term $n_{sr} z_{sr}$ appears in both the objective function to calculate total cost and Constraints (3) and (5) to limit the total transported volume by the ship's capacity and to ensure the minimum berthing requirements for each source by controlling the number of annual round trips. According to Constraint (7), $0 \leq n_{sr} \leq N_{sr} z_{sr}$, the number of round trips n_{sr} is bounded by the binary variable z_{sr} . This directly implies that when $z_{sr} = 0$, $n_{sr} = 0$; and when $z_{sr} = 1$, $n_{sr} \in [0, N_{sr}]$. In both cases, the product

$n_{sr}z_{sr}$ is exactly equal to n_{sr} . Therefore, according to the feasible region defined by Constraint (7), we have: $n_{sr}z_{sr} = n_{sr}, \forall s \in S, \forall r \in R$, which completes the proof. $\square$

This replacement eliminates unnecessary nonlinearity from both the objective and the constraints, improving computational tractability while preserving all model logic.

3.3.2. Submodular Property of the Objective Function

A set function $f : 2^Y \rightarrow \mathbb{R}$ defined on the subsets of a finite set Y is called submodular if it satisfies the diminishing returns property: for every $Y_1 \subseteq Y_2 \subseteq Y$ and every C , it holds that $f(Y_1 \cup \{y\}) - f(Y_1) \geq f(Y_2 \cup \{y\}) - f(Y_2)$ [48]. In simple terms, a submodular function is a set function that describes the relationship between a set of inputs and an output, where adding an additional input has a decreasing additional benefit (diminishing returns). This property essentially means that the marginal gain from adding an element to a smaller set is at least as great as the gain from adding the same element to a larger set. Submodular functions are particularly interesting because they help model a variety of naturally occurring phenomena, such as economies of scale, network effects, or the spread of information in social networks.

For notation convenience, we denote $\mathbf{x} = (x_i)_{i \in I}$, $\mathbf{z} = (z_{sr})_{s \in S, r \in R}$, $\mathbf{n} = (n_{sr})_{s \in S, r \in R}$ and $\mathbf{e} = (e_p)_{p \in P}$, which are restricted in the feasible region Ω . Then, the objective function can be presented by a function of $\mathbf{x}, \mathbf{z}, \mathbf{n}$ and $\mathbf{e}$, denoted by $F_\Omega(\mathbf{x}, \mathbf{z}, \mathbf{n}, \mathbf{e})$. We then give the following lemma about $F_\Omega(\mathbf{x}, \mathbf{z}, \mathbf{n}, \mathbf{e})$. The feasible region Ω is defined by Constraints (2)–(11), which include operational and capacity constraints such as construction decisions, ship chartering and route assignments, ship round-trip limits, emission allocations, and berthing frequency requirements. The objective function $F_\Omega(\mathbf{x}, \mathbf{z}, \mathbf{n}, \mathbf{e})$ represents the total cost in the planning horizon, composed of storage site construction cost, ship chartering cost, transportation cost (including fuel and berthing cost), and penalty cost for untransported CO₂ emissions, where the decision variables are constrained in the feasible region Ω .

Lemma 2. *The objective function $F_\Omega(\mathbf{x}, \mathbf{z}, \mathbf{n}, \mathbf{e})$ is submodular regarding variables $\mathbf{x}$, $\mathbf{z}$ and $\mathbf{n}$.*

Proof. The objective is to minimize the total cost of selecting a subset of locations x_i for storage site construction, a subset of ships and route assignments z_{sr} , and a subset of round-trip quantities n_{sr} , to cover all required service elements and satisfy the direct emission limits. The direct emission volume e_p is linked to the transport CO₂ volume w_{srp} via Constraint (6): $\sum_{s \in S} \sum_{r: p \in P(r)} w_{srp} + e_p = E_p, \forall p \in P$. The set covering function is known to be submodular [48], which suggests that our objective function $F_\Omega(\mathbf{x}, \mathbf{z}, \mathbf{n}, \mathbf{e})$ is also submodular with respect to the variables $\mathbf{x}, \mathbf{z}$, and $\mathbf{n}$.

In the subsequent contents, we prove this submodularity by a more rigorous step-by-step analysis.

1. Step 1: The binary variables x_i indicate whether the CO₂ storage site at location i is constructed. Define $\mathbf{x}_1$ and $\mathbf{x}_2$ as two subsets of constructed CO₂ storage sites and let $\mathbf{x}_1 \subseteq \mathbf{x}_2 \subseteq \mathbf{x}$, where each $x_i \in \mathbf{x}, \forall i \in I$, is 1. Then, for any $x_{i'} \notin \mathbf{x}_2$, the marginal gain by adding another storage site $x_{i'}$ to $\mathbf{x}_1$ can be defined as $F_\Omega(\mathbf{x}_1, \mathbf{z}, \mathbf{n}, \mathbf{e}) - F_\Omega(\mathbf{x}_1 \cup \{1\}, \mathbf{z}, \mathbf{n}, \mathbf{e})$. We note that the marginal gain is represented by $F_\Omega(\mathbf{x}_1, \mathbf{z}, \mathbf{n}, \mathbf{e}) - F_\Omega(\mathbf{x}_1 \cup \{1\}, \mathbf{z}, \mathbf{n}, \mathbf{e})$ since the objective function is the cost. Similarly, $F_\Omega(\mathbf{x}_2, \mathbf{z}, \mathbf{n}, \mathbf{e}) - F_\Omega(\mathbf{x}_2 \cup \{1\}, \mathbf{z}, \mathbf{n}, \mathbf{e})$ is the marginal gain by adding a same storage site.
2. Step 2: Since $\mathbf{x}_1 \subseteq \mathbf{x}_2$, any potential cost reduction in service coverage by adding $x_{i'}$ to $\mathbf{x}_2$ is also possible when added to $\mathbf{x}_1$, but may remain unchanged due to already covered demands with the same cost. Therefore, the marginal gain by adding another storage site $x_{i'}$ to the existing subsets of sites $\mathbf{x}_1$ is greater than or equal to adding it to $\mathbf{x}_2$.

3. Step 3: In the presence of limited injection demands, the possibility of yielding zero marginal benefit by adding a new storage site $x_{i'}$ to x_2 is higher than adding it to x_1 . In other words, when most emission sources are already well served in x_2 , adding a new site will provide negligible additional benefit, thus reinforcing the diminishing return property.

Conclusively, we always have $F_{\Omega}(x_1, z, n, e) - F_{\Omega}(x_1 \cup \{1\}, z, n, e) \geq F_{\Omega}(x_2, z, n, e) - F_{\Omega}(x_2 \cup \{1\}, z, n, e)$.

Similar submodular behavior holds for the variables z_{sr} and n_{sr} , as shown in the following analysis.

The binary variables z_{sr} indicate whether ship s is chartered and assigned to route r . Define z_1 and z_2 as two subsets of ship-route assignments and let $z_1 \subseteq z_2 \subseteq z$, where each $z_{sr} \in z, \forall s \in S, \forall r \in R$ is 1. Then, for any $z_{sr'} \notin z_2$, the marginal gain by adding another ship-route assignment $z_{sr'}$ to z_1 can be defined as $F_{\Omega}(x, z_1, n, e) - F_{\Omega}(x, z_1 \cup \{1\}, n, e)$. We note that the marginal gain is represented by $F_{\Omega}(x, z_1, n, e) - F_{\Omega}(x, z_1 \cup \{1\}, n, e)$ since the objective function is the cost. Similarly, $F_{\Omega}(x, z_2, n, e) - F_{\Omega}(x, z_2 \cup \{1\}, n, e)$ is the marginal gain by adding the same storage site. Since $z_1 \subseteq z_2$, any potential cost reduction in service coverage by adding $z_{sr'}$ to z_2 is also possible when added to z_1 , but may remain unchanged due to already covered demands with the same cost. Therefore, the marginal gain by adding another ship-route assignment $z_{sr'}$ to the existing subsets of sites z_1 is greater than or equal to adding it to z_2 . In the presence of limited transport demands, the possibility of yielding zero marginal benefit by adding a new assignment $z_{sr'}$ to z_2 is higher than adding it to z_1 . In other words, when most emission sources are already well served in z_2 , adding a new site will provide negligible additional benefit, thus reinforcing the diminishing return property. Conclusively, we always have $F_{\Omega}(x, z_1, n, e) - F_{\Omega}(x, z_1 \cup \{1\}, n, e) \geq F_{\Omega}(x, z_2, n, e) - F_{\Omega}(x, z_2 \cup \{1\}, n, e)$.

Additionally, the continuous variables n_{sr} indicate the number of round trips that ship s make on route r . Define $n^{(1)}$ and $n^{(2)}$ as two feasible round-trip decisions and let $n^{(1)} \leq n^{(2)}$. Then, the marginal gain by increasing a round trip quantity $n_{sr'}$ from the lower level $n^{(1)}$ can be defined as $F_{\Omega}(x, z, n^{(1)}, e) - F_{\Omega}(x, z, n^{(1)} + n_{sr'}, e)$. Similarly, $F_{\Omega}(x, z, n^{(2)}, e) - F_{\Omega}(x, z, n^{(2)} + n_{sr'}, e)$ is the marginal gain by increasing a round trip from the higher level $n^{(2)}$. Since $n^{(1)} \leq n^{(2)}$, any potential cost reduction brought by increasing the number of round trips from a lower level $n^{(1)}$ is also possible when increasing from the higher level $n^{(2)}$ but may remain unchanged due to already fulfilled transportation tasks. Therefore, the marginal gain by increasing the number of round trips from $n^{(1)}$ to $n^{(1)} + n_{sr'}$ is greater than or equal to increasing it from $n^{(2)}$. In the presence of limited transportable volumes, the possibility of yielding zero marginal benefit by increasing a new round trip quantity $n_{sr'}$ from $n^{(1)}$ is higher than increasing it from $n^{(2)}$. In other words, when most transport volumes are already well served by a large number of round trips $n^{(2)}$, adding more offers limited additional benefit, thus reinforcing the diminishing return property. Conclusively, we always have $F_{\Omega}(x, z, n^{(1)}, e) - F_{\Omega}(x, z, n^{(1)} + n_{sr'}, e) \geq F_{\Omega}(x, z, n^{(2)}, e) - F_{\Omega}(x, z, n^{(2)} + n_{sr'}, e)$.

The above analysis verifies the diminishing return property for the decision variables x_i , z_{sr} and n_{sr} . Therefore, the objective function is submodular regarding these variables. $\square$

3.3.3. Model Tightening via Capacity–Demand Coupling

To further strengthen the LP relaxation of the model and reduce the solution space without excluding any feasible integer solution, we examine whether constraints involving capacity limits can be tightened based on the actual amount of CO₂ allocated in the system. By exploiting the property of demand–capacity coupling, we establish Lemma 3 and Lemma 4.

Lemma 3. *The right-hand side of the annual storage capacity Constraint (4) can be tightened by replacing V_i with $\min(V_i, E_i)$, where E_i represents the total amount of CO₂ generated by all emission sources connected to the storage site i .*

Proof. Consider Constraint (4) of the model: $\sum_{s \in S} \sum_{r: i(r)=i} \sum_{p \in P(r)} w_{srp} \leq V_i x_i, \forall i \in I$, where w_{srp} indicating the amount of CO₂ transported from source p on route r by ship s per year, and V_i is the capacity of the i -th CO₂ storage site. We denote the set of emission sources that are connected to storage site i by $P(i)$. Specifically, $P(i) \subseteq P$. The total amount of CO₂ generated by all emission sources in $P(i)$ is expressed as follows: $E_i = \sum_{p \in P(i)} E_p$. From Constraint (6), we know that $\sum_{s \in S} \sum_{r: p \in P(r)} w_{srp} + e_p = E_p, \forall p \in P$, which ensure that all CO₂ generated by each source is either transported or emitted. Therefore, for each constructed site i ($x_i = 1$), the value E_i equals the sum of the CO₂ volume transported from emission sources $p \in P(i)$ to all storage sites connected to these sources, plus the total CO₂ volume directly emitted from $p \in P(i)$, that is:

$$\sum_{s \in S} \sum_{r \in R} \sum_{p \in P(i)} w_{srp} + \sum_{p \in P(i)} e_p = E_i. \quad (12)$$

Since each emission source connects to at least one storage site, the following inequality holds:

$$\sum_{s \in S} \sum_{r: i(r)=i} \sum_{p \in P(r)} w_{srp} \leq \sum_{s \in S} \sum_{r \in R} \sum_{p \in P(i)} w_{srp}. \quad (13)$$

Combined with (12), we have:

$$\sum_{s \in S} \sum_{r: i(r)=i} \sum_{p \in P(r)} w_{srp} + \sum_{p \in P(i)} e_p \leq E_i. \quad (14)$$

According to Constraint (11), e_p is non-negative for all $p \in P$, and thus we can conclude that $\sum_{r: i(r)=i} \sum_{p \in P(r)} e_p \geq 0$. Then, we have $\sum_{s \in S} \sum_{r: i(r)=i} \sum_{p \in P(r)} w_{srp} \leq E_i x_i$ for $x_i = 1$ and $\sum_{s \in S} \sum_{r: i(r)=i} \sum_{p \in P(r)} w_{srp} = 0$ for $x_i = 0$. Conclusively, we always have $\sum_{s \in S} \sum_{r: i(r)=i} \sum_{p \in P(r)} w_{srp} \leq E_i x_i, \forall i \in I$. Therefore, we can replace V_i in the original Constraint (4) with $\min(V_i, E_i)$. The modified constraints become

$$\sum_{s \in S} \sum_{r: i(r)=i} \sum_{p \in P(r)} w_{srp} \leq \min(V_i, E_i) \cdot x_i, \forall i \in I, \quad (15)$$

which completes the proof. $\square$

Lemma 4. *The right-hand side of the ship transport capacity Constraint (3) can be tightened by replacing $Q_s n_{sr}$ with $\min(Q_s n_{sr}, E_r)$, where E_r represents the amount of CO₂ generated by all emission sources on route r .*

Proof. Consider Constraint (3) of the model: $\sum_{p \in P(r)} w_{srp} \leq Q_s n_{sr} z_{sr}, \forall s \in S, \forall r \in R$, where w_{srp} indicates the amount of CO₂ transported from source p on route r by ship s per year, and Q_s is the transport capacity of ship s . According to Lemma 1, we know that the bilinear term $n_{sr} z_{sr}$ can be replaced with n_{sr} . Then, the constraints can be expressed

as follows: $\sum_{p \in P(r)} w_{srp} \leq Q_s n_{sr}$. Let the total amount of CO₂ generated by all emission sources on route r be denoted as follows: $E_r = \sum_{p \in P(r)} E_p$. From Constraint (6), we know that: $\sum_{s \in S} \sum_{r: p \in P(r)} w_{srp} + e_p = E_p, \forall p \in P$, which ensure that all CO₂ generated by each source is either transported or emitted. Therefore, for each route r which is assigned to ship s ($n_{sr} > 0$), the value E_r equals the sum of the CO₂ volume transported by ships from emission sources $p \in P(r)$ to all storage sites connected to these sources, plus the total CO₂ volume directly emitted from $p \in P(r)$, that is,

$$\sum_{s \in S} \sum_{r \in R} \sum_{p \in P(r)} w_{srp} + \sum_{p \in P(r)} e_p = E_r. \quad (16)$$

Since each emission source is included by at least one route, and is served by at least one ship on each route, the following inequality holds:

$$\sum_{p \in P(r)} w_{srp} \leq \sum_{s \in S} \sum_{r \in R} \sum_{p \in P(r)} w_{srp}. \quad (17)$$

Combined with (16), we have:

$$\sum_{p \in P(r)} w_{srp} + \sum_{p \in P(r)} e_p \leq E_r. \quad (18)$$

According to Constraint (11), e_p is non-negative for all $p \in P$, and thus we can conclude that: $\sum_{p \in P(r)} e_p \geq 0$. Then, we have $\sum_{p \in P(r)} w_{srp} \leq E_r$. Therefore, we can replace $Q_s n_{sr}$ in the original Constraint (3) with $\min(Q_s n_{sr}, E_r)$. The modified constraints become:

$$\sum_{p \in P(r)} w_{srp} \leq \min(Q_s n_{sr}, E_r), \forall s \in S, \forall r \in R, \quad (19)$$

which completes the proof. $\square$

Based on Lemmas 1–4, we establish the following main theorem, which summarizes the equivalence between the original MIP model and its modified formulation incorporating integrality constraints, linearization, and tightened capacity bounds, while preserving the submodular property of the objective function.

Theorem 1. *The original and modified MIP models are equivalent, and the submodular objective function is preserved under this modification.*

Let P be the original MIP problem with objective function $F_\Omega(x, z, n, e)$ and feasible region Ω . Let $[P]$ be the modified MIP problem with objective function $F'_\Omega(x, z, n, e)$ and feasible region Ω' (defined by (1')–(11)). Then, $[P]$ is equivalent to $[R]$. Specifically, the following applies: (i) Any feasible solution to $[P]$ can be transformed into a feasible solution to $[R]$ with the same objective value. (ii) Any feasible solution to $[R]$ can be transformed into a feasible solution to $[P]$ with the same objective value. (iii) Consequently, an optimal solution to $[P]$ corresponds to an optimal solution to $[R]$, and they share the same optimal objective value.

The reformulated model is presented as follows:

$$[R] \min \sum_{i \in I} C_i x_i + \sum_{s \in S} \sum_{r \in R} C_s^{\text{ship}} z_{sr} + \sum_{s \in S} \sum_{r \in R} T(F_{sr} + B_{sr}) n_{sr} + \sum_{p \in P} T \hat{C} e_p \quad (20)$$

$$\text{s.t.} \quad \sum_{r \in R} z_{sr} \leq 1 \quad \forall s \in S \quad (21)$$

$$\sum_{p \in P(r)} w_{srp} \leq \min(Q_s n_{sr}, E_r) \quad \forall s \in S, \forall r \in R \quad (22)$$

$$\sum_{s \in S} \sum_{r: i(r)=i} \sum_{p \in P(r)} w_{srp} \leq \min(V_i, E_i) \quad \forall i \in I \quad (23)$$

$$\sum_{s \in S} \sum_{r: p \in P(r)} n_{sr} \geq N_p \quad \forall p \in P \quad (24)$$

$$\sum_{s \in S} \sum_{r: p \in P(r)} w_{srp} + e_p = E_p \quad \forall p \in P \quad (25)$$

$$0 \leq n_{sr} \leq N_{sr} z_{sr} \quad \forall s \in S, \forall r \in R \quad (26)$$

$$x_i \in \{0, 1\} \quad \forall i \in I \quad (27)$$

$$z_{sr} \in \{0, 1\} \quad \forall s \in S, \forall r \in R \quad (28)$$

$$w_{srp} \geq 0 \quad \forall s \in S, \forall r \in R, \forall p \in P \quad (29)$$

$$0 \leq e_p \leq E_p^{\max} \quad \forall p \in P \quad (30)$$

4. Experiments

This section conducts computational experiments to verify the effectiveness of our proposed model. Specifically, we implement and solve the MIP model [R], referring to Equations (1)–(11). The experiments were conducted on a desktop computer equipped with 3.40 GHz of 13th Gen Intel Core i7 CPU and 32 GB of RAM, and the MIP model was solved by the Gurobi Optimizer 10.0.1 via the Python 3.11.5 API. In the Gurobi configuration for this model, the MIP gap parameter was explicitly set to 0.1% to specify a relative optimality gap, allowing the solver to terminate when the best feasible solution is within 1% of the optimal bound and balancing solution quality with computational efficiency. Other parameters remained at their default settings, including the utilized number of CPU cores, pre-solve, solver-integrated cuts, and heuristics, and log information. We first set initial values for parameters to obtain basic results. Furthermore, sensitivity analyses were conducted to examine the impact of these parameters.

4.1. Data Collection

The planning horizon is set to 20 years ($T = 20$), with annual calculations and constraints applied consistently to facilitate comprehensive long-term planning and system evaluation. A 600 by 600 (n mile) simulation environment is developed, representing a scale pertinent to establishing a maritime CO₂ transport network and analogous to scenarios connecting dispersed industrial clusters to offshore storage opportunities, such as those considered in European CCS planning involving continental sources and potential North Sea storage locations. Within this environment, ten emission sources and five candidate storage sites are uniformly distributed to reflect a typical regional dispersion of major industrial facilities and potential geological storage formations. Distances between any two locations are measured using the Euclidean metric, with d_{ip} denoting the distance between storage site i and emission source p , and $d_{p_1 p_2}$ indicating the distance between two emission sources p_1 and p_2 . For a route including one emission source and one storage site, the route's sailing distance l_r is calculated as $l_r = d_{ip} + d_{pi} = 2d_{ip}$. For a route including two emission sources and one storage site, the route's sailing distance is calculated as $l_r = d_{ip_1} + d_{p_1 p_2} + d_{p_2 i}$.

The parameters are categorized into three groups:

- Parameters of storage sites. Each candidate storage site is characterized by its annual storage capacity V_i , ranging from 10 to 30 million tons, derived from reported capacities of operational CCS projects across diverse geological contexts, such as those

assessed in European regions including Denmark and the North Sea [49,50]. The construction cost of storage site i is formulated as $C_i = a_i + b_i V_i$. The parameter values are informed by global CO₂ storage initiatives. Specifically, the fixed cost a_i ranges from USD 200 to 600 million, reflecting the investment required for infrastructure development in regions such as the North Sea and Norwegian continental shelf, as documented in global cost assessments and Norwegian CCS projects [50,51]. The variable cost b_i varies between 15 and 40 USD/ton, aligning with reported costs for CO₂ injection into saline aquifers and depleted oil and gas fields, particularly from Norwegian demonstration projects [51]. Specific parameter values for the five candidate storage sites are summarized in Table 4.

Table 4. The parameters for the five candidate storage sites.

Storage Site	V_i (Million Tons)	a_i (USD Million)	b_i (USD/Ton)	C_i (USD Million)
i_1	10	400	40	800
i_2	15	200	35	725
i_3	20	500	25	1000
i_4	25	300	20	800
i_5	30	600	15	1050

- Parameters of emission sources. Each emission source has an annual CO₂ emission volume E_p , ranging from 2.0 to 5.0 million tons, consistent with reported ranges for major industrial emitters such as coal-fired power plants, cement production facilities, and steel manufacturing units in heavy industrial regions globally, including Europe, the U.S., and China [52–54]. The minimum annual berthing requirement N_p is positively correlated with each source's emission volume. The specific values are detailed in Table 5.

Table 5. The parameters for the ten emission sources.

Emission Source	E_p (Million Tons)	N_p
p_1	2.0	5
p_2	2.2	5
p_3	2.5	6
p_4	2.7	6
p_5	3.0	7
p_6	3.5	8
p_7	3.8	9
p_8	4.2	10
p_9	4.7	11
p_{10}	5.0	12

All CO₂ generated by each source is either transported or emitted. The penalty cost for direct atmospheric CO₂ emissions $\hat{C}$ is set as 110 USD/ton, based on the European Union emission regulation guidelines [55], using a conversion rate of EUR 1 = USD 1.1. This standard applies uniformly across all EU ETS-covered installations, including coastal factories and power plants, as confirmed by the European Commission and national implementing agencies [56,57]. Each source is subject to an upper annual direct emission limit that cannot exceed 5% of its annual CO₂ emission amount, i.e., $E_p^{\max} = 0.05 E_p$. This specific threshold, limiting direct emissions from each CO₂ source to 5% of its total annual CO₂ generation, is established to reflect stringent environmental performance expectations and aligns with ambitious objectives, such as those indicated by the International Maritime Organization's 2023 GHG

Strategy aiming for the adoption of near-zero emission technologies. Furthermore, such high capture rates are supported by lifecycle assessments which emphasize the need for substantial emission reductions to ensure the overall environmental integrity and climate benefit of CCS operations [58–60].

- Parameters of ships. A fleet of 40 ships is considered, consisting of 10 small ships, 20 medium ships, and 10 large ships. Based on [61], small ships have transport capacities ranging from 2000 to 5000 tons and incur a total chartering cost of USD 20 million over the 20-year planning horizon. Medium ships carry 8000 to 12,000 tons with a 20-year chartering cost of USD 60 million. Large ships transport 18,000 to 25,000 tons and require a total chartering cost of USD 120 million for the same period. Each berthing operation, whether at a storage site or an emission source, incurs a cost of USD 0.05 million for small ships, USD 0.1 million for medium ships, and USD 0.2 million for large ships, in line with 2025 typical port charges for vessels of varying sizes, adjusted for the specialized requirements of CO₂ transport, and supported by BEIS port fee estimates [62]. All ships use very-low-sulfur fuel oil (VLSFO). Referring to [63], the hourly consumption (kg/h) of ship s sailing at speed v_s is expressed as follows:

$$f_s(v_s) = C_s^0 + C_s^1 \cdot v_s^{n_s}, \quad (31)$$

where n_s is the class-specific speed exponent, which is 3.5 for small ships, 4.0 for medium ships, and 4.5 for large ships. C_s^0 and C_s^1 are coefficients determined by the tonnage of ship s . The parameters for ship classes are presented in Table 6, with C_s^0 and C_s^1 derived from the median of the value ranges provided in [63]. All ships are assumed to operate at a standardized sailing speed of $v_s = 20$ knots. Each year provides $t_0 = 8500$ h effective operational hours after accounting for downtime and maintenance, allowing for a realistic estimation of the maximum number of annual round trips as $N_{sr} = t_0 / (l_r / v_s)$. The fuel cost for ship s completing one full trip on route r is calculated as $F_{sr} = C_f \cdot l_r / v_s \cdot f_s(v_s)$, where C_f is the unit fuel price, set as 0.587 USD/kg based on global bunker prices [64].

Table 6. The parameters for the three class of ships.

Ship Class	C_s^0	C_s^1	n_s	Transport Capacity (Ton)	C_s^{ship} (USD Million)	b_s (USD Million)	Sailing Speed (Knots)
Small	598.65	0.0198	3.5	2000–5000	20	0.05	20
Medium	649.65	0.0040	4.0	8000–12,000	60	0.10	20
Large	600.45	0.0009	4.5	18,000–25,000	120	0.20	20

After establishing the parameter settings, we used these values to derive the basic results, and then we conducted sensitivity analysis to examine the impacts of these parameters.

4.2. Result Interpretation

This section presents the detailed basic results derived from the proposed MIP model, covering storage site construction strategy, capacity utilization of each constructed sites, ship chartering and route assignment strategy, CO₂ transport and emission outcomes, transport operations, and a comprehensive cost analysis.

4.2.1. Storage Site Construction Strategy and Capacity Utilization

In this section, we present the results of the MIP model solved by Gurobi. The results related to the construction strategy are presented in Table 7.

Table 7. The storage sites selected for construction and their capacity utilization.

Storage Site	V_i (Million Tons)	The Amount of CO ₂ Received Annually (Million Tons)	Capacity Utilization (%)
i_1	10	8.9	89.0
i_4	25	24.7	98.8

Table 7 summarizes the selected storage sites for construction and their capacity utilization. Among the five candidate sites, i_1 and i_4 are selected for construction. Site i_1 , with an annual capacity of 10 million tons, receives 8.9 million tons of CO₂ annually, yielding an 89.0% utilization rate. Site i_4 demonstrates even higher efficiency, with an annual intake of 24.7 million tons against a capacity of 25 million tons, achieving a utilization rate of 98.8%. The average utilization rate of the two storage stations is 93.9%, indicating that the constructed sites efficiently match transportation demands.

4.2.2. Ship Route Assignment and Transport Operations

Based on the optimal layout of CO₂ storage sites derived from the MIP model, an analysis is conducted on the ship chartering, routing assignment, and how frequently each chartered ship completes round trips along its designated route per year.

Table 8 presents detailed results regarding ship chartering and route assignments. A total of 12 ships were selected for chartering from the available fleet, consisting of three small-class, seven medium-class, and two large-class ships. Each ship is assigned to a fixed route, connecting one or two emission sources to a designated storage site. Smaller ships are associated with higher values of n_{sr} , indicating more frequent annual operations. This is consistent with their lower transport capacity, which necessitates a greater number of trips to complete the required CO₂ transport. In contrast, larger ships are assigned lower values of n_{sr} , as their greater capacity allows them to meet transport demand with fewer trips. Additionally, the large ship s_{40} is assigned to a route that connects two emission sources. Its high capacity enables it to carry CO₂ from two sources in a single round trip, making it particularly suitable for joint service. This strategy enhances ship utilization and reduces the need for additional ships, further contributing to cost reduction.

Table 8. Charter of ships and route assignment.

Chartered Ship	Ship Class	Transport Capacity (Ton)	Emission Sources Included in the Route	Storage Site Included in the Route	n_{sr}
s_8	Small	4200	p_{10}	i_4	301.4
s_9	Small	4500	p_1	i_4	444.4
s_{10}	Small	5000	p_6	i_4	700.0
s_{24}	Medium	10,800	p_8	i_1	113.0
s_{25}	Medium	11,000	p_2	i_4	200.0
s_{26}	Medium	11,200	p_4	i_4	241.0
s_{27}	Medium	11,400	p_5	i_4	263.0
s_{28}	Medium	11,600	p_8	i_1	256.7
s_{29}	Medium	11,800	p_7	i_4	266.5
s_{30}	Medium	12,000	p_{10}	i_4	311.1
s_{39}	Large	24,200	p_9	i_4	194.2
s_{40}	Large	25,000	p_3, p_7	i_1	126.1

4.2.3. CO₂ Transport and Emission Analysis

Based on the optimized assignment of ships and transport plan, Table 9 presents the detailed allocation of transported CO₂ volumes and direct emissions for each emission

source. Specifically, for each emission source, the table lists the amount of CO₂ transported annually by each assigned ship.

Table 9. Transported and directly emitted CO₂ volumes for each emission source.

Emission Source	Chartered Ship	w_{srp} (Million Tons)	e_p (Million Tons)
p_1	$s = 9$	1.9996	0.0004
p_2	$s = 25$	2.1998	0.0002
p_3	$s = 40$	2.4991	0.0009
p_4	$s = 26$	2.6990	0.0010
p_5	$s = 27$	2.9980	0.0020
p_6	$s = 10$	3.5000	0
p_7	$s = 29$	3.1452	0.0014
	$s = 40$	0.6534	
p_8	$s = 24$	1.2204	0.0019
	$s = 28$	2.9777	
p_9	$s = 39$	4.7000	0
p_{10}	$s = 8$	1.2659	0.0008
	$s = 30$	3.7333	
Sum	\	33.5914	0.0086

As shown in Table 9, 33.5914 million tons of CO₂ are transported annually, with only 0.0086 million tons emitted directly, corresponding to an average emission rate of less than 0.03%. Notably, all CO₂ in emission sources p_6 and p_9 are transported without any direct emissions. The extremely low level of direct emissions can be attributed to the relatively high penalty cost, which guides the system toward sustainable emission-reducing decisions.

4.2.4. Cost Components Analysis

Based on the results derived from the MIP model, Table 10 presents a comprehensive breakdown of the total system cost in the planning horizon.

Table 10. The value and proportion of each component of the cost in the planning horizon.

Cost	Value (USD Million)	Percentage (%)
Total Construction Cost	1600	9.0
Total Chartering Cost	720	4.0
Total Fuel Cost	1137	6.4
Total Berthing Cost	14,354	80.5
Total Penalty Cost for Direct CO ₂ Emissions	19	0.1
Sum	17,830	100

The total cost amounts to USD 17,830 million, comprising expenditures related to storage site construction, ship chartering, fuel consumption, berthing operations, and penalties for direct CO₂ emissions. Among these components, the berthing cost accounts for the largest proportion of the total, reaching USD 14,354 million (80.5%). This high cost is primarily driven by the large number of berthing operations required. The construction cost totals USD 1600 million (9.0%), representing the capital investment needed to ensure adequate capacity of selected storage sites. The fuel cost is USD 1137 million (6.4%), incurred during sailing operations across the network. The chartering cost amounts to USD 720 million (4.0%), corresponding to the long-term lease of twelve ships assigned according to transport demand. The penalty cost for direct CO₂ emissions is only USD 19 million

(0.1%), which reflects the model's strong preference for transporting nearly all CO₂, driven by the relatively high penalty imposed on untransported emissions.

4.3. Sensitivity Analysis

In this section, we further analyze the sensitivity of our MIP model to the changes in input parameters, including the annual available sailing time, the sailing speed of all ships, and the capacities of the candidate storage sites.

The parameters of the experiments are set as shown in Table 11. Generally, we conduct 30 experiments with an experiment ID (EID) indexed from 0 to 29. Each of the experiments is included in the corresponding group with a group ID (GID). G0, G1, and G2 are the groups of experiments that aim to illustrate the performance of the MIP model in solving instances with different annual available sailing time (t_0), the sailing speed of ships (v_s), and the ratio of the actual capacity of each storage station to its original capacity, respectively. The ratio is denoted as α , where the original capacity is V_i , the actual capacity is V_i^{actual} , and $\alpha = V_i^{\text{actual}} / V_i$. In the " t_0 (h)", " v_s (knots)", and " α " columns, $[a, b, c]$ represents a list of numbers generated from a to b with a step size of c .

Table 11. Experiment parameter settings.

GID	EID	t_0 (h)	v_s (Knots)	α
G0	0–11	[3000, 8500, 500]	20	1
G1	12–22	8500	[10, 30, 2]	1
G2	23–29	8500	20	[0.4, 1.6, 0.2]

4.3.1. Different Annual Available Sailing Time

To investigate the impact of varying the annual available sailing time on the model's effectiveness, we design instances by altering the annual available sailing time from 3000 h to 8500 h while keeping the sailing speed ($v_s = 20$) and the capacity ratio ($\alpha = 1$) unchanged.

Figure 1 shows the optimal results of the MIP model. The x -axis represents the EID, where instances are arranged in increasing order of the annual available sailing time. The y -axis corresponds to the values of indicators, including the objective of an optimal solution (Obj), the number of storage sites constructed (SSN) and the construction cost (SSC) in the planning horizon, the average utilization rate (AUR) of all storage sites, the number of ships chartered (SCN) and the chartering cost (SCC), the amount of CO₂ emitted directly to the atmosphere (CED) in the planning horizon, and the total transportation cost (TTC) in the planning horizon and the average annual number of round trips (TRN) of all ships. Table 12 lists the explanation, calculation, and practical implication of each indicator.

Figure 1a indicates that extending the annual available sailing time from 3000 to 8500 h leads to a steady decline in Obj, before stabilizing around 8000 h. This trend is primarily driven by reductions in SSC and SCC. This observed decrease in Obj with increased sailing time reflects real-world logistics, where greater operational availability of assets typically enhances system efficiency and reduces overall costs. The model's ability to capture this fundamental relationship supports its practical applicability. A longer sailing time allows each ship to complete more round trips per year, thereby reducing the number of ships needed to meet the transport demand. Therefore, as the annual available sailing time increases, SCN and SCC are reduced, as shown in Figure 1d. This reduction in SCN and SCC is a consequence of improved fleet utilization, a primary objective in capital-intensive transport operations, thereby aligning the model with practical fleet management goals.

As shown in Figure 1b, SSN remains constant at two across all instances due to fixed CO₂ generation volumes and unchanged site capacities. However, SSC shows notable fluctuations, suggesting that although the number of constructed sites remains unchanged,

the specific site selection varies. This variation suggests that a longer annual available sailing time enables the model to reconfigure the spatial allocation of selected sites to better align with optimal-cost transport flows. This dynamic site selection, despite a constant number of sites, demonstrates the model's robustness and sophistication in adapting infrastructure choices to optimize the entire system as operational conditions change, rather than rigidly adhering to a single site configuration. Such adaptability is crucial for real-world strategic planning where operational parameters can vary.

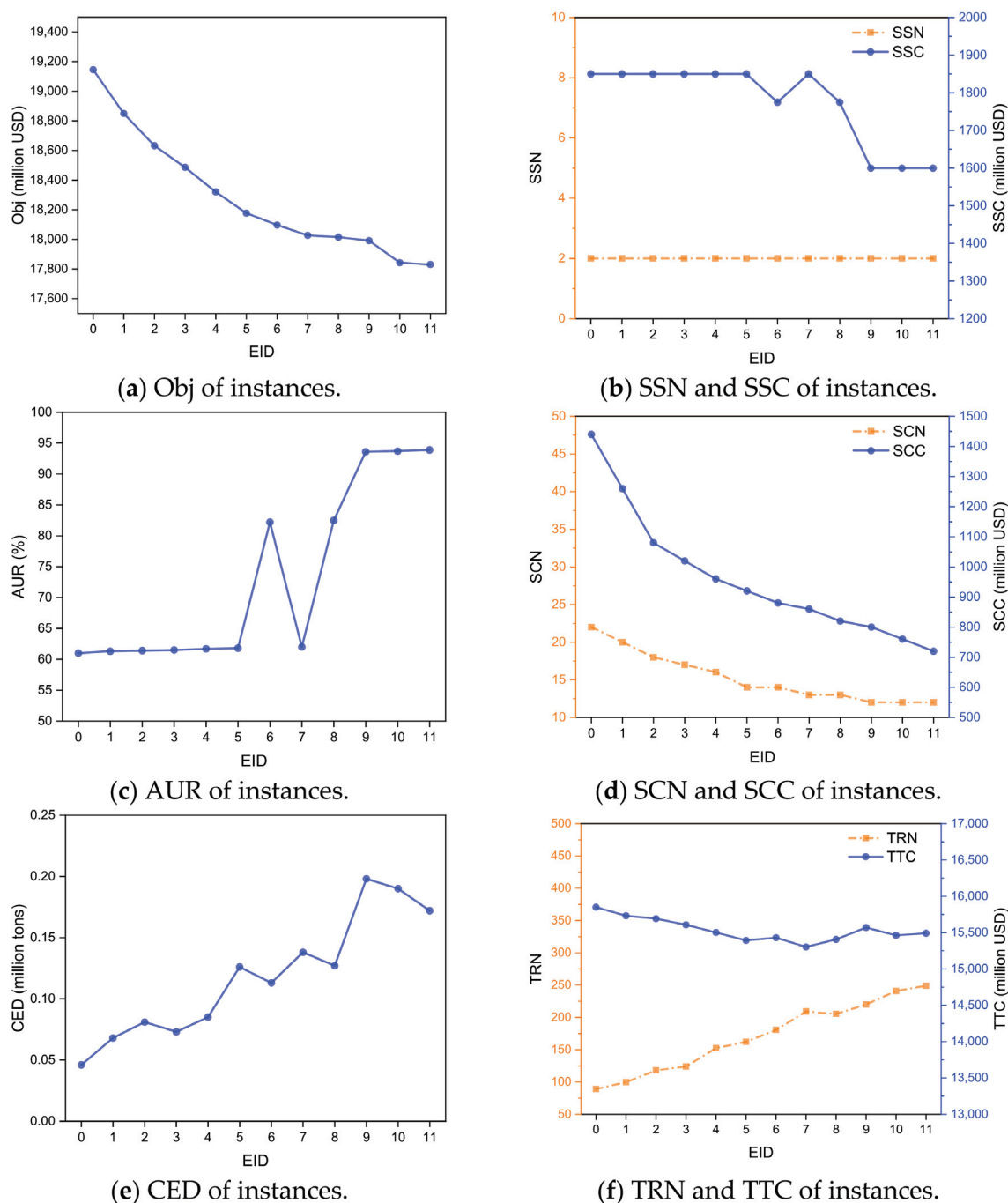


Figure 1. Optimal results of the MIP model with different annual available sailing time (EID: 0–11).

Table 12. Key performance indicators.

Indicator	Explanation	Calculation	Practical Implications
Obj	The total system cost in the planning horizon.	Equation (1): Output of the MIP model's objective function.	Primary measure of overall economic efficiency and the success of the optimization strategy.
SSN	The total count of CO ₂ storage sites selected for construction from the candidate locations.	$\sum_{i \in I} x_i$	Indicates the scale of new storage infrastructure required; impacts capital investment and land use.
SSC	The total cost for constructing all the selected storage sites.	$\sum_{i \in I} C_i x_i$	Quantifies the capital investment in storage facilities; a major component of system costs.
AUR	Average of annual capacity utilization of all constructed CO ₂ storage sites.	$\sum_{i \in I} \frac{\sum_{s \in S} \sum_{r: i(r)=i} \sum_{p \in P(r)} w_{srp}}{V_i} / \sum_{i \in I} x_i$	Measures the efficiency of capital investment in storage capacity.
SCN	The total count of ships chartered from the available fleet to transport CO ₂ .	$\sum_{s \in S} \sum_{r \in R} z_{sr}$	Reflects the size of the operational fleet required; influences overall chartering costs and operational complexity.
SCC	The total cost chartering all selected ships.	$\sum_{s \in S} \sum_{r \in R} C_s^{\text{ship}} z_{sr}$	A major component of system costs.
CED	The total amount of CO ₂ emitted directly to the atmosphere in the planning horizon.	$\sum_{p \in P} T e_p$	Direct measure of the system's environmental performance and failure to meet capture targets; incurs financial penalties.
TTC	The total transportation cost (including fuel cost and berthing cost) in the planning horizon.	$\sum_{s \in S} \sum_{r \in R} T(F_{sr} + B_{sr}) n_{sr}$	Reflects the cost for transporting CO ₂ from emission sources to storage sites.
TRN	The average number of round trips completed per year by each chartered ship.	$\frac{\sum_{s \in S} \sum_{r \in R} n_{sr}}{\sum_{s \in S} \sum_{r \in R} z_{sr}}$	Indicates the operational intensity and utilization efficiency of the chartered fleet.

Figure 1c shows that AUR does not exhibit a monotonic trend with respect to sailing time but instead shows distinct fluctuations across different EIDs. It remains around 61% under shorter sailing times (EID 0–5), rises sharply to over 80% at EID 6, drops back at EID 7, and subsequently rises to a peak from EID 9 onward. This pattern suggests that the variation in AUR is not directly driven by sailing time but is primarily influenced by the capacities of the selected storage sites. The model's ability to reflect how AUR is more closely tied to the specific characteristics of the chosen storage sites, rather than solely with sailing time, which indicates its proficiency in handling complex interactions among different system components, contributing to robust decision support.

Figure 1e illustrates that CED exhibits an overall increasing trend as sailing time increases, which is closely related to the reduction in SCN. After assigning the majority of CO₂ to ships, a small amount of residual CO₂ may remain at certain sources. Whether these residual volumes are transported depends on a cost comparison: only if the transportation cost for the remaining portion is lower than the corresponding emission penalty will the model choose to transport it. Otherwise, direct emission becomes the cost-minimizing option. As the sailing time increases, each ship can complete more round trips, allowing the model to reduce SCN significantly. Although this improves transport efficiency and

reduces SCC, having fewer chartered ships makes it more challenging to allocate ships to specific routes, thereby making it harder for the model to find a ship to transport the remaining residual CO₂ on a route at a cost lower than the direct emission, especially for routes with long distances that incur high fuel costs, or that include two emission sources, leading to high berthing costs. As a result, the model tends to opt for direct emissions more, contributing to the overall rising trend in CED, despite some fluctuations due to changing storage site configurations and chartered ships. This observed increase in CED with reduced SCN highlights a critical real-world trade-off where optimizing for lower fleet and chartering costs can lead to increased direct emissions if transporting small, difficult-to-reach CO₂ volumes becomes economically unviable compared to incurring an emission penalty. The model's capability to explicitly capture and make decisions based on this economic trade-off between transport cost and emission penalty is a strong indicator of its real-world applicability and its robustness in finding a truly minimal-cost solution considering all factors.

Figure 1f shows that TRN increases steadily with sailing time. Given that the average route length across the system does not change substantially, an extended sailing time naturally allows each ship to complete more round trips within a year. As a result, the system-level TRN rises accordingly. In contrast, TTC remains relatively stable, exhibiting only minor fluctuations. This indicates that despite the increase in TRN, the total number of transport operations and the cumulative sailing distance do not grow significantly. Since TTC is jointly influenced by SCN and TRN, the observed stability reflects a compensatory relationship: as the available sailing time increases, SCN decreases while TRN increases. The opposing movement of these two factors offsets each other, leading to limited variation in overall transportation costs. The model's demonstration of stable TTC, despite significant changes in SCN and TRN, showcases its ability to correctly balance interconnected cost factors. This realistic interplay, where increased efficiency per asset (higher TRN) allows for fewer assets (lower SCN) while maintaining overall transport activity levels, is a testament to the model's robust cost accounting and its practical relevance for comprehensive system evaluation.

Overall, extending the annual available sailing time reduces both SCC and SSC, thereby lowering Obj. However, reduced SCN results in increased CED due to difficulty in assigning ships on routes with low transportation costs for residual CO₂ volumes. The results highlight that the model minimizes Obj by jointly optimizing SSC, SCC, TTC and CED.

Table 13 shows the optimal results of the model under varying annual available sailing time, ranging from 3000 to 8500 h. As t_0 increases, Obj decreases steadily before stabilizing, indicating that a longer sailing time improves overall cost efficiency by enabling each ship to complete more round trips. Most performance indicators change with increasing sailing time. In contrast, SSN remains constant at two across all instances. This suggests that the initial two storage sites are sufficient to meet the transport demand across all levels of annual available sailing time. Attempting to build a third site would result in diminishing returns, where the additional savings in transport and emission costs would no longer offset the increased construction cost. This phenomenon aligns with Lemma 3, which states that the objective function is submodular with respect to the number of constructed storage sites. The submodular property implies that the marginal benefit of adding a storage site decreases as more sites are introduced. In this case, the first two sites contribute significantly to reducing the transportation cost, but beyond a certain point, additional sites would not provide more benefits.

Table 13. Optimal results of the MIP model with different annual available sailing time (EID: 0–11).

EID	t_0 (h)	Obj (USD Million)	SSN	SSC (USD Million)	AUR (%)	SCN	SCC (USD Million)	CED (Million Tons)	TRN	TTC (USD Million)	CPU Time (s)
0–11	3000	19,145	2	1850	61.0	22	1440	0.046	88.9	15,850	45.4
	3500	18,849	2	1850	61.3	20	1260	0.068	99.6	15,732	71.1
	4000	18,632	2	1850	61.4	18	1080	0.081	117.9	15,693	57.5
	4500	18,486	2	1850	61.5	17	1020	0.073	123.9	15,608	160.2
	5000	18,320	2	1850	61.7	16	960	0.085	152.3	15,501	95.3
	5500	18,177	2	1850	61.8	14	920	0.126	162.1	15,393	57.9
	6000	18,097	2	1775	82.2	14	880	0.113	180.5	15,430	52.2
	6500	18,027	2	1850	62.0	13	860	0.138	209.2	15,302	52.0
	7000	18,015	2	1775	82.5	13	820	0.127	205.3	15,406	41.8
	7500	17,992	2	1600	93.6	12	800	0.198	219.8	15,570	98.1
	8000	17,843	2	1600	93.7	12	760	0.190	240.9	15,462	58.8
	8500	17,830	2	1600	93.9	12	720	0.172	248.8	15,491	67.1
Avg	5750	18,285	2	1775	73.0	15	960	0.118	170.8	15,537	71.4

4.3.2. Different Sailing Speed

To investigate the impact of varying the sailing speed of all ships on the model's effectiveness, we design instances by altering the sailing speed from 10 knots to 30 knots while keeping the annual available sailing time ($t_0 = 8500$) and the capacity ratio ($\alpha = 1$) unchanged. Figure 2 shows the optimal results of the MIP model.

As shown in Figure 2a, Obj exhibits a U-shaped pattern as the sailing speed increases. At low speeds, more ships are needed to meet the transport demand, leading to high SCN and SCC. As speed increases, each ship completes more round trips, reducing SCN and SCC, as shown in Figure 2d. However, beyond a threshold, these values stabilize due to the routing constraint that limits each ship to serving at most two emission sources. This U-shaped Obj curve is characteristic of real-world speed optimization problems across various transport modes. The model demonstrates this practical reality where initial speed increases yield efficiency gains (reduced SCN/SCC), but excessive speed leads to disproportionately higher operational costs (primarily fuel), rendering the system less economical. Capturing this established trade-off is crucial for the model's credibility and applicability in determining optimal operational strategies.

According to Figure 2b, SSN remains unchanged across all instances. SSC also stays nearly constant, except for a drop at EID 17, which reflects a change in the selected storage site configuration. This change in configuration concurrently affects AUR, which is depicted clearly in Figure 2c, since each storage site differs in capacity.

Furthermore, Figure 2e illustrates a steady increase in CED with rising sailing speeds. As SCN decreases, routing flexibility becomes more limited. With fewer ships available, it becomes more difficult for the model to find a route for transporting the remaining CO₂ at a cost lower than the direct emission penalty. In addition, higher sailing speeds lead to increased fuel costs on the same routes, making it even harder to economically allocate these remaining residual volumes. As a result, the model opts for direct emissions more frequently, leading to an overall rise in CED. The model's nuanced handling of CED under varying speeds, attributing its increase to both reduced SCN (limiting routing options) and higher fuel costs per trip (making transport of residual CO₂ more expensive), demonstrates a robust understanding of the interconnected factors influencing emission decisions. This ability to weigh multiple cost drivers is vital for practical environmental and economic assessments.

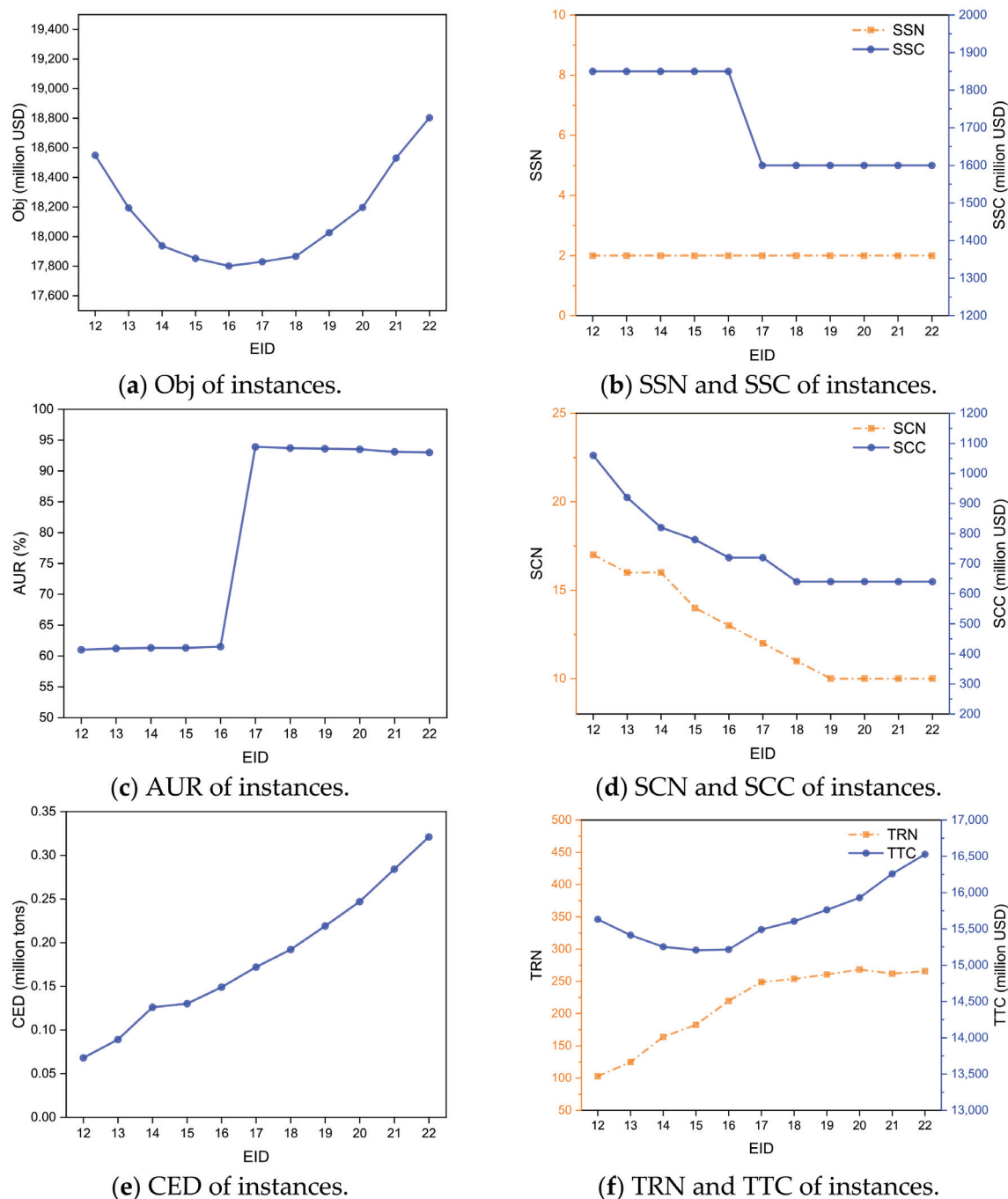


Figure 2. Optimal results of the MIP model with different sailing speeds (EID: 12–22).

Figure 2f shows that TRN increases with sailing speed since each trip takes less time, allowing ships to complete more round trips within the same annual available sailing time. However, when the sailing speed exceeds a certain threshold, SCN remains constant, and TRN also stabilizes. TTC does not follow a strictly increasing pattern: it decreases slightly at EID 12–16 due to changes in ship chartering, as ships with lower fuel and berthing costs are preferred. At EID 17–22, however, according to Equation (20), the high exponent values (3.5, 4.0, and 4.5 for the different ship classes in this study) cause disproportionately large increases in fuel consumption, resulting in a clear upward trend in TTC. The model's sensitivity to the exponential relationship between speed and fuel consumption contributes to its real-world applicability as fuel costs are a dominant factor in maritime operations. The clear upward trend in TTC at higher speeds, directly linked to this exponential fuel use,

realistically reflects the severe economic penalties of excessive speed, thereby validating the model's cost calculations and its ability to identify a practically optimal speed range.

Overall, increasing sailing speed reduces SCN and SCC by allowing each ship to complete more round trips within a year, which helps lower Obj. However, as speed continues to rise, fuel consumption rises significantly, causing TTC to increase and ultimately driving Obj back up. In addition, the reduction in SCN leads to higher CED due to limited routing flexibility. The results highlight that the model minimizes Obj by jointly optimizing SSC, SCC, TTC and CED under varying sailing speeds.

Table 14 presents the optimal results under different sailing speeds from 10 to 30 knots. Obj first decreases and then increases as speed rises, with accompanying changes in SCN, SCC, CED, TRN, and TTC. A shift in SSC and AUR reflects changes in storage site selection. SSN remains constant at two across all instances, which suggests that constructing two storage sites is sufficient to meet the transport demand and achieve significant total cost savings.

Table 14. Optimal results of the MIP model with different sailing speed (EID: 12–22).

EID	v_s (Knots)	Obj (USD Million)	SSN	SSC (USD Million)	AUR (%)	SCN	SCC (USD Million)	CED (Million Tons)	TRN	TTC (USD Million)	CPU Time (s)
12–22	10	18,549	2	1850	61.0	17	1060	0.068	102.8	15,632	76.0
	12	18,193	2	1850	61.2	16	920	0.089	124.7	15,413	69.8
	14	17,937	2	1850	61.3	16	820	0.126	163.8	15,253	57.3
	16	17,852	2	1850	61.3	14	780	0.130	182.5	15,208	81.5
	18	17,801	2	1850	61.5	13	720	0.149	219.7	15,215	52.8
	20	17,830	2	1600	93.9	12	720	0.172	248.8	15,491	78.0
	22	17,866	2	1600	93.7	11	640	0.192	253.7	15,605	43.4
	24	18,026	2	1600	93.6	10	640	0.219	260.6	15,762	62.2
	26	18,196	2	1600	93.5	10	640	0.247	268.2	15,929	58.0
	28	18,530	2	1600	93.1	10	640	0.284	262.0	16,259	34.4
	30	18,803	2	1600	93.0	10	640	0.321	265.8	16,528	34.0
Avg	20	18,144	2	1714	78.8	13	747	0.182	213.9	15,663	58.9

4.3.3. Different Capacity Ratio of Storage Sites

To investigate the impact of varying the capacity ratio of all storage sites on the model's effectiveness, we design instances by altering the capacity ratio from 0.4 to 1.6 while keeping the annual available sailing time ($t_0 = 8500$) and the sailing speed ($v_s = 20$) unchanged. Figure 3 shows the optimal results of the MIP model.

As shown in Figure 3a, Obj exhibits an overall decreasing trend as the capacity ratio increases. This is primarily due to the reduction in SSC. With a larger capacity ratio, fewer sites are needed to satisfy the CO₂ storage requirement. As a result, both SSN and SSC decrease as the capacity increases, and eventually only one storage site is selected, as reflected in Figure 3b. However, since higher-capacity sites are more expensive to construct, there are less pronounced decreases or even slight increases in SSC in some cases. The model's strategic response of reducing SSN as individual site capacities increase is a logical adaptation to changing infrastructure parameters. This ability to make high-level decisions on the scale and number of facilities based on their individual characteristics demonstrates its utility for long-term investment planning in real-world CCS projects. Furthermore, the nuanced behavior where SSC might not always decrease monotonically (if larger sites have a disproportionately higher unit construction cost) indicates the model balances economies of scale against potential cost premiums for larger infrastructure, a common real-world consideration.

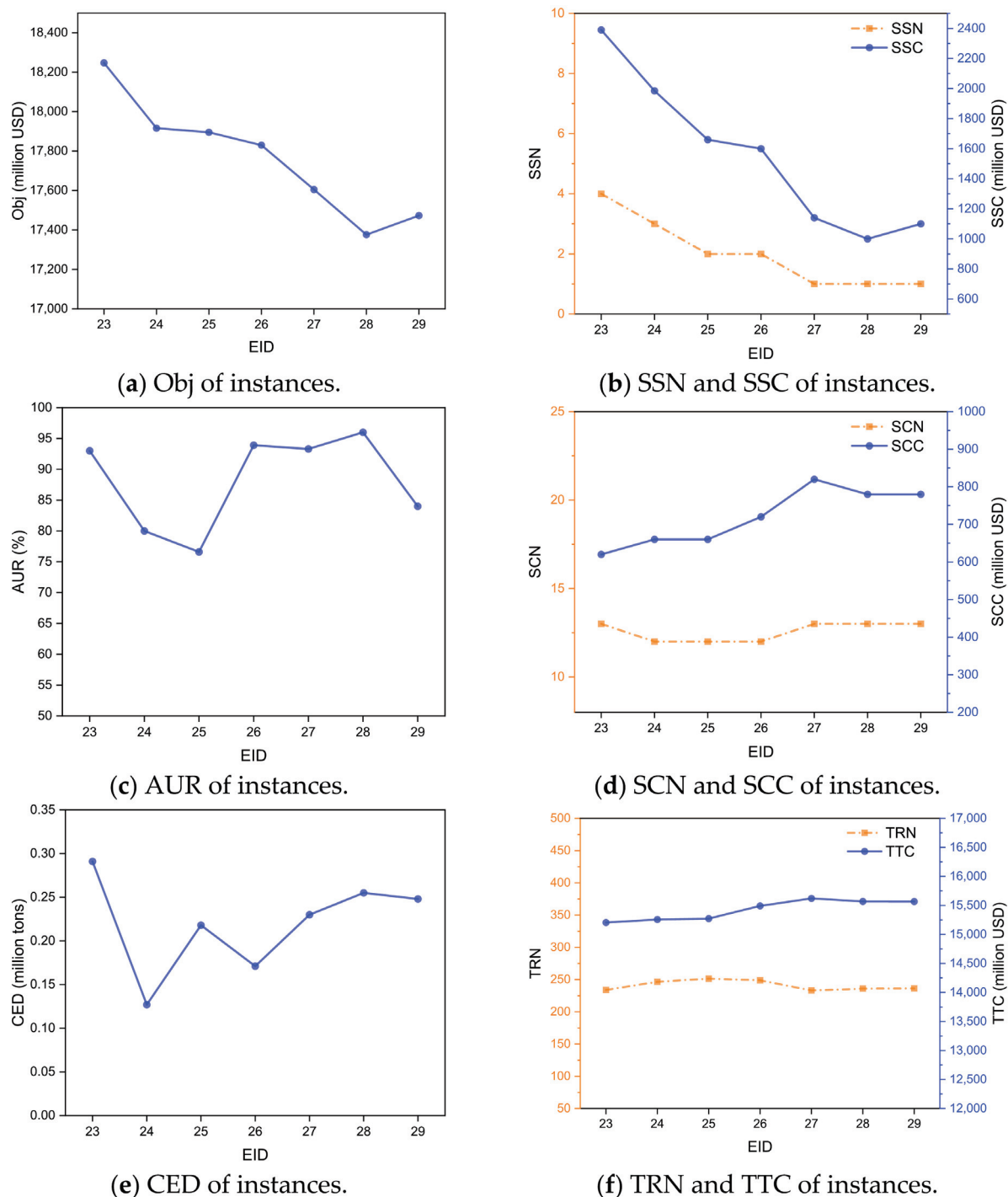


Figure 3. Optimal results of the MIP model with different capacity ratios (EID: 23–29).

Figure 3c shows that AUR fluctuates considerably across EIDs, reflecting differences in the number, location, and capacity of selected storage sites. These factors significantly influence the level of storage utilization under each scenario.

As shown in Figure 3d,f, SCN, SCC, TRN, and TTC remain relatively stable. Given that the sailing time, sailing speed, and annual CO₂ generation are fixed, transport operations do not vary significantly. The slight fluctuations in TTC are mainly driven by changes in routing distance under different storage site configurations. The observed stability in transport-related costs (SCN, SCC, TRN, TTC) when only site capacity is varied (while operational parameters like sailing time and speed are held constant) effectively demonstrates the model's capability to isolate and analyze the impact of specific infrastructure

decisions. This feature significantly enhances the capability for comprehensive scenario analysis in practical planning, enabling decision-makers to systematically evaluate the individual impacts of infrastructure versus operational changes.

Figure 3e highlights noticeable variations in CED, primarily due to differences in storage site construction strategies and routing assignments. Although ship chartering decisions vary only slightly, they still influence the routing flexibility and transport capacity, thereby affecting CED. The model's indication that CED varies based on storage site strategies, even with stable overall transport operations, emphasizes the interconnect-edness of infrastructure choices and their downstream environmental consequences via routing adjustments. This highlights the model's comprehensive view, which is crucial for integrated system planning.

Overall, when varying the capacity ratio, the variation in Obj is mainly driven by changes in SSC. Other cost components remain relatively stable, and CED variation has a limited impact on total cost. These results demonstrate the model's ability to balance construction, chartering, transport, and emission costs to achieve the minimum system cost.

Table 15 shows the numerical results under varying capacity ratios. Obj gradually decreases as the capacity increases, primarily due to the reduction in SSC, as SSN drops from four to one. AUR varies across EIDs, consistent with differences in the number and location of selected sites. Other indicators, including SCN, SCC, TRN, and TTC, remain relatively stable due to fixed sailing time and speed. CED exhibits relatively noticeable variation, reflecting the influence of storage site and routing strategies, although its absolute values remain small across all cases.

Table 15. Optimal results of the MIP model with different capacity ratios (EID: 23–29).

EID	α	Obj (USD Million)	SSN	SSC (USD Million)	AUR (%)	SCN	SCC (USD Million)	CED (Million Tons)	TRN	TTC (USD Million)	CPU Time (s)
23–29	0.4	18,247	4	2390	93.0	13	620	0.291	233.9	15,205	24.8
	0.6	17,916	3	1985	80.0	12	660	0.127	246.6	15,257	33.7
	0.8	17,895	2	1660	76.6	12	660	0.218	251.4	15,551	37.2
	1.0	17,830	2	1600	93.9	12	720	0.172	248.8	15,491	56.4
	1.2	17,605	1	1140	93.3	13	820	0.230	233.1	15,620	54.6
	1.4	17,376	1	1000	96.0	13	780	0.255	235.9	15,568	34.4
	1.6	17,473	1	1100	84.0	13	780	0.248	236.1	15,566	35.6
Avg	1.0	17,763	2	1554	88.1	13	720	0.220	240.8	15,465	39.5

To sum up, the numerical experiments, encompassing variations in annual available sailing time, sailing speed, and storage site capacity ratio (as detailed in Tables 13–15), demonstrate that the model not only effectively responds to changes in key parameters by adjusting storage site construction, ship chartering, and route assignment strategies, but also does so in a manner that reflects real-world operational logic and economic trade-offs. The results reveal how different operational and infrastructure configurations influence system-wide outcomes, particularly the Obj and its major components—SSC, SCC, and TTC. Furthermore, these analyses highlight the model's robustness in navigating complex interactions, such as balancing fleet efficiency against emission targets, or optimizing sailing speed against escalating fuel costs. The model's ability to identify these critical thresholds and trade-offs, leading to holistic, minimal-cost solutions that consider strategic infrastructure investments, operational efficiencies, and environmental implications simultaneously, supports its practical applicability as a decision-support tool for designing and managing sustainable maritime CO₂ transport systems.

The computational efficiency of the proposed MIP model was assessed by recording the CPU time required to solve each scenario in the sensitivity analyses. As summarized in Tables 13–15, the average solution times across the three sets of sensitivity analyses were 71.4 s, 58.9 s, and 39.5 s, respectively. These findings suggest that the model is computationally tractable and can be effectively solved within reasonable time, thereby confirming its practical computational efficiency.

5. Two-Stage Optimization Framework and Solution Approach

In practical scenarios, employing the comprehensive MIP model from Section 3 Problem Formulation can present significant challenges. This is primarily because the model, with its numerous binary and continuous variables and complex operational constraints governing site selection, ship assignment, and routing, inherently belongs to the class of NP-hard optimization problems. For such problems, the computational effort to find a guaranteed optimal solution can grow exponentially with network scale, such as the number of sources, sites, ships, and routes. This inherent complexity means that for large-scale, real-world instances, achieving exact optimal solutions with the integrated MIP model alone can become computationally prohibitive within practical timeframes, despite its comprehensive nature. Moreover, applying such a complex model often requires considerable mathematical optimization expertise. Therefore, to address these issues, particularly the potential computational intractability for larger systems due to its NP-hard nature and the need for a more accessible decision framework, a simplified two-stage modeling approach is proposed. This approach still addresses CO₂ storage site construction, ship chartering, route assignment, and transport operation planning by decomposing the problem into strategic site selection and tactical ship assignment and routing stages. The goal is to provide more tractable and adaptable solutions, especially for real-world applications where scalability and ease of use are critical.

The reasons for employing a two-stage approach are as follows:

1. Firstly, significantly reduced computational time for large-scale instances: While an MIP model provides optimality, its solution time can become excessively long for complex, real-world scenarios characteristic of NP-hard problems. Heuristic approaches, like the GA used in our second stage, are designed to find high-quality solutions much more rapidly, making them viable for larger problems or when quick decision-making is paramount.
2. Secondly, enhanced accessibility and reduced reliance on expensive commercial solvers: Our two-stage approach is structured such that the first stage (site selection) can potentially be addressed using simpler, perhaps even greedy, methods to obtain good initial configurations, or other less computationally intensive optimization techniques. The second stage (ship assignment and routing) then employs a genetic algorithm. Crucially, this entire two-stage framework can be implemented and executed without necessitating a commercial MIP solver like Gurobi. This is a significant practical advantage for many organizations, particularly smaller companies or those in regions with limited access to such software, for whom the substantial cost of commercial solver licenses (which can be tens or even hundreds of thousands of dollars) is a prohibitive barrier. Open-source solvers, while available, often exhibit significantly lower computational efficiency for complex MIPs compared to their commercial counterparts. Our two-stage heuristic approach therefore offers a cost-effective and more broadly accessible pathway to sophisticated decision support.

5.1. Two-Stage Modeling Approach

In this section, we describe the structure of the proposed two-stage modeling approach, which aims to simplify the original MIP model by decoupling strategic infrastructure decisions from operational transport planning. The first stage determines the optimal selection of CO₂ storage sites to ensure sufficient total capacity, while the second stage focuses on optimizing ship routing and transport operations given the selected sites.

5.1.1. First Stage: Strategic Site Construction

The first stage is formulated as a binary optimization model. It identifies which candidate storage sites should be constructed to minimize total construction costs while ensuring that the selected sites together offer enough capacity to store all generated CO₂, as shown below:

$$\min \quad \sum_{i \in I} C_i x_i \quad (32)$$

$$\text{s.t.} \quad \sum_{p \in P} E_p \leq \sum_{i \in I} V_i x_i \quad (33)$$

$$x_i \in \{0, 1\} \quad \forall i \in I \quad (34)$$

The objective function (32) minimizes the construction cost of selected CO₂ storage sites. Constraint (33) ensures the total amount of CO₂ generated annually by all emission sources does not exceed the total capacity of all selected storage sites. Constraint (34) is the domain of decision variables.

The storage site construction decisions determined in this stage are denoted by the set I^* , representing the subset of candidate locations I that are selected for construction, i.e., $I^* \subseteq I$, $x_i = 1, \forall i \in I^*$. Notably, only the storage sites $i \in I^*$ will be considered in the second stage.

Given the parameters of emission sources and candidate storage sites, the first-stage model offers a straightforward yet effective framework for identifying cost-efficient CO₂ storage site combinations. As a binary optimization problem with limited decision variables, it is computationally lightweighted and can be solved rapidly, even in large-scale scenarios. Moreover, the decision logic—selecting a subset of sites to satisfy total storage demand while minimizing cost—is highly intuitive and easily interpretable by industry practitioners. This makes the model particularly well suited for use in early-stage planning, feasibility studies, or stakeholder discussions where transparency and ease of implementation are critical. The selected sites from this stage form the foundation for the second-stage transport assignment model, which determines the ship assignments and routing strategies accordingly.

5.1.2. Second Stage: Operational Ship Assignment and Route Planning

Given the selected storage sites, this stage optimizes the chartered ships, assignment of routes, number of annual round trips, and transport volumes, with the objective of minimizing total tactical costs including chartering cost, transportation cost, and direct emission penalties. The model in this stage can be written as follows:

$$\min \quad \sum_{s \in S} \sum_{r \in R} C_s^{ship} z_{sr} + \sum_{s \in S} \sum_{r \in R} T(F_{sr} + B_{sr}) n_{sr} + \sum_{p \in P} T \hat{C} e_p \quad (35)$$

$$\text{s.t.} \quad \sum_{r \in R} z_{sr} \leq 1 \quad \forall s \in S \quad (36)$$

$$\sum_{p \in P(r)} w_{srp} \leq \min(Q_s n_{sr}, E_r) \quad \forall s \in S, \forall r \in R \quad (37)$$

$$\sum_{s \in S} \sum_{r: i(r)=i} \sum_{p \in P(r)} w_{srp} \leq \min(V_i, E_i) \quad \forall i \in I^* \quad (38)$$

$$\sum_{s \in S} \sum_{r: p \in P(r)} n_{sr} \geq N_p \quad \forall p \in P \quad (39)$$

$$\sum_{s \in S} \sum_{r: p \in P(r)} w_{srp} + e_p = E_p \quad \forall p \in P \quad (40)$$

$$0 \leq n_{sr} \leq N_{sr} z_{sr} \quad \forall s \in S, \forall r \in R \quad (41)$$

$$z_{sr} \in \{0, 1\} \quad \forall s \in S, \forall r \in R \quad (42)$$

$$w_{srp} \geq 0 \quad \forall s \in S, \forall r \in R, \forall p \in P \quad (43)$$

$$0 \leq e_p \leq E_p^{\max} \quad \forall p \in P \quad (44)$$

The objective function (35) minimizes the total cost, which consists of three components: the total chartering cost of all chartered ships, the transportation cost in the planning horizon (including fuel and berthing costs), and the penalty cost for direct CO₂ emissions into the atmosphere. Constraint (36) ensures that each ship can be assigned to at most one route, reflecting the condition that not all ships must be chartered. Constraint (37) denotes that, for each ship assigned to a route, the total amount of CO₂ transported per year does not exceed the product of its transport capacity and the number of annual round trips. Constraint (38) requires that the total injected CO₂ into each constructed storage site per year does not exceed its capacity. Constraint (39) requires that each emission source is visited by ships a sufficient number of times per year to meet its minimum berthing requirement, thereby ensuring an adequate transport frequency and enabling timely removal of the generated CO₂. Constraint (40) ensures that all CO₂ generated by each source is either transported or emitted. Constraint (41) imposes operational bounds on the annual number of round trips that each chartered ship can make on its assigned route, ensuring that this value does not exceed the maximum allowable trips based on ship speed and route distance. Constraints (42)–(44) are the domains of decision variables. Constraints (36)–(44) in this model can be viewed as a reduced and reformulated subset of the constraints originally presented in the integrated MIP model, specifically corresponding to Constraints (2)–(7) and (9)–(11). Notably, since the site selection decision is fixed in the first stage, the original storage capacity constraint (4) is reduced to (38), which only enforces capacity compliance at the predetermined sites. This reformulation leads to a more compact and computationally tractable model, facilitating faster solution times and improved scalability. Moreover, the reduced model structure enhances practical interpretability, making it more suitable for real-world applications where decision-makers require operational guidance under established infrastructure conditions.

