MACHINE LEARNING IN PORT OPERATIONS: EVALUATION, PREDICTION, AND OPTIMIZATION

CHU ZHONG

PhD

The Hong Kong Polytechnic University

2025

The Hong Kong Polytechnic University

Department of Logistics and Maritime Studies

Machine Learning in Port Operations: Evaluation, Prediction, and Optimization

CHU Zhong

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

March 2025

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____ (Signed)

CHU Zhong (Name of student)

Abstract

Port operations are critical to global trade but face challenges like surging traffic and dynamic coordination needs. Addressing these complexities requires innovative analytical approaches that can effectively measure performance, forecast vessel activities, and optimize resource allocation. Inspired by the machine learning (ML) framework of evaluation, prediction, and optimization, this thesis explores the application of ML and operations research (OR) in port operations, with a focus on vessel arrival and departure. In the evaluation phase, two novel data fusion approaches are introduced to quantify the operational status of vessel movements, providing a more comprehensive assessment of arrival and departure dynamics. The first integrating the vessel estimated time of arrival (ETA), actual time of arrival (ATA), and the corresponding data from the Automatic Identification System (AIS) to quantify vessel arrival time (VAT) delays. The analysis reveals that as the vessels approach their destination port, their reported ETA becomes increasingly accurate in both spatial and temporal dimensions. The second study integrates the vessel's estimated departure time (EDT), actual departure time (ADT), and berth entry/exit timestamps to quantify vessel turnaround time (VTT) and service time (VST). A quantitative analysis is conducted to evaluate the impact of COVID-19 on port operations, with Hong Kong Port as a case study. The findings indicate that COVID-19 and its restrictions worsened vessel arrival delays and extended turnaround time, reducing port efficiency.

The prediction phase focuses on estimating VAT, VTT, and VST. Using the established evaluation framework, relevant datasets are constructed to enable time prediction via tree-based models. This thesis is the first to integrate vessel-reported ETA and AIS data for VAT prediction of oceangoing vessels. Compared to vessel-

reported ETA, the proposed approach lowers the mean absolute error (MAE) from 6.84 to 3.11 hours, at a 54.53% reduction. For inland waterway shipping, vessel traffic flow data and the A-Star algorithm are integrated to account for river transport characteristics and estimate the remaining sailing distance for VAT prediction. Results indicate a significant improvement, reducing MAE from 17.06 to 3.49 hours: a 79.54% reduction. For VTT prediction, the proposed model enhances accuracy, lowering MAE from 5.12 to 3.94 hours compared to vessel-reported values. Likewise, for VST prediction, MAE decreases from 4.54 to 3.19 hours.

The optimization phase examines the impact of integrating VAT predictions into berth allocation planning (BAP). Leveraging the predicted VAT, a two-stage prediction-then-optimization framework is proposed. In the first stage, a VAT prediction model improves the accuracy of VAT estimates. In the second stage, the predicted VAT is incorporated into the BAP model to optimize berth scheduling. The effectiveness of VAT-based scheduling is evaluated by comparing a BAP model using predicted VAT with another based on vessel-reported ETA in both discrete and continuous berth settings. In a discrete berth scenario with 12 vessel arrivals, VAT-based scheduling reduces additional BAP costs by 64% and vessel waiting time by 73% compared to ETA-based scheduling. In a continuous berth setting, VAT-based scheduling reduces additional BAP costs by 43% and vessel waiting time by 35%. These findings highlight the effectiveness of VAT-based scheduling in improving berth allocation, reducing vessel waiting time, and optimizing resource utilization.

By systematically incorporating data-driven insights into decision-making, this study highlights the significant potential of AI-powered port management in optimizing daily port operations through vessel arrival/departure prediction models and dynamic berth scheduling optimization, advances maritime digitalization via AI-enhanced terminal operating systems with real-time nautical data integration, and accelerates decarbonization efforts through emission-aware vessel sequencing algorithms and predictive shore power allocation.

Keywords: Maritime transport; Port operation; Machine learning; Data-driven approach; Vessel arrival time prediction

Publications and Working Papers

Arising from the Thesis

Lei, J., Z, Chu*, Y. Wu, and C. Liu (2024). "Predicting vessel arrival times on inland waterways: A tree-based stacking approach". In: *Ocean Engineering* 294, p. 116838.

Peter, W., Z, Chu*, and S. Frederik (2025). *Predicting Vessel Arrival Time in Inland Waterways: Integrating ETA, AIS, and Traffic Flow Data*. Under review.

Yan, R., Z, Chu, L. Wu, and S. Wang (2024). "Predicting vessel service time: A data-driven approach". In: *Advanced Engineering Informatics* 62, p. 102718.

Z, Chu, R. Yan, and S. Wang (2024a). "Are vessel arrival and port operations affected by COVID-19? Evidence from the Hong Kong port". In: *Transport Policy* 154, pp. 157–181.

Z, Chu, R. Yan, and S. Wang (2024b). "Evaluation and prediction of punctuality of vessel arrival at port: a case study of Hong Kong". In: *Maritime Policy & Management* 51.6, pp. 1096–1124.

Z, Chu, R. Yan, and S. Wang (2024c). "Vessel Arrival Time to Port Prediction via a Stacked Ensemble Approach: Fusing Port Call Records and AIS Data". In: *Transportation Research Part C: Emerging Technologies*. Accept and in press.

Z, Chu, R. Yan, and S. Wang (2024d). "Vessel turnaround time prediction: A machine learning approach". In: *Ocean & Coastal Management* 249, p. 107021.

Z, Chu, R. Yan, and S. Wang (2025). *Enhancing Berth Allocation Efficiency Through Vessel Arrival Time Prediction*. In submission.

Acknowledgements

First and foremost, I am deeply grateful to my supervisor, Professor Shuaian (Hans) Wang, for his invaluable guidance, insightful feedback, and unwavering support throughout my four years of PhD study. At the beginning of my research journey, Professor Wang provided me with an excellent research direction, which has remained the core focus of my work over these years. His dedication to research has been a great source of inspiration to me. I deeply admire his meticulous attention to detail, rigorous approach to academic writing, and remarkable efficiency. His prompt responses to emails and messages have always been something I strive to emulate. He has also supported me in attending conferences and participating in the Attachment exchange program during my PhD studies.

Secondly, I would like to express my heartfelt gratitude to my senior and co-supervisor, Dr. Ran (Angel) Yan. Our connection began five years ago through a chance encounter on Weibo. Before I applied for my PhD, Dr. Yan provided me with a detailed introduction to Professor Wang's research group, which became one of the key reasons I decided to pursue my PhD at PolyU. Throughout these four years, Dr. Yan have been my daily mentor, guiding me

from research direction to academic writing. Dr. Yan support has been instrumental in my transformation from an inexperienced master's student to a soon-to-be PhD graduate. Every project I worked on during my PhD was made possible by Dr. Yan insightful discussions at the proposal stage and Dr. Yan patient proofreading of my papers on Overleaf. I still remember during the first ETA project, Dr. Yan used paper and pen to draw an algorithm flowchart and recorded a video to explain it to me.

I would like to express my sincere gratitude to the Board of Examiners, especially Prof. Guang Xiao, as well as the two external examiners, Prof. Yong Wu and Prof. Baozhuang Niu, for taking the time out of their busy schedules to review my PhD thesis and participate in my oral defense. Their valuable feedback and constructive suggestions have significantly contributed to improving the overall quality of my dissertation. I would also like to extend my heartfelt thanks to Prof. Bing Wu, Dr. Yuanjun Feng, and Dr. Xiaoyang Wei for their support and guidance throughout this process. Your advice four years ago helped me quickly gain an understanding of the maritime transportation field and played a crucial role in my decision to join Professor Wang's research group. My heartfelt thanks to Dr. Frederik Schulte and Peter Wenzel for supporting my exchange at TU Delft and for providing the AIS data from the Port of Rotterdam, which was essential for our research. I am also deeply grateful to Dr. Jinyu Lei for teaching me AIS data processing and tree model coding, and to Dr. Xiaohuan Lv for explaining the background knowledge of BAP. Special thanks to Dr. Lingxiao Wu for your valuable

support in our collaborative work on the VST paper. Besides, I am also grateful to PolyU and the faculty and staff of the Department of Logistics and Maritime Studies for providing me with such an excellent platform for learning, research. I appreciate Ms. Xi Luo for helping me proofreading my COVID-19 paper and Jinglin Li for the assistance during my exchange at Delft. Additionally, I am thankful to Bin Tian and Mr. & Mrs. Sun for their support throughout my PhD journey.

Finally, I dedicate this thesis with love and gratitude to my parents, Prof. Xiumin Chu and Ms. Xinyu Liu, and my girlfriend, Wenyang Zhang. To my parents, thank you for your endless and unconditional love, no matter where I am or how old I become. I am deeply grateful for the opportunity you provided me to study abroad since my undergraduate years. Mom, your patience and companionship throughout my early education and your encouragement during my PhD journey have meant the world to me. Dad, as a PhD graduate and a professor, you have always been my role model. Your dedication to academia has inspired me and set a benchmark for my own aspirations. To my girlfriend, Wenyang Zhang, five years of a long-distance relationship has not been easy, but your unwavering support and companionship during my PhD have been invaluable. Now that I am graduating, let's get married!

To all of you, I truly appreciate your guidance and help along the way.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Maritime transport stands as an indispensable pillar of global trade, serving as the backbone of globalization (UNCTAD 2025). Over 90% of worldwide trade relies on sea transport, providing a cost-effective and streamlined avenue for international commerce (UNCTAD 2025). Ports, as pivotal nodes within maritime transportation and the global supply chain, play a vital role in facilitating the seamless exchange of goods on an international scale. Port operations are planned daily to ensure efficient cargo handling, optimize vessel scheduling, and maintain smooth logistics flow.

One crucial challenge faced by terminal operators in daily port operations is the uncertainty surrounding vessel arrivals and departures. Vessels typically update their estimated time of arrival (ETA) one or several days before approaching the port; however, these estimates often differ significantly from the vessel's actual time of arrival (ATA). Similarly, during routine port operations, vessels pro-

vide an estimated departure time (EDT) before leaving the berth, but the actual departure time (ADT) frequently deviates due to unpredictable factors such as port inefficiencies and congestion. Unforeseen weather and sea conditions, along with port congestion, contribute to discrepancies between ETA and ATA, as well as EDT and ADT. These disruptions ultimately affect berth allocation planning (BAP) and other port operations, leading to reduced overall port efficiency. Managing these discrepancies is crucial for maintaining the reliability of vessel operations. According to a SeaIntelligence report, liner services are regarded as reliable if vessel delays are within 24 hours of the scheduled ETA, with global container schedule reliability averaging 51.6% for on-time arrivals as of March 2024 (Jasmina, Ovcina Mandra 2024).

However, despite the critical role of predictive accuracy in port scheduling, traditional approaches to ETA/EDT management remain largely reactive and experience-driven. Manual adjustments based on historical averages or deterministic models often fail to account for the dynamic interplay of maritime variables in daily port operations, such as cascading delays from upstream ports, real-time berth occupancy conflicts, and stochastic cargo handling rates (Filom, Amiri, and Razavi 2022). Yet, this persistent gap presents a significant opportunity to leverage machine learning (ML) for data-driven decision-making, enabling ports to dynamically adapt to operational uncertainties and optimize resource allocation in real-time (Yan, S. Wang, Zhen, and Laporte 2021).

Motivated by the inherent workflow of ML (Z.-H. Zhou 2021),

which consists of three core phases: evaluation (processing then quality assessment of training/testing datasets), prediction (constructing predictive models), and optimization (tuning model parameters to achieve optimal performance on unseen data), this study introduces a structured evaluation, prediction, and optimization framework to empower port operations through ML.

In the evaluation phase, the framework assesses and quantifies vessel arrival and departure patterns, ensuring data reliability and completeness. Additionally, it examines the impact of disruptions, such as those caused by the Coronavirus Disease 2019 (COVID-19) pandemic, on port operations, including delays, congestion, and shifts in shipping patterns. The prediction phase focuses on forecasting vessel arrival and departure time, enabling more accurate time predictions for port scheduling. Finally, the optimization phase leverages vessel arrival time (VAT) predictions to improve the following berth allocation, enhancing port operational efficiency.

1.2 Thesis Outline

The remainder of the thesis is organized as follows. Chapter 2 provides a summary of current research on VAT, vessel turnaround time (VTT) and vessel service time (VST) evaluation and prediction, the impact of COVID-19 on port operations, and data-driven approaches to the BAP. It also highlights the existing research gaps in these areas. Chapter 3 proposes two methods: one that integrates vessel ETA, ATA, and AIS data for VAT evaluation, and the other that combines ETD and ATD for VTT and VST assessment. It also quanti-

tatively evaluates vessel arrival and departure punctuality and the impact of COVID-19 on port operations. Chapter 4 incorporates scenario-specific characteristics to predict VAT for both ocean-going and inland vessels. It also forecasts VTT and VTT. Chapter 5 proposes a two-stage "prediction-then-optimization" framework to quantify the benefits of VAT prediction in the BAP. The effectiveness of VAT-based scheduling is evaluated by comparing a BAP model that utilizes predicted VAT with one that relies on vessel-reported ETA, considering both discrete and continuous BAP settings.

Chapter 2

Current Approach and Literature Review

2.1 Vessel Arrival Time Prediction

There are a number of studies evaluating the uncertainties in VAT to ports and predicting VAT to help port operators make informed decisions (Yan, S. Wang, Zhen, and Laporte 2021; Filom, Amiri, and Razavi 2022). AIS data provides diverse information on vessel voyages, encompassing static details (e.g., size, name and IMO number), dynamic attributes (e.g., location and heading), and voyage-specific information (e.g., destination and draft). These AIS data can be generated and transmitted at intervals that range from a few seconds to several minutes, providing an extensive and dynamic dataset of vessel movements (Y. Yang, Y. Liu, G. Li, Z. Zhang, and Y. Liu 2024; D. Yang, L. Wu, S. Wang, Jia, and K. X. Li 2019).

Research on VAT prediction can typically be divided into two categories. The first focuses on reconstructing vessel trajectories and uses path finding methods to estimate VAT. The second category in-

volves directly predicting VAT following data engineering. Simply put, the key difference between path-finding and data engineering models for VAT prediction is in their inputs. The prediction model uses multiple vessel trajectory date and a path-finding algorithm to predict VAT. It takes the trajectory of a single vessel voyage as input and predicts one VAT value for that trajectory. In contrast, a model using data engineering to predict VAT extracts features from vessel navigation data as input and predicts multiple VAT values, with one for each vessel navigation record. In the following subsections, we will review related literature based on these two categories.

2.1.1 VAT prediction based on vessel trajectory

VAT prediction based on vessel trajectory involves processing AIS data to reconstruct and forecast vessel voyage trajectories, facilitating more accurate predictions of VAT (Filom, Amiri, and Razavi 2022). Alessandrini, Mazzarella, and Vespe (2018) propose a data-driven path-finding method to predict vessel future voyage route and then estimate VAT. K. Park, Sim, and Bae (2021) utilize reinforcement learning on AIS data to construct vessel voyage trajectories, and then implement Bayesian sampling to estimate vessel ETA. El Mekkaoui, Benabbou, and Berrado (2023) use a deep learning sequence model to predict VAT at a bulk port of Jorf Lasfar in Morocco, considering vessel AIS data and weather features. The model's prediction results show that deep learning sequence models such as long short-term memory (LSTM) model outperforms traditional neural network models, reducing VAT error from 48 hours (from ETA data)

to 4 hours. However, this study does not evaluate the impact of reported ETA on the prediction model and only uses the reported ETA for prediction results comparison. Kolley, Rückert, Kastner, Jahn, and Fischer (2023) employs a K-Nearest Neighbors (KNN) approach to predicting VAT and integrates these predictions into subsequent berth allocation operations. The results indicate that KNN not only forecasts VAT more accurately but also significantly reduces the actual waiting time for vessels during berth operations.

Yoon, D.-H. Kim, Yun, H.-J. Kim, and S. Kim (2023) present a method to predict container VAT at Busan New Port by modeling past voyage routes using historical AIS data. They construct representative paths through spline interpolation and apply location-based criteria to determine the departure and arrival time of the vessel, achieving a reduction in mean absolute error (MAE) of approximately 3 hours compared to traditional methods. Wenzel, Jovanovic, and Schulte (2023) use advanced neural networks to predict VAT for inland waterway vessels in the Netherlands and Germany. The authors uniquely use the A* algorithm to estimate the remaining traveling distance from the vessel's AIS-reported position to its destination, and then incorporate this parameter into the neural network's input. The prediction results indicate that the accuracy of the VAT prediction can be improved by 20 6% on average for short trips compared to the industry standard approach, that is, dividing the remaining distance of the vessel by the current speed of the vessel. X. Zhang, Fu, Xiao, H. Xu, Wei, Koh, Ogawa, and Qin (2024) first apply multivariate kernel density estimation to identify vessel destina-

tion berthing areas from AIS data. Then, they implement a temporal convolutional network to predict VAT at the Singapore Port pilotage area, additionally considering pilotage booking information and meteorological data. The prediction results show an MAE ranging from 4.58 to 4.86 minutes.

There are two limitations in the above studies. The first and most significant limitation is that the prediction models in these studies often require a fixed number of AIS data points for each trajectory as input for the model. To achieve this, the authors usually preprocess the AIS data by interpolating AIS data points to standardize the length of each trajectory. However, in real-world port operations, port authorities receive AIS data from vessels with varying trajectory lengths, which greatly increases the data processing workload and complicates the model application process. The second limitation is that VAT prediction models based on vessel trajectories typically focus on individual vessels. However, in real-world operations, port authorities are required to manage and operate multiple vessels simultaneously. This presents a more complex scenario, as port authorities need to account for the collective behavior of numerous vessels, rather than just focusing on individual trajectories. Therefore, models designed to predict VAT for single vessels may not scale effectively or provide the holistic insights needed for managing port operations at an aggregate level.

2.1.2 VAT prediction via data engineering

VAT prediction through data engineering involves directly predicting VAT by mining AIS data and relevant datasets to form tabular datasets for prediction (Filom, Amiri, and Razavi 2022). Pani, Vanelslander, Fancello, Cannas, et al. (2015) implement logistic regression, classification and regression trees (CART), and random forest (RF) models to predict VAT at the Cagliari and Antwerp ports. J. Yu, G. Tang, Song, X. Yu, Qi, D. Li, and Y. Zhang (2018) employ back-propagation networks, CART, and RF models to predict delays or advances in vessel arrivals using port call data. The study also evaluates the benefits of these predictions for daily berth allocation. However, a significant limitation of these studies is that the authors only use classification methods to predict vessel arrival delay interval, not employing regression models to precisely forecast VAT. Veenstra and Harmelink (2021) introduce several novel ETA prediction evaluation metrics for a more comprehensive assessment. They also conduct a case study on a dataset of port calls at the port of Antwerp in 2019 to compute and interpret these metrics. T. Zhang, J. Yin, X. Wang, and Min (2023) utilize an eXtreme gradient boosting (XGBoost) model to assess how port congestion at the Yangshan port affects vessel in port time. Their findings reveal that congestion can cause vessel port time to vary by up to 50 hours.

In summary, current research on VAT prediction using AIS data primarily focuses on short-distance forecasts for both inland and ocean shipping, such as estimating VAT for the last 30 kilometers to a port or destination. However, this approach does not fully align with the

needs of port authorities, who are more concerned with long-distance predictions, such as forecasting VAT one day or 100 kilometers in advance. These long-range predictions are critical for efficient berth allocation and overall port operations. Moreover, from a data perspective, both port call data and vessel AIS data are crucial for port operations. AIS data provides real-time vessel movement information, with some records also including vessel-reported ETA. In contrast, ATA data, typically recorded by port authorities as part of port call data, serves as the ground truth for VAT and is essential for accurate VAT prediction. However, current research often excludes port call data, with some studies relying on the last recorded AIS data point as the vessel's arrival time, which is impractical in real-world scenarios as it uses posterior data for prior prediction. Other studies attempt to estimate the actual time and location of vessel arrival using methods like kernel density estimation (X. Zhang, Fu, Xiao, H. Xu, Wei, Koh, Ogawa, and Qin 2024), but these estimates often deviate from the actual ATA, resulting in inaccuracies in subsequent VAT predictions.

2.2 Vessel Turnaround and Service time Prediction

VTT and VST are crucial port management performance measurements. A vessel's arrival and departure contain various parts such as the duration of berthing, waiting and servicing time of the vessel and thus reflects the proficiency and efficiency of a port's overall operation. Excessive lengths and uncertainties in VTT or VST could lead to port congestion and untimely scheduling, potentially result-

ing in delays in ship schedules. This, in turn, can increase vessels' operational costs and thus diminish the port's reputation, ultimately leading to considerable economic losses. Minimizing the VTT and VST contributes to the port's capacity to handle more ships within a specific period, thereby enhancing the overall efficiency of the maritime chain. Historically, efforts to optimize VTT and VST in port operations have predominantly focused on strategies for berth scheduling. Such a focus is well-documented in the field of maritime studies, as reflected by the substantial attention it receives in the existing literature (K. H. Kim and Moon 2003; Golias, Saharidis, Boile, Theofanis, and Ierapetritou 2009). Despite this established focus, it is worth noting that in the port industry, one of the world's oldest and most consistent sectors, operations related to berth scheduling and VTT optimization often rely more heavily on expert knowledge and the established berth allocation strategies rather than innovative and data-driven scheduling strategies (Barua, Zou, and Y. Zhou 2020; Brouer, Karsten, and Pisinger 2016; Rodrigues and Agra 2022). Recent innovations in machine learning field have paved the way for developing data-driven methods in port operations management (Yan, S. Wang, Zhen, and Laporte 2021; Filom, Amiri, and Razavi 2022). These approach show considerable potential of reducing prediction inaccuracies regarding vessel arrival and departure time from a port management perspective.

Numerous studies have addressed the issues of uncertainties in the time associated with vessels' activities in ports, aiming to predict time-related factors such as VAT, VTT and departure time at

the port (Yan, S. Wang, Zhen, and Laporte 2021; Filom, Amiri, and Razavi 2022). These predictions can greatly assist port operators in their decision-making for port management. For instance, by accurately predicting VAT, the following berth allocation strategy can be optimized, which in turn enhances the operational efficiency of the port (J. Yu, G. Tang, Song, X. Yu, Qi, D. Li, and Y. Zhang 2018). Compared to the extensive literature on predicting ship ETA data (J. Yu, G. Tang, Song, X. Yu, Qi, D. Li, and Y. Zhang 2018; Yan, S. Wang, Zhen, and Laporte 2021; Filom, Amiri, and Razavi 2022), there are currently only a small number of studies on predicting ship EDT or VST using machine learning Filom, Amiri, and Razavi 2022; Yan, S. Wang, Zhen, and Laporte 2021. This discrepancy can be attributed to several reasons. First, the complexity inherent in VTT prediction caused by factors like the vessel's uncertain arrival time, unexpected port congestion and inefficiency in berth allocation that have an influence on VTT makes its precise predictions challenging. Secondly, due to privacy and security concerns of port data, acquiring detailed port operational data, such as real-time quay crane efficiency or berth occupancy rates, is difficult. The limited access to such crucial data impedes the ability to make accurate predictions on VTT.

Mokhtar and Shah (2006) are among the first to conduct research in the field of VTT evaluation. They employ a linear regression model to analyze the relationship between VTT and port facilities. To be more specific, the study concentrates on two ports in Klang, Malaysia, utilizing vessel call data and port operational data col-

lected in August 2005. The results reveal a strong correlation between VTT, crane allocation and the number of containers loaded and discharged (Mokhtar and Shah 2006). Ducruet and Merk present an overview of the VTT efficiency in world container ports in 1996, 2006, and 2011 (Ducruet and Merk 2013). Their study indicates that Ningbo port in China is the most efficient among the largest ports during the years. Štepec, Martinčič, Klein, Vladušič, and Costa (2020) introduce a machine learning based VTT prediction system, utilizing 11 years of port call data at the Port of Bordeaux. This system shows excellent performance in predicting VTT. When compared to actual VTT, the system's error rates for specific types of cargo vessels dropped below 10% (Štepec, Martinčič, Klein, Vladušič, and Costa 2020). B. Li and Y. He (2020) employ a deep neural network to predict container liner berthing time at a terminal in China using four years of container vessels berth time data. However, the advantage of the proposed model's berth time prediction results over the container berthing time reported by the vessels has not been clearly demonstrated (B. Li and Y. He 2020).

Smith (2021) analyzes factors affecting container vessels' VTT at major U.S. container ports, using vessel automatic identification system (AIS) data and port TEU volumes through statistical and comparative analysis. Results of this study show a remarkably positive correlation between the VTT and the vessel expected cargo volume per call. The research results also indicate that vessels are expected to follow the arrival and departure schedules, which make vessel schedules the primary determinants of the VTT (Smith 2021).

Abreu, Maciel, Alves, Braga, and Pontes (2023) evaluate the primary factors affecting the VTT at the port. Then they develop several classification models using data from the cargo vessel movement data in Brazilian ports in 2018 to predict the VTT. The results show that RF demonstrate promising performance, with accuracy and F1-scores above 73%. The results also suggest that RF has the potential to be applied in real-world port management scenarios (Abreu, Maciel, Alves, Braga, and Pontes 2023). Zhai, Fu, X. F. Yin, H. Xu, and W. Zhang (2022) utilize cargo operation data from tanker terminals and vessel AIS data to predict the VTT at Singapore tanker terminals. By employing linear regression and decomposed distribution methods, the prediction model demonstrates remarkable accuracy at 98.81% when evaluated by vessel historical VTT (Zhai, Fu, X. F. Yin, H. Xu, and W. Zhang 2022).

Based on a comprehensive review of prior studies, it is evident that most of the current research is focused on analyzing factors affecting VTT. The studies related to VTT or VST prediction are predominantly based on classification analysis. However, these classification models predict VTT results for vessels in the form of time intervals, which are less precise compared to predictions provided by regression models that specify a specific time-stamp. There is only one paper that has successfully employed regression machine learning models to accurately predict VTT (Štepec, Martinčič, Klein, Vladušič, and Costa 2020). In that research, the authors proposed a tree-based VTT prediction system based on port call data, vessel generic features and port tidal data. The proposed system can greatly

reduce error in VTT by at least 10% for various types of vessels compared to the original EDT data reported by vessels. The innovative aspects of the paper are primarily focused on the presentation of the first machine learning-based regression system for VTT prediction. Additionally, the author is the first to evaluate the impact of port call data, vessel generic features and tidal information on ships' VTT prediction. However, the authors of this article acknowledge that the model they proposed is validated by data from a relatively small port. They do not account for larger ports with more complex infrastructure and higher traffic volumes, where factors such as traffic congestion and vessel arrival patterns could significantly influence the VTT. Remarkably, even within this paper, current research related to VTT prediction has not taken into account the EDT data reported by vessels in predicting their VTT. Furthermore, there is a noticeable lack of attention paid to the influence of vessel generic features (such as vessel length and beam), vessels' historical performance at port (e.g., historical average VTT) and berths' historical operation performance (e.g., historical delays of the ships operated at the berth in terms of their predicted turnaround time) when these factors having a tangible influence on VTT.

Given the existing research gaps, this study develop regression models to predict both container vessels' VTT and VST at the HKP, which is one of the busiest ports in the world. Moreover, more comprehensive data are included to construct the prediction model, including vessel generic features, historical delay data of both vessels and berths and other relevant factors. Therefore, it can be expected

that our research can generate more precise predictions on the VTT and VST in a regression way and enhance the understanding of the intricate factors affecting VTT.

2.3 COVID-19 impact on Port Operations

Stated as a "black swan effect", the emergence of the COVID-19 has critically interrupted the global logistics and economy (C. Zhou, Zhu, Bell, Lee, and Chew 2022). Extensive research has been carried out to examine the repercussions of the COVID-19 pandemic on the maritime field, as shipping acts as a crucial role in ensuring the smooth functioning of the global supply chain, especially during the pandemic (L. Xu, S. Yang, J. Chen, and J. Shi 2021). These studies can be categorized based on the perspectives of individual ports, local areas, or the region and country as a whole. Xu et al. (2021), from a local area perspective, propose a regression method with data from 14 main ports in China to assess the impact of the COVID-19 on port operations and performances (L. Xu, S. Yang, J. Chen, and J. Shi 2021). Results show that the pandemic will directly affect the import and export of the ports whereas the government's COVID-19 control policy will have a positive effect on exports. Narasimha et al. (2021) examine the COVID-19 impact on the Indian maritime industry and port operation from a statistical perspective (Narasimha, Jena, and Majhi 2021). The authors also conduct a survey of experts in the maritime sector to gauge attitudes toward issues related to COVID-19 in the Indian maritime industry. Quantitative analysis indicates that India suffers a loss both in the

amount of vessel port calls and the volume of cargo traffic flows. The survey suggests that India lacks preparedness in maritime area and needs further strategy to tackle with COVID-19 related issues in the maritime area (Narasimha, Jena, and Majhi 2021). Bai et al. (2021) present a noval density based clustering algorithm based on automatic identification system (AIS) data to identify vessel voyage and port turnaround states at ports and propose corresponding measures to reduce port congestion (Bai, Jia, and M. Xu 2022). Jin et al. (2022) analyze the impact of the COVID-19 on the liner shipping route networks in China based on AIS data. The finding shows that the proportion and competitiveness of China's international liner shipping has increased since the outbreak of the COVID-19 (Jin, J. Chen, Z. Chen, Sun, and B. Yu 2022).

Wang et al. (2022) put forward a process framework to quantitatively evaluate the effect of the COVID-19 on the behaviors of visiting vessels at ports using AIS data (X. Wang, Z. Liu, Yan, H. Wang, and M. Zhang 2022). Additionally, the validity of the framework was assessed through a case study that analyzed vessel operating data in the Beibu Gulf region of China between 2019 and 2020. The findings suggest that the average berthing time and anchoring time for cargo vessel and oil tankers largely increase after the outbreak of the COVID-19 (X. Wang, Z. Liu, Yan, H. Wang, and M. Zhang 2022). Zhou et al. (2022) review the emergent technical trends and related research topic in container terminal operations in recent years and the impact of COVID-19 on them (C. Zhou, Zhu, Bell, Lee, and Chew 2022). Chen et al. (2022) take Danish waters as an example

and make a comparative analysis on the differences in the emissions and activities of passenger vessels before and after the outbreak of COVID-19 (Q. Chen, Ge, Lau, Dulebenets, Sun, Kawasaki, Mellalou, and Tao 2022). The results shows that COVID-19 has a notable influence on cruise ships: The number of cruise ships decreases by 31% and SOx emissions dropped by 51% (C. Zhou, Zhu, Bell, Lee, and Chew 2022). In port state control related area, Akyurek and Bolat (2020) conduct a comparative analysis using ship inspection records by the Paris Memorandum of Understanding (MOU) from January 2015 to May 2022. By employing grey relational analysis, they find that the amount of inspections decreased greatly, and alterations in deficiency and detention conditions were primarily attributed to the impact of COVID-19. (Akyurek and Bolat 2020). Yan et al. (2022) evaluate the impact of COVID-19 on the regional port state control inspections. They take the HKP as an example and find that the monthly number of inspections and the average number of ship deficiencies are significant lower after the outbreak of the COVID-19 (Yan, Mo, X. Guo, Y. Yang, and S. Wang 2022).

Viewed from a global perspective, Notteboom et al. (2020) assess the effect of the COVID-19 in Africa. The results show that maritime and logistics sectors in African countries have been hit seriously and will not recover quickly due to the prolonged COVID-19 (Humphreys, Dumitrescu, Biju, and Y. Y. Lam 2020). Cullinane and Haralambides (2021) give an overview of the global trends of the maritime and port economics since the outbreak of the COVID-19 and a practical outlook on maritime economics in post COVID-19

era (Cullinane and Haralambides 2021). Bai et al. (2022) quantify the impact of COVID-19 lockdown policy on global vessel port calls by combining difference-in-difference model with regression discontinuity design model (Bai, M. Xu, T. Han, and D. Yang 2022). Zhao et al.(2022) use an exponential smoothing model to evaluate the fluctuations in the dry bulk vessel and container vessel transportation during the COVID-19 period (Zhao, H.-D. He, Lu, X.-L. Han, Ding, and Peng 2022). The results show that global dry bulk transportation is significantly affected by lockdown policies.

In summary, there are tons of research focusing on evaluating the effects of COVID-19 on maritime transportation area. However, in terms of vessel arrival and port operation issues, most of the studies concentrate on a single dimension. For example, Wang et al. (2022) evaluate the COVID-19 impact on vessel port calls at the Beibu Gulf region of China between 2019 and 2020 (X. Wang, Z. Liu, Yan, H. Wang, and M. Zhang 2022). Yan et al. (2022) evaluate the impact of COVID-19 on the number of port state control inspections at the HKP (Yan, Mo, X. Guo, Y. Yang, and S. Wang 2022). Narasimha et al. (2021) show that India suffers a loss in the volume of cargo traffic flows in 2020 (Narasimha, Jena, and Majhi 2021). Current research lacks studies that evaluate the impact of COVID-19 on vessel arrivals and port operations from multiple perspectives simultaneously. Moreover, in these studies, the comparative data covers only a few months. Additionally, none of these studies explores vessel arrival punctuality or turnaround time, both of which are crucial aspects of daily port operations. To address the research gap and

provide a detailed analysis of COVID-19's impact on VAT and port operation patterns, this paper takes the HKP as an example and uses four years' historical data to conduct comparative analysis on vessel arrival and port operation status from three perspectives: vessel arrival and operation statistics, vessel arrival punctuality, and VTT before and after COVID-19 outbreak at the HKP.

2.4 Berth Allocation Problem: Transition from Model-Driven to Data-Driven Approaches

The berth allocation problem (BAP), as a critical tactical and operational decision in container terminal management, aims to optimize the assignment of berthing positions and service sequences to incoming vessels while minimizing total waiting time, service delays, and resource conflicts. Upon arrival, container vessels typically anchor on the roadstead until assigned to an available berth, a designated quayside slot for cargo handling. Once berthed, quay cranes execute loading/unloading operations, initiating a cascade of downstream logistics activities, including yard storage planning and equipment deployment. As the primary input for terminal operational planning, the berth schedule directly determines terminal competitiveness within global shipping networks by balancing service efficiency and resource fairness.

The BAP has been extensively studied since its seminal formulation as a mixed-integer programming (MIP) model by Imai, Nishimura, and Papadimitriou (2001), who pioneered the spatial-temporal deci-

sion framework for discrete dynamic berth allocation. Subsequent foundational works established key BAP modeling paradigms: Y.-M. Park and K. H. Kim (2003) introduce the formulation of discrete BAP with fixed berth layout, while Hansen, Oğuz, and Mladenović (2008) propose the solutions of the continuous BAP to accommodate variable vessel lengths. These early studies laid the groundwork for addressing BAP under deterministic assumptions, where VAT, handling rates, and operational constraints are fully known (Rodrigues and Agra 2022).

Over the past two decades, research on the BAP has expanded significantly, evolving from basic formulations toward more integrated and context-specific variants. In particular, context-driven models have been developed to address the complexities encountered in real-world port operations. These include accommodating tidal window constraints within daily scheduling processes (Du, Q. Chen, J. S. L. Lam, Y. Xu, and Cao 2015), optimizing berth allocation under stringent vessel emission control regulations (J. Yu, G. Tang, Voß, and Song 2023), and mitigating disruptions caused by extreme weather conditions (L. Guo, J. Wang, and Zheng 2021). Additionally, the scope of BAP-related research has extended to various operational and environmental scenarios, such as optimizing berth allocations for multi-terminal ports, coordinating with hinterland transportation schedules, and integrating automated terminal operations. These studies highlight the diversity and depth of scenarios currently explored in the BAP literature, demonstrating the field's ongoing evolution and its responsiveness to emerging challenges (Rodrigues and

Agra 2022).

Among these diverse scenarios, evaluating and managing VAT has emerged as a hot and critical issue in BAP related research. Inaccuracies or uncertainties in VAT can directly lead to substantial berth operation delays and increased operational costs. Existing research has extensively investigated methods for modeling uncertainty in VAT to support effective berth allocation decisions. For example, Golias, Saharidis, Boile, Theofanis, and Ierapetritou (2009) introduce a discrete berth scheduling, dynamically optimizing vessel speeds and arrival time to minimize vessel fuel consumption and overall waiting time. Zhen, Lee, and Chew (2011) propose a reactive dynamic strategy to adjust initial BAP schedules dynamically, enabling the management of realistic variations in VAT with minimal deviation cost from the original schedule. This development effectively transforms BAP models from deterministic frameworks into dynamic one. In these studies, VATs are typically represented as uncertainty sets or assumed to follow specific probability distributions, prompting innovative model adaptations and computational strategies to enhance robustness and flexibility in berth allocation planning.

Recently, data-driven machine learning methodologies have gained increasing attention in maritime operations and logistics research (Yan, S. Wang, Zhen, and Laporte 2021; Rodrigues and Agra 2022). By leveraging extensive historical data and real-time information, these approaches significantly enhance predictions of critical metrics, such as container throughput, VTT, fuel consumption, and VAT (Rodrigues and Agra 2022). In particular, numerous studies highlight that these

methods can effectively reduce uncertainties in VAT by providing more accurate VAT predictions. However, the operational value of machine learning-based VAT prediction remains insufficiently quantified within the context of BAP optimization, where precise scheduling critically depends on reliable VAT estimate. Current related studies predominantly remain model-driven, focusing on developing and comparing various BAP models or solution algorithms under a fixed VAT assumption. Few studies have adopted a data-driven perspective to examine how variations in VAT prediction accuracy impact BAP performance, a research gap that highlights the importance of accurate VAT estimation in subsequent berth scheduling.

To date, only two studies have specifically integrated VAT prediction models to optimize the subsequent BAP scheduling decisions. The first study by J. Yu, G. Tang, Song, X. Yu, Qi, D. Li, and Y. Zhang (2018) use static vessel data, including vessel-reported ETA, vessel physical features to predict VAT at Ningbo Gangji Container Terminal (GYCT). By comparing the performance of the BAP using predicted VAT against the actual VAT, the results showe that the use of predicted VAT significantly improve the subsequent quality of the BAP scheduling, achieving approximately 60% enhancement in performance metrics. Kolley, Rückert, Kastner, Jahn, and Fischer (2023) employ linear regression, k-nearest neighbors (KNN), decision tree regression, and neural networks to predict VAT with KNN ahieve best performance. Dynamic time buffers (DTBs) are dynamically adjusted based on prediction discrepancies across VAT prediction models: larger variations in VAT forecasts trigger longer

buffers to absorb uncertainty. These DTBs are integrated into a dynamic BAP model in the second stage. Compared to the benchmark BAP model, this approach significantly reduces vessel waiting time by 37% (565.71 mins vs. 902.47 min) and improves true service levels by 13.4 percentage points (78.88% vs. 65.49%), demonstrating enhanced robustness.

However, both studies have notable limitations. From the perspective of the data used for VAT prediction, J. Yu, G. Tang, Song, X. Yu, Qi, D. Li, and Y. Zhang (2018) primarily rely on static vessel-reported ETA, while Kolley, Rückert, Kastner, Jahn, and Fischer (2023)utilize dynamic AIS data; neither combines ETA and AIS data simultaneously for predicting VAT. Regarding prediction methods, both studies employ outdated models. Specifically, J. Yu, G. Tang, Song, X. Yu, Qi, D. Li, and Y. Zhang (2018) adopt a classification approach to predict arrival error intervals instead of accurately forecasting the precise VAT. Furthermore, J. Yu, G. Tang, Song, X. Yu, Qi, D. Li, and Y. Zhang (2018) method of comparing BAP performance across different VAT scenarios is methodologically questionable. Direct comparison of BAP outcomes under varying VAT predictions is inherently inappropriate, as there are always discrepancies between predicted VAT, and ATA. Thus, BAP results based on predicted VAT or ETA implicitly assume vessel ATA follows predicted schedules, making such direct comparisons invalid. Kolley, Rückert, Kastner, Jahn, and Fischer (2023) on the other hand, indirectly incorporate VAT predictions through modeling DTBs to enhance robustness in BAP optimization, rather than explicitly integrating predicted

VAT directly into the BAP formulation and performance comparison.

Chapter 3

Evaluation of Vessel and Port Performance

3.1 Evaluate Vessel Arrival Punctuality

3.1.1 HKP and data introduction

The HKP, situated in the South China Sea, specializes in handling containerized manufactured goods. It is renowned for being one of the world's busiest and most efficient international container ports (Hong Kong Maritime And Port Board 2024). In 2023, the HKP handles nearly 15 million TEUs, making it the 10th largest container port globally in terms of throughput (Hong Kong Maritime And Port Board 2024). As of December 2023, it offered approximately 300 weekly international container liner services, connecting to over 650 destinations worldwide, showcasing its critical role in supporting global maritime logistics and trade networks (Hong Kong Maritime And Port Board 2024). A detailed geographical depiction of HKP is shown

in Figure 3.1.[1] The Marine Department of Hong Kong regularly

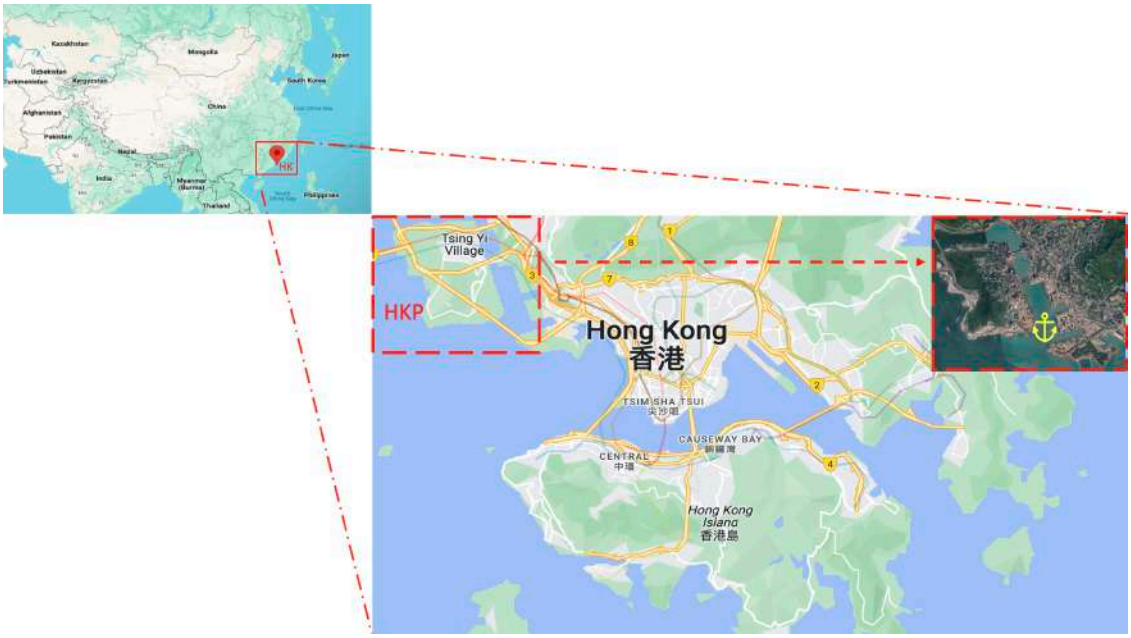


Figure 3.1: HKP location overview

(every 20 minutes) updates the arrival and departure information of ocean-going vessels on its website (Hong Kong Maritime And Port Board 2024). This dataset is publicly accessible and categorized into four types: vessels to arrive in the next 36 hours, vessels in port, vessels have arrived in the past 36 hours, and vessels have departed in the past 36 hours. AIS data includes both static details, such as the vessel's name and call sign, as well as dynamic information like latitude, longitude, speed over ground (SOG), and course over ground (COG). AIS data not only enables precise, real-time tracking but also supports historical data analysis, allowing the identification of patterns and trends over time. The variables for port call data and AIS data used are listed in Table 3.1. To evaluate and predict VAT to the

[1]The figure is generated from GoogleMap (`https://www.google.com/maps/`).

Table 3.1: Variable descriptions in port call and AIS data

Feature	Definition	Data Resource				
		Vessels arrived in last 36 hours	**Vessels due to arrive in the next 36 hours**	**Vessels departed in the last 36 hours**	**Vessel in port**	**AIS data**
Vessel name	Name of the vessel	✓	✓	✓	✓	✓
Vessel type	Type of the vessel		✓		✓	✓
Trip status	Current status of the vessel trip		✓			
Agent name	Name of the vessel's agent	✓	✓		✓	
Flag	Vessel registration country		✓		✓	
ETA	Vessel estimated time of arrival reported by the vessel caption		✓			
ATA	Vessel actual time of arrival	✓			✓	
ATD	Vessel actual time of departure			✓		
Longitude	Vessel's longitude coordinate.					✓
Latitude	Vessel's latitude coordinate.					✓
SOG	Vessel speed over ground					✓
Report time	Time of data generation	✓	✓	✓	✓	✓
Last port	Name of a vessel's last port of call before arrival		✓			
IMO number	IMO number of a vessel				✓	
Call sign	A unique alphanumeric vessel identifier	✓	✓	✓	✓	✓
Last berth	Vessel's last berthing location			✓		
Arrived location	The first location where a vessel arrives to the HKP	✓			✓	

HKP, we first collect port call and AIS data from January 1st, 2021, to February 28th, 2021. Then, we extract the targeted vessel AIS data from the AIS dataset using the vessel's call sign. It is important to note that vessels may report multiple ETA records during a single voyage to HKP, but only one ATA record is generated upon arrival. ETA and ATA port call data are stored in separate files, and a single vessel may make multiple port arrivals within a single year. Moreover, the vessel AIS data is not aligned with the corresponding ETA and ATA data, requiring additional preprocessing and pairing.

3.1.2 Port call and AIS data matching

We begin by matching the ATA data to the ETA data, followed by aligning both the ETA and ATA data with the AIS data. Before pairing the datasets, we conduct a preliminary cleaning process that involves removing duplicate ATA entries and discarding any ATA records that occur after their corresponding report time. Additionally, we filter out any AIS data entries where the latitude or longitude exceeds the maximum limits of 90° and 180°. We also standardize the time formats across the AIS, ETA, and ATA datasets. During the pairing process, AIS data serves as the baseline: we segment the AIS data by voyage based on ATA and call sign values, match the vessel's ETA and ATA values, and finally align the ETA data with the corresponding AIS records. Let i denote the identifier for a specific vessel, where $i = 1, \ldots, M$, and the dataset comprises M distinct vessels in the ETA and ATA records. Each vessel i has a total of n_i ETA records, k_i ATA records, and P_i AIS data points in the datasets. The

jth ETA of vessel i is represented as ETA_i^j, with $j = 1, \ldots, n_i$, and the kth ATA is represented as ATA_i^k, with $k = 1, \ldots, k_i$. The time of the pth AIS point is represented as AIS_i^p, with $p = 1, \ldots, P_i$. The report time of ETA_i^j is denoted as R_i^j. To pair the records, we introduce a time-based comparative interpolation method, which is described in detail in Algorithm 1. In the algorithm, the notation $a > b$ means that the time of a is later than the time of b, a 5-day threshold is set to avoid mismatches where vessel only report ETA without actual arrival, and P_i' and n_i' are the number of AIS and ETA data records corresponding to vessel i after completing the first two steps of data cleaning, respectively. The algorithm leverages temporal window matching and chronological greedy search by sorting ETA, ATA, and AIS data chronologically for each vessel and iteratively matching records based on time difference constraints to enhance computational efficiency and clarity. After pairing and cleaning the data, we obtained 314,212 AIS $\leftrightarrow$ ETA $\leftrightarrow$ ATA paired records from 1,863 vessel voyages. Figure 3.2 shows an example visualization of vessel arrival trajectories at the HKP. After pairing the data, each AIS data has a corresponding ETA and ATA data, and we present an in-depth analysis of the accuracy of ETA data reported by vessels. The accuracy of ETA data reported by vessels en route to a port can vary significantly depending on the remaining sailing time or distance to destination. For instance, an ETA reported 30 hours before arrival is likely to be less accurate than one reported 1 hour before arrival. Similarly, an ETA reported when a vessel is 100 nautical mile away is generally less precise than one reported when it is only 10 nautical

Input: ETA, ATA and AIS data
Output: AIS $\leftrightarrow$ ETA $\leftrightarrow$ ATA paired data
begin
  Unify the time format
  `// Segment vessel AIS voyage data`
  for *each vessel* $i \leftarrow 1$ **to** M **do**
    Sort the $\text{ETA}_i^{n_i}$, $\text{ATA}_i^{k_i}$, $\text{AIS}_i^{P_i}$ of vessel i chronologically
    $k \leftarrow 1$
    for $p \leftarrow 1$ **to** P_i **do**
      while $k \leq k_i$ ***and*** $ATA_i^k < AIS_i^p$ **do**
        $k \leftarrow k + 1$
      end
      if $0 < ATA_i^k - AIS_i^p \leq 48$ *hours* **then**
        AIS_i^p [ATA] $\leftarrow \text{ATA}_i^k$
      end
    end
    Remove any unpaired AIS_i^p
    `// Pair ATA data to ETA data`
    $k \leftarrow 1$
    for $j \leftarrow 1$ **to** n_i **do**
      while $k < k_i$ ***and*** $R_i^j > ATA_i^k$ **do**
        $k \leftarrow k + 1$
      end
      if $\left|ETA_i^j - ATA_i^k\right| \leq 120$ *hours* **then**
        ETA_i^j [ATA] $\leftarrow \text{ATA}_i^k$
      end
    end
    Remove any unpaired $\text{ETA}_i^{n_i}$
    `// Pair ETA data to AIS data`
    for $k \leftarrow 1$ **to** k_i **do**
      Sort $\text{AIS}_i^{P_i'}$ and $\text{ETA}_i^{n_i'}$ chronologically
      $j \leftarrow 1$
      for $p \leftarrow 1$ **to** P_i' **do**
        if $j < n_i'$ ***and*** $AIS_i^p > R_i^j$ **then**
          $j \leftarrow j + 1$
        else
          AIS_i^p[ETA] $\leftarrow \text{ETA}_i^j$
          AIS_i^p[ATA] $\leftarrow \text{ETA}_i^j$[ATA]
        end
      end
    end
    Remove any unpaired AIS_i^p
  end
end

Algorithm 1: ETA, ATA and AIS data pairing algorithm

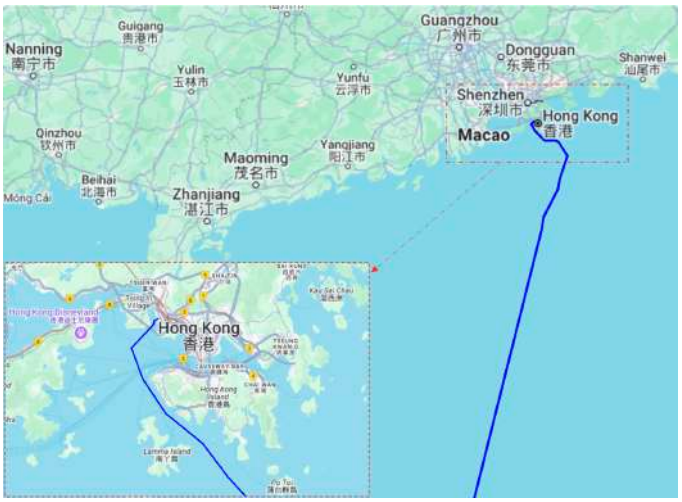


Figure 3.2: A sample visualization of vessel AIS arrival trajectories to the HKP

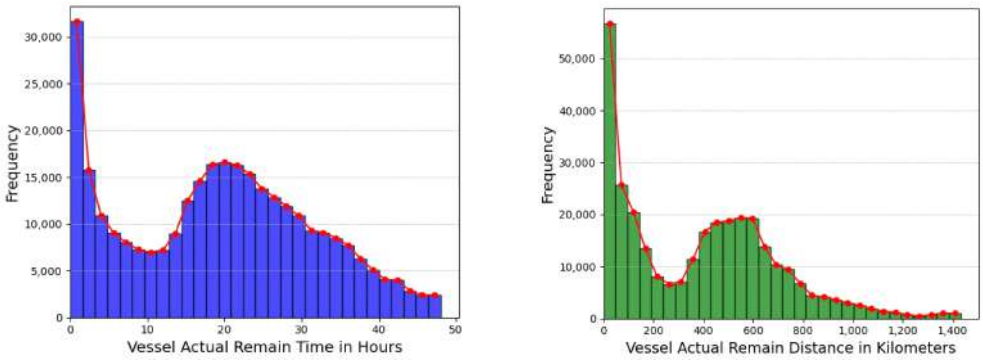
mile away. This difference in accuracy arises because uncertain factors affecting the journey become less impactful as the vessel nears the port, allowing for more precise ETA predictions. Therefore, it is not reasonable to perform a direct comparison of the accuracy of reported ETA records at different remaining time and distances, as they are influenced by different levels of uncertainty.

In the previous subsection, raw AIS data is matched with port call data, segmenting them into distinct voyages. This approach makes it possible to estimate a vessel's remaining actual sailing time and distance corresponding to each AIS point by applying a recursive method, starting from the last point and working backward through the second-to-last point in its trajectory. Specifically, to calculate the distance between two adjacent AIS points, we assume the vessel follows a great-circle route and use the Haversine formula for this estimation, with the last AIS point of each segment being considered the midpoint of the route. Denote the coordinate of two AIS records

as (φ_1, λ_1) and (φ_2, λ_2), where φ represents latitude and λ represents longitude, here, φ and λ in AIS data are in degrees and need to be converted to radians, with the formulation: $\varphi = \varphi \times \frac{\pi}{180}$, $\lambda = \lambda \times \frac{\pi}{180}$. The distance d between the two AIS points is then calculated using the Haversine formula:

$$d = 2r \arcsin\left(\sqrt{\sin^2\left(\frac{\Delta\varphi}{2}\right) + \cos(\varphi_1)\cos(\varphi_2)\sin^2\left(\frac{\Delta\lambda}{2}\right)}\right), \tag{3.1}$$

where r is the radius of the Earth. Utilizing these estimates, the AIS data can be systematically divided into discrete time slices and distance slices for the following punctuality analysis. In the evaluation section, we first provide an overview of the vessel's actual remaining voyage time and distance to port based on AIS-reported data, as visualized in Figure 3.3. The histogram in Figure 3.3(a), which shows



(a) Distribution of vessel actual remain time

(b) Distribution of vessel actual remain distance

Figure 3.3: An overview of vessel actual remain voyage time and distance

the actual remaining voyage time for each vessel AIS data in hours, reveals a left-skewed distribution with a pronounced peak at or near zero hours. Additionally, there is a secondary peak around the 20-hour mark, suggesting another notable cluster of vessel AIS data.

Beyond this point, the frequency gradually decreases. Similarly, the histogram in Figure 3.3(b), which displays the vessel actual remaining distance in kilometers, reveals a left-skewed distribution, with a high frequency of AIS data having a short remaining distance and a peak near zero km. A secondary peak occurs between 400 and 600 kilometers, after which the frequency declines steadily, extending up to approximately 1,400 kilometers. As the HKP authority mandates that vessels begin reporting the ETA data 36 hours prior to arrival, the AIS data is initially segmented into 37 time intervals, spanning from "0 hour" to "36 hours," with each interval representing a 1-hour window. For instance, the "0 hour" interval includes vessels for which the time difference between the AIS report timestamp and the ATA lies within 0 to 1 hour. The "36 hours" time slice encompasses vessels with a time difference of 36 hours or more (e.g., 38 or 40 hours).

Similarly, from a spatial perspective, we use 50 km as the interval to divide the AIS data into 21 distance slices. The first distance slice includes vessels with an actual remaining voyage distance between 0 and 50 km. The last distance slice includes all AIS data for vessels with an actual remaining voyage distance greater than or equal to 1000 km. To evaluate vessel arrival punctuality offline, the ETA data in the paired AIS will be considered as the predicted value, while the ATA data will serve as the ground truth. Four standard metrics will be utilized: root mean squared error ($RMSE$) , MAE, R-squared (R^2) and mean absolute percentage error ($MAPE$). Define n as the number of ETA or ATA data points, let y_i represent the observed ATA

values and $\hat{y}_i$ denote the vessel reported ETA values, $\bar{y}$ is the mean of the ATA for vessel $i, i = 1, \ldots, M$. The definition of four metrics is written as follows:

$RMSE$:

$$RMSE = \sqrt{\sum_{i=1}^{m} \frac{(y_i - \hat{y}_i)^2}{m}}, \tag{3.2}$$

MAE:

$$MAE = \frac{\sum_{i=1}^{m} |y_i - \hat{y}_i|}{m}, \tag{3.3}$$

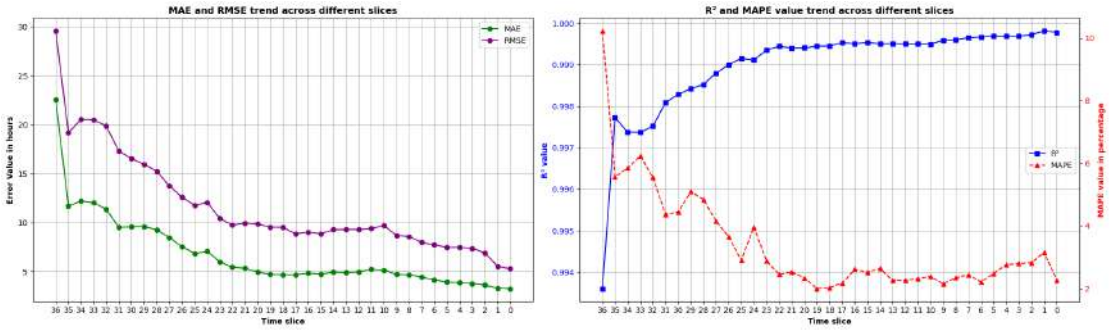
R^2:

$$R^2 = 1 - \frac{\sum_{i=1}^{m} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{m} (y_i - \bar{y})^2}, \tag{3.4}$$

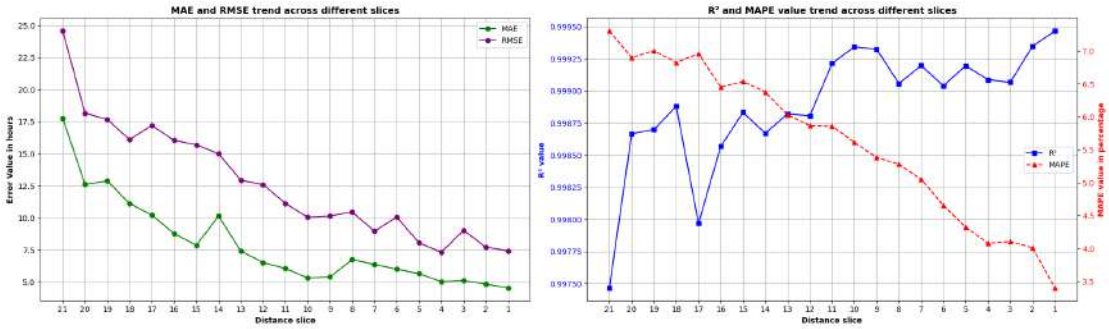
$MAPE$:

$$\text{MAPE} = \frac{1}{m} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100. \tag{3.5}$$

When assessing the R_2 value of ETA reports, it is not feasible to directly calculate it using the AIS report time and ATA timestamp. To address this issue, we first compute the time differences between these points and timestamp January 1st, 2021, 00:00. Subsequently, we perform the R^2 value calculation. The evaluation results for the four metrics of vessel arrival punctuality across various time and distance intervals are presented in Figure 3.4. Figure 3.4(a) and (b) depict the accuracy of vessel-reported ETA across different time slices using four evaluation metrics. The figures illustrate that, from both temporal and spatial perspectives, vessel-reported ETA errors consistently decrease as vessels approach the HKP across all evaluation metrics. Quantitative results further indicate that reduced remaining



(a) Quantitative analysis of vessel reported ETA in different time slices



(b) Quantitative analysis of vessel reported ETA in different distance slices

Figure 3.4: Spatial and temporal analysis of vessel reported ETA in time and distances slices

voyage time and distance correlate with increased accuracy and reliability of vessel-reported ETA data as vessels approaches the port. Moreover, it is noteworthy to observe a significant drop in all four evaluation metrics. from the “36 hours” slice to the “35 hours” slice. Similarly, a notable decline is also evident between the “21 slices” and “20 slices” in the distance slices. This is due to the fact that the “36 hours” interval includes ETA records where the time difference between the AIS report time and the ATA surpasses 36 hours, whereas the “35 hours” slice only includes records with a time difference between 35 to 36 hours. Likewise, in the distance slices, the “21 slices” category includes vessels with an actual remaining voyage distance greater than 1,000 km, while the “20 slices” category includes vessels with a remaining distance between 950 to 1,000 km. In the earlier stages of the journey, vessels tend to provide less reliable ETA data, leading to larger prediction errors.

However, there are noticeable fluctuations in error trends within certain time and distance slices, despite the general trend of decreasing error. These fluctuations are primarily caused by variations in the amount of data available for each slice. In this analysis, only two months’ vessel port call and AIS data are considered, leading to differences in the data volume for each slice. For slices with smaller data volumes, the presence of a few outliers with extreme ETA errors can significantly skew the average error for that slice, resulting in visible deviations in the trend lines. This explains the observed variations, where certain slices exhibit larger errors, despite the overall trend toward greater accuracy as vessels approach the port.

3.2 Quantitative Analysis of Vessel Turnaround and Service time

Vessel Turnaround Time (VTT) refers to the total time a ship spends in port, from its arrival at the port to the completion of all necessary services and its departure. In contrast, Vessel Service Time (VST) specifically measures the duration from the moment the vessel arrives at the berth until it leaves. Port vessel service operations encompass processes like documentation, cargo loading, and unloading, with VTT and VST having a significant impact on a port's operational performance, economic viability, and market competitiveness. During routine port operations, vessels usually provide their estimated departure time (EDT) prior to leaving the berth. Nevertheless, the EDT often diverges considerably from the actual departure time (ADT) because of unpredictable factors like port inefficiencies and congestion. Such discrepancies introduce uncertainties in VST estimation. As a result, these discrepancies may cause notable delays in vessel departures, escalate port congestion, and elevate operational costs. In this section, We aim to carry out data pre-processing and analysis, with the aim of quantitatively evaluating the VTT and VST at the HKP.