5.2. Solution Approach and Experiment Results

In this section, we introduce a heuristic solution method to address the second stage of the problem and present the corresponding computational results for CO₂ transport assignment. Given the selected storage sites from the first stage, we adopt a GA to solve the CO₂ transport assignment problem. The GA provides a metaheuristic-based alternative to exact methods, enabling near-optimal solutions under complex constraints.

5.2.1. Genetic Algorithm Design and Parameter Settings for Second Stage Model

In this section, we develop a heuristic algorithm to solve the second-stage CO₂ transportation planning problem, which focuses on assigning ships to feasible routes and

scheduling annual round trips based on the storage site construction outcomes from the first stage. This problem exhibits a complex structure characterized by binary decisions for ship chartering and ship-route assignments, continuous variables for round-trip frequencies, and additional continuous decisions related to the direct emission of untransported CO₂. These decisions are subject to multiple constraints, including storage capacity limits, ship transport capacity limits, ship operation boundaries, and direct emission limits. As the problem size scales with the number of ships and routes, the exact solution through an MIP model is computationally intractable. Therefore, we adopt a metaheuristic approach to efficiently explore the solution space while still generating feasible and interpretable results.

To this end, we design a GA tailored to the characteristics of the problem. In the GA framework, each individual (chromosome) encodes a complete assignment of ships to routes and their corresponding round-trip frequencies, ensuring that only the routes connecting to selected storage sites from the first stage are considered. Each gene in the chromosome represents a specific ship and consists of three key components: the assigned route (comprising one storage site and one or more emission sources), and the number of annual round trips n_{sr} . The assigned route also indicates whether the ship is chartered. If no route is specified, the ship is considered not chartered. The transported volume is computed based on the ship's capacity and its trip frequency and directly influences the amount of untransported CO₂, which is represented as direct emission e_p at each source. This gene structure enables the algorithm to jointly optimize ship chartering, routing, and scheduling decisions while satisfying constraints on transport capacity, storage capacity, source berthing frequency, and CO₂ emission volumes.

To initialize the population, a hybrid seeding approach is adopted to generate both feasible and cost-efficient individuals. First, for each emission source, the algorithm constructs a candidate route pool by ranking all feasible single-source and dual-source routes to each selected storage site in ascending order of expected cost, which includes fuel cost, berthing cost, and penalties for direct CO₂ emission based on maximum ship frequency. Then, the ships are sorted by transport capacity and greedily assigned to the lowest-cost routes that still have available storage capacity and unserved emission demand. For each assignment, the number of round trips is determined by the lesser of the ship's annual upper bound and the amount required to meet the remaining transport volume on that route. After the initial assignment, a repair and reallocation process is applied to resolve infeasibilities. If any emission source still has unmet transport demand, the algorithm either activates an unchartered ship with adequate capacity or reassigns a ship from an over-served source to the under-served one. Reassignment is accepted only if it leads to a reduction in total tactical cost, accounting for fuel and berthing cost increases, changes in emission penalties, and possible improvements in storage utilization. A feasibility checker is applied to ensure that all individuals comply with constraints on source berthing frequency, round trip number limits, ship transport capacity, and storage site capacity. If any constraint is violated, targeted adjustments such as reducing the round-trip frequency or switching to a nearby low-cost alternative route are applied until the individual becomes feasible. To maintain diversity, 50% of the population is generated by perturbing the greedy solution—shuffling ship orders, adjusting round trips, or replacing routes with alternatives of comparable cost. The remaining 50% are randomly generated from the space of feasible ship–route–trip combinations and subsequently repaired using the same logic. This strategy ensures that all initial individuals are feasible, cost-aware, and structurally diverse, which promotes effective exploration and fast convergence in the early generations. Figure 4 presents a sample chromosome layout where each column represents one candidate ship. The rows indicate the assigned storage site, the associated emission source, the number of round

trips, and the corresponding transport volume, respectively. Ships that are not chartered are indicated with 'NA' entries in the respective rows.

Ship	s_1	s_2	s_3	s_4	...	$s_{ S -1}$	$s_{ S }$
Assigned Storage Site	NA	NA	i_1	i_1	...	i_2	i_3
Assigned Emission Source	NA	NA	p_1	p_2	...	p_2	p_3 p_4
n_{sr}	NA	NA	300	350	...	150	200

Figure 4. Structure of a GA chromosome.

The fitness function directly evaluates the total system cost, including chartering cost, fuel and berthing cost, and penalty cost for direct emission. The algorithm starts from a randomly generated population of feasible solutions and iteratively improves them through selection, crossover, and mutation. Tournament selection is applied to favor high-quality solutions while maintaining population diversity. Crossover is performed using a two-point method to exchange route assignments between chromosomes. Mutation is conditionally applied after the crossover step, specifically targeting individuals ranking in the lowest 30% of the offspring's fitness values to enhance exploration in poorly performing regions of the solution space. For each selected individual, one ship gene is randomly chosen and mutated with a fixed probability of 0.08. The mutation involves either replacing the assigned route with another feasible alternative from the candidate route pool or adjusting the annual round-trip frequency within its allowable bounds. If the mutation causes any constraint violation, a constraint-based local repair strategy is triggered to restore feasibility without altering the rest of the chromosome. These mutated individuals remain within the same generation and are directly evaluated along with other offspring. The two-point crossover and mutation operations are illustrated in Figures 5 and 6, respectively.

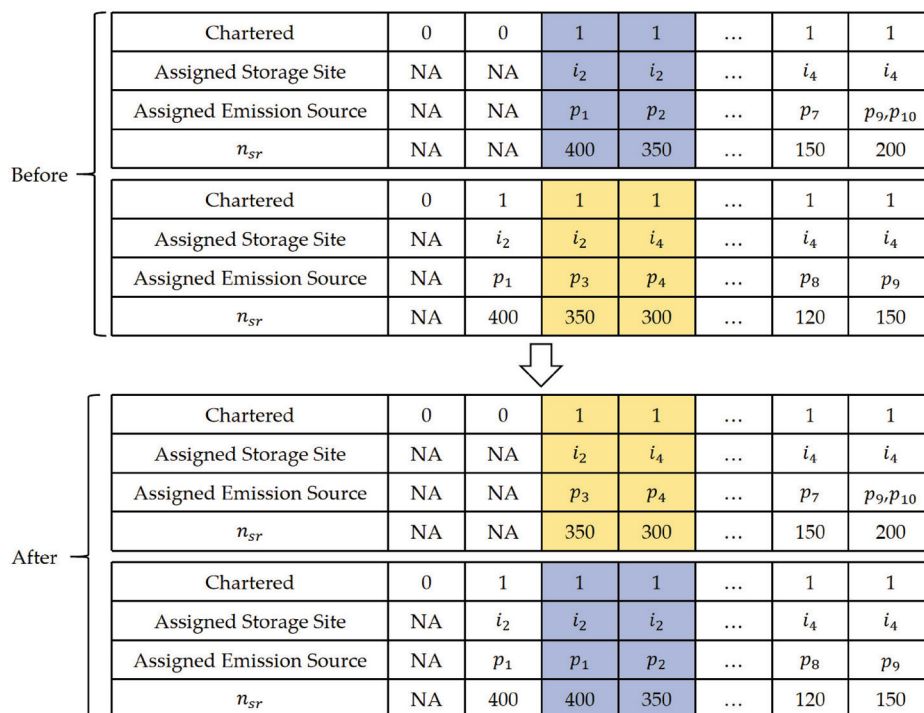


Figure 5. Two-point crossover operation (Note: Purple cells represent the original values; yellow cells represent the updated values after crossover operation.).

Before	Assigned Storage Site	NA	i_2	i_1	i_1	...	i_4	i_4
	Assigned Emission Source	NA	p_1	p_1	p_2	...	p_2	p_4
	n_{sr}	NA	400	300	350	...	120	150
↓								
After	Assigned Storage Site	NA	i_1	i_1	i_1	...	i_4	i_4
	Assigned Emission Source	NA	p_4	p_1	p_2	...	p_2	p_4
	n_{sr}	NA	400	300	350	...	120	150

Figure 6. Mutation operation (Note: Purple cells represent the original values; yellow cells represent the updated values after mutation operation.).

The offspring population replaces the current one, and each chromosome's fitness is evaluated. The best individuals are carried forward into the next generation, and the cycle of selection, crossover, mutation, and evaluation repeats until either the maximum number of iterations is reached or the best fitness value remains unchanged for a predefined number of consecutive generations.

5.2.2. Computational Results and Analysis

In this section, we present the results of applying GA to solve the second-stage CO₂ transport assignment problem. Given the parameters in Table 5, the total annual CO₂ generation from all emission sources is 33.6 million tons. Among all feasible combinations, selecting storage sites i_2 (15 million ton capacity) and i_4 (25 million ton capacity) provides a total capacity of 40 million tons, which satisfies the demand. This combination yields the lowest total construction cost of USD 1525 million. This result serves as the foundation for the second stage, where ship assignments are planned to connect all emission sources with the selected storage sites i_2 and i_4 . To solve this assignment problem, the GA is terminated after 300 generations, or earlier if the best fitness value shows no improvement over 50 consecutive generations. The algorithm employs a population size of 100, a crossover rate of 0.8, and a mutation rate of 0.1. Elitism is incorporated by retaining the top 20% of solutions in each generation. These parameter settings are selected to strike a balance between exploration and exploitation and demonstrated to be effective in producing high-quality solutions within an acceptable computational time. The results related to storage capacity utilization are presented in Table 16.

Table 16. First-stage storage site selection results and associated capacity utilization.

Storage Site	V_i (Million Tons)	The Amount of CO ₂ Received Annually (Million Tons)	Capacity Utilization (%)
i_2	15	10.6	70.7
i_4	25	23.0	92.0

Table 16 shows the two storage sites that were selected during the first-stage construction decision, along with their annual CO₂ receipts and corresponding capacity utilization rates. Site i_2 , with an annual capacity of 15 million tons, receives 10.6 million tons of CO₂ annually, yielding an 70.7% utilization rate. Site i_4 demonstrates even higher efficiency, with an annual intake of 23.0 million tons against a capacity of 25 million tons, achieving a utilization rate of 92.0%. The average utilization rate of the two storage stations is 81.4%, indicating that the GA solution is capable of generating assignments that effectively utilize the capacity of selected storage sites.

Based on the storage site layout determined by the first-stage decision, an analysis is conducted on the ship chartering, routing assignment, and how frequently each chartered ship completes round trips along its designated route per year.

Table 17 presents detailed results regarding ship chartering and route assignments. A total of 12 ships were selected for chartering from the available fleet, consisting of two small-class, seven medium-class, and three large-class ships. Each ship is assigned to a fixed route, connecting one or two emission sources to a designated storage site. Smaller ships are associated with higher values of n_{sr} and larger ships are assigned lower values of n_{sr} , as their greater capacity allows them to meet transport demand with fewer trips. Additionally, the large ships s_{38} , s_{39} , and s_{40} are assigned to routes that connect two emission sources. Compared with the result shown in Table 8, the GA solution charters one more large ship and one fewer small ship. This adjustment may be attributed to the change in storage site selection, which leads to longer average route distances and thus necessitates larger transport capacity per trip to maintain operational efficiency.

Table 17. The second-stage charter of ships and route assignment.

Chartered Ship	Ship Class	Transport Capacity (Ton)	Emission Sources Included in the Route	Storage Site Included in the Route	n_{sr}
s_9	Small	4500	p_2	i_2	488.9
s_{10}	Small	5000	p_6	i_4	327.0
s_{24}	Medium	10,800	p_1	i_4	185.1
s_{25}	Medium	11,000	p_4	i_4	245.4
s_{26}	Medium	11,200	p_9	i_2	201.4
s_{27}	Medium	11,400	p_9	i_2	214.3
s_{28}	Medium	11,600	p_5	i_4	186.7
s_{29}	Medium	11,800	p_7	i_4	211.3
s_{30}	Medium	12,000	p_{10}	i_4	268.1
s_{38}	Large	23,500	p_6, p_{10}	i_2	155.2
s_{39}	Large	24,200	p_5, p_8	i_4	207.8
s_{40}	Large	25,000	p_3, p_7	i_4	152.1

Based on the optimized assignment of ships and transport plan by solving the second-stage using GA, Table 18 presents the transport volumes and direct emissions for each emission source.

As shown in Table 15, 33.5873 million tons of CO₂ are transported annually, with only 0.0127 million tons emitted directly, corresponding to an average emission rate of less than 0.04%. Notably, all CO₂ in emission source p_6 are transported without any direct emissions. Compared to the results in Table 7, the direct emission volume under the GA solution is slightly higher. This may be attributed to the absence of one small ship with a lower transportation cost and the fact that the average transport route lengths under the GA assignment are generally longer than those in our previously proposed MIP model.

Based on the results obtained through the two-stage optimization framework and the assignment determined by solving the second stage using GA. Table 19 presents a comprehensive breakdown of the total system cost in the planning horizon.

Compared with the cost structure presented in Table 10, the results detailed in Table 19 show no significant change. The berthing cost remains the dominant contributor, accounting for over 80% of the total cost in both solutions. However, slight increases are observed in the chartering cost, fuel cost, and berthing cost under the GA-based assignment. The two-stage approach employing GA resulted in a total cost of USD 18,266 million, which is 2.4% higher than that achieved using the exact MIP model. In terms of

computational efficiency, the GA completed the second-stage assignment in an average of 28.3 s, representing a 57.9% reduction compared to the 67.2 s required by the Gurobi-solved MIP model for equivalent operational planning. These results indicate that, for the tested problem scale, the customized GA offers a notable relative decrease in computational time for the operational subproblem with only a marginal increase in cost. While the MIP model proved effective for this particular instance, the improved speed of the GA, along with its independence from costly commercial solvers and reduced demand for specialized optimization expertise, makes it a more practical option. This is especially true in scenarios where rapid decision-making is critical, the problem size increases beyond the capacity of exact methods, or resource constraints limit the deployment of sophisticated MIP tools.

Table 18. The second-stage CO₂ transport and direct emissions by each source.

Emission Source	Chartered Ship	w_{srp} (Million Tons)	e_p (Million Tons)
p_1	$s = 24$	1.9992	0.0008
p_2	$s = 9$	2.1999	0.0001
p_3	$s = 40$	2.4979	0.0021
p_4	$s = 25$	2.6992	0.0008
p_5	$s = 28$	2.1653	0.0016
	$s = 39$	0.8331	
p_6	$s = 10$	1.6349	0
	$s = 38$	1.8651	
p_7	$s = 29$	2.4933	0.0012
	$s = 40$	1.3055	
p_8	$s = 39$	4.1968	0.0032
p_9	$s = 26$	2.2562	0.0012
	$s = 27$	2.4426	
p_{10}	$s = 30$	3.2168	0.0017
	$s = 38$	1.7815	
Sum	\	33.5873	0.0127

Table 19. The value and proportion of each component of the cost in the planning horizon derived from the two-stage optimization framework.

Cost	Value (USD Million)	Percentage (%)
Total Construction Cost	1525	8.3
Total Chartering Cost	820	4.5
Total Fuel Cost	1243	6.8
Total Berthing Cost	14,650	80.2
Total Penalty Cost for Direct CO ₂ Emissions	28	0.2
Sum	18,266	100

6. Conclusions

Maritime CO₂ transport plays an important role in supporting CCS systems by connecting emission sources with suitable CO₂ storage sites. However, the transport process faces several planning and operational challenges, particularly when long sailing distances and strict emission constraints are considered simultaneously. The locations of CO₂ storage

sites significantly affect total tactical costs, as improper site selection may lead to increased sailing distances and higher transportation costs. Once storage sites are determined, further challenges arise in transport assignment, including the selection of appropriate ship types, the allocation of ship-route assignments, and the scheduling of annual sailing operations for each ship. Therefore, optimizing both the siting of CO₂ storage facilities and the assignment of transport tasks is essential to reduce total tactical costs and ensure the effective operation of maritime CCS systems.

To address these challenges, we propose an MIP model for the CSSL-TA problem. We establish a series of lemmas to analyze the mathematical structure of the model. Specifically, we prove that the decision variables associated with CO₂ storage site construction and ship-route assignments must be binary since fractional construction or half-chartered ships are not practical. We also prove that bilinear terms involving ship activity can be replaced by linear expressions, reducing model complexity. Furthermore, we demonstrate that the objective function is submodular with respect to construction, chartering, and routing variables, which reflects the diminishing marginal benefits of adding more resources. Additionally, we tighten two key sets of capacity constraints by coupling supply and demand. Instead of using fixed capacity parameters, the tightened constraints dynamically reflect the actual CO₂ generation volumes associated with each storage site and transport route. These enhancements help reduce the feasible space and accelerate convergence. These theoretical insights provide a solid foundation for enhancing solution efficiency and extending the model in future research.

In the numerical experiments, the model yields a total tactical cost of USD 17,830 million in the planning horizon. Two storage sites are selected for construction, with an average utilization rate of 93.9%. Out of a fleet of 40 available ships, 12 ships are selected for chartering, including three small-class, seven medium-class, and two large-class ships. Among all cost components, the berthing cost accounts for the largest share, reaching 80.5% of the total, while the penalty cost for direct CO₂ emissions is the smallest at only 0.1%. This outcome results from the model's preference to avoid the relatively high penalty imposed on untransported emissions. The strategy not only ensures sufficient storage capacity and effective transport coverage, but also minimizes costs by aligning storage siting and ship assignment decisions with capacity constraints, operational needs and direct emission limits. The sensitivity analysis reveals that variations in annual available sailing time, sailing speed, and storage capacity affect the total cost, storage site deployment, chartering cost, transportation cost, and direct emission volume. A longer sailing time, higher sailing speed, and greater capacity lead to fewer ships being needed. These factors, in turn, affect storage site construction strategies, transportation costs, and the volume of direct CO₂ emissions to varying degrees. In particular, higher sailing speeds cause a rapid increase in fuel cost, which significantly raises the total cost and limits the model's ability to assign ships to routes with low transportation cost for residual CO₂.

To enhance practical applicability, we further design a two-stage planning framework that separates strategic site selection and transport assignment. In the first stage, storage sites are selected based on storage capacity and construction cost. In the second stage, a GA is applied to solve the transport assignment problem given the fixed site locations. While the two-stage approach provides a computationally efficient alternative, the resulting total tactical cost is 2.4% higher than that obtained by the exact MIP model, but it reduces the computational time by 57.9%. This result demonstrates that the proposed two-stage method can serve as a reliable and efficient decision-support tool, particularly useful for real-world applications where decision-makers must balance accuracy and tractability.

The optimization framework aligns with a multi-level decision-making structure:

1. Strategic level: The framework supports long-term decisions on the construction and spatial deployment of CO₂ storage sites. These decisions involve large-scale infrastructure planning and capital investment, and are typically made by government agencies, regional CCS authorities, or public infrastructure coordinators.
2. Routing and assignment level: The framework determines which transport routes should be activated, assigns specific ships to selected routes, and specifies the number of round trips for each ship. These mid-level planning tasks are usually managed by CCS system operators or logistics planners responsible for transport design and resource allocation.
3. Operational level: It provides detailed support for day-to-day execution, including CO₂ volume allocation from emission sources, compliance with berthing frequency requirements, and control of untransported emissions. These operational tasks are implemented by maritime transport companies or CCS logistics service providers.

While the proposed optimization framework demonstrates strong potential for improving the cost effectiveness and feasibility of maritime CO₂ transport, its real-world implementation may be affected by several practical risks. These include uncertainties in CO₂ emission volumes and storage capacities, unplanned infrastructure failures at storage sites or ports, and disruptions in ship operations due to weather conditions, scheduling conflicts, or route congestion. To mitigate such risks, the framework supports proactive measures such as the use of redundant storage capacity, flexible route selection, adaptive ship assignments, and compliance-based emission control strategies. In the event of unforeseen failures, the model allows for real-time operational adjustments and the use of backup transport resources to minimize the severity of consequences. Future research should aim to more explicitly incorporate the impact of realistic factors such as weather conditions, route congestion, and berthing delays, which are crucial in maritime CO₂ transport and can significantly impact the feasibility and reliability of optimization models in real-world applications. The development of stochastic programming or robust optimization models would allow for a more direct representation of these uncertainties, leading to solutions that better ensure operational feasibility and reliability under variable conditions. In addition, the current model can be extended by incorporating real-world complexities such as uncertainty in CO₂ capture volumes, fuel prices, and system performance, which are critical to the robustness of maritime CCS planning. Recent studies have highlighted how stochastic optimization, Monte Carlo simulations, and scenario-based models can capture uncertainties in injection rates, storage efficiency, and energy market volatility, enabling more adaptive infrastructure design and avoiding stranded CO₂ scenarios [65–67]. Another promising extension involves integrating flexible pipeline transport and marine shipping options for hybrid CO₂ routing. Research comparing techno-energetic characteristics of liquefied CO₂ shipping and pipelines has demonstrated that such configurations can improve system resilience and reduce costs in maritime contexts [68]. Additionally, the model could incorporate policy-driven constraints such as dynamic emission limits or carbon-trading schemes. Emerging studies in maritime transport operations under emission-trading frameworks show that fluctuating carbon prices significantly influence routing, operational strategies, and investment planning [69]. Lastly, the model can be enhanced by adopting multi-objective optimization approaches that jointly consider cost efficiency, emission reduction, and infrastructure utilization. In particular, formulations that incorporate uncertainty can support more balanced decision-making and help identify Pareto-efficient infrastructure strategies [70]. These extensions would further enhance the model's applicability to long-term sustainable planning of maritime CCS systems.

Author Contributions: Conceptualization, S.W.; methodology, Y.T., Y.Y., Y.D. and S.W.; software, Y.T.; validation, Y.T.; formal analysis, Y.T.; investigation, Y.T., Y.Y. and S.W.; resources, S.W. and Y.D.; data curation, Y.T. and Y.Y.; writing—original draft preparation, Y.T., Y.Y., S.W. and Y.D.; writing—review and editing, Y.T., Y.Y., S.W. and Y.D.; visualization, Y.T.; supervision, S.W.; project administration, S.W.; funding acquisition, S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Change, P.C. *Global Warming of 1.5 C*; World Meteorological Organization: Geneva, Switzerland, 2018.
2. Transforming Industry through CCUS. Available online: <https://www.iea.org/reports/transforming-industry-through-ccus> (accessed on 3 April 2025).
3. Leeson, D.; Mac Dowell, N.; Shah, N.; Petit, C.; Fennell, P.S. A techno-economic analysis and systematic review of carbon capture and storage (CCS) applied to the iron and steel, cement, oil refining and pulp and paper industries, as well as other high purity sources. *Int. J. Greenh. Gas Control* **2017**, *61*, 71–84. [CrossRef]
4. Chamwudhipreacha, N.; Blunt, M.J. CO₂ storage potential in the North Sea. In Proceedings of the International Petroleum Technology Conference, Bangkok, Thailand, 7–9 November 2011.
5. Orchard, K.; Hay, M.; Ombudstvedt, I.; Skagestad, R.; Joos, M.; Nysæter, G.; Sjøbris, C.; Jarøy, A.G.; Durusut, E.; Craig, J. The Status and Challenges of CO₂ Shipping Infrastructures. In Proceedings of the 15th International Conference on Greenhouse Gas Control Technologies, GHGT-15, Abu Dhabi, United Arab Emirates, 18 March 2021.
6. Kjærstad, J.; Skagestad, R.; Eldrup, N.H.; Johnsson, F. Ship transport—A low cost and low risk CO₂ transport option in the Nordic countries. *Int. J. Greenh. Gas Control* **2016**, *54*, 168–184. [CrossRef]
7. Knoope, M.M.J.; Ramírez, A.; Faaij, A.P.C. Investing in CO₂ transport infrastructure under uncertainty: A comparison between ships and pipelines. *Int. J. Greenh. Gas Control* **2015**, *41*, 174–193. [CrossRef]
8. Roussanaly, S.; Brunsvold, A.L.; Hognes, E.S. Benchmarking of CO₂ transport technologies: Part II—Offshore pipeline and shipping to an offshore site. *Int. J. Greenh. Gas Control* **2014**, *28*, 283–299. [CrossRef]
9. Zhang, S.; Liu, L.; Zhang, L.; Zhuang, Y.; Du, J. An optimization model for carbon capture utilization and storage supply chain: A case study in northeastern China. *Appl. Energy* **2018**, *231*, 194–206. [CrossRef]
10. Bjerketvedt, V.S.; Tomasgard, A.; Roussanaly, S. Optimal design and cost of ship-based CO₂ transport under uncertainties and fluctuations. *Int. J. Greenh. Gas Control* **2020**, *103*, 103190. [CrossRef]
11. Drax Global. Why and How Is Carbon Dioxide Transported? Available online: <https://www.drax.com/carbon-capture/why-and-how-is-carbon-dioxide-transported/> (accessed on 17 May 2025).
12. Roussanaly, S.; Jakobsen, J.P.; Hognes, E.H.; Brunsvold, A.L. Benchmarking of CO₂ transport technologies: Part I—Onshore pipeline and shipping between two onshore areas. *Int. J. Greenh. Gas Control* **2013**, *19*, 584–594. [CrossRef]
13. DNV. Background on CO₂ Transport by Ship. Available online: <https://www.dnv.com/maritime/jip/ceto/background/> (accessed on 17 May 2025).
14. Chemistry Learner. Carbon Dioxide Phase Diagram Explanation. Available online: <https://www.chemistrylearner.com/co2-phase-diagram.html> (accessed on 17 May 2025).
15. Al Baroudi, H.; Awoyomi, A.; Patchigolla, K.; Jonnalagadda, K.; Anthony, E.J. A review of large-scale CO₂ shipping and marine emissions management for carbon capture, utilisation and storage. *Appl. Energy* **2021**, *287*, 116510. [CrossRef]
16. Aramco. Carbon Capture, Utilization, and Storage. Available online: <https://www.aramco.com/en/what-we-do/energy-innovation/advancing-energy-solutions/carbon-capture-utilization-and-storage> (accessed on 17 May 2025).
17. Kokal, S.L.; Sanni, M.L.; Alhashboul, A.A. Design and implementation of the first CO₂-EOR demonstration project in Saudi Arabia. In Proceedings of the SPE Annual Technical Conference and Exhibition, Dubai, United Arab Emirates, 26–28 September 2016; SPE: Richardson, TX, USA, 2016; pp. 1–18.
18. Aramco. Driving Green Initiatives. Available online: <https://china.aramco.com/en/news-media/global-news/2020/driving-green-initiatives> (accessed on 17 May 2025).
19. Northern Lights. What We Do. Available online: <https://norlights.com/what-we-do/> (accessed on 17 May 2025).
20. DNV. Northern Lights Shows the Way to Seaborne CCS Solutions. Available online: <https://www.dnv.com/expert-story/maritime-impact/Northern-Lights-shows-the-way-to-seaborne-CCS-solutions/> (accessed on 17 May 2025).

21. Equinor. Northern Lights: The World's First Cross-Border CO₂ Transport and Storage Facility. Available online: <https://www.equinor.com/energy/northern-lights> (accessed on 17 May 2025).
22. Liotta, G.; Stecca, G.; Kaihara, T. Optimisation of freight flows and sourcing in sustainable production and transportation networks. *Int. J. Prod. Econ.* **2015**, *164*, 351–365. [CrossRef]
23. Nguyen, T.B.; Leonzio, G.; Zondervan, E. Supply chain optimization framework for CO₂ capture, utilization, and storage in Germany. *Phys. Sci. Rev.* **2023**, *8*, 1685–1711. [CrossRef]
24. Koh, S.L.; Jia, F.; Gong, Y.; Zheng, X.; Dolgui, A. Achieving carbon neutrality via supply chain management: Position paper and editorial for IJPR special issue. *Int. J. Prod. Res.* **2023**, *61*, 6081–6092. [CrossRef]
25. Leonzio, G.; Bogle, D.; Foscolo, P.U.; Zondervan, E. Optimization of CCUS supply chains in the UK: A strategic role for emissions reduction. *Chem. Eng. Res. Des.* **2020**, *155*, 211–228. [CrossRef]
26. Lu, H.; Ma, X.; Huang, K.; Fu, L.; Azimi, M. Carbon dioxide transport via pipelines: A systematic review. *J. Clean. Prod.* **2020**, *266*, 121994. [CrossRef]
27. Vakili, S.; Manias, P.; Armstrong, L.M.; Turnock, S.; Teagle, D.A. Technical, economic, and environmental assessment of CO₂ ship transport in carbon capture and storage. *J. Environ. Manag.* **2025**, *373*, 123919. [CrossRef]
28. Aspelund, A.; Mølnevik, M.J.; De Koeijer, G. Ship transport of CO₂: Technical solutions and analysis of costs, energy utilization, exergy efficiency and CO₂ emissions. *Chem. Eng. Res. Des.* **2006**, *84*, 847–855. [CrossRef]
29. d'Amore, F.; Bezzo, F. Economic optimisation of European supply chains for CO₂ capture, transport and sequestration. *Int. J. Greenh. Gas Control* **2017**, *65*, 99–116. [CrossRef]
30. Ma, Z.; Chen, B.; Pawar, R.J. Phase-based design of CO₂ capture, transport, and storage infrastructure via SimCCS 3.0. *Sci. Rep.* **2023**, *13*, 6527.
31. Bachu, S. Sequestration of CO₂ in geological media: Criteria and approach for site selection in response to climate change. *Energy Convers. Manag.* **2000**, *41*, 953–970. [CrossRef]
32. Grataloup, S.; Bonijoly, D.; Brosse, E.; Dreux, R.; Garcia, D.; Hasanov, V.; Lescanne, M.; Renoux, P.; Thoraval, A. A site selection methodology for CO₂ underground storage in deep saline aquifers: Case of the paris basin. *Energy Procedia* **2009**, *1*, 2929–2936. [CrossRef]
33. Alcalde, J.; Heinemann, N.; James, A.; Bond, C.E.; Ghanbari, S.; Mackay, E.J.; Hazeldine, R.S.; Faulkner, D.R.; Worden, R.H.; Allen, M.J. A criteria-driven approach to the CO₂ storage site selection of east mey for the acorn project in the north sea. *Mar. Pet. Geol.* **2021**, *133*, 105309. [CrossRef]
34. Raza, A.; Rezaee, R.; Gholami, R.; Bing, C.H.; Nagarajan, R.; Hamid, M.A. A screening criterion for selection of suitable CO₂ storage sites. *J. Nat. Gas Sci. Eng.* **2016**, *28*, 317–327. [CrossRef]
35. Hsu, C.W.; Chen, L.T.; Hu, A.H.; Chang, Y.M. Site selection for carbon dioxide geological storage using analytic network process. *Sep. Purif. Technol.* **2012**, *94*, 146–153. [CrossRef]
36. Luo, D.; Yuan, Y.; Chen, J.; Li, Q.; Liang, J.; Zhao, H. Structural and reservoir characteristics of potential carbon dioxide storage sites in the northern south yellow sea basin, offshore eastern China. *J. Mar. Sci. Eng.* **2024**, *12*, 1733. [CrossRef]
37. Aarnes, J.E.; Selmer-Olsen, S.; Carpenter, M.E.; Flach, T.A. Towards guidelines for selection, characterization and qualification of sites and projects for geological storage of CO₂. *Energy Procedia* **2009**, *1*, 1735–1742. [CrossRef]
38. Akhurst, M.C.; Quinn, M.; Delprat-Jannaud, F.; Bossie-Codreanu, D.; Hannis, S.; Lecomte, J.C.; Frykman, P. Risk assessment-led pre-characterisation of a site for commercial-scale storage of CO₂. In Proceedings of the Fourth EAGE CO₂ Geological Storage Workshop, Stavanger, Norway, 8–10 April 2014.
39. Gruson, J.F.; Serbutoviez, S.; Delprat-Jannaud, F.; Akhurst, M.; Nielsen, C.; Dalhoff, F.; Iacobellis, S. Techno-economic assessment of four CO₂ storage sites. *Oil Gas Sci. Technol. IFP Energ. Nouv.* **2015**, *70*, 753–766. [CrossRef]
40. Elahi, N.; Shah, N.; Korre, A.; Durucan, S. Multi-stage stochastic optimisation of a CO₂ transport and geological storage in the UK. *Energy Procedia* **2017**, *114*, 6514–6525. [CrossRef]
41. Knoope, M.M.J.; Ramirez, A.; Faaij, A.P.C. The influence of uncertainty in the development of a CO₂ infrastructure network. *Appl. Energy* **2015**, *158*, 332–347. [CrossRef]
42. Gabrielli, P.; Campos, J.; Becattini, V.; Mazzotti, M.; Sansavini, G. Optimization and assessment of carbon capture, transport and storage supply chains for industrial sectors: The cost of resilience. *Int. J. Greenh. Gas Control* **2022**, *121*, 103797. [CrossRef]
43. Engebø, A.; Ahmed, N.; Garstad, J.J.; Holt, H. Risk assessment and management for CO₂ capture and transport facilities. *Energy Procedia* **2013**, *37*, 2783–2793. [CrossRef]
44. Kjærstad, J.; Morbee, J.; Odenberger, M.; Johnsson, F.; Tzimas, E. Modelling large-scale CCS development in Europe linking techno-economic modelling to transport infrastructure. *Energy Procedia* **2013**, *37*, 2941–2948. [CrossRef]
45. Morbee, J.; Serpa, J.; Tzimas, E. Optimised deployment of a European CO₂ transport network. *Int. J. Greenh. Gas Control* **2012**, *7*, 48–61. [CrossRef]

46. Bennaes, A.; Skogset, M.; Svorkdal, T.; Fagerholt, K.; Herlicka, L.; Meisel, F.; Rickels, W. Optimization of a ship-based logistics system for carbon capture and storage. In Proceedings of the International Conference on Computational Logistics, Cham, Switzerland, 1–3 September 2022; pp. 44–59.
47. Ajayi, T.; Gomes, J.S.; Bera, A. A review of CO₂ storage in geological formations emphasizing modeling, monitoring and capacity estimation approaches. *Pet. Sci.* **2019**, *16*, 1028–1063. [CrossRef]
48. Schrijver, A. *Combinatorial Optimization*; Springer: Berlin, Germany, 2003.
49. Anthonsen, K.L.; Christensen, N.P. *EU Geological CO₂ Storage Summary*; Geological Survey of Denmark and Greenland: Copenhagen, Denmark, 2021; pp. 1–28.
50. Smith, E.; Morris, J.; Kheshgi, H.; Teletzke, G.; Herzog, H.; Paltsev, S. The cost of CO₂ transport and storage in global integrated assessment modeling. *Int. J. Greenh. Gas Control* **2021**, *109*, 103367. [CrossRef]
51. Gassnova, S.F. Potential for reduced costs for carbon capture, transport and storage value chains (CCS). *Nor. Full-Scale CCS Demonstr. Proj.* **2020**, *28*, 1–61.
52. Duan, H.; Dong, X.; Xie, P.; Chen, S.; Qin, B.; Dong, Z.; Yang, W. Peaking industrial CO₂ emission in a typical heavy industrial region: From multi-industry and multi-energy type perspectives. *Int. J. Environ. Res. Public Health* **2022**, *19*, 7829. [CrossRef]
53. U.S. Environmental Protection Agency. Sources of Greenhouse Gas Emissions: Overview. Available online: <https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions> (accessed on 3 April 2025).
54. China CO₂ Emission Accounts 1997–2015. Available online: <https://www.nature.com/articles/sdata2017201> (accessed on 3 April 2025).
55. Directive, E.C. Directive of the European Parliament and of the Council on the Promotion of the Use of Energy from Renewable Sources. In Proceedings of the European Parliament and Council Conference on Renewable Energy, Brussels, Belgium, 23 April 2008.
56. EU Action on the EU Emissions Trading System (EU ETS). Available online: https://climate.ec.europa.eu/eu-action/eu-emissions-trading-system-eu-ets/about-eu-ets_en (accessed on 15 May 2025).
57. Stationary Operators-EU ETS Compliance. Available online: <https://www.epa.ie/our-services/licensing/climate-change/eu-emissions-trading-system-/emissions-trading-system---stationary-installations/stationary-operators---eu-ets-compliance/> (accessed on 15 May 2025).
58. 2023 IMO Strategy on Reduction of GHG Emissions from Ships. Available online: <https://www.imo.org/en/OurWork/Environment/Pages/2023-IMO-Strategy-on-Reduction-of-GHG-Emissions-from-Ships.aspx> (accessed on 15 May 2025).
59. Now Available: July 2024 Edition of the DNV Class Rules for Ships and Offshore Units. Available online: <https://www.dnv.com/news/now-available-july-2024-edition-of-the-dnv-class-rules-for-ships-and-offshore-units/> (accessed on 15 May 2025).
60. Global Status of CCS 2021–Global CCS Institute. Available online: <https://www.globalccsinstitute.com/resources/publications-reports-research/global-status-of-ccs-2> (accessed on 15 May 2025).
61. Harper Petersen-Container. Available online: <https://www.harperpetersen.com/container> (accessed on 3 April 2025).
62. CO₂ Shipping Study Final Report for BEIS. Available online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/761762/BEIS_Shipping_CO2.pdf (accessed on 3 April 2025).
63. Du, Y.; Chen, Q.; Quan, X.; Long, L.; Fung, R.Y. Berth allocation considering fuel consumption and vessel emissions. *Transp. Res. Part E Logist. Transp. Rev.* **2011**, *47*, 1021–1037. [CrossRef]
64. Ship; Bunker. Global 20 Ports Average Bunker Prices. Available online: <https://shipandbunker.com/prices/av/global/av-g20-global-20-ports-average> (accessed on 3 April 2025).
65. Wang, Z.; Qi, S.; Zheng, B. Uncertainty analysis of the storage efficiency factor for CO₂ saline resource estimation. *Energies* **2024**, *17*, 1297. [CrossRef]
66. Roussanally, S.; Ouassou, J.A.; Anantharaman, R.; Haaf, M. Impact of uncertainties on the design and cost of CCS from a waste-to-energy plant. *Front. Energy Res.* **2020**, *8*, 17. [CrossRef]
67. Middleton, R.; Yaw, S. The cost of getting CCS wrong: Uncertainty, infrastructure design, and stranded CO₂. *Int. J. Greenh. Gas Control* **2018**, *70*, 1–11. [CrossRef]
68. Pouloupoulos, S.G.; Tsiakaras, P.; Voutetakis, S. CO₂ marine transportation from a techno-energetic perspective. *J. Mar. Sci. Eng.* **2024**, *12*, 12.
69. Zhang, Y.; Du, B.; Wu, W.; Liao, T.W. Sustainable maritime transportation operations with emission trading. *J. Mar. Sci. Eng.* **2023**, *11*, 1647.
70. Spek, M.V.D.; Fout, T.E.; García, M.; Kuncheekanna, V.N.; Matuszewski, M.; McCoy, S.; Morgan, J.C.; Nazir, S.M.; Ramírez, A.; Roussanally, S.; et al. Uncertainty analysis in the techno-economic assessment of CO₂ capture and storage technologies. *Int. J. Greenh. Gas Control* **2020**, *100*, 103113. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Collaborative Scheduling of Yard Cranes, External Trucks, and Rail-Mounted Gantry Cranes for Sea–Rail Intermodal Containers Under Port–Railway Separation Mode

Xuhui Yu * and Cong He

College of Transportation Engineering, Dalian Maritime University, Dalian 116026, China; hecong@dlnu.edu.cn

* Correspondence: yuxuhui@dlnu.edu.cn

Abstract: The spatial separation of port yards and railway hubs, which relies on external truck drayage as a necessary link, hampers the seamless transshipment of sea–rail intermodal containers between ports and railway hubs. This creates challenges in synchronizing yard cranes (YCs) at the port terminal, external trucks (ETs) on the road, and rail-mounted gantry cranes (RMGs) at the railway hub. However, most existing studies focus on equipment scheduling or container transshipment organization under the port–railway integration mode, often overlooking critical time window constraints, such as train schedules and export container delivery deadlines. Therefore, this study investigates the collaborative scheduling of YCs, ETs, and RMGs for synchronized loading and unloading under the port–railway separation mode. A mixed-integer programming (MIP) model is developed to minimize the maximum makespan of all tasks and the empty-load time of ETs, considering practical time window constraints. Given the NP-hard complexity of this problem, an improved genetic algorithm (GA) integrated with a “First Accessible Machinery” rule is designed. Extensive numerical experiments are conducted to validate the correctness of the proposed model and the performance of the solution algorithm. The improved GA demonstrates a 6.08% better solution quality and a 97.94% reduction in computation time compared to Gurobi for small-scale instances. For medium to large-scale instances, it outperforms the adaptive large neighborhood search (ALNS) algorithm by 1.51% in solution quality and reduces computation time by 45.71%. Furthermore, the impacts of objective weights, equipment configuration schemes, port–railway distance, and time window width are analyzed to provide valuable managerial insights for decision-making to improve the overall efficiency of sea–rail intermodal systems.

Keywords: sea–rail intermodal container; port–railway separation; collaborative scheduling; synchronized loading/unloading; improved genetic algorithm; container operation time window

1. Introduction

The development of economic globalization and the growth of world trade have greatly stimulated the prosperity of container transportation. However, relying solely on one mode of transportation often fails to complete the entire journey of container shipping due to the requirement for door-to-door service, and it cannot provide a cost-effective and low-carbon solution either. As a result, container intermodal transportation, which integrates various modes of transport such as sea, rail, and road, has been proposed to streamline the transportation process, reduce transportation cost, and enhance environmental sustainability [1]. For international containers, sea–rail intermodal transportation

is the most optimal choice, combining the high capacity and low environmental impact of sea transport with the high speed of railway [2]. In developed countries, such as the Netherlands and the United States, the sea–rail mode typically accounts for 20% to 40% of total port container throughput [3]. In contrast, the adoption of the sea–rail mode in China began relatively late. According to the Annual Report on China Container Industry and Intermodal Development 2023 issued by China Container Industry Association, after years of infrastructure construction and market operations, the proportion of sea–rail intermodal containers in China has increased from just 1.2% in 2016 to 3.8% in 2023. With the vigorous promotion of multimodal transport development policies in China, the sea–rail mode is poised to experience substantial growth, driven by its untapped potential and the increasing integration of efficient logistics networks.

The significant gap between China and developed countries in container sea–rail intermodal transportation can be attributed to several factors, including the lower coverage of the railway network and the shortage of railcars. The lag in railway infrastructure development has resulted in the frequent neglect of railway access requirements during the planning of container ports, which in turn creates a spatial separation between container ports and railway hubs. Moreover, challenges such as geographical locations, space utilization, and track connection conditions further complicate efforts to achieve seamless integration between existing container ports and railway hubs. According to statistics, among the 37 container ports in China currently connected to railway hubs, only 14 have realized seamless integration between the ports and railways. Therefore, in the short term, the transportation of most sea–rail intermodal containers between ports and railway hubs will still require ET drayage to serve as an indispensable link. While optimizing the dispatching of ETs is essential for reducing drayage costs and improving drayage efficiency, it alone is insufficient to achieve overall systemic improvements. This limitation stems from the fact that the YCs in the port yard, the ETs on the road, and the RMGs in the railway yard constitute a tightly coupled tripartite system, where operational delays at any link induce cascading effects that degrade overall productivity. To address this interdependency, it is imperative to investigate the collaborative scheduling of YC, ET, and RMG under the port–railway separation mode in sea–rail intermodal systems.

Although abundant studies have concentrated on optimizing the collaborative scheduling of multiple pieces of equipment and the organization of container transshipment in sea–rail intermodal terminals, they are all limited to the port–railway integration mode [4–7]. However, the port–railway separation mode, which is the most common mode in China’s sea–rail intermodal systems, has yet to be fully explored. In this mode, the ET experiences longer transfer times than the IT due to greater road distances. The intensified temporal coupling of multiple equipment in such a spatially separated system necessitates dedicated coordination mechanisms. Furthermore, the simultaneous loading and unloading of import and export containers [8] and critical time window constraints, such as train time windows [4] and export container delivery deadlines [4,6], are often neglected in the existing literature to simplify the model. In reality, export containers can only be unloaded after the train arrives at the railway hub, and they must be unloaded before the train departs and delivered to the terminal before the vessel’s loading cutoff. Additionally, export containers must be loaded onto the train they are assigned to before its departure. Once the completion time exceeds the time window limit, it may result in the cargo missing the scheduled vessel or train, leading to delayed delivery and incurring demurrage charges. To address the research gap, this study proposes a collaborative scheduling optimization framework for RMGs, ETs, and YCs under the port–railway separation mode, integrating both train time windows and export container delivery deadlines. The key contributions are threefold:

- (1) The collaborative scheduling of YCs, ETs, and RMGs under the port–railway separation mode is mathematically modeled as an MIP model, considering synchronized loading and unloading scenarios, as well as practical constraints such as container task time windows and equipment operation sequence requirements. This model further advances the optimization framework for collaborative scheduling of multiple equipment systems;
- (2) Methodologically, this study proposes an improved GA with the First Accessible Machinery (FAM) rule to address the complexity and dependencies of the collaborative equipment scheduling problem, which integrates pickup–delivery and flexible hybrid flow shop scheduling;
- (3) Extensive computational experiments are conducted to evaluate both the efficacy and computational efficiency of the proposed model and algorithm. Furthermore, sensitivity analysis is performed to examine the impact of key factors and model parameters, providing managerial insights for sea–rail intermodal operators.

The remainder of this paper is organized as follows: Section 2 provides a comprehensive literature review. Section 3 formalizes the problem statement and presents the mathematical formulation. Section 4 details the improved GA framework. Section 5 discusses numerical experiments and analytical results. Section 6 concludes this study and proposes future research directions.

2. Literature Review

Over the past decade, sea–rail intermodal transportation has garnered significant attention due to its impact on container transshipment speed and overall port efficiency. For a comprehensive review, readers are referred to Abu-Aisha et al. [9]. The literature most relevant to our study includes works that explore the integrated scheduling of multiple equipment and the organization of container transshipment in sea–rail intermodal terminals. While the integrated scheduling of multiple equipment in automated container terminals [10] and traditional container terminals [11] has been studied, these works are not reviewed here due to their exclusion of railway systems.

2.1. Collaborative Scheduling of Multiple Equipment in Sea–Rail Intermodal Terminals

Most container terminals in developed countries and fewer than half of the container terminals in China have planned railway operation areas within the port or railway hubs adjacent to it. This arrangement is known as the port–railway integration mode, and the terminal is referred to as a sea–rail intermodal terminal. With this mode, the horizontal transportation of transshipment containers can be carried out by in-port vehicles (IVs), such as an internal truck (IT), an automatic guided vehicle (AGV), and an intelligent guided vehicle (IGV), within a relatively short time. The convenience of container transshipment in sea–rail intermodal terminals has significantly promoted the development of container intermodal transportation. For this reason, current research mainly focuses on the collaborative scheduling of multiple equipment in sea–rail intermodal terminals.

There are typically two types of railway layout modes for accessing a port: parallel and vertical access to the terminal. In a parallel railway layout, the tracks are parallel to the terminal shoreline, which is the most common type of layout and also the focus of most studies. Yan et al. [12] formulated the integrated scheduling of RMGs, ITs, and reach stackers at the railway handling area in container terminals as an MIP model and solved it using an improved GA, with the aim of minimizing the maximum completion time and the total waiting time of all equipment. They did not consider the schedule of trains and the time window of export container deliveries. Chang et al. [8] simultaneously studied the storage space allocation of the railway container yard and the job sequences

of RMGs, ITs, and QCs (quay cranes) for outbound containers, aiming at minimizing the makespan of handling operations. To solve this problem, a two-stage heuristic algorithm is proposed by introducing the rolling planning horizon and a new update strategy. In their study, the unloading of outbound containers from trains as well as train schedules were not considered, and they focused on the transshipment between the railway yard and the vessel. Li et al. [13] developed an MIP model with the objective of minimizing the makespan for the collaborative scheduling of RMGs and ITs by considering two operation modes, train-yard-vessel and train-vessel, and solved it using a GA. Liu et al. [6] presented a multi-objective optimization model to minimize the makespan, equipment waiting time, and unnecessary IT empty travel time for coordinating the scheduling of RMGs, ITs, and YCs in a sea–rail intermodal terminal and designed an adaptive large neighborhood search (ALNS) algorithm to solve it. Unlike previous studies, Li et al. [14] presented a new collaborative scheduling model of QCs, AGVs, YCs, and RMGs targeting the minimization of the makespan, the total waiting time of AGVs, and the total energy consumption of handling cranes by considering a flexible transshipment route strategy that allows containers to select the most appropriate transshipment mode during operation. To solve their problem, they designed an INSGA-II. Xia et al. [7] proposed a collaborative scheduling approach for RMGs, ITs, and YCs, considering both inter-equipment interferences and the workload balance of RMGs. They formulated the problem as a binary integer programming model intended to minimize the operation time of RMGs based on a space-time-state network and developed a Lagrangian relaxation-based heuristic to efficiently solve the reformulated model.

Although many sea–rail intermodal studies have been conducted on the integrated scheduling of multiple equipment under the parallel railway layout, only two studies have focused on the traditional vertical layout, including those by Yang et al. [15] and Yang et al. [4]. Yang et al. [15] proposed a collaborative scheduling model for the RMGs and AGVs with the goal of minimizing the energy consumption of all equipment, which was solved using a two-layer scheduling rule and a self-adaptive chaotic GA. The model was applied to a scenario involving a vertical railway accessing a port under the mixed operational mode of “train-vessel” and “train-yard-vessel”. Expanding upon their previous work, Yang et al. [4] studied the cooperative scheduling of RMGs, AGVs, and YCs under the loading and unloading mode. They developed an MIP model in order to minimize the maximum completion time and the operational energy consumption of the RMGs and AGVs. With the emergence of the U-shaped layout in container terminals, the collaborative scheduling for this layout from the perspective of sea–rail intermodality has also been investigated. Li et al. [5] established a hybrid integer programming model to collaboratively schedule RMGs, IGVs, and YCs for container transshipment between a U-shaped automated terminal yard and a railway yard, with the objective of minimizing container transit times. They applied a solution framework based on an alternating direction method of multipliers to tackle the non-crossing constraints and safe distance constraints.

2.2. Organization of Container Transshipment in Sea–Rail Intermodal Terminals

The organization of container transshipment in the sea–rail intermodal system involves the loading/unloading of container trains, handling operations of the storage yard, and the loading and unloading of container vessels. This category of research primarily focuses on the transshipment quantity of containers between different trains, storage yards, and vessels, while it does not consider the detailed operational handling time and operational sequence of various equipment. Zhao et al. [16] proposed an inbound container distribution model that considers temporal and resource constraints, with the goal of minimizing the total time import containers spend in the coordination area of the sea–rail

intermodal terminal, and presented an effective two-point crossover GA to solve it. Hu et al. [17] integrated the inter-terminal transportation of containers with the rail freight formulation and transportation process toward the hinterland to formulate an integer linear programming model, with the objective of minimizing the operating costs of the entire logistics system. Yan et al. [18] extended Zhao et al. [16] by incorporating the train schedule template decision into the inbound container transshipment operation problem. The schedules of trains were constrained to coordinate with the unloading time of vessels. The problem was formulated as an MIP model and a tailored rolling horizon approach was proposed to solve it. In contrast, Zhao et al. [19] aimed to improve the transshipment efficiency of outbound intermodal containers among vessels, storage yards, and trains. They proposed an improved GA and validated the model and algorithm via simulation. To make the model more applicable to real-life scenarios, Yan et al. [20] further investigated the bidirectional transshipment operations of import and export containers between vessels and trains by simultaneously determining the train schedule plan and the container transshipment plan among vessels, storage yards, and trains. An MIP model was developed to minimize the total cost by considering the service time window of trains.

In order to smooth container transshipment organization, some studies also concentrated on the space allocation of storage yards or the layout of sea–rail intermodal terminals. Chang and Zhu [21] designed a two-stage model for the storage space allocation problem in the railway operation area, with the objective of balancing inbound container distribution and minimizing overlapping amounts. To solve this problem, a simulated annealing algorithm was introduced for stage 1, and a rolling-horizon-based heuristic was designed for stage 2. Abu Aisha et al. [22] proposed an innovative layout for a sea–rail intermodal terminal to reduce the operation cost and accelerate container flows and established a mathematical model to validate its performance. In order to enhance the handling efficiency of double-stack trains, Zhao et al. [23] proposed a new synchronous handling technology that focused on the design of an AGV with a loading/unloading function and a new articulated flatcar while also optimizing the layout of the terminal yard. Based on the concept of sharing yard, Liu et al. [2] formulated a flexible YC scheduling model aimed at minimizing the mutual interference in the combined operations of railway and road containers, by considering time windows related to train schedule and uncertainty of ET arrivals.

2.3. Research Overview

To enhance the clarity of our presentation and better position our contribution within the existing body of research, Table 1 provides a comprehensive synthesis of the key characteristics identified in sea–rail intermodal studies addressing collaborative scheduling of multiple equipment.

Through a thorough review of the literature, significant progress has been made in the collaborative scheduling of various equipment and the organization of container transshipment in sea–rail intermodal terminals, providing valuable references for this study. However, the following limitations remain:

- (1) Existing studies on sea–rail intermodal collaborative scheduling are all limited to the port–railway integration mode, often neglecting time window constraints such as train arrival/departure schedules and the latest vessel loading deadlines for export containers. Additionally, only a few studies have incorporated both loading and unloading processes into a single mathematical model;
- (2) The port–railway separation mode in sea–rail intermodal systems has not yet been explored. As is well-known, the long distances in this mode lead to extended travel times for external trucks (ETs), introducing uncertainties in collaborative scheduling and complicating equipment coordination compared to the port–railway integration mode;

- (3) Studies on container transshipment in sea–rail intermodal systems have mainly focused on transshipment volumes between trains, storage yards, and vessels, without in-depth consideration of specific operational handling times and sequences across equipment.

To address these issues, this paper investigates the collaborative scheduling of RMGs, ETs, and YCs in the sea–rail intermodal system under the port–railway separation mode, incorporating both loading and unloading processes and practical time constraints (i.e., train schedules and export container deadlines). An MIP model is established to minimize both the maximum makespan and the total ET empty-load time, and an improved GA based on the “First Accessible Machinery” rule is designed to solve it. After verifying the correctness of the proposed MIP model and the effectiveness of the GA, this paper explores the influence of objective weights, equipment configurations, port–railway distances, and train time windows, aiming to provide managerial insights for decision-making to improve the overall efficiency of multiple equipment collaborative scheduling under the port–railway separation mode.

Table 1. Comparison of sea–rail intermodal collaborative scheduling problems.

Paper	Port–Railway Mode	Equipment Involved					Time Window	Objective	Solution Method	Max Tasks	Time (s)
		QC	IV	ET	YC	RMG					
Yan et al. [12]	Integration		✓		✓	✓	—	M, W	GA	240	—
Chang et al. [8]	Integration	✓	✓			✓	—	M	Heuristic	—	—
Li et al. [13]	Integration		✓			✓	—	M	GA	80	—
Liu et al. [6]	Integration		✓		✓	✓	TTW	M, T, W	ALNS	230	7265
Li et al. [14]	Integration	✓	✓		✓	✓	—	M, W, E	INSGA-II	1500	—
Xia et al. [7]	Integration		✓		✓	✓	—	T	Heuristic	42	3302
Yang et al. [4]	Integration		✓		✓	✓	—	M, E	GA	240	-
Li et al. [5]	Integration		✓		✓	✓	—	T	Heuristic	1000	-
This study	Separation			✓	✓	✓	TTW, ETD	M, T, W	GA	300	5045

IV: IT, AGV, IGV; TTW: train time window; ETD: export container delivery deadline; M: makespan; E: energy consumption; T: container transit time, operation time of RMGs, IV empty travel time, ET empty-load time; C: penalty costs of constraint violation; W: waiting time of equipment; Time: solving time for the maximum number of tasks.

3. Problem Description and Formulations

In this section, we provide a detailed explanation of the problem discussed in this paper, followed by a mathematical model.

3.1. Problem Description

The collaborative scheduling problem presented in this study is defined as follows: a sea–rail intermodal system in which the port terminal is geographically separated from the railway hub (as shown in Figure 1), coupled with the lack of direct rail connectivity, necessitates container transshipment between YCs in the port yard and RMGs in the railway hub via short-distance haulage by ETs (referred to as container drayage). The temporal interdependencies among RMGs, ETs, and YCs create nonlinear propagation effects, and thus, any scheduling deviation might lead to a significant decline in overall efficiency. Therefore, this paper focuses on the collaborative scheduling problem of RMGs, ETs, and YCs for the synchronized loading and unloading operations of import containers and export containers. The objective of the problem is to determine the task assignments and operation sequences for relevant equipment, while minimizing the total operating time of all tasks and total empty-load time of dispatched trucks. The planning horizon is typically one day.

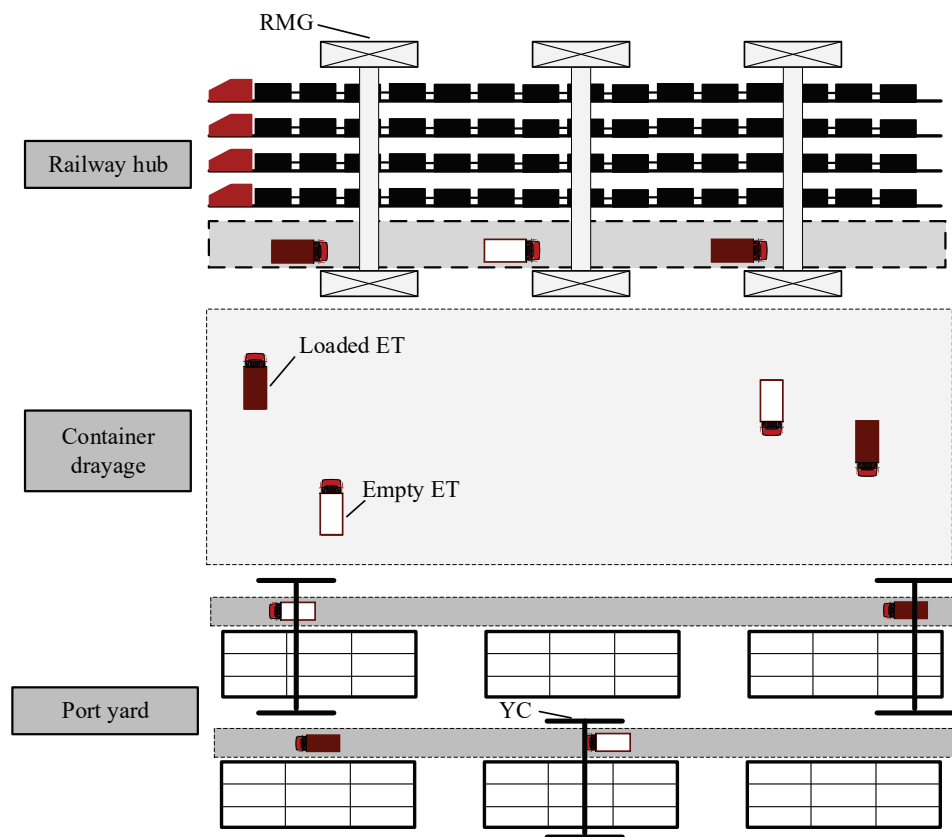


Figure 1. The layout of railway hubs and container yards in the sea–rail intermodal system with port–railway separation.

The essence of this paper is to complete all container transshipment tasks at the earliest possible time within defined time windows using a set of RMGs, ETs, and YCs. A task may refer to the unloading of an export container from the train to the port yard or the loading of an import container onto the train from the port yard. For export containers, upon arrival at the railway operation zone via inbound trains, they need to be unloaded by RMGs before the train departs, and then transferred onto assigned ETs. Next, they will be transported to designated positions at the port yard where YCs stack them at specified slots before the latest allowable vessel loading time. The process for import containers is exactly the opposite. Due to the spatial separation between the port terminal and railway hub, the extended travel distance results in longer travel times for containers transported between them via ETs. This creates a significant challenge in efficiently dispatching ETs to minimize the waiting times of RMGs and YCs, thereby improving the continuity of operations, all while being constrained by the limited number of available ETs. While the strategic pairing of import and export container tasks can somewhat enhance ET efficiency, the coupling between ETs and the operations of RMGs and YCs necessitates the simultaneous consideration of their scheduling within the model. This complexity, along with the increased computational requirements, not only escalates the difficulty of the modeling process but also places more stringent demands on solving the problem.

3.2. Mathematical Formulation

3.2.1. Assumptions

To streamline the problem under study, we make the following assumptions, adapted from Yan et al. [12] and Li et al. [14], which address a similar problem:

- (1) Within the decision-making horizon, YCs exclusively perform container retrieval and stacking for ET loading and unloading purposes. Container relocation operations and housekeeping activities within the yard are excluded from consideration;
- (2) The arrival and departure times of trains are known in advance, along with the latest vessel loading deadlines for export containers. Moreover, precise wagon positions for railcar loading and yard slot locations for container stacking are known a priori;
- (3) All equipment operates at constant speeds, excluding acceleration or deceleration during startup or turning, as well as additional travel time due to road congestion or other factors. Each type of equipment is capable of handling a single 40 ft container per operation, with identical container sizes required for all transport tasks;
- (4) Containers are loaded on trains exclusively in single-layer configurations.

3.2.2. Notations and Parameters

G	Set of RMGs, $G = \{1, \dots, g, \dots\}$
Y	Set of YCs, $Y = \{1, \dots, y, \dots\}$
K	Set of ETs, $K = \{1, \dots, k, \dots\}$
C_1	Set of import containers
C_2	Set of export containers
C	Set of all containers, $C = C_1 \cup C_2$
o	Index of virtual initial tasks
f	Index of virtual terminal tasks
i, i'	Index of containers, $i, i' \in C \cup \{o\} \cup \{f\}$
$t_{ii'g}^G$	Travel time of RMG g between the railside handling positions of container i and container i'
$t_{ii'y}^Y$	Travel time of YC y moving between the yard slot positions of container i and container i'
$t_{ii'k}^K$	Travel time of ET k between the railside handling position (or yard slot position) of container i and the yard slot position (or railside handling position) of container i' , if both containers i and i' are import (or export) containers; otherwise, travel time of ET k between the railside handling positions (or yard slot positions) of containers i and i' , if container i is an import (or export) container and container i' is an export (or import) container
t_{ig}^{GH}	Handling time of RMG g for container i
t_{iy}^{YH}	Handling time of YC y for container i
t_{ik}^K	Travel time of ET k transporting container i between the yard storage position and the railside handling position
T_{ir}^{AR}	Arrival time of train r carrying container i
T_{ir}^{DP}	Departure time of train r carrying container i
T_{is}^S	The latest permitted loading time of vessel s for container i
t	The minimum operational time interval between YCs operation in the same area
α_i	A binary parameter which takes value 1 if container i is an import container, and 0 otherwise
$\beta_{ii'}$	A binary parameter which takes value 1 if containers i and i' are loaded/unloaded within the same operational area of the yard, and 0 otherwise
$\gamma_{ii'}$	A binary parameter which takes value 1 if containers i and i' are loaded/unloaded in the same wagon of the same train, and 0 otherwise
λ_1	The objective weight coefficient assigned to the makespan
λ_2	The objective weight coefficient assigned to the empty-load time of ETs
M	A large positive integer

3.2.3. Decision Variables

$x_{ii'g}$	A binary variable that equals 1 if containers i and i' are both assigned to RMG g , and 0 otherwise
$y_{ii'y}$	A binary variable that equals 1 if containers i and i' are both assigned to YC y , and 0 otherwise
$z_{ii'k}$	A binary variable that equals 1 if containers i and i' are both assigned to ET k , and 0 otherwise
x_{ig}	A binary variable that equals 1 if container i is handled by RMG g , and 0 otherwise
y_{iy}	A binary variable that equals 1 if container i is handled by YC y , and 0 otherwise
z_{ik}	A binary variable that equals 1 if container i is handled by the truck k , and 0 otherwise
T_{istart}	The time when the container i starts to be handled
T_{iend}	The time when the handling of container i finishes
$T_{igstart}^G$	The time when RMG g starts handling container i
$T_{iystart}^Y$	The time when YC y starts handling container i
$T_{ikstart}^K$	The time when ET k starts handling container i
T_{igend}^G	The time when RMG g finishes handling container i
T_{iyend}^Y	The time when YC y finishes handling container i
T_{ikend}^K	The time when ET k finishes handling container i
T_{igk}^{GW}	The time when RMG g waits for ET k at the railside handling position of container i
T_{iyk}^{YW}	The time when YC y waits for ET k at the yard slot position of container i
T_{iyk}^{KY}	The time when ET k arrives at the yard slot position of container i , assigned to YC y
T_{igk}^{KG}	The time when ET k arrives at the railside handling position of container i , assigned to RMG g

3.2.4. Formulation

$$\min Z = \lambda_1 f_1 + \lambda_2 f_2 \quad (1)$$

$$f_1 = \min(\max(T_{iend}, \forall i \in C)) \quad (2)$$

$$f_2 = \sum_{i \in CU\{o\} \cup \{f\}} \sum_{j \in CU\{o\} \cup \{f\}} \sum_{k \in K} (z_{ijk} \cdot (T_{jkstart}^K - T_{ikend}^K)) \quad (3)$$

$$\text{s.t.} \sum_{g \in G} \sum_{i \in CU\{o\}} x_{ii'g} = 1 \quad \forall i' \in C, i \neq i' \quad (4)$$

$$\sum_{g \in G} \sum_{i' \in CU\{f\}} x_{ii'g} = 1 \quad \forall i \in C, i \neq i' \quad (5)$$

$$\sum_{k \in K} \sum_{i \in CU\{o\}} z_{ii'k} = 1 \quad \forall i' \in C, i \neq i' \quad (6)$$

$$\sum_{k \in K} \sum_{i' \in CU\{f\}} z_{ii'k} = 1 \quad \forall i \in C, i \neq i' \quad (7)$$

$$\sum_{y \in Y} \sum_{i \in CU\{o\}} y_{ii'y} = 1 \quad \forall i' \in C, i \neq i' \quad (8)$$

$$\sum_{y \in Y} \sum_{i' \in C \cup \{f\}} y_{ii'y} = 1 \quad \forall i \in C, i \neq i' \quad (9)$$

$$\sum_{i' \in C \cup \{f\}} x_{ii'g} = x_{ig} \quad \forall i \in C \cup \{o\}, i \neq i', \forall g \in G \quad (10)$$

$$\sum_{i \in C \cup \{o\}} x_{ii'g} = x_{i'g} \quad \forall i' \in C \cup \{f\}, i \neq i', \forall g \in G \quad (11)$$

$$\sum_{i' \in C \cup \{f\}} y_{ii'y} = y_{iy} \quad \forall i \in C \cup \{o\}, i \neq i', \forall y \in Y \quad (12)$$

$$\sum_{i \in C \cup \{o\}} y_{ii'y} = y_{i'y} \quad \forall i' \in C \cup \{f\}, i \neq i', \forall y \in Y \quad (13)$$

$$\sum_{i' \in C \cup \{f\}} z_{ii'k} = z_{ik} \quad \forall i \in C \cup \{o\}, i \neq i', \forall k \in K \quad (14)$$

$$\sum_{i \in C \cup \{o\}} z_{ii'k} = z_{i'k} \quad \forall i' \in C \cup \{f\}, i \neq i', \forall k \in K \quad (15)$$

$$x_{ii'g} + x_{i'ig} \leq 1 \quad \forall i, i' \in C \cup \{o\} \cup \{f\}, \forall g \in G, i \neq i' \quad (16)$$

$$y_{ii'y} + y_{i'iy} \leq 1 \quad \forall i, i' \in C \cup \{o\} \cup \{f\}, \forall y \in Y, i \neq i' \quad (17)$$

$$z_{ii'k} + z_{i'ik} \leq 1 \quad \forall i, i' \in C \cup \{o\} \cup \{f\}, \forall k \in K, i \neq i' \quad (18)$$

$$T_{iystart}^Y \geq (T_{i'yend}^Y + t_{i'iy}^Y) y_{i'iy} + T_{i'ik}^{TW} \cdot z_{ik} \quad \forall i \in C, \forall i' \in C \cup \{o\}, \forall y \in Y, \forall k \in K, i \neq i' \quad (19)$$

$$T_{igstart}^G \geq (T_{i'gend}^G + t_{i'ig}^G) x_{i'ig} + T_{igk}^{GW} \cdot z_{ik} \quad \forall i \in C, \forall i' \in C \cup \{o\}, \forall g \in G, \forall k \in K, i \neq i' \quad (20)$$

$$T_{fkend}^K \geq z_{ifk} \cdot (T_{ikend}^K + t_{ifk}^K) \quad \forall i \in C \cup \{o\}, \forall k \in K \quad (21)$$

$$y_{iy} (\alpha_i (T_{istart} - T_{iystart}^Y) + (1 - \alpha_i) (T_{iend} - T_{i'yend}^Y)) = 0 \quad \forall i \in C, \forall y \in Y \quad (22)$$

$$x_{ig} (\alpha_i (T_{iend} - T_{igend}^G) + (1 - \alpha_i) (T_{istart} - T_{igstart}^G)) = 0 \quad \forall i \in C, \forall g \in G \quad (23)$$

$$T_{i'ik}^{TW} = \max(0, y_{i'iy} (T_{ik}^{KY} - (T_{i'yend}^Y + t_{i'iy}^Y) y_{i'iy})) \quad \forall i \in C, \forall i' \in C \cup \{o\}, \forall y \in Y, \forall k \in K, i \neq i' \quad (24)$$

$$T_{igk}^{GW} = \max(0, x_{i'ig} (T_{ik}^{KG} - (T_{i'gend}^G + t_{i'ig}^G) x_{i'ig})) \quad \forall i \in C, \forall i' \in C \cup \{o\}, \forall g \in G, \forall k \in K, i \neq i' \quad (25)$$

$$T_{ik}^{KY} - \alpha_i ((T_{i'ikend}^K + t_{i'ik}^K) z_{i'ik}) - (1 - \alpha_i) ((T_{ikstart}^K + t_{ik}^K) z_{ik}) \geq 0 \quad \forall i \in C, \forall i' \in C \cup \{o\}, \forall k \in K, i \neq i' \quad (26)$$

$$T_{ik}^{KG} - \alpha_i((T_{ikstart}^K + t_{ik}^K)z_{ik}) - (1 - \alpha_i)((T_{i'kend}^K + t_{i'ik}^K)z_{i'ik}) \geq 0 \quad \forall i \in C, \forall i' \in C \cup \{o\}, \forall k \in K, i \neq i' \quad (27)$$

$$y_{iy}(T_{iyend}^Y - T_{iystart}^Y - t_{iy}^{YH}) = 0 \quad \forall i \in C, \forall y \in Y \quad (28)$$

$$x_{ig}(T_{igend}^G - T_{igstart}^G - t_{ig}^{GH}) = 0 \quad \forall i \in C, \forall g \in G \quad (29)$$

$$z_{ik}(T_{ikend}^K - T_{ikstart}^K - t_{ik}^K) \geq 0 \quad \forall i \in C, \forall k \in K \quad (30)$$

$$(\alpha_i(T_{ikstart}^K - T_{iyend}^Y) + (1 - \alpha_i)(T_{iystart}^Y - T_{ikend}^K)) \cdot (z_{ik} \cdot y_{iy}) = 0 \quad \forall i \in C, \forall y \in Y, \forall k \in K \quad (31)$$

$$(\alpha_i(T_{igstart}^G - T_{ikend}^K) + (1 - \alpha_i)(T_{ikstart}^K - T_{igend}^G)) \cdot (x_{ig} \cdot z_{ik}) = 0 \quad \forall i \in C, \forall g \in G, \forall k \in K \quad (32)$$

$$(T_{ikstart}^K - \alpha_i \cdot T_{i'gstart}^G - (1 - \alpha_i)T_{i'ystart}^Y)z_{i'ik} \geq 0 \quad \forall i' \in C \cup \{o\}, \forall i \in C, \forall k \in K, \forall g \in G, i \neq i' \quad (33)$$

$$T_{istart} - T_{i'end} + t_{ig}^{GH} + t_{i'g}^{GH} \leq M(1 - \gamma_{ii'}) \quad \forall i \in C_2, \forall i' \in C_1 \quad (34)$$

$$\beta_{ii'}(|T_{iystart}^Y - T_{i'y'start}^Y| - t \cdot y_{iy} \cdot y_{i'y'}) \geq 0 \quad \forall i, i' \in C, \forall y, y' \in Y, i \neq i' \quad (35)$$

$$T_{ir}^{AR} \leq T_{iend} \leq T_{ir}^{DP} \quad \forall i \in C_1 \quad (36)$$

$$T_{ir}^{AR} \leq T_{istart} \leq T_{ir}^{DP} \quad \forall i \in C_2 \quad (37)$$

$$T_{ir}^{AR} \leq T_{iend} \leq T_{is}^S \quad \forall i \in C_2 \quad (38)$$

The total objective of the model is given in Function (1), which aims to minimize the maximum makespan of all tasks (f_1) and the total empty-load time of ETs (f_2). Note that the total empty-load time of ETs consists of empty travel time between two consecutive tasks and waiting time at the next task position until operation begins. Constraints (4)–(9) ensure that each RMG, YC, and ET can have only one predecessor or successor task for the current task. Also, each container task must be performed by only one RMG, one ET, and one YC. Constraints (10) and (11) state that if containers i and i' are adjacent predecessor and successor tasks in the operational sequence of RMG g , both tasks must be exclusively handled by the same RMG g . Constraints (12) and (13) specify that if containers i and i' are adjacent predecessor and successor tasks in the operational sequence of YC y , both tasks must be exclusively handled by the same YC. Constraints (14) and (15) require that if containers i and i' are adjacent predecessor and successor tasks in the operational sequence of ET k , both tasks must be exclusively handled by the same ET. Constraint (16)–(18) specifies that if containers i and i' are predecessor–successor tasks in the operational sequence of a RMG, a YC or an ET, they can only be performed in one specific sequence. Constraints (19) and (20) specify the start time of operations for YC y and RMG g for container i . Constraint (21) specifies the time when the truck k returns

to the destination. Constraints (22) and (23) state that for import containers, the start time of the container handling operation matches the start time of YC, and the finish time matches the finish time of RMG, while for export containers, the start time of the container handling operation matches the start time of RMG, and the finish time matches the finish time of YC. Constraints (24) and (25) represent the waiting time of YC and RMG at the container's handling position until the corresponding ET arrives. Constraints (26) and (27), respectively, represent the time at which the ET arrives at the corresponding yard storage position and railside handling position when container i is either an import or export container. Constraints (28)~(30), respectively, specify the completion time of container i handling operation by YC y , RMG g , and ET k . Constraints (31) and (32) governed the operation time sequence that should be followed when the RMG, YC, and ET handle the same container i . Constraint (33) indicates that when the container truck handles containers i' and i in sequence, if container i' is an import container, the start time of ET operations for container i should be later than the start time of RMG operation for container i' , while if container i' is an export container, the start time of ET operations for container i should be later than the start time of YC operation for container i' . Constraint (34) regulates that if container i is an export container and container i' is an import container, and both are in the same train wagon, container i must be unloaded first, followed by the loading of container i' . Constraint (35) states that when containers i and i' are in the same yard area, the time gap between the start of their corresponding YC operations should meet the minimum required time interval. Constraints (36) and (37) mandate that the finish time of import containers and the start time of export containers should be within the arrival and departure time window of the train carrying them. Constraint (38) requires that the finish time of export containers must be after the train's arrival at the railway hub and before the latest permitted loading time of the vessel.