3.2.1 Description of port call dataset

In this study, we utilize two port call datasets for VST related research. One is the HKP vessel arrival and departure information[2]

[2]This data can be accessed online at the HK government's data portal: `https://data.gov.hk/en-data/dataset/hk-md-mardep-vessel-traffic-management-system-report`.

and the other is the VT Explorer vessel port entry and exit data.[3] Detailed descriptions of these data are as follows.

The Hong Kong port authority updates daily port call data in the website (Hong Kong Government 2022), and this information is publicly accessible. The website categorizes the data into five sections: vessels arrive in the last 36 hours, vessels will arrive in the next 36 hours, in port vessels, vessels depart in the last 36 hours, and vessels will depart HKP in the next 36 hours.

We use vessel AIS records as a supplementary data source to form our VST dataset. AIS data offers detailed insights into both static vessel information and dynamic movements. In this study, we leverage AIS data, specifically the timing of vessel entries and exits from the HKP's anchorage areas, as supplementary information. Data from the two datasets used in the VST study and their detailed descriptions are summarized in Table 3.2. Additionally, Table 3.3 illustrates the features and update frequencies of these datasets.

[3]This data can be accessed and downloaded from the VT Explorer website: `https://www.vtexplorer.com/`.

Table 3.2: Variables in the HKP dataset

Feature	Explanation	Note
Ship_type	Ship type	Totally 14 types
Trip_status	The status of ship voyage	Pending or Approved for a ship enter HK water area
ETA	Vessel estimated time of arrival (ETA) to the HKP	Reported by the vessel itself
ATA	Time of a vessel's actual arrival at the berth	Recorded by the port
EDT	Vessel estimated departure time from the HKP	Provided by the vessel operator
ADT	Actual time of a vessel's departure from the berth	Recorded when a vessel departs from the berth area in HKP
Enter_time	The time the vessel enters the anchorage area	Recorded by the port
Exit_time	The time the vessel exits the anchorage area	Recorded by the port
Upload_time	Time of data upload onto the website	Provided on the website
IMO number	The International Maritime Organization (IMO) number of a ship	A unique vessel identifier comprising seven digits
Call sign	The unique alphanumeric code identifying a ship for radio communication	A unique vessel identifier
Berth	The vessel's present location of berthing	\
Arrive_location	The initial location where a vessel stays upon arriving in HK waters	\
Agent_name	Vessel's agent name	\

Table 3.3: Specific distribution of variables in the dataset

	Vessels arrived in the last 36 hours	Vessels will arrive in the next 36 hours	Vessels departed in the last 36 hours	In port vessels	Vessels will depart in the next 36 hours	Vessel entry
Update_frequency	Daily	Daily	Daily	20 minutes	Daily	Daily
Ship_type	✓	✓	✓	✓		
Trip_status		✓				
ETA		✓				
ATA	✓			✓		
EDT					✓	
ADT			✓			
Enter_time						✓
Exit_time						✓
Upload_time	✓	✓	✓	✓	✓	✓
IMO number			✓			✓
Call sign	✓	✓	✓	✓	✓	✓
Berth	✓		✓	✓	✓	
Agent_name		✓		✓	✓	✓

For the subsequent quantitative analysis and prediction of VTT and VST, we have collected all related data from January 1, 2022 to March 31, 2023 for our analysis. The ATA, EDT and ADT datasets contain 28,126, 18,365 and 19,430 records, respectively. As indicated in the data description above, the ATA, EDT and ADT data records reported by a vessel are stored in different files. The expected VTT for a port call is the difference between the EDT and ATA for that particular voyage, while the actual VTT is the difference between the ADT and ATA. VST is defined as the time difference between the departure time and the arrival time of the vessel at berth. These records are scattered across different files. Therefore, in order to quantify the specific duration of the VTT and VST, further data pre-processing is required.

3.2.2 Data preprocessing

Normally, the process for a ship arriving and departing from the berth goes as follows: First, the ship reports its ETA data before reaching the port, and the port authority allocates a berth for the ship based on this ETA. Subsequently, the ship enters the port's anchorage, and the port logs the time of entry, referred to as the Enter_time data in this study. After that, the ship reaches its berth, and the port records the ATA data. Sometimes, if a berth is available, the ship does not need to wait in the anchorage and can proceed directly to the berth. In such cases, the ship's Enter_time and ATA are the same. After the port completes its operations, during the ship's departure process, the ship first leaves the berth, at which point the port records the ADT data. Finally, as the ship leaves the anchorage area, the port logs the Exit_time data. In addition, the ship may report its EDT data at any time after reporting the ETA and before leaving the berth. Below is an illustration of the process of reporting data for a ship's arrival and departure from the port. Among these data, ETA and EDT are estimates reported by the ship, while the remaining data records are recorded by the port authority.

$$\text{ETA} \xrightarrow{(\text{ETD})} \text{Enter\_time} \xrightarrow{(\text{ETD})} \text{ATA} \xrightarrow{(\text{ETD})} \text{ATD} \longrightarrow \text{Exit\_time}$$

To match and quantify the VST data, we use the Enter_time and Exit_time data as the foundational elements, and match other relevant data with them. The overview of the data processing is shown in Figure 3.5, and the description of each step is as follows:

1. Collect original vessel port call data

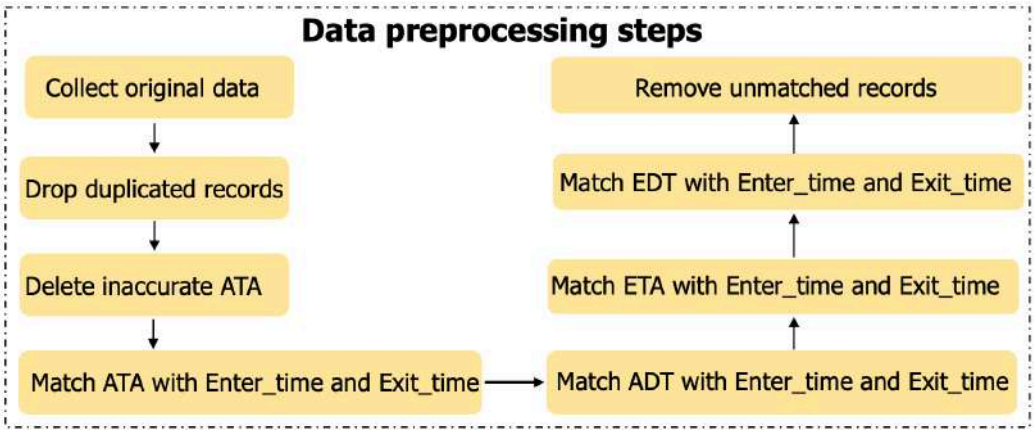


Figure 3.5: The overview of the data preprocessing steps

The initial step in creating the dataset involves aggregating vessel arrival and departure data from various folders. Once collected, the time formats are converted to the "Year, Month, Day, Hour " standard.

2. Remove duplicate records

   The HKP authority provides daily port call updates. This includes creating records for ETA, ATA, EDT, ADT and berth entry and exit time for each vessel. However, after consolidating the data, we find multiple duplicate ETA, ATA, EDT, ADT, and berth entry/exit records for a vessel's single port call, each with varying upload time. In these cases, we retain only the first record of ATA, ADT, and berth movement data based on the upload time. For ETA and EDT data, if a vessel reports several different ETA and EDT records, we preserve each unique ETA and EDT records, as they provide valuable insights into alterations in the vessel's arrival and departure timetable, which is crucial for the following VST analysis. Nevertheless, when the

ETA or EDT record remains unchanged over multiple uploads with just the upload time varying, we remove these duplicates, retaining only the first record.

3. Delete ATA records that are later than their upload time
   The HKP logs the ATA of a vessel after its arrival, ensuring the upload time is always subsequent to the ATA. If the ATA appears before its upload time, it likely indicates system errors, rendering the data inaccurate. Utilizing such erroneous ATA data can adversely affect vessel arrival analyses and predictions. Thus, any record where the ATA precedes its upload time is excluded from the dataset.

4. Match vessel ATA data with the Enter_time and Exit_time records
   During vessel operations at the port, vessel ATA, ADT, Enter_time and Exit_time records are unique for a specific visit, with Enter_time and Exit_time forming a pair. The ATA always occurs no earlier than the Enter_time. To match the ATA with the Enter_time record, we first filter the dataset by call sign to identify records of ATA and Enter_time/Exit_time belonging to the same vessel. Then, for each Enter_time record, we identify the ATA that follows the Enter_time and has the smallest time difference. Furthermore, the time difference between the ATA and Enter_time should not exceed 6 hours for a match to be considered valid.[4] If there is no ATA that satisfies both criteria, the

[4]We enforce a 6-hour time threshold to prevent instances where the ATA is recorded at the vessel's arrival, but the Enter_time data is not recorded when the ship enters the berth. Without this 6-hour restriction, the Enter_time of the next voyage might be incorrectly matched with the current voyage's ATA, adversely affecting subsequent data matching.

Enter_time record is considered not to have the corresponding ATA data match.

5. Match the vessel ADT data with the Enter_time and Exit_time records

   The ADT data is recorded later than both the ATA and the Enter_time time but earlier than the Exit_time for each voyage. To pair the ADT data to Exit_time data, we initially filter the ADT data from the dataset using the call sign for the specific vessel. Then, for every Exit_time record, we find the closest ADT record that occurs before the Exit_time time. If the time difference between this ADT record and the Exit_time record is less than 24 hours, we pair the ADT record with the respective Exit_time record.

6. Match vessel ETA and EDT data with the Enter_time and Exit_time records

   Like in the previous procedure, we begin by sorting the ETA records in the dataset using the call sign of the target vessel. Considering that vessels often report multiple ETAs as they approach a port, and taking into account that the HKP requires ships to report data 36 hours before arriving, we adopt a specific strategy to pair ETA data with Enter_time/Exit_time records. To determine the most appropriate match, we focus on a specific time range surrounding the ATA data that has already been matched with a particular Enter/Exit record. This time frame extends to two days before, one day before, the same day, one day after, and two days after the ATA date. Within this time-

frame, we go through all covered ETA records and select the earliest ETA record where the ETA report time differs from the ATA by less than 36 hours. This ETA is then considered as the matched ETA to the Enter/Exit record.[5] The matching process of the EDT data is similar to that of the ETA data. Since EDT reports can be made either before the vessel reaches its berth or after arrival, and considering that multiple EDT reports might be submitted post-arrival, we follow a dual strategy. First, we select the first EDT that is reported within 36 hours before the ATA, capturing the earliest departure estimate before the vessel's arrival. Second, we also record the first EDT reported after the vessel has reached the port, irrespective of the exact report time. This approach ensures comprehensive coverage of EDT data, reflecting estimates both prior to and following the vessel's arrival at the port. Notably, among the ETA and EDT data, there are records for 14 different types of vessels arriving at HKP. However, the Enter_time specifically pertains to vessels with a designated berth allocation, specifically container vessels. Consequently, after processing this data, the retained records exclusively concern container vessels.

7. Remove unmatched Enter_time and Exit_time records:
   After finishing the above processing, any Enter_time/Exit_time

[5]Selecting the first ETA within a 36-hour window is based on operational considerations for port management. Opting for the ETA closest to the time of docking often lacks significant value for port operations. Similarly, the first reported ETA might be highly inaccurate and too far away in advance of reaching the port. Given the HKP requirement that vessels report data 36 hours before arrival, we use this 36-hour timeframe as a critical boundary. Within this limit, we record the earliest reported ETA data.

records without the corresponding ATA, EDT, or ADT are deemed erroneous and are deleted. This situation can be caused by system error logs or missing record data.

During the data processing phase, the number of records for ETA, ATA, EDT, and ATA, as well as Enter_time and Exit_time data, both before and after preprocessing, is summarized in Table 3.4.

Table 3.4: Summary of record counts in the data pre-processing scheme

Step	Method	ETA	ATA	EDT	ADT	Enter/Exit
1)	Data collection and time format unification	53,221	48,237	50,237	47,314	39,455
2)	Drop duplicated data	49,001	44,993	45,188	42,325	38,268
3)	Delete Abnormal ATA data	49,001	41,126	45,188	42,325	38,268
4)	Pair ATA data to Enter_time/Exit_time data	49,001	21,356	45,188	42,325	21,356
5)	Pair ADT data to Enter_time/Exit_time data	49,001	21,356	45,188	18,444	18,444
6)	Pair ETA and EDT data to Enter_time/Exit_time data	12,003	16,216	12,003	18,444	12,003
7)	Remove unmatched data	11,782	11,782	11,782	11,782	11,782

After matching and cleaning the data, a total of 6,899 records are successfully matched with ship-reported EDT data prior to port arrival, and 4,883 records are matched with post-arrival reported ETD data.

3.2.3 Quantitative analysis of VTT and VST at the HKP

This section presents a quantitative analysis of vessel arrivals and departures at HKP from September 2020 to March 2023, based on processed VTT and VST data. The analysis provides a detailed assessment of expected and actual VTT and VST, along with the associated delays. These insights are essential for port authorities to evaluate operational efficiency and formulate strategies for optimizing port service management. We begin by analyzing VTT.

Table 3.5: Statistic analysis of the actual VTT at the HKP

Item	Number of records	Mean (hours)	Median (hours)	Maximum (hours)	Minimum (hours)	Standard deviation (hours)
All types of vessels in the dataset	14,975	18.20	14.63	63.61	0.20	18.59
All types of vessels in 2022	11,938	18.80	15.16	63.61	0.21	18.91
All types of vessels in 2023	3,037	16.83	13.82	56.65	0.11	17.33
Total container vessels in the dataset	9,998	17.66	13.88	49.12	1.21	14.52
Container vessels in 2022	8,178	18.29	13.90	49.12	1.39	14.96
Container vessels in 2023	1,820	15.89	12.30	43.28	1.21	13.59

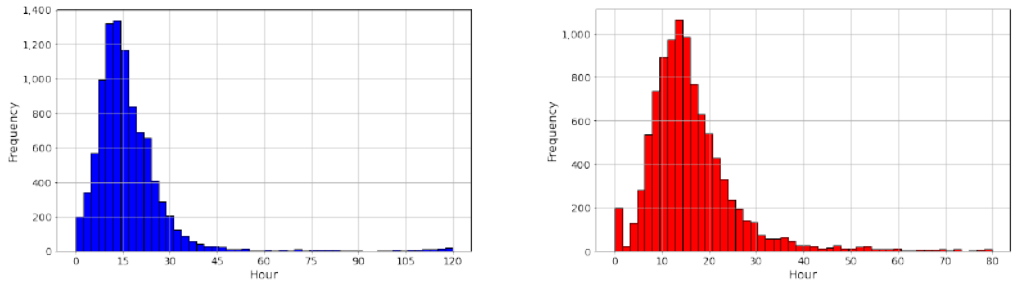
In Table 3.5, we find that the mean, median and variance values of actual VTT for container vessels are lower than those of all other types of vessels. This may be attributed to the fact that operations involving container vessels at the port are primarily automated, resulting in fewer disruptions from human factors compared to other types of vessels. Additionally, we find that the maximum value of the actual VTT of all types of vessels is significantly larger than that of container vessels. The reason is that these maximum values in VTT correspond to the arrival operations of cement vessels at the HKP. Next, we evaluate the expected VTT at the HKP. The expected VTT for each vessel is defined as the difference between its EDT and its ATA. Similar to the analysis of actual VTT, we have categorized the data into two groups: all types of vessels and the container vessels. The vessel expected VTT evaluation results in hours are presented in Table 3.6.

Table 3.6: Statistic analysis of the expected VTT at the HKP

Item	Total number of records	Mean (hours)	Median (hours)	Maximum (hours)	Minimum (hours)	Standard deviation (hours)
Total_vessel	14,975	17.43	13.65	119.98	−17.94	17.62
2022_vessel	11,938	17.85	14.17	119.98	−16.61	17.79
2023_vessel	3,037	15.82	11.95	156.65	−17.94	16.81
Total_container	9,998	16.34	14.63	69.95	−17.94	14.27
2022_container	8,178	14.68	14.68	69.20	−16.61	14.33
2023_container	1,820	14.48	12.08	69.95	−17.94	13.83

In Table 3.6, we encounter an anomaly where the minimum value of the expected VTT is less than zero. The expected VTT is calculated as EDT minus ATA, and the negative value anomaly arises from

vessels that report their EDT data within 36 hours before departing from the port. However, according to our previous analysis shown in Table 3.5, the average actual VTT at HKP is approximately 18 hours. In some rare instances, vessels report their EDT before they arrive at the port. Nonetheless, the reported EDT data often contain significant errors and might be earlier than the ATA at the port. This discrepancy results in a negative value by deducting EDT by ATA, which, in turn, leads to a negative expected VTT value in these scenarios. These instances could potentially be outliers in the dataset, which may skew prediction outcomes. To avoid this, we will exclude these outliers during the visualization and prediction phases. Figure 3.6 illustrates the distribution of both the expected and the actual VTT for container vessels at the HKP. In Figure 3.6, we ob-



(a) Distribution of the expected VTT of container vessels at the HKP

(b) Distribution of the actual VTT of container vessels at the HKP

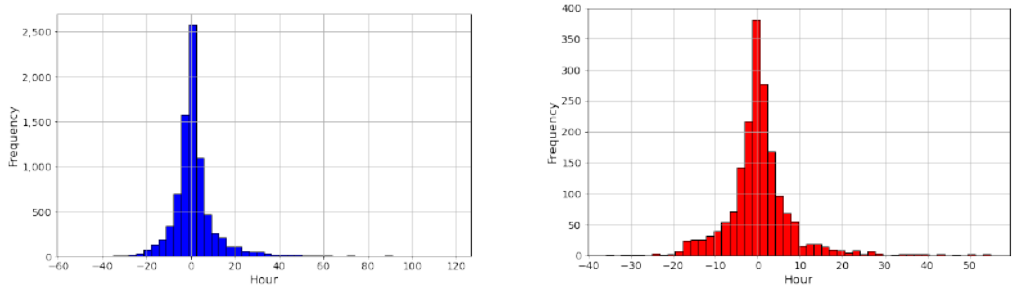
Figure 3.6: Visualization of the expected and actual VTT of container vessels at the HKP

serve that the distributions of both the expected and actual VTT for container vessels at the HKP roughly follow a Gaussian distribution, which is consistent with the results in Table 3.5 and 3.6, indicating that the peak of the expected VTT distribution is approximately 14 hours and the peak of the actual VTT is around 15 hours. Next, we

give an analysis of vessel departure delay at the HKP, which is evaluated by the difference between EDT and ADT. A negative value of vessel delay shows that a vessel departs latter than expected, while a positive value suggests that the vessel departs earlier than expected. The value of 0 indicates that the vessel departed on time. There are 69 data records categorized as "on-time". The vessel departure delay analysis results (in hours) for the other two classes are summarized in Table 3.7.

Table 3.7: VTT delay analysis in hours

Item	Total number of records	Minimum (hours)	Maximum (hours)	Median (hours)	Mean (hours)	Standard deviation (hours)
Early	7,637	0.02	118.25	3.45	7.44	6.02
Late	7,269	0.10	58.17	3.48	5.58	11.04
Early_container	5,261	0.16	52.02	3.27	4.93	5.40
Late_container	4,685	0.11	76.25	3.25	6.90	10.63



(a) Distribution of container VTT delay at the HKP in 2022

(b) Distribution of container VTT delay at the HKP in 2023

Figure 3.7: Visualization of the container VTT delay at the HKP

In Figures 3.7a and 3.7b, we observe that the container VTT delay data for the years 2022 and 2023 roughly follow a Gaussian distribution centered around zero. This is in line with the requirements for pre-processing the prediction set. Furthermore, we find that the VTT delay distribution for 2022 in Figure 3.7a aligns more closely with a Gaussian distribution compared to that of 2023 in Figure 3.7b. This

can be attributed to the 2022 dataset encompassing data from an entire year, thus being larger, while the 2023 distribution only contains data from the first three months. In addition, both 2022 and 2023's VTT delays for container vessels exhibit several instances of extreme delays, which are responsible for the outlier points observed in the distribution figures. Next, we analyze the VST at the HKP. Table 3.8 presents the numerical analysis of actual VST at the HKP.

Table 3.8: Actual VST statistic analysis

	Records number	Mean (hours)	Median (hours)	Maximum (hours)	Minimum (hours)	Standard deviation (hours)
Vessels in 2020	1,519	10.74	10.34	37.22	4.02	5.52
Vessels in 2021	4,298	11.91	11.01	58.21	4.35	6.73
Vessels in 2022	4,285	11.21	10.95	43.59	3.98	6.27
Vessels in 2023	1,680	10.41	10.21	32.51	4.40	5.33
Total vessels in the dataset	11,782	11.33	10.85	58.21	3.98	6.30

The complete dataset in Table 3.8 encompasses 11,782 vessel records, with fewer data points for 2020 and 2023 due to the dataset only covering four months of 2020 and three months of 2023. The overall average actual VST at the HKP is 11.33 hours, with a median of 10.85 hours. The longest actual VST observed is 58.21 hours, while the shortest is 3.98 hours and both of these records are associated with cargo vessels. The actual VST data also exhibit a standard deviation of 6.30 hours, indicating variability in the actual VST.

Comparing data by year, we can observe that in 2023, the actual VST for vessels shows the lowest mean, median, and standard deviation values compared to other years. This indicates an improvement in the port service efficiency of HKP in 2023 over the other years. This improvement is likely due to the impact of COVID-19 on port handling efficiency from 2020 to 2022, which saw a decline. However, in 2023, following the full reopening from COVID-19 related restrictions, the HKP handling efficiency rebounded. Conversely, 2021 records the highest mean, median, and standard deviation values, indicating the lowest efficiency in berth operations for HKP during that year. This reduction in efficiency can be attributed to the multiple impacts of the local pandemic outbreak on Hong Kong, which adversely affect the port operations in 2021 (Yan, Mo, X. Guo, Y. Yang, and S. Wang 2022). Next, we analyze the HKP expected VST, defined as the difference between each vessel's EDT and ATA. Like the actual VST analysis, the data is categorized into five groups. The analysis results are shown in Table 3.9.

Table 3.9: Statistical analysis of the expected VST at the HKP

	The number of records	Mean (hours)	Median (hours)	Maximum (hours)	Minimum (hours)	Standard deviation (hours)
Vessels in 2020	1,519	13.18	14.63	33.47	0.47	6.56
Vessels in 2021	4,298	14.32	13.97	34.85	0.35	6.91
Vessels in 2022	4,285	14.02	14.00	35.88	0.35	6.85
Vessels in 2023	1,680	13.10	12.85	35.87	0.35	6.06
Total vessels in the dataset	11,782	14.02	13.88	35.87	0.35	6.75

Table 3.9 offers a statistical analysis of the expected VST at the HKP from 2020 to 2023. Within this period, the data shows fluctuations but relatively similar average service time, ranging between 13.10 and 14.32 hours. The median values of service time closely align with the average, suggesting a balanced distribution of data. Meanwhile, the year of 2020 is associated with the highest median at 14.63 hours and the year of 2023 the lowest at 12.85 hours. The maximum service time peaks at 35.88 hours in the year of 2023, whereas the minimum service time consistently stood at 0.35 hours across all years. The standard deviation indicates the greatest variability in 2020 at 6.56 hours and the least in 2023 at 6.06 hours. Notably, the self-reported expected VST data for 2023 show the lowest mean, median, and standard deviation, similar to the actual VST scenarios. This pattern indicates an enhancement and optimization in vessel service management, leading to more efficient and consistent service time for vessels in 2023. Finally, we analyze the delay in VST,

which is calculated by subtracting the EDT data from the ADT data for each vessel. The results are presented in Table 3.10.

Table 3.10: VST delay analysis at the HKP

	The number of records	Mean (hours)	Median (hours)	Maximum (hours)	Minimum (hours)	Standard deviation (hours)
Vessels in 2020	1,519	-3.31	-1.85	17.21	-29.48	6.37
Vessels in 2021	4,298	-2.40	-1.41	26.68	-32.18	6.12
Vessels in 2022	4,285	-2.81	-1.76	18.18	-34.32	5.86
Vessels in 2023	1,680	-2.69	-1.55	14.12	-31.50	5.52
Total vessels in the dataset	11,782	-2.68	-1.58	26.68	-34.32	5.98

Table 3.10 presents a statistical analysis of VST delays at HKP from 2020 to 2023. A negative value indicates that the vessel departs eariler from the berth than scheduled. Analyzing the VST delay at the HKP from 2020 to 2023 as a whole, there is a discernible pattern of improvement in the service efficiency, with delays decreasing over the years. The mean value of the delay decreases from -3.31 hours in 2020 to -2.69 hours by 2023, indicating that the delays in VST are decreasing annually. This improvement is mirrored in the median delay times, decreasing from -1.85 to -1.55 hours during the same period. The standard deviation decrease from 6.37 to 5.52 hours over the years, indicating more consistent port service. Overall, from 2020 to 2023, there has been a general reduction in the VST delay at HKP, with mean and median delays decreasing over

the years. The maximum delay has significantly decreased, while the minimum delay has become slightly more negative. The standard deviation has gradually decreased, indicating a reduction in the variability of delay time.

3.3 COVID-19 Impact on Vessel Arrivals and Port Operations

3.3.1 Analysis of vessel arrivals and port operations at the HKP before and after the COVID-19 outbreak

The statuses of vessel arrival and port operation are crucial elements in port management. Specifically, "vessel arrival status" refers to the number of ship arrivals and the deviation in their scheduled arrival times at the port within a given period. Similarly, the "port operational status" indicates the efficiency of cargo handled by the port within a specific period (Authority 2023). In this study, We use four indicators to assess the HKP's vessel arrivals and port operations across various timeframes: the number of vessels of different types arriving at HKP, the total vessel NT, the average GT for container vessels, and HKP's throughput. The first indicator measures the vessel arrival status, while the remaining three assess the port operation status. In this subsection, a comprehensive evaluation of the vessel arrival and port operation status at the HKP before and after the COVID-19 outbreak is provided. Initially, vessel arrival data and port operation data at the HKP, ranging from January 2019 to December 2022, which cover the time periods prior to and follow-

ing the outbreak of COVID-19, are collected and analyzed.

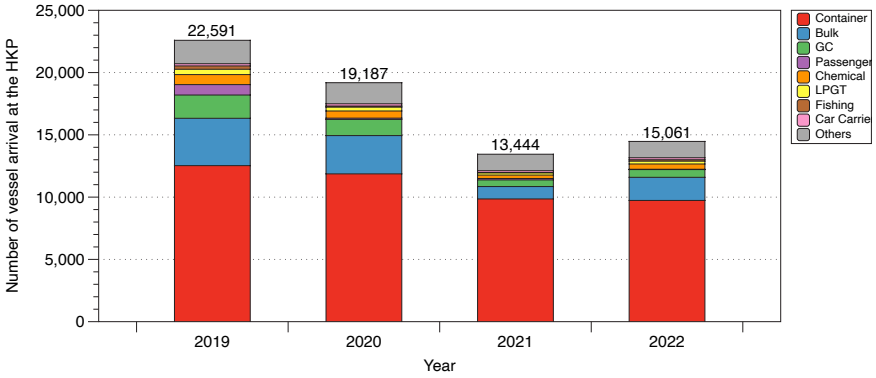
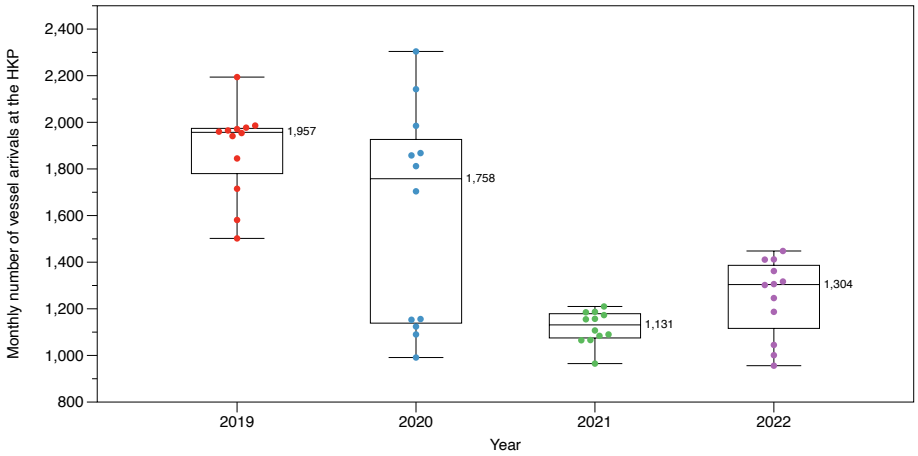


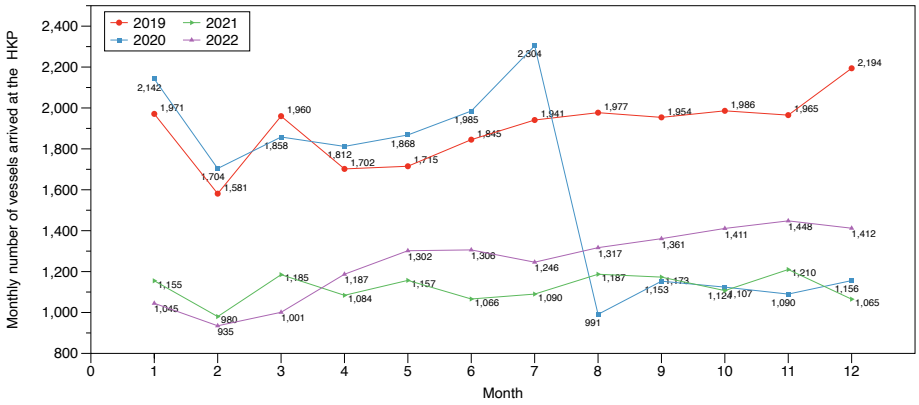
Figure 3.8: The annual number of vessels arriving at the HKP

Figure 3.8 depicts the annual amount of vessels arriving at the HKP in 2019, 2020, 2021, and 2022. The data reveals a downward trend in the number of vessel arrivals, which amounted to 22,591 in 2019, and decreased by 15.3% to 19,187 in 2020. In 2021, the decline in vessel arrivals was more substantial, recording a drop of 29.9% from 2020, with only 13,444 vessels arriving at the port. However, 2022 witnessed a modest improvement in the annual total vessel arrivals, registering an increase of 12.3% from 2021, with a total of 15,061 vessels arriving at the port. The decline in vessel arrivals at the HKP over the past two years could be attributed to the COVID-19 pandemic's unparalleled impact on the worldwide economy and commerce (Narasimha, Jena, and Majhi 2021; Notteboom, Pallis, and Rodrigue 2021). The crisis has led to significant disturbances to global supply chains, leading to a reduction in the demand for goods and services worldwide, which highly affected the number of vessel arrivals at the HKP. The difficulties faced by the shipping companies due to the pandemic-related regulations and restrictions further ex-

acerbated the decline in vessel arrivals. The modest increase in the amount of vessel arrivals in 2022 may indicate the gradual recovery of global trade and supply chains, as the pandemic related restrictions began to be lifted in Hong Kong from the end of April (Bureau 2022).