4. Solution Algorithm

The model described in this study involves three main operations: train loading/unloading in the railway hub, container transportation on the road, and retrieval and storage of containers at the port yard. Import and export containers follow the reverse sequence in their workflow. Additionally, in the container transportation and handling operations, each task can be performed by multiple types of equipment, while the container transportation is essentially a pickup and delivery problem. Therefore, this problem shares the characteristics of both flexible hybrid flow shop scheduling (FHFS) [24] and pickup and delivery problems with time windows (PDPTW) [25], both of which are NP-hard, making the problem itself NP-hard. The proposed MIP model can be solved to optimality using commercial solvers, such as Gurobi, but only for small-scale instances. To overcome this challenge, we propose an improved GA to solve this problem quickly and effectively.

The GA, originally introduced by Holland [26], has become widely used for solving complex optimization problems [27,28]. The main advantage of GA is its ability to carry out parallel searches in the solution space through population-based evaluation, selection, crossover, and mutation. To improve the algorithm's performance and address deadlock issues in FHFS, we propose an enhanced GA that includes a First Accessible Machinery (FAM) rule, which is specifically tailored for the problem at hand. The detailed steps of the algorithm are shown in Figure 2.

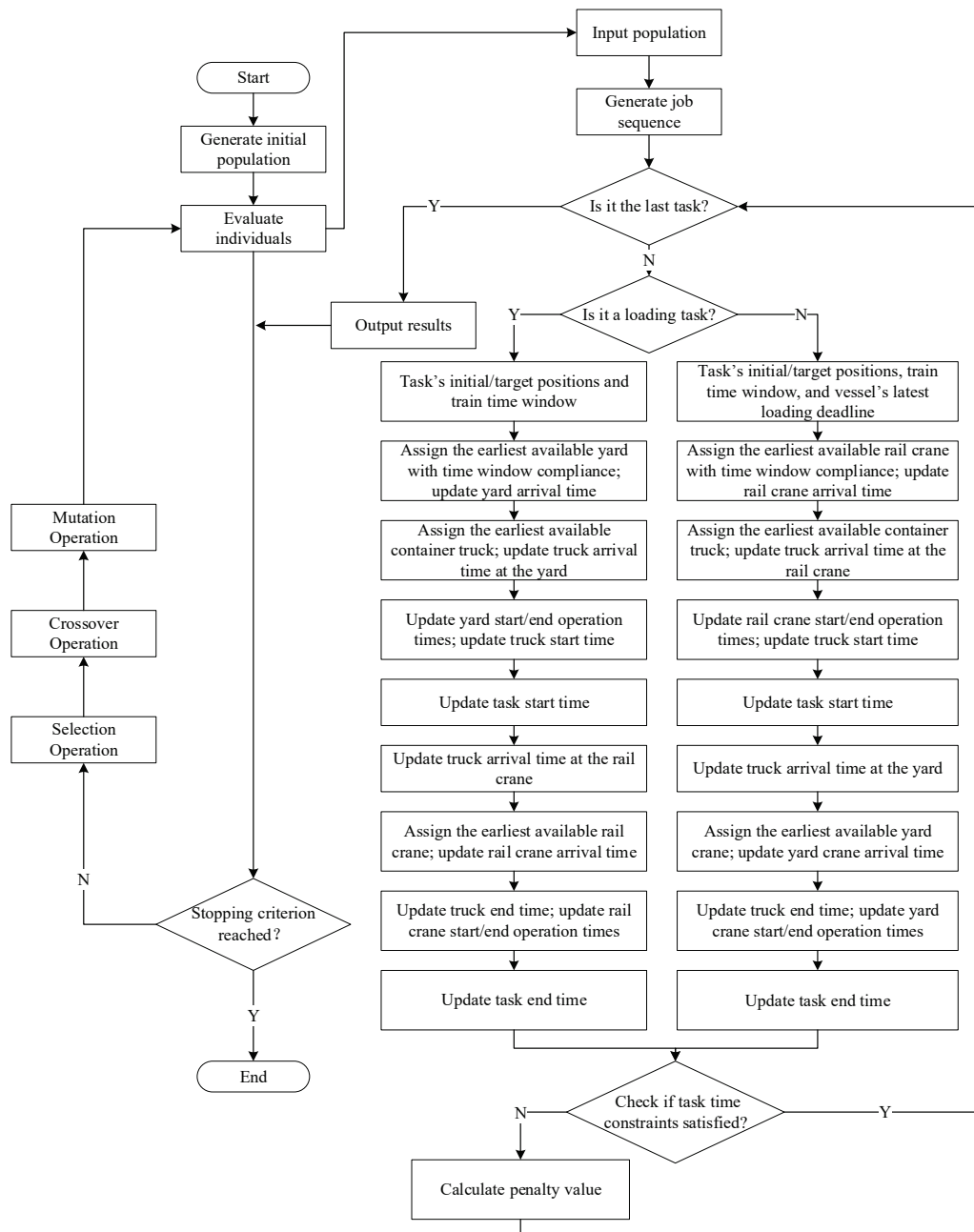


Figure 2. Flowchart of the improved GA.

4.1. Chromosome Encoding and Decoding

Tasks are encoded using a natural number encoding scheme, where the chromosome length equals the total number of tasks, i.e., the total number of containers that need to be transshipped. For example, assuming that there are 10 tasks (5 export and 5 import containers), the chromosome consists of 10 genes. The sequence of genes represents the operational order of the containers. For export containers, the unloading operation by the RMG is denoted as “D” (Discharge), while for import containers, the loading operation is denoted as “L” (Load). An exemplary chromosome encoding is illustrated in Figure 3.

Tasks	1	5	3	6	2	8	4	7	9	10
Operational types	D	D	D	L	D	L	D	L	L	L

Figure 3. Chromosome encoding.

The objective of the proposed model is to minimize the maximum makespan and truck empty-load time. To achieve this, the FAM rule is employed, which assigns RMGs, YCs, and ETs by minimizing equipment waiting times. Additionally, to meet practical operational requirements, priority is given to assignment schemes where ETs wait for RMGs and YCs. The decoding process of a task, i.e., the operational sequence of a task, is detailed using an export container i as an example:

Step 0: Based on the starting position and time window information of the export container i , select the RMG that completes its previous task earliest and can finish handling the export container i within its time window. Calculate the time when the RMG arrives at the wagon position corresponding to the export container i , denoted as T_{ig}^G , with the constraint $T_{ig}^G \leq T_{ir}^{DP}$. Initialize the start time of the export container i as $T_{istart} = T_{ir}^{AR}$.

Step 1: Initialize $S_1 = \emptyset$ and $T_1 = \emptyset$. Here, S_1 represents the set of ETs arriving at the wagon position of the container i before or at the RMG arrival time T_{ig}^G , i.e., $S_1 = \{k | T_{igk}^{KG} \leq T_{ig}^G, k \in K\}$. T_1 records the waiting time of ETs for the RMG, defined as $T_1 = \{T_{ig}^G - T_{igk}^{KG} | k \in S_1\}$.

Similarly, initialize $S_2 = \emptyset$ and $T_2 = \emptyset$. Here, S_2 represents the set of ETs arriving at the wagon position of container i after the RMG arrival time T_{ig}^G , i.e., $S_2 = \{k | T_{igk}^{KG} > T_{ig}^G, k \in K\}$. T_2 records the waiting time of the RMG for ETs, defined as $T_2 = \{T_{igk}^{KG} - T_{ig}^G | k \in S_2\}$.

Step 2: ET time calculation.

Step 2.1: Initialize the ET set K_0 and set the ET index $k = 1$.

Step 2.2: Calculate the arrival time of ET k at the wagon position of container i after completing its previous task.

Step 2.3: If $T_{igk}^{KG} \leq T_{ig}^G$, update $S_1 = S_1 \cup \{k\}$, $T_1 = T_1 \cup \{T_{igk}^{KG}\}$; else, update $S_2 = S_2 \cup \{k\}$, $T_2 = T_2 \cup \{T_{igk}^{KG}\}$.

Step 2.4: Set $k = k + 1$. If $k \geq |K_0| + 1$, proceed to Step 3; otherwise, repeat Step 2.2.

Step 3: The FAM rule is applied for ET assignment. If $S_1 \neq \emptyset$, select the ET k' with the earliest arrival time (i.e., $T_{ig}^G - T_{igk'}^{KG} = \max T_1$). Update the start time $T_{istart} = \max(T_{ir}^{AR}, T_{ig}^G)$ of export container i , and update the start time of the corresponding RMG. If $S_1 = \emptyset$, select the truck k'' with the shortest RMG waiting time (i.e., $T_{igk''}^{KG} - T_{ig}^G = \min T_2$). Update the start time $T_{istart} = \max(T_{ir}^{AR}, T_{igk''}^{KG})$ of export container i , and update the start time of the corresponding RMG.

Step 4: YC assignment is finally determined by sequentially updating the RMG's start/end times and the ET's start time. Calculate the ET's arrival time at the yard using Constraint (26), and then select the idle YC that arrives earliest at the container's storage area. Update the YC's start/end times, ET's end time, and the end time of export container i according to Constraints (19), (22), (24), (28), and (31).

The decoding process of import containers is reversed compared to export containers, as illustrated in Figure 2.

4.2. Selection Process

The substantial number of new individuals in the initial and newly generated populations requires an effective selection process to identify qualified parents. In this study, a tournament selection strategy is adopted as the individual selection mechanism in genetic operations. Although the original model enforces hard time windows for tasks, soft time windows are introduced during algorithm design to enhance solution feasibility and expand the explorable solution space. To avoid violations of time window constraints, a penalty term is incorporated into the fitness function, which constrains infeasible solutions and mitigates excessive deviations from time window requirements. The penalty function is formulated as follows:

$$penalty(i) = \begin{cases} penalty_factor * \left((T_{iend} - \alpha_i \cdot T_{ir}^{DP} - (1 - \alpha_i) \cdot T_{is}^S)^2 \right), & T_{iend} > T_{ir}^{DP} (T_{is}^S) \\ 0, & other \end{cases} \quad (39)$$

Herein, $penalty(i)$ denotes the penalty value of container i and $penalty_factor$ represents the penalty factor. Since the equipment assignment process already accounts for constraints on the task start time window, the time window penalty is calculated solely based on the task completion time. It should be noted that the two optimization objectives—minimizing the total makespan and minimizing the empty-loaded time of trucks—sometimes differ significantly in magnitude. To ensure the effectiveness of the fitness function, both metrics must undergo normalization. Therefore, the fitness function of an individual j is defined as follows:

$$f_j = 1 / \left(\lambda_1 \cdot \frac{f_{1j} - f_{1min}}{f_{1max} - f_{1min}} + \lambda_2 \cdot \frac{f_{2j} - f_{2min}}{f_{2max} - f_{2min}} + \sum_{i \in C} penalty(i) \right) \quad (40)$$

The fitness value of the current individual j is denoted by f_j , where λ_1 and λ_2 are the weight coefficients assigned to the makespan and empty-loaded truck time, respectively. Herein, f_{1j} represents the makespan (total completion time) of the current individual and f_{2j} corresponds to its empty-loaded truck time. f_{1max} (f_{1min}) and f_{2max} (f_{2min}), respectively, denote the maximum (minimum) values of makespan and the empty-loaded truck time in the current population. The total penalty value incurred by the current individual due to violations of time window constraints is expressed as $\sum_{i \in C} penalty(i)$.

4.3. Crossover Operation

In order to explore more solutions in the search space, thereby enhancing the algorithm's global search capability, this study employs the partially mapped crossover for individual crossover operations. During the procedure, two crossover points cx_1 and cx_2 satisfying $0 \leq cx_1 < cx_2 < L$ are randomly generated on the parent chromosomes to define the crossover interval $[cx_1, cx_2)$, where L is the chromosome length. The gene segments within this interval are directly exchanged between parent 1 and parent 2 to construct initial offspring 1 and offspring 2. A bidirectional mapping relationship is established between the genes of the two parents within the crossover interval. Subsequently, conflicting genes outside the crossover interval are repaired through iterative replacement: if a gene exists in the mapping relationship, it is replaced by the corresponding mapped gene until no conflicts remain.

As shown in Figure 4, randomly generated crossover points are two and six, resulting in the interval $[2, 6)$, i.e., from the third to the sixth genes. Initial offspring are generated as offspring 1: 1-3-1-7-5-4-2-8 and offspring 2: 3-2-5-6-4-7-8-6. A bidirectional mapping is established: for offspring 1, the mappings are 1-5, 7-6, 5-4, 4-7, while for offspring 2, the mappings are 5-1, 6-7, 4-5, 7-4. Both offspring initially contain duplicate genes, requiring

conflict resolution. For offspring 1, genes outside the crossover interval are sequentially checked by its index. At index 0 (task 1), the mapping triggers replacements following the chain $1 \rightarrow 5 \rightarrow 4 \rightarrow 7 \rightarrow 6$. The repair process of index 0 in offspring 1 terminates as task 6 is not in the mapping table, while other indices remain unchanged. The final repaired offspring 1 becomes 6-3-1-7-5-4-2-8. Similarly, for offspring 2, the conflict repair process yields the final offspring 2: 3-2-5-6-4-7-8-1.

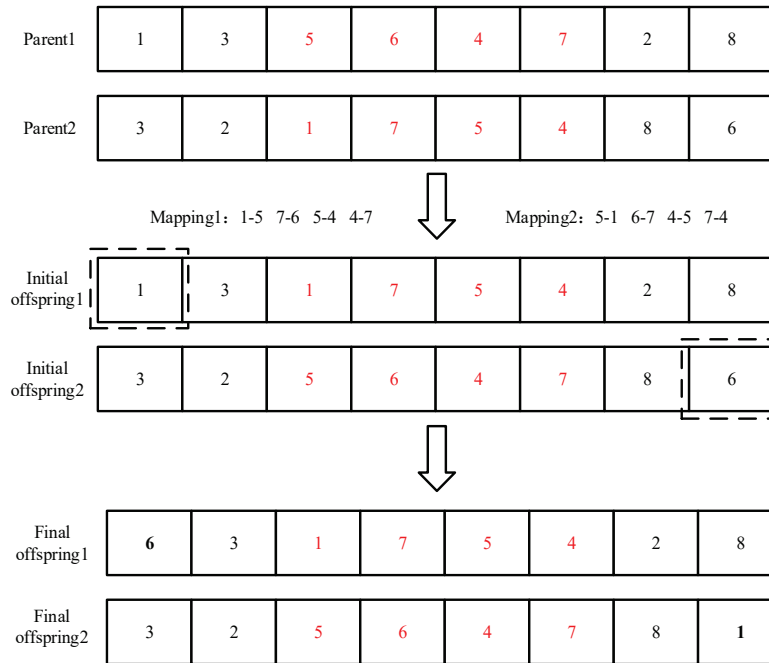


Figure 4. Schematic diagram of partially mapped crossover.

4.4. Mutation Operation

In the design of a GA, the incorporation of mutation operators enhances local search capability and maintains population diversity, thereby mitigating premature convergence to local optima. To ensure algorithmic efficacy, this study employs two types of mutation strategies: inversion mutation and swap mutation. The inversion mutation operator randomly selects two gene positions and reverses the sequence of genes between these positions, as illustrated in Figure 5. Conversely, the swap mutation operator randomly selects multiple gene positions and exchanges their positions sequentially, as demonstrated in Figure 6.

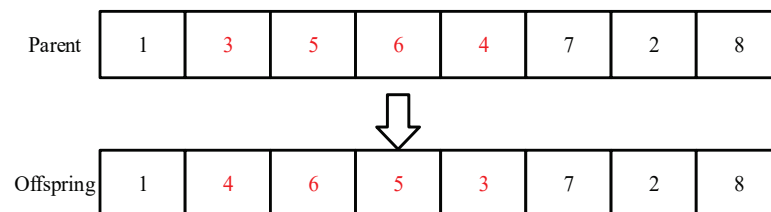


Figure 5. Schematic diagram of an inversion mutation.

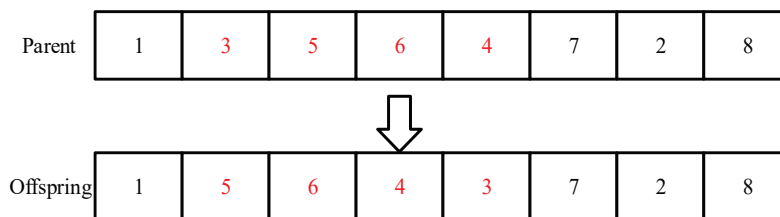


Figure 6. Schematic diagram of a swap mutation.

4.5. Stopping Criterion

The algorithm terminates and outputs the current optimization results when the iteration count reaches the predefined maximum number of evolutionary generations. Although the proposed GA, guided by heuristic rules, does not guarantee the discovery of a globally optimal solution, it provides a near-optimal solution adhering to fundamental scheduling rules for large-scale scheduling problems within a reasonable timeframe. This solution serves as a valuable reference for practical field operations.

5. Computational Experiments

This section validates and evaluates the mathematical model and GA through numerical experiments. All experiments were conducted on a desktop computer equipped with a 12th Generation Intel Core i7-12700KF CPU @ 3.60 GHz and 32.0 GB of RAM, running a 64-bit Windows 11 operating system. The mathematical model was solved using the Gurobi 10.0.1 solver invoked via PyCharm 2021.3, while the GA was implemented in Python 3.10 within the same PyCharm 2021.3 environment.

5.1. Experiment Settings and Instances Generation

To our knowledge, no benchmark instances for the studied problem are publicly available in the literature. Therefore, the experimental instances are generated randomly. Specifically, a port terminal yard and a railway hub are generated with a benchmark distance of 10 km. There are two loading/unloading tracks in the railway hub. Container tasks are randomly generated based on the storage strategies in the port terminal yard and the train schedules in the railway hub. Time windows for the tasks are randomly assigned within [0, 1440] minutes, with each time window duration set to 120 min. Containers assigned to the same train share identical time windows. For trains parked on the same loading/unloading track, the arrival time of a subsequent train is scheduled 10 min after the departure time of the preceding one. Key parameters of the mathematical model and operational equipment are summarized in Table 2 [8,15]. Since the preferences of sea–rail intermodal operators are unknown, both weighting coefficients λ_1 and λ_2 are set to 0.5 in the result analysis section.

Although the eight studies in Table 1 have already investigated problems similar to the one in this research, with the maximum task number reaching 1500, most of them do not provide solution times. Therefore, to assess the effectiveness of the proposed GA, we compare its performance against both the results obtained by directly solving the model with Gurobi and those from the ALNS algorithm in Liu et al. [6], which can handle an instance scale similar to that of this study. In our experiments, the maximum runtime is limited to 7200 s per instance. Given the complex solution space and the tendency to become trapped in local optima inherent to this problem, the parameters of the GA are configured based on preliminary experiments and problem-specific characteristics to enhance global search efficiency and convergence stability. The algorithm-related parameters are configured as follows: population size $p = 100$, maximum iteration count $G = 500$, crossover probability $pc = 0.9$, mutation probability $pm = 0.1$, and probability of selecting

inversion mutation $pw = 0.8$. Based on these settings, small-scale, medium-scale, and large-scale instances are designed and solved to validate the performance of the proposed model and the improved GA. The instances are characterized by the number of export containers and import containers. For example, 4 (2, 2) denotes an instance containing 2 export containers and 2 import containers.

Table 2. Parameters of the model and equipment.

Parameter	Value	Parameter	Value
Trolley travel speed of RMG	85 m/min	Highway travel speed of ET (loaded)	500 m/min
Gantry travel speed of RMG (unloaded)	80 m/min	Gantry travel speed of YC (unloaded)	100 m/min
Hoisting speed of RMG spreader (unloaded)	24 m/min	Trolley travel speed of YC	45 m/min
Hoisting speed of RMG spreader (loaded)	15 m/min	Hoisting speed of YC (unloaded)	35 m/min
Hoisting height of RMG	10 m	Hoisting speed of YC (loaded)	20 m/min
Intra-yard travel speed of ET (unloaded)	600 m/min	Railcar length	15 m
Intra-yard travel speed of ET (loaded)	300 m/min	Length of loading/unloading tracks	1050 m
Highway travel speed of ET (unloaded)	750 m/min	Distance between tracks and port yard	10 km

5.2. Results on Small-Scale Instances

Firstly, 10 small-scale instances are randomly generated to validate the correctness of the proposed MIP and the effectiveness of GA, with detailed results provided in Table 3. The performance metrics in the table are as follows: f_1 denotes the makespan of all tasks, $f_2^{(1)}$ represents the empty travel time of ETs between two consecutive tasks, $f_2^{(2)}$ indicates the waiting time of ETs at the next task position until operation begins, and Time specifies the solving time. $f_2^{(1)}$ and $f_2^{(2)}$ together constitute the total empty-load time of ETs f_2 . The decomposition of f_2 aims to thoroughly analyze the system's bottleneck by examining the detailed composition of the total empty-load time of ETs.

The results of instances 1 to 10 indicate that Gurobi could solve the proposed MIP model to optima or sub-optima, demonstrating the correctness of the proposed model. However, as the size of the instances increased, the computation time of Gurobi exhibits exponential growth, which indirectly suggests that the problem belongs to the NP-hard category. Specifically, Gurobi outperforms GA in terms of computation time in instances 1 and 2, while in instances 4, 6, and 8, it achieves slightly better solution quality than GA but with significantly longer computation times. From instance 10, it can be observed that when the total number of container tasks reached 16, the solution quality of Gurobi is lower than that of GA, and the computation time has reached 7200 s, implying that Gurobi is unable to obtain an exact solution within the specified time limit. This highlights that when container quantities reach 16 or higher, the escalating problem complexity renders Gurobi inadequate in balancing solution quality and computation time, making it impractical for real-world applications. As a result, the comparison between GA and Gurobi is limited to the small-scale instances.

The performance of GA is also evaluated based on the results in Table 3. Notably, for small-scale instances, the GA could produce solutions that are either identical to or closely approximate those obtained by Gurobi, while the maximum solving time is 132.53 s, much less than the 7200 s of Gurobi. In the small-scale instances, the average values of GAP1 and GAP2 between GA and Gurobi are, respectively, -1.15% and -5.91% . The reason for the lower average objective value of GA compared to Gurobi is primarily due to Gurobi's inability to find the optimal solution for instance 10, which results in a significantly poorer solution quality. In addition, although the GA's solving time for instances 1 and 2 is slightly longer than that of Gurobi, its average solving time is only 89.27 s, which is far lower than that of Gurobi. The results for small-scale instances exhibits GA's ability to find optimal or

near-optimal solutions within an acceptable solving time, validating the effectiveness and efficiency of the proposed algorithm, while showcasing superior capability in handling large-scale problems.

Table 3. Performance comparison between Gurobi and GA for small-scale instances.

Ins.	Tasks	Gurobi				GA				GAP1 (%)	GAP2 (%)
		f_1	$f_2^{(1)}$	$f_2^{(2)}$	Time (s)	f_1	$f_2^{(1)}$	$f_2^{(2)}$	Time (s)		
1	4 (2, 2)	85.02	1.70	0	2.32	85.02	1.70	0	25.82	0.00	0.00
2	5 (2, 3)	64.14	14.91	0	55.82	64.14	14.91	0	74.90	0.00	0.00
3	6 (3, 3)	69.13	2.20	0	692.20	69.13	2.20	0	86.92	0.00	0.00
4	8 (4, 4)	82.19	3.00	0	1752.01	84.99	3.30	0	52.34	3.41	10.00
5	10 (5, 5)	110.90	4.41	0	6236.00	110.90	4.41	0	119.60	0.00	0.00
6	12 (5, 7)	97.37	17.05	12.23	6525.10	98.27	18.25	13.28	78.04	0.92	7.68
7	15 (7, 8)	112.54	32.63	0	6953.26	112.54	32.63	0	128.32	0.00	0.00
8	15 (8, 7)	84.99	33.63	0	7000.20	84.99	32.63	2.45	95.35	0.00	4.31
9	15 (10, 5)	98.32	45.92	0	7051.33	98.32	45.92	0	98.86	0.00	0.00
10	16 (8, 8)	115.63	36.62	37.35	7200.00	97.33	6.62	7.35	132.53	−15.83	−81.11
Avg	—	91.29	18.00	3.74	4335.61	89.74	14.93	0.98	89.27	−1.15	−5.91

$$GAP1 = (GA(f_1) - Gurobi(f_1)) / Gurobi(f_1), GAP2 = (GA(f_2^{(1)} + f_2^{(2)}) - Gurobi(f_2^{(1)} + f_2^{(2)})) / Gurobi(f_2^{(1)} + f_2^{(2)}).$$

5.3. Results on Medium- and Large-Scale Instances and Stability Tests

Given Gurobi's inability to deliver optimal solutions for instances with more than 16 container tasks within the designated time limit, subsequent analyses focus on comparing the proposed GA and the ALNS across 10 randomly generated medium-scale instances and 10 randomly generated large-scale instances. The comparative experimental results between ALNS and GA for the medium- and large-scale instances are presented in Tables 4 and 5, respectively. Figure 7 displays the iteration convergence curves for instances 25 and 27. The figure shows that the GA converged within a relatively short time in the two instances, denoting that the proposed GA has excellent convergence properties.

The results in Tables 4 and 5 indicate that as the instance scale increases, both the objective values and the computation times of the GA and ALNS also grow accordingly, although some fluctuations are observed. For three-quarters of the 20 instances, excluding instances 16, 21, 22, 23, and 25, the maximum makespan of all tasks and the total empty-load time of ETs of GA are both superior to those of ALNS. For instances 16, 21, 22, and 25, either the maximum makespan of all tasks or the total empty-load time of ETs of GA outperforms that of ALNS. Only for instance 23 do both the objective values of GA fall short of those of ALNS. The overall average values of GAP1 and GAP2 for medium-scale instances are −0.83% and −2.69%, respectively, while for large-scale instances, they are −0.23% and −1.31%, respectively. Moreover, GA exhibits a significant advantage over ALNS in terms of computation time. When the number of container tasks reaches 270, the solving time of ALNS exceeds 7200 s, and at 300 tasks, it increases to as high as 8165.70 s. In comparison, the maximum solving time of GA is only 5044.53 s, which aligns with the algorithm's time efficiency requirements. The average solving times of GA for medium- and large-scale instances are 844.56 s and 2887.25 s, respectively, which are significantly smaller than the corresponding times of ALNS, 1519.48 s and 5354.81 s. These results demonstrate the performance advantage of the proposed GA, particularly its significant advantage in computational time.

From the column f_1 in the tables, it can be found that the expansion of instance scale exerts a moderate influence on the makespan of all tasks, though its variation remains relatively stable compared to the more pronounced fluctuations in ET's empty travel time

and waiting time. Notably, ET waiting time exhibits a significant upward trend, reflecting intensified resource competition caused by task accumulation in large-scale scenarios, which inherently leads to increased waiting time. Additionally, the increase in the number of tasks also results in a higher number of trains in the schedule. If the tasks between two adjacent trains are not properly coordinated, it leads to ET waiting. This becomes a significant factor contributing to the increased ET waiting time.

Table 4. Performance comparison between ALNS and GA for medium-scale instances.

Ins.	Tasks	ALNS				GA				GAP1 (%)	GAP2 (%)
		f_1	$f_2^{(1)}$	$f_2^{(2)}$	Time (s)	f_1	$f_2^{(1)}$	$f_2^{(2)}$	Time (s)		
11	40 (20, 20)	239.32	213.70	90.23	810.00	234.02	205.31	52.60	263.47	−2.21	−15.14
12	60 (30, 30)	340.55	293.12	260.86	869.40	338.01	282.13	271.40	343.07	−0.75	−0.08
13	70 (30, 40)	421.69	653.89	350.10	920.71	419.33	652.80	303.21	499.76	−0.56	−4.78
14	80 (40, 40)	440.25	524.30	230.45	1065.02	437.69	518.91	232.31	573.15	−0.58	−0.47
15	95 (45, 50)	549.40	449.10	550.25	1523.90	547.33	447.53	546.74	611.82	−0.38	−0.51
16	100 (50, 50)	550.31	490.56	532.50	1392.07	547.52	490.14	549.77	949.21	−0.51	1.65
17	110 (50, 60)	619.93	936.56	341.52	1696.45	615.73	933.94	330.30	1793.42	−0.68	−1.08
18	115 (60, 55)	618.69	636.25	475.00	1933.20	616.22	631.22	474.93	855.96	−0.40	−0.46
19	120 (60, 60)	649.36	590.89	600.99	2415.00	641.65	873.14	293.50	1620.43	−1.19	−2.12
20	125 (60, 65)	649.22	639.22	1006.56	2569.00	642.15	636.19	945.00	935.26	−1.09	−3.92
Avg	—	507.87	542.76	443.85	1519.48	503.97	567.13	399.98	844.56	−0.83	−2.69

Table 5. Performance comparison between ALNS and GA for large-scale instances.

Ins.	Tasks	ALNS				GA				GAP1 (%)	GAP2 (%)
		f_1	$f_2^{(1)}$	$f_2^{(2)}$	Time (s)	f_1	$f_2^{(1)}$	$f_2^{(2)}$	Time (s)		
21	130 (60, 70)	682.01	1256.20	1129.31	3023.50	679.33	1236.06	1170.02	1782.08	−0.39	0.86
22	130 (65, 65)	638.36	650.21	1536.29	3625.10	638.90	664.41	1519.50	1909.99	0.08	−0.12
23	130 (70, 60)	672.36	1229.05	1170.36	3596.41	674.93	1230.81	1174.37	1836.54	0.38	0.24
24	150 (70, 80)	724.85	1301.54	2295.01	3942.05	719.50	1219.42	2302.75	2303.77	−0.74	−2.07
25	190 (100, 90)	863.20	1802.12	4208.22	5116.00	875.42	1816.51	4110.90	2763.91	1.42	−1.38
26	190 (95, 95)	868.02	1269.45	4659.08	5027.69	867.26	1107.57	4624.51	2389.29	−0.09	−3.31
27	220 (110, 110)	889.58	1826.08	4932.56	6397.02	886.26	1814.20	4788.44	3106.89	−0.37	−2.31
28	230 (110, 120)	893.15	1864.78	7002.46	7131.01	884.17	1791.50	6858.96	3988.39	−1.01	−2.44
29	270 (140, 130)	890.58	2241.31	8936.20	7523.64	880.78	2196.52	8719.85	3747.10	−1.10	−2.34
30	300 (150, 150)	882.36	2102.88	8103.00	8165.70	877.78	2053.69	8126.05	5044.53	−0.52	−0.26
Avg	—	800.45	1554.36	4397.25	5354.81	798.43	1513.07	4339.54	2887.25	−0.23	−1.31

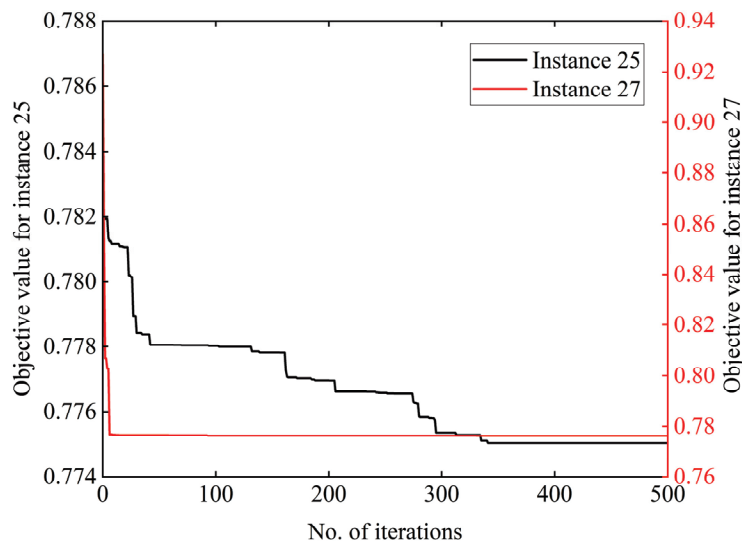


Figure 7. Convergence curve of GA.

Furthermore, the results in the columns of $f_2^{(1)}$ and $f_2^{(2)}$ for GA reveal that when the numbers of import and export containers are balanced, the empty travel time can be significantly decreased, although the waiting times increase to some extent. For instance, under the same number of tasks, due to the balanced number of import and export containers, instance 22 (26) has 571.65 (708.94) minutes less of empty travel time compared to instance 21 (25), but the waiting time only increases by 349.48 (513.61) minutes. By analyzing the detailed scheduling results, it is found that the underlying reason for this phenomenon is that task balancing can reduce the truck's empty travel time by pairing more import and export containers. Moreover, the makespan of instance 22 (26) is also 40.43 (8.16) minutes shorter than instance 21 (25). Since the fuel consumption during the truck's travel is much higher than that during waiting (whether idling or parking), it can be concluded that balancing the number of import and export containers during rail-sea intermodal transportation scheduling is critical to enhancing transfer efficiency and optimizing equipment utilization. In practice, the decision-makers should reasonably determine the transshipped number of import and export containers during a certain time period to reduce the empty-load time and improve the dispatching efficiency of ETs, rather than transporting them immediately as soon as a demand arises.

To examine the stability of the proposed GA, four instances are selected and repeatedly run 10 times, as shown in Table 6. For the makespan, the largest GAP value is 0.96% and the average GAP value is 0.30%, while for the empty-load time, the largest GAP value is 5.27% and the average GAP value is 2.56%. These results demonstrate the high stability of GA.

Table 6. Statistical results of GA from 10 computational runs.

Ins.	$\min(f_1)$	$\max(f_1)$	$\text{avg}(f_1)$	$\min(f_2)$	$\max(f_2)$	$\text{avg}(f_2)$	GAP1 (%)	GAP2 (%)
13	419.33	419.33	419.33	955.77	956.44	956.04	0.00	0.07
14	437.69	437.69	437.69	731.28	769.80	748.52	0.00	5.27
25	873.98	875.97	875.58	5882.40	5986.00	5927.40	0.23	1.76
27	885.91	894.43	888.93	6517.20	6722.40	6602.80	0.96	3.15

$$GAP1 = (\max(f_1) - \min(f_1)) / \min(f_1), \quad GAP2 = (\max(f_2) - \min(f_2)) / \min(f_2).$$

5.4. Sensitivity Analysis

Intuitively, in the spatially separated sea–rail intermodal system, factors such as equipment configuration schemes at the port yard and railway hub, the distance between port yard and railway hub, and the time window width of tasks are supposed to have an influence on the system efficiency. To investigate how these factors impact the maximum makespan of all tasks and the empty-load time of ETs, this section conducts a sensitivity analysis by comparing multiple sets of experiments.

5.4.1. Effect of Objective Weights

For the numerical experiments in Sections 5.2 and 5.3, the weights λ_1 and λ_2 are set to equal values, both being 0.5. However, it can be found that the empty-load time of ETs is obviously greater than the makespan in most instances due to the long distance between the port yard and railway hub. In real-world scenarios, the decision-makers may have varying preferences regarding the makespan and empty-load time of ETs, which can significantly influence the objective value. In order to provide valuable managerial insights, this section discusses the impact of objective weights.

The objective weights ranging from 0 to 1 are investigated for instances 14, 17, and 19, with an increment of 0.1 each time. When modifying λ_1 , λ_2 is correspondingly adjusted as $\lambda_2 = 1 - \lambda_1$. For each instance, 10 repeated runs are conducted with identical parameter

settings, and the final results for each parameter group are determined by the average value, as shown in Figure 8. As can be seen, when the weights are set to $\lambda_1 = 0.6$ and $\lambda_2 = 0.4$, the three instances all achieve relatively small values for both makespan and the empty-load time of trucks, indicating that the algorithm exhibits strong optimization capability and effectiveness at this setting. Therefore, the objective function weights are set as $\lambda_1 = 0.6$ and $\lambda_2 = 0.4$ for the experiments in the following sensitivity analysis.

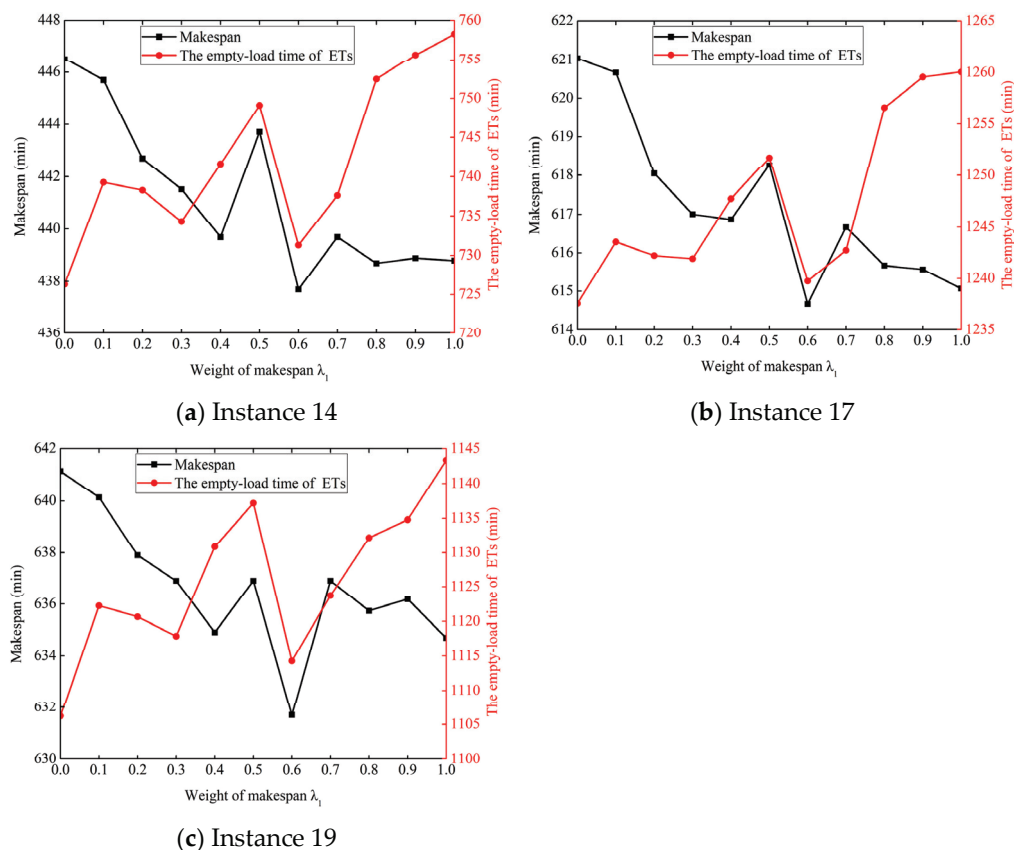


Figure 8. Experimental results with different objective weights.

5.4.2. Effect of Equipment Configuration Schemes

The effect of equipment configuration schemes on operational efficiency is investigated by analyzing randomly selected instances: 14, 19, 24, 25, and 27. For each instance, five varying equipment configurations are deployed, with detailed results presented in Table 7. The configuration in the first line of each instance is used as the initial configuration. Due to the long distance of ETs, the increased ratio of ETs is set to be three times that of RMGs or YCs. Considering the fluctuation in GA optimization results, each configuration is independently executed 10 times to ensure stability and accuracy.

It can be seen from Table 7 that increasing the number of a single equipment type or all three types can reduce the makespan of all tasks to some extent, but their effects on ET's empty travel time and waiting time display different patterns. This implies that under the initial configuration, the capability of all equipment is sufficient to complete all the container transshipment tasks. However, it does not mean that increasing the number of equipment is pointless. Compared to increasing the number of ETs, exclusively increasing the number of YCs or RMGs results in a smaller reduction in the makespan, but the reduction in ET empty travel time is more significant, and the increase in ET waiting time is also smaller. This finding reveals that the increase in handling equipment in port yards or railway hubs helps improve the dispatching flexibility of ETs. In contrast, although increasing the number of ETs can significantly shorten the makespan, this comes at the cost

of sacrificing the efficiency of the ETs, that is, increasing their waiting time. The more tasks there are, the more the waiting time of ETs increases. It should be noted that the purchase cost of different equipment varies significantly. Considering the high purchase cost of RMGs and YCs, as well as the limited impact of their quantity on reducing the makespan, increasing the number of ETs may be a more cost-efficient solution in practice, especially when the availability of RMGs or YCs is constrained.

Table 7. Effect of different equipment configuration schemes.

Ins.	RMGs/ETs/YCs	f_1	$f_2^{(1)}$	$f_2^{(2)}$
14	4/6/4	446.52	452.54	307.00
	4/9/4	435.69	397.76	1210.86
	4/6/5	437.69	346.29	406.33
	5/6/4	437.69	386.13	381.48
	5/9/5	435.68	450.89	985.39
19	4/6/4	631.69	594.14	663.52
	4/9/4	621.07	685.45	1781.67
	4/6/5	629.24	607.42	509.34
	5/6/4	629.24	501.17	580.08
	5/9/5	600.60	619.05	1704.55
24	4/10/4	717.50	913.95	2711.32
	4/13/4	713.50	978.70	4396.57
	4/10/5	715.50	900.67	2473.50
	5/10/4	715.50	887.40	2377.35
	5/13/5	712.10	912.30	4448.70
25	4/12/4	890.68	1266.95	4666.28
	4/15/4	861.55	1291.86	6836.88
	4/12/5	875.97	1041.16	5018.75
	5/12/4	875.78	1147.41	4550.59
	5/15/5	858.90	1172.32	6774.95
27	4/13/4	886.90	1442.32	5123.84
	4/16/4	881.42	1135.20	7624.37
	4/13/5	885.91	1336.07	5294.51
	5/13/4	883.42	1136.85	5551.52
	5/16/5	863.90	1493.79	7390.17

To enhance clarity, Figure 9 illustrates the trend of changes in makespan, empty travel time, and waiting time of ETs for instance 27. Although increasing the number of ETs reduces the makespan, the results of instance 27 present a significant rise in ET waiting time. For instance, when the number of ETs increases to 16, their waiting time exceeds that of the scenarios where only RMGs and yard cranes are added. The increase in waiting time primarily stems from improved operational efficiency within the same time window due to additional ETs, thereby shortening the makespan. Taking two trains on the same track as an example, the first with a time window of [0, 120] and the second with [130, 250], the initial configuration yields a maximum makespan of 120 min for the first train. With increased equipment, the makespan is shortened to 100 min. However, since the handling of tasks is required to be started once the train arrives, increased equipment extends the waiting time of ETs for the second train to 30 min, which is 20 min longer than before. To avoid prolonged waiting of ETs, decision-makers are suggested to assign an appropriate number of ETs to provide drayage operation or reasonably determine the start time of task handling.

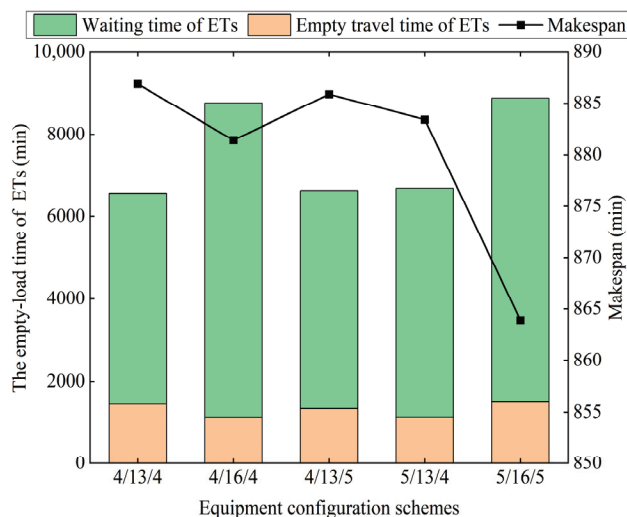


Figure 9. Effect of equipment configuration schemes, for instance, 27.

5.4.3. Effect of Distances Between the Port Yard and Railway Hub

In practical operations, the distance between the port yard and rail hub varies significantly, and this variation critically plays a vital role in transshipment efficiency. Railway hubs located near the port typically have distinct efficiency advantages during transportation. However, for railway hubs located farther away, ETs require substantially more time to complete drayage operations, thereby complicating overall equipment scheduling. Therefore, this section evaluates the effect of increasing the port–railway distance from 5 km to 30 km on transshipment efficiency based on instance 19. The results are presented in Table 8, and the trend of normalized objective values is shown in Figure 10.

Table 8. Effect of distances between the port yard and railway hub for instance 19.

Distances	RMGs/ETs/YCs	f_1	$f_2^{(1)}$	$f_2^{(2)}$
5	4/20/4	556.05	331.54	7874.32
10	4/20/4	576.04	666.10	6933.36
15	4/20/4	596.03	967.82	5812.52
20	4/20/4	616.03	1027.57	5341.99
25	4/20/4	641.03	1519.54	3906.93
30	4/20/4	670.03	1465.01	3396.85

From Table 8, it can be observed that under the same equipment configuration and task quantity, an increase in port–railway distance directly extends the makespan of all tasks. Specifically, when the distance increases from 5 km to 30 km, the makespan is prolonged by nearly 2 h. Meanwhile, it can be seen that although the ET empty travel time and waiting time are both considerable and show opposite trends, they are inevitable due to the long distance. These findings suggest that as the distance between the port yard and the railway hub increases, a reasonable arrangement of equipment quantity is necessary to minimize the makespan. Furthermore, due to the significantly greater negative effects caused by empty travel of ETs over long distances, it is even more crucial to optimize the assignment of ETs for import and export container tasks.

It is important to note that when the port–railway distance reaches 30 km, the makespan of 670.03 min fails to meet operational demands within the predefined time windows, leading to significant penalties for some tasks, as calculated in Constraint (39). This, in turn, results in an exceptionally high objective value. To maintain consistency with the other objective values, the objective value is set to 1 for the 30 km port–railway distance

in Figure 10. From this figure, it can be seen that as the distance increases, the objective values generally decrease, reaching the minimum when the distance is 25 km. However, this does not imply that longer distances from the port yard to the railway hub always result in better transshipment efficiency. Rather, it reflects that for efficient transshipment over long distances, an adequate number of initial equipment must be provided. For example, a configuration of 4/20/4, or similar, may be required. Under the 4/20/4 configuration scheme, when the distance is less than 25 km, a shorter distance results in a significant increase in ET waiting time under fixed time window widths, which far outweighs the benefits from reduced makespan and ET empty travel time, ultimately causing an increase in the overall objective value. This not only demonstrates the effect of port–railway distance but also supports the premise that the rational deployment allocation of truck quantities is crucial for specified port–railway distance.

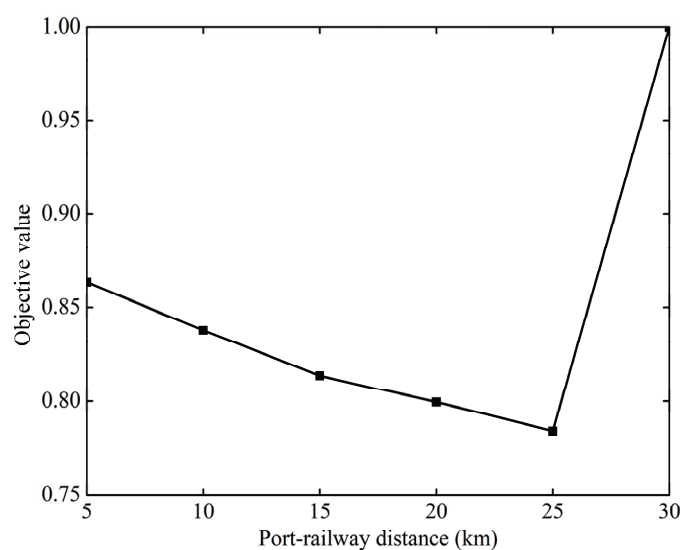


Figure 10. Objective values under different port–railway distances for instance 19.

5.4.4. Effect of Time Window Widths

Based on the analysis in Section 5.4.2, the 4/20/4 equipment configuration (4 RMGs, 20 ETs, and 4 YCs) is unable to complete all tasks for instance 19 within the predefined time window when the port–railway distance reaches 30 km, indicating that the transport capacity is insufficient under this configuration. Therefore, one may wonder whether increasing the width of the time window would allow the task to be completed successfully. To address this, the following section further examines the effect of time window width on transshipment efficiency using instance 19, under the settings of a 30 km port–railway distance and the 4/20/4 equipment configuration.

Table 9 reports the detailed results. As can be seen, the makespan of all tasks, ET empty travel time, and ET waiting time progressively increase with expanding time window width. This indicates that a wider time window enables more scheduling flexibility but sacrifices ET waiting time to reduce empty travel. Notably, the objective value indicates that a time window width of 150 min yields the optimal comprehensive result under the 4/20/4 configuration and 30 km port–railway distance. Once the time window width exceeds 150 min, sufficient equipment results in a rapid increase in ET waiting time, and thus is detrimental to the scheduling management of the overall system. Therefore, when the time window is relatively large, it is possible to reduce the dispatched equipment quantities to avoid resource waste. Figure 11 further illustrates the optimal number of ETs for instance 19 under varying time window widths at the 30 km port–railway distance. The objective

value of 1 implies that the transport capacity of ETs is insufficient to complete all tasks within predefined time windows, and it is set to 1.

Table 9. Effect of time window widths for instance 19 with a 30 km port–railway distance.

Time Window Width	RMGs/ETs/YCs	f_1	$f_2^{(1)}$	$f_2^{(2)}$	Obj.
120	4/20/4	670.02	1465.01	3396.85	333,712.600
150	4/20/4	783.72	1545.57	5374.60	0.786
180	4/20/4	896.03	1626.13	7530.03	0.791
210	4/20/4	1016.03	1948.38	9072.26	0.797
240	4/20/4	1136.03	2109.51	11,274.98	0.806

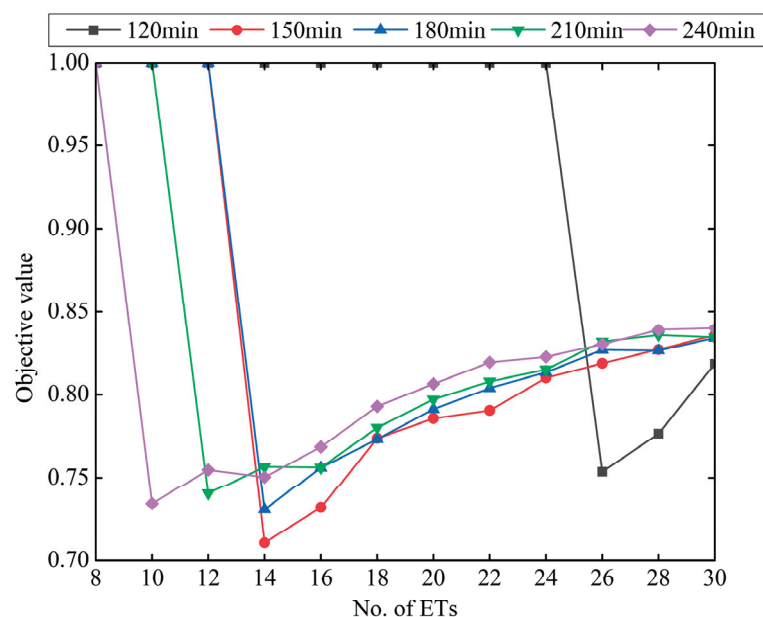


Figure 11. Effect of time window widths under various numbers of ETs for instance 19 with a 30 km port–railway distance.

Figure 11 demonstrates that different time window widths correspond to distinct optimal truck quantities. Specifically, a 120-min time window requires 26 ETs to ensure task completion, as fewer than 26 ETs fails to meet operational deadlines. For 150-min and 180-min time windows, the optimal number of ETs decreases to 14. As the time window expands further, the optimal ET count continues to decline: 12 ETs are needed for a 210-min window, and this number drops to 10 ETs for a 240-min window. These findings suggest that larger time window widths allow a reduction in the required number of ETs. Among these configurations, the 150-min time window strikes a balance, delivering favorable makespan and truck empty-load time with fewer ETs. Thus, when port–railway distances are relatively long, adjusting time window widths and scientifically planning train schedules based on the number of tasks is essential. However, excessively large time windows may reduce the number of trains that can be processed daily, undermining transshipment efficiency and potentially disrupting the overall sea–rail intermodal workflow.

6. Conclusions

This study addresses the collaborative scheduling optimization of RMGs, ETs, and YCs for container transshipment in sea–rail intermodal transportation under port–railway separation mode. An MIP model is developed to minimize the makespan of all tasks and the empty-load time of ETs, incorporating constraints such as train arrival/departure time

windows, export container delivery deadlines, and equipment operation sequences. To enhance computational efficiency, an improved GA is proposed. The results of numerical experiments at varying scales demonstrate the correctness of the model and the effectiveness of the GA. Compared to Gurobi, the GA achieves a 6.08% improvement in solution quality and a 97.94% decrease in computation time for small-scale instances. Additionally, for medium to large-scale instances, it surpasses the ALNS algorithm by 1.51% in solution quality and reduces computation time by 45.71%. Furthermore, sensitivity analysis provides valuable insights as follows: (1) The objective weights have a significant impact on makespan and ET empty-load time, and thus decision-makers should determine the appropriate weightings based on their preferences. (2) Appropriately increasing the number of ETs may be a more cost-efficient solution in practice, especially when the availability of RMGs or YCs is constrained. (3) The longer port–railway distance necessitates more ETs to ensure timely task completion within predefined time windows, which in turn makes it even more crucial to optimize the assignment of ETs for import and export container tasks. (4) Widening time windows can reduce the required number of ETs for the scenario of long port–railway distance, but excessive time window expansion may prolong makespan and compromise overall transshipment efficiency.

However, the current study has certain limitations. While it assumes pre-known information about containers requiring transportation, practical operations often encounter demand fluctuations during implementation, necessitating dynamic adjustments. For instance, port operations may experience uncertainty in container transportation demand due to factors such as vessel arrival delays, last-minute changes in container orders, or unexpected incidents. Future research should address the dynamic tasks arising from these uncertainties. Additionally, in drayage operations, ETs may encounter situations such as traffic congestion, leading to variations in speed, which can affect the timeliness of the scheduling plan. To account for this, future studies could introduce appropriate time buffers for ET travel times and further explore this cooperative scheduling problem by considering the characteristics of time-varying road networks.

Author Contributions: Conceptualization, X.Y.; methodology, X.Y. and C.H.; software, C.H.; validation, X.Y. and C.H.; formal analysis, X.Y. and C.H.; investigation, C.H.; resources, X.Y.; data curation, X.Y. and C.H.; writing—original draft preparation, C.H.; writing—review and editing, X.Y.; visualization, C.H.; supervision, X.Y.; project administration, X.Y.; funding acquisition, X.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 42001122.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zou, Y.; Xiao, G.; Li, Q.; Biancardo, S.A. Intelligent Maritime Shipping: A Bibliometric Analysis of Internet Technologies and Automated Port Infrastructure Applications. *J. Mar. Sci. Eng.* **2025**, *13*, 979. [CrossRef]
2. Liu, W.; Zhu, X.; Wang, L.; Li, S. Flexible Yard Crane Scheduling for Mixed Railway and Road Container Operations in Sea-Rail Intermodal Ports with the Sharing Storage Yard. *Transp. Res. E Logist. Transp. Rev.* **2024**, *190*, 103714. [CrossRef]
3. Seo, Y.J.; Chen, F.; Roh, S.Y. Multimodal Transportation: The Case of Laptop from Chongqing in China to Rotterdam in Europe. *Asian J. Shipp. Logist.* **2017**, *33*, 155–165. [CrossRef]
4. Yang, Y.; Sun, S.; He, S.; Jiang, Y.; Wang, X.; Yin, H.; Zhu, J. Research on the Multi-Equipment Cooperative Scheduling Method of Sea-Rail Automated Container Terminals under the Loading and Unloading Mode. *J. Mar. Sci. Eng.* **2023**, *11*, 1975. [CrossRef]
5. Li, J.; Yan, L.; Xu, B. Research on Multi-Equipment Cluster Scheduling of U-Shaped Automated Terminal Yard and Railway Yard. *J. Mar. Sci. Eng.* **2023**, *11*, 417. [CrossRef]

6. Liu, W.; Zhu, X.; Wang, L.; Zhang, Q.; Tan, K.C. Integrated Scheduling of Yard and Rail Container Handling Equipment and Internal Trucks in a Multimodal Port. *IEEE Trans. Intell. Transp. Syst.* **2024**, *25*, 2987–3008. [CrossRef]
7. Xia, T.; Wang, L.; Zhang, Q.; Dong, J.; Song, D.; Zhu, X. Integrated Scheduling of Gantry Cranes, Container Trucks and Yard Cranes in On-dock Railway Operation Areas at Multimodal Container Ports. *IET Intell. Transp. Syst.* **2025**, *19*, e12600. [CrossRef]
8. Chang, Y.; Zhu, X.; Haghani, A. Modeling and Solution of Joint Storage Space Allocation and Handling Operation for Outbound Containers in Rail-Water Intermodal Container Terminals. *IEEE Access* **2019**, *7*, 55142–55158. [CrossRef]
9. Abu-Aisha, T.; Audy, J.-F.; Ouhimmou, M. Toward an Efficient Sea-Rail Intermodal Transportation System: A Systematic Literature Review. *J. Shipp. Trade* **2024**, *9*, 23. [CrossRef]
10. Li, Y.; Sun, Z.; Hong, S. An Exact Algorithm for Multiple-Equipment Integrated Scheduling in an Automated Container Terminal Using a Double-Cycling Strategy. *Transp. Res. E Logist. Transp. Rev.* **2024**, *186*, 103565. [CrossRef]
11. Hsu, H.-P.; Chou, C.-C.; Wang, C.-N. Heuristic/Metaheuristic-Based Simulation Optimization Approaches for Integrated Scheduling of Yard Crane, Yard Truck, and Quay Crane Considering Import and Export Containers. *IEEE Access* **2022**, *10*, 64650–64670. [CrossRef]
12. Yan, B.; Zhu, X.; Wang, L.; Chang, Y. Integrated Scheduling of Rail-Mounted Gantry Cranes, Internal Trucks and Reach Stackers in Railway Operation Area of Container Terminal. *Transp. Res. Rec. J. Transp. Res. Board* **2018**, *2672*, 47–58. [CrossRef]
13. Li, W.; Wu, Z.; Yang, P.; Cai, L. Collaborative Scheduling Optimization of Equipment in Multimodal Transport Harbor Considering Hybrid Operation Mode of “Train-Yard-Vessel” and “Train-Vessel”. In Proceedings of the 2022 IEEE 18th International Conference on Automation Science and Engineering (CASE), Mexico City, Mexico, 20–24 August 2022; IEEE: Mexico City, Mexico, 2022; pp. 86–91.
14. Li, W.; Zhong, L.; He, L.; Guo, W. Integrated Scheduling in Sea–Rail Intermodal Terminals with Flexible Transshipment Routes. *Swarm Evol. Comput.* **2025**, *93*, 101852. [CrossRef]
15. Yang, Y.; He, S.; Sun, S. Research on the Cooperative Scheduling of ARMGs and AGVs in a Sea–Rail Automated Container Terminal under the Rail-in-Port Model. *J. Mar. Sci. Eng.* **2023**, *11*, 557. [CrossRef]
16. Zhao, J.; Zhu, X.; Liu, Y.; Wang, L.; Yan, B. A Practical Model for Inbound Container Distribution Organization in Rail-Water Transshipping Terminal. *J. Control Sci. Eng.* **2018**, *2018*, 9148405. [CrossRef]
17. Hu, Q.; Corman, F.; Wiegmanns, B.; Lodewijks, G. A Tabu Search Algorithm to Solve the Integrated Planning of Container on an Inter-Terminal Network Connected with a Hinterland Rail Network. *Transp. Res. C Emerg. Technol.* **2018**, *91*, 15–36. [CrossRef]
18. Yan, B.; Zhu, X.; Lee, D.-H.; Jin, J.G.; Wang, L. Transshipment Operations Optimization of Sea-Rail Intermodal Container in Seaport Rail Terminals. *Comput. Ind. Eng.* **2020**, *141*, 106296. [CrossRef]
19. Zhao, J.; Zhu, X.; Wang, L. Study on Scheme of Outbound Railway Container Organization in Rail-Water Intermodal Transportation. *Sustainability* **2020**, *12*, 1519. [CrossRef]
20. Yan, B.; Jin, J.G.; Zhu, X.; Lee, D.-H.; Wang, L.; Wang, H. Integrated Planning of Train Schedule Template and Container Transshipment Operation in Seaport Railway Terminals. *Transp. Res. E Logist. Transp. Rev.* **2020**, *142*, 102061. [CrossRef]
21. Chang, Y.; Zhu, X. A Novel Two-Stage Heuristic for Solving Storage Space Allocation Problems in Rail–Water Intermodal Container Terminals. *Symmetry* **2019**, *11*, 1229. [CrossRef]
22. Abu Aisha, T.; Ouhimmou, M.; Paquet, M.; Montecinos, J. Developing the Seaport Container Terminal Layout to Enhance Efficiency of the Intermodal Transportation System and Port Operations—Case of the Port of Montreal. *Marit. Policy Manag.* **2022**, *49*, 181–198. [CrossRef]
23. Zhao, Z.; Wang, X.; Cheng, S.; Liu, W.; Jiang, L. A New Synchronous Handling Technology of Double Stack Container Trains in Sea-Rail Intermodal Terminals. *Sustainability* **2022**, *14*, 11254. [CrossRef]
24. Teng, Y.; Li, X.; Gao, L.; Xie, J.; Li, Y. An Improved Co-Evolutionary Memetic Algorithm Based on Novel Schedule Type and Unconditional Feasibility for Hybrid Flow-Shop Scheduling Problem. *Comput. Ind. Eng.* **2024**, *193*, 110324. [CrossRef]
25. Pankratz, G. A Grouping Genetic Algorithm for the Pickup and Delivery Problem with Time Windows. *OR Spectr.* **2005**, *27*, 21–41. [CrossRef]
26. Holland, J.H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*; The MIT Press: Cambridge, MA, USA, 1992; ISBN 978-0-262-27555-2.
27. Chen, G.; Govindan, K.; Golias, M.M. Reducing Truck Emissions at Container Terminals in a Low Carbon Economy: Proposal of a Queueing-Based Bi-Objective Model for Optimizing Truck Arrival Pattern. *Transp. Res. E Logist. Transp. Rev.* **2013**, *55*, 3–22. [CrossRef]
28. Huang, Y.; Jin, Z.; Liu, P.; Wang, W.; Zhang, D. Container Drayage Problem Integrated with Truck Appointment System and Separation Mode. *Comput. Ind. Eng.* **2024**, *193*, 110307. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Enhancing Port Shipping Synergy Through Bayesian Network: A Case of Major Chinese Ports

Siqian Cheng, Jiankun Hu *, Youfang Huang and Zhihua Hu *

Institute of Logistics Science and Engineering, Shanghai Maritime University, Shanghai 201306, China; chengsiqian0003@stu.shmtu.edu.cn (S.C.)

* Correspondence: jkhu@shmtu.edu.cn (J.H.); zhhu@shmtu.edu.cn (Z.H.)

Abstract: Port shipping collaboration is vital to greener, more resilient trade, yet decisions remain siloed and uncertain. This study develops a Bayesian network model grounded in empirical data from major Chinese ports, aiming to systematically analyze and enhance port shipping collaborative capacity. The methodology integrates expert knowledge and structural learning algorithms to construct a Directed Acyclic Graph (DAG), representing complex multi-stakeholder interactions among port enterprises, shipping companies, customers, and governmental bodies. Through forward and backward probabilistic inference, the study quantifies how coordinated improvements yield substantial synergistic benefits. Five leverage points stand out: customer engagement in green supply chains, perceived service quality, port digital information integration, multilateral trading maturity, and strict policy enforcement. A newly revealed feedback loop between digital integration and enforcement extends Emerson et al.'s collaborative governance framework, highlighting “digital-era connectivity” as a critical governance dimension and offering managers a focused, evidence-based action agenda.

Keywords: Bayesian network; port shipping collaboration; digital transformation; probabilistic modeling

1. Introduction

Driven by rapid growth in trade volumes, strengthened decarbonization regulations, and accelerating maritime digital transformation, global maritime trade has entered a period of dynamic transformation [1]. As a cornerstone of the global logistics system, the port shipping industry not only facilitates efficient cross-regional movement of goods but also plays a crucial role in promoting sustainability, innovation, and resilience in global supply chains. Over the past decade, numerous global ports have focused on enhancing physical handling capacities and expanding service scopes. However, with the continuously rising complexity of international trade, these conventional approaches struggle to effectively address multidimensional challenges, such as volatile market demand and stringent environmental requirements. They are also inadequate for fostering coordinated solutions among key stakeholders. Recent research has indicated that isolated upgrades to operational components, such as adding berths or deploying standalone digital platforms, do not necessarily enhance overall system performance. Conversely, substantial evidence demonstrates that collaboration among port enterprises, shipping companies, customers, and governments yields significant and sustainable benefits [2]. These benefits include logistical process optimization, reduced carbon footprints, and improved policy coordination efficiency [3]. Such a systemic perspective aligns closely with the concept of “collaborative governance”, which emphasizes shared responsibilities among all stakeholders.

Nevertheless, promoting effective collaboration among multiple stakeholders in the port-shipping industry faces considerable challenges [4]. For instance, although port enterprises and shipping companies rely heavily on integrated scheduling and energy-efficiency management, they often lack quantifiable decision support models [5]. Meanwhile, customers are increasingly demanding greener, timelier, and more reliable services, yet existing customer engagement mechanisms remain underdeveloped [6]. While governments hold influential positions through fiscal subsidies and regulatory policy-making, insufficient recognition of stakeholders' differentiated demands hinders expected collaborative outcomes [7]. Given these complex interactions, traditional linear analysis or single-factor assessments fall short in uncovering inherent nonlinearities and uncertainties within the system.

To address these gaps, this research introduces the Bayesian network (BN)—a powerful modeling and inference tool—to analyze multi-agent collaborative networks. The objective is to explore effective strategies for enhancing collaborative capacity in the port shipping industry under dynamic conditions. Specifically, this research aims to answer the following questions:

- (1) How can BNs be utilized to construct accurate models representing the complex structure of collaborative systems within the port shipping industry?
- (2) How can probability distributions simulate the dynamic and uncertain nature of governance factors, and how can Bayesian inference be applied to reveal underlying collaboration mechanisms?
- (3) Under multi-stakeholder participation, which key factors significantly impact collaborative performance?
- (4) How can empirical data and inference results be effectively combined to propose viable optimization pathways and policy recommendations for collaborative development in the port shipping industry?

The aim of this study is to develop a BN-based approach to enhance port shipping collaborative capacity. In summary, we construct a BN model with empirical data and find that synchronized, moderate improvements across stakeholders yield greater synergy than isolated large-scale actions.

The main innovations of this research include:

- (1) **Introducing the concept of collaborative development based on a BN:** The research constructs and visualizes a multi-agent collaboration system centered around four major subsystems—port enterprises, shipping companies, customers, and governments—to clearly depict their complex interactions.
- (2) **Proposing a data knowledge-driven structural learning approach:** By integrating expert prior knowledge with empirical data, this method addresses BN limitations associated with small datasets and the risk of overfitting.
- (3) **Simulating dynamic collaborative processes using probability distributions and Bayesian inference:** Under uncertainty, this approach quantitatively assesses cumulative effects resulting from improvements at individual or combined nodes within the network.
- (4) **Providing practical value through empirical case studies:** This research conducts empirical analyses on representative Chinese ports, delivering tailored policy and managerial recommendations that offer feasible pathways toward green, digital, and collaborative industry development.

Based on these objectives and innovations, the remainder of this research is structured as follows: Section 2 reviews the relevant literature on collaborative development in the port shipping industry; Section 3 describes the BN modeling and inference methodologies

employed; Section 4 develops a collaborative capability evaluation model and conducts BN inference analysis, using major Chinese ports as empirical cases; Section 5 synthesizes results from forward and backward inferences, analyzing key influencing factors and their underlying mechanisms.; Section 6 proposes collaborative development strategies and policy implications tailored specifically to different stakeholders; and Section 7 concludes the research and outlines directions for future study.

2. Literature Review

This section synthesizes the scholarship most relevant to port shipping collaboration, highlighting the research gaps our study addresses.

2.1. Port Shipping Industry

The port shipping industry plays a critical role in global trade and regional economic development. It serves as a vital hub that integrates multiple resource flows, including goods, information, and capital. With the continuous expansion of maritime trade and concurrent drivers of decarbonization and digital transformation, demands from governments, port enterprises, and shipping companies are steadily increasing. These demands focus on infrastructure construction, innovation in operational models, and environmental protection.

Previous studies have predominantly examined the coupling relationship between port throughput capacity and shipping network efficiency. Typical topics include berth configuration strategies, automated container terminal operations, channel dredging, and environmental governance within port areas [8,9]. Traditionally, business interactions between port enterprises and shipping companies have been largely unidirectional. These interactions have emphasized improvements in cargo-handling efficiency and berth utilization while neglecting systematic considerations of customer needs and policy environments [10]. Although infrastructure expansions and technological upgrades driven by single-objective optimization can temporarily improve operational efficiency, such approaches often encounter limitations. These limitations arise from inadequate attention to external stakeholders and inefficient policy coordination across regions [11].

Addressing these challenges involves examining interaction mechanisms among shipping companies, customers, and governments. It also necessitates in-depth exploration of data sharing and information transparency. Given the simultaneous progression of digitalization and decarbonization, balancing port throughput capacity with ecological requirements has become an urgent and pivotal issue [12].

2.2. Collaborative Development in the Port Shipping Industry

The concept of collaborative development in the port shipping industry has emerged partly as a response to limitations inherent in traditional “vertical management” or “single-entity centralized decision-making” models. Existing research confirms that under rapid market fluctuations, diverse regional policies, and increasingly stringent environmental standards, unilateral actions by either governments or enterprises seldom lead to significant system-wide improvements [13]. In contrast, multi-agent collaboration can better leverage shipping networks, regulatory policies, and customer-related information. This approach opens new opportunities for resource integration, data interoperability, and innovative benefit-sharing mechanisms [14].