(a) Monthly number of vessel arrivals at the HKP[6]



(b) Monthly comparison of total vessel arrivals at the HKP

Figure 3.9: Monthly statistics of the total number of vessel arrivals at the HKP

Next, we will analyse the monthly status of the proportion of of

[6]Note: the numbers on the right of each box is the second quartile, i.e., median of the proportion of monthly vessel arrivals at the HKP.

vessel calls at the HKP in Figure 3.9. Figure 3.9a reveals that the effects of the COVID-19 are evident on monthly vessel arrival pattern at the HKP. In 2020, the range of monthly vessel arrivals exhibited greater variability compared to other years. Vessel arrivals during the beginning months of the year at the HKP were in a normal state. However, as the COVID-19 intensified and governments implemented various containment strategies, the maritime sector faced unprecedented challenges, including disrupted supply chains, labor shortages, and operational restrictions (Akyurek and Bolat 2020). These factors subsequently led to a decrease in vessel arrivals at the HKP, resulting in an expanded range of monthly vessel arrivals. The range in 2021 indicated a more unified pattern of vessel arrivals as the HKP was affected by the COVID-19 throughout the year. In 2022, the range of monthly vessel arrivals at the HKP increased compared to 2021, with an overall increase in the number of vessel arrivals. These findings suggest a gradual recovery following the beginning outbreak of the COVID-19, and that Hong Kong has also experienced a shift from the "zero-COVID" policy to "live with COVID" as pandemic related restrictions are gradually being lifted and the impact on ports is decreasing. When considering the median values for each year, there was a sharp decline in 2020 compared to 2019, which is as a result of the detrimental effects of the COVID-19 pandemic on the shipping industry. The sustained low median in 2021 suggests that the COVID-19 impact on the HKP had intensified, as reflected by the decline in the overall number of vessel arrivals. This situation is not unique to the HKP, as ports worldwide

have been similarly affected by the pandemic (Millefiori, Braca, Zissis, Spiliopoulos, Marano, Willett, and Carniel 2021). However, in 2022, the median showed a slight recovery compared to 2021, indicating a gradual recovery of vessel traffic at the HKP.

Figure 3.9b shows the trend of the monthly number of vessel arrivals at the HKP. The data for 2019, which represents pre-COVID data, shows a generally unified trend of high vessel arrivals throughout the year. There is a noticeable decline in arrivals in February, associated with the Chinese New Year holidays, which often leads to a slowdown in port activities. Starting in April, the number of vessel arrivals begins to recover and shows a gradual upward trajectory, culminating in a peak during December. This peak may be attributed to increased maritime traffic in preparation for the holiday season and year-end commercial activities. The overall pattern from the Figure 3.9b indicates a consistent level of vessel activity throughout the year 2019, accompanied by temporary surges during the summer and winter seasons. After the COVID-19 pandemic began, the number of vessel arrivals at the HKP was initially unaffected. Even though, there was an increase in vessel arrivals from March to July 2020, with July having the highest number of arrivals in 2020 at 2,304. The reasons for such notable circumstance could be attributed to the following factors:

1. Hong Kong and other port cities in China were able to quickly control the initial outbreak of the pandemic, resulting in fewer disruptions to shipping and trade (Grinter 2023d).

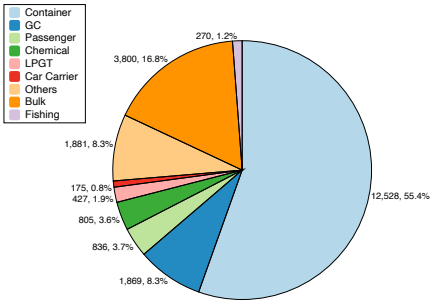
2. Before the onset of the pandemic's third wave in July 2020,

> Hong Kong allowed crew members to disembark at the HKP and return home without being subject to mandatory quarantine. Many vessels took advantage of this and stopped at the HKP to drop off crew members before continuing on their routes (Grinter 2023e; Grinter 2023g).

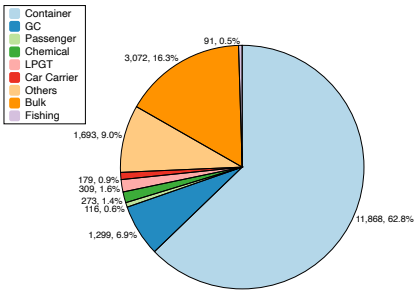
Subsequently, from August 2020 to March 2022, the pandemic had a significant impact on the amount of vessel arrivals, with the amount of arrivals being far lower than the same period in 2019. This period marked the peak of the pandemic's impact on the port. The significant decrease in the number of vessel arrivals at HKP in August 2020 can be attributed to the third wave of the COVID-19 pandemic that hit Hong Kong during that month. Clusters of COVID-19 infections were reported at the port, leading to partial shutdown of some of the port terminals for disinfection to prevent further spread of the virus (Grinter 2023c). This had a direct impact on the number of vessel arrivals at the port. To halt the continued spread of the COVID-19, the port authorities implemented several policies, including prohibiting non-working crew members from disembarking, requiring daily testing for all personnel, and temporarily suspending crew changes (Grinter 2023c; Post 2022).

Then, we analyze vessel arrivals at the HKP by vessel types. The annual number and proportion of different types of vessels arriving at the HKP are shown in Figure 3.10.[7] The analysis of vessel arrivals of different types of ships at the HKP from 2019 to 2022 reveals several trends in terms of the arrival number and proportion. As shown

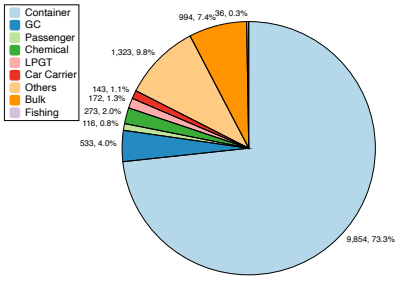
[7]GC is short for general cargo vessel and LPGT is short for liquefied petroleum gas tanker.



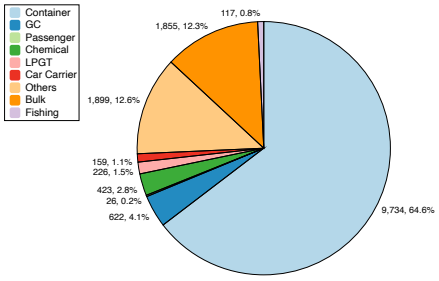
(a) Statistics of the number and proportion of ships arriving at the HKP by type in 2019



(b) Statistics of the number and proportion of ships arriving at the HKP by type in 2020



(c) Statistics of the number and proportion of ships arriving at the HKP by type in 2021



(d) Statistics of the number and proportion of ships arriving at the HKP by type in 2022

Figure 3.10: The annual number and proportion of different types of vessels arriving at the HKP

in Figure 3.10, container vessels consistently constituted the largest number and proportion of vessels arriving at the HKP throughout the four years, which echos the fact that the HKP primarily serves as a container port and emphasizes the importance of container shipping to Hong Kong's economy. Although the number of container vessels decreased from 12,528 arrivals in 2019 to 9,734 in 2022 due to the pandemic, the proportion of container vessels among all vessel arrivals increased. In 2019, container vessels accounted for 55.4% of total vessel arrivals, and by 2021, this proportion increased to its highest value at 73%. This trend demonstrates that amidst the pandemic, the relative frequency of non-container vessels arriving at ports diminished, while the proportion of container ships facilitating the conveyance of indispensable goods and cargo experienced an upturn. Bulk and GC vessels seem to be more heavily influenced by the COVID-19 pandemic compared to container vessels. Since the pandemic began, the number of both bulk and GC vessels arriving at the port has decreased. The greater impact on bulk and GC vessels compared to container vessels can be attributed to two main reasons. Firstly, container vessels operate with a higher degree of automation in loading and unloading operations, while bulk and GC vessels require more manual labor and are thus more vulnerable to the impact of the pandemic during during which there is a reduction in port personnel. Secondly, the items transported by container vessels are more time-sensitive than those transported by bulk or GC vessels. Goods transported by container vessels are given higher priority in port handling compared to goods carried by bulk

and general cargo vessels. Bulk and GC vessels are more susceptible to the effects of the COVID-19 crisis and the functioning of the port. Furthermore, the HKP has placed greater emphasis on ensuring the normal operation of container vessels during the pandemic period, given its role as a container port (HKSAR 2023). This circumstance has also affected the arrival of bulk and GC vessels. Passenger vessels have been severely impacted by the pandemic, as the majority of such vessels are tourist boats from Mainland China. The amount of passenger vessel arrivals significantly declined from 836 in 2019 to only 26 in 2022. Meanwhile, car carrier and chemical vessels are the least influenced by the pandemic, with relative stable numbers of vessel arrivals. This can be attributed to the fact that the arrivals of these vessel types are primarily driven by local consumer demand, rather than external factors such as travel restrictions or disruptions to global supply chains.

Then, we conduct an analysis of the monthly vessel arrival times for different types of vessels at the HKP under the impact of COVID-19. We classify vessel types into three categories: vessels unaffected by COVID-19 (car carrier), vessels affected by the COVID-19 with immediate impact (passenger, fishing), and vessels with delayed impact from the pandemic (container, bulk, GC, tanker, LGPT), in terms of their number of arrivals. The illustration and enumeration of these three categories are presented as follows:

Vessel arrivals with no impact from COVID-19:

The monthly vessel arrival data for car carriers at the HKP is shown in Figure 3.11, which indicates that car carriers have rarely been in-

fluenced by the COVID-19 and the related policies. In Hong Kong,

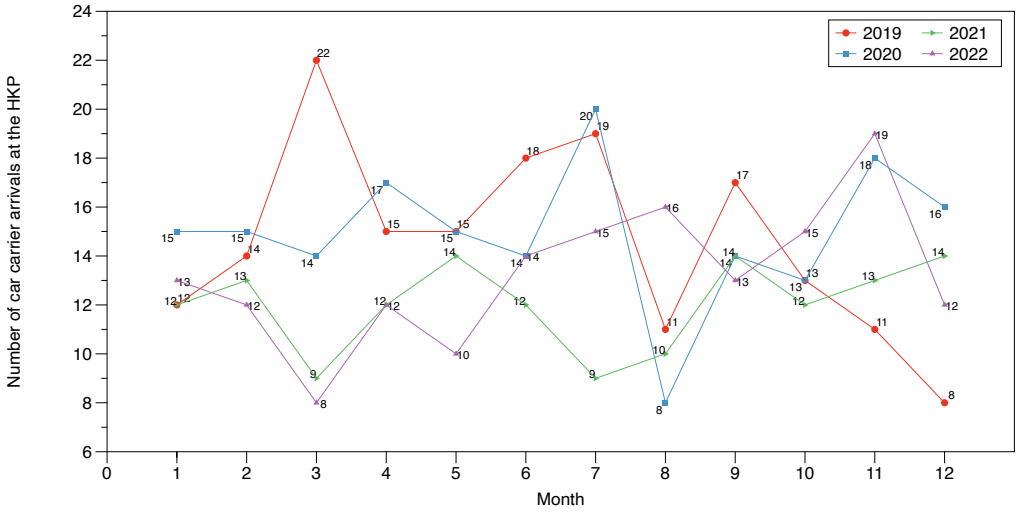


Figure 3.11: Monthly comparison of car carrier arrivals at the HKP

all brand-new vehicles are imported from overseas transported by car carriers. The sales volume of imported cars is influenced by brand marketing strategies and government transport policies rather than being significantly affected by the COVID-19 pandemic. Consequently, the number of car carrier arrivals at the HKP has suffered rare effect from the pandemic. Vessel arrivals with immediate impact from COVID-19.

The arrival numbers of both passenger and fishing vessels have been influenced by the COVID-19 pandemic, with an instantaneous manifestation. Figure 3.12 demonstrates a substantial decline in the amount of tourist vessels arriving at the port since February 2020. The passenger vessels that arrive at the HKP are primarily tourist ships from Guangdong Province and Macau. Since February 2020, as a result of the COVID-19 outbreak in major cities in the Mainland China, travel between Hong Kong and the Mainland China was impacted by poli-

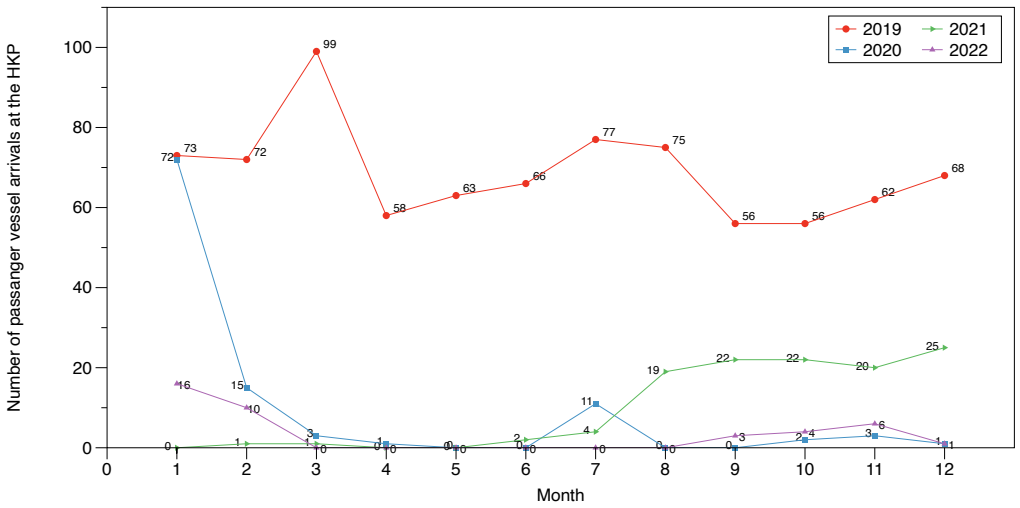


Figure 3.12: Monthly comparison of passenger vessel arrivals at the HKP

cies mandating quarantine, leading to a considerable decrease in the amount of passenger vessels operating and arriving at the port (Bureau 2022). From July 2021 to February 2022, passenger vessels arriving at the port were cruise ships catering to local Hong Kong residents engaged in high-sea staycation tourism, with no stopovers at other ports. However, after the fifth wave of COVID-19 outbreaks in March 2022, the number of ship arrivals once again fell to zero. Passenger vessels represent the tourism sector, which stands among the industries that have been most affected by the COVID-19 (Tsui, Fu, T. Chen, Lei, and H. Wu 2021). As shown in Figure 3.13, until May 2022, fishing and passenger vessels both experienced a decrease in the number of arrivals. Whereas from May 2022 afterwards, the number of fishing vessel arrivals began to recover and eventually returned to normal levels at the end of the year. This is likely due to the Hong Kong government's decision to lift additional quarantine measures for incoming vessels in May 2022 (Bureau 2022).

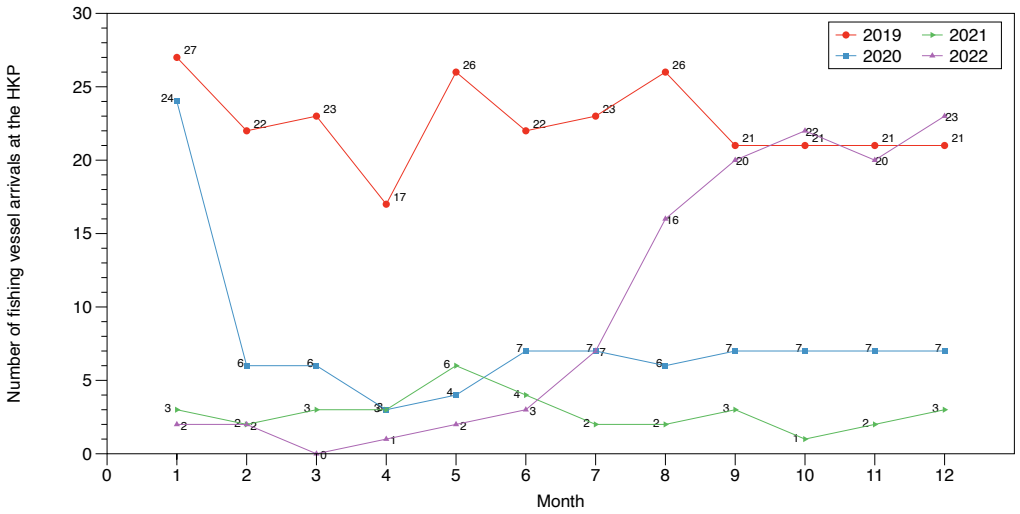


Figure 3.13: Monthly comparison of fishing vessel arrivals at the HKP

Vessel arrivals with delayed impact from the pandemic.

As a transit port, the number of vessels arriving at the HKP is not only related to local policies but also affected by external economic conditions. Five categories of vessels, namely container ship, GC ship, bulk carrier, LPGT, and oil tanker, have been impacted by the COVID-19 pandemic with a delayed manifestation. These five categories of vessels are able to be broadly divided into two groups: one related to cargo and the other related to energy. The monthly arrival circumstances of these five categories of vessels are depicted as follows.

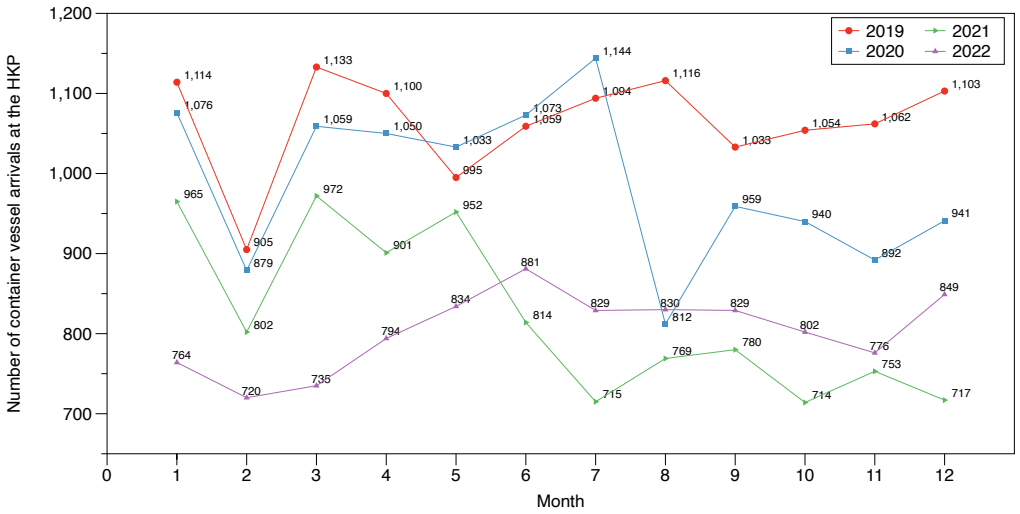


Figure 3.14: Monthly comparison of container vessel arrivals at the HKP

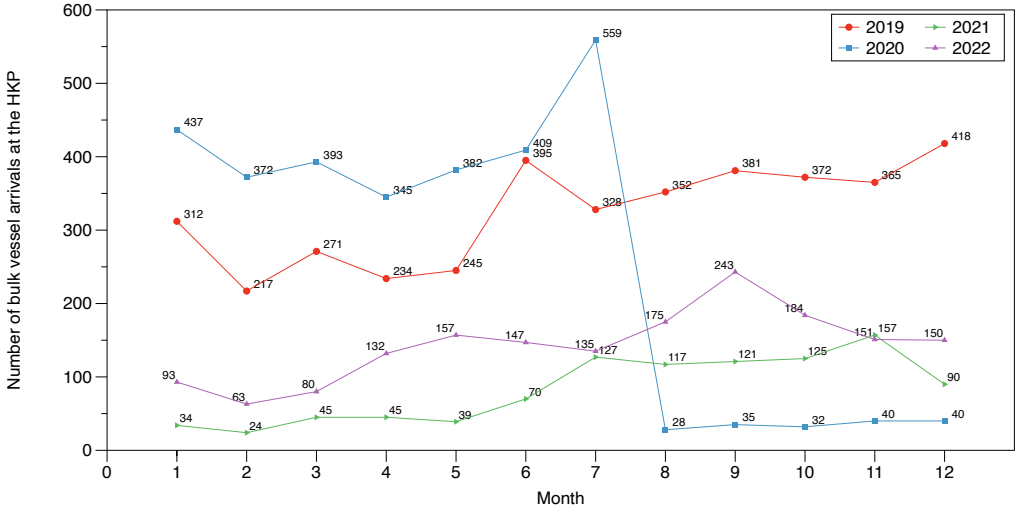


Figure 3.15: Monthly comparison of bulk carrier arrivals at the HKP

First, we analyze the vessel arrival patterns that are related to goods transportation. Figures 3.14 and 3.15 demonstrate a similar arrival pattern for bulk carriers and container vessels. Prior to July 2020, the number of arrivals of both categories was unaffected by the COVID-19 pandemic, and there was even a temporary growth

from May to July 2020 due to a brief increase in external demand and the HKP authorities' allowance of crew disembarkation (Bureau 2022; Post 2022). Starting from August 2020, there was a significant decline in vessel arrivals due to the implementation of pandemic prevention policies at the port and the related economic impact of the pandemic on vessels, leading to a reduction in demand (List 2023b; List 2023a). Subsequently, the number of arrivals of bulk carriers and container vessels at HKP remained low until June 2021. In June 2021, the port authorities eased its control requirements for crew members, allowing them to disembark after undergoing quarantine (Bureau 2022). Furthermore, in the second half of 2021, the global economy began to enter a recovery period, contributing to a gradual recovery of the number of arrivals of bulk carriers and container vessels. After Hong Kong transferred from "Zero-COVID" mode to "live with COVID" in May 2022, the growth rate of vessel arrivals was limited, and the amount of arrivals did not recover to pre-pandemic levels. This is attributed to several factors, including the lingering influences of the COVID-19 fallout in Mainland China and the reduced demand fromthe United States and Europe due to rampant inflation (Grinter 2023b). Another factor contributing to this is the impact of the ongoing Russia-Ukraine conflict on global trade (Grinter 2023a). The monthly arrival number of GC vessels at the HKP is presented in Figure 3.16. In 2019, the arrival numbers for GC vessels remained relatively stable from January to November, with a significant increase observed in December due to the approaching of Christmas and New Year. In 2020, unlike the con-

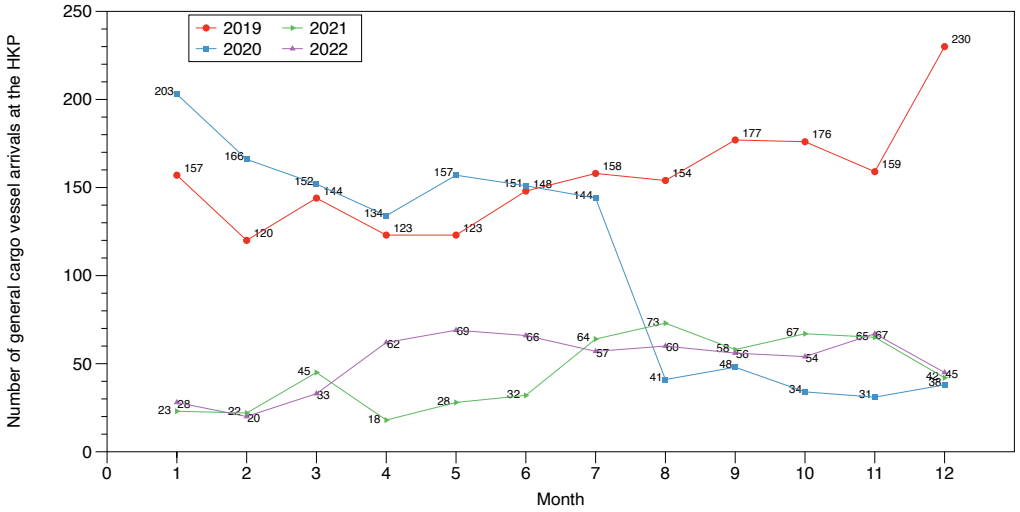


Figure 3.16: Monthly comparison of GC vessel arrivals at the HKP

tainer and bulk ships, the number of GC vessel arrivals from January to July is not significantly influenced. Instead, a substantial decline was observed in August, which was then followed by a low vessel arrivals period. Additionally, the December surge in previous years was not observed during the pandemic period.

Next, we analyze the arrival pattern of vessels that are related to energy. The number of monthly arrivals of LPGT is shown in Figure 3.17. LPGT in Hong Kong is mainly utilized for household and local transportation purposes (Statista 2023a). The COVID-19 and the related policy restrictions have adversely impacted residents' travel, leading to a significant decline in LPGT consumption in Hong Kong. From August 2020, coinciding with the pandemic's rapid spread and the related policy restrictions, LPGT consumption remained at a low level. However, since Hong Kong's reopening in May 2022, LPGT consumption has been gradually recovering and has nearly reached pre-pandemic levels as of September 2022. The

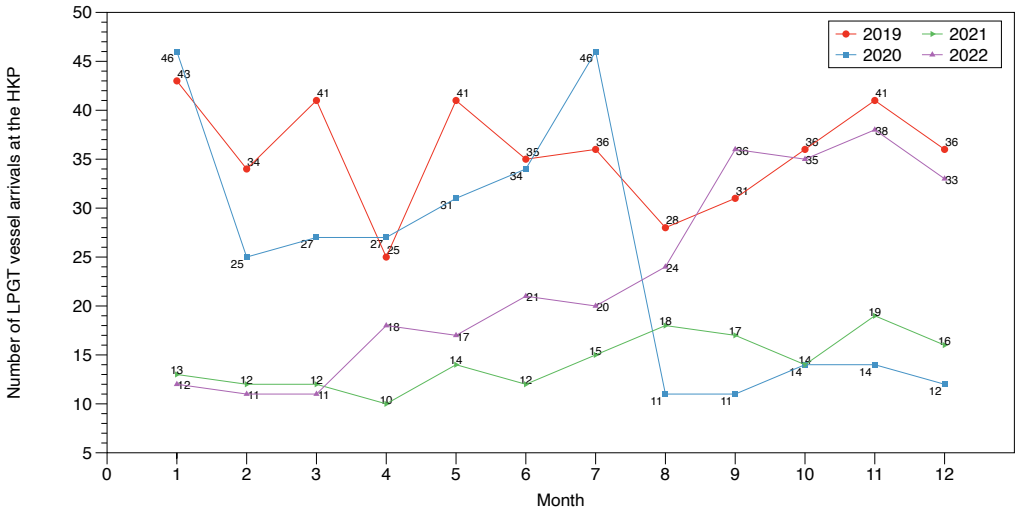


Figure 3.17: Monthly comparison of LPGT vessel arrivals at the HKP

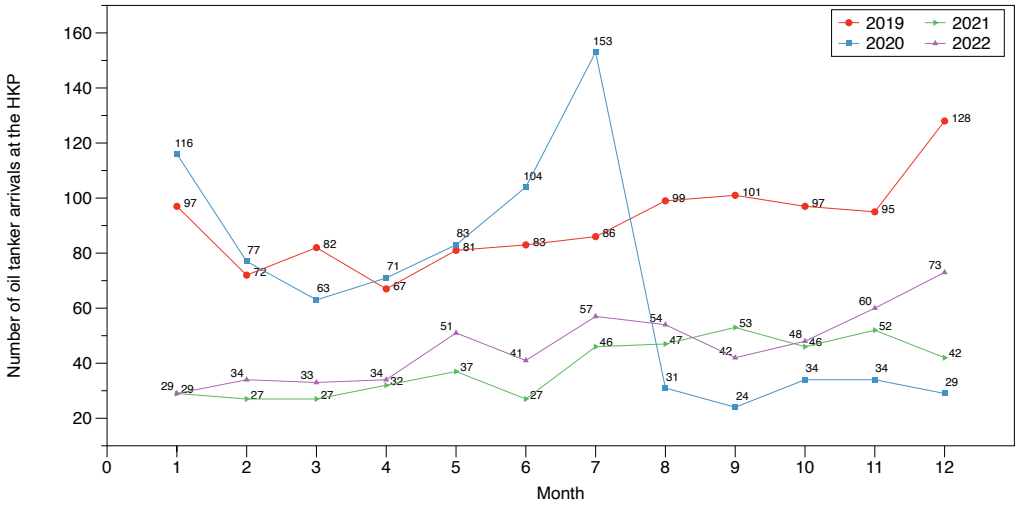


Figure 3.18: Monthly comparison of oil tanker arrivals at the HKP

arrival pattern of oil tankers, as shown in Figure 3.18, is critical to Hong Kong's oil supply, as the region relies entirely on foreign imports. Hong Kong's crude oil consumption comprises three components: aviation and shipping fuel consumption, household petroleum consumption, and petroleum consumption in civil industries (Statista 2023b). The correlation between the economy and oil consumption is highly positive. Hong Kong serves as a hub for air transportation, and the aviation industry was severely affected during the pandemic, with a significant reduction in flights and thus in aviation fuel consumption (Statista 2023b; Statista 2023a). Additionally, restrictions on residents' movements and a reduction in the number of vessel arrivals during the pandemic resulted in a decline in oil consumption. Since Hong Kong's reopening in May 2022, flight numbers have gradually recovered, and the local economy has resumed growth. Consequently, the arrival number of oil tankers has gradually increased, although they have not yet reached the pre-pandemic levels.