Drawing insights from the collaborative governance literature, this research defines the collaborative capability system within the port shipping industry into four subsystems: port enterprises, shipping companies, customers, and governments. A collaborative devel-

opment framework (Figure 1) is established to further identify critical factors within each subsystem clearly and systematically.

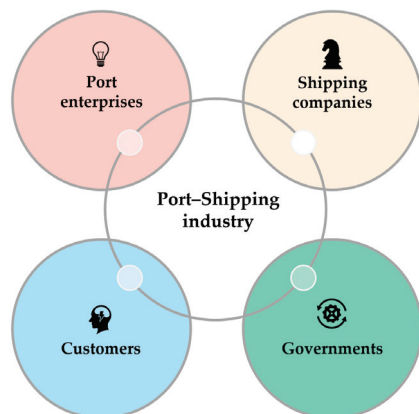


Figure 1. Collaborative development framework.

Port enterprises are widely regarded as key hubs within global logistics and maritime trade networks, performing critical functions such as cargo consolidation, warehousing, and multimodal transportation integration [15]. Previous studies at the port level have often focused on indicators like throughput capacity and operational efficiency. Conventional wisdom suggests that improving port infrastructure or expanding capacity directly enhances cargo-handling efficiency [16] and service timeliness [17]. However, given the increasing complexity of global logistics, reliance solely on infrastructure (“hard power”) is insufficient. Ports today must also address pressures from artificial intelligence applications [18], environmental compliance, and cross-regional collaboration [19]. Shipping companies typically have a deeply interdependent relationship with port enterprises. Nevertheless, effective collaboration is frequently hindered by information asymmetry or unequal benefit distribution [20]. Furthermore, because shipping companies primarily prioritize reducing operational costs [21] and maintaining schedule reliability [22], improvements in energy efficiency or service innovation are often confined internally without an effective coordination mechanism. Consequently, these internal optimizations rarely achieve significant benefits across the entire supply chain [23]. Customers represent a crucial demand-side driver for collaborative development in the port shipping industry. They encompass cargo owners and logistics service providers and ultimately extend to end consumers [24]. Research highlights that proactively responding to customers’ growing demands for greener and smarter logistics solutions can stimulate coordinated transformations across port infrastructure and shipping service models [25]. Conversely, neglecting these evolving demands could reinforce path dependency, encouraging continued reliance on traditional operational methods [26]. Governments hold a pivotal role in the port shipping industry through macro-level regulation and institutional development. They influence corporate behavior and market dynamics primarily through cross-regional policy coordination, regulatory enforcement, and financial incentives such as subsidies [27]. Previous research extensively discusses how governments provide public services to support port shipping collaboration. Typical measures include developing green shore-power infrastructure and enforcing vessel emissions control [28]. However, within complex multi-stakeholder interactions, governmental interventions risk limited effectiveness if policymakers underestimate implementation challenges faced by enterprises or overlook the diverse demands of customers. This oversight may inadvertently result in ineffective policy implementation or new administrative barriers [29]. The identification of development factors within this collaborative framework is summarized in Table 1.

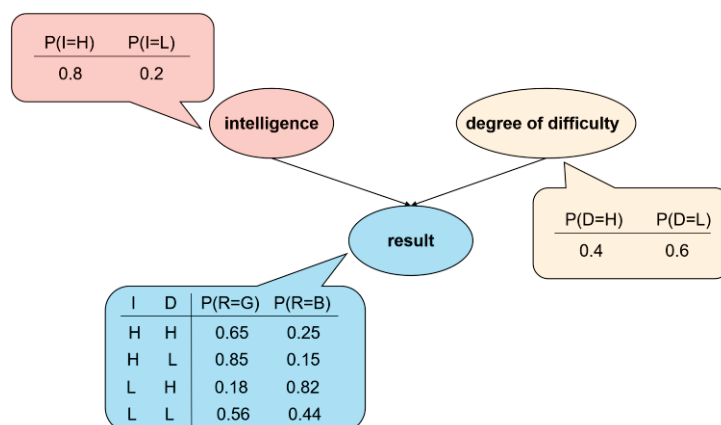
Table 1. A summary of collaborative system and factors.

Collaborative System	Collaborative Factors	Year	Ref.
Port enterprises	Cargo handling efficiency	2009	[16]
	Service timeliness	2020	[17]
	Infrastructure management and technological innovation	2022	[18]
	Environmental sustainability	2025	[19]
Shipping companies	Shipping route and capacity optimization	2014	[21]
	Ship management	2024	[22]
	Service quality	2022	[23]
	Carbon emission accounting	2025	[19]
Customers	Green supply chain	2017	[25]
	Customer engagement approach	2020	[24]
	Customer satisfaction	2022	[26]
Governments	Policy support	2023	[27]
	Public resource services	2024	[28]
	Multilateral collaboration	2024	[29]

2.3. Bayesian Network Model

A BN is an effective inference tool based on known network structures [30]. It provides a Directed Acyclic Graph (DAG), describing dependency relationships among variables [31]. This network can be represented as a binary tuple $B = \langle G, \theta \rangle$. $G = \langle V(G), E(G) \rangle$ denotes a DAG, where $V(G)$ is a set of nodes representing random variables. These variables abstract any phenomena, outcomes, or states within a given problem. $E(G)$ is a set of directed edges indicating causal dependencies between nodes. The parameter set θ specifically refers to the Conditional Probability Tables (CPTs), quantitatively expressing the dependencies among nodes via conditional probabilities.

For illustration, Figure 2 portrays a rudimentary Bayesian network with only three nodes, in which “intelligence” and “degree of difficulty” serve as the parent nodes to “result” [32]. Model development necessitates the establishment of both network topology and conditional probabilities associated with each variable, frequently grounded in empirical data [33] or expert opinion. Panahi et al. [34] utilized Bayesian networks to introduce an evaluation methodology for gauging the resilience of port infrastructure. Nonetheless, this methodology is encumbered by certain limitations, such as data intensiveness and the presupposition of variable independence.

**Figure 2.** Bayesian network structure.

2.4. Summary

In summary, there is consensus within academia and industry regarding the necessity of collaborative development in the port shipping industry. Nevertheless, existing research exhibits gaps in fully elucidating collaborative mechanisms, identifying key influencing factors, and dynamically modeling system interactions. Most studies either focus on single-entity improvements related to environmental impacts or operational efficiency, or they investigate the partial impacts of specific technological interventions. Such approaches struggle to account comprehensively for complex causal relationships and systemic uncertainties among multiple stakeholders. Methodologically, techniques such as interviews, qualitative approaches, and decision science tools demonstrate limited capacity in handling nonlinear and uncertain characteristics inherent in collaborative systems [35–37]. Furthermore, given the increasingly dynamic nature of collaborative development in the port shipping industry, there is a pressing need for research methodologies capable of capturing multi-agent interaction mechanisms, simulating evolutionary trajectories, and accurately assessing the influences of key nodes.

Addressing these identified gaps, this research employs a BN to construct a comprehensive model of the collaborative development system in the port shipping industry. The research provides an in-depth analysis of how port enterprises, shipping companies, customers, and governments interact across dimensions such as environmental compliance, digital transformation, policy frameworks, and service innovation. Through forward and backward inference analyses, critical influencing factors and evolutionary mechanisms are identified and assessed. This approach not only introduces a novel quantitative research paradigm but also offers actionable decision-making references for managers and policy-makers in the port shipping industry. These references aim to enhance resource allocation, technological upgrades, and stakeholder coordination.

3. Methodology

Section 3 outlines the methodological pipeline adopted in this study—from data collection and Bayesian network (BN) construction to parameter estimation and inference experiments. Each step is explicitly aligned with the aim of quantifying multi-stakeholder interactions under uncertainty.

3.1. Research Framework

Figure 3 illustrates the research framework developed in this research. First, through a literature review and expert interviews, the functional roles and key indicators for multiple stakeholders—including port enterprises, shipping companies, customers, and governments—are identified. An improved K2 algorithm, combined with expert refinement, is then employed to construct a DAG, representing the multi-agent collaborative mechanism. Subsequently, parameter estimation for nodes' prior and conditional probabilities is performed, using questionnaire surveys and field investigation data. This facilitates capturing multidimensional causal relationships under uncertainty via the BN. Next, forward inference is applied to analyze the collaborative development mechanism. Backward inference is conducted to identify key factors that must be prioritized to achieve high-level collaborative capability. Finally, based on empirical results and inference insights, practical recommendations are provided for industry and policy. Moreover, future research directions are discussed, considering model applicability and data limitations.

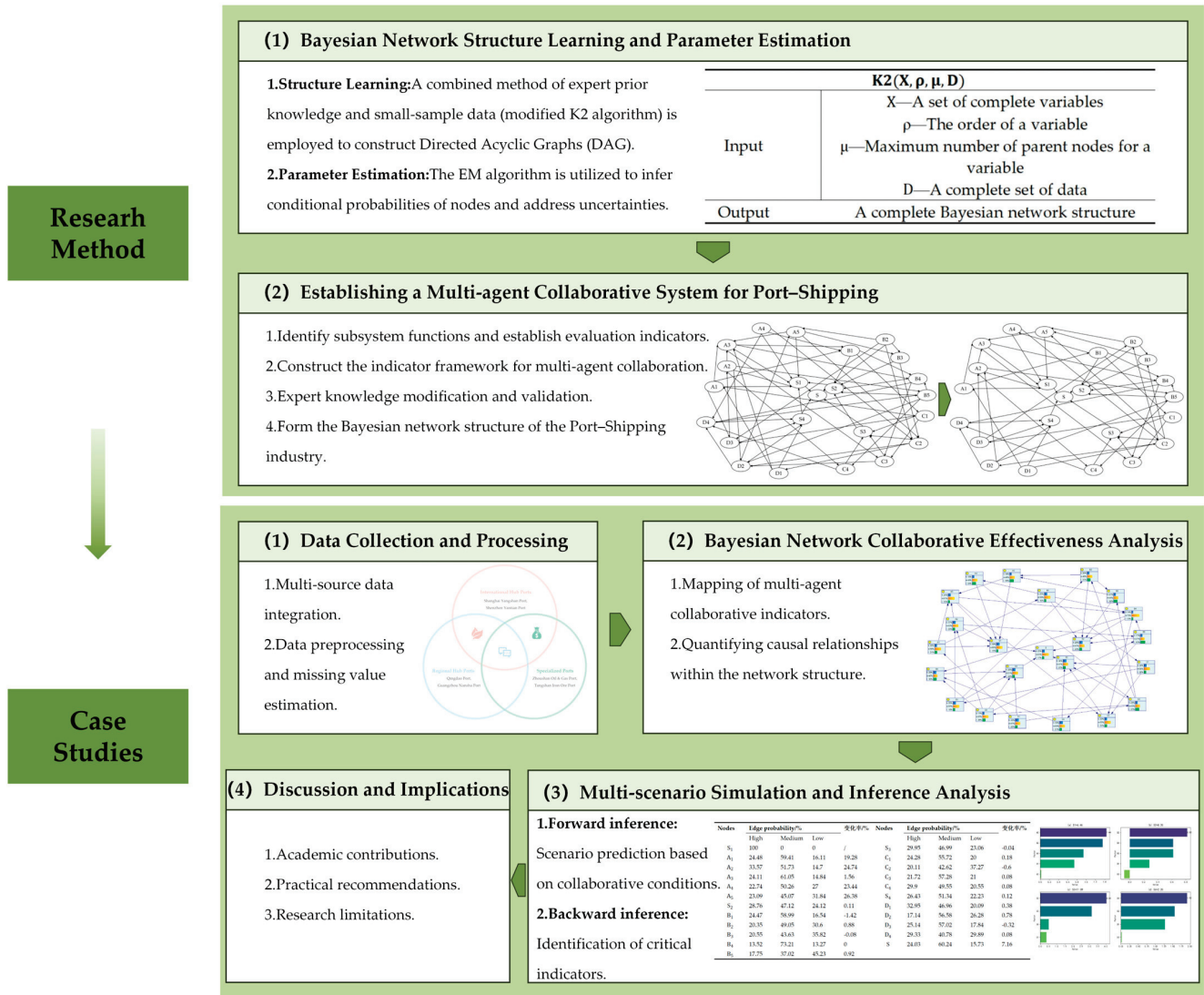


Figure 3. The developed research framework.

3.2. BN Fundamentals

Governed by Bayesian principles, the Joint Probability Distribution (JPD) of a BN comprises n variables, as described in Equation (1):

$$P(A_1, A_2, A_3, \dots, A_n) = P(A_1 | A_2, A_3, \dots, A_n) P(A_2 | A_3, A_4, \dots, A_n) P(A_{n-1} | A_n) P(A_n) \quad (1)$$

Equation (1) can be simplified in the following manner, as shown in Equation (2):

$$P(A_1, A_2, A_3, \dots, A_n) = \prod_{i=1}^n P(A_i | A_{i+1}, A_{i+2}, \dots, A_n) = \prod_{i=1}^n P(A_i | \text{Parents}(A_i)) \quad (2)$$

The internal causal mechanisms and their corresponding magnitudes within the BN are captured through the use of IF-THEN logical functions, elaborated in Equation (3).

$$R_k : \text{IF} \{C_1^k, C_2^k, \dots, C_p^k\}, \text{ THEN} \{(\beta_1^k, S_1), (\beta_2^k, S_2), \dots, (\beta_q^k, S_q)\} \quad (3)$$

In Equation (3), R_k denotes the k th rule of the system. Confidence levels, represented by β_d ($d = 1, 2, \dots, q$), are mapped to states C_d ($d \in q$), thereby converting p conditional attributes $\{C_1, C_2, \dots, C_p\}$ into q distinct states $\{S_1, S_2, \dots, S_q\}$.

In this research, particular emphasis is placed on two essential functions of BN:

- (1) Modeling Function: Decomposing and characterizing the complex collaborative system within the port shipping industry.
- (2) Inference Function: Capturing dynamic interaction mechanisms and investigating factors influencing the collaborative capability of the port shipping industry.

Compared with deterministic system dynamics simulations or black box, purely data-driven algorithms (e.g., random forest or XGBoost classifiers), BNs provide three unique advantages for multi-stakeholder port shipping problems: (i) explicit causal directionality that managers can interpret and validate; (ii) seamless fusion of sparse empirical observations with structured expert priors, which is essential when historical data are patchy or proprietary; and (iii) exact probabilistic inference that simultaneously propagates aleatory and epistemic uncertainty through the network. These properties allow BNs to capture the systemic complexity of port shipping collaboration more faithfully than alternative modeling families, even when formal performance benchmarks are not feasible within the scope of a single study.

By analyzing four primary stakeholders—port enterprises, shipping companies, customers, and governments—a BN model is constructed. The network intuitively reveals the complex internal structure of the port shipping collaborative system, enhancing comprehension and visualization. Due to inherent uncertainties and dynamic characteristics of collaboration within this industry, BNs effectively manage stochastic uncertainties arising from individual factors and their interactions via prior and conditional probabilities. Through adjusting node probabilities, the network dynamically simulates stakeholder behaviors in collaborative governance. It also quantitatively associates influencing factors through probabilistic inference, uncovering uncertain relationships. Therefore, a BN is particularly suitable for studying collaborative capability in the port shipping industry.

3.3. Hybrid Data-Knowledge Structure Learning

BN structure learning refers to the process of determining network structures using datasets [38]. Because empirical port shipping data are often sparse or proprietary, integrating expert knowledge is essential for robust BN structure learning. Accordingly, we employ a data-knowledge hybrid approach to build the DAG that represents port shipping collaborative governance. Section 4.3 provides a detailed explanation of this method.

The algorithm identifies the optimal parent node set for each node by employing a search strategy in conjunction with a scoring function, thereby yielding the most robust network architecture. The algorithm requires the specification of two parameters: the nodal sequence and the maximum allowable number of parent nodes. These constraints prevent a preceding node from acting as a parent to subsequent nodes and restrict the number of parent nodes to a predefined maximum. The intricacies of the scoring function are elaborated in Equation (4).

$$\max[P(G, D)] = C \prod_{i=1}^n \max_{\Pi^i} \left[\prod_{j=1}^{q_j} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)} \prod_{k=1}^{r_i} N_{ijk}! \right] \quad (4)$$

In Equation (4), the variable G symbolizes the network's structural framework, while D represents the dataset employed for structural learning. r_i specifies that node X_i can exist in r different states, and q_j enumerates the q possible state combinations for its parent nodes. $\prod_i$ delineates the set of parent nodes corresponding to the i th node X_i . Table 2 details the input and output variables associated with the K2 algorithm.

Table 2. Inputs and outputs of the K2 algorithm.

K2 (X, ρ, μ, D)	
Input	X—A set of complete variables
	ρ—The order of a variable
	μ—Maximum number of parent nodes for a variable
	D—A complete set of data
Output	A complete Bayesian network structure

The K2 algorithm was chosen because it strikes a practical balance between model fit and parsimony while allowing expert guidance. Constrained parent limits tame over-fitting on small datasets, and expert-defined node ordering narrows the search space, so K2 can recover plausible structures with modest computation.

3.4. Parameter Learning

The Expectation-Maximization (EM) algorithm is widely used in BN parameter learning for such incomplete data because it iteratively maximizes likelihood while inputting gaps, yielding unbiased conditional probability estimates before convergence. This research employs the EM algorithm to transform incomplete data into complete data through a two-stage process: the E-step and the M-step. In the E-step, expected values are calculated, whereas the M-step focuses on optimization. After several iterations, the algorithm converges, providing estimates for unknown parameters. Specifically, the E-step involves calculating the lower bound of the log-likelihood function. Let X represent the unknown variables, Y the observed variables, and D the training set. By defining $q(X = x|Y)$ as the probability associated with $X = x$ when Y is the observed value, we can derive Equation (5):

$$\sum_x q(X = x|Y) = 1 \quad (5)$$

Suppose the log-likelihood function L is defined as in Equation (6):

$$L(\theta) = \sum_m \log \sum_n P(X = x_i, Y = D) \quad (6)$$

Let L denote the log-likelihood function, X the set of unknown variables, Y the observed variables, and D the dataset. Assuming $P(X = x_i, Y = D)$ is a convex function with an extremum, Jensen's inequality yields Equation (7):

$$\begin{aligned} L &= \sum_m \log \sum_n q(X|Y) \times \frac{P(X = x_i, Y = D)}{q(X|Y)} \\ &\geq \sum_m \sum_n q(X|Y) \times \frac{P(X = x_i, Y = D)}{q(X|Y)} \end{aligned} \quad (7)$$

The EM algorithm effectively calculates a lower bound for the parameters during the E-step and updates them to their maximum values during the M-step. Let $q(X = x|Y) = P(X|\theta_t)$ and θ_t represent the unknown parameters in the current iteration, and θ_{t+1} signify the parameters for the next iteration. Solving for the lower bound L essentially involves determining the parameters for the next iteration under the given current parameters. Let the lower bound of L be $Q(\theta_{t+1}|\theta)$, expressed in Equation (8):

$$Q(\theta_{t+1}|\theta) = \sum_m \sum_n P(X|\theta_t) \log P(X, D|\theta_{t+1}) \quad (8)$$

In accordance with the methodology in maximum likelihood estimation, the parameters $\hat{\theta}_t$ are found when $Q(\theta_{t+1}|\theta)$ is maximized, as shown in Equation (9):

$$\hat{\theta}_t = \frac{EN_{ijk}}{EN_{ij}} \quad (9)$$

In Equation (9), EN_{ijk} represents the data in the dataset D that satisfies $X_i = x_{ik}$ and $P_a(X_i) = P_a(X_i)_j$, the data corresponding to the parent node j when node i takes its k th value.

Take node X_i with conditional probability $P(X_i|\text{Parents}(X_i))$ as an example:

(1) Frequency counting:

Count occurrences N_{ijk} of each possible state x_{ik} of node X_i under each combination of parent nodes π_{ij} .

For instance, if X_i has parent nodes A and B , count all frequencies for combinations such as $(A = a, B = b)$.

(2) Calculating conditional probabilities:

The conditional probability is estimated by normalizing the frequencies, as shown in Equation (10):

$$P(X_i = x_{ik}|\pi_{ij}) = \frac{N_{ijk} + \alpha}{\sum_{k'} (N_{ijk'} + \alpha)} \quad (10)$$

Here, α is a smoothing factor (Laplace smoothing), introduced to avoid zero probability problems. If data are sufficient, α can be set to 0.

(3) Constructing CPT:

The computed probabilities are inserted into the CPT. For example, if node X_i has 3 states and its parent nodes have 4 combinations, the resulting CPT is a 4×3 matrix.

3.5. Probabilistic Reasoning and Scenario Simulation

BN inference calculates the posterior probabilities of target nodes based on prior probabilities and available evidence. It includes two primary methods: forward inference and backward inference. When identifying critical subsystems and collaborative mechanisms between subsystems, forward inference elucidates how multiple subsystems interact and collectively influence the ultimate target. It also clarifies subsystem roles within urban community governance more quickly and accurately by altering subsystem states. Conversely, backward inference simplifies data analysis, reducing uncertainty and complexity when identifying critical parameters and conducting holistic system analysis. Backward inference begins from known targets and rapidly narrows the problem scope, making it easier to identify parameters directly influencing these targets. Thus, analyzing every parameter and their complex relationships becomes unnecessary. GeNIe is a software tool for BN modeling and inference, offering a user-friendly interface that significantly facilitates inference processes. BN inference in this research is conducted using GeNIe.

4. BN Model for Port Shipping Collaboration

The BN developed here is the study's core instrument for tackling fragmented collaboration—the key challenge identified in Section 1. By representing ports, shipping lines, customers, and regulators as interconnected nodes, the model (i) captures how decisions by any stakeholder ripple through the system and (ii) quantifies which joint actions yield the largest synergy gains. The remainder of Section 4 explains how the BN was built and how each component links to real-world governance levers.

4.1. Order Parameter Framework

4.1.1. Unpacking Synergistic Development Competencies

Collaborative development, a key paradigm for modernizing the port shipping system, illustrates how multiple stakeholders—including port enterprises, shipping companies, customers, and governments—can achieve significant improvements in system efficiency through dynamic collaboration. This is particularly important given the dual challenges of globalized supply chains and low-carbon transitions. However, ambiguity in defining the specific objectives, pathways, and practical outcomes of this development model presents substantial barriers. These barriers affect mechanism design, policy tool selection, and effectiveness evaluation in both theoretical and practical contexts.

Based on the Integrated Framework for Collaborative Governance (IFCG) proposed by Emerson et al., several scholars have attempted to analyze collaborative governance in the port shipping industry. They emphasize that collaborative capability should encompass institutionalized participation mechanisms, incentives for shared benefits, and joint action capabilities [39]. Cleave et al. further suggested that the goals of collaborative port development include governance restructuring, optimized resource allocation, and integrated economic, environmental, and social benefits [40]. Nonetheless, practical challenges—such as mismatches in stakeholder responsibilities, data barriers, or conflicting policies—often hinder the full realization of expected collaborative effects.

Drawing upon the collaborative governance framework in Figure 4 and existing research, this research defines the collaborative development capability of the port shipping industry as follows: the integrated capacity of the network governance structure—comprising port enterprises, shipping companies, customers, and governments—to achieve collaborative development, address industry challenges, optimize resource aggregation, and create public value through dynamic cross-subsystem interactions. As illustrated in Figure 4, this framework clarifies the logic and constraints of potential collaborative effects. It also provides target-oriented decision support for optimizing port shipping policies.

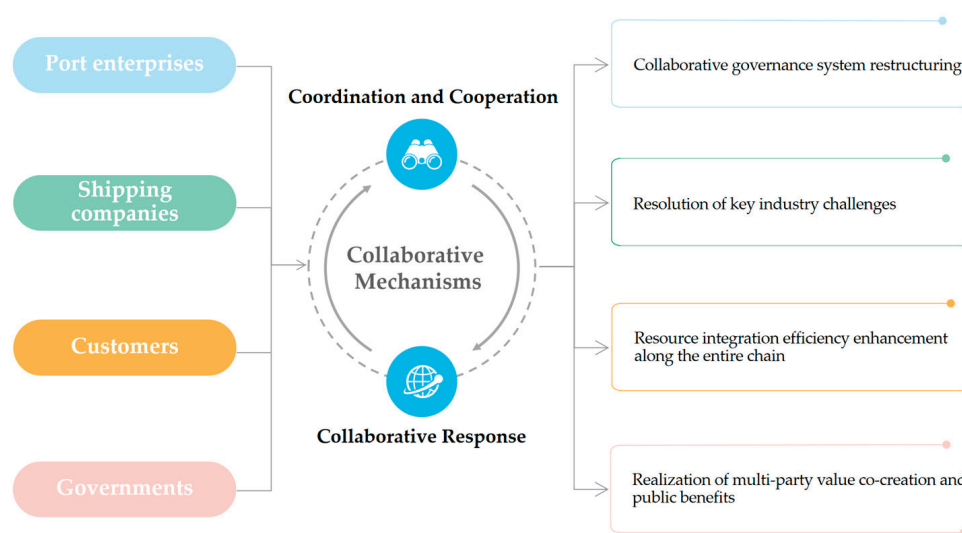


Figure 4. Collaborative interaction framework in the port shipping system.

4.1.2. Linking Order Parameters to a Graphical Model

In BN modeling, nodes represent critical factors driving collaborative capability. This research identifies influencing factors of collaborative capability through an order parameter system, mapping these factors onto network nodes. The construction process for the order parameter system is as follows:

(1) Preliminary Screening

Based on the collaborative framework of port enterprises, shipping companies, customers, and governments shown in Figure 1, and the port shipping development factors summarized in Table 1, 21 order parameters were initially extracted.

(2) Delphi Expert Knowledge Integration

A two-round Delphi survey was conducted with 15 purposely selected experts (five senior port managers, four liner-shipping strategists, three academics, and three government regulators; mean experience = 15 years). Round 1 collected open-ended factor suggestions; Round 2 asked the same panel to rate the 23 preliminary factors on a 5-point relevance scale. After two rounds, a consensus was reached. The resulting 18 factors are listed in Table 3.

Table 3. Governance factors and indicator system for port shipping collaboration.

Collaborative System	Subsystems	Order Parameters	Ref.
S	Port enterprises (S ₁)	Port infrastructure capacity (A ₁)	[41]
		Port digital information level (A ₂)	[20]
		Port operational efficiency (A ₃)	[18]
		Port diversified service capability (A ₄)	[19]
		Application rate of low-carbon technologies (A ₅)	[21]
	Shipping companies (S ₂)	Compatibility of shipping networks (B ₁)	[23]
		Ship energy efficiency management level (B ₂)	[24]
		Transparency of carbon emission accounting (B ₃)	[21]
		Reliability and timeliness of shipping services (B ₄)	[19]
		Coordination of regional blockchain transportation (B ₅)	[42]
	Customers (S ₃)	Logistics demand forecast accuracy (C ₁)	[43]
		Green supply chain participation (C ₂)	[21]
		Customer satisfaction and service quality evaluation (C ₃)	[28]
		Willingness to engage in port shipping optimization (C ₄)	[26]
	Governments (S ₄)	Cross-regional policy coordination (D ₁)	[30]
		Completeness and enforcement of policies (D ₂)	[44]
		Fiscal support and subsidy levels (D ₃)	[29]
		Maturity of multilateral trading mechanisms (D ₄)	[36]

(3) Final System Output

The finalized order parameter system for collaborative capability in the port shipping industry (S) includes 18 order parameters across four subsystems, as summarized in Table 3.

4.2. Data Acquisition and Pre-Processing Procedures

Collaborative development data in the port shipping industry were collected from multiple heterogeneous sources, primarily including expert ratings from international maritime organizations and structured data from operational annual reports of shipping companies. To ensure objectivity, a three-level scale (“Low”, “Medium”, and “High”) was employed to evaluate 23 key indicators, minimizing potential bias and subjectivity.

The research team selected 6 representative ports from three major port clusters (Yangtze River Delta, Guangdong–Hong Kong–Macao region, and Bohai Rim), as shown in Figure 5.

A total of 317 valid data samples were collected, covering multiple stakeholders, including port enterprises (23%), shipping companies (35%), logistics service providers (28%), and customs regulatory authorities (14%). A data quality control matrix and anomaly detection algorithm (Figure 6) were applied, resulting in the retention of 298 high-quality datasets with an effective rate of 94.01%. This comprehensive dataset captures the multi-scale characteristics of port shipping collaborative development, providing robust support

for subsequent modeling. BN modeling based on this dataset thus demonstrates broad applicability within the port shipping industry.

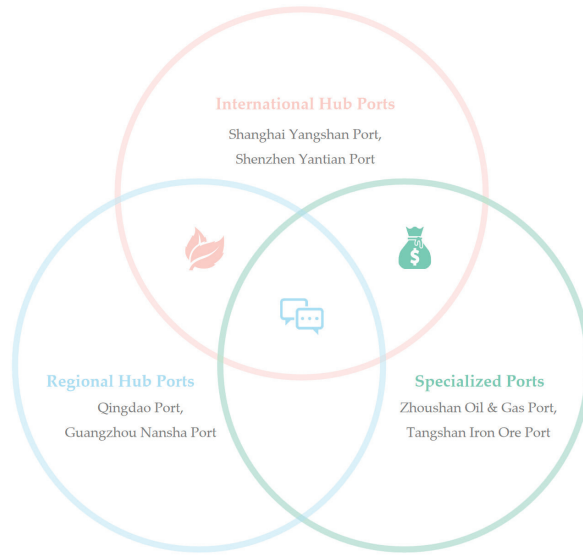


Figure 5. Representative ports.

```

1  def maritime_data_cleaning(df):
2
3      iso_forest = IsolationForest(contamination=0.05)
4      outliers = iso_forest.fit_predict(df)
5
6      df_clean = df[(outliers == 1) & (df.completeness > 0.85)]
7
8      maritime_std = MaritimeSTDChecker()
9      return maritime_std.transform(df_clean)

```

Figure 6. Anomaly detection algorithm.

4.3. Experimental Design for Structure Learning

4.3.1. Node Sequence and Prior Structural Information Design

The order parameter system of collaborative development capability in the port shipping industry also exhibits a hierarchical structure, where lower-level indicators significantly influence higher-level indicators. By extracting effective information from the order parameter system to design the node sequence and prior structural information, the structure learning performance of the K2 algorithm can be optimized, enhancing model accuracy and data utilization efficiency.

(1) Node Sequence

In the BN for port shipping collaborative capacity, allowing subsystem nodes (e.g., S_1, S_4) to directly point to order parameter nodes (e.g., A_1, B_1) conflicts with actual collaborative development logic. Consequently, this research prescribes a node-ordering approach based on the priority of these order parameters to circumvent unproductive network structure searches and thereby improve search efficiency, model accuracy, and data utilization. In practice, order parameter nodes (e.g., port operational efficiency A_3 or completeness and enforcement of policies D_2) are placed before subsystem nodes, while the target node (overall port shipping synergy S) is positioned at the same hierarchical level as the subsystem nodes (S_1, S_2, S_3, S_4) but sorted randomly. The final node sequence is thus defined as:

$$(A_1, \dots, A_5, B_1, \dots, B_5, C_1, \dots, C_4, D_1, \dots, D_4), (S, S_1, S_2, S_3, S_4), S$$

(2) Prior Structural Information

Port shipping systems often exhibit cyclical or iterative dependencies, yet a BN—being a DAG—cannot directly capture such circular relationships. To address this limitation, this research utilizes a tiered structure derived from the order parameter system (Figure 7) as the input framework when initializing the DAG via the K2 algorithm. Concretely, subsystem nodes and their corresponding order parameter nodes are assigned to different layers, with each subsystem directly depending on its own order parameter nodes. These order parameter nodes function as intermediary variables for complex inter-subsystem interactions, re-routing such dependencies to the order parameter layer, thus mitigating circular reliance and streamlining the network’s overall structure.

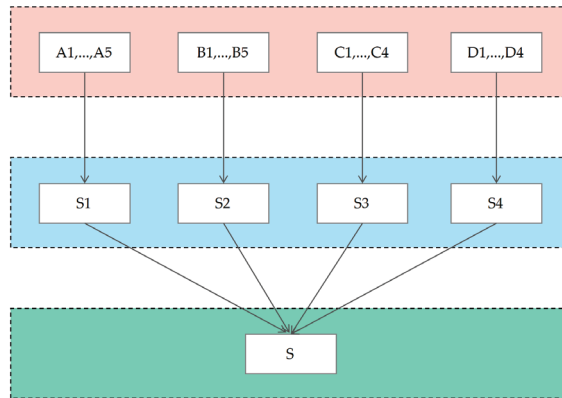


Figure 7. Tiered structure derived from the order parameter system, used to initialize the Bayesian network.

4.3.2. Improved Design for Node Ordering and Prior Structural Information

Traditional K2 algorithms cannot directly incorporate prior structural information into their search strategy, nor can they fully utilize expert insights. To integrate expert knowledge with data-driven approaches more effectively, this research refines both the algorithm’s input and search strategy, proposing the “K2-PSI” (K2 with Prior Structure Integration) method. This approach resolves the issue of representing iterative dependencies when constructing a BN for port shipping collaborative capacity. Specifically:

(1) Input Enhancement: Layered Prior Structure Embedding

Let the node set $Z = X_1, X_2, \dots, X_{23}$ correspond to the 23 order parameters, each having three states—low, medium, and high ($r_i = 3$). The dataset D comprises $m = 298$ records. Each core node can have up to five parent nodes, whereas auxiliary nodes are capped at three parents.

At the input stage, the prior tiered structure is introduced to the K2 algorithm in the form of “prior node sets.” Denote by PN_i the set of prior order parameter nodes for a given subsystem i . As illustrated in Figure 7, $PN_{S_1} = [A_1, A_2, A_3, A_4, A_5]$, $PN_{S_2} = [B_1, B_2, B_3, B_4, B_5]$, $PN_{S_3} = [C_1, C_2, C_3, C_4]$, and $PN_{S_4} = [D_1, D_2, D_3, D_4]$ serve as “preceding” node groups for S_1, S_2, S_3 , and S_4 , respectively. Cross-subsystem, or “overstepping”, pointers are disallowed to prevent violations of the DAG’s acyclicity constraint.

(2) Search Strategy Enhancement: K2-PSI-port

Improvement of search strategy: Before executing the greedy search strategy, the algorithm checks if the node X_i has prior parent nodes. If it does, these nodes are added to the parent node set of X_i before executing the search strategy. If not, the search strategy is executed directly. This process seeks the optimal parent node set for X_i . Finally, the algorithm outputs the parent node sets for all nodes. The process of the K2 algorithm for

integrating prior hierarchical structural information is shown in the K2-PSI-port algorithm (Figure 8).

Input: Node set, node order, dynamic parent node limit u_{core}/u_{aux} , prior parent node set PN_i , multi-source database D , weight matrix W
Output: Parent node set and weights for each node

```

1: Procedure K2-PSI-Port
2: For  $i = 1 \rightarrow n$  do
3:   If  $i \in \text{core nodes}$ , then  $u = u_{core}$ ; else  $u = u_{aux}$ 
4:    $\pi_i := PN_i$  // Initialize prior parent nodes
5:    $P_{old} := g_{new}(i, \pi_i, W)$ 
6:   OKToProceed := true
7:   While OKToProceed and  $|\pi_i| < u$  do
8:      $z := \text{argmax}_{z \in \text{Pred}(X_i) - \pi_i} [g_{new}(i, \pi_i \cup \{z\}, W)]$ 
9:      $P_{new} := g_{new}(i, \pi_i \cup \{z\}, W)$ 
10:    If  $P_{new} > P_{old}$  then
11:       $P_{old} := P_{new}$ 
12:       $\pi_i := \pi_i \cup \{z\}$ 
13:    Else
14:      OKToProceed := false
15:    End If
16:  End While
17:  If conflict exists ( $\pi_i$  does not conform to expert rule base) then
18:    Trigger expert review, update  $PN_i$  or  $W$ 
19:  End If
20:  Output node  $x_i$ 's parent nodes  $\pi_i$  and weight  $W$ 
21: End For
22: End Procedure

```

Figure 8. Flow diagram of the K2-PSI port algorithm.

Visualize the output results of K2-PSI port algorithm, which yield the initial DAG, as shown in Figure 9. From the connections between the subsystem nodes in Figure 9, it can be observed that there are only three direct dependency relationships among these four subsystems (S_1 – S_4), and there is a close relationship between the order parameters. This result confirms the effectiveness of the hierarchical prior structure designed in this paper, which successfully guides some of the cyclic dependencies between subsystems to the sequence parameter level.

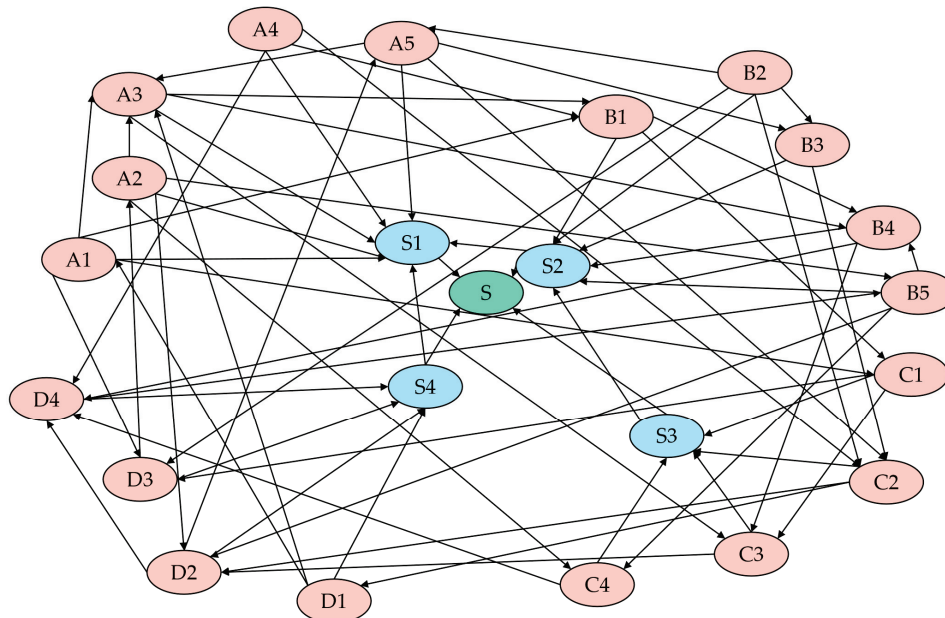


Figure 9. Initial Bayesian network structure generated by the K2-PSI port algorithm.

To enhance accuracy in reflecting real-world conditions, expert knowledge was utilized to refine the initial DAG. Five experts individually evaluated each directed edge shown in Figure 9. Edges identified by at least three experts as needing modification were adjusted accordingly. The experts unanimously agreed that overly strong interactions among order parameters in the initial DAG weakened direct causal relationships among subsystems.

Thus, these direct subsystem causalities, recognized as weak causal links, were removed. Additionally, inconsistencies with domain knowledge suggested potential overfitting due to insufficient data.

Accordingly, unreasonable directed edges were removed, and improperly directed edges were reversed based on expert feedback. Specifically, 11 edges were removed, and 11 edges were reversed. The final refined DAG, as shown in Figure 10, clearly illustrates the complex relationships among governance entities and factors, effectively depicting the internal structure of the collaborative system.

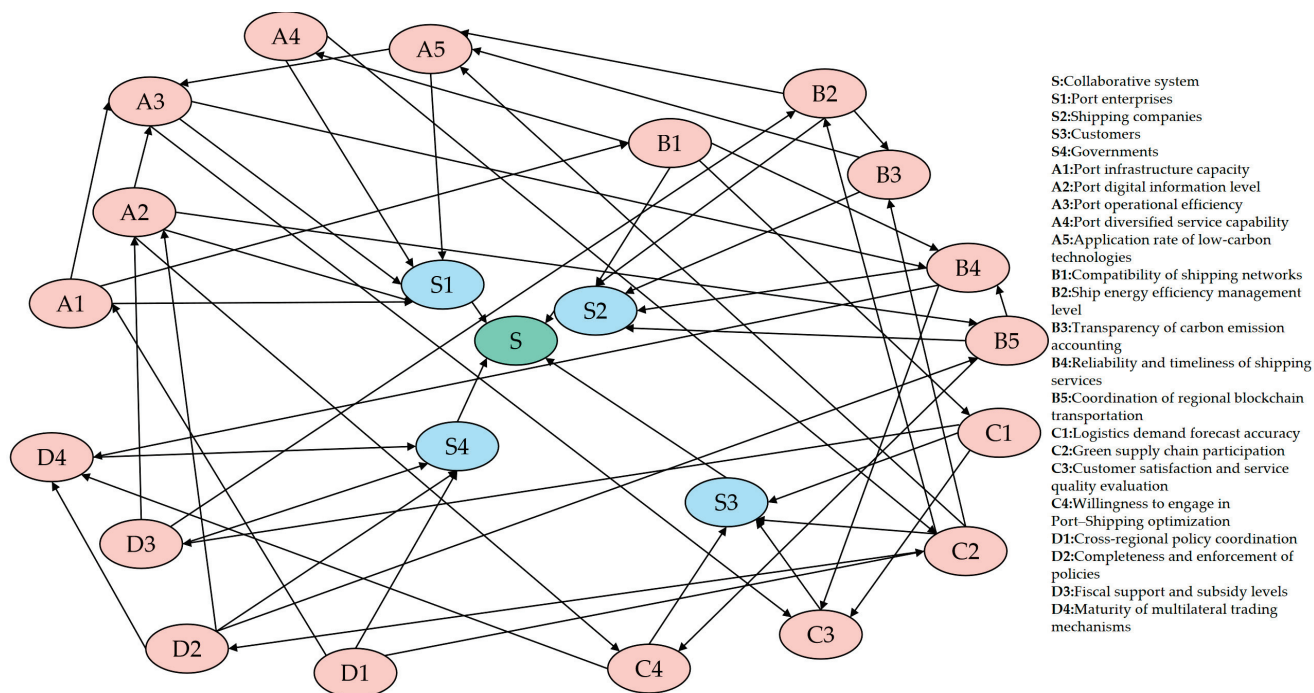


Figure 10. Final refined Bayesian network (DAG) after incorporating expert feedback and adjustments.

4.3.3. DAG Analysis

To provide deeper insights into the internal structure of port shipping synergy, this research constructs a DAG that visualizes the multilayered interactions among port enterprises, shipping companies, customers, and governments. Each subsystem node (S_1 – S_4) is linked to its respective order parameter nodes (e.g., A_1 – A_5 , B_1 – B_5), and certain key variables also exhibit cross-connections, reflecting the complex coupling mechanisms in collaborative governance. For instance, an improvement in the port enterprise subsystem's digital information integration (A_2) strongly influences other subsystems by facilitating initiatives like regional blockchain-based logistics (shared with the shipping subsystem) and enabling cross-regional policy coordination (connecting to the government subsystem). A higher level of port digitalization thus leads to more synchronized operations between ports and shipping lines, which in turn boosts overall collaborative performance. Another important causal path originates from the customers subsystem: high customer participation in green supply chain programs (C_2) and demand for better service quality (C_3) put positive pressure on ports and shipping companies to adopt greener practices and enhance service standards. This demand-pull effect is represented in the DAG by directed links from customer-related nodes to improvements in port operations and shipping services. On the policy side, strict environmental policy enforcement (D_2) and robust fiscal support (D_3) in the government subsystem emerge as top-down drivers that propagate through the network, compelling corresponding enhancements in both port and shipping subsystems.

According to Bayesian theory, if $P(S_1|A_1, \dots, A_5)$ denotes the probability distribution of S_1 under specific order parameter states—and similarly $P(S_2|B_1, \dots, B_5)$, $P(S_3|C_1, \dots, C_4)$, and $P(S_4|D_1, \dots, D_4)$ represent prior distributions for other subsystems—then the overall joint probability for the target node SSS can be expressed as shown in Equation (11):

$$\begin{aligned} & P(S, S_1, S_2, S_3, S_4, A_1, \dots, A_5, B_1, \dots, B_5, C_1, \dots, C_4, D_1, \dots, D_4) \\ = & P(A_1, \dots, A_5)P(S_1|A_1, \dots, A_5)P(B_1, \dots, B_5)P(S_2|B_1, \dots, B_5)P(C_1, \dots, C_4)P(S_3|C_1, \dots, C_4)P(D_1, \dots, D_4) \\ & P(S_4|D_1, \dots, D_4)P(S|S_1, S_2, S_3, S_4) = P(A_1, \dots, A_5, B_1, \dots, B_5, C_1, \dots, C_4, D_1, \dots, D_4) \\ & P(S_1|A_1, \dots, A_5)P(S_2|B_1, \dots, B_5)P(S_3|C_1, \dots, C_4)P(S_4|D_1, \dots, D_4)P(S|S_1, S_2, S_3, S_4) \end{aligned} \quad (11)$$

Such a decomposition effectively captures inter-subsystem coupling and, through forward and backward inference, identifies both the pathways and magnitudes of critical interactions. This network design retains hierarchical structure and clear visualization, overcoming the limitations of conventional BNs in handling cyclical dependencies and improving explanatory power in real-world systems. In the DAG, each edge denotes a probabilistic dependency derived from expert-informed structure learning; it signals influence, not deterministic causation.

From an internal perspective within each subsystem, every specific indicator not only directly affects that subsystem but may also exert indirect influence through interactions with other indicators. For instance, within the port enterprise subsystem (S_1), port operational performance (A_1) and port digital information level (A_2) exhibit a strong probabilistic dependency; through their interlinked effects, A_1 and A_2 together shape port efficiency (A_3) via a nonlinear interaction.

4.4. BN Parameter Estimation

The Expectation-Maximization (EM) algorithm was adopted for parameter estimation. Given the known network structure, this approach determines the conditional probability values of network nodes based on survey data on port shipping collaborative governance. Specifically, using survey data on collaborative governance capabilities, probabilities of each node in “Low”, “Medium” and “High” states were calculated to form a CPT. CPTs effectively quantify causal relationships within the network, explicitly indicating the impacts of governance factors on collaborative capability.

To ensure alignment with practical conditions in port shipping operations, this research integrates expert interviews and empirical data to determine the marginal probability of each node’s state. These aggregated results are summarized in Figure 11, which reveals the baseline probability distribution for each indicator. As shown in Figure 11, several core capability indicators already have a high likelihood of being in the “High” state (reflecting strong baseline performance in those areas), whereas other indicators remain predominantly in a “Medium” state, suggesting room for improvement. This probabilistic profile provides a useful context for the inference analysis that follows.

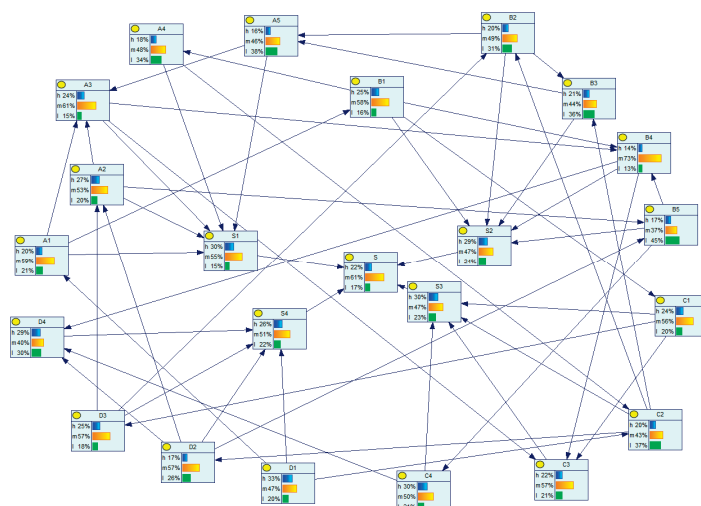


Figure 11. The marginal probabilities of nodes in the Bayesian network.

5. BN Model for Collaborative Development in the Port Shipping Industry

With the BN calibrated, Section 5 applies the model to analyze collaborative performance. We report forward and backward inference results and translate them into practical leverage points for managers and policymakers.

5.1. Scoring Framework for Node State Transition Analysis

Previous practices often relied solely on a simplistic “low-medium-high” classification to define node states, which fails to capture the distinct contributions of various node indicators to overall collaborative capacity. To thoroughly investigate the key drivers of port shipping collaborative development, this research builds upon conventional BN inference methods and incorporates the unique characteristics of port shipping operations and collaborative mechanisms. Accordingly, an innovative approach to quantifying node states is proposed, including a differentiated scoring system that more accurately reflects the impact of node state changes on overall synergy (Table 4).

Table 4. Quantitative indicators for node state changes.

Node Type	Low	Medium	High	Nodes
Core Capability Indicators	2	4	6	S, S_1, S_2, S_3, S_4
Technology/Policy/Collaboration Indicators	1	3	5	$A_1, A_2, \dots, D_3, D_4$

For technology- and policy-related indicators, this research applies a scoring scheme of 1, 3, and 5 points (low–medium–high). However, for core capability indicators—given their more pronounced effect on synergy—the scoring scale is adjusted to 2, 4, and 6 points (low–medium–high).

In real-world port shipping collaboration, each node exhibits varying degrees of importance. To quantify their overall scores, this research proposes a novel weighting and state transition method grounded in operational priorities. The core formulas are as follows:

- (1) The Pre-Adjustment Node Score (Score_{before}) is calculated according to Equation (12):

$$\text{Score}_{\text{before}} = \sum_{i=1}^n (P_{\text{high}} \times W_{\text{high}} + P_{\text{medium}} \times W_{\text{medium}} + P_{\text{low}} \times W_{\text{low}}) \quad (12)$$

Here, P_{high} , P_{medium} , and P_{low} denote the probabilities of a node being in high, medium, or low states before adjustment, whereas W_{high} , W_{medium} , and W_{low} represent the weights assigned to these respective states.

(2) The Post-Adjustment Node Score (Score_{after}) is given by Equation (13):

$$\text{Score}_{\text{after}} = \sum_{i=1}^n (P'_{\text{high}} \times W_{\text{high}} + P'_{\text{medium}} \times W_{\text{medium}} + P'_{\text{low}} \times W_{\text{low}}) \quad (13)$$

Here, P'_{high} , P'_{medium} , and P'_{low} denote the probabilities of the node being in each state after adjustment. Increasing the probability of the high state (for instance) would require a corresponding decrease in the medium or low states, reflecting the benefits of improving the node's condition.

(3) The Node Score Change Rate (Changes) can be computed according to Equation (14):

$$\text{Change}_S = \frac{\text{Score}_{\text{after}} - \text{Score}_{\text{before}}}{\text{Score}_{\text{before}}} \times 100\% \quad (14)$$

By comparing node scores before and after state optimization, it is possible to evaluate each node's relative contribution to enhancing overall port shipping synergy.

5.2. Forward Inference Analysis

Utilizing evidence-based causal inference helps reveal the dynamic relationships between subsystems and assess how changes in subsystem behavior impact the collaborative development capability of the port shipping industry. To achieve this, this research employs forward inference to examine how modifications to specific subsystem inputs influence the overall system. The rate of score variation indicates the extent to which each subsystem affects overall collaborative capability.

5.2.1. Analysis of the Port Enterprise Subsystem (S_1)

The inference results for changes in the port enterprise subsystem are presented in Table 5. Within the port enterprise subsystem (S_1), improvements typically bring the most immediate benefits to shipping companies (S_2) and governments (S_4), while exerting relatively limited influence on customers (S_3). This outcome partly reflects the tendency of traditional port enterprises to focus on close collaboration with the shipping sector and regulatory agencies in order to secure throughput, vessel turnover, and compliance. However, they pay insufficient attention to customer needs and engagement, resulting in minimal changes to C_3 and C_4 in the inference results.

Table 5. Reasoning for port enterprise subsystem change.

Nodes	Edge Probability/%			Change Ratio/%	Nodes	Edge Probability/%			Change Ratio/%
	High	Medium	Low			High	Medium	Low	
S_1	100	0	0	/	S_3	29.95	46.99	23.06	− 0.04
A_1	24.48	59.41	16.11	19.28	C_1	24.28	55.72	20	0.18
A_2	33.57	51.73	14.7	24.74	C_2	20.11	42.62	37.27	− 0.6
A_3	24.11	61.05	14.84	1.56	C_3	21.72	57.28	21	0.08
A_4	22.74	50.26	27	23.44	C_4	29.9	49.55	20.55	0.08
A_5	23.09	45.07	31.84	26.38	S_4	26.43	51.34	22.23	0.12
S_2	28.76	47.12	24.12	0.11	D_1	32.95	46.96	20.09	0.38
B_1	24.47	58.99	16.54	− 1.42	D_2	17.14	56.58	26.28	0.78
B_2	20.35	49.05	30.6	0.88	D_3	25.14	57.02	17.84	− 0.32
B_3	20.55	43.63	35.82	− 0.08	D_4	29.33	40.78	29.89	0.08
B_4	13.52	73.21	13.27	0	S	24.03	60.24	15.73	7.16
B_5	17.75	37.02	45.23	0.92					

When the port enterprise subsystem (S_1) significantly enhances its capabilities, the level of port digital information (A_2) accordingly rises, thereby more effectively boosting the coordination of regional blockchain transportation (B_5). Higher degrees of digitalization facilitate real-time data sharing and intelligent scheduling, reducing information asymmetry between shipping companies and governments and prompting more cross-regional and multiagency initiatives. Against the backdrop of rising decarbonization and environmental priorities, an elevated application rate of low-carbon technologies (A_5) not only helps secure government subsidies (D_3) and policy support but may also attract deeper collaboration from shipping companies, promoting more eco-friendly routing and vessel retrofitting.

5.2.2. Analysis of the Shipping Company Subsystem (S_2)

The inference results for changes in the shipping company subsystem are presented in Table 6.

Table 6. Reasoning for shipping company subsystem change.

Nodes	Edge Probability/%			Change Ratio/%	Nodes	Edge Probability/%			Change Ratio/%
	High	Medium	Low			High	Medium	Low	
S_1	29.72	55.32	14.96	0.12	S_3	29.94	47	23.06	−0.06
A_1	20.18	58.01	21.81	−0.72	C_1	24.03	55.98	19.99	−0.32
A_2	26.78	53.09	20.13	0.3	C_2	20.1	43.04	36.86	0.14
A_3	23.69	61.37	14.94	0.52	C_3	21.45	57.55	21	−0.52
A_4	17.99	47.93	34.08	−0.22	C_4	30	49.39	20.61	0.16
A_5	16.00	46.88	37.12	1.64	S_4	26.43	51.31	22.26	0
S_2	100	0	0	/	D_1	33.11	46.50	20.39	0.1
B_1	28.83	53.74	17.43	5.52	D_2	16.99	56.52	26.49	0
B_2	23.33	52.6	24.07	19.9	D_3	25.17	57.29	17.54	0.34
B_3	20.85	43.98	35.17	1.82	D_4	28.84	40.81	30.35	−1.22
B_4	16.08	73.39	10.53	10.66	S	24.98	59.67	15.35	9.78
B_5	19.93	40.5	39.57	16.6					

Shipping companies must interface with port enterprises (S_1) for berthing, cargo handling, and intermodal transport details, while also attending to customers' (S_3) sensitivity regarding timeliness, safety, and cost. Their operational efficiency and reliability not only directly affect port workflows but also indirectly shape customer satisfaction and engagement in overall supply chain services. When S_2 achieves a high-level state, it provides robust support for upstream port enterprises (S_1) and often drives government agencies toward green-shipping policies, while simultaneously stimulating greater demand for high-quality, sustainable logistics among customers (S_3).

Shipping operations are widely considered a major source of carbon emissions and energy consumption. When ship energy efficiency management (B_2) and the transparency of carbon emission accounting (B_3) both reach elevated levels, governments are more inclined to offer policy or subsidy support (D_3) and to encourage broader involvement in monitoring and refining green shipping services. Maintaining high performance in timeliness and reliability (B_4) further raises customer satisfaction and participation in port shipping optimization (C_3 , C_4). In addition, shipping companies should prioritize the coordination of regional blockchain transportation (B_5) by enabling real-time sharing and verification of shipping documents with port enterprises, customers, and governments, thus achieving end-to-end visibility and strengthening collaborative efficiency.

5.2.3. Analysis of the Customer Subsystem (S_3)

The inference results for changes in the customer subsystem are presented in Table 7. The customer subsystem (S_3) exerts the most significant positive impact on node S , reaching 10.2%—underscoring its critical role in modern supply chains. Customers' (S_3) preferences regarding delivery speed, costs, and eco-friendly attributes directly influence the operational and policy priorities of port enterprises (S_1), shipping companies (S_2), and governments (S_4). When customers actively connect with port and shipping sectors by furnishing more precise forecasts (C_1) and specifying clearer green service needs (C_2), port enterprises and shipping companies can better identify improvement pathways, collectively advancing the overall collaborative system.

When the customer subsystem (S_3) operates at its maximum potential, green supply chain participation (C_2) can surge by 54%, while customer satisfaction and service quality evaluation (C_3) may rise by 43%. This finding shows that stronger customer commitment to eco-friendly transport and carbon reductions (C_2) frequently motivates governments (S_4) to introduce targeted incentives or regulations—such as green subsidies or tax incentives—and spurs port and shipping sectors to enhance low-carbon technologies (A_5 , B_2 , and B_3). Customer satisfaction and service quality evaluation (C_3) directly shape how rapidly and in what direction port shipping enterprises refine their operations; if satisfaction remains low for an extended period, port and shipping enterprises lack clear insights into service defects, and government agencies (S_4) also lack reference data for timely policy adjustments. Moreover, more accurate logistics demand forecasts (C_1) enable port enterprises (S_1) to optimize berth allocation, scheduling, and equipment usage in advance, thereby mitigating congestion, reducing empty-load rates, and elevating overall turnover efficiency.

Table 7. Reasoning of customer subsystem change.

Nodes	Edge Probability/%			Change Ratio/%	Nodes	Edge Probability/%			Change Ratio/%
	High	Medium	Low			High	Medium	Low	
S_1	29.68	55.38	14.94	0.08	S_3	100	0	0	/
A_1	20.04	58.73	21.23	0.16	C_1	27.05	53.57	19.38	6.94
A_2	27.28	52.46	20.26	1.04	C_2	33.51	43.16	23.33	54.02
A_3	22.86	61.77	15.37	−2	C_3	34.38	53.69	11.93	43.48
A_4	17.88	47.38	34.74	−1.76	C_4	33.08	47.37	19.55	8.44
A_5	15.98	46.12	37.9	0.04	S_4	27.80	51.45	20.75	5.82
S_2	28.72	47.11	24.17	0	D_1	32.71	46.60	20.69	−1.3
B_1	24.92	58.41	16.67	−0.78	D_2	20.9	58.28	20.82	19.22
B_2	20.13	49.04	30.83	−0.02	D_3	25.06	57.37	17.57	0
B_3	20.36	43.73	35.91	−0.64	D_4	29.66	40.23	30.11	0.9
B_4	12.92	73.44	13.64	−1.88	S	25.48	58.88	15.64	10.2
B_5	17.24	37.52	45.24	−0.12					

5.2.4. Analysis of the Government Subsystem (S_4)

The inference results for changes in the government subsystem are presented in Table 8.

Table 8. Reasoning of government subsystem change.

Nodes	Edge Probability/%			Change Ratio/%	Nodes	Edge Probability/%			Change Ratio/%
	High	Medium	Low			High	Medium	Low	
S ₁	29.7	55.37	14.93	0.14	S ₃	30.38	46.7	22.92	1.10
A ₁	19.79	58.75	21.46	−0.8	C ₁	24.24	55.67	20.09	−0.1
A ₂	27.09	52.63	20.28	0.62	C ₂	21.64	42.24	36.12	4.7
A ₃	23.53	61.46	15.01	0.06	C ₃	21.77	57.23	21	0.12
A ₄	17.97	48.05	33.98	−0.06	C ₄	29.75	49.59	20.66	−0.44
A ₅	16.00	46.1	37.9	0.08	S ₄	100.00	0	0	/
S ₂	28.74	47.10	24.16	0.06	D ₁	52.38	33.64	13.98	51.46
B ₁	25.15	58.34	16.51	0	D ₂	24.08	50.87	25.05	17.12
B ₂	20.28	48.8	30.92	0.1	D ₃	24	58.29	17.71	−2.34
B ₃	20.58	43.62	35.8	0.02	D ₄	34.64	47.24	18.12	34.84
B ₄	13.28	73.23	13.49	−0.86	S	24.38	59.21	16.41	6.46
B ₅	17.21	37.5	45.29	−0.28					

Among the four subsystems, the government subsystem (S₄) exhibits the smallest positive effect, at only 6.46%. In a port shipping synergy system, changes in government policies or actions often benefit the customer subsystem (S₃) first—such as incentives for green logistics or mechanisms to boost customer engagement, which must be carried out by port enterprises and shipping companies. At the same time, infrastructure projects, technology subsidies, and legal frameworks also favor port enterprises (S₁) and shipping companies (S₂), as these measures align closely with industry needs. However, focusing excessively on collaboration with S₁ and S₂ while disregarding customer demands may constrain C₄ and hamper the full synergistic potential of D₁–D₄.

When the government subsystem (S₄) reaches its highest state, cross-regional policy coordination (D₁) significantly increases, making governments more inclined to coordinate interregional logistics processes, port oversight, and shipping strategies to better address both industrial and public needs. The maturity of multilateral trading mechanisms (D₄) rises by 34.84%, second only to the shift in D₁. By proactively engaging with port enterprises, shipping companies, and customers—often through multilateral consultation platforms—governments can promote joint technological innovations and low-carbon initiatives, thereby expanding both commercial and societal value. A well-established regulatory framework (D₂), combined with rigorous enforcement, further strengthens the confidence of port and shipping enterprises in the business environment, encouraging them to invest in port infrastructure and green upgrading (A₁, A₅, B₂). Nonetheless, if fiscal support and subsidies (D₃) do not increase proportionally, governments may need to refine funding mechanisms to better incentivize active participation from social organizations and sustain momentum for port shipping collaboration.

5.3. Backward Inference Analysis

To determine which factors are most critical for enhancing port shipping collaborative capacity (S), we set the overall synergy node S to a “high” state (100%) in the Bayesian network model, then applied backward inference. By observing how each indicator node’s state probabilities changed under S = “high” (as illustrated in Figure 12), we identified the elements that contribute the most to overall synergy. The results show that green supply chain participation (C₂), customer satisfaction and service quality (C₃), the maturity of multilateral trading mechanisms (D₄), completeness and enforcement of policies (D₂), and the port digital information level (A₂) exhibit the most significant changes. Prioritizing these five components is associated with the largest expected uplift in overall synergy according to the model.

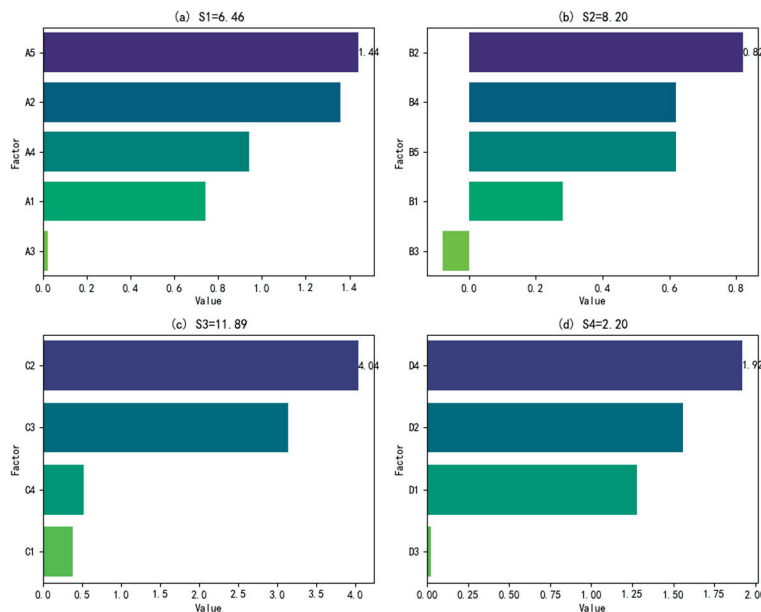


Figure 12. Reasoning for the collaborative ability of port shipping change.

Specifically, C_2 and C_3 reflect the customer subsystem's (S_3) pull for more eco-friendly logistics and higher service quality; when both factors reach high levels, they effectively incentivize port and shipping enterprises to decarbonize, improve efficiency, and reduce costs. Meanwhile, D_4 and D_2 originate from the government's subsystem (S_4), with D_4 offering more avenues for new business and subsidies through multilateral platforms, and D_2 ensuring a stable, enforceable institutional environment that underpins collaborative development. Finally, the port digital information level (A_2) in the port enterprise subsystem (S_1) was associated with a measurable improvement in the precision of information sharing and operational scheduling.

Although individual nodes exhibit score change rates mostly between 0.5% and 4%, suggesting relatively minor absolute shifts, small adjustments across multiple nodes can accumulate into a notable synergy, significantly raising S overall. Consequently, port enterprises, shipping companies, customers, and governments should aim to form a collective force rather than adopting fragmented, localized strategies. When shipping companies (S_2) attain a high level, they interact positively with active customer feedback (S_3) and specialized government support (S_4). However, fully unleashing the potential of port enterprises (S_1)—and thus stimulating system-wide synergy—requires simultaneous enhancements of key factors such as B_1 , B_2 , and B_5 , along with improvements in C_3 and C_4 . Consistent with the forward inference findings, the customer subsystem (S_3) stands out for its particularly strong influence on S , further affirming the central position of customers in port shipping collaboration. Small, simultaneous adjustments across multiple stakeholders accumulate into notable synergy gains, which implies that ports, shipping companies, customers, and governments should act in concert rather than in isolation to maximize collaborative benefits.

6. Discussion and Policy Implications

Building on the quantitative findings, this section interprets their managerial and policy significance and situates them within existing collaborative governance research.

6.1. Establishing a Unified Objective and Enhancing Synergistic Effects

Building on the inference results, we find that improvements in basic throughput capacity, digitalization, and service diversification at port enterprises (subsystem S_1) directly

promote overall synergy (S). However, simply expanding facilities or upgrading technologies on the port side alone does not fully boost collaboration with shipping companies (S₂) and customers (S₃). This underscores the need for ports to integrate capacity expansion with digital data-sharing in close coordination with shipping companies and customers, in order to fully leverage technological advancements. To maximize contributions to total collaborative capacity, port enterprises must strengthen both “hardware” (physical infrastructure) and “software” (digital and organizational mechanisms) while actively linking with the other subsystems.

First, when planning new terminals or expanding berths, port enterprises should coordinate schedules and routing with shipping companies—such as harmonizing shipping timetables and dispatch information—and leverage cross-regional policy support from governments (S₄) to balance construction with environmental requirements.

Second, improvements in port digital information level (A₂) must go beyond internal logistics efficiency and provide transparent, convenient data interfaces for customers (S₃), enabling them to swiftly track port operations and offer feedback. This in turn fosters greater demand-side acceptance of green and intelligent logistics. When ports and customers exchange information seamlessly, port operational efficiency (A₃) and diversified service capability (A₄) can be more accurately validated and refined.

Finally, ports should align their application rate of low-carbon technologies (A₅) with government ecological initiatives, seeking policy subsidies and piloting emerging technologies—such as blockchain—to bolster energy efficiency management connections with shipping companies (S₂). Under the momentum of green logistics, this also helps translate environmental values into market gains alongside customers. Only by placing equal emphasis on physical infrastructure (“hardware”) and collaborative mechanisms, and by forging common goals with other stakeholders, can ports truly serve as the cornerstone of overall port shipping synergy.

6.2. Identifying Stakeholder Differences and Optimizing Resource Allocation

Shipping companies (S₂) demonstrate a notably strong impact on overall collaborative capacity in forward inference, particularly through key indicators such as compatibility of shipping networks (B₁), ship energy efficiency management (B₂), and coordination of regional blockchain transportation (B₅). If shipping companies focus solely on internal scheduling or vessel dispatch without effectively coordinating with port enterprises (S₁) and governments (S₄), or without addressing the transport demands and service feedback of customers (S₃), they cannot sustain overall synergy enhancement. This reinforces the need for shipping companies to not only optimize energy usage but also to actively respond to client expectations in terms of personalized services and sustainable logistics solutions.

According to backward inference outcomes, shipping companies should deepen collaboration in two areas. First, they need close interaction with port enterprises to share information and synchronize schedules, translating the operational advantages of high-efficiency vessels into efficient port operations and higher customer satisfaction. Second, they should proactively accommodate customers’ preferences for green supply chains (C₂), service quality (C₃), and specialized requirements in their route planning and strategic development, steering shipping networks and bills of lading data toward lower carbon emissions and greater transparency. This approach not only attracts government incentives—such as fiscal support and subsidies (D₃)—but also garners broader social approval. By maintaining strong communication with S₁, S₃, and S₄, and by harnessing digital tools like big data, artificial intelligence, and blockchain, shipping companies can elevate their own operational performance while advancing industry-wide decarbonization and upgrading efforts.

6.3. Flexibly Adjusting Customer Engagement Mechanisms

Within BN inferences, the customers subsystem (S_3) exhibits a strong “demand-pull” effect, especially through indicators such as green supply chain participation (C_2) and customer satisfaction and service quality evaluation (C_3), which exert substantial pressure on port enterprises (S_1) and shipping companies (S_2) to innovate in technologies and services. However, the large and diverse customer base also complicates communication and governance. Achieving a genuine boost from S_3 thus requires addressing both “breadth” and “depth.” This dynamic can be further strengthened if customers actively use established feedback channels to express expectations for service excellence and sustainability.

On one hand, governments (S_4) and enterprises can gather and respond to customer input via multiple channels, such as online platforms, periodic surveys, and complaint-reporting mechanisms, thereby enhancing customer influence in areas like low-carbon logistics and digital operations. On the other hand, during critical decision-making and implementation—for instance, launching new green routes or piloting blockchain transportation—stakeholders should involve key customer segments or primary interest groups in more focused discussions to ensure practical viability and market alignment. By adopting a “broad-first, then-deep” participation strategy, organizations can capture diverse demands and identify hidden challenges without compromising decision efficiency. When customers maintain high satisfaction (C_3) and continue to engage (C_2 , C_4), their cumulative push for port shipping synergy grows stronger, making improvements by S_1 and S_2 more targeted and driving positive feedback loops through market choices and reputation.

6.4. Strengthening the Government’s Coordinating Role in Collaborative Governance

In both forward and backward inference results, the government subsystem (S_4) plays a clear “macro-level coordinator” role. Governments leverage policy and regulatory tools—including fiscal support, subsidies, and cross-regional coordination—to streamline resource flows and align stakeholder interests in port shipping systems. By offering cross-regional platforms and legal guarantees through well-crafted incentives and regulations, governments foster collaboration among all parties.

First, governments should employ top-level design to introduce tiered and differentiated support policies that ensure adequate investment in infrastructure, environmental protection, and digital transformation for port enterprises and shipping companies, while also meeting green logistics (C_2) and service demands (C_3) from the customer subsystem.

Second, governments need to strengthen legal and regulatory frameworks (D_2) to create a stable environment for innovations—such as blockchain transportation, energy efficiency management, and digital operations—and to use cross-regional policy coordination (D_1) to reduce local administrative barriers, allowing route planning and port expansion to achieve both low-carbon and high-efficiency objectives on a broader scale.

Finally, governments should appropriately redefine their role at the operational level: rather than micromanaging the activities of enterprises and customers, they can utilize policy instruments and public platforms to facilitate multi-stakeholder collaboration, assist S_1 and S_2 in upgrading technologies and operations, and help S_3 establish structured feedback mechanisms and participation channels. Only by integrating macro-level planning with grassroots flexibility can governments maintain stability, efficiency, and sustainability in this complex collaborative system.

7. Conclusions

Addressing challenges in multi-stakeholder coordination, resource misallocation, and dual pressures of low-carbon and digital transformations within port shipping, this re-

search develops a comprehensive BN model. Through forward and backward inferences, it elucidates the key interactive mechanisms among port enterprises (S_1), shipping companies (S_2), customers (S_3), and governments (S_4). The findings indicate that when stakeholders undertake moderate yet synchronized enhancements in hardware, technology, and policy, they generate stronger synergistic effects than large-scale investments made by any single actor in isolation. Moreover, core indicators—such as green supply chain participation, maturity of multilateral trading mechanisms, and port digital information level—demonstrate high sensitivity and amplification effects, underscoring the importance of prioritizing these aspects in policy-making and corporate reforms. Cross-regional policy coordination and rigorous regulatory enforcement also prove critical to balancing economic benefits with environmental goals; effectively integrating these factors with enterprise upgrades and customers' low-carbon demands can notably accelerate the greening and digitization of port shipping systems.

Based on the above inferences, this research proposes feasible recommendations for multi-stakeholder governance. Port enterprises should integrate capacity expansion with digital data-sharing in collaboration with shipping companies and customers to fully leverage technological dividends. Shipping companies must focus on energy efficiency management and strengthen responses to customer demands—both in customization and green logistics. Customers, in turn, should actively utilize feedback channels to advocate for higher service quality and environmental performance. Governments can offer cross-regional platforms and legal guarantees through differentiated incentives and regulatory frameworks, thus fostering collaboration among all parties. From a theoretical standpoint, our results align with collaborative governance scholarship that stresses the superiority of coordinated, cross-boundary action over single-actor initiatives. The BN reveals that digital information integration at ports and stringent policy enforcement reinforce one another, illustrating the shared motivation and joint capacity loop described in the IFCG. This coupling between digital infrastructure and policy extends the IFCG by highlighting connectivity in the data era as an emerging dimension of collaborative governance.

While BNs are well suited for capturing the nonlinear, uncertain interactions of multi-stakeholder systems, limitations remain concerning data completeness and the risk of overfitting. Future research should enlarge sample sizes, experiment with dynamic BNs, and undertake rigorous sensitivity or learning curve analyses to enhance predictive accuracy and overall model robustness. Furthermore, we recognize the importance of aligning the BN model with tangible operational performance metrics and validating its predictions against real-world outcomes. Although external validation of our BN has not been carried out in this study, we have now added a discussion of how the model's synergy insights relate to key port performance indicators. For example, improved port shipping coordination predicted by the BN would likely shorten vessel turnaround times and increase berth utilization by reducing delays and optimizing schedules. Similarly, higher collaborative efficiency—such as through synchronized operations and green initiatives—should translate into lower carbon emissions, as ships spend less time idling and resources are used more effectively. We emphasize that future work should validate these expectations by comparing the BN's probabilistic predictions with actual operational data (e.g., observed turnaround times, berth occupancy rates, and emissions), which will be crucial to confirm the model's practical applicability. Overall, by implementing relatively small yet simultaneous improvements across multiple dimensions, the port shipping industry can more rapidly achieve efficiency, resilience, and sustainability, thereby laying a solid foundation for a green and efficient global maritime ecosystem.

Author Contributions: Conceptualization, S.C., J.H. and Y.H.; methodology, S.C., J.H. and Z.H.; writing—original draft preparation, S.C. and J.H.; writing—review and editing, S.C., J.H. and Y.H.

and Z.H.; Visualization, S.C., J.H. and Z.H.; funding acquisition, J.H. and Y.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Project 2022YFC2302804.

Data Availability Statement: Readers can access our data by sending an email to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Christiansen, M.; Fagerholt, K.; Pisinger, D. Fifty Years on Maritime Transportation. *EURO J. Transp. Logist.* **2025**, *14*, 100148. [CrossRef]
2. Zhang, J.; Luo, M.; Yang, D. Are Efficient Ports for Port Operators Also Those for Shipping Companies? A Meta-Frontier Analysis of Global Top 80 Container Ports. *Ocean Coast. Manag.* **2025**, *263*, 107616. [CrossRef]
3. Tong, H. The Future Development and Restructuring of the International Shipping Industry: Conference Report. *Mar. Policy* **2022**, *137*, 104956. [CrossRef] [PubMed]
4. Shang, M.; Zhang, L.; Duan, H.; Wang, L.; Xiao, N. Trade Policy Uncertainty, Shipping Risk, and Commodity Markets. *Finance Res. Lett.* **2025**, *73*, 106604. [CrossRef]
5. Baştuğ, S.; Haralambides, H.; Esmer, S.; Eminoğlu, E. Port Competitiveness: Do Container Terminal Operators and Liner Shipping Companies See Eye to Eye? *Mar. Policy* **2022**, *135*, 104866. [CrossRef]
6. Zhang, J.; Yang, D.; Luo, M.F. Port Efficiency Types and Perspectives: A Literature Review. *Transp. Policy* **2024**, *156*, 13–24. [CrossRef]
7. Song, Z.; Tang, W.; Zhao, R.; Zhang, G. Implications of Government Subsidies on Shipping Companies' Shore Power Usage Strategies in Port. *Transp. Res. Part E Logist. Transp. Rev.* **2022**, *165*, 102840. [CrossRef]
8. Wang, N.; Yuen, K.F.; Gao, X.; Nie, Y. Resilience Assessment of Global Container Shipping Network via Port Communities. *Transp. Res. Part D Transp. Environ.* **2025**, *141*, 104649. [CrossRef]
9. Liu, J.; Fu, Y.; Han, K.; Cui, Y. Resilience Analysis of the RCEP Shipping Network. *Ocean Coast. Manag.* **2024**, *257*, 107317. [CrossRef]
10. Rødseth, K.L.; Wangsness, P.B.; Schøyen, H.; Førsund, F.R. Port Efficiency and Emissions from Ships at Berth: Application to the Norwegian Port Sector. *Marit. Econ. Logist.* **2020**, *22*, 585–609. [CrossRef]
11. Brooks, M.R.; Schellinck, T. Measuring Port Effectiveness: What Really Determines Cargo Interests' Evaluations of Port Service Delivery? *Marit. Pol. Manag.* **2015**, *42*, 699–711. [CrossRef]
12. Zhang, Z.; Song, C.; Zhang, J.; Chen, Z.; Liu, M.; Aziz, F.; Kurniawan, T.A.; Yap, P.S. Digitalization and Innovation in Green Ports: A Review of Current Issues, Contributions and the Way Forward in Promoting Sustainable Ports and Maritime Logistics. *Sci. Total Environ.* **2024**, *912*, 169075. [CrossRef] [PubMed]
13. Wang, L.; Peng, H. Collaboration or Competition? A Science Map for Promoting the Sustainable Transition of Port Relationships. *Mar. Policy* **2023**, *155*, 105795. [CrossRef]
14. Senarak, C. Shipping-Collaboration Model for the New Generation of Container Port in Innovation District: A Case of Eastern Economic Corridor. *Asian J. Ship. Logist.* **2020**, *36*, 65–77. [CrossRef]
15. Xu, L.; Xie, F.; Wang, C. Passive or Proactive Capacity Sharing? A Perspective of Cooperation and Competition between Two Regional Ports. *Marit. Policy Manag.* **2022**, *49*, 492–509. [CrossRef]
16. Yan, J.; Sun, X.; Liu, J.J. Assessing Container Operator Efficiency with Heterogeneous and Time-Varying Production Frontiers. *Transp. Res. Part B Methodol.* **2009**, *43*, 172–185. [CrossRef]
17. Feng, M.; Shaw, S.-L.; Peng, G.; Fang, Z. Time Efficiency Assessment of Ship Movements in Maritime Ports: A Case Study of Two Ports Based on AIS Data. *J. Transp. Geogr.* **2020**, *86*, 102741. [CrossRef]
18. Mi, C.; Huang, S.; Zhang, Y.; Zhang, Z.; Postolache, O. Design and Implementation of 3-D Measurement Method for Container Handling Target. *J. Mar. Sci. Eng.* **2022**, *10*, 1961. [CrossRef]
19. Ma, Q.; Li, S.; Jia, P.; Kuang, H. Is Port Integration a Panacea for Regions' Green Development: An Empirical Study of China Port City. *Transp. Policy* **2025**, *160*, 15–28. [CrossRef]
20. Jeh, J.; Nam, J.; Sim, M.; Kim, Y.; Shin, Y. A Study on the Efficiency Analysis of Global Terminal Operators Based on the Operation Characteristics. *Sustainability* **2022**, *14*, 536. [CrossRef]
21. Moon, D.S.-H.; Woo, J.K. The Impact of Port Operations on Efficient Ship Operation from Both Economic and Environmental Perspectives. *Marit. Pol. Manag.* **2014**, *41*, 444–461. [CrossRef]
22. Xu, L.; Wu, J.; Zhao, Q.; Chen, J.; Liu, J.; Gao, F.; Liu, X.; Xu, J. Spatial-Temporal Characteristics of Port Infrastructures on Sulfur-Oxide Concentrations of Coastal Port in China. *Ocean Coast. Manag.* **2024**, *258*, 107399. [CrossRef]

23. Meng, L.; Wang, J.; Yan, W.; Han, C. A Differential Game Model for Emission Reduction Decisions between Ports and Shipping Enterprises Considering Environmental Regulations. *Ocean Coast. Manag.* **2022**, *225*, 106221. [CrossRef]
24. Munim, Z.H. Does Higher Technical Efficiency Induce a Higher Service Level? A Paradox Association in the Context of Port Operations. *Asian J. Ship. Logist.* **2020**, *36*, 157–168. [CrossRef]
25. Wang, Z.; Subramanian, N.; Abdulrahman, M.D.; Cui, H.; Wu, L.; Liu, C. Port Sustainable Services Innovation: Ningbo Port Users' Expectation. *Sustain. Prod. Consum.* **2017**, *11*, 58–67. [CrossRef]
26. Munim, Z.H.; Hasan, K.R.; Schramm, H.-J.; Tusher, H.M. A Port Attractiveness Assessment Framework: Chittagong Port's Attractiveness from the Users' Perspective. *Case Stud. Transp. Policy* **2022**, *10*, 463–471. [CrossRef]
27. Feng, L.; Li, X.; Ng, A.K.Y.; Jia, P.; Kuang, H.; Zhang, X. Can Port Systems Be Enhanced by Government Promotion? The Case of Port Cooperation along the Maritime Silk Road. *Ocean Coast. Manag.* **2023**, *242*, 106669. [CrossRef]
28. Chen, S.; Miao, C.; Zhang, Q. Understanding the Evolution of China's Green Shipping Policies: Evidence by Social Network Analysis. *J. Clean. Prod.* **2024**, *482*, 144204. [CrossRef]
29. Zhao, C.; Peng, L.; Dong, K.; Yang, H. Government Policies on Port Resilience Amid Strike Events—A Two-Stage Van Damme Based Tripartite Evolutionary Game. *Transp. Res. Part A Policy Pract.* **2024**, *188*, 104196. [CrossRef]
30. De Campos, L.M. A Scoring Function for Learning Bayesian Networks Based on Mutual Information and Conditional Independence Tests. *J. Mach. Learn. Res.* **2006**, *7*, 2149–2187.
31. Song, G.; Khan, F.; Yang, M. Integrated Risk Management of Hazardous Processing Facilities. *Process Saf. Prog.* **2019**, *38*, 42–51. [CrossRef]
32. Li, X.; Chen, G.; Zhu, H. Quantitative risk analysis on leakage failure of submarine oil and gas pipelines using Bayesian network. *Process Saf. Environ. Prot.* **2016**, *103*, 163–173. [CrossRef]
33. Zhang, L.; Wu, X.; Ding, L.; Skibniewski, M.J.; Yan, Y. Decision support analysis for safety control in complex project environments based on Bayesian networks. *Expert Syst. Appl.* **2013**, *40*, 4273–4282. [CrossRef]
34. Panahi, R.; Gargari, N.S.; Lau, Y.Y.; Ng, A.K. Developing a resilience assessment model for critical infrastructures: The case of port in tackling the impacts posed by the COVID-19 pandemic. *Ocean Coast. Manag.* **2022**, *226*, 106240. [CrossRef]
35. Xiao, G.; Lu, Q.; Ni, A.; Zhang, C. Research on Carbon Emissions of Public Bikes Based on the Life Cycle Theory. *Transp. Lett.* **2023**, *15*, 278–295.
36. Lian, X.; Li, D.; Di, W.; Oubibi, M.; Zhang, X.; Zhang, S.; Xu, C.; Lu, H. Research on Influential Factors of Satisfaction for Residents in Unit Communities-Taking Ningbo City as an Example. *Sustainability* **2022**, *14*, 6687. [CrossRef]
37. Alshehri, S.A.; Rezugui, Y.; Li, H. Disaster Community Resilience Assessment Method: A Consensus-Based Delphi and AHP Approach. *Nat. Hazards* **2015**, *78*, 395–416. [CrossRef]
38. Li, M.; Ren, Z.; Hong, M.; Bai, C. Improved Structure Learning Algorithm of Bayesian Network Based on Information Flow. *Syst. Eng. Electron.* **2018**, *40*, 1385–1390.
39. Grootjans, S.J.M.; Stijnen, M.M.N.; Kroese, M.; Ruwaard, D.; Jansen, M.W.J. Collaborative Governance at the Start of an Integrated Community Approach: A Case Study. *BMC Public Health* **2022**, *22*, 1013. [CrossRef]
40. van Cleave, K. Narratives of Collaborative Governance: An Exploration of the National Innovation Council. Master's Thesis, Uppsala University, Uppsala, Sweden, 2020.
41. Beuren, M.M.; Andriotti, R.; Vieira, G.B.B.; Ribeiro, J.L.D.; Neto, F.J.K. On Measuring the Efficiency of Brazilian Ports and Their Management Models. *Marit. Econ. Logist.* **2018**, *20*, 149–168. [CrossRef]
42. Zheng, S.; Jiang, C. Consortium Blockchain in Shipping: Impacts on Industry and Social Welfare. *Transp. Res. Part A Policy Pract.* **2024**, *183*, 104071. [CrossRef]
43. Castelein, R.B.; Geerlings, H.; van Duin, J.H.R. Divergent Effects of Container Port Choice Incentives on Users' Behavior. *Transp. Policy* **2019**, *84*, 82–93. [CrossRef]
44. Xiao, Z.; Lam, J.S.L. Effects of Project-Specific Government Involvement Actions on the Attractiveness of Port Public-Private Partnerships Among Private Investors. *Transp. Policy* **2022**, *125*, 59–69. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Digital Transformation and Business Model Innovation: Enhancing Productivity in the Croatian Maritime Transport Sector

Marija Jović Mihanović ¹, Saša Aksentijević ², Edvard Tijan ^{2,*} and Gregor Lenart ³

¹ Institute of Shipping Economics and Logistics, 28359 Bremen, Germany; jovic@isl.org

² Faculty of Maritime Studies, University of Rijeka, 51000 Rijeka, Croatia; sasa.aksentijevic@pfri.uniri.hr

³ Faculty of Organizational Sciences, University of Maribor, 4000 Kranj, Slovenia; gregor.lenart@um.si

* Correspondence: edvard.tijan@pfri.uniri.hr; Tel.: +385-51-338-411

Abstract: This research investigates how business model changes induced by digital transformation impact productivity within the maritime transport sector in Croatia. Given the limited existing literature addressing digital transformation's productivity implications, specifically in maritime contexts, this study aims to identify and analyze key mediating factors. An online survey conducted among Croatian maritime transport stakeholders resulted in 82 valid responses, which were statistically analyzed using descriptive statistics, Spearman's correlation, and principal component analysis (PCA). The study identifies two primary dimensions of business model changes—innovation and process digitalization—that significantly correlate with increased productivity. Key influencing factors include the digitalization of internal and external business processes, development of new digital revenue streams, introduction of innovative services, and novel pricing models. Results underscore the importance of targeted digital transformation initiatives and serve as a valuable reference for maritime transport stakeholders, aiming to enhance their productivity and competitiveness through digital innovation.

Keywords: digital transformation; maritime transport; productivity; business model innovation; process digitalization; Croatia; principal component analysis

1. Introduction

The maritime sector in Croatia plays a vital strategic and economic role, reflecting the country's extensive Adriatic coastline, rich maritime heritage, and integration into broader European logistics networks. It encompasses a wide range of stakeholders, including state maritime administrations, port authorities, shipping companies, freight forwarders, sea brokers, terminal operators, and supporting hinterland transport providers, such as road and rail carriers. Croatia's major seaports—such as Rijeka, Split, and Ploče—serve as critical nodes not only for national trade flows but also for trans-European transport corridors, particularly Corridor Vb, linking Central Europe with the Adriatic Sea. However, in comparison to more digitally mature Western European ports like Rotterdam, Hamburg, or Antwerp, Croatian maritime logistics exhibit several structural and technological gaps. These include fragmented information systems, the reliance on manual procedures, limited interoperability between stakeholders, and slower adoption of advanced digital platforms, such as port community systems (PCSs), Internet of Things (IoT)-enabled monitoring, and AI-based predictive analytics. The institutional and regulatory inertia, lack of investment in digital infrastructure, and a general organizational conservatism in public and private

maritime actors have historically constrained innovation. Nevertheless, recent years have witnessed a growing momentum in digital transformation efforts driven by EU funding, national strategic frameworks (e.g., Smart Specialisation Strategy), and pressures to align with the European Green Deal and Fit-for-55 initiatives. Notably, Croatian maritime stakeholders are increasingly recognizing the productivity and efficiency gains available through the digitalization of business processes, integration of multimodal logistics platforms, and deployment of innovative business models. The emergence of “smart port” initiatives, albeit at early stages, mirrors European trends towards data-driven, sustainable, and customer-centric logistics operations. Yet, challenges remain in harmonizing stakeholder collaboration, developing digital skills within the maritime workforce, and securing resilient digital infrastructures. As a result, Croatia finds itself at a critical crossroads: while it possesses the geographical advantages and policy incentives to elevate its maritime sector into a competitive regional logistics hub, it must overcome systemic barriers and adopt a holistic, innovation-led approach to fully capitalize on the opportunities of digital transformation in line with broader European maritime logistics evolution.

The transformative potential of digital technologies in reshaping organizational productivity and the competitive advantage is gaining increasing attention from researchers and industry practitioners [1,2]. Digital transformation encompasses more than merely adopting new technologies; it also involves fundamentally rethinking business models, processes, and corporate strategies to adapt to evolving market demands and consumer expectations [3,4]. While numerous industries have rapidly integrated digital advancements, the maritime transport sector faces unique challenges, causing slower progress in digital transformation compared to other transportation sectors [3,5]. Maritime transport stakeholders often experience inefficiencies due to traditional manual processes, insufficient coordination among stakeholders, and the inadequate integration of digital solutions [6,7]. These inefficiencies lead to increased operational costs, reduced reliability, and longer transit times [5]. Addressing these challenges through targeted digital initiatives, such as the digitalization of internal and external processes and the adoption of innovative business models, is essential for enhancing productivity and operational efficiency within maritime transport stakeholders [1,3,6].

The maritime sector’s digital transformation journey involves significant organizational and cultural changes, requiring new business models, advanced digital solutions, and comprehensive integration strategies. These transformations streamline cargo handling, improve document management, enhance financial transaction efficiency, and optimize overall operational processes [5–7]. However, adopting digital practices in maritime transport stakeholders is hindered by several barriers, including entrenched organizational cultures, limited resource allocation, and regulatory complexities [3,4]. Additionally, challenges such as cybersecurity threats and fragmented regulations further complicate the transition to digital processes, potentially stifling innovation [6,7]. For decision-makers in enterprises and governments, it is important to understand how enterprises are approaching the adoption of digital technologies and how successful they are in the digital transformation journey. This paper builds on the existing body of literature on digital transformation within maritime sectors, particularly utilizing the Technology–Organization–Environment (TOE) framework introduced by Tornatzky and Fleischer (1990) [8], which examines technological adoption, organizational preparedness, and external environmental influences as key determinants in successful digital transformation initiatives [1,7].

The aim of our research is to investigate how digital transformation influences business model changes and productivity within Croatian maritime transport stakeholders.

2. Methodology

The research is structured in the following way, as shown in the Figure 1.

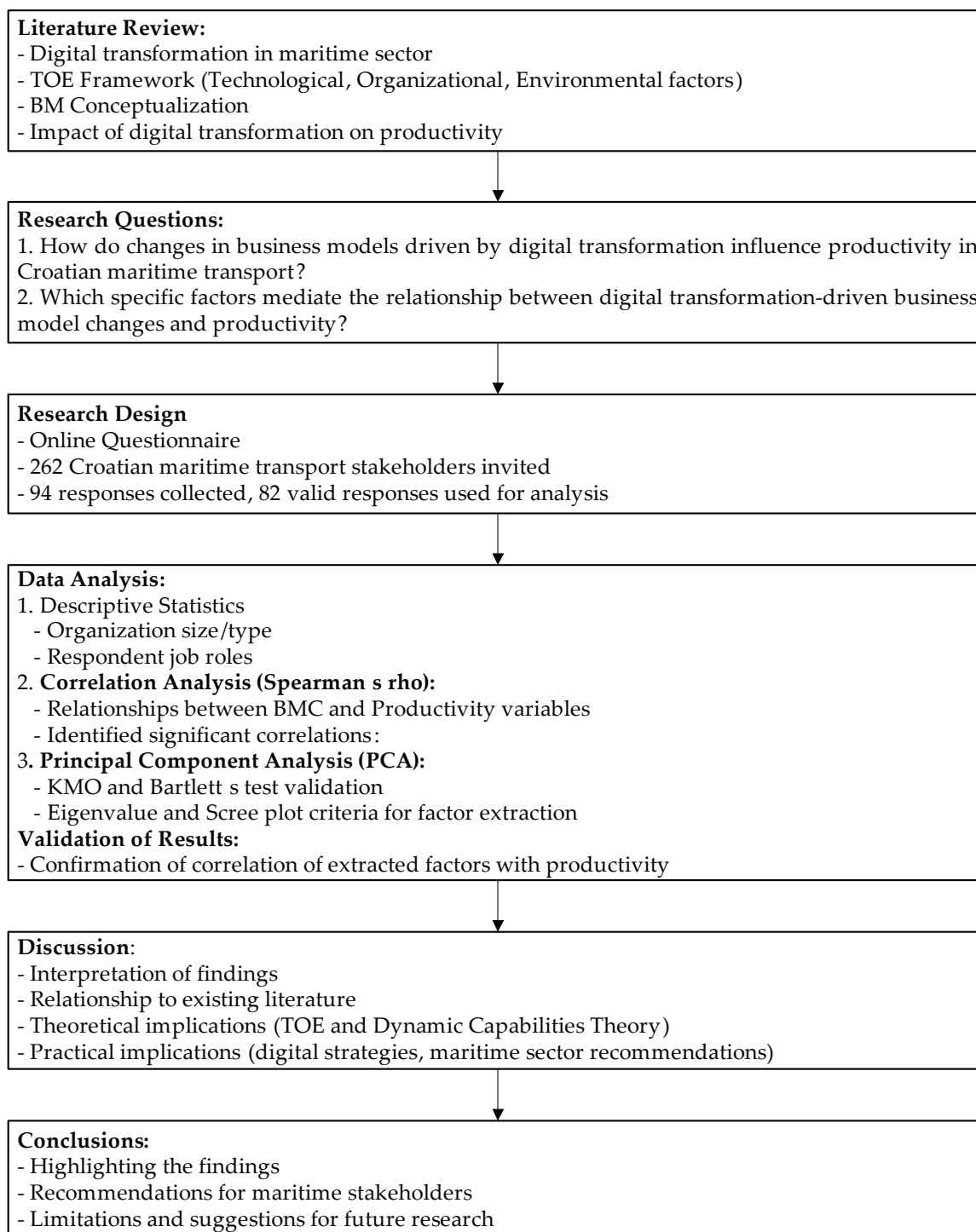


Figure 1. Research methodology.

2.1. Literature Review

To ensure the inclusion of the most recent findings, the literature review was initially focused on literature published within the past three years. This timeframe ensures that the review captures the latest developments and trends, reflecting the dynamic nature of the sector's transformation. This review utilized the Web of Science database, employing specific search terms to ensure a focused retrieval of relevant literature. The Web of Science

database was considered as it represents the world's leading scientific citation search and analytical information platform [9].

The keywords used in the search included the following:

- Digital Transformation AND Maritime Transport
- Digital Transformation AND Seaport
- Digital Transformation AND Shipping

Recognizing the strategic role of seaports as pivotal nodes in the transport chain [10] and acknowledging that their competitiveness and development are strongly influenced by connectivity with the hinterland [11,12], the authors also included rail and road carriers in both the survey and the literature search. These actors play a crucial role in the functioning of maritime logistics chains and were thus relevant for a holistic understanding of digital transformation impacts across the transport ecosystem.

The search criteria within the Web of Science database were limited to either the topic or title of the articles, and only those published in the English language were considered. Subsequently, manual screening of the papers was undertaken to assess their relevance to the topic at hand.

The statistical part of this study builds upon acquired data and research findings by Jović et al. [3], which encompasses various factors related to digital transformation initiatives and productivity outcomes in maritime transport stakeholders.

This study further explores what are the digital transformation business model changes key factors to increase productivity in maritime transport stakeholders. For the purpose of this study, an online survey was conducted among 262 maritime transport stakeholders in Croatia that were invited to participate in the research. In total, 94 organizations responded to the survey. Out of the 94 that filled out questionnaires, 82 questionnaires were usable for analysis in this study, while 12 were incomplete.

The questionnaire consisted of 51 questions, out of which the initial 7 were demographics questions, followed by 30 questions related to the technological, organizational, and environmental factors (TOE framework) of digital transformation. The second part of the questionnaire consisted of 8 questions about digital transformation business model changes (Table 1) and 6 questions about digital transformation, which also included the question about the increase in productivity due to digital transformation (Table 1).

A questionnaire was used to measure two main underlying constructs for this study, namely the following:

1. Digital transformation business model changes (BMCs);
2. The productivity increase (P).

A focused literature review was conducted to deepen the understanding of the relationship among digital transformation, business model characteristics, and productivity in the maritime transport sector.

Sooprayen et al. [13] provided a structured framework identifying 17 key factors influencing innovation adoption in ports, including the organizational support, financial capacity, network embeddedness, and perceived usefulness. These factors conceptually align with variables examined in this study, particularly collaboration with partners (BMC1), internal and external process digitalization (BMC2, BMC3), and the perceived impact of digital transformation on productivity (PROD).

Kaczerska et al. [14] investigated the adoption of digital tools in ferry shipping, focusing on innovations such as online ticket sales, chatbot integration, and social media engagement. These initiatives represent a shift towards new service delivery models and customer interaction channels, aligning directly with variables such as new services (BMC6), new sales channels (BMC7), and new charging methods (BMC8). Moreover,

the strategic use of external platforms and digital interfaces implies cross-organizational collaboration (BMC1).

Pang et al. [15] investigated the impact of digital transformation on sustainable management performance in shipping firms, focusing on the moderating role of coordination mechanisms (both cross-functional and customer-based) and social norms. Although their primary outcome is sustainability, their study implicitly addresses efficiency and process performance improvements—concepts closely related to productivity (PROD). The findings are particularly relevant to this research, as they reinforce the idea that the effectiveness of digital transformation depends not only on technology adoption, but also on internal organizational capabilities (BMC2) and stakeholder collaboration (BMC1).

Belmoukari et al. [16] provided a comprehensive systematic literature review of the smart port concept, emphasizing digital transformation as a multidimensional shift involving technological innovation, stakeholder collaboration, and operational efficiency. Their classification of smart port features—such as real-time information sharing, intelligent infrastructure, and process automation—aligns closely with several business model change variables in this research (e.g., internal/external process digitalization—BMC2, BMC3; new services—BMC6; new revenue streams—BMC4). Notably, they identify operational efficiency and productivity improvement as key goals of smart port transformation, reinforcing the validity of using productivity (PROD) as a performance indicator in evaluating the outcomes of digital transformation initiatives in maritime transport stakeholders.

Guerrero-Molina et al. [17] provided a scientometric review of Industry 4.0 technologies in port operations, highlighting key development trajectories, such as smart port development, sustainability, and digital innovation. Their findings emphasize how technologies such as IoT, AI, and digital twins are transforming port processes and value chains, enhancing productivity, automation, and stakeholder integration. These insights align closely with the variables analyzed in this research—particularly internal and external process digitalization (BMC2, BMC3), the introduction of new services (BMC6), and increased productivity as a performance outcome (PROD).

Basulo-Ribeiro and Teixeira [18] presented a systematic literature review on the application of Industry 4.0 technologies in seaport logistics, emphasizing their role in improving operational efficiency, stakeholder collaboration, and sustainability. Their findings support several key aspects analyzed in this study, particularly the digitalization of internal and external processes (BMC2, BMC3), cooperation with partners in digital development (BMC1), and the enhancement of productivity (PROD). The review also highlights benefits such as real-time operations and improved and personalized customer service, aligning with the examined variables on innovative service provision and revenue generation (BMC4, BMC6, BMC8).

Durán et al. [19] developed the DMLBC method—an integrated system combining blockchain and machine learning—to improve management control and decision-making in Chilean seaports. Their model demonstrates how digital transformation can increase operational efficiency through real-time monitoring, the automation of documentation, and data-driven forecasting. These aspects strongly align with the variables in this research, particularly in relation to internal and external process digitalization (BMC2, BMC3), productivity gains (PROD), and enhanced stakeholder collaboration (BMC1). Furthermore, the study highlights the potential of smart contracts and predictive analytics to support new service delivery and revenue models (BMC4, BMC6, BMC8), reinforcing the role of innovation as a key driver of digital transformation in maritime logistics.

Nguyen [1] used the Technology–Organization–Environment (TOE) framework to identify key drivers and outcomes of digital transformation in Vietnamese shipping companies. The study finds that digital transformation, supported by digital technology

development, skills, and top management commitment, significantly improves economic performance—aligning with this research’s findings on productivity (PROD). Moreover, Nguyen emphasizes digital interconnectivity, process automation, and real-time data sharing, which resonate with the analyzed variables related to internal and external process digitalization (BMC2, BMC3). The study also confirms that digital transformation facilitates new business models and services, linking conceptually to variables concerning innovative services, revenue streams, and pricing models (BMC4, BMC6, BMC8). These insights confirm that digital innovation plays an important role in improving productivity in maritime transport stakeholders.

Gao et al. [20] explored how a digital platform in container shipping—Duckbill Technology—built digital resilience during the COVID-19 crisis through evolving orchestration roles. Their longitudinal case study highlights how digital transformation supported internal and external process digitalization (BMC2, BMC3), stakeholder collaboration (BMC1), and the development of new services, pricing models, and market expansion (BMC4–BMC8). By applying dynamic strategies such as activation, buffering, and leapfrogging, the company managed to enhance operational flexibility and efficiency, ultimately increasing productivity (PROD).

Yang and Lin [21] investigated how digitalization and digital logistics platform adoption affect digital transformation and organizational performance in Taiwan’s maritime logistics sector. Their results show that both digitalization and digital transformation significantly enhance organizational performance, including productivity enhancement, value creation, and customer interaction—which closely align with variables used in this research (e.g., productivity improvement—PROD, new revenue streams—BMC4, new services—BMC6). Notably, while platform adoption alone had no direct effect on performance, its positive impact was fully mediated by internal digital transformation, highlighting the importance of digitalizing internal processes (BMC2) before engaging in broader collaborative innovations (BMC1).

Yu et al. [22] evaluated the competitiveness of container shipping companies in the era of sustainability and digitalization, emphasizing the critical role of technological innovation, customer satisfaction, and the service scope. Their operational framework incorporates both direct and indirect performance indicators—such as the digitalization level, environmental protection, and service reach—which conceptually align with the constructs used in this study (e.g., internal and external process digitalization, new service development, and cooperation with partners). Their findings suggest that higher levels of digital maturity contribute to enhanced operational performance, which resonates with this study’s observation that changes in the digital transformation business model are positively associated with a productivity increase (PROD).

Bucak [23] highlighted digital transformation as the most influential trend in the liner shipping sector, emphasizing its role in gaining a competitive advantage. This supports the findings of the present study, where business model changes such as process digitalization (BMC2, BMC3), new services (BMC6), and pricing models (BMC8) are linked to productivity improvements (PROD). The study reinforces the idea that digital adaptation enhances both operational efficiency and market positioning in maritime logistics.

Lam and Tang [24] analyzed the implementation of robotic process automation (RPA) in the cold chain management of a freight forwarding company. Their case study demonstrates significant improvements in operational efficiency and data processing speeds—up to 97% time savings—after adopting RPA bots for shipment and temperature tracking. These results align closely with the internal process digitalization variable (BMC2) and support the observed positive effect of digital transformation on productivity (PROD) in

this study. Moreover, the improved workflow and automated reporting enhance service delivery, corresponding to business model innovations such as new services (BMC6).

Aerts and Mathys [25] analyzed digitalization trends in maritime transport based on over 500 industry presentations. They identify key focus areas such as smart ports, AI, blockchain, and digital supply chains. These insights support this study's focus on internal and external process digitalization (BMC2, BMC3) and suggest that digital innovation contributes to improved operational performance, aligning with the observed relationship between business model changes and productivity (PROD).

Evmides et al. [26] developed advanced machine-learning-based models to improve the accuracy of vessel arrival time predictions using AIS data in the Eastern Mediterranean. By enhancing ETA reliability, the study demonstrates tangible benefits in resource planning and port efficiency. These operational improvements align closely with the business model change variables in this study—especially the digitalization of internal and external processes (BMC2, BMC3)—and support the notion that targeted digital initiatives contribute to productivity gains (PROD) in maritime transport stakeholders.

Su et al. [27] highlighted the deployment of advanced digital technologies—such as digital twin systems, blockchain, and automated customs clearance—as enablers of service innovation within port operations. While not explicitly labeled as “new services”, these tools represent enhanced service offerings aligned with the construct of BMC6, showcasing how digital transformation facilitates value-added services in the maritime domain. The emphasis on efficiency also indirectly supports the link with productivity improvements (PROD).

Gündoğan and Keçeci [28] explored digital transformation adoption in maritime logistics using the Technology Acceptance Model. The findings of the study indicate that the degree to which employees in the industry embraced digital transformation technologies was influenced by various factors, including their perception of how easy these technologies were to use and the perceived benefits they offered. Their findings therefore provide additional insight into the organizational dynamics that can drive or hinder productivity (PROD) improvements in maritime logistics.

Melnyk et al. [29] proposed a probabilistic model for assessing cybersecurity risks in shipboard systems, emphasizing the increasing vulnerability of digitalized maritime assets. While the study is primarily technical and focuses on risk modeling, its insights are relevant to the broader context of digital transformation in the maritime sector. Secure and resilient digital systems are a prerequisite for the successful implementation of internal and external process digitalization (BMC2, BMC3), as well as for sustaining productivity gains (PROD) and developing trustworthy digital service offerings (BMC6). Their findings underscore the need to consider cybersecurity as a strategic component of digital business model changes.

2.2. Research Questions

Based on this theoretical foundation, the current study aims to empirically investigate how digital transformation influences business model changes and productivity within Croatian maritime transport stakeholders. Employing a survey methodology, this study examines the state of digital transformation, categorizing influencing factors according to the TOE framework and assessing their impacts on productivity outcomes. Croatia serves as an appropriate context for this research due to its strategic geographic location, significant role in European maritime logistics networks, and increasing investments in digital transformation initiatives, particularly within its major ports. This research addresses two central questions:

1. How do changes in business models driven by digital transformation influence productivity in Croatian maritime transport stakeholders?
2. What specific factors mediate the relationship between digital transformation-driven business model changes and productivity?

The results from this study provide valuable insights into the key factors facilitating digital transformation, contributing to the global understanding of digital innovation's role in maritime sector productivity. These findings can assist stakeholders and decision-makers in designing effective digital strategies in order to increase productivity in their respective organizations.

TOE mapping will clearly demonstrate that the research aligns its measured variables directly with each dimension of the TOE framework, supporting a comprehensive analysis and reliable conclusions about the relationship among digital transformation, business model innovation, and productivity within the maritime transport sector. In case of our research, TOE mapping is done as follows:

1. Technological dimension (BMC2, BMC3): Digitalization processes directly involve the implementation and usage of technological systems, solutions, or infrastructures.
2. Organizational dimension (BMC1, BMC4, BMC6, BMC7, BMC8): These variables reflect internal strategic, managerial, or innovative capabilities essential for effectively adopting and utilizing digital transformation.
3. Environmental dimension (BMC5): Focuses explicitly on external market orientation, competition, and environmental pressures.

3. Results

The BMC construct consisted of eight questions and the P construct of one question. The scale had a high level of internal consistency, as determined by a Cronbach's alpha of 0.904. A Cronbach's alpha value equal or higher than 0.7 shows a good level of internal consistency [30].

For the analysis of gathered data, the SPSS 29 software package was used. The analysis included the following:

- Descriptive Statistics: To summarize the data characteristics and establish a foundational understanding of the dataset;
- Correlation Analysis: To identify potential relationships between the variables associated with business model changes and productivity increases, the Spearman rho coefficient analysis was used due to the use of an ordinal Likert scale (1-totally disagree, 5-totally agree). All variables had paired observations ($N = 82$), and a test of monotonic relationships among variables was conducted based on a scatter plot diagram in SPSS;
- Principal Component Analysis (PCA): To reduce dimensionality and identify the underlying structures influencing productivity stemming from digital transformation efforts. PCA requires 5–10 cases per variable. Since for the purpose of this research study PCA was run on eight variables, our sample size $N = 82$ was of a sufficient size ($N > 80$).

Table 1 presents the descriptive statistics of the variables, which were analyzed in this research study.

Table 1. Variable short and long name.

Short Name	Variable Name	N	Mean	Std. Deviation
BMC1	The organization cooperates with new partners with the aim of developing new digital solutions	82	3.46	1.091
BMC2	The organization has digitalized internal business processes	82	3.70	1.039
BMC3	The organization has digitalized external business processes	82	3.37	1.025
BMC4	The organization generates additional revenue from new sources as a result of the implementation of digital technologies	82	2.79	1.108
BMC5	The organization has entered new markets as a result of digitalization and digital transformation	82	2.60	1.076
BMC6	The organization provides new services as a result of digitalization and digital transformation	82	3.01	1.232
BMC7	The organization has introduced new sales channels as a result of digitalization and digital transformation	82	2.90	1.273
BMC8	The organization has introduced new ways of charging for services as a result of digitalization and digital transformation	82	3.02	1.296
PROD	The organization has increased productivity by introducing digital transformation	82	3.54	0.905

The survey participants were from small, medium-sized, and large maritime transport stakeholders of various types as shown in Table 2. Most participants, 67.1%, were from small organizations with up to 49 employees; 17.1% of respondents were from medium-sized organizations with 50–249 employees, while 14.6% were from large organizations that have 250 or more employees.

Table 2. Survey participants' organization sizes.

	N	%
No answer	1	1.2%
Small organization (up to 49 employees)	55	67.1%
Medium organization (50–249 employees)	14	17.1%
Large organization (250+ employees)	12	14.6%

The largest group of organizations were state maritime transport organizations and administration (25.6%), followed by shipping and logistics stakeholders (each account for 14.6%), sea brokers (13.4%), port operators and terminals, along with the 'Other' category, each comprising 11.0%. The remaining participants were involved in road transport, 3.7%, seaport agents, 3.7%, and railways transport, 1.2% (Table 3).

Table 3. Survey participant organization type.

	N	%
State organization and administration	21	25.6%
Shipping	12	14.6%
Road transport	3	3.7%
Railways transport	1	1.2%
Sea port agent	3	3.7%
Sea brokers	11	13.4%
Logistics operators	12	14.6%
Port operators and terminals	9	11.0%
Other	9	11.0%
No answer	1	1.2%

Most respondents held a position of executive directors (35.4%), followed by IT executive directors (9.8%) and project managers cover (7.3%); 4.9% of participants were sales executive directors. A small percentage of respondents did not provide an answer to this question (2.4%) (Table 4).

Table 4. Respondent function in organizations.

	N	%
Executive director	29	35.4%
IT executive director	8	9.8%
Project manager	6	7.3%
Sales executive director	4	4.9%
No answer	2	2.4%
Other	33	40.2%

A Spearman's rank-order correlation was run (Table 5) to assess the relationship between digital transformation business model changes and the productivity increase. Eighty-two survey participants provided valid responses that were included in this analysis. The preliminary analysis showed the relationship to be monotonic, as evidenced by the visual inspection of a scatterplot, which enabled further analysis.

Table 5. Spearman rho correlation coefficient between digital transformation business model changes and productivity increases.

		PROD	BMC1	BMC2	BMC3	BMC4	BMC5	BMC6	BMC7
Spearman's rho	BMC1	Corr. Coef.	0.200						
		Sig. (2-tailed)	0.072						
	BMC2	Corr. Coef.	0.487 **	0.476 **					
		Sig. (2-tailed)	<0.001	<0.001					
	BMC3	Corr. Coef.	0.373 **	0.466 **	0.729 **				
		Sig. (2-tailed)	<0.001	<0.001	<0.001				
	BMC4	Corr. Coef.	0.333 **	0.342 **	0.343 **	0.394 **			
		Sig. (2-tailed)	0.002	0.002	0.002	<0.001			
	BMC5	Corr. Coef.	0.186	0.352 **	0.271	0.282	0.755 **		
		Sig. (2-tailed)	0.094	0.001	0.014	0.010	<0.001		
	BMC6	Corr. Coef.	0.322 **	0.391 **	0.392 **	0.413 **	0.715 **	0.741 **	
		Sig. (2-tailed)	0.003	<0.001	<0.001	<0.001	<0.001	<0.001	
	BMC7	Corr. Coef.	0.230	0.345 **	0.322 **	0.312 **	0.740 **	0.621 **	0.690 **
		Sig. (2-tailed)	0.038	0.001	0.003	0.004	<0.001	<0.001	<0.001
	BMC8	Corr. Coef.	0.306 **	0.421 **	0.416 **	0.468 **	0.627 **	0.535 **	0.569 **
		Sig. (2-tailed)	0.005	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

** Correlation is significant at the 0.01 level (2-tailed). N = 82.

There was a statistically significant, strong positive correlation between the digitalization of internal business processes (BMC2) and a productivity increase (PROD), $r_s = 0.487$, $p < 0.01$. Furthermore, there was also a similar statistically significant, strong positive correlation between the digitalization of external business processes (BMC3) and a productivity increase (PROD), $r_s = 0.373$, $p < 0.01$. Additionally, there was also a statistically significant, strong positive correlation between generating additional revenue from new sources as a result of the implementation of digital technologies (BMC4) and an increase in productivity (PROD), $r_s = 0.333$, $p < 0.01$; between providing new services as a result of digital transformation (BMC6) and an increase of productivity (PROD), $r_s = 0.322$, $p < 0.01$; and between introducing new ways of charging for services as a result of digital transformation (BMC8) and an increase of productivity (PROD), $r_s = 0.306$, $p < 0.01$.

Altogether, we can sum up that five different variables for digital transformation business model changes are statistically significantly correlated with the increase in productivity. Additionally, it is also noticeable from Table 5 that there are many statistically significant correlations among other digital transformation business model change variables, which were taken into consideration for further principal components analysis when selecting appropriate method for rotating components.

Since there are many statistically significant correlated variables, we performed principal component analysis in order to extract digital transformation business model change key factors that correlate with the productivity increase. Principal component analysis (PCA) is a technique to extract key factors from a larger set of variables by analyzing variance–covariance differences among a set of variables. The extracted factors preserve large proportions of the original set of variables while reducing the number of variables in order to support an understanding of the structure of latent extracted factors [31].

Principal component analysis was conducted for variables BMC1 to BMC8 with the direct oblimin rotation method since we encountered correlated variables during the correlation analysis (Table 5). The suitability of the PCA was tested prior to analysis. An inspection of the correlation matrix showed that all analyzed variables had at least one correlation coefficient higher than 0.3. The overall Kaiser–Meyer–Olkin (KMO) measure was 0.869 and Bartlett’s test of sphericity was statistically significant, chi-square (28) = 485, $p < 0.001$, which confirmed that the sample data were appropriate for PCA analysis.

The number of extracted factors was determined based on Kaiser’s criteria of an eigenvalue > 1 (Table 6) and the Cattell scree plot (Figure 2).

Table 6. Initial eigenvalues and total variance explained.

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings ¹
	Total	% of Variance	Cumulative %	Total
1	5.047	63.086	63.086	4.634
2	1.233	15.408	78.494	3.595
3	0.507	6.337	84.831	
4	0.410	5.128	89.959	
5	0.268	3.346	93.305	
6	0.200	2.505	95.809	
7	0.190	2.375	98.184	
8	0.145	1.816	100.000	

Extraction method: principal component analysis. ¹ When components are correlated, the sums of squared loadings cannot be added to obtain the total variance.

PCA revealed two components that have eigenvalues greater than one and which explained 63.08% and 15.41% of the total variance (Table 6).

The eigenvalue-one criterion is a method for establishing how many components to retain in a principal components analysis. An eigenvalue less than one indicates that the component explains less variance than a variable would and hence should not be retained. The visual inspection of the scree plot (Figure 2) also indicated that the research should proceed with two-component solution [32,33]. In addition, a two-component solution met the interpretability criterion since the research aimed to extract factors from the variables set in order to explain which digital transformation business model changes factors and how they impact the productivity increase.

The two-component solution explained 78.49% of the total variance. Component loadings of the rotated pattern matrix solution are presented in Table 7. To contrast the

solution, we set a cutoff value of 0.50 in order to have single variable representing each principal component [31].

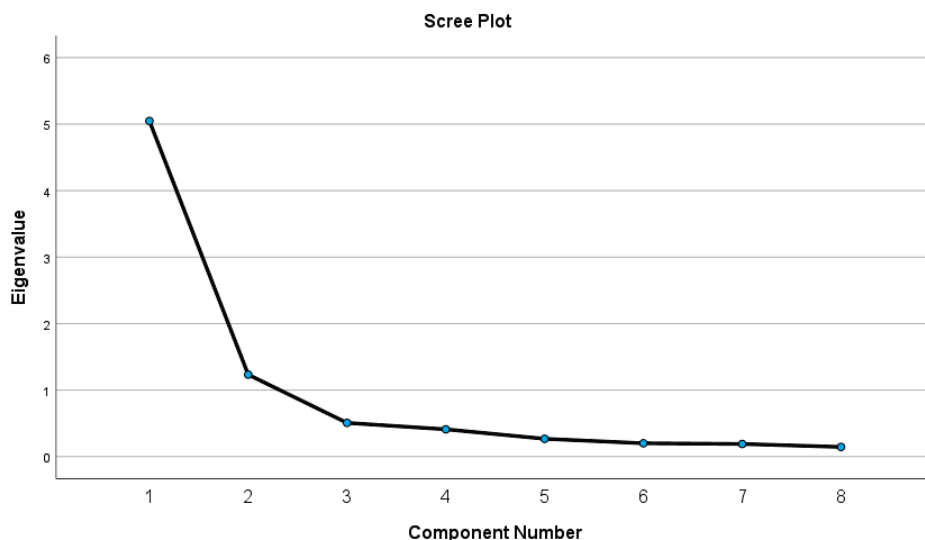


Figure 2. Scree plot.

Table 7. Principal component analysis pattern matrix rotated solution.

		Component	
		1	2
BMC5	The organization has entered new markets as a result of digitalization and digital transformation	0.931	
BMC4	The organization generates additional revenue from new sources as a result of the implementation of digital technologies	0.927	
BMC7	The organization has introduced new sales channels as a result of digitalization and digital transformation	0.910	
BMC6	The organization provides new services as a result of digitalization and digital transformation	0.860	
BMC8	The organization has introduced new ways of charging for services as a result of digitalization and digital transformation	0.673	
BMC2	The organization has digitalized internal business processes		0.982
BMC3	The organization has digitalized external business processes		0.901
BMC1	The organization cooperates with new partners with the aim of developing new digital solutions		0.689

Extraction method: principal component analysis. Rotation method: oblimin with Kaiser normalization ¹.
¹ Rotation converged in four iterations.

The rotated solution presents a simple structure of the two main factors. The interpretation of the extracted two-factor solution is consistent with the digitalization and innovation attributes in the questionnaire, and both exhibit strong loadings (>0.90). Based on the pattern-matrix-rotated solution, latent variables for component 1 as innovation and process digitalization for component 2 were defined.

As a result of PCA analysis, regression factor scores were calculated for each sample case for both solution factors, which were then taken for a further correlation analysis to confirm the correlation between two new component solution factors (innovation and process digitalization) and the productivity increase. The results of this final correlation analysis are presented in Table 8. Since the PROD variable was of the ordinal type, Spearman rho correlation analysis was used. Preliminary analysis showed the relationship to be monotonic, as assessed by a visual inspection of the scatterplot.

Table 8. Correlations.

			PROD	REGR Factor Score 1	REGR Factor Score 2
Spearman's rho	The organization has increased productivity by introducing digital transformation (PROD)	Correlation coefficient Sig. (2-tailed) N	82		
	REGR factor score 1 (innovation)	Correlation coefficient Sig. (2-tailed) N	0.314 ** 0.004 82	82	
	REGR factor score 2 (process digitalization)	Correlation coefficient Sig. (2-tailed) N	0.443 ** <0.001 82	0.453 ** <0.001 82	82

** Correlation is significant at the 0.01 level (2-tailed).

The Spearman's rank-order correlation analysis results confirmed the statistically significant correlation between innovation and the productivity increase, $r_s = 0.314$, $p = 0.004$, and between process digitalization and the productivity increase $r_s = 0.443$, $p < 0.001$.

4. Discussion

The transformation of the maritime industry through digital technologies is no longer a future ambition but a present necessity, shaped by the rapid evolution of logistics networks, stakeholder expectations, and global economic pressures. The empirical results of this study add a valuable dimension to our understanding of how digital transformation—when implemented through business model changes—can serve as a mechanism for increase of productivity in maritime sector organizations.

The results confirm the statistically significant correlation between digital transformation business model changes and an increase in productivity. The study identified two key factors that mediate the relationship between digital transformation and productivity in maritime organizations: innovation and Process digitalization. The factors were extracted through the principal component analysis of the eight variables related to the digital transformation business model changes. The correlation analysis also confirmed the significant relationship between these two factors and the increase in productivity. The findings of the study correspond with existing recent literature, which suggests that digital transformation can lead to improved operational efficiency and customer engagement, thereby increasing productivity [1].

The innovation factor, which encompasses the organization's ability to generate additional revenue from new sources, enter new markets, introduce new sales channels, and provide new services as a result of digital transformation, is confirmed to be significantly correlated with a productivity increase. This suggests that maritime organizations that innovate in their digital transformation are likely to experience a significant increase in productivity. This finding is in line with the broader literature on digital transformation, which suggests that innovation is a critical component of successful digital transformation efforts [2].

The process digitalization factor, which includes the digitalization of internal and external business processes and collaboration with new partners to develop new digital solutions, also correlates significantly with increased productivity. This indicates that maritime organizations that effectively digitalize their business processes can achieve a higher level of productivity. This indication is also consistent with the existing literature, which indicates that the digitalization of processes can lead to improved operational efficiency and productivity [2].

A key insight emerging from this research is the multidimensional nature of the transformation itself. Rather than being an isolated or uniform change, digital transformation manifests itself across a spectrum of activities, ranging from shifts in organizational behavior and external partnerships to the reimagining of revenue logic. This complexity supports the view expressed by [4,5] that transformation is not simply a question of adopting new technologies, but also involves a reevaluation of organizational logic, structures, and interactions with external ecosystems.

In the context of the maritime sector—often portrayed as conservative and resistant to rapid change—the identification of several statistically meaningful correlations between transformation variables and productivity provides a promising outlook. These correlations underscore the tangible effects that properly directed digital initiatives can have, even within highly regulated and operationally intensive industries such as port operations, shipping, and multimodal logistics.

The broader literature echoes the importance of this finding. For instance, successful innovation adoption in ports is frequently contingent on a combination of the internal capacity (such as management support and resource availability) and external enablers (like stakeholder alignment and policy frameworks) [13]. The alignment of such multidirectional efforts is reflected in this study's findings, particularly in the interconnectedness of variables across internal operations, the service design, and market interactions.

Furthermore, the methodological framework adopted in this research—leveraging the TOE model—proves effective in dissecting the drivers behind transformation. The technological, organizational, and environmental categories outlined in the TOE framework reveal how interdependent elements shape transformation outcomes. For example, the presence of technological infrastructure without accompanying organizational readiness, or vice versa, may limit the effectiveness of any digital strategy. These interrelations are supported by the researchers who found that digital maturity in shipping firms was not only a matter of infrastructure but also a function of managerial commitment and institutional adaptability [1].

One important area for interpretation lies in the evolution of service delivery in the maritime transport domain. The literature further demonstrates that digitalization efforts—such as mobile ticketing systems, chatbot deployment, or digital freight platforms—represent more than incremental process enhancements; they constitute foundational shifts in how value is delivered to customers and partners [11]. When mapped to the data in this study, these examples highlight the strategic dimension of transformation, where innovation is deployed not merely for efficiency but also for differentiation and competitiveness.

Additionally, the significance of networked innovation—wherein organizations collaborate across traditional boundaries to co-develop or co-deliver value—cannot be overstated. The data analyzed here suggest that inter-organizational cooperation plays a critical role in facilitating access to new markets, developing tailored digital solutions, and unlocking synergies in logistics chains. This is particularly relevant in maritime clusters, where port authorities, carriers, customs agencies, and third-party logistics providers must operate in coordination to achieve true digital cohesion. Prior research supports this notion by illustrating how the smart port concept relies on such collaborations to achieve shared efficiencies and responsiveness [16,18].

From a theoretical standpoint, the results may also be viewed through the lens of the Dynamic Capabilities Theory, which posits that an organization's ability to sense, seize, and reconfigure resources in a rapidly changing environment determines its competitive position. In this context, digital transformation is a means of enhancing dynamic capabilities, especially in relation to strategic alignment, resource orchestration, and customer responsiveness.

Interestingly, while the study employed a quantitative approach, the patterns it reveals point toward deeper organizational dynamics that might be more fully unpacked through qualitative methodologies. For instance, what motivates certain organizations to adopt more aggressive transformation strategies than others? What internal cultural or leadership factors facilitate or hinder this process? These are questions best addressed through longitudinal case studies or ethnographic inquiry, offering a richer understanding of the change process.

Another avenue for reflection is the role of resilience in the digital strategy. As the maritime industry faces mounting pressures from climate change, global disruptions (e.g., pandemics), and cybersecurity threats, digital transformation must also be framed as a resilience-building activity. The literature suggests that digital orchestration mechanisms during crises can help firms navigate uncertainty more effectively [20]. This view encourages future research to explore how transformation not only boosts productivity in stable conditions but also fortifies organizations against volatility.

5. Conclusions

The findings of this study underscore the significance of business model changes, particularly innovation and process digitalization, as key factors of productivity enhancement in Croatian maritime transport stakeholders undergoing digital transformation. The BMC construct consists of eight distinct questions because digital transformation in maritime organizations is multifaceted, involving numerous interrelated changes in business models. Digital transformation is not a singular technological shift but involves comprehensive innovation and process alterations that impact various organizational aspects. The construct thus includes questions covering key dimensions, such as cooperation with partners for digital solutions (BMC1), internal (BMC2) and external (BMC3) business process digitalization, the creation of new revenue streams (BMC4), market entry through digital transformation (BMC5), the provision of new services (BMC6), the introduction of new sales channels (BMC7), and the implementation of novel pricing models (BMC8). Each of these questions represents a specific business model innovation or operational area directly impacted by digital transformation. The thoroughness of this construct, validated by a high Cronbach's alpha (0.904), demonstrates its relevance and internal consistency, capturing the comprehensive effects digital transformation has on business models within maritime transport stakeholders.

The strong positive correlations identified between productivity and specific transformation activities—such as internal and external process digitalization (BMC2, BMC3), revenue generation from new sources (BMC4), and the introduction of new services and pricing models (BMC6, BMC8)—reaffirm the multifaceted benefits of digital innovation across operational and strategic levels.

These results align with the broader literature emphasizing that digital transformation, when coupled with business model innovation, leads to superior organizational performance and efficiency. The identification of two principal components is consistent with previous studies that suggest that digital business model innovation involves both the reconfiguration of value creation mechanisms and the implementation of enabling technologies. These two main components—**innovation** (e.g., new services, markets, revenue streams) and **process digitalization** (e.g., internal/external process automation, digital partnerships)—were identified through principal component analysis as key mediators of this relationship. The research demonstrates strong statistical correlations between these components and productivity gains, highlighting that digital transformation must go beyond technology adoption to include the strategic rethinking of value delivery. Compared to Western European ports, Croatia is still in the early stages of digital maturity, facing

challenges such as fragmented systems and conservative organizational cultures. However, there is growing momentum, driven by EU initiatives and national policies, to modernize operations and better integrate into the European digital maritime logistics ecosystem.

Specifically, the positive relationship between process digitalization and productivity supports the growing body of evidence that digital technologies such as IoT, AI, and integrated ERP systems streamline operations, reduce errors, and enhance coordination. Likewise, the innovation component—characterized by new services, markets, and revenue streams—mirrors trends in digitally mature organizations that pursue service diversification and customer-centric strategies to stay competitive.

Importantly, the strong statistical significance found in correlations suggests that digital transformation cannot be effective in isolation; it must be accompanied by strategic rethinking of how value is delivered and captured. This is especially relevant in the maritime sector, where traditional practices often hinder agility and integration. The results thus echo calls for a holistic transformation framework that combines technological readiness, organizational capability, and a conducive external environment—core pillars of the TOE framework used in this study.

In practice, this means that maritime transport stakeholders should prioritize investments not just in digital tools but also in organizational change management, workforce upskilling, and collaborative innovation with external partners. As digital ecosystems become more interconnected, fostering interoperability and secure data exchange becomes crucial for realizing productivity gains across the value chain.

While this study provides valuable insights into the relationship among digital transformation, business model innovation, and productivity in maritime transport stakeholders, several limitations should be acknowledged. First, the research is based on a cross-sectional survey design, which captures data at a single point in time and limits the ability to establish causality between digital transformation efforts and productivity outcomes. Longitudinal studies would be more suitable for tracking the dynamic and evolving nature of digital transformation. Second, the sample is geographically restricted to a single country, Croatia, which, although strategically significant in European maritime logistics, may limit the generalizability of findings to other regions with different regulatory, technological, and economic environments. Furthermore, the sample size is fairly limited, although it is of an appropriate size for the statistical methods used. On the other hand, it has limitations as it includes organizations of different sizes. To overcome this issue the sample size could be larger to include more medium and large organizations, or it could be revenue- or turnover-weighted. Third, the reliance on self-reported data may introduce subjective bias, particularly in responses concerning productivity improvements and strategic outcomes. Additionally, the study uses a single-item measure for productivity, which, although supported by the internal consistency of related constructs, may not capture the full complexity of performance metrics in maritime transport stakeholders and may limit the robustness of the findings. The Technology–Organization–Environment (TOE) framework was particularly relevant to this study because it offers a structured approach for examining the complexity of digital transformation initiatives. We have employed the TOE framework introduced by Tornatzky and Fleischer precisely due to its ability to effectively analyze three critical dimensions: technological adoption, organizational preparedness, and environmental influences. In the maritime sector, these dimensions play crucial roles, as digital transformation often faces technological barriers (e.g., legacy systems, cybersecurity), organizational resistance (e.g., cultural inertia, lack of management support), and environmental complexities (e.g., regulatory compliance, stakeholder coordination). By utilizing the TOE framework, influencing factors have been systematically identified, categorized, and analyzed. They mediate the relationship between digital transformation-

induced business model changes and productivity outcomes. This structured analysis thus makes the TOE framework especially suitable for the comprehensive evaluation of digital transformation's multidimensional impact on maritime transport organizations, providing clarity and coherence to the empirical findings of the research. However, while the TOE framework provides a structured approach to analyzing digital transformation, it may not fully account for the nuanced, industry-specific drivers of innovation in maritime contexts, such as geopolitical factors or evolving sustainability regulations.

For port authorities and maritime firms, the findings suggest that investing in digital transformation should prioritize not only technology upgrades but also organizational innovation and collaboration. Implementing internal and external process digitalization—such as automating documentation, adopting real-time tracking, or integrating ERP systems—can significantly boost efficiency. Moreover, developing new digital services, revenue models, and partnerships enables firms to remain competitive and responsive to changing logistics demands. These strategic shifts are essential for aligning with EU digitalization goals and improving integration into trans-European transport networks.

Future research could involve expanding the scope of productivity metrics beyond self-reported perceptions to include objective performance indicators (e.g., cargo throughput, turnaround time, cost savings), which would enhance the robustness of future analyses. Qualitative methods such as case studies or interviews could complement the survey findings by capturing the lived experiences of stakeholders and providing a deeper understanding of organizational change dynamics. Additionally, integrating advanced analytical methods—such as structural equation modeling (SEM) or machine learning—could offer more precise models of causality and prediction. Finally, future studies might examine the role of digital skills, leadership, and cross-sector collaboration as mediators or moderators in the relationship between digital transformation and business outcomes, further enriching the theoretical and practical contributions of this research stream. As digital transformation advances, maritime organizations must address the growing complexity of securing digital infrastructures and navigating inconsistent regulatory frameworks across jurisdictions. A more thorough analysis of these risks would provide valuable guidance for mitigating barriers to innovation and ensuring resilient, compliant operations.

Overall, the study contributes to a growing understanding of how targeted digital strategies can yield tangible productivity benefits in the maritime sector. It also provides a structured empirical foundation for policy-makers and industry leaders aiming to accelerate digital maturity and competitive positioning in a sector that is increasingly under pressure to innovate sustainably.

Author Contributions: Conceptualization, S.A. and M.J.M.; methodology, G.L.; software, G.L.; validation, E.T.; formal analysis, G.L. and S.A.; investigation, S.A. and M.J.M.; resources, G.L. and M.J.M.; data curation, M.J.M. and G.L.; writing—original draft preparation, S.A. and M.J.M.; writing—review and editing, E.T.; visualization, G.L.; supervision, E.T.; project administration, S.A. and E.T.; funding acquisition, E.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the project line ZIP UNIRI of the University of Rijeka, Croatia, for the project UNIRI-ZIP-2103-3-22.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

TOE	Technology–Organization–Environment (framework)
PCA	Principal Component Analysis
SPSS	Statistical Package for the Social Sciences
BMC	Business Model Change (used as BMC1–BMC8 in variables)
PROD	Productivity (variable representing perceived productivity increase)
ETA	Estimated Time of Arrival
IoT	Internet of Things
AI	Artificial Intelligence
ERP	Enterprise Resource Planning
RPA	Robotic Process Automation
DMLBC	Dependable Machine Learning for Seaports Using Blockchain (a method)
SEM	Structural Equation Modeling
AIS	Automatic Identification System (used for vessel tracking)

References

1. Nguyen, T.-H. Investigating Driving Factors of Digital Transformation in the Vietnam Shipping Companies: Applied for TOE Framework. *Sage Open* **2024**, *14*, 21582440241301210. [CrossRef]
2. Teng, X.; Wu, Z.; Yang, F. Research on the Relationship between Digital Transformation and Performance of SMEs. *Sustainability* **2022**, *14*, 6012. [CrossRef]
3. Jović, M.; Tijan, E.; Vidmar, D.; Pucihar, A. Factors of Digital Transformation in the Maritime Transport Sector. *Sustainability* **2022**, *14*, 9776. [CrossRef]
4. Verhoef, P.C.; Broekhuizen, T.; Bart, Y.; Bhattacharya, A.; Qi Dong, J.; Fabian, N.; Haenlein, M. Digital Transformation: A Multidisciplinary Reflection and Research Agenda. *J. Bus. Res.* **2021**, *122*, 889–901. [CrossRef]
5. Fruth, M.; Teuteberg, F. Digitization in Maritime Logistics—What Is There and What Is Missing? *Cogent Bus. Manag.* **2017**, *4*, 1411066. [CrossRef]
6. Heilig, L.; Lalla-Ruiz, E.; Voß, S. Digital Transformation in Maritime Ports: Analysis and a Game Theoretic Framework. *NETNOMICS Econ. Res. Electron. Netw.* **2017**, *18*, 227–254. [CrossRef]
7. Tijan, E.; Jović, M.; Aksentijević, S.; Pucihar, A. Digital Transformation in the Maritime Transport Sector. *Technol. Forecast. Soc. Change* **2021**, *170*, 12087. [CrossRef]
8. Drazin, R. The Processes of Technological Innovation. *J. Technol. Transf.* **1991**, *16*, 45–46. [CrossRef]
9. Li, K.; Rollins, J.; Yan, E. Web of Science Use in Published Research and Review Papers 1997–2017: A Selective, Dynamic, Cross-Domain, Content-Based Analysis. *Scientometrics* **2018**, *115*, 1–20. [CrossRef]
10. Behdani, B.; Wiegman, B.; Roso, V.; Haralambides, H. Port-Hinterland Transport and Logistics: Emerging Trends and Frontier Research. *Marit. Econ. Logist.* **2020**, *22*, 1–25. [CrossRef]
11. Notteboom, T.; Pallis, A.; Rodrigue, J.-P. Port Hinterlands, Regionalization and Corridors. In *Port Economics, Management and Policy*; Routledge: London, UK, 2021; ISBN 9780429318184. Available online: <https://Porteconomicsmanagement.Org/Pemp/Contents/Part2/Port-Hinterlands-Regionalization/> (accessed on 20 May 2025).
12. Wide, P.; Roso, V. Information on Resource Utilisation for Operational Planning in Port Hinterland Transport. *Trans. Marit. Sci.* **2021**, *10*, 477–487. [CrossRef]
13. Sooprayen, K.; Van de Kaa, G.; Pruyn, J.F.J. Factors for Innovation Adoption by Ports: A Systematic Literature Review. *J. Ocean Eng. Mar. Energy* **2024**, *10*, 953–962. [CrossRef] [PubMed]
14. Kaczerska, D.; Kizielewicz, J.; Skrzyszewska, K. Digital Transformation in Ferry Shipping—Case Study in the Baltic Sea Region. *TransNav Int. J. Mar. Navig. Saf. Sea Transp.* **2023**, *17*, 553–560. [CrossRef]
15. Pang, Q.; Cai, L.; Wang, X.; Fang, M. Digital Transformation as the Fuel for Sailing toward Sustainable Success: The Roles of Coordination Mechanisms and Social Norms. *J. Enterp. Inf. Manag.* **2024**, *37*, 1069–1096. [CrossRef]
16. Belmoukari, B.; Audy, J.-F.; Forget, P. Smart Port: A Systematic Literature Review. *Eur. Transp. Res. Rev.* **2023**, *15*, 4. [CrossRef]
17. Guerrero-Molina, M.-I.; Vázquez-Suárez, Y.-A.; Valdés-Mosquera, D.-M. Smart, Green, and Sustainable: Unveiling Technological Trajectories in Maritime Port Operations. *IEEE Access* **2024**, *12*, 47713–47723. [CrossRef]
18. Basulo-Ribeiro, J.; Teixeira, L. Industry 4.0 Supporting Logistics towards Smart Ports: Benefits, Challenges and Trends Based on a Systematic Literature Review. *J. Ind. Eng. Manag.* **2024**, *17*, 492–515. [CrossRef]

19. Durán, C.; Fernández-Campusano, C.; Carrasco, R.; Carrillo, E. DMLBC: Dependable Machine Learning for Seaports Using Blockchain Technology. *J. King Saud Univ.—Comput. Inf. Sci.* **2024**, *36*, 101918. [CrossRef]
20. Gao, X.; Kong, Y.; Cheng, L. Strategies and Mechanisms for Building Digital Resilience of Container Shipping Platform in Crisis Situations: A Network Orchestration Perspective. *Ocean Coast. Manag.* **2023**, *246*, 106887. [CrossRef]
21. Yang, C.-S.; Lin, M.S.-M. The Impact of Digitalization and Digital Logistics Platform Adoption on Organizational Performance in Maritime Logistics of Taiwan. *Marit. Policy Manag.* **2023**, *51*, 1884–1901. [CrossRef]
22. Yu, H.; Zhang, M.; Cui, C.; Xu, L.; Lin, S.; Xu, J. Evaluating Competitiveness of Container Shipping Operators in the Sustainability and Digitalization Era. *Sustainability* **2024**, *16*, 4228. [CrossRef]
23. Bucak, U. Expected Utilities of Liner Shipping Market Trends: How Can Companies Benefit? *Int. J. Shipp. Transp. Logist.* **2024**, *18*, 92–110. [CrossRef]
24. Lam, H.; Tang, V. Digital Transformation for Cold Chain Management in Freight Forwarding Industry. *Int. J. Eng. Bus. Manag.* **2023**, *15*, 18479790231160857. [CrossRef]
25. Aerts, G.; Mathys, G. Discovering Trends in the Digitalization of Shipping: An Exploratory Study into Trends Using Natural Language Processing. *J. Mar. Sci. Eng.* **2024**, *12*, 618. [CrossRef]
26. Evmides, N.; Aslam, S.; Ramez, T.T.; Michaelides, M.P.; Herodotou, H. Enhancing Prediction Accuracy of Vessel Arrival Times Using Machine Learning. *J. Mar. Sci. Eng.* **2024**, *12*, 1362. [CrossRef]
27. Su, Z.; Liu, Y.; Gao, Y.; Park, K.-S.; Su, M. Critical Success Factors for Green Port Transformation Using Digital Technology. *J. Mar. Sci. Eng.* **2024**, *12*, 2128. [CrossRef]
28. Gündoğan, O.; Keçeci, T. A TAM-Based Study on the Adoption of Digital Transformation in the Maritime Transportation Logistics Sector. *J. ETA Marit. Sci.* **2024**, *12*, 92–105. [CrossRef]
29. Melnyk, O.; Onyshchenko, S.; Onishchenko, O.; Lohinov, O.; Ocheretna, V. Integral Approach to Vulnerability Assessment of Ship's Critical Equipment and Systems. *Trans. Marit. Sci.* **2023**, *12*, 1–10. [CrossRef]
30. DeVellis, R.F. *Scale Development: Theory and Applications*, 2nd ed.; Sage Publications: Thousand Oaks, CA, USA, 2003.
31. Dunteman, G.H. *Principal Components Analysis*; SAGE Publications: Thousand Oaks, CA, USA, 1989; ISBN 9780803931046.
32. Cattell, R.B. The Scree Test for the Number of Factors. *Multivar. Behav. Res.* **1966**, *1*, 245–276. [CrossRef] [PubMed]
33. Kaiser, H.F. The Application of Electronic Computers to Factor Analysis. *Educ. Psychol. Meas.* **1960**, *20*, 141–151. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Optimizing Stack-Yard Positioning in Full Shoreline Loading Operations

Xueqiang Du ¹, Bencheng Luo ¹, Jing Wang ¹, Jieting Zhao ¹, Dahai Li ¹, Qian Sun ¹ and Haobin Li ^{2,*}

¹ China Waterborne Transport Research Institute, Beijing 100088, China; duxueqiang@wti.ac.cn (X.D.); luobc@wti.ac.cn (B.L.); wangjing@wti.ac.cn (J.W.); zhaojieting@wti.ac.cn (J.Z.); lidahai@wti.ac.cn (D.L.); sunqian@wti.ac.cn (Q.S.)

² Centre of Excellence in Modelling and Simulation for Next Generation Ports, National University of Singapore, Singapore 119077, Singapore

* Correspondence: li_haobin@nus.edu.sg

Abstract: Loading operations are a crucial part of container terminal activities and play a key role in influencing shoreline operation efficiency. To overcome the challenge of mismatched local ship decisions and global yard decisions during single-vessel operations, which often result in conflicts related to container retrieval in the yard, a novel intelligent decision-making model for stack-yard positioning in full shoreline loading operations is proposed. This model seeks to optimize the balance between yard operation instructions and quay crane operation instructions. An enhanced Constrained Optimization Genetic Algorithms-Greedy Randomized Adaptive Search (COGA-GRASP) algorithm is introduced to tackle this decision-making issue, and it is applied to identify the most optimal bay configuration for full shoreline loading operations. The proposed model's effectiveness is validated through testing and solution outcomes.

Keywords: full shoreline; loading operations; intelligent decision-making; yard location optimization; instruction sequence

1. Introduction

Loading operations are a critical component of container terminal operations, directly influencing the efficiency of shoreline activities. As a dynamic process, loading operations involve continuous changes in the yard state due to tasks such as loading, unloading, container gathering, pickup, and relocation. Simultaneously, the pending tasks of vessels evolve, increasing the computational complexity of decision-making. To maintain operational continuity, instructions must be activated promptly to avoid stagnation in yard or shoreline operations. Therefore, developing intelligent decision-making models and algorithms that balance computational speed and solution quality is essential for optimizing full shoreline loading operations.

Currently, as shown in Figure 1, container terminal loading operations typically adopt a single-ship decision-making approach, where the containers to be loaded onto the vessel are distributed throughout the entire yard. This creates a mismatch between the local decision-making for the vessel and the global decision-making for the yard, leading to insufficient overall control, which can negatively impact loading efficiency.