3.3.2 Port operation status at the HKP

This subsection delves into the effect of the COVID-19 on the operational aspects of the HKP. We assess the port operation status at HKP using three indicators: the total vessel NT, HKP's throughput, and the average GT for container vessels. Vessel NT is a measurement to calculate the capacity of a ship, which is calculated by the difference between the volume of a vessel's cargo spaces and its non-revenue-generating spaces. The total NT handled by a port within a

specific period reflects the port's capacity utilization and infrastructure capabilities, indicating how effectively it can accommodate and service vessels. A high total NT suggests the port is a significant node in global shipping networks, capable of handling large vessels efficiently (Branch 2012). Similarly, the throughput of a port reveals the volume of cargo handled, serving as a direct measure of the port's activity level, its economic impact, and its efficiency in cargo operations (Rødseth, Wangsness, and Schøyen 2018). Lastly, the average GT for container vessels provides insights into the size of container ships visiting the port and signifies the port's operational efficiency and infrastructure readiness to manage larger, more modern vessels (Tchang 2020). Collectively, these metrics shed light on the port's operational efficiency, its strategic importance in the maritime industry, and its ability to meet the demands of international trade efficiently (Tai, J. Guo, Guan, and Q. Shi 2021). We begin by investigating the total vessel NT at the HKP. The monthly comparison of total vessel NT is shown in Figure 3.19. The pattern of

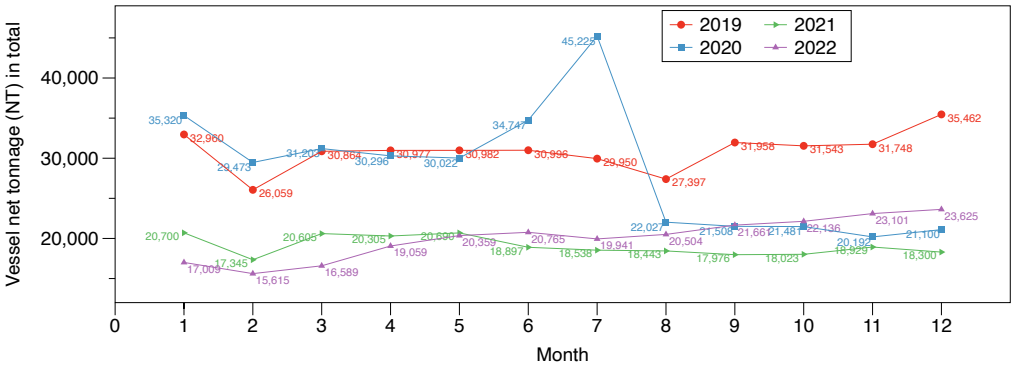


Figure 3.19: Monthly comparison of total vessels NT at the HKP

vessel NT presented in Figure 3.19 shows a similar trend to the to-

tal vessel arrival data presented in Figure 3.9b, which suggests that the total NT of a vessel is closely related to the number of vessel arrivals. Staff can also utilize NT as an indicator to assess vessel arrival circumstances. Next we analyze the average GT of arriving container vessels and the HKP container throughput. The annual container throughput of the HKP from 2019 to 2022 is depicted in Figure 3.20. From a yearly perspective as shown in the Figure 3.20,

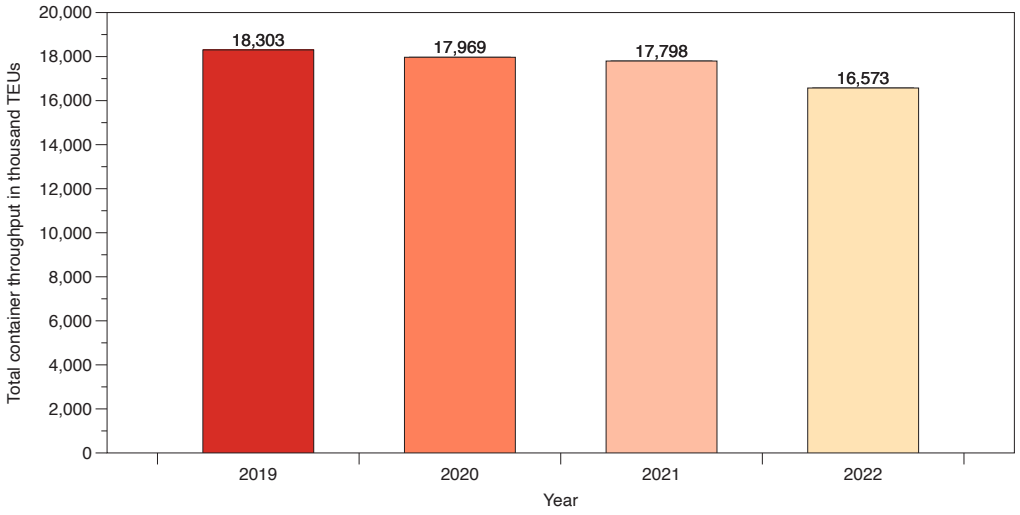


Figure 3.20: HKP throughput on a yearly basis

the HKP container throughput has gradually decreased from 2019 to 2021, albeit by a small margin. However, the data for 2022 shows a significant decline compared to 2021. Prior to 2019, the throughput of the HKP already begun to decrease, and in the first three years, the decline was mainly attributed to a decrease in Hong Kong's competitive edge compared to other ports in Mainland China instead of the COVID-19. HKP container throughput was normal in the first five months of 2022. Whereas starting from June, the port container throughput began to decline due to the external economic and policy

environment. The increase in interest rates in Europe and the United States, as well as the Russia-Ukraine conflict, have led to a decrease in foreign trade demand, resulting in a reduction in the volume of goods passing through the HKP as a transit hub, thereby reducing the port's throughput (Grinter 2023a; Grinter 2023f). The monthly

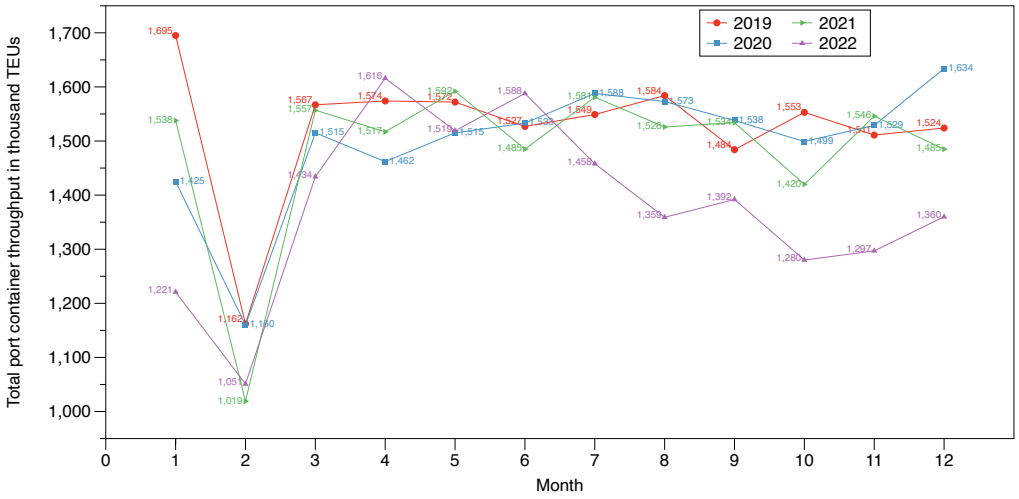


Figure 3.21: Monthly comparison of port total container throughput at the HKP

comparison of port container throughput at the HKP is shown in Figure 3.21. It is noticed that the container throughput at HKP is usually at its lowest in February each year, as the month has the fewest number of days and may include the Chinese New Year holiday. The port container throughput in April and May of 2020 decreased compared to the same period of 2019, primarily due to the COVID-19 outbreak that affected the port's operations and resulted to the work-from-home of staff, which lowered port handing efficiency (Post 2022). Finally, we analyze the average GT of container ships arriving at the HKP on a monthly basis. Container GT serves as an indicator of the cargo-carrying capacity of container ships. It is a

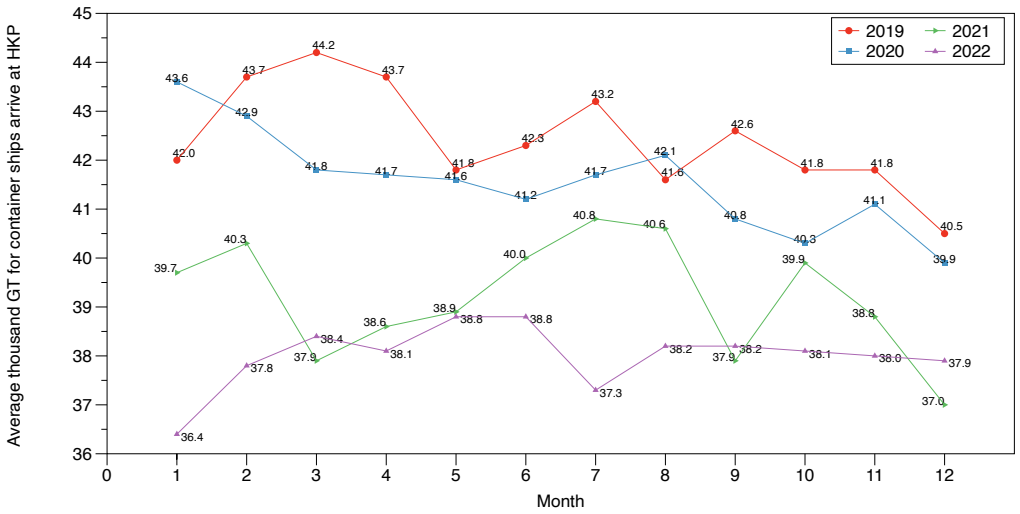


Figure 3.22: Monthly comparison of average GT of container ships at the HKP

widely used metric that measures a ship's capacity to carry cargo and helps evaluate the operational efficiency of ports in handling containerized cargo. The monthly comparison of the average GT of container ships at the HKP is shown in Figure 3.22. From 2020 afterwards to the end of 2022, the average GT of container ships arriving at Hong Kong gradually decreases, and despite the re-opening of Hong Kong in post-pandemic, the average GT of container ships has not yet fully recovered to its pre-pandemic levels. During the initial phase of the COVID-19, the crew members of long-haul container ships were inevitably susceptible to infection during voyages. If infected, container voyages would be delayed, resulting in an increase in the total shipping costs. In addition, Hong Kong's strict COVID-19 prevention measures, such as prohibiting crew members from going ashore and banning positive cases from approaching the port, could affect the scheduling of large container ships. Usually, large container vessels are deployed for long-distance journeys, such

as those connecting China with the United States or Europe. These factors led to longer completion times for large ships on the same routes when comparing the pandemic period to pre-pandemic times, resulting in a reduction in the number of large-tonnage ships arriving and a decrease in the average container GT. After HK reopening, the impact of the pandemic on crew members and cargo handlers continues, as crew members getting sick can still affect ship operations. Furthermore, in the latter part of 2022, due to economic recession and declined demand, the demand for large long-haul container ships decreased, and the average container GT remained to be low. The beam, length, and GT value of ships are positively correlated. Ultimately, the analysis of changes in the beam and length of ships visiting HKP under the COVID-19 impact is similar to the GT situation. To sum up, our analysis reveals that the COVID-19 and related restrictions result in a reduction on vessel calls across various types of vessels, as well as a decrease in port total throughput, and thus vessel total NT and vessel average GT.

3.3.3 Vessel arrival and container rate correlation analysis

Container freight rate is the price at which a TEU container is delivered from one port to its destination (Placek 2022). It is a shipping price index that reflects the trend of price changes in the global container transport market (Placek 2022). Container freight rates have been greatly affected by the COVID-19 pandemic, as it disrupted global supply chains and caused imbalances in the availability of shipping containers (Saeed, Nguyen, Cullinane, Gekara, and Chhetri

2023). In this section, we aim to examine the correlation between vessel arrival and departure times in the HKP and container freight rates. We primarily refer to the global container freight rate established by Statista,[8] which monitors the expenses of transporting 40-foot containers across eight significant global routes, covering both extended spot rates and short-term contract rates (Placek 2022). The global container freight rate index established by Statista represents the average container freight rate across the last full week's five business days of each month. This calculation method offers a clear, standardized benchmark, providing a market snapshot that is both recent and stable. It minimizes the impact of short-term price fluctuations and anomalous events that could distort a full-month average. Figure 3.23 presents the global container freight rate index from January 2019 to December 2022.

[8]https://www.statista.com/statistics/1250636/global-container-freight-index/.

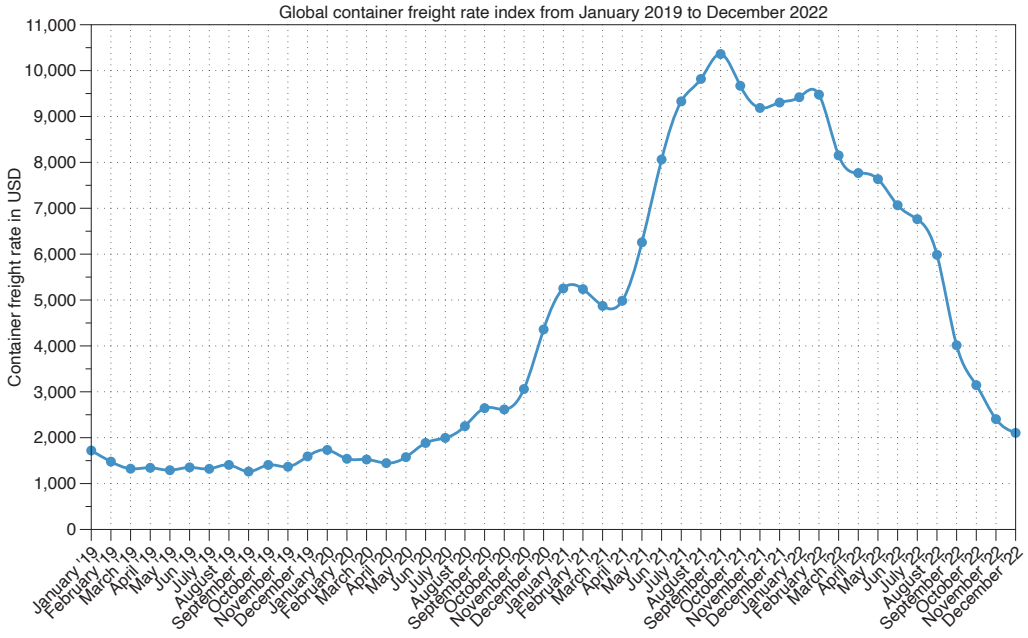


Figure 3.23: Global container freight rate overview

Container freight rate remained relatively stable in 2019. However, following the outbreak of the COVID-19, it exhibited an upward trend, reaching its peak in September 2021. Subsequently, it gradually declined and approached normal values by the end of 2022. At the onset of the pandemic, numerous factories in China and other parts of Asia experienced a decline in working efficiency, leading to a reduction in the volume of goods being transported. Subsequent to the worldwide lockdown in 2020, China succeeded in reinvigorating its economy at a swifter pace compared to the United States and Europe (Saeed, Nguyen, Cullinane, Gekara, and Chhetri 2023). Nevertheless, the lockdown in Europe and North America led to a significant backlog of shipping containers, causing disruptions and increased shipping costs. This unforeseen consequence had

far-reaching impacts on the global maritime transportation, as containers remain immobilized in these two continents. Consequently, China has experienced a shortage of shipping containers, and shipping companies have been unable to meet the demand due to the scarcity of available containers. Such circumstance has exacerbated the supply-demand imbalance, culminating in considerable surges in container freight rates. Shipping companies had to raise their prices to compensate for the expenses of repositioning empty containers and managing the disruptions caused by the pandemic (Jin, J. Chen, Z. Chen, Sun, and B. Yu 2022). In 2021, there was a significant surge in global freight rates, with prices reaching an all-time high of almost 10,400 U.S. dollars in September. However, by December 2022, the global freight rate index dropped to 2,100 U.S. dollars (Placek 2022). The shipping rates on the China-US and China-Europe routes have a significant impact on global container rates, as these trade lanes are among the busiest and most important in international trade (Slack and Gouvernal 2011). The HKP serves as the vital transit port for logistics and trade between China and the world. Extreme delay and turnaround time for container in the HKP not only affect local logistic operations but also undermine the stability of world supply chain (Akyurek and Bolat 2020). We use the Pearson correlation coefficient (PCC) and Spearmans correlation coefficient (SCC) as indicators to analyze the correlation of vessel arrival and departure at the HKP to global container freight rate. PCC and SCC are two statistical measurements of the linear correlation and similarity between two variables (J. Benesty, J. Chen, Huang, and Cohen

2009). Suppose we have n samples of the vessel arrival delay or vessel turnaround time as X and container freight rate Y. For each month $i, i = 1, \ldots, n$, X_i is the value of time factor for month i and $\bar{X}$ is the mean value of X. Y_i is freight rate for month i and $\bar{Y}$ is the average value of freight rate Y. R (X) is the ranked variables of X and R (Y) is the ranked variables for Y. The definitions of PCC and SCC are as follows:

$$PCC = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2} \sqrt{\sum_{i=1}^{n} \left(Y_i - \bar{Y}\right)^2}}, \tag{3.6}$$

$$SCC = r_s = \frac{\text{cov}(\text{R}(X), \text{R}(Y))}{\sigma_{\text{R}(X)} \sigma_{\text{R}(Y)}} = 1 - \frac{6 \sum d_i^2}{n\left(n^2 - 1\right)}, \tag{3.7}$$

where $cov(X, Y)$ is short for covariance of X and vector Y, σ_X is the standard deviation of X, $\mathbb{E}(X)$ is the expectation of X, μ_X is the mean of X, $d_i = \text{R}\left(X_i\right) - \text{R}\left(Y_i\right)$ is the difference between the two ranks of each variables. The range of PCC and SCC values is from -1 to 1. A value of -1 represents a perfect negative correlation, while a value of 0 indicates no correlation, and a value of 1 indicates a perfect positive correlation between the variables being studied. PCC is ideal for quantifying the linear correlation between two variables, offering insights that are scale and unit independent. Conversely, SCC offers a non-parametric alternative, robust against outliers and capable of identifying monotonic relationships in variables, irrespective of linearity. Leveraging these advantages, we employ the PCC and SCC indicators to more precisely depict the relationship between global freight rates, arrival delays, and the turnaround time

of all vessels, including container vessels at the HKP, as well as examining the correlation between HKP's total throughput and global freight rates, with our analysis encompassing related data from January 2019 to December 2022. The analysis results are shown in Tables 3.11 and 3.12.

Table 3.11: Correlation analysis between the global freight rate and the vessel arrival status at the HKP

Indicators	Arrival time delay of all types of vessels	Container vessel arrival delay	Turnaround time of all types of vessels	Container vessel turnaround time	HKP total throughput
PCC value	0.675	0.774	0.569	0.814	-0.164
SCC value	0.690	0.741	0.558	0.817	-0.152

Table 3.11 demonstrates a positive correlation between the arrival time delay and turnaround time of containers at the HKP and global freight rate. Conversely, there is a weak negative correlation between the throughput of the HKP and the global freight rate. Furthermore, container vessel arrival delay and turnaround time are more closely related to container rate compared to all types of vessel. Subsequently, we examine the correlation between the global freight rate and vessel arrival status at the HKP amidst the pandemic. The timeframe spanning from February 2020 to May 2022 is designated as the COVID-19 period for this analysis. The analysis results are shown in Table 3.12. Table 3.12 discloses that throughout the COVID period, a positive correlation is observed between the vessel arrival delay and turnaround time of container at the HKP and the global freight

Table 3.12: Correlation analysis between the global freight rate and the vessel arrival status at the HKP during the COVID-19 period

Indicators	Vessel arrival delay	Container veesel arrival delay	Vessel turnaround time	Container vessel turnaround time
PCC value	0.601	0.715	0.623	0.835
SCC value	0.632	0.744	0.657	0.841

rate. This observation aligns with the conclusions drawn from Table 3.11. This finding suggests that widespread container vessel arrival delay and prolonged turnaround time at the HKP exhibit a positive correlation with the escalation of global container rates. Several factors can attribute to this positive correlation. For example, when delays or prolonged turnaround time occur in the HKP, container vessels may have to spend more time waiting for cargo loading and unloading, which could affect the operational plans and transport capacity of shipping companies. This could result in further vessel arrival and departure delays, leading to the disruption of container liner shipping and global supply chain. Furthermore, as an essential transit port, the HKP plays a crucial role as a central hub for long-haul routes connecting China with the United States and Europe as well as other distant regions. As a key shipping hub, any delays or disruptions at the HKP can have ripple effects on other ports and regions as well as the global container rate.

Chapter 4

Vessel Time Prediction: Arrival, Turnaround, and Service Estimations

4.1 Predictive Model Introduction

ML is a specialized branch of artificial intelligence that centers on the use of data and algorithms to simulate the process of human learning (Z.-H. Zhou 2021). An ML model autonomously extracts knowledge within the dataset and employs these latent patterns to predict future data. Typically, ML involves three primary components: data preparation, algorithmic representation and model optimization (Z.-H. Zhou 2021; Filom, Amiri, and Razavi 2022). Data preparation refers to the process of preparing the dataset for the training. Algorithmic representation denotes the structure and formulation of the algorithms used in the ML process. Lastly, model optimization involves fine-tuning the model's hyperparameters to achieve the most favorable outcome. Tabular data serves as the fundamental

data format for port operations (Filom, Amiri, and Razavi 2022), in which each row presents an observation or a sample and every column represents a feature. Tree-based models, such as XGBoost and random forest (RF), are often utilized for predictive analysis. They tend to be more adept than neural networks at extracting valuable features and information from the dataset through techniques like bagging and ensemble learning (Grinsztajn, Oyallon, and Varoquaux 2022). The performance of these tree-based models often exceeds that of neural networks in many scenarios, especially when dealing with tabular data. In our research, our prediction dataset is a prime example of the tabular data format, thus we mainly use tree-based model to predict vessel related time. In addition, for comparison of models performance, we utilize various other prediction models including CART, XGBoost, RF, back propagation neural networks (BPNN), light gradient-boosting machine(LightGBM), LSTM to forecast the target value. We will introduce prediction models in detail regarding model definition, starting with an introduction to the CART model.

4.1.1 Introduction to basic prediction models

CART, standing for Classification and Regression Tree, is a methodology used for constructing both RF and XGBoost. This algorithm is capable of generating models for both types of trees. In our project, we are focusing on regression problems, and the CART model serves as a foundational element for the subsequent XGBoost model. This subsection will delve into the specific composition of the CART re-

gression model, exploring how it functions and its application in our regression-focused project. Suppose our dataset D has n samples and each sample has m features: $D = \{(\mathbf{x}_i, y_i), i = 1, \ldots, n\}, \mathbf{x}_i \in R^m, y_i \in R$, where $\mathbf{x}_i$ is the vector of features for sample i and y_i is the label or the corresponding real target value.

The process of splitting the tree begins at the root node. Initially, a feature m_i is chosen along with its corresponding value s_i. This pair, represented as (m_i,s_i), serves as a potential splitting point. Utilizing this point, the entire dataset D is divided into two distinct sub-regions, R_1 and R_2 and these two areas are expressed as: $R_1 = \{y_i \mid x_{i,m_i} \leq s_i\}, R_2 = \{y_i \mid x_{i,m_i} > s_i\}$. In the CART model, the average target values for all samples in regions R_1 and R_2 are determined and assigned as the predicted targets for the samples in their respective regions. The effectiveness of each candidate splitting point is evaluated by calculating the sum of the MSE for both R_1 and R_2. The algorithm iterates over all features and their corresponding values to generate various candidate splitting points. The aim is to identify the splitting point that results in the lowest combined MSE for the two regions. This optimal pair, which minimizes the sum of MSE, is then chosen as the final point to split the current node. These steps are repeated for each node until a pre-defined stopping condition for halting the growth of the tree is met. This systematic approach ensures the development of an efficient and accurate model (Lewis 2000).

To convey the process of constructing a CART regression tree more precisely, we can present it in mathematical terms as follows:

1. Beginning at the root node, a feature-value pair (m_1, s_1) is chosen to divide the dataset D into two distinct regions, R_1 and R_2.

2. Calculate the mean targets of samples in the two sub-regions R_1 and R_2, which are denoted by C_1 and C_2, respectively:

$$c_1 = \frac{1}{n_1} \sum_{x_i \in R_1} y_i, c_2 = \frac{1}{n_2} \sum_{x_i \in R_2} y_i. \tag{4.1}$$

3. Iterate through all features d_i and their corresponding values s_i to identify the split pair that minimizes the loss function in Equation (4.2). The optimal split pair of the current node is denoted as (m^*, s^*).

$$(m^*, s^*) = \min_{m_i, s_i} \left\{ \min_{x_i \in R_1(m_i, s_i)} (y_i - c_1)^2 + \min_{x_i \in R_2(m_i, s_i)} (y_i - c_2)^2 \right\}. \tag{4.2}$$

4. Use the optimal splitting pair (m^*, s^*), the samples are divided into two new areas. These two new areas are defined as follows:

$$R_1(m^*, s^*) = \{y_i \mid \boldsymbol{x}_{i,m_i} \leq s^*\}, R_2(m^*, s^*) = \{y_i \mid \boldsymbol{x}_{i,m_i} > s^*\}. \tag{4.3}$$

5. Carry out steps 1 and 2 on each node until any one of the predefined tree growth conditions is met, and no further splitting of nodes is allowed. The nodes in the final layer become leaf nodes. As a result, the entire training set is divided into K regions, $R_1, \ldots\ldots R_k$, where k also represents the number of leaf nodes. The model generated from the iterations can be written

as:

$$f(\boldsymbol{x}) = \sum_{i=1}^{K} c_i \mathbb{I}\left(x \in R_i\right), \tag{4.4}$$

where $k = 1, \ldots, K$, and $\mathbb{I}$ is an indicator function with the following form:

$$\mathbb{I} = \begin{cases} 1 & \text{if } (\boldsymbol{x} \in R_k) \\ 0 & \text{if } (\boldsymbol{x} \notin R_k)\,. \end{cases} \tag{4.5}$$

However, traditional CART models suffer from the problem of overfitting, leading to weak generalization capability (Breiman 2001), as they are sensitive to extreme data and subtle changes. To overcome this issue, a bootstrap aggregating (bagging) method is proposed to create divergence in the training set by using an ensemble of CART models to construct a unified model. The basic idea of CART with the bagging method is presented as follows:

1. Suppose that we have an original training set with n samples. To form a bootstrap sample, n samples from the original training set are randomly extracted with replacement. Then, this process is repeated k times (Breiman 2001), and we have a total of k bootstrap samples after the resampling process (Breiman 2001).

2. Train k CART models using the K bootstrap samples. In the regression problem, the final output is given by averaging the outputs of the k CART models.

The RF improves on the bagging method based on the CART model.

The only difference between RF and CART models with bagging is the manner in which each node in the tree is split. The optimal split pair is selected from a random subset of features instead of all of the features in the RF model, and the number of selected features is preset. With this characteristic, the RF model can handle high-dimensional data without feature selection and is more robust against overfitting (Breiman 2001).

In maritime studies, the RF model is widely used for vessel fuel consumption prediction, ship energy efficiency prediction, and the efficient inspection of vessels, among other topics (J. Yu, G. Tang, Song, X. Yu, Qi, D. Li, and Y. Zhang 2018; Y. Yang, Y. Liu, G. Li, Z. Zhang, and Y. Liu 2024).

XGBoost is a powerful tree-based ML algorithm built on the boosting framework. Unlike RF, which trains trees independently and averages their predictions, XGBoost employs a sequential boosting approach, where each tree corrects the errors of its predecessors. This leads to higher predictive accuracy and better handling of complex patterns. Additionally, XGBoost incorporates advanced regularization techniques, such as L1 and L2 penalties, to prevent overfitting, and it efficiently handles missing values and large-scale datasets through optimized parallel computing (T. Chen, T. He, M. Benesty, Khotilovich, Y. Tang, Cho, K. Chen, Mitchell, Cano, T. Zhou, et al. 2015). XGBoost achieves this by integrating numerous foundational CART models, creating a robust prediction model.

In regression task, XGBoost comprises of K basic CART regression models and functions cumulatively to predict outcomes. We can

represent the model's output as:

$$\hat{y}_i = \phi\left(\mathbf{x}_i\right) = \sum_{t=1}^{K} f_t\left(\mathbf{x}_i\right), f_t \in F, \tag{4.6}$$

where $\hat{y}_i$ is the predicted value by the XGBoost, f_t is the t-th basic CART tree, F is the set of functions for all K CART trees. And the training loss can be represented as:

$$L = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i\right). \tag{4.7}$$

A practical choice for the loss function is the MSE, where:

$$\sum_{i=1}^{n} l\left(y_i, \hat{y}_i\right) = \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2. \tag{4.8}$$

The accuracy of a prediction model is determined by both its bias and variance (Z.-H. Zhou 2021). The bias of the model is represented by the loss function, L, while Ω, is utilized to penalize the complexity of the model and evaluate the variance of the output. In this way, the object function (Obj) of the XGBoost is:

$$Obj = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i\right) + \sum_{t=1}^{K} \Omega\left(f_t\right). \tag{4.9}$$

The first term in Eq. (4.9) is the total training loss of n samples in the training set The second term in Eq. (4.9) represents the sum of the complexities of K trees. The complexity for each individual tree

is expressed as follows:

$$\Omega\left(f_t\right) = \gamma T_t + \frac{1}{2}\lambda \sum_{j=1}^{T_t} w_j^2, \tag{4.10}$$

where γ and λ are pre-set hyperparameters, T_t is the number of leaves for t-th tree and w_j is the weight in the j leave. To minimize the object function in Eq. (4.9), we cannot directly implement the gradient descent method as is traditionally utilized in boosting model. Instead, we formulate and train the model in an additive approach. Suppose the model after $t,\ t = 1, \ldots, K$ iterations, the XGBoost model currently has t trees and the prediction value for the i-th sample by the current t trees is:

$$\hat{y}_i^{(t)} = \sum_{k=1}^{t} f_k\left(\mathbf{x}_i\right) = \hat{y}_i^{(t-1)} + f_t\left(\mathbf{x}_i\right). \tag{4.11}$$

In this way, we can rewrite our object function in Eq. (4.9) as:

$$\begin{aligned} Obj^{(t)} &= \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{t} \Omega(f_k) \\ &= \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{t-1} + f_t\left(\mathbf{x}_i\right)\right) + \sum_{k=1}^{t} \Omega\left(f_k\right) \\ &= \sum_{i=1}^{n} \left(y_i - \left(\hat{y}_i^{(t-1)} + f_t\left(\mathbf{x}_i\right)\right)\right)^2 + \Omega\left(f_t\right) + \sum_{k=1}^{t-1} \Omega\left(f_k\right). \end{aligned} \tag{4.12}$$

Recall that the second order Taylor expansion:

$$f(x + \Delta x) \simeq f(x) + f'(x)\Delta x + \frac{1}{2} f''(x)\Delta x^2. \tag{4.13}$$

Following the rule, by viewing $\hat{y}_i^{t-1}$ as x and $f_t(x_i)$ as Δx, Eq. (4.12) can be rewritten as:

$$Obj^{(t)} \simeq \sum_{i=1}^{n} \left[L\left(y_i, \hat{y}_i^{t-1}\right) + g_i f_t(\mathbf{x}_i) + \frac{1}{2} h_i f_t^2(\mathbf{x}_i) \right] + \Omega(f_t) + \sum_{k=1}^{t-1} \Omega(f_k), \tag{4.14}$$

where g_i and h_i are the first and second order gradients of the Eq. (4.12): $g_i = \partial_{\hat{y}^{(t-1)}} L\left(y_i, \hat{y}^{(t-1)}\right) = 2\hat{y}_i^{(t-1)} - 2y_i$, $h_i = \partial^2_{\hat{v}^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right) = 2$. When training the t-th iteration, as $\hat{y}_i^{(t-1)}$ has already been determined, the first term in Eq. (4.14): $l\left(y_i, \hat{y}_i^{t-1}\right)$, which is the training loss of the $(t-1)$-th iteration, is a constant. And the last term in Eq. (4.14): $\sum_{k=1}^{t-1} \Omega(f_k)$, which represents the total penalty complexity of previous $t-1$-th iterations, is also a constant. In this way, we can rewrite the approximated object function in Eq. (4.14) as:

$$\begin{aligned} Obj^{(t)} &= \sum_{i=1}^{n} \left[g_i f_t(\mathbf{x}_i) + \frac{1}{2} h_i f_t(\mathbf{x}_i)^2 \right] + \Omega(f_t) \\ &= \sum_{i=1}^{n} \left[g_i f_t(\mathbf{x}_i) + \frac{1}{2} h_i f_t(\mathbf{x}_i)^2 \right] + \gamma T_t + \frac{1}{2} \lambda \sum_{j=1}^{T_t} w_j^2. \end{aligned} \tag{4.15}$$

Suppose the sample set in leaf j is defined as:

$$I_j = \{ i \mid q(\mathbf{x}_i) = j \}, \tag{4.16}$$

where $q(\mathbf{x}_i)$ is the given fixed tree structure, we can rewrite Eq.