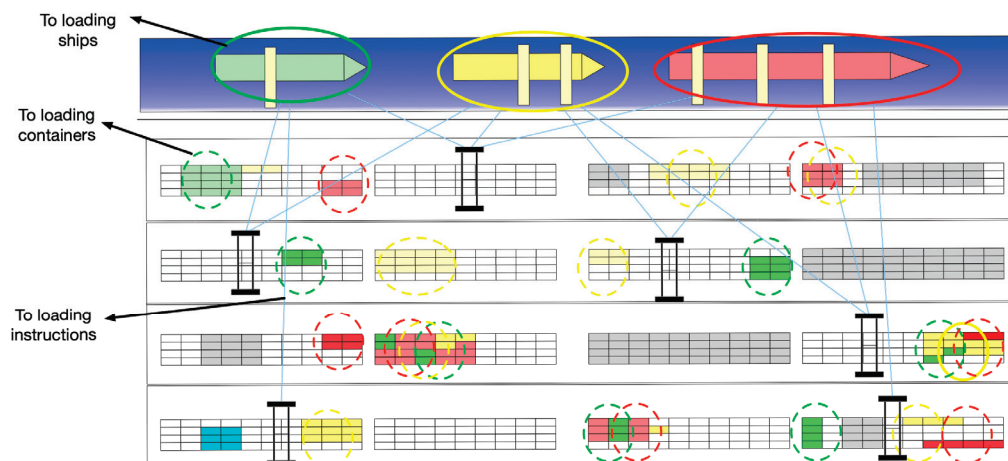


Figure 1. Schematic of loading decision.

To address the issue of the mismatch between local and global strategies, this paper conducts a systematic analysis of terminal loading operations and proposes the concept of full-quay container ship loading decision-making. As shown in Figure 2, it considers the vessel at berth as a unified whole, thereby aligning the global decision-making for the quay and the global decision-making for the yard. However, to implement full-quay loading decision-making, the first challenge is to address the selection of operation locations within the yard for all handling routes. In the loading process, there are multiple available operation locations for each route, which are unevenly distributed across the yard. This requires collaborative computation across all handling routes to select the appropriate yard positions for each. As the number of handling routes increases, the computational difficulty also rises.

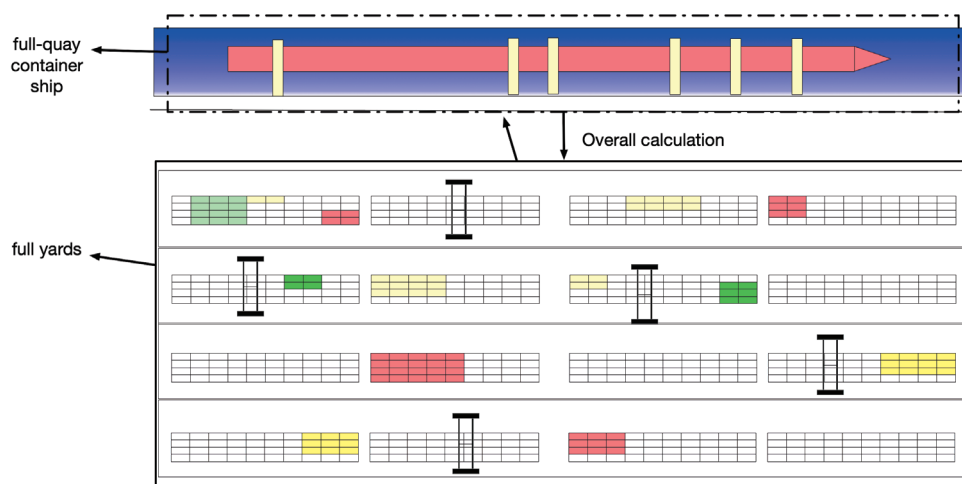


Figure 2. Loading decision-making along the entire coastline.

To address the critical issue of misalignment between local operation strategies and global resource scheduling in port logistics systems, this study conducts systematic modeling analysis on container terminal handling operations based on operations research and combinatorial optimization theories. By constructing a multi-objective optimization model, we innovatively propose the theoretical framework of “Full-quay Collaborative Loading Decision-making”. This framework transcends the limitations of traditional segmented scheduling models by considering the berthed vessel cluster as a dynamic operational entity with spatiotemporal coupling characteristics. The research focuses on overcoming the collaborative optimization challenges between terminal front operations and rear yard

resources, achieving integrated optimization of terminal-wide scheduling decisions and yard space allocation strategies through a planning model. At the algorithmic implementation level, this paper specifically addresses the approximation of optimal solutions for the NP-Hard problem of yard operation location selection. Notably, when concurrent operation routes increase exponentially with growing port throughput, the solution space dimension expands drastically to posing severe challenges to traditional optimization algorithms.

Based on the above analysis, this paper proposes a full-quay container ship loading decision-making model for yard location selection, which is solved using the improved COGA and GRASP algorithms. By applying the above algorithm, the efficiency of the container terminal's quay crane loading operation has increased by about 2%, and the yard operation has become smoother. The contributions of this paper are as follows:

- This paper develops a full-quay container ship loading yard location selection model to dynamically address the location selection problem during the loading process.
- By integrating the COGA and GRASP algorithms, an improved COGA-GRASP algorithm is designed to enhance both solution efficiency and quality.
- Through comparative analysis of different weightings and algorithms, the robustness of the proposed algorithm is validated.

Section 2 reviews previous research and provides a comparative analysis. Section 3 develops the full-quay container ship loading yard location selection model. Section 4 presents the improved algorithm design for solving the problem. Section 5 validates the model and algorithm's effectiveness through comparisons of different weightings and solution algorithms. Section 6 concludes the paper and proposes directions for future research.

2. Literature Review

To enhance the efficiency of container terminal loading operations, several researchers have investigated the optimization of operation instruction sequences [1–19]. Some scholars have proposed optimization models that integrate quay cranes, yard cranes, and horizontal transportation for improved operational efficiency [20–45]. Its research objectives include operational efficiency and consumption, and the research scenarios include traditional terminals and U-shaped terminals.

2.1. Instruction Optimization

Wen et al. [1] investigated the scheduling problem of container terminal loading operations, modeling and solving the problem by considering the impact of the quay crane spreader trajectory on the operation sequence.

Zhou et al. [2] addressed a combined optimization problem for continuous loading and unloading operations on a floating cargo handling platform. While the research does not directly focus on the container terminal loading process, the established goal is to minimize vessel turnaround time using a Markov decision process model. Additionally, a tabu-based tree search algorithm was developed and improved to control the depth and breadth of the local tree structure, which remains valuable for reference.

Rahman et al. [3] investigated the optimization of port container handling, integrating dual-cycle crane operations and dockyard rehandling minimization. A unified model was proposed, encompassing two operations: one for unloading and loading containers via quay cranes, and another for reducing rehandling at the dock during loading. A hybrid genetic algorithm (GA), QCDC-DR-GA, composed of one-dimensional and two-dimensional GAs, was developed and validated for solving the problem.

Yan et al. [4] proposed a deep reinforcement learning algorithm employing directed search for the CRP problem, where a deep neural network is trained by DRL to learn

container repositioning strategies. Directed search is integrated into the algorithm to improve the quality of the solution and hence the efficiency of the loading operation.

Zhang et al. [5] develops a data-driven model to predict the Probabilistic Container Repositioning Problem (PCRP) for the overturning of containers during ship loading, and proposes an adaptive Monte Carlo tree search algorithm. The expected number of container repositionings is minimized by integrating a new heuristic: local safety and global flexibility.

Parreño-Torres et al. [6] developed two exact solution methods to minimize crane operation time. One method was an integer linear model, and the other a branch-and-bound algorithm. These methods introduced new upper and lower bounds, dominance criteria, and heuristic programs, providing optimal solutions for problems of practical scale.

Xia et al. [7] proposed a Double Deep Q-Network (DDQN) with priority experience replay to address the container terminal loading sequence problem. In this approach, simulations with varying complexity are introduced to model the loading process. The proposed PERDDQN, trained on real-world data, can solve the loading sequence problem in seconds. The results show that the algorithm is highly efficient and versatile.

Cai et al. [8] designed a container rehandling algorithm for ship loading based on inverse reinforcement learning, aiming to minimize the number of container relocations in the yard and the waiting time of horizontal transportation vehicles at the quay.

Xia et al. [9], based on sequential decision theory, decomposed the ship loading process into container retrieval and instruction sequencing decisions, solving each sequentially.

Yue et al. [10] introduced a dynamic scheduling process to solve the AGV scheduling and path planning problem, in which the scheduling scheme determines the starting and ending nodes of the path, and the selection of the paths between nodes affects the scheduling of subsequent AGVs. Under the loading and unloading time constraints, a two-stage mixed integer optimization model is proposed with the objective of minimizing the AGV transportation cost. A dynamic optimization algorithm including an improved rule-based heuristic algorithm and a dynamic optimization algorithm integrating Dijkstra's algorithm and Q-Learning algorithm are designed to solve the optimal AGV scheduling and path schemes. A novel conflict avoidance strategy based on graph theory is also proposed to reduce the probability of path conflicts among AGVs.

Fibrianto et al. [11], in the context of automated container terminals equipped with overhead shuttle cranes (OS), proposed a job scheduling model and heuristic methods. By categorizing each operation as a main operation, pre-unloading, or re-unloading task, the model aims to reduce total delay times for flatbeds and external trucks. The Operation Sequencing Job Scheduling Problem (OSJSPS) is formulated as a Mixed Integer Programming (MIP) model, which simultaneously sequences a set of jobs and explores the possibility of separating them into pre-sorting and re-sorting tasks. A two-stage genetic algorithm (TGA) is proposed, incorporating two local improvement procedures: an iterative local search and an opportunistic task separation procedure. Case studies validate the effectiveness of the model.

Oladugba et al. [12] proposed a novel algorithm based on the Modified Johnson Algorithm (JA) for the job instruction scheduling problem of a Twin Automatic Stacking Crane (Twin-ASC), which employs a Mixed Integer Programming (MIP) model for solving the Twin-ASC scheduling problem in order to minimize the scheduling time span.

Liu et al. [13] considering the fluctuation of the arrival time of external trucks and the change in the loading and unloading volume of yard cranes, a stochastic programming model is established for the yard crane scheduling problem, and a customized genetic algorithm is used to find the optimal scheme with good adaptability to uncertain scenarios.

Yu et al. [14] built an integer model for the problem of space allocation for incoming containers at an automated container terminal with the goal of minimizing the waiting time for container trucks, and developed a genetic algorithm embedded in simulation to solve it.

Yang et al. [15] considered the integration and optimization of yard resources in light of the uncertainties associated with container demand, establishing a stochastic optimization model to enhance the robustness of daily yard management. He designed a harmony search algorithm based on a two-stage decomposition strategy to solve the proposed model. To accelerate the solving speed, a scoring method was introduced to generate initial space allocation plans, and a tabu list was added to improve the quality of solutions.

Gulić et al. [16] studied the two-dimensional, static, offline, and constrained container repositioning problem in actual-sized yard bays for the box-changing issue. He proposed a new method using a Genetic Algorithm (GA) to solve the container relocation problem, aiming to minimize the number of container relocations.

Yang et al. [17] addressed the scheduling problem of electric AGVs from the perspective of energy conservation and emission reduction by proposing a speed control strategy that takes into account the port traffic environment. Additionally, the charging capacity of the port was discretized to simulate the limited processing capability of the port and avoid congestion at battery swapping stations. To minimize the delay costs and carbon emissions associated with AGV operations, a mixed-integer programming model was established. This model optimizes job efficiency and carbon emissions by assigning and prioritizing container transportation tasks and AGV battery swapping tasks. An improved genetic algorithm-based method was designed, incorporating a greedy strategy to obtain better initial solutions, while simulated annealing was used for population selection to prevent the algorithm from becoming trapped in local optima. Furthermore, an adaptive adjustment strategy for crossover and mutation probabilities was employed to enhance the convergence of the algorithm.

Xu et al. [18] developed a mixed-integer programming model for the integrated scheduling of quay cranes, AGVs, and dual cantilever rail cranes during the unloading process at a U-shaped automated container terminal. The model's objectives are to minimize the loading and unloading times of various tasks and avoid AGV conflicts. It integrates discrete event dynamics models and continuous time dynamics models. Additionally, an improved Genetic Seagull Optimization Algorithm (GSOA) was designed to solve this complex optimization problem.

Wang et al. [19] studied the scheduling and routing problem of Automated Guided Vehicles (AGVs) with multiple bidirectional paths to generate conflict-free routes. A mixed-integer programming (MIP) model was established to minimize the makespan, i.e., the completion time of all operations. A customized branch-and-bound (B&B) algorithm was developed to solve the problem within a reasonable time frame. The B&B process was used to obtain job assignments and job sequences, while a heuristic algorithm was proposed to generate conflict-free routes.

2.2. Scheduling Optimization

Lu Y et al. [20] proposed an improved genetic algorithm (GA)-based integrated optimization method for automated container terminal scheduling. The method aims to minimize the total operation time of container loading and unloading tasks at automated terminals by constructing a three-stage integrated optimization model. An improved genetic algorithm, named PGA (Probabilistic Genetic Algorithm), was developed to solve the model, which alters the population distribution to accelerate convergence. The constructed model is of significant reference value for determining the optimal allocation of

QC (quay cranes), AGV (automated guided vehicles), and ASC (automated stacking cranes) in automated container operations.

Niu et al. [21] presented a multi-objective integrated container terminal scheduling problem, considering berth allocation, yard crane assignment, and horizontal transportation. The objective is to shorten ship operation time and reduce quay crane operation costs. He proposed a multi-objective bacterial colony optimization algorithm (MOBCO) incorporating concepts such as multiple populations, topology, personal optimal, and global optimal. The algorithm was verified, and it demonstrated superior performance, especially for large-scale container terminals.

Luo et al. [22] integrated the scheduling of automated container terminal handling equipment with the goal of minimizing ship berth time. He constructed a mixed integer programming (MIP) model, which can optimize small-scale problems using existing solvers. For large-scale problems, an adaptive heuristic algorithm was developed to adjust the parameters of the genetic algorithm (GA) based on observed performance, aiming to approximate the optimal solution.

Li et al. [23] utilized a double-cycle container loading and unloading strategy to coordinate quay cranes, yard cranes, and horizontal transport equipment. He proposed a mixed integer linear programming (MILP) model to minimize the operation time for multiple devices. Additionally, an enhanced Benders decomposition algorithm was designed to solve the problem, along with several acceleration measures to speed up convergence.

Yin et al. [24] studied the coordinated scheduling problem of quay cranes and shuttle vehicles, which involves decisions on assigning tasks from the apron to the quay crane (QC), ordering QC operations, allocating tasks to shuttle vehicles (SV), and ordering operations for the SV. A mixed integer linear programming (MILP) model was developed to minimize operation time. The relationship and interaction between four constraint sections (berth scheduling QC, task ordering by QC, task scheduling for SV, and apron buffer capacity constraint per berth) were analyzed, leading to the development of three rules: berth position rule, near-to-far rule, and full-buffer rule. Finally, sequence insertion, greedy insertion, and an improved genetic algorithm were proposed to solve medium and large-scale cases.

Yue et al. [25] addressed the configuration and scheduling problem of dual 40-foot dual-trolley quay cranes (QC) and automated guided vehicles (AGVs) in container terminals. He proposed a two-stage bi-objective mixed integer programming model to maximize customer satisfaction, minimize AGV idle time, and reduce QC delay time. In the proposed model, the first stage's fragmentation function was linearized by adding special ordered set constraints, and the GUROBI solver version 9.5 was used. An improved Non-Dominated Sorting Genetic Algorithm III (NSGA-III) was applied to efficiently solve the second stage. Numerical experiments validated the model and algorithm's effectiveness.

Yu et al. [26] investigated the potential yard space conflicts during the ship loading process and, to minimize conflicts, established a loading cluster selection model. A more flexible allocation strategy was derived to organize the yard space for outbound containers. A bi-objective model was constructed, considering both transportation distance and the balance between loading and unloading across blocks. Additionally, a model was developed to minimize all possible loading and unloading scenarios in the export yard during the loading process.

Zheng et al. [27] addressed the coordinated scheduling problem of quay cranes, yard cranes, and horizontal transportation equipment in container terminal operations. A mixed integer linear programming (MILP) model was constructed, and an improved three-dimensional chromosome representation was designed based on the genetic algorithm (GA) to solve the problem.

He et al. [28] studied the berth allocation and quay crane assignment problem, considering factors such as QC driver costs, diurnal operational efficiency, and performance-based wages. The study analyzed how factors related to QC drivers influence vessel schedules and formulated the objective function to include QC driver costs. A hybrid integer programming model with an accelerated algorithm was developed to address the proposed problem, and a metaheuristic framework incorporating a three-stage algorithm was introduced to solve the problem.

Xu et al. [29] proposed an integrated scheduling optimization model based on mixed integer programming to analyze the dual-trolley quay cranes, conflict-free automated guided vehicles (AGVs), and dual cantilever cranes in U-shaped automated container handling modes. Unlike previous studies, which primarily focused on traditional container terminal coordination, this model is tailored to automated systems. A super-heuristic genetic algorithm based on reinforcement learning was applied to solve the model. The proposed algorithm was compared with dual-layer genetic algorithms, adaptive genetic algorithms, hybrid genetic algorithms, and cuckoo search algorithms to validate its effectiveness.

Jonker et al. [30] considering QCs, automated guided vehicles and shore bridges simultaneously during ship loading, a Hybrid Flow Shop (HFS) approach is used to construct a synergistic schedule, which is solved by a simulated annealing algorithm.

Lu [31] took maximizing the efficiency of automated container terminals as his goal and loading and unloading time as his research object. On the premise of balancing the operating efficiency of each container ship at the berth, he established a comprehensive optimization scheduling model for automated container terminals with proportional fairness priority by setting priority parameters.

Gao et al. [32] established a mathematical models for multi-AGV dispatching, and a QL-CNP algorithm is proposed to tackle the multi-AGV dispatching problem (MADP). The distribution of traffic load is balanced for multiple AGVs performing tasks in the road network. The proposed model is validated using a Gurobi solver with a small experiment. Then, QL-CNP is used to conduct experiments with different sizes.

Cahyono et al. [33] uses a finite state machine (FSM) framework to dynamically model the integrated (end-to-end) operations of a container terminal, where each state machine is represented by a discrete event system (DES). The hybrid model includes the operations of quay cranes, internal trucks and yard cranes, as well as the storage location selection of container yards and ship slots. Using the integrated container terminal hybrid model, we propose a model prediction algorithm to obtain a near-optimal solution to the integrated terminal operation problem.

Shouwen et al. [34] established a bi-level programming model for the integrated scheduling problem with the objective of minimizing the completion time of all containers. He designed two bi-level optimization algorithms based on conflict resolution strategies: a Bi-level Adaptive Genetic Algorithm based on Conflict Resolution Strategy (CRS-BAGA) and a Bi-level Genetic Algorithm based on Conflict Resolution Strategy (CRS-BGA). Finally, numerical experiments were conducted to verify the effectiveness of the two optimization algorithms.

Hsu et al. [35] addressing the collaborative scheduling problem of quay cranes, yard cranes, and container trucks, proposed a simulation-based optimization framework considering vessel stowage factors. This framework employs four heuristic/metaheuristic algorithms: Sorting by Bay (SBB), Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Multi-Group Particle Swarm Optimization (MGPSO). It simultaneously handles the Yard Crane Scheduling Problem (YCSP), Yard Truck Scheduling Problem (YTSP), and Quay Crane Scheduling Problem (QCSP) for export containers, while taking operational constraints into account.

Gong et al. [36] proposed a hybrid multi-AGV scheduling algorithm aimed at minimizing energy consumption in the Automated Container Terminal (ACT) yard and the total completion time of AGVs. This framework models the AGV scheduling process as a Markov Decision Process (MDP). Furthermore, he introduced a novel scheduling algorithm named Multi-Agent Deep Deterministic Policy Gradient for AGV Scheduling (MDAS), based on Multi-Agent Deep Deterministic Policy Gradient (MADDPG), to facilitate online real-time scheduling decisions.

Xin et al. [37], with the primary goal of energy savings for quay cranes and AGVs, established a Mixed-Integer Nonlinear Programming (MINLP) model. He developed an efficient Genetic Algorithm (GA) incorporating a weighted strategy to solve the problem. Additionally, he devised an ϵ -constraint algorithm to analyze the Pareto frontier.

Wang et al. [38] investigates an integrated optimization problem concerning berth allocation, quay crane assignment, and truck deployment in container terminals, with a specific focus on incorporating quay crane maintenance into the integrated models. A nonlinear integer programming model is proposed, and then a number of equivalent or relaxed models are developed to ease the model. To solve the model on large-scale instances, a SWO-GA is developed to provide a solution.

Shen et al. [39], from the point of AGV road blocking, the scheduling mode of group operation area is proposed. In order to minimize the maximum completion time of AGV, an AGV scheduling optimization model is established considering the interference constraints and AGV congestion in the actual operation of the terminal. Hybrid Genetic Algorithm with Fuzzy Logic Controller (HGA-FLC) is used to simulate the behavior of AGVs, and different scale examples are designed to solve the problem.

Fan et al. [40] designed a novel relocation strategy and established a collaborative optimization model to address the yard allocation for inbound and outbound containers and the scheduling of dual automated stacking cranes at automated container terminals. The model considers constraints such as safety distances and buffer zone capacities, aiming to minimize the total completion time of all storage requests for outbound containers and the total reshuffling time during retrieval, according to the priority defined by early retrieval and early loading rules for inbound and outbound containers. A hybrid genetic algorithm based on variable neighborhood search was developed to solve the model.

Jian et al. [41] addressed AGV optimization issues by considering different congestion scenarios along the routes, employing symmetric triangular fuzzy numbers to describe the distribution of AGV operation times. With the objectives of minimizing quay crane delay risk and reducing AGV operation time, a multi-objective scheduling optimization model was established. An improved genetic algorithm was designed to solve this model. The effectiveness of the model and algorithm was validated by comparing the results of AGV scheduling and container storage optimization models under fixed congestion coefficients across different instance scales.

Liu et al. [42] focused on the integrated scheduling of horizontal transportation equipment and handling equipment, as well as AGV path planning, based on a newly constructed U-shaped yard layout at a sea-rail intermodal container terminal. A bi-level programming model was established with the objectives of minimizing the maximum completion time and total waiting time for both AGVs and trucks. The model aims to derive the scheduling plan for yard cranes and determine the specific running trajectories of AGVs. Additionally, a bi-level genetic algorithm based on spatiotemporal greedy strategies was designed to handle task scheduling and conflict resolution.

Hsu et al. [43] proposes a hybrid model, termed WOA + PSO, which combines the Whale Optimization Algorithm (WOA) and Particle Swarm Optimization (PSO) as a novel

approach to address the integrated scheduling problem of automated quay cranes (AQC), automated lift vehicles (ALVs), and automated stacking cranes (ASCs) in an ACT.

Fontes et al. [44] formulates a mixed-integer linear programming (MILP) model and proposes a bi-objective multi-population biased random key genetic algorithm (mp-BRKGA) for the joint scheduling of quay cranes and speed adjustable vehicles in container terminals considering the dual-cycling strategy.

Hsu et al. [45] tackled the joint scheduling problem of yard cranes and container trucks by applying and comparing three optimization algorithms: Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Subgroup PSO (SGPSO). The performance and effectiveness of each algorithm were evaluated to determine the most suitable approach for optimizing the scheduling of yard cranes and container trucks.

2.3. Summary

Table 1 summarizes the reviewed literature. In ship loading operations, many scholars focus on single-ship or single-quay crane as the research objective, often using completion time minimization as the objective function. There is relatively little consideration given to yard operation conflicts or workload balancing, which hampers the ability to effectively ensure the overall loading efficiency of container terminals. In terms of the choice of solution algorithm, scholars use GA and its deformation, multi-objective solution algorithm, deep learning and other methods to solve problems.

Table 1. Literature summary and analysis.

Algorithm Type	Literature	Key Features
Genetic Algorithm (GA) and Variants	[1,3,8,11–14,16–18,20,22,23,25,27,29,31,34,37–41,44]	Includes standard GA, hybrid GA (e.g., embedded tabu search), adaptive GA, multi-objective GA (e.g., NSGA-III), etc.
Swarm Intelligence Algorithms	[21] (MOBCO), [30] (Simulated Annealing), [35] (PSO/MGPSO), [36] (MADDPG), [43] (WOA+PSO), [45] (GA/PSO/SGPSO)	Particle Swarm Optimization (PSO), Whale Optimization Algorithm (WOA), Bacterial Colony Optimization (BCO), etc., often used for multi-objective problems.
Deep Reinforcement Learning (DRL)	[4,7,36]	Combines deep learning and reinforcement learning for dynamic decision-making (e.g., container loading sequences, AGV scheduling).
Tree Search Algorithms	[2] (tabu-based tree search), [5] (Monte Carlo Tree Search)	Tree-structured search strategies with heuristic rules or probabilistic extensions.
Exact Algorithms and Decomposition Methods	[6] (Branch-and-Bound), [19] (Branch-and-Bound), [23] (Benders Decomposition), [28] (Mixed-Integer Programming)	Suitable for small-scale problems or model decomposition (e.g., accelerated convergence).
Hybrid Heuristic Algorithms	[10] (Dijkstra + Q-Learning), [15] (tabu list), [24] (sequence insertion+ greedy insertion), [33] (Model Predictive Control), [42] (Spatiotemporal Greedy Strategy)	Integrates rules, graph theory, or simulation optimization to address complex constraints (e.g., path conflicts, dynamic scheduling).
other	[9] (sequential decision), [26] (not explicitly mentioned), [32] (Gurobi solver)	

This study aims to address these gaps by proposing an intelligent decision-making model and algorithm for full shoreline loading operations, ensuring balanced yard and quay crane operations while minimizing conflicts and delays. The principal conclusion highlights the model's effectiveness in improving terminal efficiency and reducing operational bottlenecks.

3. Mathematical Model

3.1. Assumptions

The following assumptions are made to simplify the model while maintaining its practical relevance: (1) All equipment operates normally and is fully attended during production operations. (2) The target bays for quay crane operations are known during production. (3) The unloading plan for vessels has been confirmed. (4) The stowage plan

for vessels has been confirmed and will not be revised. (5) The distance between the quay crane's operating position and the target bays in the yard is known. (6) Only loading and unloading operations are considered in the decision-making process. (7) The current loading and unloading instructions for each yard crane are known.

3.2. Symbol Specifications

3.2.1. Dimensions

The following dimensions (Table 2) are defined to describe the key components of the model.

Table 2. Dimensions.

Symbol	Description
Q	Set of quay cranes operating normally during production, $q, q' \in Q$
Y	Set of yard cranes operating normally during production, $y, y' \in Y$
C	Set of all containers requiring operations on the berthed vessel, $c, c' \in C$
V	Set of all vessel bays requiring operations (categorized into deck bays and hold bays), $v, v' \in V$;
B	Set of all yard bays requiring operations, $b, b' \in B$
CH	Set of all yard channels requiring operations, $ch, ch' \in CH$

3.2.2. Model Parameter

The key parameters of the model are listed in Table 3, including variables related to quay cranes, yard cranes, containers, vessel bays, yard bays, and channels.

Table 3. Model parameters.

Symbol	Description
WT_q	Working state of quay crane q (−1: unloading, 0: simultaneous, 1: loading)
W_{qv}	Binary variable : whether quay crane q
CT_c	Set of all containers requiring operations on the berthed vessel, $c, c' \in C$
BL_{cb}	Set of all vessel bays requiring operations (categorized into deck bays and hold bays), $v, v' \in V$;
VL_{cv}	Set of all yard bays requiring operations, $b, b' \in B$
CL_{bch}	Set of all yard channels requiring operations, $ch, ch' \in CH$
$VB_{cc'}$	Binary variable : whether containers c and c' belong to the same vessel bay
HL_{ch}	Number of loading instructions in channel ch
HD_{ch}	Number of unloading instructions in channel ch
HL_q	Number of loading instructions for quay crane q
HD_q	Number of unloading instructions for quay crane q
len_c	Total number of containers requiring operations
len_q	Total number of quay cranes
len_{ch}	Total number of channels
$T_{cc'}$	Binary variable : whether container c is above container c' in the same column
$CH_{bb'}$	Binary variable : whether yard bays b and b' belong to the same channel

3.2.3. Decision Variables

x_{qc} : A binary decision variable indicating whether quay crane q selects container c , A value of 0 means no, and 1 means yes.

3.3. Objective Function

3.3.1. Yard Channel Instruction Balancing

The yard is a core resource in container terminals, and its efficiency directly impacts the overall productivity of the terminal. As a central hub in container logistics, all inbound and outbound containers are temporarily stored in the yard. Uneven distribution of yard operation instructions may lead to congestion in some areas and idle resources in others, hindering container flow. By balancing yard operation instructions, container flow efficiency can be optimized, congestion reduced, and overall operational efficiency

improved. The number of loading instructions can be expressed as in Equation (1), and the number of unloading instructions can be expressed as in Equation (2).

$$IL_{ch} = HL_{ch} + \sum_c \sum_b \sum_v \sum_q CT_c \cdot BL_{cb} \cdot VL_{cv} \cdot W_{qv} \cdot x_{qc} \quad (1)$$

$$ID_{ch} = HD_{ch} + \sum_c \sum_b \sum_v \sum_q (1 - CT_c) \cdot BL_{cb} \cdot VL_{cv} \cdot W_{qv} \cdot x_{qc} \quad (2)$$

For all channels in the yard, the objective is to minimize the difference in instruction quantities between any two channels, with weights assigned to loading and unloading instructions. Let α_1 and α_2 represent the weights for loading and unloading instructions, respectively, as shown in Equation (3):

$$\min F_1 = \sum_{ch} \sum_{ch'} \left(\frac{\alpha_1 |IL_{ch} - IL_{ch'}| + \alpha_2 |ID_{ch} - ID_{ch'}|}{2} \right) \forall ch, ch' \in CH, ch \neq ch' \quad (3)$$

3.3.2. Quay Crane Instruction Balancing

Quay crane workload balancing is critical for terminal efficiency. Imbalanced workloads can lead to overburdened or idle cranes, disrupting operations and increasing dwell time. Balancing involves ensuring equitable distribution of loading and unloading tasks across all cranes. The number of loading instructions can be expressed as in Equation (4), and the number of unloading instructions can be expressed as in Equation (5).

$$IL_q = HL_q + \sum_c \sum_v \sum_q CT_c \cdot VL_{cv} \cdot W_{qv} \cdot (1 - |WT_q|) \cdot x_{qc} \quad (4)$$

$$ID_q = HD_q + \sum_c \sum_v \sum_q (1 - CT_c) \cdot VL_{cv} \cdot W_{qv} \cdot (1 - |WT_q|) \cdot x_{qc} \quad (5)$$

The balance of quay crane instructions across the entire shoreline within a unit time can be expressed as:

$$\min F_2 = \sum_q \sum_{q'} \left(\frac{\beta_1 |L_q - L_{q'}| + \beta_2 |U_q - U_{q'}|}{2} \right) \forall q, q' \in Q, q \neq q' \quad (6)$$

3.3.3. Model Objective Function

The intelligent decision-making problem for stack-yard positioning in loading operations involves balancing instructions for quay cranes and yard cranes, aiming to determine an optimal container pickup point. This is a multi-objective combinatorial optimization problem. To address this, the two sub-objectives are first normalized to ensure they are on the same scale. Then, a weighted average method is applied to integrate these normalized sub-objectives into a single objective function, providing a unified evaluation criterion for decision-making.

1. Normalization of Sub-Objective Functions

The sub-objective F_1 representing channel instruction balancing is normalized using Equation (7):

$$F'_1 = \frac{F_1}{len_c \cdot len_{ch}} \quad (7)$$

The sub-objective F_2 representing quay crane instruction balancing, is normalized using Equation (8):

$$F'_2 = \frac{F_2}{len_c \cdot len_q} \quad (8)$$

2. Total Objective Function

After normalization, the sub-objective functions F_1 and F_2 are scaled to the range [0,1]. Let λ_1 and λ_2 denote the weights of the sub-objectives, which can be determined using expert scoring methods. The total objective function is expressed as:

$$\min F = \lambda_1 \cdot F'_1 + \lambda_2 \cdot F'_2 \quad (9)$$

3.4. Constraint Condition

Loading operations must strictly follow the stowage plan to ensure the final loading result aligns with the plan. During operations, containers cannot be picked up unsupported, as expressed in Equation (10).

$$\sum_q \sum_b \sum_{c'} T_{cc'} \cdot BL_{cb} \cdot DH_b \cdot x_{qc} \leq 0, \forall c \in C \quad (10)$$

Each container can only be stored in one yard bay, as shown in Equation (11).

$$\sum_b BL_{cb} = 1, \forall c \in C \quad (11)$$

Each container can only be stowed in one vessel bay, as shown in Equation (12).

$$\sum_v VL_{cv} = 1, \forall c \in C \quad (12)$$

Each container can only be handled by one quay crane, as shown in Equation (13).

$$\sum_q x_{qc} = 1, \forall c \in C \quad (13)$$

Each yard bay can only belong to one channel, as shown in Equation (14).

$$\sum_{ch} CL_{bch} = 1, \forall b \in B \quad (14)$$

The binary variables defined in the model must satisfy the following constraint:

$$x_{qc} \in \{0, 1\} \quad (15)$$

4. GOGA-GRASP Algorithm

The intelligent decision-making for berth allocation in full-quay container ship stowage is an NP-Hard problem. As the problem size increases, the solution time for conventional algorithms grows exponentially. Given the inherent complexity of this problem, traditional exact solution methods face challenges when addressing large-scale issues. Therefore, to efficiently solve this problem, this paper adopts the Constrained Optimization Genetic Algorithm with Greedy Randomized Adaptive Search Procedures (COGA-GRASP) algorithm. COGA-GRASP is a hybrid algorithm that integrates Genetic Algorithms (GA) and Greedy Randomized Adaptive Search Procedures (GRASP). It is primarily designed to address constrained optimization problems by synergistically combining the global search capabilities of GA with the local search efficiency of GRASP. This dual-mechanism approach ensures that the algorithm not only maintains the global quality of solutions

but also rapidly identifies high-quality solutions within local search spaces when tackling optimization problems with constraints.

4.1. Gene Coding

The objective of the intelligent decision-making problem for berth allocation in full-quay container ship stowage is to calculate the optimal container dispatch points. Here, x_{qc} represents whether quay crane q is assigned to work on container c .

This paper defines a chromosome as a combination of three gene points: the first gene point represents the quay crane ID, the second gene point represents the position in the yard, and the third gene point represents the number of containers to be handled. This chromosome does not represent the execution order of quay cranes, but rather whether the crane will perform the operation. The specific execution sequence will be determined in the later stages of the problem-solving process.

As shown in Figure 3, the green part of the content forms a gene that indicates quay crane Q01 handles three containers at yard location 1A46. Therefore, this chromosome can represent quay crane Q01 handling three containers at 1A46, quay crane Q01 handling four containers at 2B23, quay crane Q02 handling five containers at 2D34, and quay crane Q03 handling one container at 3C28.

Q01	1A46	3	Q01	2B23	4	...	Q02	2D34	5	Q03	3C28	1
-----	------	---	-----	------	---	-----	-----	------	---	-----	------	---

Figure 3. Gene coding.

4.2. Initial Solution Generation

The initial solution is one of the key components of a genetic algorithm. The quality and diversity of the initial solution directly impact the algorithm's output quality and computational efficiency. Currently, there are various strategies for generating the initial solution, among which the greedy-random strategy combines both greediness and randomness. This approach aims to quickly find a good starting solution while satisfying the problem constraints. In this paper, the greedy-random strategy is employed to construct a set of initial solutions, as shown in Table 4.

Table 4. InitializePopulation.

```

S ← ∅
...Function GreedyRandomInitializePopulation()
    s0 ← Empirical design
    S ← S + s0
    While S.size ≤ Initial population size do
        n ← GreedyRandomSelection(seed)
        If n Satisfy constraint
            S ← S + n
    End While
Return S
Function GreedyRandomSelection(seed)
    Based on seed, generate a list of potential solutions in a greedy manner
    Randomly select a solution from the list
Return selected solution

```

4.3. Fitness Value

When using genetic algorithms to solve constrained optimization problems, the design of the fitness function is crucial. The COGA algorithm used in this paper must consider both the objective function value and the constraint conditions when calculating the fitness function. To ensure that the solution satisfies all constraints, a penalty function is typically

introduced to penalize solutions that violate the constraints. In this paper, the fitness value is designed using the penalty function method, as shown in Equation (16):

$$R = F + \theta \cdot \text{constraint} \quad (16)$$

where θ represents the penalty coefficient matrix and *constraint* represents the matrix of constraint violations.

4.4. Select, Crossover and Mutation

This paper adopts Tournament Selection (TS) as the selection algorithm. TS randomly samples m individuals from the population and selects the optimal solution based on fitness values. The best individual is then passed on to the next generation.

This paper adopts the Partially Matched Crossover (PMX, as shown in Table 5) strategy. PMX ensures that each gene in a chromosome appears only once, preventing duplicate genes, thus guaranteeing that the chromosome satisfies the constraints and is feasible. As shown in Figure 4.

Table 5. PMX.

Function PMX ()	
$P_1, P_2 \leftarrow$	random select 2 individual
$\text{cross}_{start}, \text{cross}_{end} \leftarrow$	random select 2 index
$P'_1, P'_2 \leftarrow$	change from cross_{start} to cross_{end}
$\text{Child}_1, \text{Child}_2 \leftarrow$	remove duplication
Return $\text{Child}_1, \text{Child}_2$	

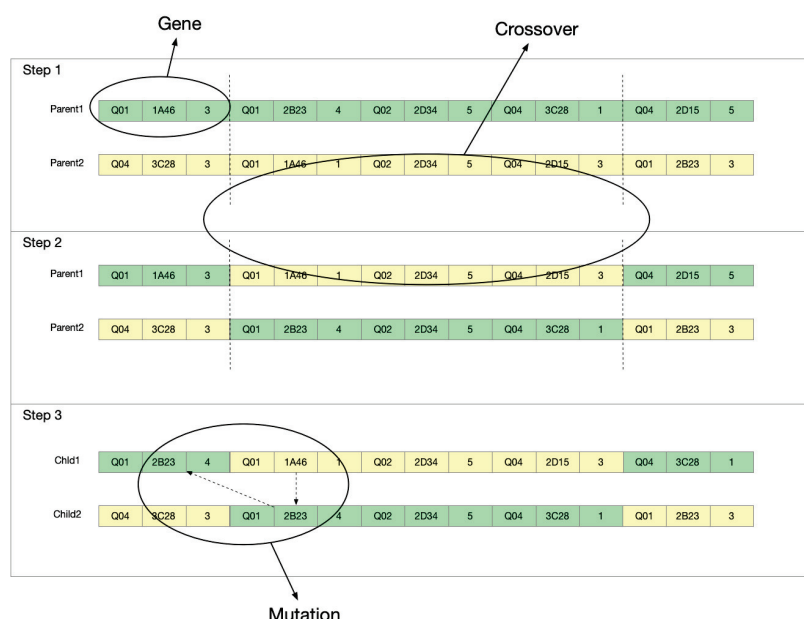


Figure 4. PMX.

First, a pair of parent chromosomes is randomly selected, and crossover points are chosen. When selecting crossover points, it is important to ensure that gene points are intact (e.g., the gene {Q01, 2B23, 4} cannot be split). The crossover points should be the same for both parent chromosomes. Next, the crossover genes are exchanged, and finally, the feasibility of the child chromosomes is verified to ensure there are no duplicates. Duplicates are determined based on whether the quay crane and yard location are the same, while the number of containers is not used as a criterion for duplication.

In this paper, Self-adaptive Mutation (SAM) is chosen as the mutation strategy. SAM adjusts the mutation parameters based on the algorithm's evolutionary state or the population characteristics, thereby enhancing the algorithm's performance. During the mutation process, ω gradually decreases with the number of iterations. Initially, the mutation rate is high to allow for extensive exploration. Then, based on the evolutionary progress, the mutation rate is gradually reduced to better exploit the information in the search space. For a given individual P , the mutation probability is adjusted according to the fitness value r . The larger r is, the lower the mutation rate, increasing the likelihood of retaining that individual. Conversely, the smaller r is, the higher the mutation rate, increasing the probability of searching for a better solution.

5. Experiment

This chapter conducts a case study on the intelligent decision-making problem of stack-yard positioning for full shoreline loading operations to validate the correctness of the model, the effectiveness and applicability of the algorithm, and the practicality of the final results. Through result analysis and algorithm comparison, it is demonstrated that the proposed COGA algorithm with tournament selection and adaptive mutation (COGA-GRASP) outperforms the traditional GA algorithm with roulette wheel selection and single-point mutation in terms of computational efficiency and robustness.

5.1. Experiment Description

The source of the basic data for the test case is with Shanghai Mingdong Container Terminal Company Limited (Shanghai, China), as shown in Figure 5. It consists of three main parts: containers to be loaded and unloaded, existing quay crane loading and unloading instructions and existing channel loading and unloading instructions. A total of 500 General Purpose Containers (GP) are distributed along three channels, each channel contains existing loading and unloading instructions. The types of quay crane operations and the number of quay crane operation instructions are also taken from the actual production data of MingDong Terminal to ensure the validity of the test case. This experiment is coded in Python 3.10 language and runs on a Mac computer with an M3 pro chip and 18 GB of RAM.

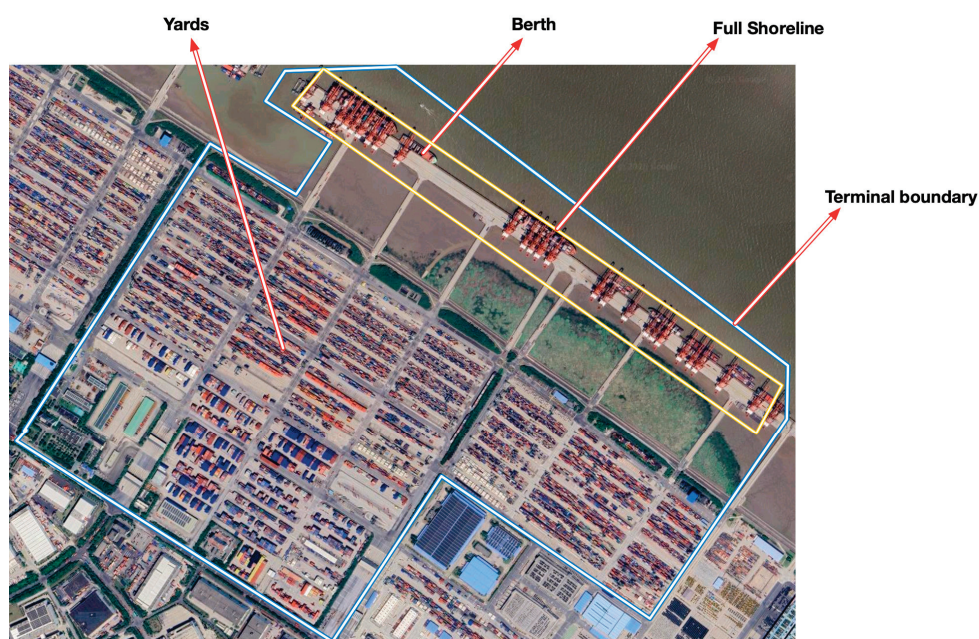


Figure 5. Boundary of Shanghai Mingdong Container Terminal Co.

5.2. Experiment Analysis

In this paper, multiple computations were performed with different population sizes, iteration counts, and algorithmic strategies. The following are the computational results:

- Using Tournament Selection and Self-adaptive Mutation algorithms, with 100 iterations and population sizes of 50 and 100, the convergence curves are shown in Figure 6. In this example, the fitness function values are calculated with the parameters for loading and unloading instructions in Equations (3) and (6) set to $\alpha_1 = 0.8, \alpha_2 = 0.2, \beta_1 = 0.8, \beta_2 = 0.2$. For the total objective function in Equation (9), the parameter values are set to $\lambda_1 = 0.5, \lambda_2 = 0.5$.

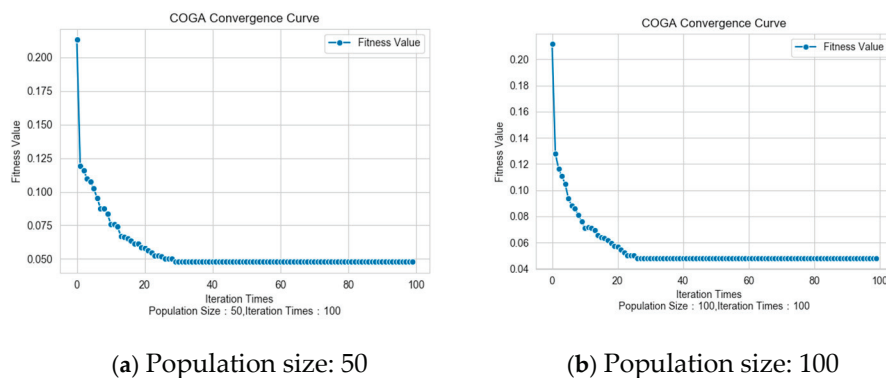


Figure 6. Convergence Comparison for Different Population Sizes.

From the Figure 6, it can be observed that in this problem, the difference in fitness values between a population size of 50 and 100 is relatively small. The algorithm achieves good convergence around 40 iterations in both cases. Table 6 presents the final container handling results and the corresponding fitness values. Specifically, quay crane Q01:21:19 indicates that quay crane Q01 has 21 loading instructions and 19 unloading instructions, while yard 3B:18:45 indicates that yard 3B has 18 loading instructions and 45 unloading instructions. The fitness values are 0.0958 and 0.0944, both of which yield good solutions.

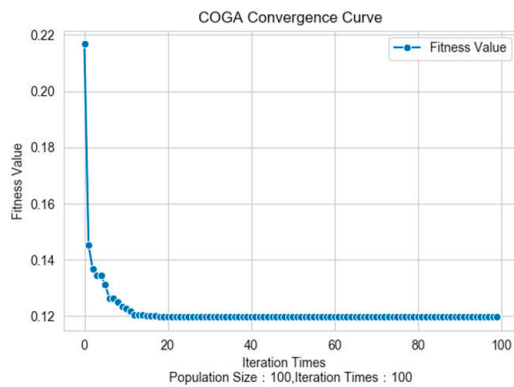
Table 6. Result.

Population Size:50			Population Size: 100		
Quay INS	Yard INS	Fitness Value	Quay INS	Yard INS	Fitness Value
Q01:21: 19	3B:18:45	0.0958	Q01:20:19	3B:22:45	0.0944
Q02:21: 14	3A:36:20		Q02:20: 14	3A:36: 20	
Q03:21: 15	2B:38: 30		Q03:20: 15	2B:38: 30	
Q04:21: 15	1A:34:19		Q04:20: 15	1A:22: 19	
Q05: 21: 0	2A:42:34		Q05: 20: 0	2A:35: 34	
Q06:21: 12	1C:26: 46		Q06:20: 12	1C:26: 46	
Q07:21: 15	3C:38: 32		Q07:20: 15	3C:39: 32	
Q08:21: 15	2C:29: 24		Q08:20: 15	1B:33: 39	
Q09: 21: 4	1B:27: 39		Q09: 20: 4	2C:22: 24	
Q10: 21: 16			Q10: 20: 16		
Q11: 21: 4			Q11: 20: 4		
Q12: 21: 10			Q12: 20: 10		
Q13: 21: 6			Q13: 20: 6		
Q14: 21: 5			Q14: 20: 5		
Q15: 21: 4			Q15: 20: 4		

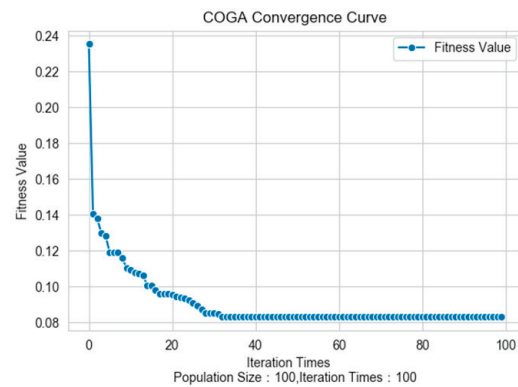
- Due to the varying emphasis placed on quay crane balance and yard balance by different container terminals, the importance of loading and unloading instructions differs. Therefore, the parameters $\alpha_1, \alpha_2, \lambda_1, \lambda_2$ are set differently. To analyze the

impact of different parameters on the algorithm's convergence, several experiments were conducted, and some of the results are summarized below.

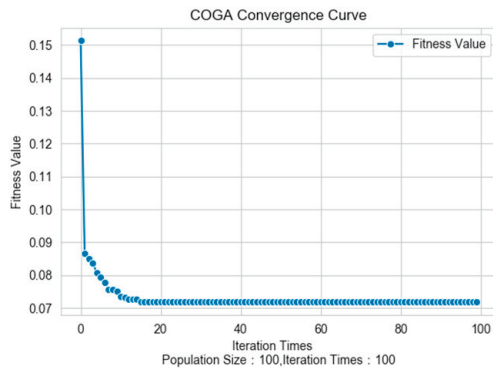
- Experiment 1: The parameters are set as $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $\beta_1 = 0.5$, $\beta_2 = 0.5$, $\lambda_1 = 0.5$, $\lambda_2 = 0.5$. The fitness value is 0.11975, as shown in Figure 7a.
- Experiment 2: The parameters are set as $\alpha_1 = 0.3$, $\alpha_2 = 0.7$, $\beta_1 = 0.3$, $\beta_2 = 0.7$, $\lambda_1 = 0.5$, $\lambda_2 = 0.5$. The fitness value is 0.08304 as shown in Figure 7b.
- Experiment 3: The parameters are set as $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $\beta_1 = 0.5$, $\beta_2 = 0.5$, $\lambda_1 = 0.3$, $\lambda_2 = 0.7$. The fitness value is 0.07185, as shown in Figure 7c.
- Experiment 4: The parameters are set as $\alpha_1 = 0.5$, $\alpha_2 = 0.5$, $\beta_1 = 0.5$, $\beta_2 = 0.5$, $\lambda_1 = 0.8$, $\lambda_2 = 0.2$. The fitness value is 0.07185, as shown in Figure 7d.



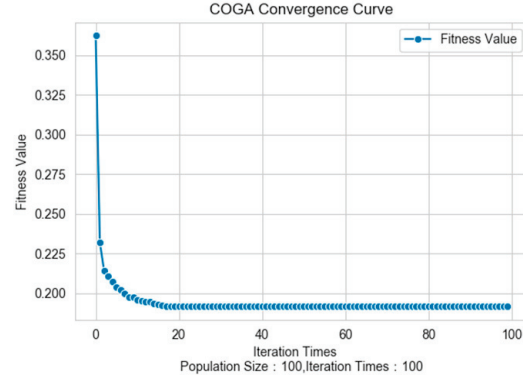
(a) Experiment 1



(b) Experiment 2



(c) Experiment 3



(d) Experiment 4

Figure 7. Experimental comparison.

From the experimental convergence curves, it can be observed that for different parameter settings, the algorithm demonstrates high convergence, quickly reaching a good solution. Its solving capability is stable, with both feasibility and practicality.

3. This paper presents a statistical analysis of different strategies and population sizes, as shown in Figure 8. It can be observed that the COGA algorithm, utilizing Tournament Selection (TS), Partially Matched Crossover (PMX), and Self-Adaptive Mutation (SAM), achieves faster solution speeds and higher solution quality. A comparison between the COGA and GA algorithms indicates that COGA offers higher solving efficiency and greater stability, which is of significant importance for practical production.

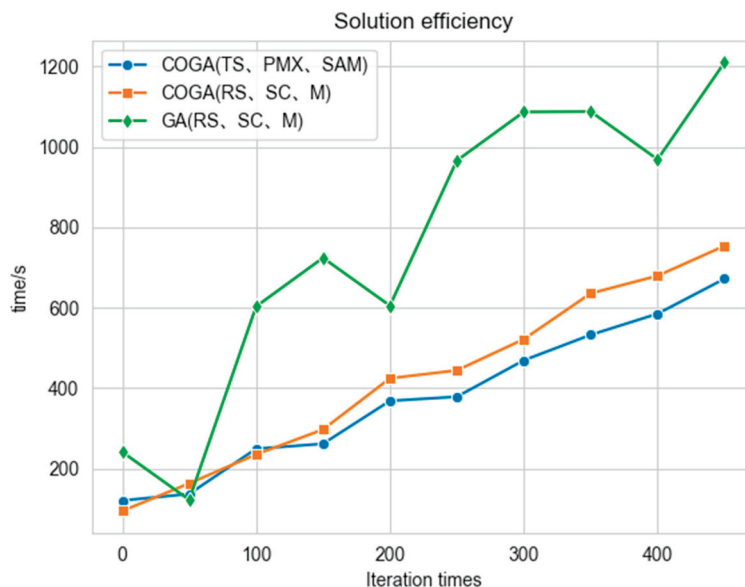


Figure 8. Comparison of solving efficiency.

5.3. Applicationt Analysis

The Mingdong Terminal primarily accommodates scheduled liner services with fixed shipping routes. To conduct a comprehensive benchmark analysis of full-quay vessel loading efficiency and operational capacity, comparative evaluations were implemented through two distinct approaches: the terminal’s proprietary digital twin platform versus conventional loading operations.

As shown in Figure 9, the quay crane productivity metrics reveal three experimental configurations: Dataset 1 (Baseline): Conventional loading operations from 7 to 13 December 2024. Dataset 2 (Field Implementation): Mathematical model-driven operations from 14 to 20 December 2024. Dataset 3 (Digital Twin Simulation): Model-enhanced operations within the digital environment during 7–13 December 2024.

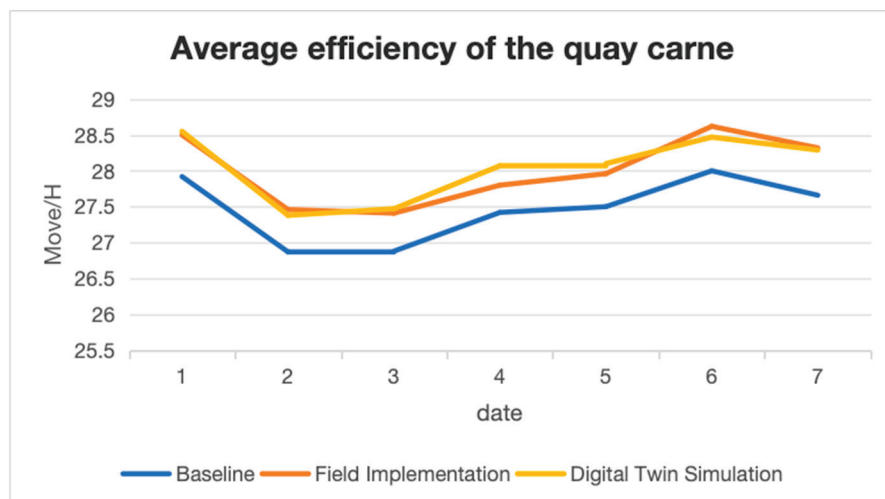


Figure 9. Average efficiency of the quay crane.

The comparative analysis substantiates the operational efficacy of the proposed model across multiple validation scenarios, with quantified results demonstrating a 2.12% productivity increase in quay crane operations through Digital Twin Implementation (Dataset 3) and a 1.98% efficiency improvement in Physical Terminal Deployment (Dataset 2). This dual-environment verification confirms statistically significant enhancements in

both simulated and real-world settings, wherein the digital twin platform's capabilities yielded marginally superior optimization outcomes compared to physical implementations, thereby systematically validating the algorithm's cross-domain adaptability and technical robustness.

6. Conclusions

This study investigates the container loading instruction allocation problem within full shoreline operational areas, establishing an intelligent decision-making model through refined optimization objectives and constraint formulations. A novel COGA-GRASP hybrid algorithmic framework is proposed to address this NP-hard combinatorial optimization challenge.

The experimental configuration utilizes an initial dataset comprising 500 pending containers from Shanghai Port's Mingdong Container Terminal. The methodology incorporates tournament selection mechanisms, partially matched crossover operators, and adaptive mutation strategies to enhance solution quality. Comparative analysis of multiple parameter configurations and fitness evaluations demonstrates the superior solution optimality achieved by COGA-GRASP across operational scenarios.

Benchmark testing against conventional GA implementations reveals COGA-GRASP's enhanced adaptability to varying dataset scales and problem complexities. The algorithm exhibits robust performance characteristics with rapid solution convergence, maintaining computational efficiency under different operational scales. However, given the real-time dynamic nature of terminal operations and strong interdependency with horizontal transportation systems, future research should prioritize tighter integration with vehicular scheduling models to develop comprehensive loading optimization frameworks.

Author Contributions: Conceptualization, X.D.; methodology, X.D.; software, J.W.; validation, B.L.; formal analysis, Q.S.; investigation, D.L.; resources, B.L.; data curation, H.L.; writing—original draft, X.D.; Writing—review and editing, J.Z.; supervision, H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Integrated maritime guidance/navigation high confidence performance prediction and information services, grant number 2021YFB3901502.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wen, B.; Kim, K.H.; Feng, X. Sequencing ship operations considering the trajectory of the quay crane spreader. *Transp. Res. Part B Methodol.* **2024**, *186*, 102987. [CrossRef]
2. Zhou, C.; Yuan, M.; Zhang, J.; Zhang, W. A tree search algorithm for uncertainty-considered consecutive discharging and loading operations between ship and offshore platform. *Eur. J. Oper. Res.* **2024**, *315*, 729–749. [CrossRef]
3. Rahman, M.M.; Jahin, M.A.; Islam, M.S.; Mridha, M.F. Optimizing Container Loading and Unloading through Dual-Cycling and Dock-yard Rehandle Reduction Using a Hybrid Genetic Algorithm. *arXiv* **2024**, arXiv:2406.08534.
4. Yan, Q.; Song, R.; Kim, K.H.; Wang, Y.; Feng, X. Optimizing container relocation operations by using deep reinforcement learning. *Marit. Policy Manag.* **2024**, 1–23. [CrossRef]
5. Zhang, Z.; Tan, K.C.; Qin, W.; Chew, E.P.; Li, Y. A data-driven approach to solving the container relocation problem with uncertainties. *Adv. Eng. Inform.* **2025**, *65*, 103112. [CrossRef]
6. Parreño-Torres, C.; Alvarez-Valdes, R.; Ruiz, R.; Tierney, K. Minimizing crane times in pre-marshalling problems. *Transp. Res. Part E Logist. Transp. Rev.* **2020**, *137*, 101917. [CrossRef]
7. Xia, M.; Li, Y.; Shen, Y.; Zhao, N. Loading sequencing problem in container terminal with deep Q-learning. *J. Coast. Res.* **2020**, *103*, 817–821. [CrossRef]
8. Cai, M.D. Intelligent Decision-Making for Multi-Bay Container Loading in Yards Based on Inverse Reinforcement Learning. Ph.D. Thesis, Wuhan University of Technology, Wuhan, China, 2023. [CrossRef]

9. Xia, M.J. Research on Sequential Decision-Making for Intelligent Ship Control in Container Terminals. Ph.D. Thesis, Shanghai Maritime University, Shanghai, China, 2021. [CrossRef]
10. Yue, L.; Fan, H. Dynamic scheduling and path planning of automated guided vehicles in automatic container terminal. *IEEE/CAA J. Autom. Sin.* **2022**, *9*, 2005–2019. [CrossRef]
11. Fibrianto, H.Y.; Kang, B.; Hong, S. A job sequencing problem of an overhead shuttle crane in a rail-based automated container terminal. *IEEE Access* **2020**, *8*, 156362–156377. [CrossRef]
12. Oladugba, A.O.; Gheith, M.; Eltawil, A. A new solution approach for the twin yard crane scheduling problem in automated container terminals. *Adv. Eng. Inform.* **2023**, *57*, 102015. [CrossRef]
13. Liu, W.; Zhu, X.; Wang, L.; Yan, B.; Zhang, X. Optimization approach for yard crane scheduling problem with uncertain parameters in container terminals. *J. Adv. Transp.* **2021**, *2021*, 5537114. [CrossRef]
14. Yu, M.; Liang, Z.; Teng, Y.; Zhang, Z.; Cong, X. The inbound container space allocation in the automated container terminals. *Expert Syst. Appl.* **2021**, *179*, 115014. [CrossRef]
15. Yang, X.; Hu, H.; Jin, J.; Luo, N. Joint optimization of space allocation and yard crane deployment in container terminal under uncertain demand. *Comput. Ind. Eng.* **2022**, *172*, 108425. [CrossRef]
16. Gulić, M.; Maglić, L.; Krljan, T.; Maglić, L. Solving the container relocation problem by using a metaheuristic genetic algorithm. *Appl. Sci.* **2022**, *12*, 7397. [CrossRef]
17. Yang, X.; Hu, H.; Cheng, C.; Wang, Y. Automated guided vehicle (agv) scheduling in automated container terminals (acts) focusing on battery swapping and speed control. *J. Mar. Sci. Eng.* **2023**, *11*, 1852. [CrossRef]
18. Xu, B.; Jie, D.; Li, J.; Zhou, Y.; Wang, H.; Fan, H. A hybrid dynamic method for conflict-free integrated schedule optimization in U-shaped automated container terminals. *J. Mar. Sci. Eng.* **2022**, *10*, 1187. [CrossRef]
19. Wang, Z.; Zeng, Q. A branch-and-bound approach for AGV dispatching and routing problems in automated container terminals. *Comput. Ind. Eng.* **2022**, *166*, 107968. [CrossRef]
20. Lu, Y. The Three-Stage Integrated Optimization of Automated Container Terminal Scheduling Based on Improved Genetic Algorithm. *Math. Probl. Eng.* **2021**, *2021*, 67921. [CrossRef]
21. Niu, B.; Liu, Q.; Wang, Z.; Tan, L.; Li, L. Multi-objective bacterial colony optimization algorithm for integrated container terminal scheduling problem. *Nat. Comput.* **2021**, *20*, 89–104. [CrossRef]
22. Luo, J.; Wu, Y. Scheduling of container-handling equipment during the loading process at an automated container terminal. *Comput. Ind. Eng.* **2020**, *149*, 106848. [CrossRef]
23. Li, Y.; Sun, Z.; Hong, S. An exact algorithm for multiple-equipment integrated scheduling in an automated container terminal using a double-cycling strategy. *Transp. Res. Part E Logist. Transp. Rev.* **2024**, *186*, 103565. [CrossRef]
24. Yin, Y.Q.; Zhong, M.; Wen, X.; Ge, Y.E. Scheduling quay cranes and shuttle vehicles simultaneously with limited apron buffer capacity. *Comput. Oper. Res.* **2023**, *151*, 106096. [CrossRef]
25. Yue, L.; Fan, H.; Ma, M. Optimizing configuration and scheduling of double 40 ft dual-trolley quay cranes and AGVs for improving container terminal services. *J. Clean. Prod.* **2021**, *292*, 126019. [CrossRef]
26. Yu, H.; Zhang, M.; He, J.; Tan, C. Choice of loading clusters in container terminals. *Adv. Eng. Inform.* **2020**, *46*, 101190. [CrossRef]
27. Zheng, Y.; Xu, M.; Wang, Z.; Xiao, Y. A Genetic Algorithm for Integrated Scheduling of Container Handling Systems at Container Terminals from a Low-Carbon Operations Perspective. *Sustainability* **2023**, *15*, 6035. [CrossRef]
28. He, J.; Wang, Y.; Tan, C.; Yu, H. Modeling berth allocation and quay crane assignment considering QC driver cost and operating efficiency. *Adv. Eng. Inform.* **2021**, *47*, 101252. [CrossRef]
29. Xu, B.; Jie, D.; Li, J.; Yang, Y.; Wen, F.; Song, H. Integrated scheduling optimization of U-shaped automated container terminal under loading and unloading mode. *Comput. Ind. Eng.* **2021**, *162*, 107695. [CrossRef]
30. Jonker, T.; Duinkerken, M.B.; Yorke-Smith, N.; de Waal, A.; Negenborn, R.R. Coordinated optimization of equipment operations in a container terminal. *Flex. Serv. Manuf. J.* **2021**, *33*, 281–311. [CrossRef]
31. Lu, Y. The optimization of automated container terminal scheduling based on proportional fair priority. *Math. Probl. Eng.* **2022**, *2022*, 7889048. [CrossRef]
32. Gao, Y.; Chen, C.H.; Chang, D. A machine learning-based approach for multi-agv dispatching at automated container terminals. *J. Mar. Sci. Eng.* **2023**, *11*, 1407. [CrossRef]
33. Cahyono, R.T.; Kenaka, S.P.; Jayawardhana, B. Simultaneous allocation and scheduling of quay cranes, yard cranes, and trucks in dynamical integrated container terminal operations. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 8564–8578. [CrossRef]
34. Shouwen, J.; Di, L.; Zhengrong, C.; Dong, G. Integrated scheduling in automated container terminals considering AGV conflict-free routing. *Transp. Lett.* **2021**, *13*, 501–513. [CrossRef]
35. Hsu, H.P.; Wang, C.N.; Fu, H.P.; Dang, T.T. Joint scheduling of yard crane, yard truck, and quay crane for container terminal considering vessel stowage plan: An integrated simulation-based optimization approach. *Mathematics* **2021**, *9*, 2236. [CrossRef]
36. Gong, L.; Huang, Z.; Xiang, X.; Liu, X. Real-time AGV scheduling optimisation method with deep reinforcement learning for energy-efficiency in the container terminal yard. *Int. J. Prod. Res.* **2024**, *62*, 7722–7742. [CrossRef]

37. Xin, J.; Meng, C.; D'Ariano, A.; Wang, D.; Negenborn, R.R. Mixed-integer nonlinear programming for energy-efficient container handling: Formulation and customized genetic algorithm. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 10542–10555. [CrossRef]
38. Wang, T.; Gao, G.; Wang, K.; Shi, J. Integrated operation models with quay crane maintenance in a container terminal. *Ocean Coast. Manag.* **2024**, *251*, 107101. [CrossRef]
39. Shen, Z.; Liang, C.; Gen, M. Scheduling of AGV with group operation area in automated terminal by hybrid genetic algorithm. In Proceedings of the Fifteenth International Conference on Management Science and Engineering Management, Toledo, Spain, 2–3 August 2021; Springer International Publishing: Berlin/Heidelberg, Germany, 2021; Volume 115, pp. 427–442.
40. Fan, H.; Peng, W.; Ma, M.; Yue, L. Storage space allocation and twin automated stacking cranes scheduling in automated container terminals. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 14336–14348. [CrossRef]
41. Jian, W.; Zhu, J.; Zeng, Q. An Optimization Model of Integrated AGVs Scheduling and Container Storage Problems for Automated Container Terminal Considering Uncertainty. *Symmetry* **2021**, *13*, 1904. [CrossRef]
42. Liu, W.; Zhu, X.; Wang, L.; Wang, S. Multiple equipment scheduling and AGV trajectory generation in U-shaped sea-rail intermodal automated container terminal. *Measurement* **2023**, *206*, 112262. [CrossRef]
43. Hsu, H.P.; Wang, C.N.; Nguyen, T.T.T.; Dang, T.T.; Pan, Y.J. Hybridizing WOA with PSO for coordinating material handling equipment in an automated container terminal considering energy consumption. *Adv. Eng. Inform.* **2024**, *60*, 102410. [CrossRef]
44. Fontes, D.B.M.M.; Homayouni, S.M. A bi-objective multi-population biased random key genetic algorithm for joint scheduling quay cranes and speed adjustable vehicles in container terminals. *Flex. Serv. Manuf. J.* **2023**, *35*, 241–268.
45. Hsu, H.P.; Tai, H.H.; Wang, C.N.; Chou, C.C. Scheduling of collaborative operations of yard cranes and yard trucks for export containers using hybrid approaches. *Adv. Eng. Inform.* **2021**, *48*, 101292. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Truck Lifting Accident Detection Method Based on Improved PointNet++ for Container Terminals

Yang Shen ^{1,2}, Xintai Man ³, Jiaqi Wang ³, Yujie Zhang ⁴ and Chao Mi ^{2,3,*}¹ Higher Technology College, Shanghai Maritime University, Shanghai 201306, China; yangshen@shmtu.edu.cn² Shanghai SMUVision Smart Technology Ltd., Shanghai 201306, China³ Logistics Engineering College, Shanghai Maritime University, Shanghai 201306, China; 202430210134@stu.shmtu.edu.cn (X.M.); 202230210297@stu.shmtu.edu.cn (J.W.)⁴ School of Technology and Architecture, ISCTE-Instituto Universitário de Lisboa, 1649-026 Lisbon, Portugal; yzgeu@iscte-iul.pt

* Correspondence: chaomi@shmtu.edu.cn

Abstract: In container terminal operations, truck lifting accidents pose a serious threat to the safety and efficiency of automated equipment. Traditional detection methods using visual cameras and single-line Light Detection and Ranging (LiDAR) are insufficient for capturing three-dimensional spatial features, leading to reduced detection accuracy. Moreover, the boundary features of key accident objects, such as containers, truck chassis, and wheels, are often blurred, resulting in frequent false and missed detections. To tackle these challenges, this paper proposes an accident detection method based on multi-line LiDAR and an improved PointNet++ model. This method uses multi-line LiDAR to collect point cloud data from operational lanes in real time and enhances the PointNet++ model by integrating a multi-layer perceptron (MLP) and a mixed attention mechanism (MAM), optimizing the model's ability to extract local and global features. This results in high-precision semantic segmentation and accident detection of critical structural point clouds, such as containers, truck chassis, and wheels. Experiments confirm that the proposed method achieves superior performance compared to the current mainstream algorithms regarding point cloud segmentation accuracy and stability. In engineering tests across various real-world conditions, the model exhibits strong generalization capability.

Keywords: multi-line LiDAR; PointNet++; mixed attention mechanism; point cloud semantic segmentation; truck lifting accidents

1. Introduction

With the rapid development of container terminals and the widespread adoption of automation technology, the efficiency of loading and unloading operations at terminals has significantly improved. The container truck, as the key link between the container and the port machinery, directly affects the smooth operation of the entire terminal. During unloading operations at the terminal, when the lock on the truck's chassis does not separate from the corner fittings of the container, the truck may be lifted together with the container, a situation commonly referred to as a truck lifting accident. These truck lifting accidents not only result in damage to containers and equipment failure but also pose safety threats to workers. Therefore, the prevention and management of such accidents have become important tasks in the management of modern ports. Figure 1 shows nine different types of truck lifting accidents during the operation of 40-foot containers.

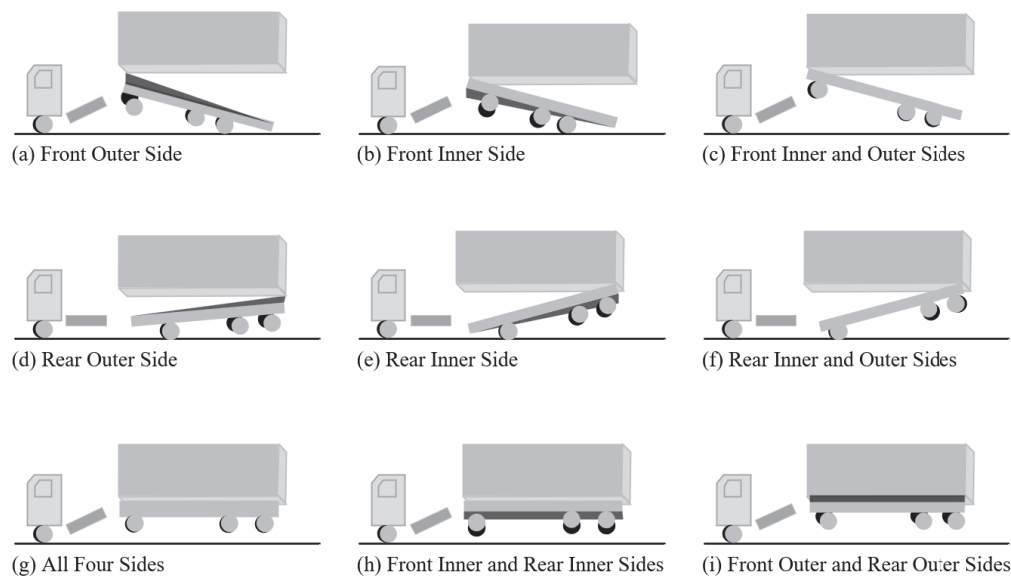


Figure 1. Nine different types of truck lifting accidents.

Currently, the traditional truck lifting accident detection methods mainly rely on visual cameras and single-line LiDAR, both of which exhibit specific limitations that this study aims to address. Vision-based detection systems [1,2] process image data, which lack three-dimensional spatial information, making it difficult to accurately capture the geometric characteristics of objects. Additionally, vision-based systems are highly sensitive to environmental lighting conditions, further reducing their reliability in real-world operational environments. On the other hand, single-line LiDAR-based detection systems [3] acquire a limited number of points due to their single scanning line. These systems can be categorized into horizontal and vertical scanning types and are prone to installation effects. The impact of installation can lead to incomplete or inconsistent data capture, resulting in high rates of missed detections. Therefore, while these two methods are applicable in specific situations, they exhibit shortcomings regarding precision and stability.

Compared to the above two methods, multi-line LiDAR provides richer three-dimensional raw feature data through multiple scanning channels, offering higher spatial resolutions and depth information. This provides a new approach for addressing truck lifting accident detection.