(4.15) as:

$$Obj^{(t)} = \sum_{j=1}^{T_t} \left[\left(\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T_t. \tag{4.17}$$

As the tree structure $q(\mathbf{x}_i)$ is fixed, $\sum_{i \in I_j} g_i$, $\sum_{i \in l_j} h_i$ and T_t are also fixed. To obtain the optimal w_j^* for leaf j, we can set the first derivative of the objective function to be 0, and the optimal value of w_j^* can be derived as follows:

$$w_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda}. \tag{4.18}$$

By substituting Eq. (4.18) into the objective function Eq. (4.17), we obtain the following optimal value of the objective function:

$$Obj^{(t)*} = -\frac{1}{2} \sum_{j=1}^{T_t} \frac{\left(\sum_{i \in I_j} g_i \right)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T_t. \tag{4.19}$$

During the derivation of the optimal value for the object function, we make the assumption that the tree structure $q(\mathbf{x}_i)$ is predetermined. Therefore, our next step involves determining the tree structure of the XGBoost model under this optimal value. Practically, the XGBoost algorithm implements a greedy method that starts from a single node and progressively adds branches to the tree to form its structure. Suppose I_L and I_R are the samples of the left and right nodes of the tree after splitting and $I = I_L \cup I_R$ is the set of nodes before splitting, then we can write the object function before and after the splitting as:

before splitting:

$$Obj_{L+R}^{(t)} = -\frac{1}{2}\frac{\left(\sum_{i\in I} g_i\right)^2}{\sum_{i\in I} h_i + \lambda} + \lambda, \tag{4.20}$$

after splitting:

$$Obj_{L}^{(t)} + obj_{R}^{(t)} = -\frac{1}{2}\left[\frac{\left(\sum_{i\in I_L} g_i\right)^2}{\sum_{i\in I_L} h_i + \lambda} + \frac{\left(\sum_{i\in I_R} g_i\right)^2}{\sum_{i\in I_R} h_i + \lambda}\right] + 2\lambda. \tag{4.21}$$

Then, the gain of the splitting is expressed as:

$$\begin{aligned}\text{gain} &= obj_{L+R}^{(t)} - \left(obj_{L}^{(t)} + obj_{R}^{(t)}\right) \\ &= \frac{1}{2}\left[\frac{\left(\sum_{i\in I_L} g_i\right)^2}{\sum_{i\in I_L} h_i + \lambda} + \frac{\left(\sum_{i\in I_R} g_i\right)^2}{\sum_{i\in I_R} h_i + \lambda} - \frac{\left(\sum_{i\in I} g_i\right)^2}{\sum_{i\in I} h_i + \lambda}\right] - \gamma.\end{aligned} \tag{4.22}$$

When splitting the node, we consider all candidates that make the splitting gain in Eq. (4.22) larger than 0 and select the value of features that corresponding to the largest value of gain in Eq. (4.22) to split the node.

Building on the strengths of XGBoost, we next introduce LightGBM, another widely used tree-based ML algorithm known for its efficiency in prediction tasks. Compared to XGBoost, LightGBM leverages a histogram-based approach and a unique leaf-wise tree growth strategy, allowing it to train faster and handle large datasets more efficiently. Additionally, LightGBM is optimized for lower memory usage and better scalability, making it particularly suitable for high-dimensional data (Ke, Meng, Finley, T. Wang, W. Chen, Ma, Ye, and T.-Y. Liu 2017).

Similar to the XGBoost model, LightGBM consists of K basic CART regression models, and the model's output can be represented as:

$$\hat{y}_i = \phi\left(\mathbf{x}_i\right) = \sum_{t=1}^{K} f_t\left(\mathbf{x}_i\right), \tag{4.23}$$

where $\hat{y}_i$ is the value predicted by LightGBM, f_t is the t-th basic CART tree. The training loss is expressed as:

$$L = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i\right). \tag{4.24}$$

A common loss function selection is the MSE, where:

$$L = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2, \tag{4.25}$$

and n represents the number of training samples in the current node during a split decision. To determine the features that most effectively separate the data and predict the target variable, we need to split the tree to minimize the training loss (Z.-H. Zhou 2021). For an individual CART tree, the loss of node j before a split is represented as:

$$L_j = \sum_{x_i \in \text{leaf}\, j} l\left(A_j, y_j\right), \tag{4.26}$$

where A_j is the average value of the prediction value on node j and y_j is the prediction target of node j in the tree. The improvement from the split on the CART tree's loss function, mirroring the gain

in LightGBM's split decision, is quantified as:

$$\Delta \text{ loss } = L_j - L_{\text{left}} - L_{\text{right}}\,, \tag{4.27}$$

where L_{left} denotes the loss value of the left node and L_{right} represents the loss value of the right node after splitting. Thus, the objective function for each split can be expressed as:

$$\begin{aligned} \underset{f,t}{\text{argmax}} \Bigg(& \frac{1}{n} \sum_{x_i \in \text{ leaf } j} (y_j - A_j)^2 \\ & - \frac{1}{n_{\text{left}}} \sum_{x_i \in \text{leaf}_{\text{left}}} (y_j - A_{\text{left}})^2 \\ & - \frac{1}{n_{\text{right}}} \sum_{x_i \in \text{leaf}_{\text{right}}} (y_j - A_{\text{right}})^2 \Bigg) \end{aligned} \tag{4.28}$$

where A_{left} is the average value of the prediction value in the left node after splitting and A_{right} is the average value of the prediction value in right node after splitting. n_{left} is the number of samples in left node and n_{right} is the number of samples in right node after splitting. Equation (4.28) identifies the optimal split point that maximizes the reduction of training loss during the tree's branching. The optimization principle of LightGBM's loss function closely aligns with the foundational strategy of Gradient Boosted Decision Trees (GBDT). LightGBM utilizes the negative gradient of the loss function to optimize the objective function. This effectively approximates the residual values for the current model. During each iteration, this approach fits a new decision tree to the approximated gra-

dients. A new submodel is added in each iteration, with the original model remaining unchanged. Suppose $f_i(x)$ is the ith submodel in the ith iteration. Thus, the ensemble LightGBM model $F_i(x)$ after i iterations is written as:

$$F_i(x) = \partial_0 f_0(x) + \partial_1 f_1(x) + \cdots + \partial_i f_i(x), \tag{4.29}$$

where ∂_i is fixed parameters. The loss function is written as $L\left[F_i(x), Y\right]$ and by comparing with the previous iteration we get:

$$L\left[F_i(x), Y\right] < \left[F_{i-1}(x), Y\right]. \tag{4.30}$$

This method ensures continuous improvement of the loss function, leading to predicted values that converge closer to the ground truth. (Ke, Meng, Finley, T. Wang, W. Chen, Ma, Ye, and T.-Y. Liu 2017). Furthermore, LightGBM sets itself apart from traditional GBDT and other tree-based models by utilizing a leaf-wise growth strategy, in contrast to the level-wise growth strategy typical in GBDTs. This method prioritizes splitting the node that maximizes loss function reduction, thereby concentrating the model's development on the most beneficial segments of the decision tree. Unlike the level-wise strategy which evaluates all possible split points across every level of the tree simultaneously, LightGBM's leaf-wise method concentrates on identifying the optimal split point for each leaf individually. This targeted approach enables LightGBM to achieve greater efficiency and accuracy, as it directly aligns with the framework's overarching goal of minimizing the loss more rapidly. Additionally, by concentrating growth on the most informative nodes, LightGBM requires less

data to achieve comparable or superior performance levels, thereby enhancing its scalability and effectiveness on large datasets. To mitigate the risk of overfitting, LightGBM incorporates mechanisms to limit the maximum depth of the trees. This balance between aggressive optimization and careful regularization encapsulates LightGBM's ability to offer a highly efficient, accurate, and scalable solution for gradient boosting, leveraging its leaf-wise growth strategy to stand out among tree-based learning algorithms.

(a) Level-wise splitting schematic diagram

(b) Leaf-wise splitting schematic diagram

Figure 4.1: Splitting schematic comparison

Besides, to enhance both training speed and memory efficiency, LightGBM leverages an innovative histogram-based algorithm that discretizes continuous feature values into k discrete intervals. This approach significantly streamlines the process of finding the optimal split points during tree construction by limiting the search to these predefined intervals. By reducing the granularity of the feature space, the algorithm effectively minimizes the computational complexity of model training from $O(\#\text{ data} \times \#\text{ feature})$ to $O(k \times \#\text{ feature})$, where # data represents the total number of data points and # feature denotes the total number of features. Given that the number of data points vastly exceeds k (i.e., $\#\text{ data} \gg k$), this method

significantly improves training efficiency and reduces the space required to store the model. This histogram-based feature discretization not only accelerates the training process by simplifying the evaluation of potential splits but also contributes to LightGBM's overall efficiency and scalability, particularly when dealing with large datasets and high-dimensional feature spaces. Additionally, LightGBM's exclusive feature bundling method merges mutually exclusive features within a specific conflict ratio, enabling dimensionality reduction without losing information (Ke, Meng, Finley, T. Wang, W. Chen, Ma, Ye, and T.-Y. Liu 2017).

LSTM is a special type of RNN designed to address the vanishing and exploding gradient problems that traditional RNNs face when dealing with long sequences. LSTM achieves this by introducing a memory cell and three gates—forget, input, and output gates—that regulate the flow of information, making it highly effective for tasks involving sequential data, such as time series prediction and natural language processing.

Mathematically, an LSTM cell operates as follows. At each time step t the forget gate determines which part of the previous memory to retain, calculated as:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right), \tag{4.31}$$

where, f_t is the forget gate output, h_{t-1} is the previous hidden state, x_t is the current input, W_f and b_f are the forget gate's weight matrix and bias, respectively, and σ is the sigmoid activation function. The input gate decides which new information to add to the memory. It

has two components: the candidate memory update and the input gate itself:

$$\begin{aligned} \tilde{C}_t &= \tanh\left(W_c \cdot [h_{t-1}, x_t] + b_c\right) \\ i_t &= \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \end{aligned} \tag{4.32}$$

The memory cell is then updated as:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{4.33}$$

The output gate determines the next hidden state based on the updated memory:

$$\begin{aligned} o_t &= \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \\ h_t &= o_t \cdot \tanh\left(C_t\right) \end{aligned} \tag{4.34}$$

In these equations, $\tilde{C}_t$ is the candidate memory, i_t is the input gate's activation, C_t is the current memory cell, o_t is the output gate's activation, and h_t is the current hidden state. The three gates: forget, input, and output, work together to regulate the flow of information, enabling LSTM to capture long-term dependencies effectively. This structure makes LSTM robust in handling complex sequential tasks like language modeling, speech recognition, and financial forecasting.

4.1.2 Stacking method introduction

Since our prediction tasks for VAT, VTT, and VST all involve tabular data, tree-based models are well-suited as our fundamental predictive models due to their ability to handle structured data effectively.

Unlike neural networks, which are typically data-driven and require large volumes of data to learn complex representations, tree-based models are feature-driven—they explicitly capture hierarchical feature interactions through recursive splitting. This makes them particularly effective for tabular data with heterogeneous features, missing values, and non-linear relationships. To enhance the accuracy of the prediction results from a single model, an advanced stacking model has been proposed, which incorporates RF, XGBoost, and LightGBM as its base learners. These tree-based models collectively leverage their strengths to improve overall predictive performance (Z.-H. Zhou 2012). This approach offers several advantages, such as improving predictive accuracy by combining the diverse learning capabilities of each model, reducing the risk of overfitting by balancing out the individual model errors (Z.-H. Zhou 2012).

To illustrate the principles of the stacking method, we take the VAT prediction task as an example. The proposed stacking model structure for VAT prediction, illustrated in Figure 4.2, consists of two main components. The training dataset is first divided into five folds for cross-validation (CV). Each base model is trained on four folds and validated on the fifth, producing out-of-fold predictions (e.g., "Predict_L1" to "Predict_L5" for LightGBM). This process is repeated five times for each base model while ensuring that every fold serves as a validation set once.

In each iteration of the stacking method, each base model makes predictions on an unseen test set. These individual predictions are then averaged to produce the final prediction for the test set. After

completing CV for all base models, the predictions from each fold are stacked to form a new dataset. This stacked dataset is then used as input to train the second layer linear regression meta-learner, which combines the base models' predictions to generate the final output. This stacking process effectively leverages the strengths of different models, leading to improved predictive performance.

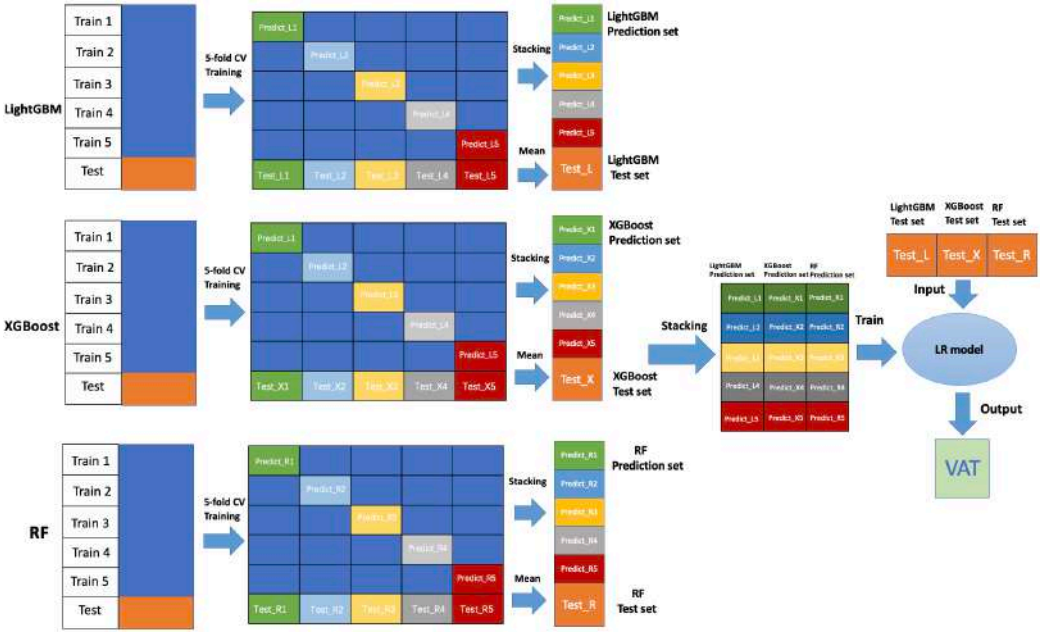


Figure 4.2: The architecture of the proposed stacking model used for VAT prediction

4.1.3 Model evaluation metrics

For offline evaluation of VTT prediction, the model's output is considered as the predicted value, while the actual corresponding value is treated as the ground truth. we adopt the following common regressor evaluation metrics: RMSE, MSE, MAE, MAAPE, R^2 and $MAPE$. Suppose that a total of n vessels arrive at the HKP within

the specified period and the actual corresponding value is y_i for vessel $i, i = 1, \ldots, n$. The predicted value, as predicted by the model or reported by the vessel, is denoted by $\hat{y}_i$ for vessel i. The definitions of RMSE, MSE, MAE, MAAPE, R^2 and MAPE are expressed as:
RMSE:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}. \tag{4.35}$$

MSE:

$$MSE = \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{n}. \tag{4.36}$$

MAE:

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n}. \tag{4.37}$$

MAAPE:

$$MAAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|. \tag{4.38}$$

R^2:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}. \tag{4.39}$$

$MAPE$:

$$\text{MAPE} = \frac{1}{m}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100. \tag{4.40}$$

In evaluation, lower values of RMSE, MSE, MAE, MAAPE, MAPE indicate better performance. Conversely, a higher R^2 value, typically ranging from 0 to 1, is preferable, indicating greater predictive accuracy and improved overall model performance.

4.2 VAT Prediction

4.2.1 Ocean-going VAT prediction

Port operations are planned daily to reduce vessel waiting time and cargo handling durations, thus maximizing efficiency. One crucial challenge faced by terminal operators in daily port operations is the uncertainty of vessel arrivals. Vessels normally upload their ETA one or several days before approaching the port, but these estimates often significantly differ from the vessel's ATA to the port. Factors such as unforeseen weather, sea conditions, and port congestion can cause discrepancies between ETA and ATA, reducing port efficiency (Y. Yang, Yan, and S. Wang 2024). Managing these discrepancies is crucial for maintaining the reliability of vessel operations. According to a SeaIntelligence report, liner services are regarded as reliable if vessel delays are within 24 hours of the scheduled ETA, with global container schedule reliability averaging 51.6% for on-time arrivals as of March 2024 (Jasmina, Ovcina Mandra 2024). To mitigate the uncertainty associated with vessel arrivals, data-driven models can be leveraged to provide more accurate predictions of vessel arrival time (VAT) compared to the traditionally reported ETA. These models offer greater reliability than vessel-reported estimates, thereby enhancing decision-making processes in port operations. Port call data and automatic identification system (AIS) data are both critical datasets for VAT prediction models and are essential sources of vessel operation information for port authorities (Yan, S. Wang, Zhen, and Laporte 2021). Port call data comprises details such as vessel

ETA and ATA data, serving as crucial references for vessel management and port operation. Meanwhile, AIS data provides real-time movement trajectories of vessels (D. Yang, L. Wu, S. Wang, Jia, and K. X. Li 2019; Filom, Amiri, and Razavi 2022). Although vessel arrival uncertainty in ports is a well-recognized issue by port authorities, vessel operators, and the maritime research community, current research on addressing this challenge is extremely limited, and the problem continues to pose difficulties for port operators (D. Yang, L. Wu, S. Wang, Jia, and K. X. Li 2019; Filom, Amiri, and Razavi 2022). Moreover, existing studies on VAT prediction often rely on data from a single source and focus primarily on short-distance forecasting. For instance, many studies exclusively utilize AIS data to predict VAT within a range of 10 to 50 kilometers from the destination, frequently overlooking the importance of port call data in improving VAT predictions for both short- and long-distance scenarios (Filom, Amiri, and Razavi 2022; X. Zhang, Fu, Xiao, H. Xu, Wei, Koh, Ogawa, and Qin 2024; T. Zhang, J. Yin, X. Wang, and Min 2023). Previous studies commonly set the vessel's final AIS point of its approach to the port as the arrival point, using the AIS timestamp at this location to determine the vessel's ATA (Filom, Amiri, and Razavi 2022). However, this approximation introduces inaccuracies when compared to the vessel's true ATA, as precise ATA data is often derived from port call records maintained by the port authorities. Some research has employed static port call data for VAT prediction (J. Yu, G. Tang, Song, X. Yu, Qi, D. Li, and Y. Zhang 2018), but port call data lacks real-time vessel trajectory information. More-

over, the limited frequency of ETA reports during a voyage reduces the effectiveness of ETA-based predictions when relying solely on port call data (J. Yu, G. Tang, Song, X. Yu, Qi, D. Li, and Y. Zhang 2018). To enhance VAT prediction accuracy, a promising approach is to integrate static port call data with dynamic AIS data. The ATA data derived from port call records can serve as the ground truth for predictions, while AIS data provides the foundation for building the predictive model. However, research on integrating port call records and AIS data for VAT predictions is currently lacking. Consequently, the key research questions for port authorities are as follows:

Q1: What is the trend of accuracy of the vessel-reported ETA during the vessel's voyage to port?

Q2: How to integrate vessel AIS data with the corresponding port call data to improve VAT prediction?

Q3: What models are suitable for VAT prediction when considering both AIS data and port call data?

Q4: Which features have a significant contribution to VAT prediction?

In this study, we focus on forecasting the VAT at the HKP using integrated data. Specifically, we predict a more accurate one based on the reported ETA and additional features generated from feature engineering. Figure 4.3 illustrates the VAT evaluation and prediction process. With the matched VAT data, we present a tree-based stacking model for VAT prediction, conduct VAT predictions, and perform

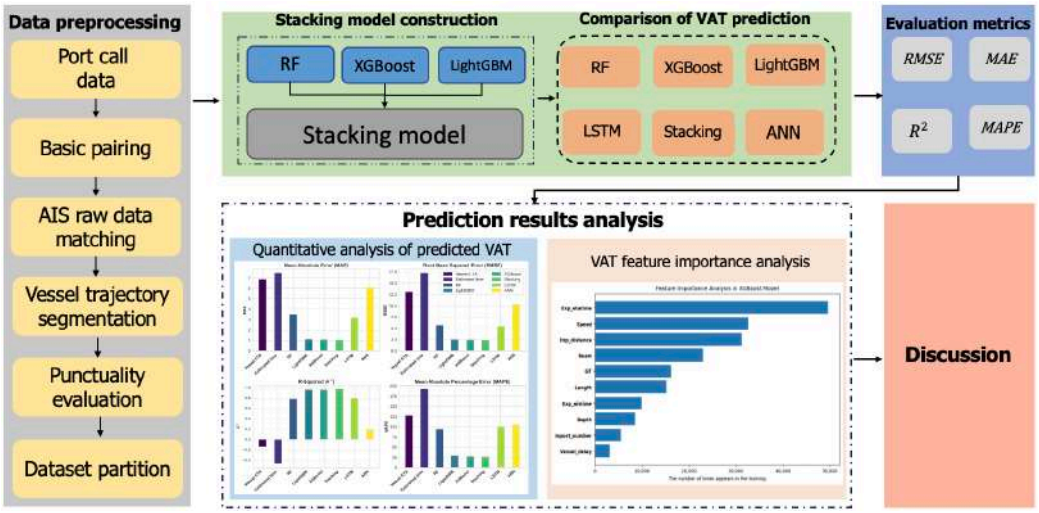


Figure 4.3: The flowchart of vessel arrival evaluation and prediction

a quantitative analysis of the results, including an evaluation of feature importance.

4.2.1.1 Feature engineering

Building on domain expertise, feature engineering is a crucial ML process that transforms raw data into meaningful features, improving model performance. By selecting, transforming, or combining variables, it enables algorithms to better capture essential patterns in the data, thus strengthening the model's predictive power and accuracy (Z.-H. Zhou 2012). The main stages of feature engineering in the study include feature selection and fusion, focused on identifying and incorporating the most relevant variables for VAT prediction. Our approach enhances VAT prediction accuracy by integrating features across four key categories: vessel physical features, temporal features, berth operational features, and spatial features. The actual remaining voyage time, calculated as the difference between the ATA

and the AIS report time for each vessel record, serves as the ground truth in the VAT prediction task. The specific features used for VAT prediction are explained as follows and summarized in Table 4.1.

Vessel physical features:

Vessel physical features, encompassing the geometric dimensions of a vessel such as length, depth, and draft, are pivotal in predicting VAT at ports (J. Yu, G. Tang, Song, X. Yu, Qi, D. Li, and Y. Zhang 2018). These geometric characteristics of vessel directly impact subsequent berth allocation, thereby influencing overall port operations. These geometric characteristics directly affect vessel maneuverability, berthing requirements, and port congestion levels, thereby playing a crucial role in determining VAT. Larger vessels with greater draft and beam require deeper water channels and larger berth spaces, often leading to longer turnaround time due to more complex docking and cargo handling processes. Conversely, smaller vessels with shallower drafts may experience more flexible berthing options, reducing overall VAT variability.

Furthermore, these physical attributes influence vessel speed, fuel consumption, and navigational constraints, all of which contribute to variations in VAT under different port conditions. To enhance our predictive model, we integrate external vessel physical data from the World Register of Ships (WRS) into our AIS dataset using vessel call sign value as an identifier (WRS 2023). This approach enables the compilation and integration of vessel physical characteristics, including beam, gross tonnage (GT), depth, and length, which are essential for capturing the operational complexity of different vessel

types. Additionally, vessel type information is sourced directly from port call data, ensuring a comprehensive understanding of vessel behavior across different shipping routes and port infrastructures.

Spatial features:

AIS data contains real-time vessel latitude and longitude information, which can be used to estimate the precise distance of a vessel from the port, serving as a spatial feature for VAT prediction. These spatial features are crucial as they capture real-time vessel movement, reflecting navigational constraints, congestion, and speed variations that impact arrival time. Unlike static vessel attributes, AIS-based spatial data allow for dynamic updates, improving VAT accuracy by incorporating actual voyage conditions. Previous research in estimating the remaining voyage distance typically calculates the distance between the vessel's current AIS position and the final destination of its voyage. However, this approach has limitations in practical port operations because ports typically have multiple berths, and the port authority does not know the last point of the vessel's voyage when predicting the ETA. Therefore, we propose a method for estimating the remaining distance to the port for vessels based on historical AIS data. This method involves the following steps.

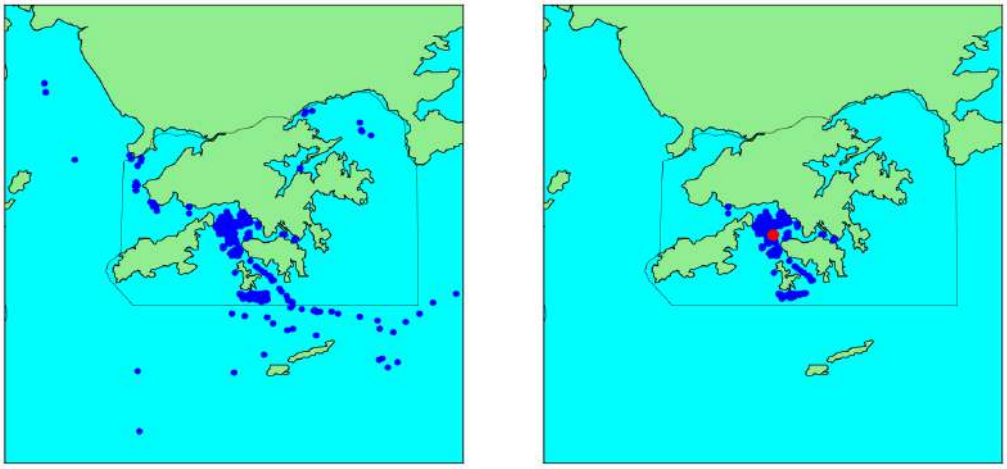
(1) Collect Final Points: Compile the coordinates of the final points for each voyage from the AIS dataset.

(2) Determine the Coordinates: Calculate the mean value of latitude and longitude of these final points to determine the mean center coordinates.

(3) Identify the Nearest Points: Find the 90th percentile of the final

coordinates points closest to this mean center and compute their average coordinates. This average is considered the estimated endpoint for each voyage.

(4) Compute the Expected Remaining Distance: Finally, we calculate the expected remaining distance from each AIS point to the estimated endpoint using the great circle distance, applying the Haversine formula.



(a) Unprocessed collection of final voyage points

(b) Processed collection of final voyage points with estimated endpoint

Figure 4.4: Visualization of the final voyage AIS data points

Temporal features:

When predicting the VAT for vessels, incorporating temporal features is crucial for capturing the nuances of vessel movement and external factors over time. These time-based attributes offer valuable insights into the progression of a vessel's journey, helping to identify patterns, delays, and deviations from expected routes. For instance, analyzing historical AIS timestamps can reveal recurrent de-

lays during peak hours, while comparing real-time movement trends with ETA data helps adjust predictions dynamically. Additionally, incorporating temporal trends allows for better adaptation to seasonal variations, weather disruptions, and operational constraints, ultimately enhancing the accuracy and reliability of VAT forecasts. In our prediction model, we primarily consider five temporal features: vessel speed, estimated remaining voyage time from ETA data ($Exp\_etatime$), estimated remaining voyage time from AIS data ($Exp\_aistime$), the weekday of the AIS report time (AIS_day), the weekday of the ETA time (ETA_day), and the hour shifts of the AIS report time and ETA (AIS_hour_shift, ETA_hour_shift).

The vessel speed, obtained from real-time AIS data, is used to calculate the remaining time for the vessel to reach the port. The estimated remaining voyage time from ETA data refers to the time remaining when the vessel reports its ETA, calculated by subtracting the AIS report time from the ETA. The estimated remaining voyage time from AIS data is determined by dividing the estimated voyage distance by the vessel's current speed. The weekday of the AIS report time and the ETA are considered because port operational efficiency varies between weekdays and weekends, and this factor is incorporated into our VAT prediction model using one-hot vector encoding. Additionally, since port efficiency also fluctuates throughout different time shifts of a day, one day is divided into three shifts: midnight to 8:00 AM, 8:00 AM to 4:00 PM, and 4:00 PM to midnight. Both the ETA and the AIS report time are one-hot encoded to account for these temporal variations in the model.

Berth operational features:

Berth operational features are vital in determining VAT. For instance, the number of vessels present at the port at the time of an AIS report serves as a key indicator of the port's operational intensity. When vessel traffic exceeds the port's berthing capacity, limited berthing availability can lead to prolonged waiting time at anchorages, increasing overall port congestion and disrupting planned schedules. These factors collectively impact VAT by affecting in-port vessel turnaround time and port efficiency, making berth operational features an essential component in predictive modeling. In predicting VAT, two berth operational features are considered: the number of vessels in the port at the AIS report time, which is directly accessible from the port's records, and the vessel's historical average delay. The latter is particularly relevant to vessels that frequently berth at HKP, as their average delay is used to assess the reliability of their reported ETA, with the single-berth delay calculated as the difference between the vessel's ATA and ETA.

Prediction target:

For the VAT prediction task, the prediction model cannot directly predict timestamp values. Instead, the model aims to predict the remaining sailing time of the current voyage, where the ground truth is calculated as the difference between the vessel's ATA and the AIS report time. The predicted VAT for the vessel is then calculated by adding the predicted remaining voyage time to the AIS timestamp.

Table 4.1: Description of features for ocean going VAT prediction

Feature category	Detailed features	Explication	Note
Vessel physical features	*Beam*	The width of a vessel at its widest point	Physical feature set, matched from WRS
	GT	The overall internal volume of a vessel	Physical feature set, matched from WRS
	Depth	From the top of the keel to the deck of a vessel	Physical feature set, matched from WRS
	Length	The measurement of the vessel from the front to the back	Physical feature set, matched from WRS
	Type	Vessel type	Physical feature set, obtained from vessel due to arrival file, 7 different vessel types, one-hot vector encoding
Spatial features	*Exp_distance*	Expected remaining voyage distance	AIS feature set, estimated from AIS data
Temporal features	*Exp_aistime*	Expected remaining voyage time from AIS data	AIS feature set, calculated by dividing estimated remaining distance by real time speed
	Speed	Vessel real time speed	AIS feature set, obtained from AIS data
	Exp_etatime	Expected remaining voyage time from ETA data	Port call feature set, ETA time minus AIS report time
	AIS_day	Week day of the AIS report time	AIS feature set, from Monday to Sunday, one-hot vector encoding
	ETA_day	Week day of the ETA time	Port call feature set, from Monday to Sunday, one-hot vector encoding
	AIS_hour_shift	Hour shift of the AIS report time	AIS feature set, three different shifts, one-hot vector encoding
	ETA_hour_shift	Hour shift of the ETA	Port call feature set, three different shifts, one-hot vector encoding
Berth operational features	*Inport_number*	Te number of vessels in port when the vessel reports AIS data	Port call feature set, counted from in port vessels file
	Vessel_delay	The mean value of vessel historical VAT delay	Port call feature set, ATA mins ETA
Prediction target/ Ground truth value	*Actual_remaining_time*	The actual remaining voyage time when vessel reports AIS data	ATA minus AIS report time

4.2.1.2 VAT Prediction results analysis

A comprehensive evaluation of VAT prediction is conducted, structured into five distinct parts. Initially, the standard procedure is applied, where the dataset is divided into training and test sets with a 4:1 split to facilitate baseline model training and evaluation, as shown in Table 4.2. However, the traditional method of dataset partitioning has limitations, as port authorities in real-world operations cannot access data for vessels that have not yet arrived. To better address these operational constraints, the dataset is divided based on temporal and voyage-based criteria. For the temporal criteria, the data is split chronologically, with earlier data used for training and later data reserved for predictions. In the voyage-based criteria, the data is partitioned by vessel voyages, ensuring that entire voyages are included either in the training set or the prediction set, without overlap.

The dataset spans from January 1st, 2021 to February 28th, 2021 For the time-based partitioning, data from January 1st, 2021 to February 13th, 2021 is used as the training set, while the remaining data serves as the test set, with the results presented in Table 4.3. In the voyage-based partitioning, we first segment the matched AIS dataset into different vessel voyages based on vessel call signs and ATA values. Subsequently, 80% of the voyage data, along with the corresponding AIS data, is allocated to the training set, while the remaining 20% is reserved for testing, with the results presented in Table 4.4.

To better illustrate the contribution of different features to the VAT prediction, we also evaluate the VAT prediction performance of

different models across distinct feature sets, based on the note in Table 4.1, and across the three training scenarios. Four distinct feature sets are considered: the first set included all features; the second excluded port call features ($Exp\_etatime$, ETA_day, ETA_hour_shift, and $Inport\_number$); the third excluded AIS-related features; and the fourth excluded vessel physical characteristics ($Beam$, GT, $Depth$, $Length$, $Type$). Moreover, to assess models predictive performance at varying vessel remaining distances and time, VAT prediction models are initially trained on the training sets defined in Table 4.2. The test set is subsequently segmented into three time slices and three distance slices. The segmentation is performed at intervals of 12 hours for time slices and 200 kilometers for distance slices.

Specifically, time slice 1 includes matched AIS data from the test set where the actual remaining voyage time is within 2 hours; time slice 2 includes data where the actual remaining voyage time is between 2 and 4 hours; and time slice 3 includes data where the actual remaining voyage time is greater than 4 hours. Similarly, distance slice 1 comprises data where the actual remaining voyage distance is within 200 kilometers; distance slice 2 includes data where the actual remaining voyage distance is between 200 and 400 kilometers; and distance slice 3 includes data where the actual remaining voyage distance is greater than 400 kilometers. The trained models are then evaluated across these various test slices to determine their performance in different operational contexts. The results of these evaluations are presented in Table 4.5 and Table 4.6.

Table 4.2: Prediction performances of the models on the original test dataset

Model	**All variables**				**All variables except port call features**				**All variables except AIS features**				**All variables except physical features**			
	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE
Vessel ETA	6.88	13.11	−0.14	128.16	–	–	–	–	6.88	13.11	−0.14	128.16	6.88	13.11	−0.14	128.16
Estimated time	7.52	17.31	−0.47	192.16	7.52	17.31	−0.47	192.16	–	–	–	–	7.52	17.31	−0.47	192.16
RF*	3.53	5.74	0.78	94.85	4.02	6.22	0.74	94.53	4.18	6.42	0.73	219.67	3.84	6.30	0.74	63.26
LightGBM	1.11	2.52	0.96	29.53	1.90	3.92	0.90	41.13	2.66	5.31	0.80	120.43	2.90	5.30	0.80	46.99
XGBoost	1.08	2.48	0.96	27.56	1.87	3.93	0.90	39.29	2.58	5.13	0.83	113.87	2.84	5.30	0.81	46.87
Stacking	**1.06**	**2.37**	**0.97**	**26.91**	**1.86**	**3.89**	**0.91**	**39.01**	**3.35**	**5.81**	**0.79**	**108.99**	**2.80**	**5.30**	**0.82**	**45.55**
LSTM‡	3.22	5.53	0.80	100.99	2.96	5.36	0.81	75.18	4.10	6.47	0.72	155.89	3.50	6.07	0.71	58.18
ANN	6.09	10.43	0.19	105.91	4.01	6.00	0.65	99.57	4.51	6.65	0.51	240.12	4.43	8.27	0.51	92.19

* The best performance model in J. Yu, G. Tang, Song, X. Yu, Qi, D. Li, and Y. Zhang (2018).
‡ The best-performing model in the study by Wenzel, Jovanovic, and Schulte (2023).

Table 4.2 compares the performance of RF, LightGBM, XGBoost, Stacking, LSTM, and ANN prediction models, alongside the best-performing model from the previous studies, on the HKP vessel arrival dataset under different metrics and feature combinations. Port operations rely on two key estimates: the Vessel ETA, based on the vessel's reported arrival time, and the estimated time, obtained by evaluating the remaining voyage distance from AIS data and dividing it by the vessel's speed. Both of these sources exhibit large MAE and RMSE values, negative R^2, and high MAPE, highlighting their significant inaccuracies and unreliability for VAT estimation. Among the VAT prediction models, Stacking, LightGBM, and XGBoost consistently outperform the others, with the stacking model achieving the lowest errors across all variables.

Compared to the vessel ETA, Stacking reduces the MAE by 84.59% (from 6.88 hours to 1.06 hours) and the RMSE by 81.92% (from 13.11 hours to 2.37 hours). Additionally, the R^2 improves dramatically from -0.14 to 0.97, indicating a much better fit. MAPE also drops by 79.00%, from 128.16% to 26.91%. These reductions in error metrics demonstrate that advanced tree-based models like XG-

Boost and Stacking significantly enhance prediction accuracy over the vessel-reported ETA. The results for LSTM and ANN, however, are less impressive compared to the tree-based models. LSTM shows moderate performance with an MAE of 3.22 hours and an RMSE of 5.53 hours, while ANN performs even worse, with an MAE of 6.09 hours and an RMSE of 10.43 hours. The primary reason for their underperformance is that the VAT prediction task is not a typical time series prediction problem, where LSTM typically excels (Z.-H. Zhou 2012). Instead, this problem involves handling tabular data with mixed features, where neural networks generally struggle to capture complex feature relationships effectively (Grinsztajn, Oyallon, and Varoquaux 2022)

In the analysis of VAT prediction across different feature combinations, the exclusion of port call or AIS features leads to a significant increase in VAT prediction errors, underscoring the critical importance of integrating ETA or AIS data for accurate VAT predictions. This impact is particularly evident in models like RF and ANN, which show a marked decline in performance compared to more robust tree models like LightGBM, XGBoost, and stacking approach. LSTM and ANN, in particular, exhibit consistently higher MAE and MAPE across all feature combinations, further emphasizing their limitations in handling tabular data with limited features. In contrast, the stacking approach demonstrates superior resilience against variations in input features, maintaining strong performance even when key features are unavailable, highlighting its ability to manage feature variability and provide more reliable predictions.

Table 4.3: Prediction performance of models on the time-ordered test dataset

Model	**All variables**				**All variables except port call features**				**All variables except AIS features**				**All variables except physical features**			
	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE
Vessel ETA	6.84	10.61	0.05	71.37	–	–	–	–	6.84	10.61	0.05	71.37	6.84	10.61	0.05	71.37
Estimated time	7.02	13.56	−0.13	111.17	7.02	13.56	−0.13	111.17	–	–	–	–	7.02	13.56	−0.13	111.17
RF	3.53	5.64	0.74	61.63	3.98	6.10	0.68	68.49	4.22	6.88	0.61	60.72	3.62	5.64	0.69	49.31
LightGBM	3.19	5.33	0.77	44.71	3.74	6.22	0.68	50.13	4.39	6.80	0.59	60.84	3.62	5.69	0.70	47.91
XGBoost	3.16	5.32	0.78	40.42	3.77	6.07	0.69	44.68	4.30	6.67	0.63	60.14	3.59	5.65	0.73	46.27
Stacking	**3.11**	**5.29**	**0.81**	**38.77**	**3.74**	**6.00**	**0.71**	**42.19**	**4.19**	**6.51**	**0.65**	**55.81**	**3.52**	**5.49**	**0.75**	**43.22**
LSTM	3.91	6.01	0.53	56.91	4.51	7.11	0.50	69.10	4.57	7.59	0.33	70.91	4.01	7.09	0.49	66.19
ANN	5.89	9.23	0.03	65.91	6.11	9.43	0.04	100.71	4.71	9.71	0.11	72.39	5.17	9.00	0.21	70.71

Table 4.4: Prediction performance of models on the voyage-split test dataset

Model	**All variables**				**All variables except port call features**				**All variables except AIS features**				**All variables except physical features**			
	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE
Vessel ETA	6.53	13.09	−0.18	118.36	–	–	–	–	6.53	13.09	−0.18	118.36	6.53	13.09	−0.18	118.36
Estimated time	7.31	12.93	−0.39	99.13	7.31	12.93	−0.39	99.13	–	–	–	–	7.31	12.93	−0.39	99.13
RF	4.14	6.80	0.68	94.85	4.65	7.22	0.64	100.83	4.94	8.10	0.55	228.55	4.16	6.63	0.65	74.07
LightGBM	3.80	6.75	0.69	55.63	4.43	7.04	0.76	72.01	4.88	7.89	0.59	101.22	4.02	6.50	0.75	51.24
XGBoost	3.75	6.45	0.71	62.90	4.46	7.03	0.75	72.40	4.89	7.94	0.57	100.69	4.03	6.54	0.72	59.70
Stacking	**3.70**	**6.42**	**0.72**	**52.33**	**4.31**	**6.99**	**0.80**	**70.91**	**4.84**	**7.88**	**0.61**	**99.03**	**3.99**	**6.49**	**0.78**	**51.31**
LSTM	4.02	7.00	0.63	102.73	4.63	7.44	0.62	95.67	5.18	8.53	0.50	262.34	4.15	6.72	0.68	82.69
ANN	6.00	10.13	0.43	105.79	5.35	8.41	0.44	130.99	5.73	9.93	0.31	191.19	4.52	7.00	0.51	103.31

Tables 4.3 and 4.4 summarize the prediction performance of various models on both the time-ordered and voyage-split test datasets. Consistent with the findings in Table 4.2, the vessel-reported ETA and the estimated time from AIS data exhibit relatively poor performance, as indicated by the high MAE and RMSE values, particularly for AIS data. This emphasizes the unreliability of these basic estimates for VAT prediction in real-world scenarios.

In terms of model performance, the stacking approach, along with XGBoost and LightGBM, consistently performs the best overall. The stacking model shows a slight advantage over the others across most scenarios in both the time-ordered and voyage-split cases. Additionally, it is important to note that, compared to VAT prediction on a randomly split dataset, shown in Table 4.2, the model errors are significantly larger across all metrics and feature exclusion scenarios

when using time-based and voyage-based splits (as demonstrated in Tables 4.3 and 4.4). This discrepancy arises because the voyages are divided based on a 4:1 split, but the corresponding integrated data is not split in the same 4:1 ratio since the amount of data for each voyage varies. Moreover, the time-based split better reflects real-world port operations, as ports do not have access to a vessel's exact future route information, making this approach a more realistic evaluation scenario.

Table 4.5: Prediction performance of the models on the different time slices

Model	**All data**				**Time slice 1**				**Time slice 2**				**Time slice 3**			
	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE
Vessel ETA	6.88	13.11	−0.14	128.16	3.67	6.56	−0.03	168.58	4.54	8.55	−0.06	24.82	11.91	19.51	−0.10	34.84
Estimated time	7.52	17.31	−0.47	192.16	6.59	9.48	−0.07	110.61	8.77	19.64	−0.32	83.09	22.93	36.81	−0.58	113.58
RF	3.53	5.74	0.78	94.85	2.49	5.23	0.09	100.01	2.34	3.31	0.09	26.36	4.66	8.09	0.01	13.69
LightGBM	1.11	2.52	0.96	29.53	1.09	2.06	0.64	39.98	0.79	1.48	0.79	4.27	0.99	2.43	0.83	3.05
XGBoost	1.08	2.48	0.96	27.56	0.91	1.76	0.74	54.53	0.71	1.31	0.83	3.93	0.86	2.08	0.88	2.68
Stacking	**1.01**	**2.37**	**0.97**	**26.91**	**0.90**	**1.72**	**0.75**	**48.13**	**0.68**	**1.29**	**0.84**	**3.74**	**0.82**	**2.01**	**0.89**	**2.55**
LSTM	3.22	5.53	0.80	100.99	2.00	4.19	0.30	75.70	2.22	4.21	−0.11	13.72	4.83	11.29	0.13	21.39
ANN	6.09	10.43	0.19	105.91	3.51	6.55	0.01	110.77	3.19	7.03	−0.19	20.11	8.93	17.33	−0.10	30.18

Table 4.6: Prediction performance of the models on the different distance slices

Model	**All data**				**Distance slice 1**				**Distance slice 2**				**Distance slice 3**			
	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2	MAPE
Vessel ETA	6.88	13.11	−0.14	28.16	5.63	11.86	−0.03	19.16	6.40	12.40	−0.09	24.27	6.36	7.05	−0.14	19.45
Estimated time	7.52	17.31	−0.47	192.16	7.00	16.64	−0.26	182.80	13.37	19.01	−0.54	103.09	19.52	22.26	−0.54	117.69
RF	3.53	5.74	0.78	94.85	4.94	7.32	0.49	28.04	2.90	5.04	0.51	12.17	2.03	3.11	0.78	6.51
LightGBM	1.11	2.52	0.96	29.53	1.41	2.79	0.93	18.64	0.70	1.41	0.96	3.25	0.52	0.91	0.98	1.74
XGBoost	1.08	2.48	0.96	27.56	1.16	2.33	0.95	16.76	0.67	1.31	0.97	3.13	0.49	0.89	0.97	1.64
Stacking	**1.01**	**2.37**	**0.97**	**26.91**	**1.16**	**2.32**	**0.95**	**16.70**	**0.61**	**1.22**	**0.97**	**2.86**	**0.47**	**0.84**	**0.98**	**1.55**
LSTM	3.22	5.53	0.80	100.99	3.81	6.84	0.52	20.99	2.83	4.09	0.55	8.91	1.83	2.49	0.69	5.55
ANN	6.09	10.43	0.19	105.91	5.93	10.19	0.11	30.11	6.06	10.51	0.19	20.12	3.91	5.81	−0.20	10.12

The performance of VAT prediction models across time and distance slices (Table 4.5 and Table 4.6) shows that the stacking model consistently provides stable and accurate predictions, outperforming other models. Across all data, the stacking model achieves a low MAE of 1.01 hours, an RMSE of 2.37 hours, and a high R^2 of 0.97,

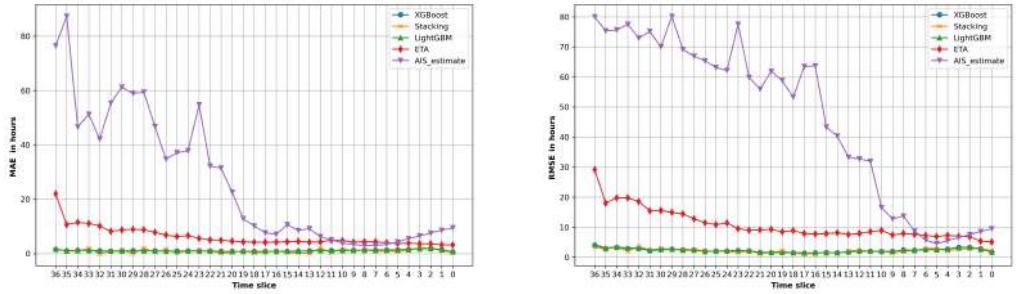
surpassing models such as RF and ANN, which exhibit higher errors.

In time slice 1, the stacking model performs well with an R^2 of 0.90, while models such as LSTM and ANN exhibit higher errors. In time slices 2 and 3, the stacking model continues to show lower errors than ANN and LSTM, demonstrating better stability over time. For distance slices, the stacking model also outperforms. To be specific, in distance slice 1, its MAE and RMSE are lower than those of XGBoost and ANN. In distance slices 2 and 3, the stacking model maintains lower errors, particularly in MAE and MAPE, compared to LSTM and ANN, which show significantly higher RMSE values.

Notably, the accuracy of VAT prediction models does not necessarily improve as the vessel approaches to port, either temporally or spatially. It is noteworthy that the models demonstrate optimal performance in time slice 2 and distance slice 2. Since the models are trained on the entire dataset and then predict across three slices, the greater amount of data in the slice allows for more thorough training, leading to better performance in this slice. Overall, the stacking model consistently outperforms others, showing superior accuracy and stability across time and distance slices.

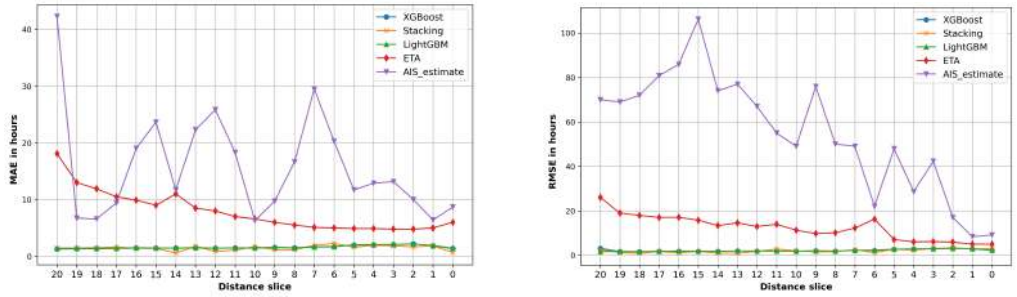
We also divide the original prediction set into different time slices and distance slices, following the partitioning method described in the earlier sections. Using the VAT models, we perform predictions within these slices to highlight the changes in error magnitude across time and distance slices, measured by MAE and RMSE. The specific results are visualized in Figures 4.5 and 4.6. In the figures, the "ETA" legend refers to the error in the vessel-reported ETA values, while

"AIS_estimate" represents the VAT time estimated directly based on AIS data.



(a) MAE trend for the time slices (b) RMSE trend for the time slices

Figure 4.5: Comparison of VAT prediction error trends across different time slices



(a) MAE trend for the distance slices (b) RMSE trend for the distance slices

Figure 4.6: Comparison of VAT prediction error trends across different distance slices

The results presented in Figures 4.5 and 4.6 align with the previous analysis. Overall, both the vessel-reported ETA and the VAT estimated using AIS data show improved accuracy as the vessel approaches the port. However, the AIS-based VAT estimates exhibit greater variability and larger errors, particularly across different slices. In contrast, the predictive tree-based models, including XGBoost and LightGBM, consistently perform well across various time and distance slices. Among these models, the stacking model stands

out, delivering the best overall performance with significantly lower MAE and RMSE errors compared to both the vessel-reported ETA and AIS-based VAT estimates. The consistent underperformance of vessel-reported ETA and AIS-based VAT estimates highlights the inherent limitations of traditional methods in accurately forecasting VAT. This underscores the necessity of adopting advanced predictive models to improve VAT predicting accuracy. Notably, the robust performance of tree-based models, particularly the stacking model, across diverse conditions demonstrates their adaptability and reliability for real-world port operation, making them a valuable tool for dynamic and uncertain port operational environments.

In addition to the numerical analysis of VAT prediction results, a feature importance analysis is conducted to leverage the interpretability of the tree-based model by using XGBoost's built-in function to calculate the scores based on the frequency of each feature's occurrence during model inference. Specifically, the feature importance is measured by the total number of times a particular feature appears in any tree in the model, used as a criterion for a split (T. Chen, T. He, M. Benesty, Khotilovich, Y. Tang, Cho, K. Chen, Mitchell, Cano, T. Zhou, et al. 2015). A higher score indicates that the feature is more frequently used to make decisions in the tree-building process. The top ten most important features are depicted in Figure 4.7.

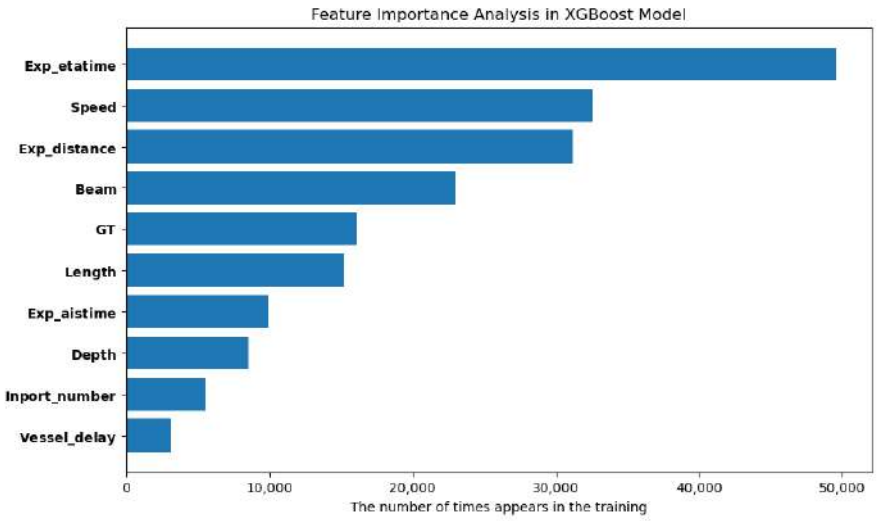


Figure 4.7: Feature importance analysis in the XGBoost model

To further enhance model interpretability, we also compute SHapley additive exPlanations (SHAP) values for each parameter of the XGBoost model. SHAP values, which are based on cooperative game theory (Scott, Su-In, et al. 2017), provide a consistent and locally accurate attribution of feature contributions to individual predictions. By applying the SHAP framework, we obtain both global and instance-level insights into how each feature influences the model output. This dual analysis, which employs both traditional feature importance scores and SHAP values, offers a more comprehensive understanding of the decision making process in VAT prediction models and helps identify the key features that drive VAT predictions. The visualization of the SHAP value derived from the XGBoost prediction model are presented in Figure 4.8.

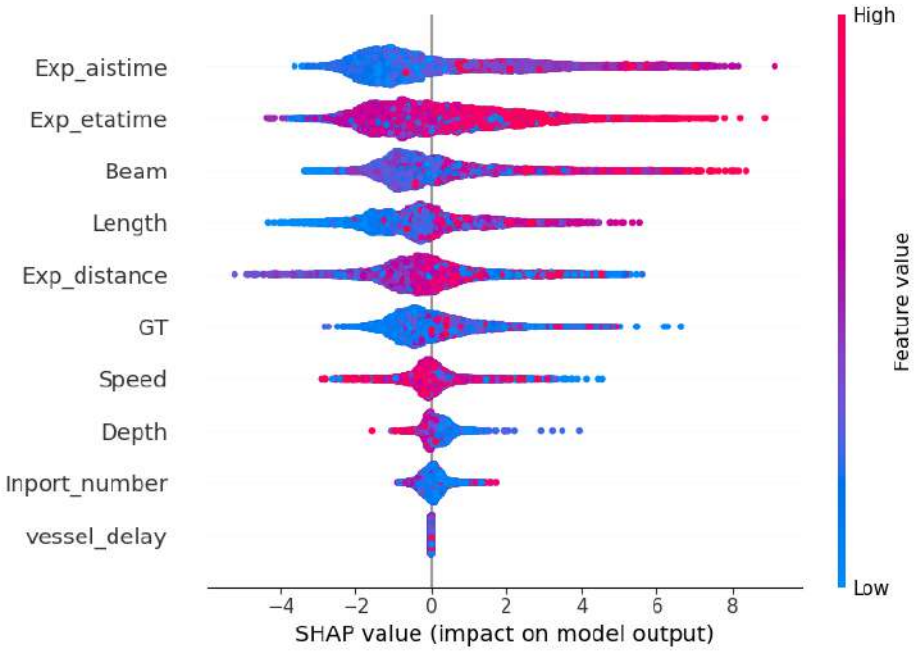


Figure 4.8: Analysis of SHAP values in the XGBoost model for VAT prediction

The expected remaining voyage time derived from ETA data emerges as the most important feature, highlighting its substantial contribution to VAT prediction, even there are discrepancies between the ETA values and the ATA values. This is followed by features of vessel sailing speed and the expected remaining voyage distance as derived from AIS data, where both of which are key indicators of VAT. Therefore, these three features stand out with considerably higher importance scores than the others, highlighting their crucial role in accurately predicting VAT.

Vessel physical attributes, such as $Beam$, GT, $Length$, and $Depth$, contribute to VAT prediction, although to a lesser extent than ETA and AIS data. These characteristics directly influence terminal and berth allocation at the port. For instance, larger vessels with greater

length and beam may encounter limitations when entering certain navigation channels or docking areas at specific berths, as not all facilities such as terminals or berthing facilities can accommodate vessels of such dimensions. This can result in delays while waiting for suitable berths or navigating restricted port areas, ultimately affecting VAT predictions.

Features $Inport\_number$ and $Vessel\_delay$ also influence VAT prediction. $Inport\_number$, which represents the number of vessels currently at the port, can contribute to potential delays due to congestion or limited berth availability. When there are more vessels at the port, it can cause delays in berthing or increase the waiting time for vessels to dock. $Vessel\_delay$, reflecting the historical average delay in vessel arrivals, provides some insight into potential vessel arrival performance. However, its impact is limited as it offers a retrospective view rather than serving as an active predictor. The arrival time of ships is influenced by numerous factors, each of which may vary for individual voyages, making past delays not necessarily applicable to the current route. Overall, feature importance analysis highlights the significance of contextual features beyond ETA and AIS data in VAT prediction, which is sometimes overlooked in previous VAT prediction studies.

From the SHAP value visualization, further insights can be drawn to complement the above analysis. The SHAP values (horizontal axis) from Figure 4.8 illustrate both positive and negative contributions of features to VAT prediction. For instance, $Exp\_aistime$ and $Exp\_etatime$ exhibit wide SHAP value distributions, indicating

high variability in their influence on model predictions. This aligns with their critical role in VAT prediction. Meanwhile, vessel physical attributes such as $Beam$ and $Length$ display narrower SHAP distributions, suggesting their more limited while relevant impact.

The color gradient in the SHAP plot (from blue to red) represents the magnitude of feature values. For example, higher feature values (in red) for $Speed$ and $Exp\_distance$ are associated with positive SHAP values, indicating that increase in speed or voyage distance tends to increase the values of predicted VAT. In contrast, lower characteristic values (in blue) for $Vessel\_delay$ result in negative SHAP values, highlighting that smaller historical delays generally lead to increase VAT predictions.

Lastly, features such as $Inport\_number$ and $Vessel\_delay$ exhibit relatively limited SHAP value variation, aligning with their role as supplementary predictors. However, their long-tail distributions indicate that under specific conditions, these features can exert a significant influence on VAT predictions. The broad distribution of SHAP values for $Exp\_etatime$ further underscores the high sensitivity of the prediction target to this feature, reinforcing its critical importance in port operations and VAT forecasting.

4.2.1.3 Prediction results summary

Uncertainty in VAT presents significant challenges in daily port operations, resulting in operational inefficiencies and economic losses. This study is the first to evaluate and predict VAT by integrating port call data and AIS data, using HKP as a case study. Specifically,

a framework is firstly introduced to integrate ETA and ATA with AIS data. A detailed quantitative analysis is conducted to assess the accuracy of vessel-reported ETA as vessels approach the port, considering both temporal and spatial factors. The results show a clear reduction in vessel arrival delays as vessels approach the port. Furthermore, a state-of-the-art tree-based stacking model is proposed to enhance the precision of VAT predictions. The performance of the stacking model and other base models is rigorously evaluated across various datasets using four metrics, offering a comprehensive analysis of their strengths and limitations.

The prediction results indicate that the stacking model outperforms other models across various evaluation metrics and datasets. Specifically, the stacking model reduces the MAE from 6.88 hours, as reported in the vessel operator's ETA data, to 1.06 hours in the test set, reflecting a reduction rate of 84.6%. Moreover, the proposed model achieves an 81.9% reduction in RMSE, with the RMSE decreasing from 13.11 hours to 2.37 hours. Additionally, the stacking model achieves an R^2 value of 0.97, indicating it explains 97% of the variance in VAT predictions and highlighting its high reliability. Beyond basic comparative predictions, an analysis of feature importance derived from the prediction model is also conducted. The analysis indicates that the vessels' reported ETA is the primary feature influencing the VAT prediction accuracy. Additionally, the vessel's real-time speed, expected remaining distance from AIS data, and physical characteristics significantly contribute to predicting VAT. Our findings highlight the potential of integrating ETA and AIS data

to predict VAT and demonstrate that the proposed tree-based models offer improved performance over traditional models in VAT prediction.

This study demonstrates the advantages of combining port call and AIS data to quantitatively assess vessel arrival punctuality, showing that the tree-based stacking model yields more accurate VAT predictions than both the vessel's reported ETA and the AIS-estimated ETA. However, several research questions remain for future exploration, particularly regarding prediction and subsequent optimization. From a predictive perspective, this study does not include real-time weather data, a critical factor affecting vessel arrival accuracy. Unexpected conditions such as heavy fog or storms can significantly disrupt the precision of vessel arrivals. In future research, if the necessary data becomes available, weather information could be integrated with AIS datasets based on time and geographic coordinates to account for weather factors in VAT prediction. Additionally, due to data acquisition challenges, this study only considered the real-time number of vessels in port as a port operation feature, neglecting other real-time operational data like berth availability and terminal handling efficiency. Future studies could incorporate these additional features to improve prediction accuracy.

4.2.2 Inland waterway VAT prediction

Accurate VAT prediction is critical for port operational efficiency, as discrepancies between captains' ETA and ATA in inland waterway transportation (IWT) cause economic losses. Most previous research

on VAT prediction has predominantly focused on ocean-going ships, typically relying on static port call data (e.g., ETA, ATA) or dynamic AIS data, with limited attention to additional features or waterway-specific conditions. According to a literature review by H. Li, Jiao, and Z. Yang (2023), a key limitation is the insufficient integration of multisource information beyond AIS data, highlighting opportunities to enhance predictive accuracy and robustness. Furthermore, while most studies have centered on ocean-going VAT, inland waterways present unique characteristics. Unlike ocean-going vessels, inland waterways feature fixed routes and higher traffic density. In these contexts, port authorities can more accurately estimate the remaining sailing distance of vessels due to the predictable nature of the routes. To address these research gaps, we propose a novel ML approach to estimate the VAT for IWT at the Port of Rotterdam. We integrate multiple data sources, including the remaining sailing distance of the vessel (calculated using the A* algorithm), real-time traffic flow on the waterway, vessel-reported ETA records, weather data (temperature, wind speed, wind direction and water levels), ship dimensions (length and width) and AIS data to apply ensemble tree models including RF, XGBoost, and LightGBM, with tailored data pre-processing and feature engineering. In summary, this part makes four specific contributions:

1. It introduces new long-distance VAT predictions for inland waterway vessels, marking a significant shift from the traditional focus on ocean going forecasting;

2. It proposes several features that are novel for VAT prediction,

including the ETA and AIS records, the remaining travel distance, maritime traffic flow, weather conditions, and the vessel's dimensions;

3. It suggests an A* algorithm for precisely segmenting the navigation route, enhancing the accuracy of the remaining vessel travel distance calculations and evaluating the maritime traffic flow;

4. It evaluates multiple prediction models on real-world vessel arrival data and derives new insights from feature correlation and importance based on VAT prediction models;

4.2.2.1 Introduction to IWT shipping and data description

The process for inland barge arrivals at the Port of Rotterdam begins with the captain reserving a specific time slot, ensuring the vessel arrives within this allocated period. However, challenges arise when a barge misses its scheduled slot or the reserved time slot becomes impractical, complicating port operations. Unlike sea-going vessels, inland freight often has a lower priority. When an inland vessel misses its space, its cargo, usually containers, may not transfer to the intended seagoing ship, leading to delays until the next available vessel. While precise scheduling is less critical for bulk cargo, improved prediction methods could enhance efficiency in cargo handling and optimize berth utilization for inland container barges.