This paper proposes a detection method based on multi-line LiDAR and an improved PointNet++ algorithm to address the constraints of the current detection techniques and methodologies. This method uses multi-line LiDAR to collect real-time point cloud data from operational lanes and enhances the PointNet++ model by integrating a multi-layer perceptron (MLP) and a mixed attention mechanism (MAM), enhancing the model's capacity to capture both local and global features, enabling high-precision semantic segmentation and accident detection of critical structures such as containers, truck chassis, and wheels.

The contributions of this paper are as follows:

- (1) A truck lifting accident detection system based on multi-line LiDAR is designed to overcome the limitations of traditional vision cameras and single-line LiDAR in terms of three-dimensional information acquisition and environmental lighting conditions.
- (2) An improved PointNet++ network model that combines an MLP and MAM is proposed, addressing the limitations of the traditional PointNet++ in extracting features from blurry boundary regions and capturing global features.
- (3) Building upon the improved PointNet++ network, this paper introduces a new approach for the accurate and efficient detection of truck lifting accidents by computing the shortest distance between the point clouds of the container and the truck chassis.

2. Related Work

With the rapid development of 3D point cloud machine learning, LiDAR-based detection and measurement technologies have been widely applied in industrial fields, especially in port logistics. Examples include container lock pin detection [4], unmanned truck navigation and obstacle avoidance [5], container ship guide rail segmentation [6], and bulk carrier cargo hold and material detection [7]. These applications show that LiDAR can stably provide reliable and accurate detection capabilities in complex environments with unstable lighting conditions.

In order to meet the demand for high-precision three-dimensional information for containers and truck chassis in truck lifting accident detection, point cloud segmentation algorithms have become a key research direction. Traditional unsupervised clustering algorithms, such as K-means [8], are limited in handling complex and high-dimensional point cloud data. Support Vector Machines (SVMs) [9], as a supervised learning method, perform well in classification tasks but still face challenges in point cloud segmentation tasks. With the advent of deep neural networks (DNNs) [10], PointNet [11], PointNet++ [12], and PointCNN [13], the processing capabilities regarding point cloud data have been significantly improved. In particular, PointNet++ can directly handle raw point cloud data through hierarchical feature learning and local feature extraction, enabling both global classification and local segmentation tasks. This has greatly advanced the research on point cloud object recognition and semantic segmentation. Recently, PointNet++ has been widely applied across various fields. Zhang and Wang [14] proposed a convergence segmentation method based on PointNet++ and edge information, effectively improving the accuracy of cervical vertebra segmentation. Fan et al. [15] addressed the difficulty of tree species classification due to uneven point cloud quality and quantity from airborne LiDAR and validated the feasibility of using PointNet++ in addressing these issues. Hu et al. [16], facing the challenge of recognizing complex regional terrain features, proposed a multi-information fusion method based on PointNet++, effectively improving the accuracy of Digital Elevation Model (DEM) construction. Lin and Knudby [17] developed and tested an automated method based on PointNet++ for the global extraction of bathymetric photons from ICESat-2 ATLAS data. Patel et al. [18] used multiple deep learning models, including PointNet++, for plant organ segmentation and extracted phenotypic features based on segmentation results, verifying the potential application of LiDAR point cloud data in plant phenotyping. Therefore, given the numerous advantages of PointNet++, including its ability to effectively process unstructured point cloud data, capture both local and global features, adapt to point clouds of varying densities, and its extensive successful applications in engineering fields, this paper selects PointNet++ as the base architecture.

Despite the good performance of PointNet++ in point cloud segmentation tasks, it still has some limitations, such as sensitivity to point cloud density and inadequate feature extraction ability for blurry boundary regions. Researchers have proposed several improvement methods to address these issues. Jing et al. [19] introduced the Squeeze-and-Excitation PointNet++ (SE-PointNet++) architecture, incorporating an SE module to enhance the selectivity of feature channels. Guo et al. [20] proposed an improved PointNet++ model that combines deep learning with clustering algorithms, optimizing the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm and incorporating an improved attention module to enhance semantic segmentation accuracy. Chen et al. [21] proposed a point cloud segmentation and depth measurement approach based on PointNet++ and the Depth Measurement for Cigarette Perforation (DMCP) algorithm, using clustering and an attention mechanism to solve issues such as the identification of small holes and depth calculation. Cheng and Gu [22], to address the limited ability of PointNet++, proposed the Shuffle Attention and Offset-Attention Point-

Net++ (SO-PointNet++) architecture, combining Shuffle Attention and Offset-Attention to enhance feature extraction ability and solve the gradient vanishing problem. Yu et al. [23] introduced the Double Self-Attention Mechanism PointNet++ (DSPNet++) network using a double self-attention mechanism for airborne LiDAR point cloud segmentation, leveraging local feature aggregation modules and self-attention convolution networks to adaptively extract semantic features from the point cloud. However, their approach primarily focused on improving the segmentation accuracy of specific types of point cloud data, such as airborne LiDAR or small-scale objects, and did not fully address the challenges of handling complex high-dimensional point cloud data with blurry boundary features, which are common in industrial environments like container terminals. Additionally, while these improvements have enhanced the performance of PointNet++ in certain scenarios, they often lack the ability to effectively balance the extraction of local and global features, which is crucial for accurate semantic segmentation tasks in dynamic and complex operational environments. This limitation becomes particularly evident in applications such as truck lifting accident detection.

In summary, researchers have partially solved the limitations of PointNet++ in handling different types of point cloud data by introducing attention mechanisms, optimizing loss functions, and combining clustering algorithms. However, there is still room for improvement in feature extraction capabilities for blurry boundary regions and the fusion of local and global features. To fill the gaps in the current research, this paper proposes a truck lifting accident detection method based on multi-line LiDAR and an improved PointNet++, which aims to further enhance the model's feature extraction ability and segmentation accuracy for key structures by integrating an MLP and a mixed attention mechanism, providing reliable technical support for the safe and efficient operation of container terminals.

3. Truck Lifting Accident Detection Method

3.1. Hardware System

This paper designs a truck lifting accident detection system for container terminals based on multi-line LiDAR. The system's hardware platform consists of a single multi-line LiDAR and an edge computing module, which can be mounted on various container handling machinery at container terminals, such as gantry cranes and rail-mounted gantry cranes. As shown in Figure 2, the LiDAR is mounted on the rigid leg of a rail-mounted gantry crane, facing the truck operational lane for scanning. Compared to the flexible leg, the rigid leg structure is more stable, making it suitable for installing precision equipment such as LiDAR, avoiding sensor malfunctions caused by vibrations. The edge computing module is installed in the electrical room above the rigid leg and is responsible for receiving and processing point cloud data collected by the LiDAR during container lifting operations. Additionally, it listens to the status information of the gantry crane's Programmable Logic Controller (PLC), such as the spreader lock and unlock status and the height encoder, to accurately determine when the system triggers accident detection and alarms.

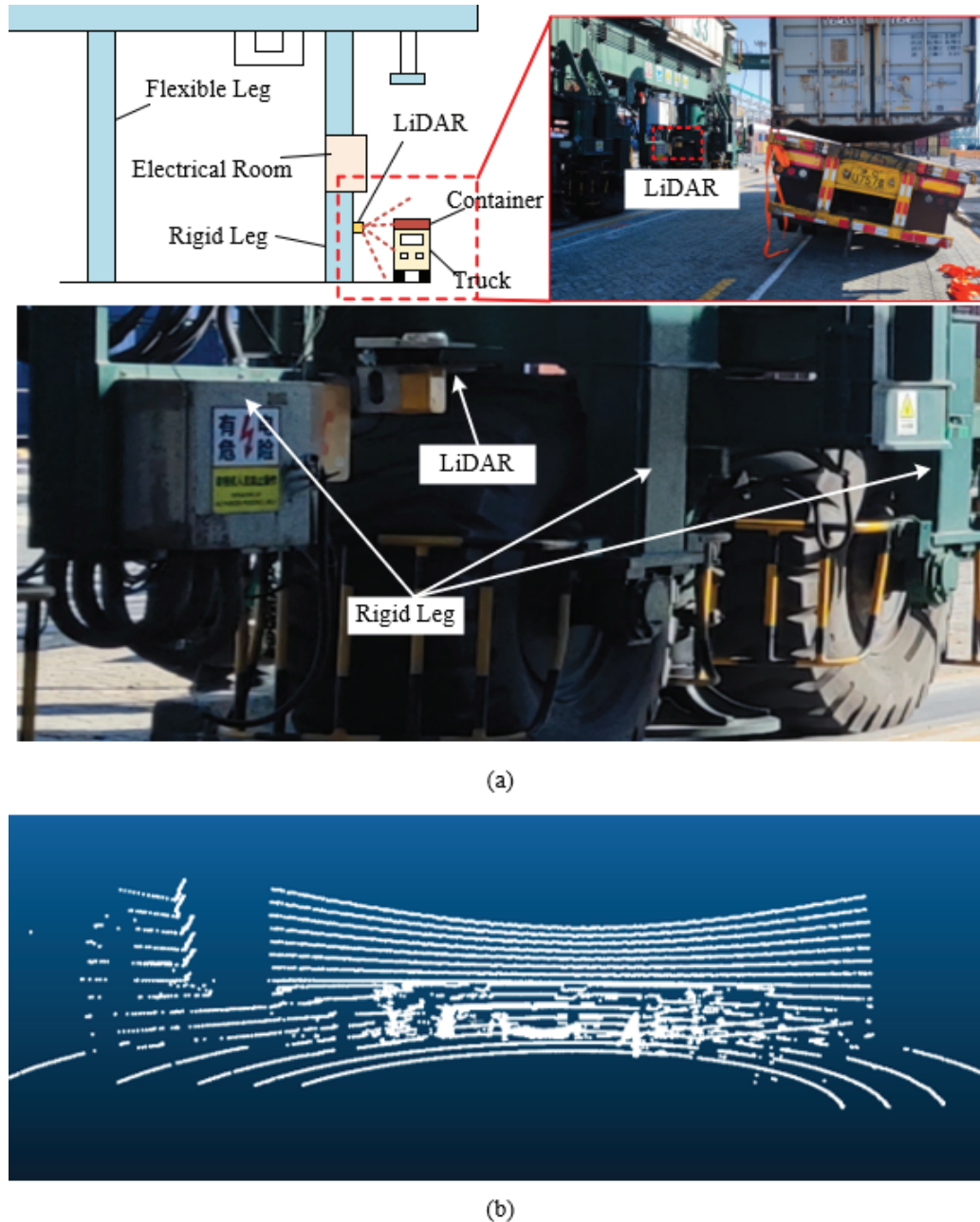


Figure 2. LiDAR installation and data collection. (a) Shows LiDAR installation; (b) shows a point cloud of a container being lifted as captured by the LiDAR.

3.2. Algorithm Design

3.2.1. Truck Lifting Accident Detection Process Framework

In the detection of truck lifting accidents, the quantification of boundary features for accident-related objects (such as containers, truck chassis, and wheels) is based on the spatial distribution and geometric relationships of points in the point cloud, including geometric attributes such as curvature, normal vectors, and the density of neighboring points. These features are crucial for distinguishing between different objects or components, especially in regions where boundaries are blurred. In this paper, an improved PointNet++ model is used to extract these boundary features from the point cloud and represent them as high-dimensional feature vectors, thereby enhancing the model's ability to recognize blurry boundary regions.

The truck lifting accident detection process based on multi-line LiDAR and the improved PointNet++ model is shown in Figure 3. First, multi-line LiDAR collects real-time

3D point cloud data from the scene. Through 3D Region of Interest (ROI) cropping, the point cloud data are preprocessed to extract the point cloud from the container lifting operation area, reducing data redundancy while retaining key features. The processed point cloud is then input into the improved PointNet++ model, which incorporates MLP to enhance the high-dimensional feature expression of the point cloud. It also combines MAM to optimize feature extraction for blurry boundary regions from both local and global perspectives. Through point cloud segmentation, the model can accurately identify key structures such as containers, truck chassis, and wheels, and further detect if there is any adhesion between the truck and the container after the container is lifted based on a distance threshold method, thus determining whether a truck lifting accident has occurred.

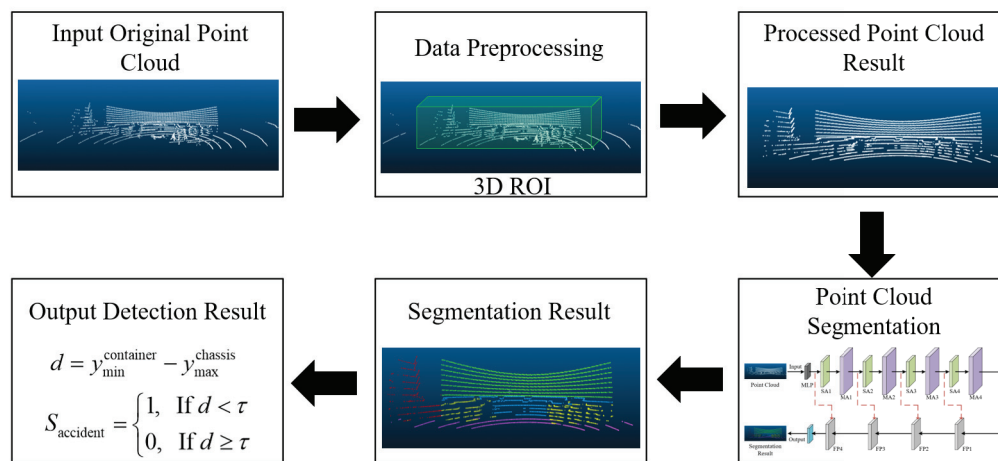


Figure 3. Truck lifting accident detection process.

3.2.2. Improved PointNet++ for Point Cloud Segmentation Model

PointNet++ effectively captures local features on multi-scale structures through its hierarchical feature extraction mechanism, making it suitable for general point cloud segmentation tasks. However, when processing point cloud data from complex container terminal scenarios, PointNet++ still suffers from insufficient dependence on local and global features, especially in segmenting point clouds with blurry boundary features in truck lifting accident detection. Therefore, this paper proposes the following improvements to address these issues:

(1) At the input stage, an MLP module [24] is introduced to enhance the model's ability to learn high-dimensional features, particularly when handling complex geometric structures. This enables the model to better capture subtle differences in boundary regions, such as minor relative positional changes between container and truck chassis, thereby improving segmentation accuracy. Additionally, through adaptive feature learning, the MLP enables the model to learn more suitable feature representations based on the varying scales of objects, further enhancing overall performance.

(2) At the feature abstraction stage, a feature extraction module based on MAM is introduced. MAM can perform weighted processing on different modalities of input data, such as point cloud positional data and boundary features, enabling the model to automatically focus on critical regions. This enhances the extraction of boundary features and strengthens the model's global perception capabilities. The MAM includes a channel attention mechanism (CAM) [25] for channel feature enhancement and a self-attention mechanism (SAM) [26] for global feature extraction. The CAM learns the correlations between different channels, assigning higher weights to important channels, thus improving the model's detection accuracy in truck lifting accident regions. The SAM enhances the model's ability

to capture global point cloud features, enabling more comprehensive analysis of boundary features between containers and truck chassis.

As shown in Figure 4, the improved PointNet++ network introduces MLP and MAM into the classic PointNet++ model, specifically adapting and optimizing it for complex container lifting scenarios to enhance point cloud segmentation performance.

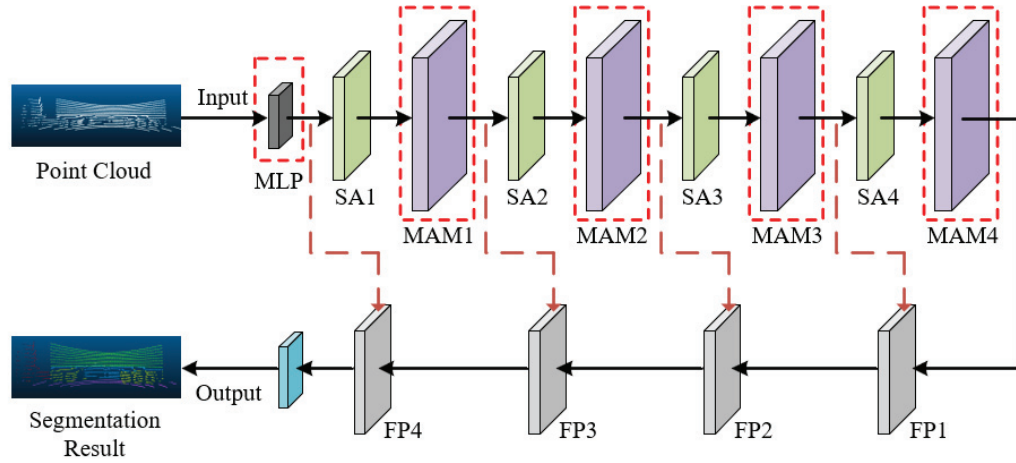


Figure 4. Improved PointNet++ network structure.

(1) Multi-layer Perceptron

In this paper, the input point cloud features are represented by the position coordinates of each point. Through the nonlinear transformations of hidden layers, the MLP can map these raw features into a higher-dimensional space, resulting in more abstract and discriminative feature representations.

As shown in Figure 5, the input point cloud feature matrix $X \in \mathbf{R}^{N \times d_0}$, where N is the number of points and d_0 is the initial feature dimension of each point. After processing by the MLP, the high-dimensional transformed features $F \in \mathbf{R}^{N \times d}$ are obtained, where d is the feature dimension after the high-dimensional transformation. The MLP consists of L fully connected layers, and the input and output features of each layer are represented by $z^{(l-1)}$ and $z^{(l)}$. The output of layer l is calculated as follows:

$$z^{(l)} = \sigma \left(W^{(l)} z^{(l-1)} + b^{(l)} \right) \quad (1)$$

where $W^{(l)} \in \mathbf{R}^{C_l \times C_{l-1}}$ represent the weight matrix and C_{l-1} and C_l represent the feature dimensions of the previous and current layers, respectively. $b^{(l)} \in \mathbf{R}^{C_l}$ is the bias, σ is the activation function used to introduce nonlinear transformations, and $z^{(0)} = X$ represents the input features of the first layer.

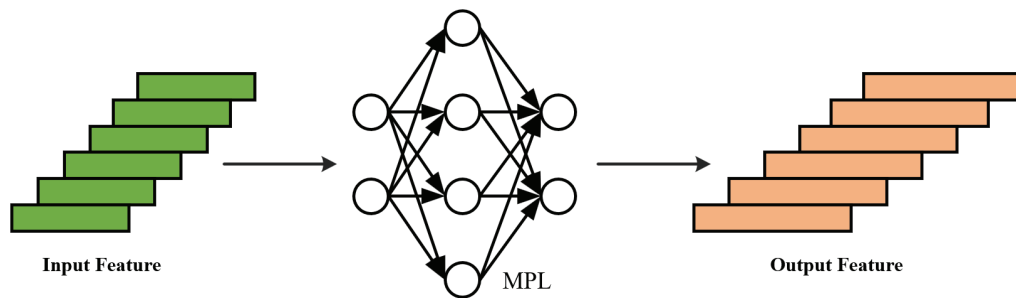


Figure 5. Multi-layer perceptron.

(2) Mixed Attention Mechanism

The classic PointNet++ model struggles to capture local and global feature dependencies effectively when processing point cloud data from complex container terminal scenarios. This is mainly due to the complex and variable structural features of containers and trucks. Therefore, we propose a feature extraction module based on MAM that combines both channel attention and self-attention to enhance the model's overall feature learning ability.

The implementation mechanism of MAM is shown in Figure 6, with the specific steps as follows:

Step One: Input the feature matrix $F \in \mathbf{R}^{B \times N \times (d+C)}$, where B is the batch size, N is the number of points in each sample, and C is the point set feature dimension.

Step Two: Enhance the cross-channel feature learning through the channel attention module, generating the feature vector $F_c \in \mathbf{R}^{B \times N \times (d+C_c)}$, where C_c is the feature dimension output by CAM.

Step Three: Simultaneously, the self-attention module strengthens the global feature dependencies between distant points in the point cloud, generating the feature vector $F_s \in \mathbf{R}^{B \times N \times (d+C_s)}$, where C_s is the feature dimension output by SAM.

Step Four: By concatenating F_c and F_s , the feature representation capability of the point cloud features is enhanced while compensating for any geometric feature loss. The enhanced feature can be represented as $F' = \text{Concat}(F_c, F_s)$.

Step Five: Output the new feature matrix $F' \in \mathbf{R}^{B \times N \times (d+C')}$, where $C' = C_c + C_s$.

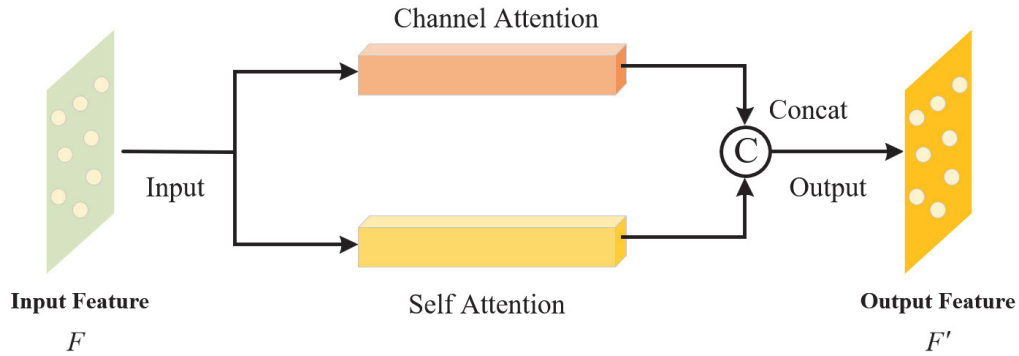


Figure 6. Feature extraction module based on mixed attention mechanism.

(a) Channel Attention Mechanism

CAM assigns different weights to different channels, enabling the model to prioritize critical features while suppressing irrelevant ones. This mechanism helps the model to better distinguish between overlapping or closely positioned objects. CAM achieves this by performing global max pooling and global average pooling on the input features, followed by a shared-weight MLP to generate channel attention weights. These weights are then applied to the original features, enhancing the representation of important channels. The main principle is illustrated in Figure 7, with the specific steps as follows:

The input point cloud feature F has dimensions $B \times N \times C$, where C is the number of channels. F is subjected to global max pooling and global average pooling, resulting in two feature vectors with dimensions $B \times 1 \times C$:

$$F_{\max} = \text{MaxPool}(F) \quad (2)$$

$$F_{\text{avg}} = \text{AvgPool}(F) \quad (3)$$

These two feature vectors are then input into a shared-weight MLP for dimensionality reduction as follows:

$$M_c = \sigma(\text{MLP}(F_{\text{avg}}) + \text{MLP}(F_{\text{max}})) \quad (4)$$

where σ is the Sigmoid activation function; M_c is the channel attention weight.

M_c is then multiplied with the original input feature F to obtain the enhanced channel feature F_c .

$$F_c = F \odot M_c \quad (5)$$

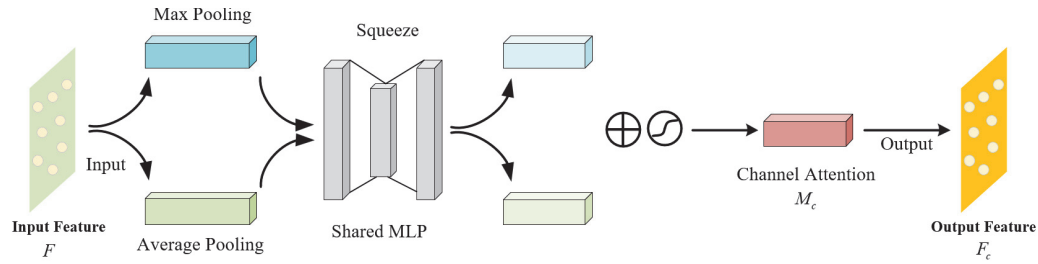


Figure 7. Channel attention mechanism.

(b) Self-Attention Mechanism

SAM complements the function of CAM by capturing long-range dependencies within the point cloud data. SAM calculates attention weights for each point by learning the internal relationships between points, thereby adaptively adjusting the importance of each point. This mechanism is particularly effective in handling irregular and complex structures. By strengthening global feature dependencies, SAM ensures that the model can accurately capture the spatial relationships between distant points. The workflow of SAM is shown in Figure 8 and is described as follows:

Similar to the CAM described earlier, the input point cloud feature F has dimensions $B \times N \times C$. Through three independent linear layers, the query matrix Q , key matrix K , and value matrix V are generated, which are expressed as follows:

$$Q = FW_q \quad (6)$$

$$K = FW_k \quad (7)$$

$$V = FW_v \quad (8)$$

where W_q , W_k , and W_v represent the weight matrices for Q , K , and V , each with dimensions $C \times C'$, and are shared across all samples in the batch.

The dot product similarity between Q and k is calculated, normalized using the Softmax function to obtain the attention weight matrix, which is then multiplied with the value matrix V to generate the weighted feature matrix F' :

$$F' = \text{Softmax}\left(\frac{QK^T}{\sqrt{C'}}\right)V \quad (9)$$

The point cloud feature F' is subsequently processed through convolutional linear transformations, batch normalization, and the ReLU activation function. Finally, the processed feature is added to the original feature vector, and the sum is computed to obtain the final output point cloud feature matrix F_s .

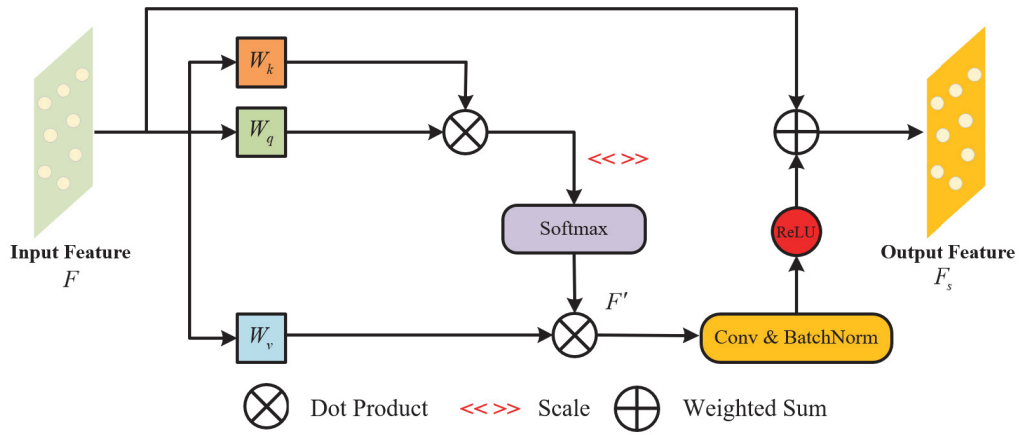


Figure 8. Self-attention mechanism.

3.2.3. Truck Lifting Accident Detection Algorithm

Based on the segmentation results from the network model, truck lifting accidents are precisely detected by analyzing the geometric relationship between the container and the truck chassis. The judgment logic is as follows: when the container is lifted to a detection height, if the container point cloud and the truck chassis point cloud still have adhesion, it may result in a truck lifting accident. By detecting the minimum distance between the two point clouds, we can effectively determine if an accident has occurred.

Assume the point set for the container lifting operation is represented by $P = \{p_i | p_i \in \mathbf{R}^3, i = 1, 2, \dots, N\}$, where $p_i = (x_i, y_i, z_i)$ is a point in the point set. After processing through the above point cloud segmentation network, we obtain the point set for container $P_{\text{container}}$ and the point set for truck chassis P_{chassis} . The minimum value of $P_{\text{container}}$ on the y-axis and the maximum value of P_{chassis} on the y-axis can be expressed as

$$y_{\min}^{\text{container}} = \min_{p_j \in P_{\text{container}}} y_j \quad (10)$$

$$y_{\max}^{\text{chassis}} = \max_{p_k \in P_{\text{chassis}}} y_k \quad (11)$$

The minimum distance between $P_{\text{container}}$ and P_{chassis} along the y-axis can be represented as

$$d = y_{\min}^{\text{container}} - y_{\max}^{\text{chassis}} \quad (12)$$

By setting a safety threshold τ , when the container is lifted to the detection height, the accident detection status S_{accident} is as follows:

$$S_{\text{accident}} = \begin{cases} 1, & \text{If } d < \tau \\ 0, & \text{If } d \geq \tau \end{cases} \quad (13)$$

This means that, if the minimum distance d between the container and the chassis is less than the safety threshold τ , a potential accident is detected ($S_{\text{accident}} = 1$); otherwise, if the distance exceeds or equals the threshold, the situation is deemed safe ($S_{\text{accident}} = 0$).

4. Point Cloud Segmentation Experiment Validation and Results Analysis

4.1. Container Lifting Operation Dataset

The improved PointNet++ model for truck lifting accident detection is applied to real container terminal handling machinery to effectively prevent truck lifting accidents. The hardware equipment includes the WLR-720 16-line LiDAR from Wanji Technology and

an edge computing module. Figure 9 shows the point cloud data collection process for 20-foot and 40-foot container lifting accidents.

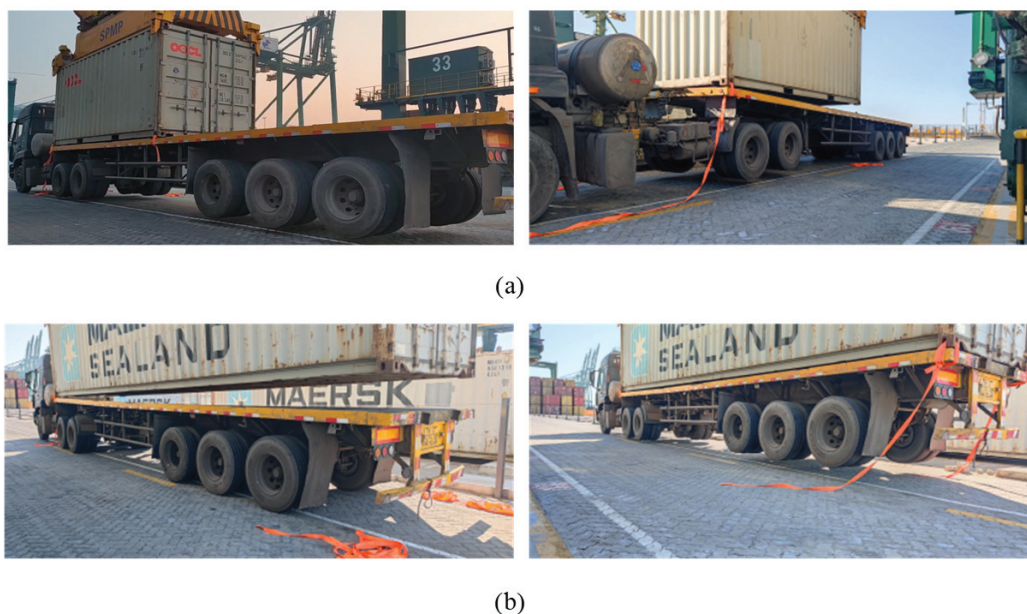


Figure 9. Point cloud dataset collection process. (a) Shows the collection process for a truck lifting accident involving a 20-foot container; (b) shows the collection process for a truck lifting accident involving a 40-foot container.

The main parameters of the LiDAR are shown in Table 1. Additionally, its maximum ranging distance is 60 m. To ensure point cloud density while covering the height of the truck unloading operation (0–3.5 m vertically), the LiDAR should be positioned at a horizontal distance of 2–8 m from the truck.

Table 1. Main parameters of the LiDAR.

Parameter	Value
Horizontal Scan Resolution	0.2 °
Vertical Scan FOV	−16°~14°
Vertical Scan Resolution	2°
Scan Frequency	10 Hz

Each scene is divided into six categories for point cloud segmentation: “Ground”, “Container”, “Truck Cab”, “Truck Chassis”, “Truck Wheel”, and “Others”, with the data annotated using CloudCompare software.

4.2. Experiment Environment Configuration

To verify the effectiveness of the proposed model, experiments were conducted based on a custom container lifting operation point cloud dataset. Table 2 outlines the hardware and software configurations utilized in the experiment.

The experimental training involved a total of 8911 point cloud samples, divided into training and test sets in an 8:2 ratio. Among these, 6230 samples represent normal scenarios (before and after normal container lifting), while 2681 samples represent anomalous scenarios (multiple types of truck lifting accidents). The training set contains 2145 accident samples, whereas the test set contains 536 accident samples. This distribution ensures a balanced representation of both normal and anomalous conditions during model training and evaluation.

Table 2. Experiment training environment configuration.

Hardware/Software	Configuration Parameters
CPU	Intel Xeon Processor E5 2680 v4
GPU	NVIDIA GeForce RTX 3090
Memory	64GB
Operating System	Ubuntu 20.04
Deep Learning Framework	Pytorch 1.8

4.3. Model Evaluation Metrics

The evaluation of boundary features is conducted after the point cloud segmentation. Based on the extracted boundary features, the model segments the point cloud into different categories. The accuracy of the segmentation is evaluated using the mean Intersection over Union (mIoU), where a higher mIoU value indicates better accuracy in boundary feature extraction and segmentation. Additionally, the Overall Accuracy (OA) and mean Class Accuracy ($mAcc$) are used to further quantify the model's overall performance in boundary feature extraction and segmentation tasks. These metrics are used to comprehensively evaluate the performance of point cloud segmentation. The formulas for calculation are as follows:

$$OA = \frac{\sum_{i=1}^C TP_i}{N} \quad (14)$$

$$mAcc = \frac{1}{C} \sum_{i=1}^C \frac{TP_i}{TP_i + FN_i} \quad (15)$$

$$mAcc = \frac{1}{C} \sum_{i=1}^C \frac{TP_i}{TP_i + FN_i} \quad (16)$$

where C is the number of categories. TP_i , FP_i , and FN_i represent the true positive, false negative, and false positive for category i .

4.4. Comparison Experiments and Analysis

In the experiment environment configuration used in this study, the following training parameters were set during model training: a batch size of 8, a total of 4096 sampling points, and the Adam optimization algorithm. The initial learning rate was set to 0.001, and the number of epochs was 200.

To explore the performance differences between the proposed optimized model and mainstream point cloud segmentation networks, comparison experiments were conducted using the same training parameters and sample training data. The comparison results of the training are shown in Table 3.

Table 3. Segmentation test results of different models.

Model	OA (%)	mAcc (%)	mIoU (%)	Training Time (h)	Throughput (ins./s.)
PointNet	85.4	73.5	67.9	13.4	158
PointCNN	87.6	79.1	74.1	19.6	143
Point Transformer	90.3	81.2	76.2	26.3	71
PointNet++	88.7	79.6	74.5	17.9	136
Ours	91.4	82.3	77.4	22.7	119

The experimental results show that PointNet, as a pioneering model in point cloud segmentation, performs well in extracting global features but has limitations in handling local geometric features. Its mIoU is relatively low at only 67.9%, showing a clear gap compared to other models. Point Transformer performs better than PointNet, achieving a mIoU of 76.2%, which surpasses PointNet++ (74.5%) and PointCNN (74.1%). However, it still falls 1.2% behind the improved model in terms of mIoU. The proposed improved model outperforms existing mainstream point cloud semantic segmentation models in terms of performance metrics. Its OA, mAcc, and mIoU reach 91.4%, 82.3%, and 77.4%, respectively. Compared to PointNet, PointCNN, Point Transformer, and PointNet++, the improved model shows mIoU increases of 9.5%, 3.3%, 1.2%, and 2.9%, respectively.

In terms of training time, the proposed model requires 22.7 h, which is longer than PointNet (13.4 h) and PointNet++ (17.9 h) but shorter than Point Transformer (26.3 h). Despite the increased training time, the proposed model achieves higher accuracy and robustness, making it suitable for truck lifting accident detection tasks. Additionally, the proposed model achieves a throughput (119 instances per second) that surpasses Point Transformer (71 instances per second) but slightly trails behind PointNet++ (136 instances per second). Despite this slight drop compared to PointNet++, the proposed model still maintains solid efficiency for real-time applications.

These results confirm that the proposed model has significant advantages in terms of segmentation accuracy and robustness, making it suitable for truck lifting accident detection tasks.

Figure 10 presents the visualized results of the segmentation comparison experiments. Under various container lifting operation scenarios, the proposed model performs better in accurately and clearly segmenting truck chassis and containers. In contrast, the baseline PointNet++ model exhibits incorrect segmentation in some edge regions. This difference is mainly due to PointNet++'s insufficient feature extraction ability in blurry boundary regions and limited ability to capture global features. The proposed model, however, enhances the learning of local detailed features and strengthens the interaction of features across channel and spatial dimensions, thus achieving finer segmentation in the boundary areas between containers and chassis.

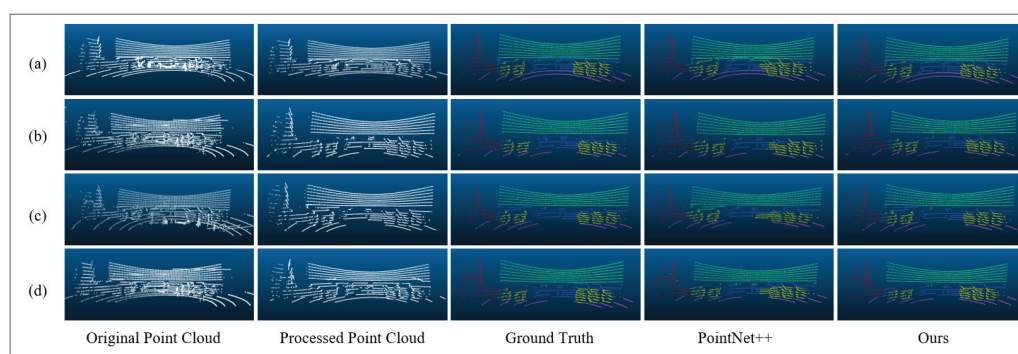


Figure 10. Visualization of point cloud segmentation results. (a) Shows a container remaining stationary without being lifted; (b) shows a container being lifted successfully under normal conditions; (c) shows a lifting accident occurring with the front lock engaged; (d) shows a lifting accident occurring with the rear lock engaged.

4.5. Ablation Experiment

To verify the improvement in point cloud segmentation performance achieved by the MLP and MAM modules, a series of ablation experiments were conducted, using PointNet++ as the baseline model. The experiments progressively added the improvement modules, and the performance was recorded. The results are shown in Table 4.

Table 4. Ablation experiment results comparison.

Model/Combination	MLP	CAM	SAM	OA(%)	mAcc (%)	mIoU (%)
PointNet++	×	×	×	88.7	79.6	74.5
+ MLP	✓	×	×	89.2	80.1	75.0
+ CAM	×	✓	×	89.5	80.5	75.4
+ SAM	×	×	✓	90.1	81.0	76.9
+ MLP + CAM + SAM (Ours)	✓	✓	✓	91.4	82.3	77.4

As shown in Table 4, the OA, mAcc, and mIoU of PointNet++ are 88.7%, 79.6%, and 74.5%, respectively. When MLP, CAM, and SAM are individually introduced, they enhance the baseline model's performance, demonstrating that these modules can improve the ability to extract both local and global features to a certain extent. Among them, MLP improves feature representation, resulting in a slight increase in mIoU to 75.0%, although the improvement in mIoU is relatively small. CAM enhances the model's focus on important channels, improving feature selection, and increasing mIoU to 75.4%. SAM performs the best among these three modules, leading to a significant increase in mIoU to 76.9%. When MLP, CAM, and SAM are combined, the OA, mAcc, and mIoU reach 91.4%, 82.3%, and 77.4%, respectively, showing a significant improvement compared to the baseline model. These results fully validate the effectiveness and complementarity of MLP and MAM modules in enhancing point cloud segmentation performance.

4.6. Engineering Application Comparison Test

To further study the generalization capability of the proposed improved PointNet++ model in real-world scenarios, multi-condition engineering application tests were conducted in an actual container terminal environment. The tests focused on examining the impact of different lighting and weather conditions on truck lifting accident detection. In the experiment, a server with identical hardware configurations was used to perform segmentation and detection on real-time point cloud data. To ensure comprehensive evaluation, each of the nine types of truck lifting accidents were artificially replicated under every test condition. The test results are shown in Table 5, where each condition is tested with a sample size of 500.

Table 5. Comparison of multi-condition real-world scene tests.

Condition	PointNet++				Ours			
	Acc (%)	FPR (%)	FNR (%)	Inf. Time (ms)	Acc (%)	FPR (%)	FNR (%)	Inf. Time (ms)
Daytime	88.6	13.0	9.0	105	97.4	1.7	4.0	120
Nighttime	89.2	12.3	8.6	106	97.8	1.5	3.0	122
Light Rain	86.8	13.7	12.5	112	96.1	2.6	5.5	125
Heavy Rain	82.8	17.5	18.5	116	93.0	4.7	10.5	130
Foggy	83.2	16.0	18.0	113	94.5	4.4	7.0	128

As shown in Table 5, the proposed model demonstrates superior performance compared to PointNet++ across all conditions. In stable conditions such as daytime and nighttime, the detection accuracy (Acc) of the proposed model is 97.4% and 97.8%, respectively, with false positive rates (FPRs) as low as 1.7% and 1.5% and false negative rates (FNRs) of 4.0% and 3.0%. In contrast, PointNet++ achieves lower Acc of 88.6% and 89.2% under the same conditions, with significantly higher FPRs of 13.0% and 12.3% and FNRs of 9.0% and 8.6%. The mean inference (Inf.) time of the proposed model is slightly

longer (120 ms and 122 ms) compared to PointNet++ (105 ms and 106 ms), but the improvement in accuracy and reduction in false positives and false negatives justify the additional computational cost.

As weather conditions deteriorate, the detection performance of both models decreases due to the reduction in the quality of the return signal from the LiDAR. During light rain, the Acc of the proposed model drops slightly to 96.1%, while PointNet++ achieves an Acc of 86.8%. In more adverse conditions such as heavy rain and fog, the detection performance is significantly affected. The proposed model maintains Acc of 93.0% and 94.5% in heavy rain and fog, respectively, with FPRs of 4.7% and 4.4% and FNRs of 10.5% and 7.0%. In comparison, PointNet++ achieves lower Acc of 82.8% and 83.2% in heavy rain and fog, with higher FPRs of 17.5% and 16.0% and higher FNRs of 18.5% and 18.0%.

Despite the slight increase in inference time, the proposed model demonstrates superior Acc, lower FPRs, and lower FNRs across all the conditions. By leveraging a continuous multi-frame detection mechanism, the system can effectively mitigate the impact of adverse weather conditions, maintaining its usability even in severe environments such as heavy rain and fog.

5. Conclusions

This paper addresses the safety issue of truck lifting accidents during container terminal operations by proposing a detection method based on multi-line LiDAR and the improved PointNet++ model. By incorporating a multi-layer perceptron and a mixed attention mechanism, the feature extraction and point cloud segmentation performance of the model are optimized, effectively improving the model's ability to segment blurry boundary regions. The experimental results show that this method significantly outperforms the traditional methods in terms of point cloud segmentation accuracy and robustness, with all the performance metrics showing substantial improvements. The real-world engineering tests further validate the model's reliability and stability under various conditions, especially under stable weather conditions. Compared with the existing detection methods based on visual cameras or single-line LiDAR, the proposed method fully utilizes the rich three-dimensional spatial information from multi-line LiDAR, providing higher precision and stronger adaptability for truck lifting accident detection.

Author Contributions: Conceptualization, Y.S. and X.M.; methodology, Y.S. and X.M.; software, J.W. and Y.Z.; validation, Y.S. and X.M.; formal analysis, Y.S. and C.M.; investigation, X.M. and J.W.; resources, C.M., J.W., and Y.Z.; data curation, Y.S.; writing—original draft preparation, Y.S., X.M., and J.W.; writing—review and editing, Y.S. and X.M.; visualization, Y.S. and C.M.; supervision, Y.S., X.M., and J.W.; project administration, Y.S. and J.W.; funding acquisition, Y.S. and C.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by National Natural Science Foundation of China (No. 52472435), the Education Science Research Project of Shanghai Municipality (No. B2023003), and the Science and Technology Commission of Shanghai Municipality (No. 22ZR1427700).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data. Data were obtained from Shanghai SMUVision Smart Technology Ltd. and are available from the authors with the permission of Shanghai SMUVision Smart Technology Ltd.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Huang, Q.F.; Huang, Y.G.; Zhang, Z.W.; Zhang, Y.J.; Mi, W.J.; Mi, C. Truck-Lifting Prevention System Based on Vision Tracking for Container-Lifting Operation. *J. Adv. Transp.* **2021**, *2021*, 9612480.
- Ji, Z.Y.; Zhao, K.K.; Liu, Z.X.; Hu, H.; Sun, Z.H.; Lian, S.G. A Novel Vision-Based Truck-Lifting Accident Detection Method for Truck-Lifting Prevention System in Container Terminal. *IEEE Access* **2024**, *12*, 42401–42410. [CrossRef]
- Asborn, M.I.; Burris, C.G.; Hernandez, S. Truck Body-Type Classification using Single-Beam Lidar Sensors. *Transp. Res. Rec.* **2019**, *2673*, 26–40. [CrossRef]
- Zhang, L.; Zhang, W.; Li, Y.; Li, B.; Li, C.; Sun, X. Acquisition Method of Container Lock Pin Model Based on Point Cloud Data. In Proceedings of the 2022 China Automation Congress (CAC), Xiamen, China, 25–27 November 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 490–494. [CrossRef]
- Ma, Z.; Postolache, O.; Yang, Y. Obstacle Avoidance for Unmanned Vehicle based on a 2D LIDAR. In Proceedings of the 2019 International Conference on Sensing and Instrumentation in IoT Era (ISSI), Lisbon, Portugal, 29–30 August 2019; IEEE: Piscataway, NJ, USA; pp. 1–6. [CrossRef]
- Zong, C.J.; Wan, Z.B. Container Ship Cell Guide Accuracy Check Technology Based on Improved 3D Point Cloud Instance Segmentation. *Brodogradnja* **2022**, *73*, 23–35. [CrossRef]
- Mi, C.; Shen, Y.; Mi, W.J.; Huang, Y.F. Ship Identification Algorithm Based on 3D Point Cloud for Automated Ship Loaders. *J. Coast. Res.* **2015**, *73*, 28–34. [CrossRef]
- Saglam, A.; Makineci, H.B.; Baykan, Ö.; Baykan, N.A. Clustering-Based Plane Refitting of Non-planar Patches for Voxel-Based 3D Point Cloud Segmentation Using K-Means Clustering. *Trait. Signal* **2020**, *37*, 1019–1027. [CrossRef]
- Zhang, J.X.; Lin, X.G.; Ning, X.G. SVM-Based Classification of Segmented Airborne LiDAR Point Clouds in Urban Areas. *Remote Sens.* **2013**, *5*, 3749–3775. [CrossRef]
- Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507. [CrossRef] [PubMed]
- Qi, C.R.; Su, H.; Mo, K.; Guibas, L.J. Pointnet: Deep learning on point sets for 3d classification and segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; IEEE: Piscataway, NJ, USA; pp. 652–660.
- Qi, C.R.; Yi, L.; Su, H.; Guibas, L.J. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017.
- Li, Y.; Bu, R.; Sun, M.; Wu, W.; Di, X.; Chen, B. Pointcnn: Convolution on x-transformed points. In *Advances in Neural Information Processing Systems*; Neural Information Processing Systems Foundation, Inc. (NeurIPS): La Jolla, CA, USA, 2018; Volume 31.
- Zhang, L.; Wang, H. A novel segmentation method for cervical vertebrae based on PointNet plus plus and converge segmentation. *Comput. Methods Programs Biomed.* **2021**, *200*, 105798. [CrossRef] [PubMed]
- Fan, Z.M.; Wei, J.H.; Zhang, R.Y.; Zhang, W.X. Tree Species Classification Based on PointNet plus plus and Airborne Laser Survey Point Cloud Data Enhancement. *Forests* **2023**, *14*, 1246. [CrossRef]
- Hu, H.; Zhang, G.H.; Ao, J.F.; Wang, C.L.; Kang, R.H.; Wu, Y.L. Multi-information PointNet plus plus fusion method for DEM construction from airborne LiDAR data. *Geocarto Int.* **2023**, *38*, 2153929. [CrossRef]
- Lin, Y.W.; Knudby, A.J. Global automated extraction of bathymetric photons from ICESat-2 data based on a PointNet plus plus model. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *124*, 103512. [CrossRef]
- Patel, A.K.; Park, E.S.; Lee, H.; Priya, G.G.L.; Kim, H.; Joshi, R.; Arief, M.A.A.; Kim, M.S.; Baek, I.; Cho, B.K. Deep Learning-Based Plant Organ Segmentation and Phenotyping of Sorghum Plants Using LiDAR Point Cloud. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 8492–8507. [CrossRef]
- Jing, Z.W.; Guan, H.Y.; Zhao, P.R.; Li, D.L.; Yu, Y.T.; Zang, Y.F.; Wang, H.Y.; Li, J. Multispectral LiDAR Point Cloud Classification Using SE-PointNet plus. *Remote Sens.* **2021**, *13*, 2516. [CrossRef]
- Guo, R.C.; Xie, J.L.; Zhu, J.X.; Cheng, R.F.; Zhang, Y.; Zhang, X.H.; Gong, X.J.; Zhang, R.W.; Wang, H.; Meng, F.F. Improved 3D point cloud segmentation for accurate phenotypic analysis of cabbage plants using deep learning and clustering algorithms. *Comput. Electron. Agric.* **2023**, *211*, 108014. [CrossRef]
- Chen, R.; Yang, M.; Guan, A.; Lan, Z.; Wang, L.; Wang, Y.; Niu, Q. Cigarette Perforation Point Cloud Segmentation and Hole Depth Calculation Based on the Improved PointNet++ Network and DMCP Algorithm. *IEEE Sensors J.* **2024**, *24*, 21048–21061. [CrossRef]
- Cheng, G.; Gu, C. Enhancing Point Cloud Classification and Segmentation With Attention-Enhanced SO-PointNet++. *IEEE Access* **2024**, *12*, 195986–195995. [CrossRef]
- Yu, L.; Yu, H.; Yang, S. A Deep Neural Network Using Double Self-Attention Mechanism for ALS Point Cloud Segmentation. *IEEE Access* **2022**, *10*, 29878–29889. [CrossRef]
- Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **1958**, *65*, 386–408. [CrossRef] [PubMed]

25. Hu, J.; Shen, L.; Albanie, S.; Sun, G.; Wu, E.H. Squeeze-and-Excitation Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 2011–2023. [CrossRef] [PubMed]
26. Vaswani, A. Attention is all you need. In *Advances in Neural Information Processing Systems*; Neural Information Processing Systems Foundation, Inc. (NeurIPS): La Jolla, CA, USA, 2017. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

A Decomposed-Ensemble Prediction Framework for Gate-In Operations at Container Terminals

Yifan Shen ¹, Beng Xuan ^{2,*}, Hongtao Hu ^{3,*}, Yansong Wu ⁴, Ning Zhao ³ and Zhen Yang ²

¹ Ministry of Education Engineering Research Center for Container Supply Chain Technology, Shanghai Maritime University, Shanghai 201306, China

² Institute Logistics Science and Engineering, Shanghai Maritime University, Shanghai 201306, China

³ Logistics Engineering College, Shanghai Maritime University, Shanghai 201306, China

⁴ Xiamen Port Holding Group Co., Ltd., Xiamen 361013, China

* Correspondence: xuanbeng0001@stu.shmtu.edu.cn (B.X.); hth@shmtu.edu.cn (H.H.); Tel.: +86-150-6858-7206 (B.X.); +86-138-1832-4958 (H.H.)

Abstract: Container terminals play a crucial role in global logistics and trade, with gate-in operations significantly impacting overall terminal efficiency and cargo turnover speed. This paper analyzes a series of problems caused by the randomness of the arrival of export containers at the container yard, including wastage of yard space, excessive waiting time for external trucks, and conflicts with other production operations. To address these issues, a method based on a decomposed ensemble framework is proposed to predict short-term container quantities for gate-in operations at container terminal gates. The experiment compares the autoregressive integrated moving average (ARIMA) algorithm, the prophet algorithm, and the Long Short-Term Memory (LSTM) algorithm, with results indicating the clear advantage of Long Short-Term Memory in decomposed time series modeling. The introduction of this method is expected to enhance the accuracy and flexibility of terminal production planning, optimizing resource utilization. Contributions of this paper include the proposal of predictive models, a shipping route-based decomposed-ensemble framework, and confirmation of the superiority of Long Short-Term Memory in prediction through comparative analysis. These contributions are expected to improve terminal operational efficiency, reduce resource wastage, and better adapt to the highly stochastic gate-in operation environment.

Keywords: container terminal; gate-in operations; randomness; time series prediction; decomposed ensemble framework; LSTM algorithm

1. Introduction

Container terminals play a crucial role in global logistics and trade by serving as pivotal nodes for container imports and exports, leading to increasing attention on the research of port operational efficiency [1]. These terminals connect various segments of the supply chain, with gate-in operations for export containers in the yard occupying a considerable proportion of terminal operations, significantly impacting overall terminal efficiency and cargo turnover speed. Efficient gate-in operations directly influence on-site production efficiency and customer satisfaction, prompting terminals to allocate substantial resources to optimize this process [2].

The figure below (Figure 1) illustrates the cycle of containers within a terminal [3]. Initially, containers enter the yard through the gate, referred to as gate-in operations. Subsequently, equipment loads containers onto the vessel, known as loading operations.

Following this, containers unloaded from the vessel enter the yard, termed unloading operations. Finally, customers transport containers unloaded from the vessel away from the terminal through the gate, referred to as gate-out operations. During these processes, the loading and unloading operations can be accurately planned by the terminal since the vessel's schedule is determined, and stowage information is available. Conversely, gate-in and gate-out operations are inherently stochastic for the terminal due to various unpredictable factors, such as uncertain times for customers to deliver or pick up containers. This study specifically focuses on predicting gate-in operations for containers entering the terminal. Accurate forecasting of gate-in quantities is essential for optimizing stacking arrangements and resource allocation within the yard. By employing predictive modeling techniques, such as the proposed decomposed ensemble framework, this research aims to improve the accuracy of forecasts associated with gate-in operations. Such enhancements are critical for mitigating the uncertainties inherent in gate-in activities, thereby enabling terminal operators to make informed decisions and streamline operational efficiency.

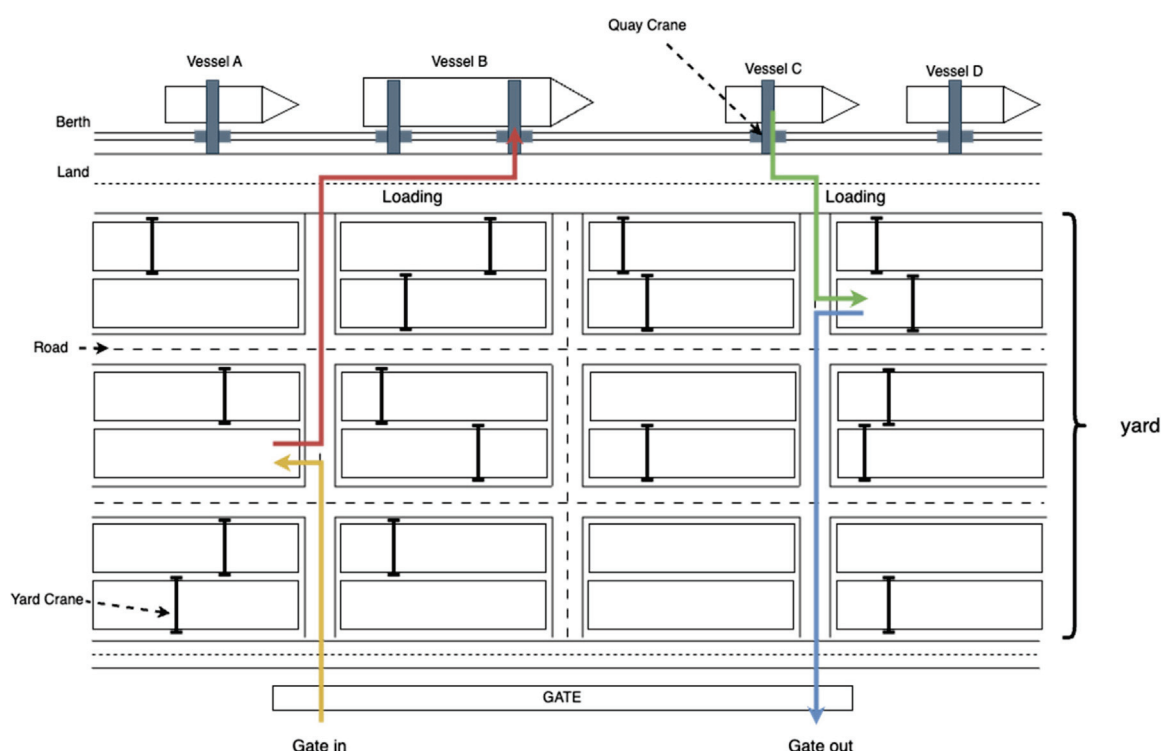


Figure 1. The cycle of containers.

Typically, the vessel planning personnel at container terminals develop monthly, weekly, and daily berthing plans for vessels [4]. Once the berthing plan and dates for a vessel are determined, yard planning personnel formulate the gate-in plan for that vessel [5]. In this plan, they usually set the opening date and closing dates of the gate. Customers are required to deliver containers to the terminal yard via external trucks between these two dates. Subsequently, yard production personnel carry out the gate-in operations using rubber-tired gantry cranes. This process ensures that containers requiring loading onto vessels can arrive at the terminal within the specified timeframe.

Despite the formulation of gate-in plans by yard planning personnel, the arrival of customers delivering containers to the terminal yard constitutes a stochastic event. This randomness gives rise to various challenges during on-site gate-in operations in the yard. Firstly, during the gate-in process for export containers, there is significant wastage of yard space. Since planners cannot allocate real-time positions for each container, a particular area in the yard is usually chosen for the gate-in operations of a specific vessel. Consequently,

yard space is substantially wasted during this gate-in process. As illustrated in Figure 2, the area below the dashed line represents the available slots for stacking arrangements over time, while the area below the solid line indicates the cumulative gate-in container quantities over time. The area marked by orange dots signifies the allocation of wasted slots resources. The black dotted line indicates the time of the gate-in containers start loading operations. A significant overlap between the two areas suggests that the stacking plan is well-structured and minimizes unnecessary waste. Therefore, it can be inferred that the effectiveness of stacking arrangements largely depends on the accuracy of container quantity forecasting. Secondly, during the gate-in process for export containers, external trucks experience prolonged waiting times in the yard. Congestion is prone to occur in the yard during peak gate-in operations, posing safety hazards. Additionally, gate-in operations in the yard can easily impact other production operations in the yard, particularly exerting a significant influence on loading operations. Typically, the container area designated for export container plans is also the loading area, where a considerable number of loading operations take place. As a result, conflicts may arise between gate-in operations and loading operations.

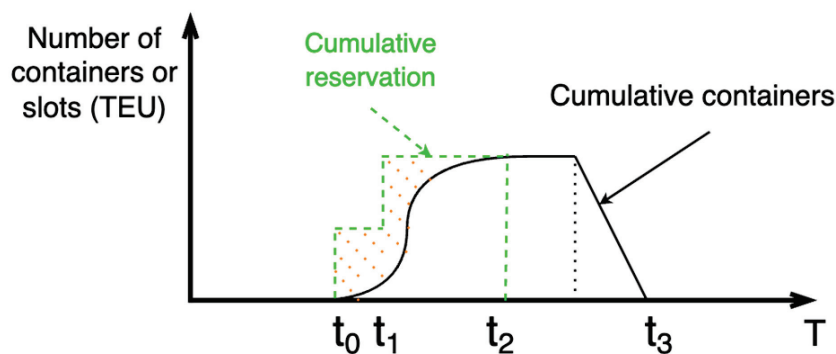


Figure 2. Wasted space during gate-in process.

The fundamental issue leading to the challenges mentioned above is the stochastic nature of the arrival process for export containers in the yard. Despite yard planning personnel being able to formulate gate-in plans and determine opening and closing dates, the exact distribution of the arrivals of export containers during this period cannot be accurately determined. Due to the inability to pinpoint the precise arrival distribution of export containers, yard production personnel face a dilemma when devising yard production plans: they cannot ascertain the production operation requirements for export containers. Overinvestment in production resources would result in wastage, while insufficient resource allocation would fail to meet production demands. Yard production personnel struggle to predefine feasible production plans, and even if plans are devised in advance, they frequently prove ineffective during on-site production. The stochastic nature of gate-in operations leads to numerous conflicts with other production operations.

The majority of container terminals in China operate on a 24 h production system. The Figure 3 below illustrates the statistical data for gate-in (export) container operations in the yard at a container terminal in Xiamen over a 24 h period. The chart provides a clear depiction of the volume of gate-in operations in the yard for each hour.

This data serves as a valuable resource for optimizing resource allocation, enhancing operational efficiency, and addressing challenges associated with the stochastic nature of gate-in operations. Yard production personnel can use this information to better anticipate and plan for the varying demand for gate-in operations throughout the day.

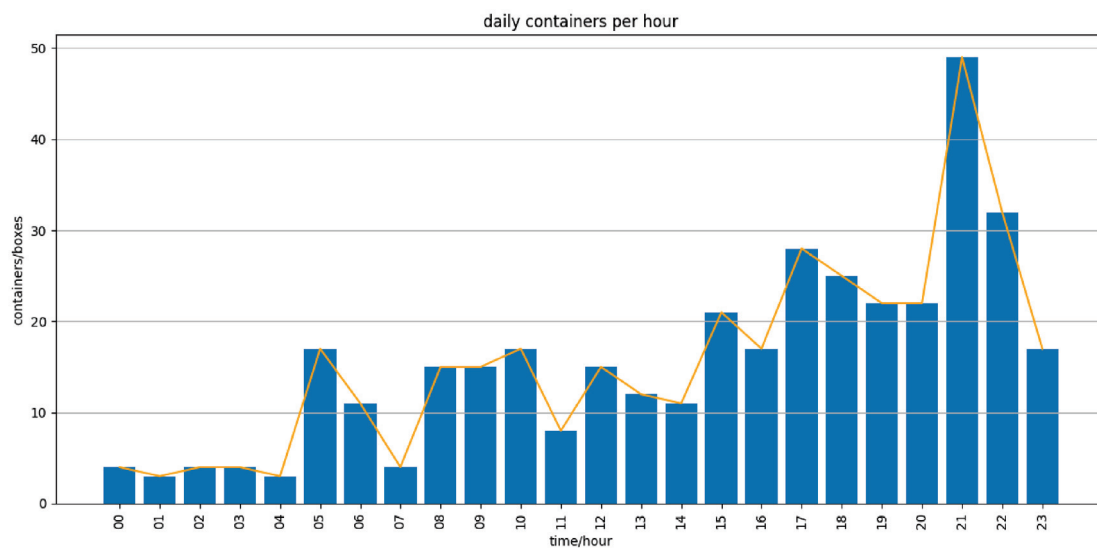


Figure 3. Description of the statistical data for 24 h gate-in operations.

Based on the analysis above, this paper presents a method that employs a decomposed ensemble framework to predict the volume of short-term gate-in operations. The primary aim is to forecast the hourly gate-in operation volume for the upcoming 24 h, thereby determining the production demand for gate-in activities within the yard for the next cycle. To achieve this, the study utilizes three machine learning algorithms: the classical ARIMA algorithm, the prophet algorithm, and the LSTM algorithm. The contributions of the paper are summarized as follows:

The study introduces a model for the short-term prediction of gate-in container volumes at container terminals, assisting planners by providing hourly forecasts of gate-in operations.

A shipping route-based decomposed ensemble framework is proposed, which disaggregates the distribution of container volumes in the gate-in process, thereby enhancing the precision of the prediction results.

By employing three time series forecasting techniques and conducting comparative analysis, the research identifies LSTM as being superior in predicting gate-in operations compared to traditional linear time series forecasting methods.

The next section reviews previous relevant works, analyzing both long-term and short-term container throughput forecasting research. Section three introduces the shipping route-based decomposed-ensemble framework. Section four details the methods applied in this study. Section five presents experimental validation, analyzing and comparing the effectiveness of the proposed method using historical gate-in operation data. Section six concludes the work and suggests directions for further research.

2. Literature Review

Numerous researchers and scholars have conducted studies on the prediction of container throughput, examining the topic from various perspectives. Some focus on the port level [6–8], while others approach it from the container terminal perspective [9–11]. Additionally, there are studies considering the viewpoints of vessels and shipping companies [12]. Some researchers also examine the throughput of gate-out export containers from the side, predicting it through truck forecasts [13]. The literature is categorized into long-term and short-term predictions, with weekly, monthly, quarterly, and yearly predictions considered long-term, while daily and hourly predictions are considered short-term.

2.1. Long-Term Prediction

Many researchers predicted annual port throughput. Twrady and Batista [14] used a Markov Chain model, time series trend models, time series trend models with periodic terms, and Grey System models to predict container annual container throughput in the North Adriatic Sea ports. Rashed et al. [15] proposed a three-step method combining autoregressive distributed lag models with economic scenarios to forecast container throughput in major ports within the Hamburg-Le Havre range. Cuong et al. [16] introduced a multi-criteria decision-making strategy based on a neural network predictive controller and an adaptive fractional-order super-twisting sliding mode controller, showing its effectiveness in the recovery strategy of container throughput growth in Vietnamese ports.

Schulze and Prinz [7] studied quarterly container throughput in German ports using SARIMA and Holt-Winters exponential smoothing methods. SARIMA slightly outperformed Holt-Winters in forecasting container throughput, especially in quarterly data with seasonal behavior.

Several studies focus on long-term container throughput prediction, often based on monthly data. In monthly throughput prediction, Tian et al. [17] proposed a TEI@I-based model for comprehensive monthly container throughput prediction, demonstrating effectiveness with Hong Kong port data. Pang and Gebka [18] developed a prediction method for the monthly throughput of Indonesia's Tanjung Priok port based on single-container terminal data, showing that SARIMA excelled in overall port throughput forecasting, while VAR/VECM was more applicable for terminal-level predictions. Niu et al. [19] proposed a hybrid decomposed ensemble model (VMD-ARIMA-HGWO-SVR), achieving low and high-frequency mode predictions through variational mode decomposition (VMD) and integrating ARIMA and Grey Wolf Optimized Support Vector Regression (HGWO-SVR) models. Sanguri et al. [20] introduced a cross-time coordination framework for hierarchical container throughput (CT) prediction, significantly improving predictive performance.

Apart from traditional methods, Chen and Chen [6] utilized heuristic algorithms and established optimal models for accurate monthly port container throughput predictions, comparing genetic programming (GP), decomposition (X-11), and seasonal autoregressive integrated moving average (SARIMA), with GP proving to be superior.

With the development of machine learning and deep learning algorithms, researchers employ these techniques for container throughput prediction. Yang and Chang [8] proposed a method for monthly container throughput demand prediction using a hybrid precision neural network architecture, demonstrating higher predictive performance compared to traditional machine learning methods. Shankar et al. [21] used Long Short-Term Memory (LSTM) networks for container throughput prediction, confirming LSTM's advantages over common time series forecasting methods through monthly throughput data for Singapore's Port of Singapore Authority (PSA). Lee et al. [22] presented a focus on a time series decomposition-based LSTM model, outperforming traditional LSTM models and aligning with trends using monthly container throughput data from South Korea's Busan port.

2.2. Short-Term Prediction

Research on short-term prediction is less common compared to long-term prediction. Gao, Chang, Chen, and Fang [10] employed Long Short-Term Memory (LSTM) networks to forecast daily container intake at the yard. In subsequent research, Gao, Chang, Fang, and Fan [11], preprocessed historical data using boxplots, and the LSTM model was trained for daily average container volume prediction. Comparative results with other methods such as ARIMA and BP neural networks demonstrated LSTM's superior performance and lower prediction errors. Jin, Ma, Jin, Cui, and Bai [9] introduced a decomposed-ensemble framework and used the XGBoost machine learning method to predict daily

gate movements at container terminals. Through research on historical data from China's Ningbo-Zhoushan Port, the proposed framework and XGBoost method showed significant advantages in error metrics.

Wang, Qin, and Shen [2] achieved precise prediction of hourly outbound container arrivals at container terminals by combining KNN with deep learning. The study validated the method's effectiveness using real-world automated terminal data. Shen et al. [23] designed a voyage prediction system, utilizing the KNN machine learning algorithm and statistical methods to predict the distribution of container arrival times for given voyages in specific ports within 0–120 h. Wang, Shi, and Sun [12] proposed a model for predicting the number of containers on a voyage, using a genetic algorithm (GA) to optimize SVM parameters. Nadi, Sharma, Snelder, Bakri, van Lint, and Tavasszy [13] proposed a short-term (hourly) truck traffic prediction method based on feedforward neural networks and logistics information, indirectly reflecting gate-out container throughput.

2.3. Summary

Table 1 summarizes the reviewed literature. Despite extensive research on container throughput prediction, studies specifically focusing on gate-out export container throughput are relatively scarce. Existing short-term predictions primarily address daily gate throughput, offering limited assistance to terminal production planning decisions. Literature related to hourly predictions often focuses on voyages or vessel quantities, leading to disruptions in time series continuity due to the uncertainty of voyages or vessels, resulting in significant prediction errors and limits applicability to specific voyages or vessels.

Table 1. Literature summary.

Article	Method	Data	Period
Twrdy and Batista [14]	Markov Chain model, time series trend models, time series trend models with periodic terms, and grey system models	North Adriatic Sea ports	Yearly
Rashed, Meersman, Sys, Van de Voorde, and Vanelslander [15]	Three-step method	Hamburg-Le Havre Range of Port	Yearly
Cuong, Kim, You, and Nguyen [16]	Multi-criteria decision-making strategy	Vietnam Ports	Yearly
Schulze and Prinz [7]	SARIMA, Holt-Winters	German Ports	Seasonally
Tian, Liu, Lai, and Wang [17]	TEI@I-based Model	Hong Kong Port	Monthly
Pang and Gebka [18]	SARIMA, VAR/VECM	Tantung Priok Port	Monthly
Niu, Hu, Sun, and Liu [19]	VMD-ARIMA-HGWO-SVR	PSA, Shanghai Port	Monthly
Sanguri, Shankar, Punia, and Patra [20]	Cross-time coordination framework	Los Angeles Port	Monthly
Chen and Chen [6]	Heuristic algorithms GP, X-11, SARIMA	Taiwan Port	Monthly
Yang and Chang [8]	Hybrid precision neural network architecture	Taiwan Port	Monthly
Shankar, Ilavarasan, Punia, and Singh [21]	Long Short-Term Memory (LSTM) Networks	PSA	Monthly
Lee, Kim, and Bae [22]	Time series decomposition-based LSTM model	South Korea's Busan Port	Monthly
Gao, Chang, Chen, and Fang [10]	Long Short-Term Memory (LSTM)	-	Daily
Gao, Chang, Fang, and Fan [11]	Long Short-Term Memory (LSTM), Boxplots	-	Daily
Jin, Ma, Jin, Cui, and Bai [9]	Decomposed-ensemble framework, XGBoost	Ningbo-Zhoushan Port	Daily
Wang, Qin, and Shen [2]	Combining KNN with deep learning	Shanghai Port	Hourly
Shen, Cao, Yang, and Guo [23]	KNN, statistical methods	Shanghai Port	Hourly
Wang, Shi, and Sun [12]	GA to optimize SVM parameters	Simulation Data	Hourly
Nadi, Sharma, Snelder, Bakri, van Lint, and Tavasszy [13]	Feedforward neural networks	Rotterdam Port	Hourly

Considering the analysis, this paper proposes a method based on the examination of container shipping route characteristics, identifying certain patterns in the temporal sequence for each shipping route. The proposed approach involves decomposing gate-out container intake by shipping route, predicting the hourly intake for each shipping route, and accumulating the intake for each shipping route to obtain the total hourly gate-out intake, aiding real-time coordination of yard resources for on-site production personnel.

3. Decomposed-Ensemble Framework

In this section, we introduce the proposed method, which is a decomposed-ensemble framework based on the characteristics of container shipping routes. The prediction problem for the total hourly intake is decomposed into the prediction of the intake for each shipping route on an hourly basis. The research framework is illustrated in the following Figure 4.

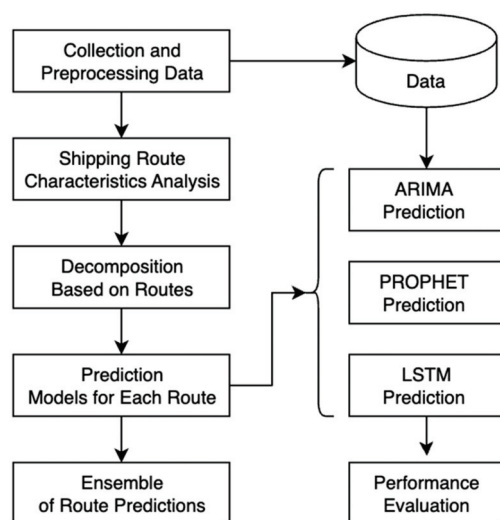


Figure 4. Research framework.

The proposed framework offers a systematic approach to addressing the challenges posed by the stochastic nature of gate-out container intake. By decomposing the prediction task according to each container shipping route, the model captures the specific patterns and variations associated with different shipping routes, providing a more precise and granular forecast.

3.1. Shipping Route Characteristics Analysis

Shipping routes [24] refer to the predetermined course of a vessel on the sea or inland waterways, typically including the starting point, destination, and intermediate waypoints. Liners [24] are vessels engaged in regular shipping along fixed shipping routes, adhering to a schedule and specific freight rates. Consequently, vessels on the same shipping route exhibit noticeable periodicity over time.

As depicted in Figure 1, the containers received at the gate undergo subsequent operations at the terminal, primarily involving loading onto ships. Hence, there is a certain correlation between the arrival time of containers at the terminal and the arrival time of the vessels at the port. For a given shipping route, the arrival time of containers at the terminal also demonstrates apparent periodicity, as illustrated in the following Figure 5, where the curve represents the varying quantity of container arrivals over time. Additionally, most container terminals establish a drop-off cycle for containers based on shipping routes, typically ranging from 5 to 7 days. Within this cycle, the arrival of containers at the terminal follows a certain distribution pattern.

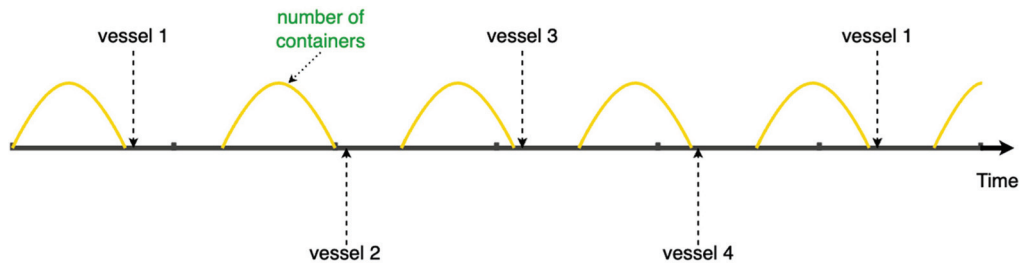


Figure 5. Quantity of containers in time series.

From the aforementioned analysis, it is evident that there is a correlation between the volume of containers received at the gate for each shipping route and time. Container shipping routes are analyzed to identify patterns and trends associated with each shipping route. This analysis includes understanding the temporal distribution of container entries for different shipping routes, considering factors such as peak hours. Therefore, predicting the quantity of container arrivals at the storage yard based on time is feasible.

3.2. Decomposition Based on Shipping Routes

Directly predicting the total volume of outbound containers for the entire storage yard poses certain challenges. As illustrated in Figure 6, it is assumed that the volume of containers received for 10 shipping routes follows a normal distribution over time, and the total volume of containers received within each time period is the result of the cumulative reception across multiple shipping routes. However, different shipping routes have distinct distributions. If predictions are made by stacking the volumes of multiple shipping routes, it significantly increases the complexity of the data and the difficulty of prediction.

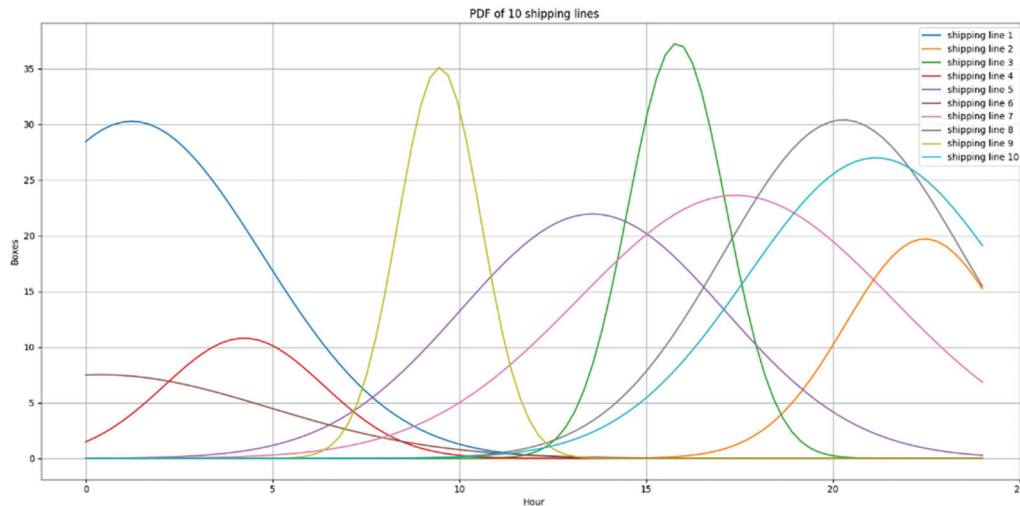


Figure 6. Volume of containers received for 10 shipping routes.

The total hourly intake prediction task is decomposed into sub-tasks, each focused on predicting the intake for a specific container shipping route on an hourly basis. This decomposition allows the model to address the unique characteristics and variations present in each shipping route's intake.

Therefore, this study has opted for a method that decomposes the total gate-in container volume based on container shipping route attributes, independently forecasting the outbound container volume for each shipping route. The decomposition formula is as follows:

$$C_{t+\Delta t} \rightarrow S_{t+\Delta t}^1, S_{t+\Delta t}^2, S_{t+\Delta t}^3, \dots, S_{t+\Delta t}^n \quad (1)$$

where $C_{t+\Delta t}$ represents the total volume of gate-in containers during the $t + \Delta t$ time period, and $S_{t+\Delta t}^n$ represents the total volume of containers received for shipping route n during the $t + \Delta t$ time period. The symbol $\rightarrow$ indicates the decomposition.

3.3. Ensemble of Shipping Route Predictions

The individual predictions for each shipping route are combined into an ensemble prediction for the total hourly intake. This ensemble approach leverages the strengths of each shipping route-specific model, providing a holistic and accurate forecast for the entire gate-out container intake.

The container prediction model can be described by the following Formula (2):

$$\hat{s}_{t+\Delta t} = f(X_t) + \varepsilon_t \quad (2)$$

where $\hat{s}_{t+\Delta t}$ is the predicted value, Δt is the prediction time period, $X_t = \{x_{t-l}, x_{t-l+1}, \dots, x_{t-1}, x_t\}$, where x_t represents time series data, l denotes the range of the time series, and ε_t is the time prediction error.

Based on the above formula, the prediction for the outbound container volume for each shipping route in each time period can be calculated. The results can be ensembled to obtain the total predicted volume of outbound containers for a specific time period, as shown in the following Formula (3):

$$\hat{C}_{t+\Delta t} = \sum_{i=1}^n \hat{S}_{t+\Delta t}^i \quad (3)$$

where $\hat{C}_{t+\Delta t}$ represents the total predicted volume of outbound container reception for the time period Δt .

4. Methodology Formulation

This paper employs three classic machine learning algorithms for prediction: the Autoregressive Integrated Moving Average (ARIMA) algorithm [25], the prophet algorithm [26], and the Long Short-Term Memory (LSTM) algorithm [27].

4.1. Autoregressive Integrated Moving Average (ARIMA)

ARIMA is a statistical analysis model commonly used for time series forecasting, comprising three components: AR, MA, and I. Autoregression (AR) is a model that describes the relationship between the current value and historical values. It predicts a variable using its own historical data. The mathematical expression is as follows (4):

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \varepsilon_t \quad (4)$$

where y_t represents the observation at time t , γ_i denotes the regression coefficients, ε_t is the error term, and p represents the order of the regression. From the formula, it can be inferred that the value at time t is correlated with its past values. Therefore, for AR models to be effective, the data must exhibit correlation, and the AR model is effective when the correlation coefficient is greater than 0.5.

Moving Average (MA) represents the time series as a linear combination of unrelated random variables, known as white noise. MA accumulates the error terms from the autoregressive model, effectively eliminating random fluctuations in predictions. The Formula (5) is as follows:

$$y_t = \mu + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (5)$$

where y_t is the observation at time t , θ_i represents the weighting coefficients, and $\{\varepsilon_{t-i}\}$ represents the white noise sequence.

Adding Formulas (4) and (5) yields the $ARMA(p, q)$ model [25]. However, if the observed time series data is non-stationary, ARMA cannot be used. Non-stationary sequences need to be transformed into stationary sequences; a process known as differencing. ARMA with differencing is referred to as $ARIMA(p, d, q)$ [25], where d is the differencing order.

In most cases, observed datasets exhibit noticeable periodicity, and ARIMA performs poorly in predicting periodicity. To address this, the Seasonal ARIMA (SARIMA) model $(p, d, q)(P, D, Q)$ is proposed to account for the seasonality. The formula (additive mode) is as follows:

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i} + \sum_{i=1}^P \Phi_i y_{t-i \cdot s} + \sum_{i=1}^Q \Theta_i \epsilon_{t-i \cdot s} \quad (6)$$

where P is the order of the seasonal autoregressive part, Q is the order of the seasonal moving average part, s is the length of the seasonal cycle, Φ_i represents the seasonal autoregressive coefficients, and Θ_i represents the coefficients of the seasonal moving average part.

4.2. Prophet

Prophet [26] is a time series forecasting algorithm developed by Facebook, designed specifically to handle time series data with seasonality, holiday effects, and trends. The prophet model decomposes time series data into three main components: Trend, which describes the long-term trend of the time series; seasonality, which captures the periodic changes that repeat within a year; and the holiday effect, which considers irregular events that may occur during holiday periods. The formula is as follows:

$$y_t = g(t) + s(t) + h(t) + \epsilon_t \quad (7)$$

where y_t represents the observation at time t , $g(t)$ represents the trend component, $s(t)$ represents the seasonality component, $h(t)$ represents the holiday effect component, and ϵ_t is the error term.

4.3. Long Short-Term Memory (LSTM)

Long Short-Term Memory [27] is a deep learning model used for processing and predicting time series data. It is a variant of the Recurrent Neural Network (RNN) [28] specifically designed to address the issues of vanishing and exploding gradients in traditional RNNs. Figure 7 below illustrates the structure of an LSTM unit.

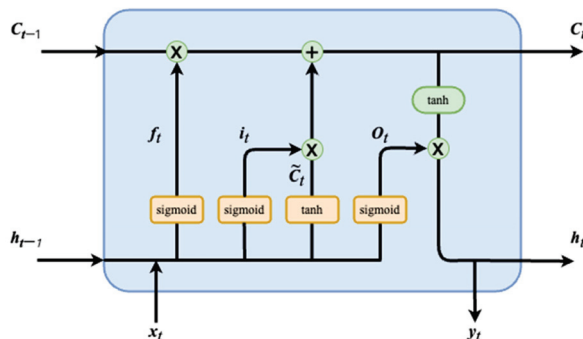


Figure 7. The structure of an LSTM unit.

LSTM has a core structure called a memory cell, which allows the network to capture and store long-term dependencies. The memory cell controls the flow of information through three main gates: Input Gate (i_t), which controls whether information should enter the memory cell; Forget Gate (f_t), which controls which information in the memory cell

should be forgotten; and Output Gate (O_t), which controls the flow of information from the memory cell to the output. The specific steps are described as follows:

(1) Forget: The forget gate applies a *sigmoid* function to determine which information from the previous cell state C_{t-1} needs to be forgotten. The sigmoid function output varies between 0 and 1, where 0 indicates “complete forgetfulness” of previous information, and 1 indicates “complete acceptance” of the previous information at time step $t - 1$. The forget gate vector is obtained from the following equation:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (8)$$

where f_t is the forget gate vector, W_f is the weight vector of the forget gate, h_{t-1} is the output value of the hidden layer from the previous unit, x_t is the input value of the current unit, and b_f is the bias vector of the forget gate.

(2) Input: The input gate has two parts: the information update of the input gate, determined by a sigmoid function, and the creation of a new cell state $\tilde{C}_t$, which is added to the state. The formulas are as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (9)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (10)$$

where i_t is the input gate vector, W_i is the weight vector of the input gate, b_i is the bias vector of the input gate, W_C is the weight vector for creating the new cell state, and b_C is the bias vector for creating the new cell state.

(3) Update: Update the old cell state by combining the output of the forget gate and the output of the input gate to obtain the new cell state. The formula is as follows:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (11)$$

(4) Output: The output gate combines the cell state to produce the final result. The formulas are as follows:

$$O_t = \sigma(W_O \cdot [h_{t-1}, x_t] + b_O) \quad (12)$$

$$h_t = O_t * \tanh(C_t) \quad (13)$$

where O_t is the output gate vector, W_O is the weight vector of the output gate, b_O is the bias vector of the output gate, and h_t is the current unit's hidden state.

LSTM, with its gating mechanism and memory cell structure, enables the network to better handle and capture long-term dependencies, demonstrating excellent performance in modeling time series and sequence data.

4.4. Evaluation Metrics for Prediction Results

This paper employs three metrics to assess the performance of the prediction methods, namely mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) [29].

(1) Mean absolute error (MAE) measures the average absolute difference between predicted values and observed values. MAE directly quantifies the average magnitude of prediction errors without considering their direction. A smaller MAE indicates smaller average prediction errors at the overall level. The formula is as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (14)$$

where y_i is the observed value of the sample, and $\hat{y}_i$ is the predicted value.

(2) Mean squared error (MSE) is a metric that measures the average of the squared differences between predicted and actual values. MSE calculates the squared prediction errors for each sample and averages them, ensuring that negative and positive errors do not cancel each other out. The computation formula is as follows:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (15)$$

(3) Root mean squared error (RMSE) is the square root of MSE and aims to provide a measure in the same units as the original data. RMSE is more sensitive to larger values of prediction errors. The formula is:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (16)$$

5. Experiment

The algorithms used in this paper were executed on a system with an AMD Ryzen 7 3750H processor, 2.30 GHz, and 16 GB of RAM. All algorithms were implemented in the Python programming language. The SARIMA algorithm utilized in this paper is from the Statsmodels package in Python, version 0.12.2. The prophet algorithm is sourced from the open-source time series forecasting tool prophet, developed by Facebook, with a version of 1.1.5. The LSTM algorithm is implemented using the Keras package, version 2.6.0. Except for the prophet algorithm, which runs on Python version 3.8, all other algorithms operate in a Python environment with version 3.6.

5.1. Data Source and Experimental Design

The data for this experiment is sourced from the production data of the Xiamen Port, covering the period from 1 July 2020 to 30 June 2021. The data includes information such as shipping route, vessel name, voyage, container number, unloading port, container attributes, container position in the yard, and entry time. An example of the data is shown in Table 2 below. The dataset only includes loaded container data, and empty containers are not considered in this experiment due to specific business considerations. The dataset encompasses export container data from multiple shipping routes and multiple vessels with various voyages, providing diversity in the data.

Table 2. Container information.

Shipping Route	Vessel Name	Voyage	No.	Pod	ISO Code	Weight (kg)	Position	Record Time
FE2	UASMZ	012W	HLXU3226987	DEHAM	22GP	21,271	3A0090102	2020/7/1 0:05
NEU4-OA	CGFET	0FL6VW1	SDDU2003406	DEHAM	22GP	5062	4D0710402	2020/7/1 0:07
NEU4-OA	CGFET	0FL6VW1	OOLU1018613	DEHAM	22GP	11,249	4D0430601	2020/7/1 0:19
NEU3-OA	EMCTN	0497B	IMTU3080876	DEHAM	22GP	6203	2A0210201	2020/7/1 1:02
NEU4-OA	CGFET	0FL6VW1	VICU0400050	GBSOU	22GP	11,911	4D0490301	2020/7/1 1:04

The experimental design involves the following steps based on the historical export container time series data:

- (1) Data Preprocessing and Analysis: Convert the raw data into time series data, decompose the data based on shipping routes, and obtain historical export container time series data for each shipping route. Randomly cut the preprocessed dataset into data of consistent length and split the dataset into training and testing sets. Analyzing a subset of the data through time series decomposition yields results indicating that the data demonstrates a discernible pattern over time.

- (2) **Training and Prediction:** Utilize time series algorithms to train and predict the data. Three time series algorithms are employed in this paper: the classical SARIMA algorithm, the prophet algorithm with holiday features, and the LSTM algorithm based on deep learning.
- (3) **Evaluation:** Apply the three-evaluation metrics (MAE, MSE, and RMSE) to assess the accuracy and precision of the predictions generated by each algorithm. These metrics provide quantitative measures of the prediction errors.
- (4) **Comparative Analysis:** Compare the performance of the three algorithms (SARIMA, prophet, and LSTM) in terms of prediction accuracy and computational efficiency. Analyze the strengths and weaknesses of each algorithm under different scenarios and identify any specific patterns or trends in the results.
- (5) This experimental design aims to evaluate the performance of the prediction methods using three different time series algorithms.

5.2. Data Preprocessing and Analysis

The original dataset comprises records of export containers from multiple shipping routes. Firstly, we employed the robust Pandas tool to process this dataset. By categorizing and summarizing the data based on shipping routes, we obtained the statistical results illustrated in Figure 8 below, revealing a total of 21 shipping routes. The y-axis represents the volume of gate-in containers for each shipping route.

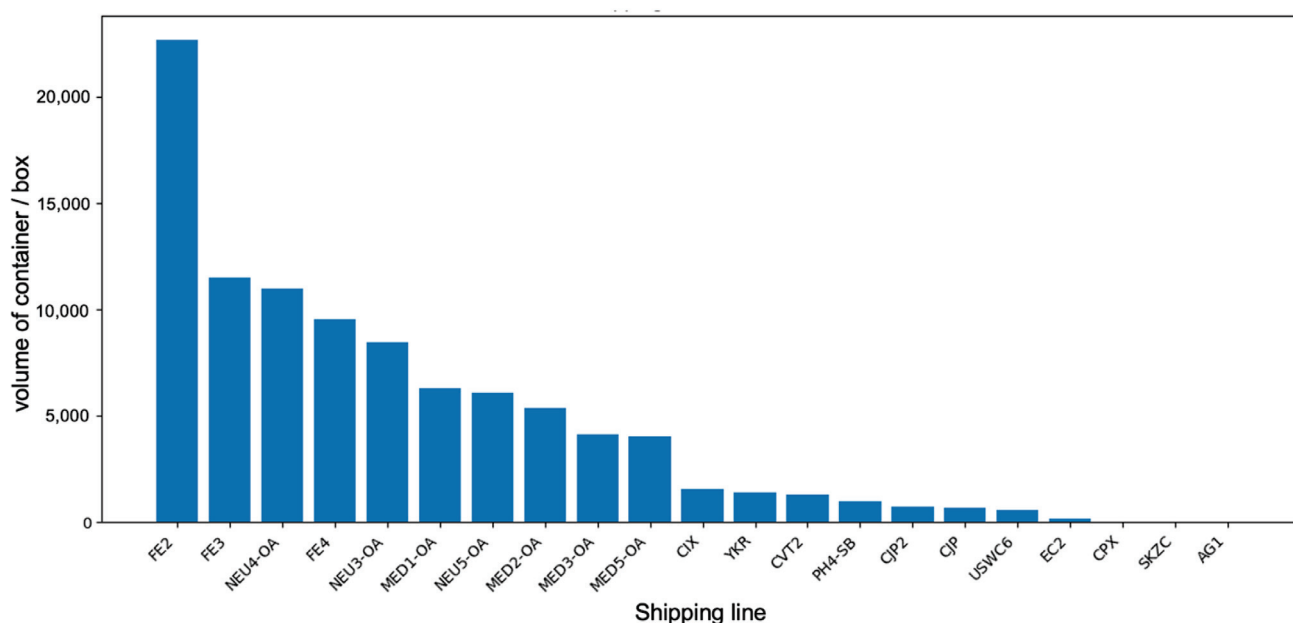


Figure 8. Shipping routes distribution.

Subsequently, we conducted a more detailed analysis for each shipping route, aggregating the data on an hourly basis. Figure 9 below presents the original historical export container records for the FE2 shipping route. This step facilitates a better understanding of the data distribution and trends for each shipping route.

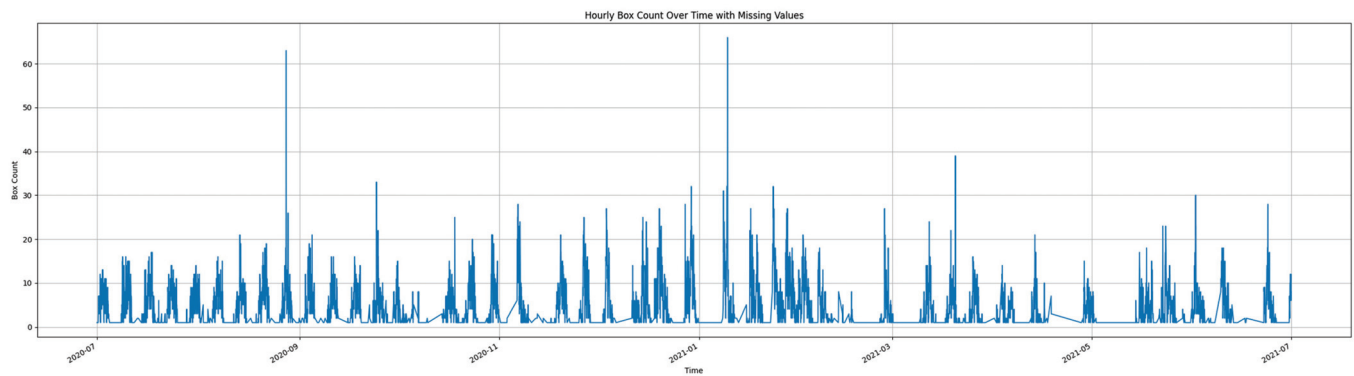


Figure 9. The original historical records of the FE2 shipping route.

Finally, addressing missing values became essential. Given that the experiment operates on an hourly basis, numerous time points were missing in the data. To ensure the continuity of the time series, we filled in the missing time points, assigning them a value of 0. The result is depicted in Figure 10 below.

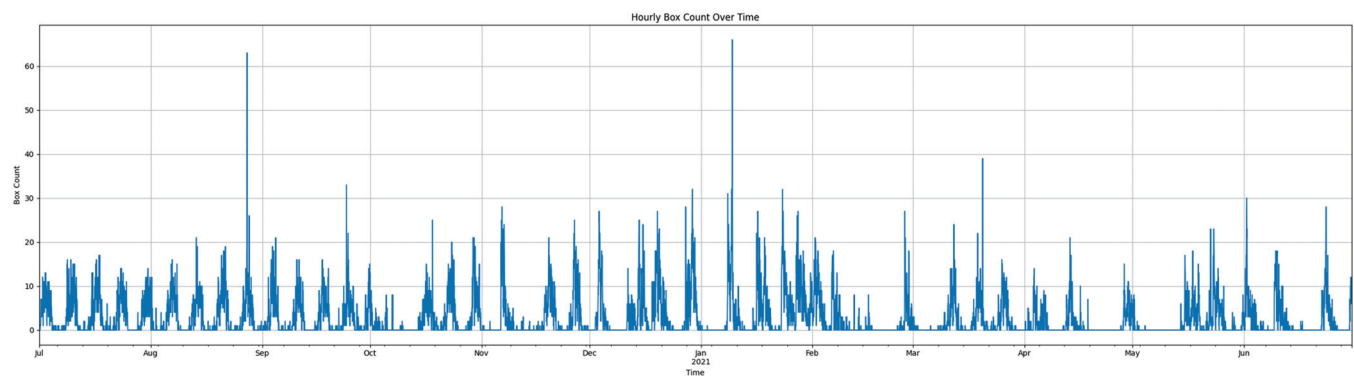


Figure 10. The dataset of the FE2 shipping route.

From Figure 10 above, it is evident that the data exhibits a noticeable periodicity. Excluding a few individual outliers, the majority of peak fluctuations are quite similar. Analyzing a one-month subset of the data through time series decomposition yields Figure 11 below.

During time series decomposition, the typical approach involves breaking down the time series into three main components: trend, seasonality, and residuals. Analyzing the trend subgraph from Figure 11, it is observed that the dataset does not display a clear upward or downward trend but rather demonstrates a fluctuating pattern. The seasonality subgraph indicates a distinct cyclic pattern with a cycle length of approximately 24 h. In the residual subgraph, there is no apparent pattern or trend.

Based on the analysis above, the preprocessed dataset is both stationary and exhibits a clear periodicity. It demonstrates a discernible pattern over time. Therefore, employing time-series-related algorithms for prediction with this dataset is deemed feasible. This series of data preprocessing steps not only ensures the availability and continuity of the data but also provides a reliable foundation for subsequent training and prediction using time series algorithms.

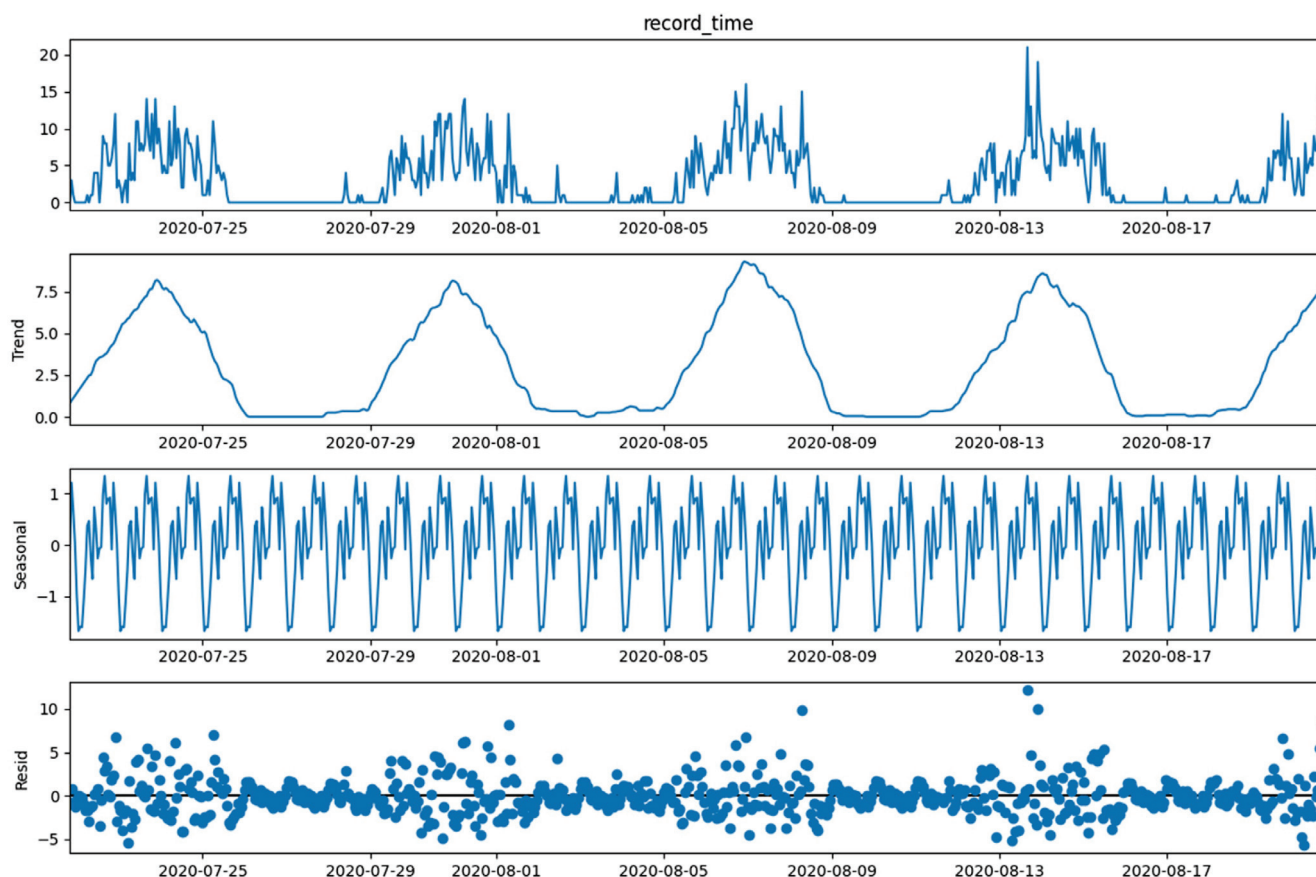


Figure 11. Time series decomposition yields.

5.3. Experimental Results

In this experiment, three algorithms were employed for time series prediction, and the following procedures are illustrated using the FE2 shipping route as an example.

5.3.1. Results and Analysis of ARIMA

The ARIMA time series model served as the baseline for this study. ARIMA is a standard and easily implementable time series model. It only requires the time series data (target values) of the target problem and does not depend on other exogenous variables. SARIMA, a seasonal time series model, performs well in predicting data with periodicity.

The key to the SARIMA model lies in determining the model parameters: autoregressive order (p), differencing order (d), moving average order (q), as well as seasonal autoregressive order (P), seasonal differencing order (D), seasonal moving average order (Q), and seasonal period (s). Figure 12a illustrates the autocorrelation function (ACF) for daily count data. The horizontal axis represents lag periods, while the vertical axis indicates autocorrelation coefficients. It is noticeable that the autocorrelation coefficient is 1 at lag 0, and as the lag period increases, the coefficients gradually decrease, indicating that the temporal dependence of the data weakens over time. Furthermore, the blue shaded area in the figure marks the confidence interval for the autocorrelation coefficients, offering a criterion for assessing the significance of autocorrelation. Overall, this figure reveals the correlational characteristics of the data and provides important insights for further time series analysis. Figure 12b illustrates the Partial Autocorrelation Function (PACF) for daily counts. The x-axis represents the lag times, while the y-axis denotes the autocorrelation coefficients. The plotted blue lines indicate the values of the PACF, with the size of the points reflecting the strength of the correlation. A notable characteristic of the graph is the

envelope line at lag 0, which equals 1. As the lag increases, the autocorrelation coefficients diminish and approach 0, suggesting that the correlation with the current value becomes negligible at greater lags. The shaded blue region highlights the statistical significance range for the PACF values, indicating the intervals within which the PACF is considered statistically significant.

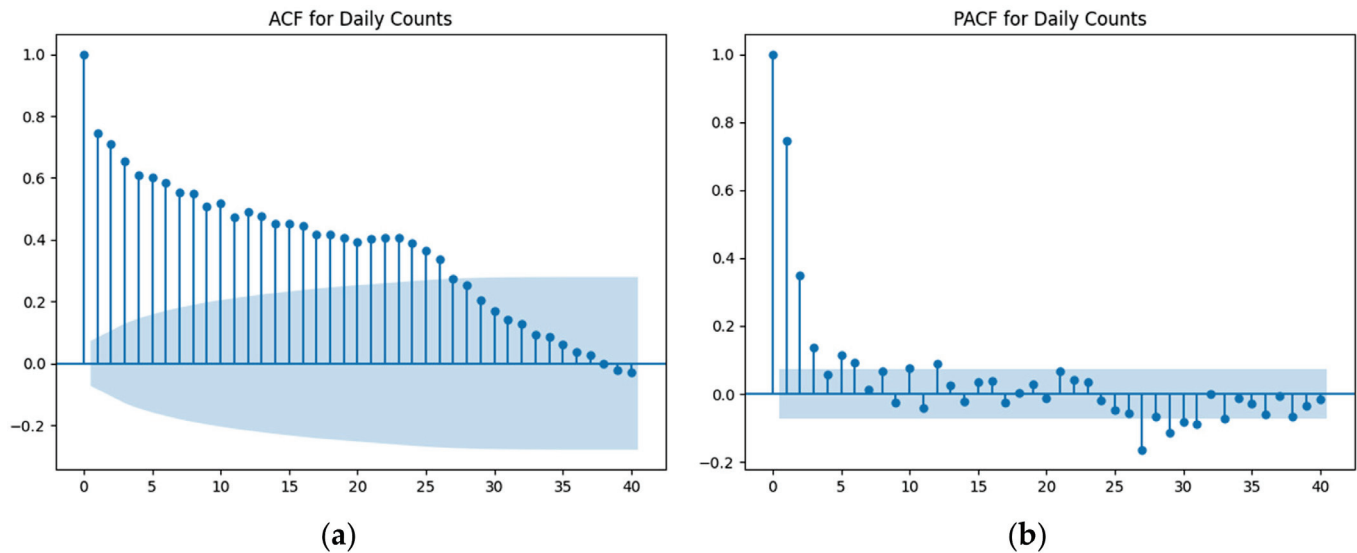


Figure 12. (a) ACF and (b) PACF plots.

Based on ACF and PACF analysis, combined with the analysis in Section 5.2, the experimental research data exhibited a clear periodicity. It is challenging to determine the regression order and moving average order from ACF and PACF alone. Therefore, this study performed AIC ($AIC = 2k - 2\ln(\hat{L})$) and BIC ($BIC = k \ln(n) - 2\ln(\mathcal{L})$) validations [30] on multiple parameter sets. The selected parameters were compared based on the values of AIC and BIC, with the results shown in Table 3 below.

Table 3. Values of AIC and BIC.

Parameters (p, d, q) (P, D, Q, s)	AIC	BIC	Tit	F
(1, 0, 1) $\times$ (1, 1, 1, 24)	2574.195	2595.762	43	2.226
(1, 0, 1) $\times$ (2, 1, 1, 24)	2572.120	2598.001	40	2.222
(1, 0, 1) $\times$ (3, 1, 1, 24)	2569.062	2599.257	47	2.218
(1, 0, 1) $\times$ (4, 1, 1, 24)	2567.895	2602.403	20	2.215
(1, 0, 1) $\times$ (5, 1, 1, 24)	2558.646	2597.468	28	2.205
(1, 0, 1) $\times$ (6, 1, 1, 24)	2547.983	2591.118	29	2.194
(1, 0, 1) $\times$ (7, 1, 1, 24)	2547.456	2594.905	38	2.192
(1, 0, 1) $\times$ (8, 1, 1, 24)	2542.822	2594.585	50	2.186
(1, 0, 1) $\times$ (9, 1, 1, 24)	2544.277	2600.353	50	2.186
(1, 0, 1) $\times$ (8, 1, 2, 24)	2545.863	2601.939	50	2.187
(2, 0, 1) $\times$ (8, 1, 1, 24)	2541.431	2597.507	50	2.183
(3, 0, 1) $\times$ (8, 1, 1, 24)	2542.244	2602.634	50	2.183
(1, 1, 1) $\times$ (8, 1, 1, 24)	2550.593	2602.334	16	2.193
(2, 1, 1) $\times$ (8, 1, 1, 24)	2544.152	2600.204	50	2.186

From Table 3, it can be analyzed that when the parameter values are (2, 0, 1) $\times$ (8, 1, 1, 24), the AIC is minimized, and F (residual sum of squares) is also minimized. Therefore, the chosen parameters for the experimental process are (2, 0, 1) $\times$ (8, 1, 1, 24). The reason for selecting $s = 24$ is that the experimental data's time unit is in hours, and 24 h constitutes one cycle; hence, s is set to 24.

The SARIMA model was trained and predicted, and the diagnostic plots of the model are shown in Figure 13.

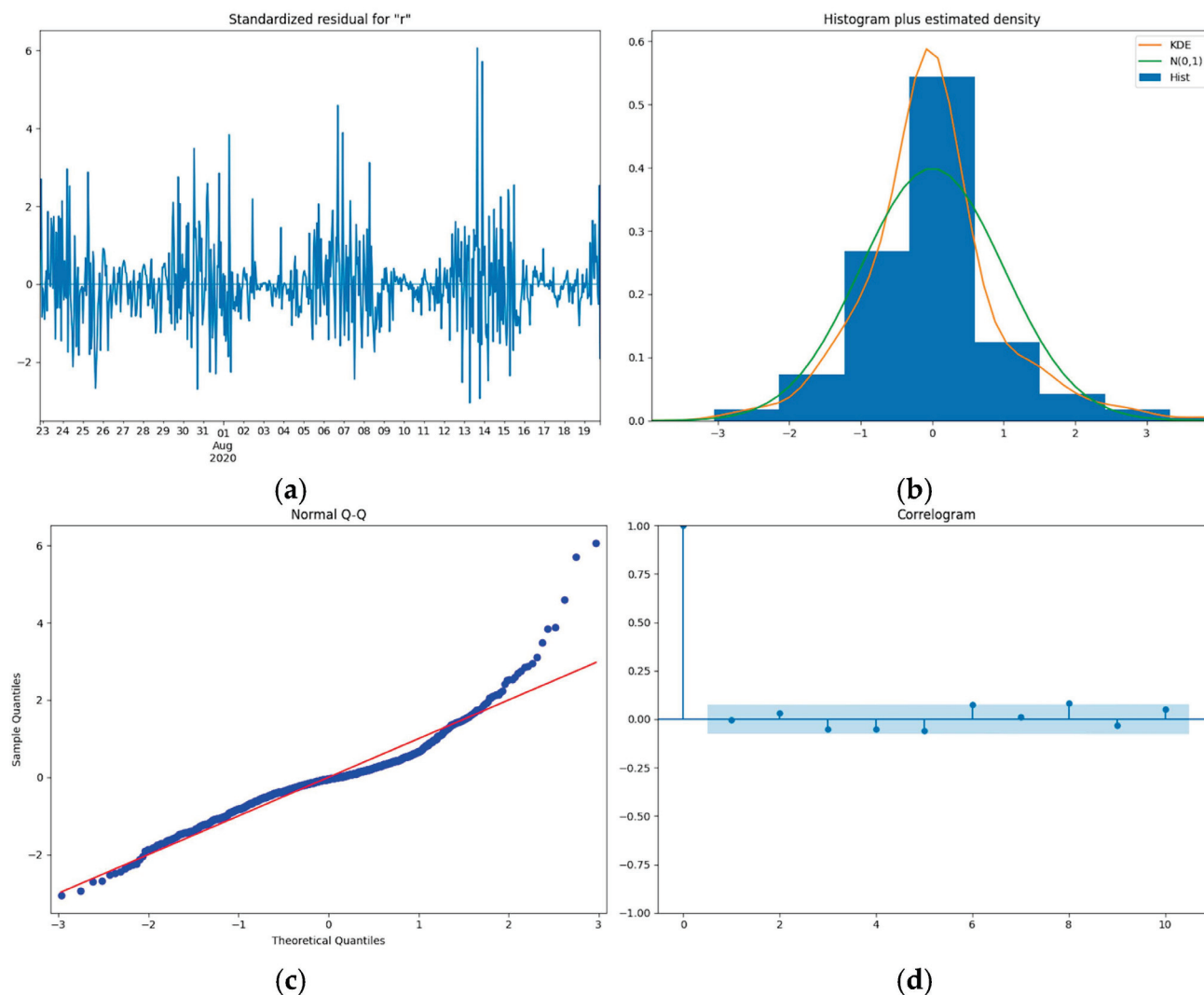


Figure 13. The diagnostic plots of the model.

Figure 13a of the distribution of residuals over time, the residuals are randomly distributed around zero. Figure 13b represents the histogram and kernel density estimate of residuals; the residuals and kernel density are close to a normal distribution. Figure 13c is a normal Q-Q plot. Blue dots represent the sample quantiles of the data being analyzed. Red line is a reference line representing the expected trend if the data follows a normal distribution. The initial data points lie on the line, deviating slightly towards the end which indicates the data is approximately normally distributed. Figure 13d shows the autocorrelation of residuals and indicates no autocorrelation. From the above description, it can be analyzed that the SARIMA model performs well in predicting the data.

Figure 14 represents the overall trend, and Figure 15 represents the predicted part of the trend. From both graphs, it can be observed that the SARIMA model consistently predicts the trend of the data and fits the predicted trend very well. Therefore, it can be concluded that the SARIMA model is feasible for predicting the time series data in this experiment.

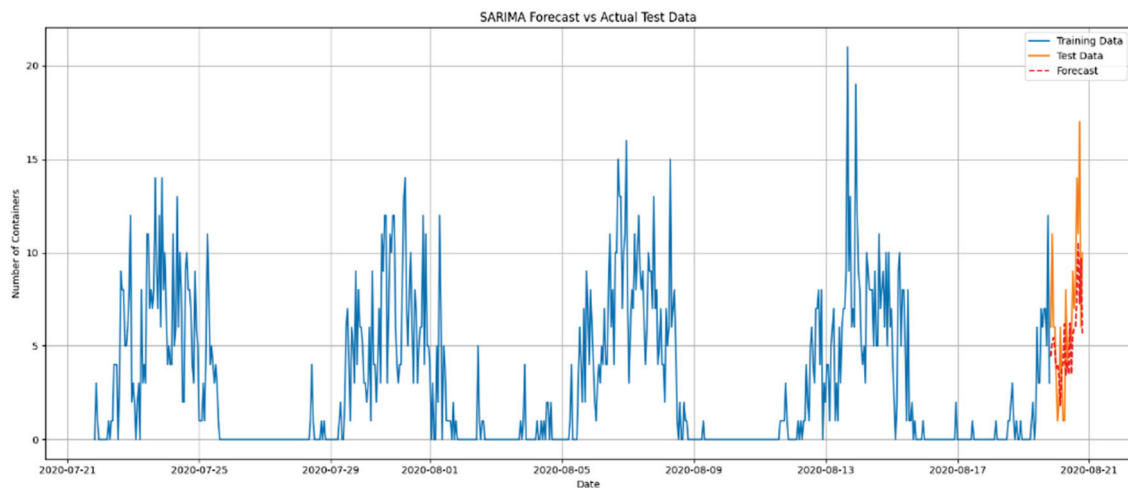


Figure 14. Predicted result of SARIMA for the overall trend.

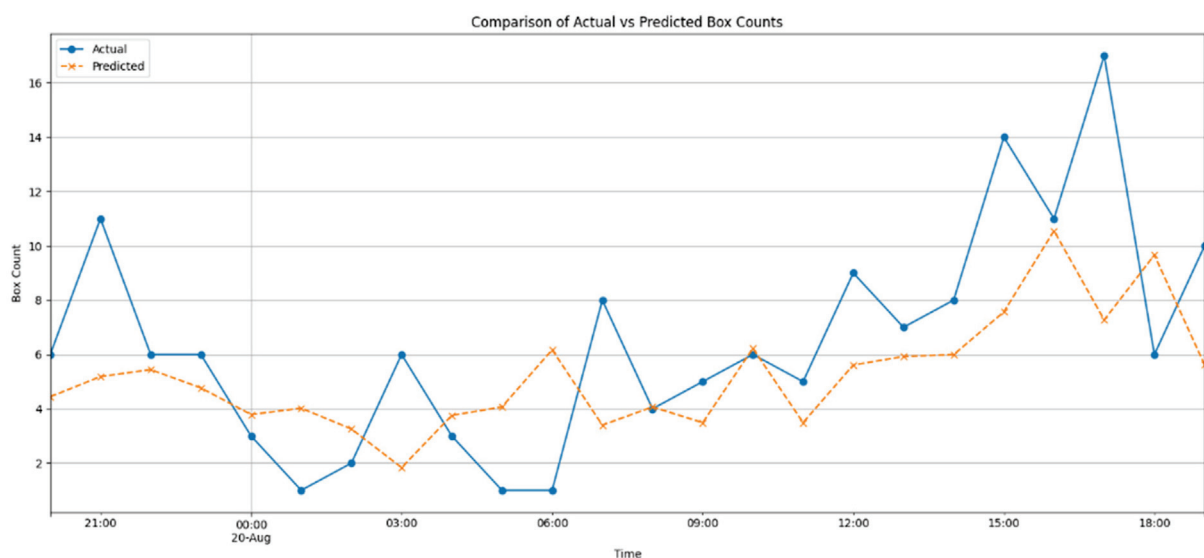


Figure 15. Predicted result of SARIMA for part of trend.

5.3.2. Results and Analysis of Prophet

Compared to the SARIMA model's prediction of time series data, Facebook's proposed prophet model has further improved the periodicity component. The parameters used for prophet in this experiment are shown in Table 4 below.

Table 4. The parameters of the prophet.

Parameters	Value
growth	'logistic'
yearly_seasonality	False
weekly_seasonality	5
daily_seasonality	24
interval_width	0.95
country_holidays	country_name = 'CN'

The Figure 16 below displays the decomposition plot for the prophet model in this experiment. Figure 16a illustrates the overall trend. Solid Line represents the predicted trend component of the time series. Dashed Line represents the confidence interval of the predicted trend. This subfigure indicating a lack of significant fluctuations within the observed period, suggesting stability in the dataset. Figure 16b appears to depict the

seasonal component, characterized by uniformity across the time points, which may imply the absence of seasonal variations in the data. Figure 16c displays weekly periodicity, with the x-axis representing the days of the week and the y-axis denoting a specific measurement, such as event counts or sales figures. The data reveals pronounced peaks during midweek, particularly on Wednesday and Thursday, contrasted by lower values on weekends, indicating a clear weekly cyclical pattern. Figure 16d examines daily periodicity, with the x-axis illustrating the hours of the day. This subplot evidences marked fluctuations over the course of a day, highlighting particular times at which the data exhibits peaks and troughs, thus indicating potential concentrations or activity patterns at specific times throughout the day. From the graph, it can be observed that there is no clear trend in the data, consistent with the analysis results in Section 5.2. Additionally, the experiment used Chinese holidays in the prophet model due to the data's origin from Xiamen Port. However, from the graph, it can be seen that the experiment's data does not exhibit holiday effects. The weekly and daily subgraphs reveal some periodic patterns in the experimental data. The weekly subgraph indicates a peak period every Friday, while the daily subgraph shows relatively fewer container arrivals in the early morning compared to the afternoon and early evening.

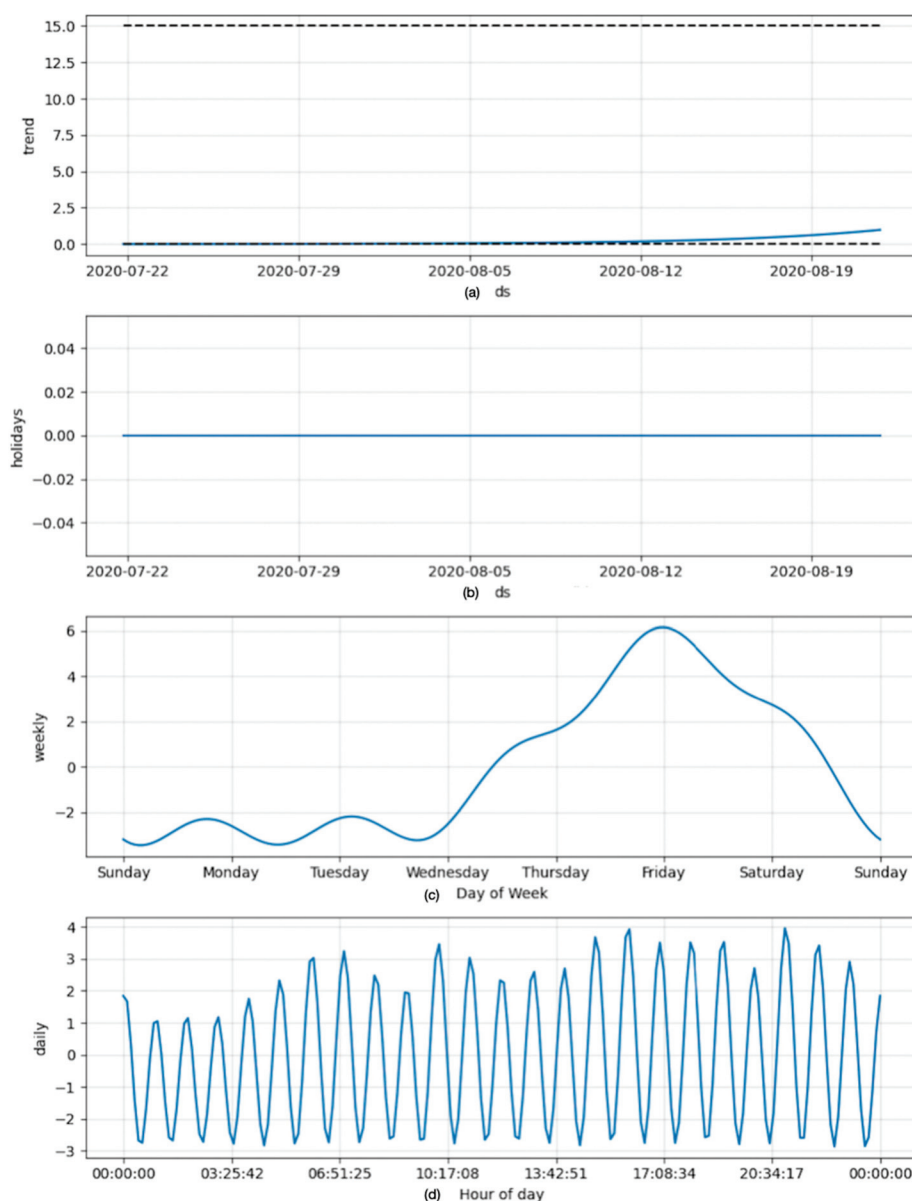


Figure 16. The decomposition plots.

The following figures show the trend predictions based on the prophet model. Figure 17 represents the overall trend, while Figure 18 represents the predicted results for 24 h. The purple shading in both figures represents the 95% confidence intervals around the forecast, indicating the range within which the actual values are expected to lie with a certain level of confidence. From the overall trend graph, it is evident that the predicted trend is consistent with the actual observations. The prophet model effectively predicts the changing trend in the experimental data. However, there are some shortcomings in the prediction results. The model shows certain deviations when predicting some large values and zero values within the data. In the 24 h prediction results graph, the prophet model performs well, with most observed values falling within the predicted confidence interval. The predicted trend is generally consistent with the observed values.

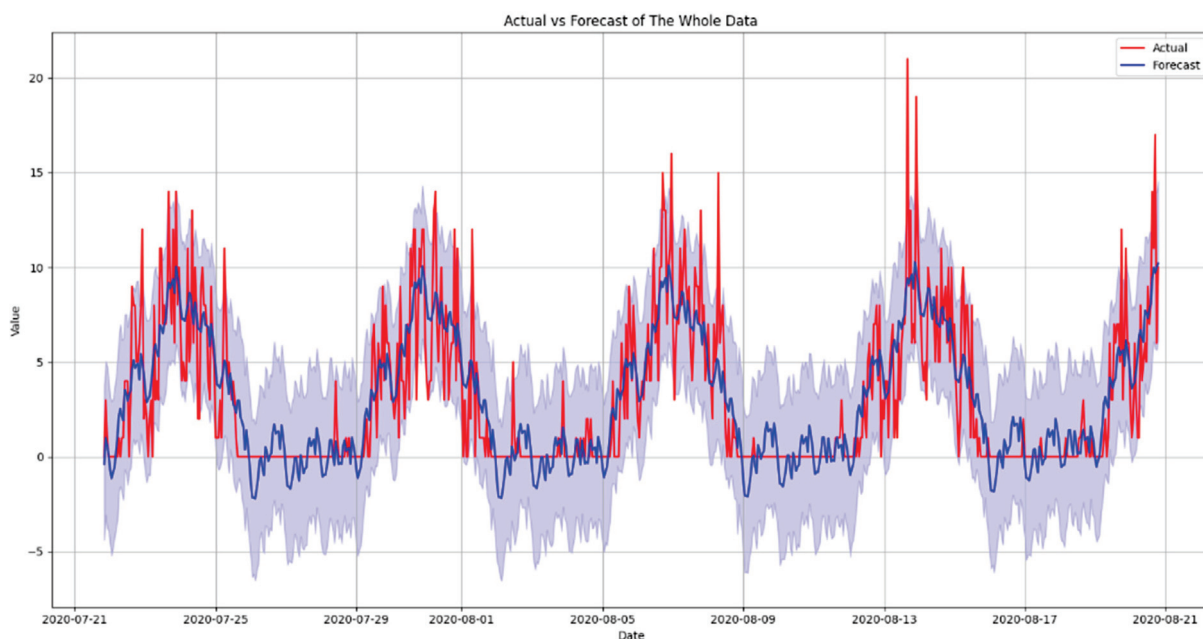


Figure 17. Predicted result of prophet for the overall trend.

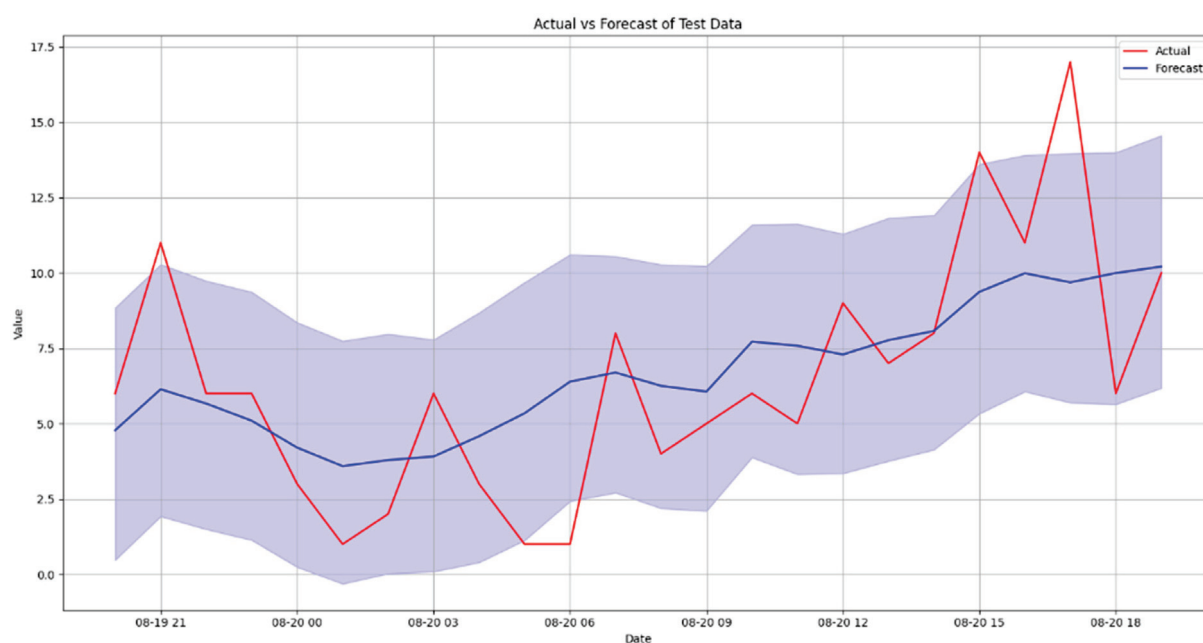


Figure 18. Predicted result of prophet for part of trend.

5.3.3. Results and Analysis of LSTM

Due to the sensitivity of the LSTM model to the values of input data, this experiment preprocesses the data for LSTM by normalizing it. The min-max scaling method is employed for data normalization, as defined by the following Formula (16):

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (17)$$

To enhance the accuracy of LSTM predictions on time series data, the experiment transforms the time series data into input-output pairs suitable for LSTM. The transformation function is defined by the following formula:

$$f(X, lb) = \{(X_i, y_i)\}_{i=1}^{n-lb+1} \quad (18)$$

Here, $X = [x_1, x_2, \dots, x_n]$ is a time series of length n , lb is a given parameter representing the length of the input sequence created, $X_i = [x_i, x_{i+1}, \dots, x_{i+lb-1}]$ is the input sequence vector for the LSTM model, and y_i is the target value corresponding to x_{i+lb} . Through this function, the original time series data is transformed into a time series vector, converting the time series problem into a supervised learning problem, enabling LSTM to better learn information in the sequence and improve the model's ability to predict future values.

The parameters used for LSTM in this experiment are shown in the table below (Table 5). The experimental results are presented in the figures below.

Table 5. The parameters of LSTM.

Parameters	Value
units	50
input_shape (lb, 1)	(24, 1)
Dense	1
loss	mean_squared_error
optimizer	Adam
epochs	20

Figures 19 and 20 show the predictions of LSTM on the training and test datasets. Firstly, from the graphs, it can be analyzed that the LSTM model performs well in predicting the overall trend of the training dataset, accurately capturing the general trend of the data and showing good performance in predicting zero values. Secondly, in the test set predictions, LSTM exhibits high sensitivity to data fluctuations, capturing the changing trends when the data undergoes variations. Whether on the training set or test set, the LSTM model demonstrates excellent prediction results, and the test set results indicate the absence of overfitting. Additionally, LSTM has certain limitations in predicting extreme values, showing a tendency to predict the trend rather than accurately predicting large fluctuations.

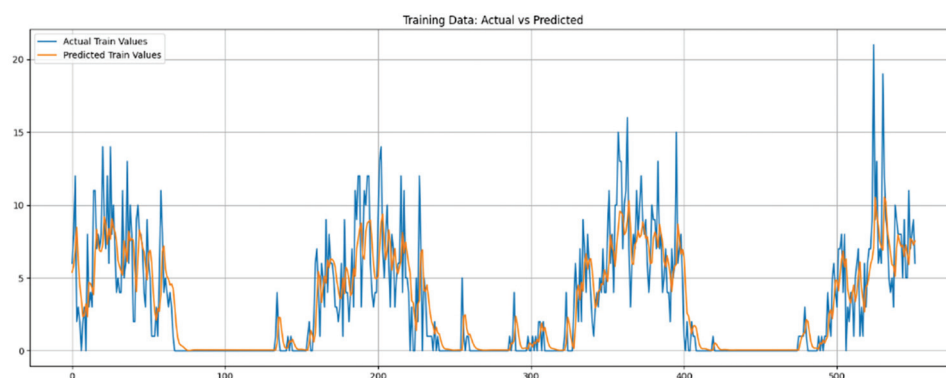


Figure 19. Predicted result of LSTM on the training dataset.

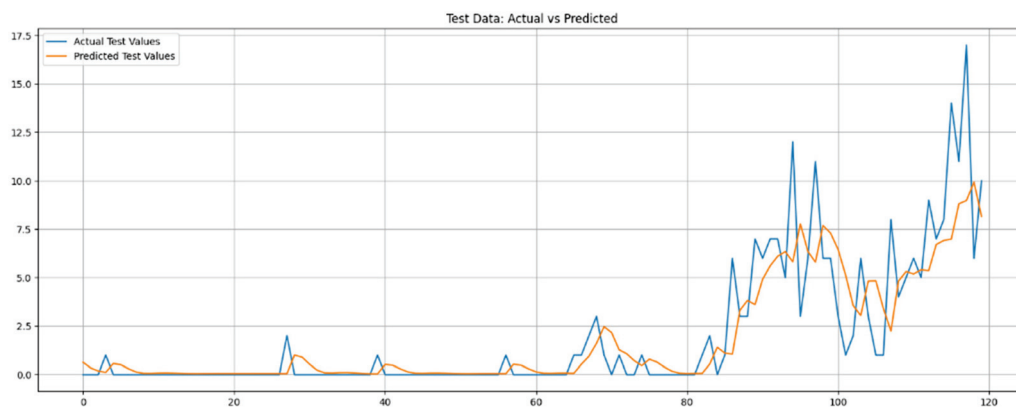


Figure 20. Predicted result of LSTM on the testing dataset.

5.4. Comparative Analysis

In this experiment, three different time series data forecasting methods—ARIMA, prophet, and LSTM—were applied to the same dataset. The performance of these methods was quantitatively analyzed using three commonly used evaluation metrics: mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE). The results are presented in Table 6 below.

Table 6. Evaluation metrics.

Metric	ARIMA	Prophet	LSTM
MAE	2.767	1.944	1.006
MSE	13.113	6.350	3.451
RMSE	3.621	2.443	1.858

Analyzing Table 6, it is evident that LSTM exhibits a significant advantage, particularly in terms of MAE, where its lower value (1.006) indicates that its predicted results are relatively closer to the true values. This reflects the outstanding performance of LSTM in accurately modeling time series data. In comparison, ARIMA and prophet have MAE values of 2.767 and 1.944, respectively, indicating relatively larger errors. Examination of the MSE metric further confirms the excellent performance of LSTM in time series data prediction, with its smaller value (3.451) reflecting better robustness to outliers. In contrast, ARIMA and prophet have MSE values of 13.113 and 6.350, highlighting their relatively higher model complexity. Regarding RMSE, LSTM also excels with a value of 1.858 compared to ARIMA (3.621) and prophet (2.443), emphasizing its sensitivity to prediction errors and accurate capture of trends.

In addition to the quantitative analysis, Table 7 below presents the predicted results of the three algorithms. Firstly, for the ARIMA method, it is observed that its predicted values deviate significantly from the actual values at most time points. Particularly during peak hours, ARIMA fails to accurately capture the fluctuations in the actual observed values. Secondly, the prophet method demonstrates relatively good overall predictive performance. The differences between predicted and actual values are relatively small, with a total sum of 160, close to the actual sum of 156. Although there is slight prediction error during some peak periods, the overall performance remains relatively stable. Lastly, the LSTM method performs exceptionally well in predicting values at various time points, achieving precise predictions at many time points. Notably, during peak hours, LSTM successfully captures the peaks of actual observed values, showcasing its sensitive modeling characteristics for time series data.

Table 7. Results of the three algorithms.

Time	Actual	ARIMA Predicted	Prophet Predicted			LSTM Predicted
			Yhat	Yhat_Lower	Yhat_Upper	
2020/8/14 20:00	6	4.442	4.779	0.454	9.049	5.262
2020/8/14 21:00	11	5.186	6.139	2.019	10.119	5.031
2020/8/14 22:00	6	5.442	5.666	1.455	9.839	6.994
2020/8/14 23:00	6	4.771	5.097	0.797	9.155	6.224
2020/8/15 0:00	3	3.790	4.203	−0.040	8.046	5.595
2020/8/15 1:00	1	4.022	3.589	−0.414	7.717	4.266
2020/8/15 2:00	2	3.257	3.790	−0.054	8.216	2.724
2020/8/15 3:00	6	1.836	3.909	−0.462	8.025	2.363
2020/8/15 4:00	3	3.756	4.581	0.066	8.348	4.113
2020/8/15 5:00	1	4.070	5.342	1.420	9.454	3.868
2020/8/15 6:00	1	6.172	6.391	2.325	10.405	2.529
2020/8/15 7:00	8	3.399	6.695	2.838	10.738	1.650
2020/8/15 8:00	4	4.070	6.252	1.837	10.031	4.393
2020/8/15 9:00	5	3.488	6.061	1.804	10.430	4.436
2020/8/15 10:00	6	6.224	7.717	3.334	11.995	4.386
2020/8/15 11:00	5	3.492	7.584	3.383	11.539	4.759
2020/8/15 12:00	9	5.613	7.292	3.166	11.157	4.641
2020/8/15 13:00	7	5.925	7.768	3.778	12.221	6.019
2020/8/15 14:00	8	5.993	8.073	3.877	12.069	6.031
2020/8/15 15:00	14	7.567	9.371	5.250	13.632	6.230
2020/8/15 16:00	11	10.552	9.988	5.600	14.050	8.252
2020/8/15 17:00	17	7.272	9.686	5.598	13.696	8.154
2020/8/15 18:00	6	9.663	9.994	5.877	14.325	9.583
2020/8/15 19:00	10	5.619	10.208	6.176	14.325	7.255
24 h Sum	156	126	160			125

6. Conclusions and Future Work

In this paper, a novel approach for predicting hourly gate-in container quantities was presented. By analyzing raw data from the Xiamen Container Terminal, periodic characteristics of time series data were identified, and three forecasting methodologies—ARIMA, prophet, and LSTM were employed—to predict the container reception quantities. Experimental results demonstrate that the LSTM model outperformed both ARIMA and prophet in predictive accuracy, showcasing its effectiveness in capturing complex relationships within time series data.

The following innovations are highlighted in this study:

1. This research introduces a decomposed ensemble framework tailored for predicting hourly container reception quantities. By segmenting the total container quantities based on distinct shipping routes, the framework allows for customized forecasting, addressing the inherent variability in port operations. This approach represents a significant enhancement over traditional forecasting methods that often apply a generalized model.
2. The results obtained from this study suggest that the LSTM model may offer improved predictive performance when compared to classical forecasting methods. Its capability to capture underlying trends, seasonal fluctuations, and extreme values appears to enhance the reliability of forecasts, which could be beneficial for terminal operators in their operational planning and decision-making processes. Further research may be warranted to explore these findings in diverse operational contexts and evaluate the generalizability of LSTM's advantages in various forecasting scenarios.
3. This study enhances the accessibility of advanced predictive modeling techniques, such as LSTM, for terminal operators without specialized AI knowledge. By utilizing advanced machine learning frameworks, the implementation of these models becomes more user-friendly, enabling terminal personnel to effectively integrate robust forecasting methods into their daily operations. The predictive outputs from LSTM can

inform better stacking space allocation and gantry crane scheduling, demonstrating the practical applicability of AI-driven insights in improving operational efficiency.

Future research directions may explore the integration of pre-recorded data to enhance predicted gate reception quantities, especially considering the increasing trend toward requiring advance appointments for entry at terminals. While the reliability of such pre-recorded data may vary, it holds potential for improving the accuracy of forecasting methodologies when combined with historical data.

Additionally, there are numerous avenues for extending the application of our research findings. For example, the predicted container reception quantities can facilitate more effective stacking space allocation, allowing terminal operators to devise optimized storage plans that better match anticipated demand. This proactive approach can help minimize congestion and improve space utilization within the terminal.

Overall, the conversion of AI-driven predictive insights into practical tools empowers terminal operators to make informed decisions, thereby improving operational efficiency and responsiveness to dynamic market conditions. Future studies could further evaluate these applications and explore additional frameworks that integrate AI technologies into routine terminal operations.

Author Contributions: Conceptualization, Y.S.; methodology, B.X.; software, B.X.; validation, N.Z.; formal analysis, H.H.; investigation, Y.W.; resources, N.Z.; data curation, Y.W.; writing—original draft, Y.S.; visualization, Z.Y.; supervision, H.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: Author Yansong Wu was employed by the company Xiamen Port Holding Group Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Belcore, O.M.; Di Gangi, M.; Polimeni, A. Connected Vehicles and Digital Infrastructures: A Framework for Assessing the Port Efficiency. *Sustainability* **2023**, *15*, 8168. [CrossRef]
2. Wang, W.; Qin, W.; Shen, S. Short-term Prediction of Outbound Container Arrival Quantity Based on KNN with Deep Learning in Container Terminals. In Proceedings of the 2022 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Kuala Lumpur, Malaysia, 7–10 December 2022.
3. Voß, S.; Stahlbock, R.; Steenken, D. Container terminal operation and operations research—A classification and literature review. *OR Spectr.* **2004**, *26*, 3–49. [CrossRef]
4. Nishimura, E.; Imai, A.; Papadimitriou, S. Berth allocation planning in the public berth system by genetic algorithms. *Eur. J. Oper. Res.* **2001**, *131*, 282–292. [CrossRef]
5. Zhang, C.; Liu, J.; Wan, Y.-W.; Murty, K.G.; Linn, R.J. Storage space allocation in container terminals. *Transp. Res. Part B Methodol.* **2003**, *37*, 883–903. [CrossRef]
6. Chen, S.-H.; Chen, J.-N. Forecasting container throughputs at ports using genetic programming. *Expert Syst. Appl.* **2010**, *37*, 2054–2058. [CrossRef]
7. Schulze, P.M.; Prinz, A. Forecasting container transshipment in Germany. *Appl. Econ.* **2009**, *41*, 2809–2815. [CrossRef]
8. Yang, C.-H.; Chang, P.-Y. Forecasting the Demand for Container Throughput Using a Mixed-Precision Neural Architecture Based on CNN-LSTM. *Mathematics* **2020**, *8*, 1784. [CrossRef]
9. Jin, J.; Ma, M.; Jin, H.; Cui, T.; Bai, R. Container terminal daily gate in and gate out forecasting using machine learning methods. *Transp. Policy* **2023**, *132*, 163–174. [CrossRef]

10. Gao, Y.; Chang, D.; Chen, C.-H.; Fang, T. Deep Learning with Long Short-Term Memory Recurrent Neural Network for Daily Container Volumes of Storage Yard Predictions in Port. In Proceedings of the 2018 International Conference on Cyberworlds (CW), Singapore, 3–5 October 2018.
11. Gao, Y.; Chang, D.; Fang, T.; Fan, Y. The Daily Container Volumes Prediction of Storage Yard in Port with Long Short-Term Memory Recurrent Neural Network. *J. Adv. Transp.* **2019**, *2019*, 5764602. [CrossRef]
12. Wang, Y.; Shi, G.; Sun, X. A Forecast Model of the Number of Containers for Containership Voyage. *Algorithms* **2018**, *11*, 193. [CrossRef]
13. Nadi, A.; Sharma, S.; Snelder, M.; Bakri, T.; van Lint, H.; Tavasszy, L. Short-term prediction of outbound truck traffic from the exchange of information in logistics hubs: A case study for the port of Rotterdam. *Transp. Res. Part C Emerg. Technol.* **2021**, *127*, 103111. [CrossRef]
14. Twrdy, E.; Batista, M. Modeling of container throughput in Northern Adriatic ports over the period 1990–2013. *J. Transp. Geogr.* **2016**, *52*, 131–142. [CrossRef]
15. Rashed, Y.; Meersman, H.; Sys, C.; Van de Voorde, E.; Vanelslender, T. A combined approach to forecast container throughput demand: Scenarios for the Hamburg-Le Havre range of ports. *Transp. Res. Part A Policy Pract.* **2018**, *117*, 127–141. [CrossRef]
16. Cuong, T.N.; Kim, H.-S.; You, S.-S.; Nguyen, D.A. Seaport throughput forecasting and post COVID-19 recovery policy by using effective decision-making strategy: A case study of Vietnam ports. *Comput. Ind. Eng.* **2022**, *168*, 108102. [CrossRef] [PubMed]
17. Tian, X.; Liu, L.; Lai, K.K.; Wang, S. Analysis and forecasting of port logistics using TEI@I methodology. *Transp. Plan. Technol.* **2013**, *36*, 685–702. [CrossRef]
18. Pang, G.; Gebka, B. Forecasting container throughput using aggregate or terminal-specific data? The case of Tanjung Priok Port, Indonesia. *Int. J. Prod. Res.* **2017**, *55*, 2454–2469. [CrossRef]
19. Niu, M.; Hu, Y.; Sun, S.; Liu, Y. A novel hybrid decomposition-ensemble model based on VMD and HGWO for container throughput forecasting. *Appl. Math. Model.* **2018**, *57*, 163–178. [CrossRef]
20. Sanguri, K.; Shankar, S.; Punia, S.; Patra, S. Hierarchical container throughput forecasting: The value of coherent forecasts in the management of ports operations. *Comput. Ind. Eng.* **2022**, *173*, 108651. [CrossRef]
21. Shankar, S.; Ilavarasan, P.V.; Punia, S.; Singh, S.P. Forecasting container throughput with long short-term memory networks. *Ind. Manag. Data Syst.* **2019**, *120*, 425–441. [CrossRef]
22. Lee, E.; Kim, D.; Bae, H. Container Volume Prediction Using Time-Series Decomposition with a Long Short-Term Memory Models. *Appl. Sci.* **2021**, *11*, 8995. [CrossRef]
23. Shen, S.; Cao, J.; Yang, Y.; Guo, Y. A kNN Based Voyage's Containers' Entering Time Distribution Prediction System. In Proceedings of the 2021 IEEE International Conference on Progress in Informatics and Computing (PIC), Shanghai, China, 17–19 December 2021.
24. Branch, M.A.E.; Branch, A.E. *Elements of Shipping*; Routledge: London, UK, 2014.
25. Wulff, S.S. Time Series Analysis: Forecasting and Control, 5th edition. *J. Qual. Technol.* **2017**, *49*, 418–419. [CrossRef]
26. Taylor, S.J.; Letham, B. Forecasting at Scale. *Am. Stat.* **2018**, *72*, 37–45. [CrossRef]
27. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
28. Deng, L. *Deep Learning: Methods and Applications*; Now Publishers: Hanover, MA, USA, 2014. [CrossRef]
29. Du, P.; Wang, J.; Yang, W.; Niu, T. A novel hybrid model for short-term wind power forecasting. *Appl. Soft Comput.* **2019**, *80*, 93–106. [CrossRef]
30. Balaguer, E.; Palomares, A.; Soria, E.; Martín-Guerrero, J.D. Predicting service request in support centers based on nonlinear dynamics, ARMA modeling and neural networks. *Expert Syst. Appl.* **2008**, *34*, 665–672. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Journal of Marine Science and Engineering Editorial Office

E-mail: jmse@mdpi.com
www.mdpi.com/journal/jmse



Disclaimer/Publisher's Note: The title and front matter of this reprint are at the discretion of the Guest Editors. The publisher is not responsible for their content or any associated concerns. The statements, opinions and data contained in all individual articles are solely those of the individual Editors and contributors and not of MDPI. MDPI disclaims responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-8206-9