This study focuses on the Port of Rotterdam, one of Europe's major ports, which handles approximately 438 million tonnes of cargo annually. The port receives around 28,000 sea-going ships and 90,000

inland vessels each year (Havenbedrijf Rotterdam B.V. 2023)Terminal operations for inland barges are highly dependent on accurate ETA. For this analysis, AIS data from the entire year of 2022 has been used. Figure 4.9 below illustrates a typical inland vessel docking at a port berth, often requiring stops at multiple terminals to unload its cargo.

Figure 4.9: Inland barge undocking at a berth

The Figure 4.10 below depicts the dimensions of the terminal. Each terminal typically includes designated areas for docking inland barges as well as ocean-going vessels. The area with the dotted line is the area for the sea-going vessels, and the rectangle with the solid line is the area for barge operations. The equipment used for loading and unloading these different types of vessels often varies, which is why the berths for sea-going and inland barges are separate. Consequently, the two types of vessels can be treated independently in planning operations. The following Figure 4.11 shows a heat map of

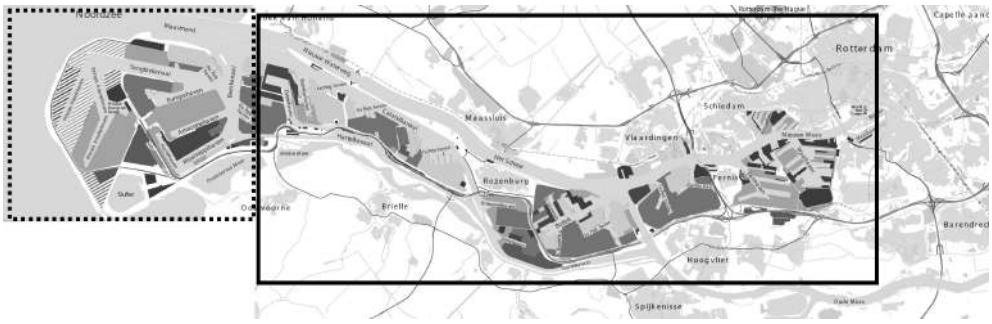


Figure 4.10: Map of terminals in the Port of Rotterdam

all the vessels that are going to the port of Rotterdam. Since multiple routes are taken towards the port area, this picture underlines the importance reliable predictions.

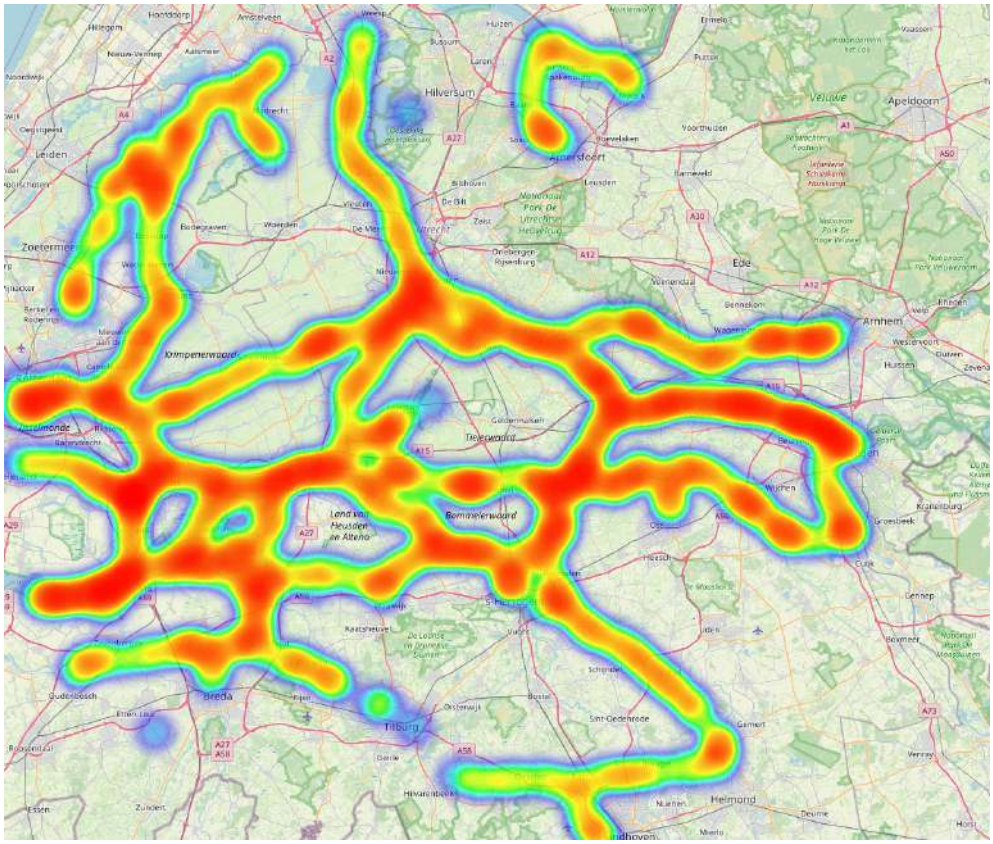


Figure 4.11: Heatmap of barges traveling to the port of Rotterdam

Figure 4.11 also illustrates the distribution of inland vessel traffic. Some traffic originates from the south, with vessels traveling from Antwerp towards Rotterdam. From the east, cargo is primarily transported along the Alpine-Rhine Corridor. Northern routes handle inland cargo within the Netherlands, connecting Rotterdam with Amsterdam and the country's northeastern regions.

4.2.2.2 Inland waterway segmentation

Compared to ocean shipping, inland shipping operates on fixed waterways, offering stable routes less affected by ocean tides. Its dense network allows access deep into the interior and the shorter distances make it ideal for regional freight transport. Compared to ocean shipping, inland shipping operates on fixed waterways, offering stable routes less affected by ocean tides. Its dense network allows access deep into the interior, and the shorter distances make it ideal for regional freight transport. However, in inland shipping, calculating straight-line or great-circle distances between two points using latitude and longitude often results in unreliable results due to the meandering nature of rivers, which causes ships to deviate from direct paths. To address this and improve accuracy, we first divide the inland waterways into segments represented by nodes within 2 km radius squares. The following Figure 4.12 shows the segments created based on the waterways in the Netherlands.

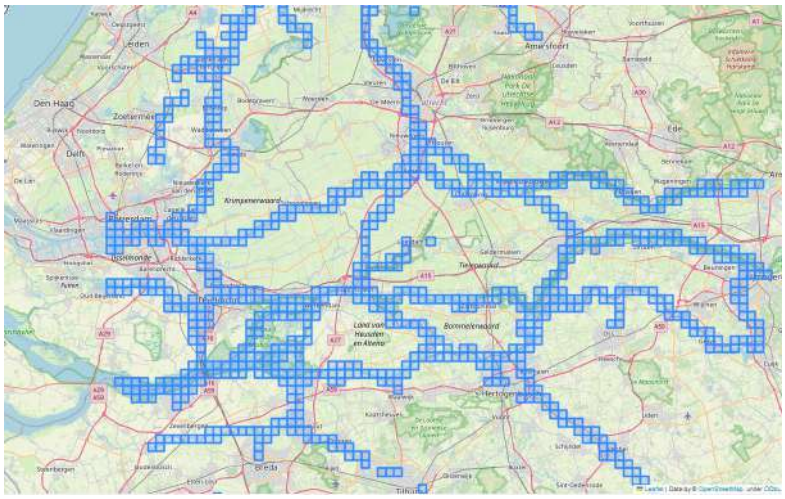


Figure 4.12: Detailed view of the river segments to the port of Rotterdam

Following the segmentation of inland waterways, the A* algorithm is employed to evaluate the remaining sailing distance for vessels. A* algorithm is a widely recognized heuristic search algorithm for pathfinding and graph traversal, combining the thoroughness of Dijkstra's algorithm with the efficiency of greedy search, thereby ensuring optimal path discovery while reducing computational overhead (Wenzel, Jovanovic, and Schulte 2023) In this context, the A* algorithm is utilized to compute the distance between the vessel's current position and the centroid of the next node along the intended route. These distances are incrementally summed to estimate the vessel's remaining travel distance. Furthermore, the segmented nodes facilitate the analysis and monitoring of vessel traffic flow within each node region, offering valuable insights into localized traffic patterns.

4.2.2.3 Destination coordinate estimation

When predicting VAT, exact destination coordinates are unknown, so our study uses historical AIS navigation data to estimate the endpoint of the trip. First, the final coordinates of each voyage are extracted from the AIS dataset, and their mean latitude and longitude are calculated to establish a central reference point. The 90th percentile of final coordinates closest to this reference point is then identified, and their average is computed to determine the estimated endpoint. Finally, the expected remaining distance from each AIS point to this estimated endpoint is calculated using the great-circle distance based on the Haversine formula. This method provides an effective means of estimating the endpoint of the vessel and the remaining distance. Figure 4.13a illustrates the visualization of the voyage endpoints' latitude and longitude from historical data, while Figure 4.13b shows the schematic diagram of the estimated endpoints after processing.

(a) The collection of endpoint coordinates in the map

(b) A schematic diagram of the estimated destination

Figure 4.13: Comparison of endpoint coordinates and estimated destination

4.2.2.4 Feature extension

Feature engineering involves selecting, transforming, or creating new features from raw data to enhance the performance of predictive models. This process includes techniques such as scaling, encoding categorical variables, and generating domain-specific features. Feature extension for prediction builds on this by expanding the feature set through methods like polynomial features, interaction terms, or temporal lags, enabling the model to capture complex relationships and patterns for improved predictive accuracy.

In our VAT prediction task, we aim to extend the foundational AIS dataset by mining and integrating additional features to create a comprehensive set of factors to predict vessel ETAs. Specifically, we focus on four categories of features: vessel physical attributes, temporal aspects, spatial elements, and waterway characteristics. The ground truth for prediction is the actual voyage time for each AIS report, calculated as the difference between the vessel's ATA and the AIS report time. Below are detailed explanations of these features:

We integrate vessel physical features into the AIS dataset, as these attributes directly influence the berth at which a vessel docks, affecting VAT prediction. Specifically, we consider the length and width of the vessel, representing its dimensions, gross tonnage of the vessel, indicating its cargo capacity, and type of the vessel, as different types exhibit distinct docking and operational patterns. These features are crucial for VAT prediction, impacting maneuverability, berthing requirements, and potential waiting time. Incorporating these factors helps capture both the intrinsic characteristics of the vessel and the

external logistical constraints, improving the accuracy and robustness of the model.

Next, we consider two spatial features based on the AIS data and the predicted destination estimated in the previous subsection. The first is the great-circle distance from each AIS point to the predicted destination, which provides a direct spatial metric. We use the Haversine formula to calculate the distance between these two points. The second spatial feature is the cumulative path distance calculated using the A* algorithm. It involves determining the distances between consecutive nodes and summing them to compute the total distance from the current AIS point to the estimated destination. These spatial features capture both the shortest and the navigable paths, offering valuable insights for improving the accuracy of VAT prediction.

Next, we focus on time-related features. First, we consider the speed of each vessel reported in the AIS data. Using the current speed, we calculate the estimated remaining voyage time of the vessel in three ways: one based on the great circle distance, another derived from the distance calculated using the A* algorithm, and finally, by subtracting the AIS reported time from the reported ETA of the vessel. These features provide diverse perspectives on the vessel's remaining journey time.

Finally, we introduce three innovative waterway features: vessel traffic flow, the number of nodes a vessel passes through, and the current water level at the vessel's location—factors not covered in previous VAT research. Vessel traffic flow refers to the number

of vessels passing through a specific area within a given time frame. This feature significantly impacts the arrival time of vessels in inland waterways due to its influence on navigation efficiency and congestion. High traffic density often leads to delays caused by limited channel capacity, longer waiting time at locks, and increased risks of traffic conflicts, especially in narrow or heavily trafficked sections. In contrast, smoother traffic flow with effective scheduling and real-time traffic management can reduce delays and enhance the predictability of VAT. In our study, we analyze traffic flow data from the past 10, 20, 30 minutes for the nodes where vessels report AIS data. The number of nodes through which a vessel passes represents the total number of nodes along the future route of the vessel, reflecting the distance traveled and potentially influencing VAT. A longer route with more nodes could indicate a longer remaining travel time, which would affect the estimated arrival time. Lastly, changes in water levels directly impact the navigability of waterways. Higher water levels generally increase channel depth, reducing the risk of grounding and enabling faster navigation, thus reducing VAT. However, excessively high water levels can cause strong currents, making navigation more challenging and potentially causing delays. In contrast, low water levels can restrict channel depth, requiring vessels to reduce their load or speed and increasing VAT. Fluctuations in water levels can also affect the efficiency of locks and port operations, further influencing VAT. The features used to predict VAT, along with the ground truth values, are summarized in Table 4.7.

Table 4.7: VAT prediction features description

Feature category	Detailed features	Explication	Note
Vessel dimensions	$Length$	The measurement of the vessel from the front to the back	Matching based on vessel MMSI value
	$Width$	The width of a vessel at its widest point	Matching based on vessel MMSI value
	$Grosstonnage$	The overall internal volume of the vessel	Matching based on vessel MMSI value
	$Type$	The type of the vessel	Obtain from AIS data, one-hot vector encoding
Spatial features	$Exp\_directdistance$	Expected remaining voyage distance	Estimate from AIS data
	$Exp\_astardistance$	Expected remaining voyage distance based on A* algorithm	Estimate from AIS data
Temporal features	$Exp\_aistime$	Expected remaining voyage time from AIS data	Estimated remaining distance divide real-time speed
	$Exp\_astartime$	Expected remaining voyage time from A-star algorithm	Estimated remaining distance from A-star algorithm divide real-time speed
	$Speed$	Vessel real-time speed	From AIS data
	$Exp\_etatime$	Expected remaining voyage time from reported ETA data	ETA time minus AIS report time
Environmental features	Temperature	The temperature of Rotterdam port	Obtain from meteorological station
	River speed	The current river velocity in the vicinity of the vessel	Obtain from meteorological station
	Wind speed	The current wind speed in the vicinity of the vessel	Obtain from meteorological station
Waterway features	$Trafficflow$	The number of vessels in port when the vessel reports AIS data	Count from A* algorithm
	$Nodenumber$	The number of nodes traversed by a vessel along its remaining voyage	Count from A* algorithm
	$Waterlevel$	The height of the water surface	Obtain from meteorological station
Prediction target/ Ground truth value	$Actual\_remaining\_time$	The actual remaining voyage time when vessel reports AIS data	ATA minus the report time of AIS

4.2.2.5 VAT prediction result analysis

In this subsection, we present a comprehensive analysis of VAT prediction. Following the traditional approach, we first split the integrated data set into training and testing sets in a 4:1 ratio. As discussed in previous sections, a VAT prediction model is trained for each record in the integrated dataset, and its performance is compared against the original ATA. To emphasize the significance of various feature parameters, we also perform feature ablation by selectively removing specific feature groups and retraining the model. Additionally, the dataset is divided into time slices (based on remaining voyage time) and distance slices (based on remaining voyage distance) to evaluate the model's VAT prediction performance under different scenarios.

To enhance the realism of port operations simulation, we introduce an innovative data segmentation approach based on time and distance. In real-world scenarios, randomly splitting the data could result in AIS data close to the port being included in the training set, while distant data is assigned to the testing set. This setup is impractical for actual port operations, as it implies using future data to train models on past data. To address this issue, we segment the dataset by AIS report time and remaining voyage distance. Specifically, the first 80% of the AIS data, which corresponds to earlier report time and greater distances, is assigned to the testing set, while the remaining 20% forms the training set. This ensures a more realistic and practical application of the model in real-world port scenarios.

Lastly, to provide a baseline for comparison with the prediction

model, we include two additional methods: the vessel-reported ETA (denoted as Vessel ETA in the table) and the estimated arrival time (Estimated Time), which is calculated using the great-circle distance between the current AIS point and the port. We evaluate all approaches using four metrics: MAE, RMSE, R^2, and Median Absolute Deviation (MAD). This comprehensive evaluation framework highlights the model's predictive performance and robustness.

Table 4.8: Prediction performance of the models on the original test dataset

Model	**All variables**				**All variables except ETA features**				**All variables except AIS data**				**All variables except Weather**			
	MAE	RMSE	R^2	MAD	MAE	RMSE	R^2	MAD	MAE	RMSE	R^2	MAD	MAE	RMSE	R^2	MAD
Vessel ETA	18.19	31.56	-0.08	18.20	-	-	-	-	18.19	31.56	-0.08	18.20	18.19	31.56	-0.08	18.20
Estimated time	2.24	10.54	-0.02	5.23	2.24	10.54	-0.02	5.23	-	-	-	-	2.24	10.54	-0.02	5.23
RF	0.20	0.61	0.90	3.73	0.33	0.94	0.77	0.33	1.21	5.22	0.53	1.22	0.25	0.88	0.66	0.23
LightGBM	**0.10**	**0.29**	**0.99**	**3.00**	**0.38**	**1.01**	**0.98**	**3.03**	**0.87**	**3.32**	0.85	**3.00**	**0.88**	3.23	0.87	**3.01**
XGBoost	0.11	0.35	0.98	3.00	0.39	0.38	1.02	3.00	0.88	3.45	**0.87**	**3.00**	0.88	**3.22**	**0.88**	**3.01**
LSTM	0.18	0.44	0.89	4.18	4.17	0.99	0.78	0.30	1.24	5.32	0.43	1.24	0.30	0.80	0.64	0.30
ANN	0.21	0.44	0.80	5.21	0.31	0.94	0.67	0.31	1.22	5.44	0.44	1.21	0.32	0.89	0.49	0.32

Table 4.8 provides a comprehensive analysis of prediction performance across different models and feature inclusion scenarios for VAT prediction in the original integrated dataset. The vessel-reported ETA performs poorly, with high errors (e.g., MAE: 18.19, $RMSE$: 31.56) and a negative R^2 value, indicating unreliable self-reported data. In contrast, the estimated time, derived from vessel coordinates, speed, and destination distance, achieves significantly better accuracy (MAE: 2.24, RMSE: 10.54), demonstrating that geometric estimations are far more reliable than self-reported values. Among the ML models, LightGBM and XGBoost consistently outperform others, achieving the lowest errors (e.g., MAE: 0.10, $RMSE$: 2.99, R^2: 0.99) when all variables are included. This performance remains robust even when certain features are excluded, with a moderate increase in errors (e.g., MAE: 0.12, RMSE: 3.08 without ETA fea-

tures).

Excluding AIS data causes the most significant performance degradation across all models, as AIS data likely captures dynamic information (e.g., speed, direction) crucial for accurate predictions. LightGBM, while affected, remains relatively stable (MAE: 0.33, RMSE: 3.22), highlighting its adaptability. Weather data exclusion has a comparatively smaller impact on performance, suggesting that weather information plays a less critical role than AIS or ETA features. Comparatively, XGBoost performs competitively under full-feature scenarios (e.g., MAE: 0.11, RMSE: 3.05), but its performance deteriorates more sharply than LightGBM when key features are excluded.

Table 4.9: Prediction performance of models on the distance-split test dataset

Model	**All variables**				**All variables except ETA features**				**All variables except AIS data**				**All variables except Weather**			
	MAE	RMSE	R^2	MAD	MAE	RMSE	R^2	MAD	MAE	RMSE	R^2	MAD	MAE	RMSE	R^2	MAD
Vessel ETA	22.38	36.70	-0.06	22.28	-	-	-	-	22.38	36.70	-0.06	22.28	22.38	36.70	-0.06	22.28
Estimated time	2.35	10.34	-0.02	4.23	2.35	10.34	-0.02	4.23	-	-	-	-	2.35	10.34	-0.02	4.23
RF	2.01	8.33	0.11	3.83	2.17	9.94	0.32	4.22	4.82	14.17	0.11	9.73	3.11	11.91	0.33	2.75
LightGBM	**1.80**	7.70	0.73	**3.01**	**1.90**	**8.55**	**0.71**	**3.06**	3.52	**10.52**	0.63	2.06	**2.35**	9.54	**0.62**	1.92
XGBoost	**1.80**	**7.65**	**0.75**	3.06	1.91	8.68	0.70	3.66	**3.48**	10.76	**0.66**	**2.03**	2.37	**9.45**	**0.62**	**1.46**
LSTM	1.99	8.71	0.29	4.11	2.10	9.57	0.30	4.00	5.33	15.32	0.23	11.73	3.33	11.99	0.14	2.30
ANN	2.23	10.11	0.05	4.20	2.11	10.11	0.03	4.30	7.22	17.41	0.04	19.02	4.18	12.12	0.09	3.32

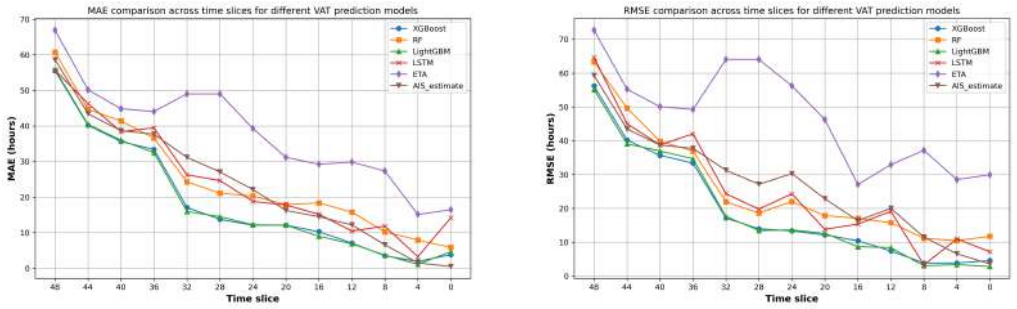
Table 4.10: Prediction performance of models on the time-order test dataset

Model	**All variables**				**All variables except ETA features**				**All variables except AIS data**				**All variables except Weather**			
	MAE	RMSE	R^2	MAD	MAE	RMSE	R^2	MAD	MAE	RMSE	R^2	MAD	MAE	RMSE	R^2	MAD
Vessel ETA	17.06	29.99	-6.41	5.20	-	-	-	-	17.06	29.99	-6.41	5.20	17.06	29.99	-6.41	5.20
Estimated time	4.98	14.03	-0.02	2.26	4.98	14.03	-0.02	2.26	-	-	-	-	4.98	14.03	-0.02	2.26
RF	5.20	13.99	0.21	2.33	4.91	13.24	0.41	2.11	5.91	15.95	0.20	3.88	5.51	15.21	0.21	2.91
LightGBM	3.53	**11.03**	**0.61**	**1.41**	4.07	11.48	**0.52**	1.68	4.23	**11.32**	0.60	2.31	3.56	11.01	0.61	1.51
XGBoost	**3.49**	11.04	0.69	1.40	**3.53**	**11.21**	**0.52**	**1.59**	**4.14**	11.92	**0.63**	**1.98**	**3.55**	**10.80**	**0.62**	**1.41**
LSTM	5.33	15.04	0.09	3.00	5.73	15.73	0.33	3.30	6.44	17.23	0.12	4.33	5.48	15.81	0.05	3.21
ANN	7.21	18.19	-0.01	3.33	7.83	18.19	0.03	4.11	8.02	19.15	0.00	5.27	7.51	16.92	-0.05	4.01

The analysis of the distance-split and time-order test datasets in Table 4.9 and 4.10 reveal distinct impacts on the model performance for the prediction of the arrival time of the vessel. In both cases,

vessel-reported ETA performs poorly, with high errors and no sensitivity to feature inclusion. At the same time, the estimated time offers moderate improvement but remains less effective than ML models. LightGBM consistently achieves the best results across both splits, with the lowest errors and highest R^2. However, its performance declines in the time-order dataset (e.g., MAE increases from 1.80 to 3.53) due to increased temporal variability. Compared to LightGBM, XGBoost performs slightly worse, while neural network models like LSTM and ANN struggle more in the time-order dataset, showing significant increases in error metrics. Feature importance analysis highlights AIS data as the most critical input, with its exclusion causing the largest performance degradation across all models, especially in the time-order split where temporal trends are more pronounced. In contrast, the exclusion of weather and ETA features has a more minor impact.

To comprehensively evaluate the practical predictive capabilities of different VAT prediction models across various time intervals, we segment the dataset based on time order following the previously adopted division approach. Specifically, the dataset is divided into 49 time slices at 1-hour intervals ranging from 0 to 48 hours, where the 0-th slice contains instances with an actual remaining time to arrival of 0 to 1 hour, while the 48-th slice includes cases where the remaining time is greater than or equal to 48 hours.



(a) MAE error trends for different time slices

(b) RMSE error trends for different time slices

Figure 4.14: Comparison of prediction errors across time slices for different models

The results presented in Figure 4.14 illustrate that the MAE and RMSE of the prediction models generally decrease as the remaining time to arrival shortens, suggesting that VAT predictions become more reliable as vessels approach their destinations. Among the ML models, XGBoost, and LightGBM demonstrate superior predictive performance, with consistently lower errors across all time slices, particularly for shorter horizons. LSTM and RF exhibit greater fluctuations, indicating potential instability in certain time intervals, which may stem from its sensitivity to sequence dependencies and variations in training data. In contrast, traditional estimation methods, including ETA (vessel-reported ETA records) and AIS_estimate (derived by dividing AIS-reported distance by vessel speed), perform notably worse in long-term predictions, with significantly higher MAE and RMSE values. The high error rates associated with ETA suggest that self-reported arrival estimates may be subject to operational uncertainties, human biases, or unaccounted delays.

However, an interesting trend emerges as vessels get closer to the

port: AIS_estimate, despite its overall underperformance in long-term predictions, demonstrates increasing accuracy in short-term VAT assessments, sometimes outperforming ML models and vessel-reported ETA. This phenomenon can be attributed to the nature of AIS-based distance evaluation, which provides a more direct measurement of a vessel's position relative to the destination. As the vessel nears the port, factors such as real-time navigation adjustments, external influences (e.g., port congestion, weather conditions), and vessel speed fluctuations have a diminished effect on distance-based estimations, making AIS-based calculations more precise. In contrast, ML models, while trained on historical data, may be influenced by overall trends in the dataset, leading to suboptimal performance in specific short-distance cases where real-time dynamic factors play a crucial role. The vessel-reported ETA also tends to improve as arrival nears, but its accuracy remains dependent on external factors, operational schedules, and reporting delays. The analysis underscores the effectiveness of ML models, particularly XGBoost and RF, in providing reliable VAT predictions across most time slices, while also highlighting the short-term reliability of AIS-based estimations as vessels approach the port. The observed error trends suggest that a hybrid approach, integrating ML models for long-term VAT forecasting and AIS-based estimations for short-term corrections, could further enhance practical VAT prediction accuracy in real-world maritime operations.

Next, we present and analyze the importance of the trained LightGBM model for VAT prediction. Feature importance is evaluated

using two approaches: First, the feature importance score provided by LightGBM, which quantifies the importance of each feature based on the number of splits. Second, the SHAP values, which explain the model's predictions by attributing the contribution of each feature to the output (Scott, Su-In, et al. 2017). The 10 most important characteristics and their corresponding scores are shown in Figure 4.15.

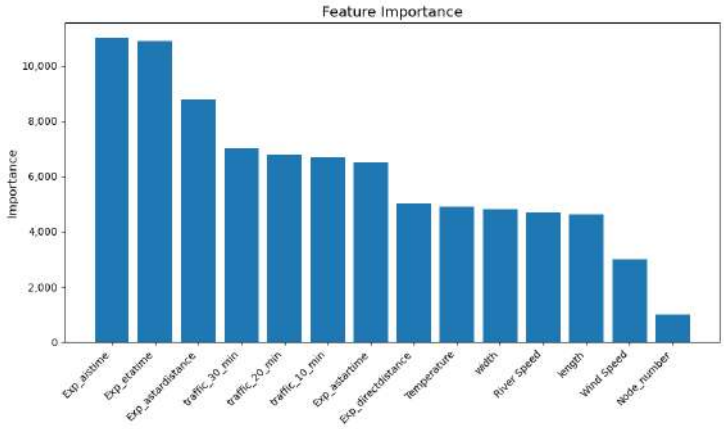


Figure 4.15: Feature importance score analysis for VAT prediction

The feature importance analysis in Figure 4.15 reveals that "exp_aistime" and "Exp_etatime" are the most influential features. Notably, "exp_aistime" represents the estimated voyage remaining time calculated from the vessel's latitude and longitude, while "Exp_etatime" is derived from the vessel-reported ETA. These features highlight the critical role of temporal and spatial factors in VAT prediction. "exp_astardistance" and "exp_atartime also rank highly, further emphasizing the significance of spatial metrics such as remaining voyage distance and time from the a star algorithm. Traffic flow-related features (traffic_30_min, traffic_20_min, and traffic_10_min) contribute meaningfully, reflecting the impact of

traffic flow and the port congestion within different time windows. Additionally, “Exp_directdistance”, which represents the estimated voyage distance calculated from AIS coordinate, shows moderate importance. Environmental and vessel-specific characteristics, such as temperature, river speed, wind speed, vessel width, and length, also exhibit importance, indicating their influence on VAT prediction.

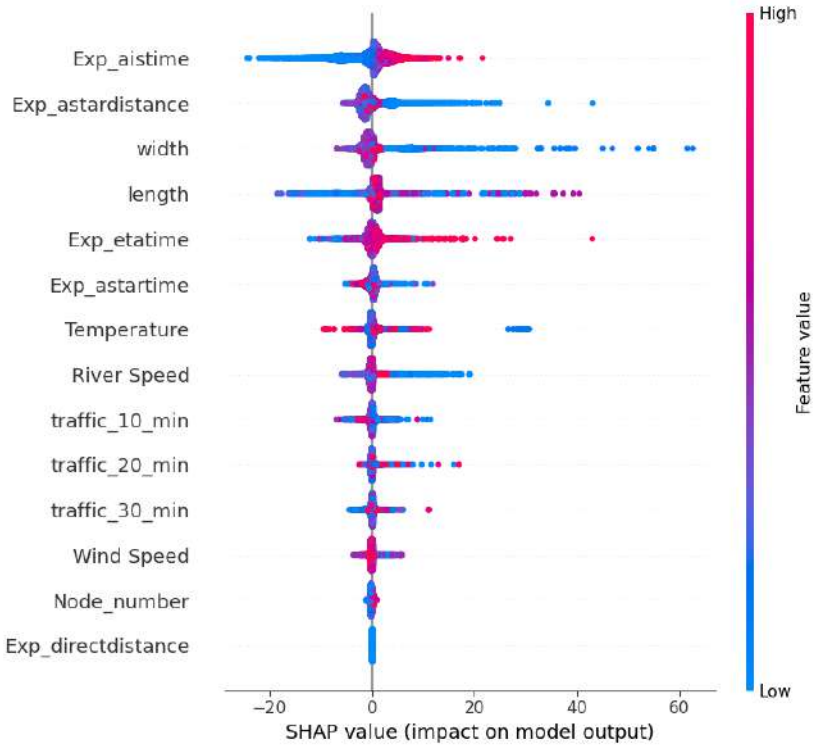


Figure 4.16: Analysis of feature SHAP value

The SHAP value visualization provides a comprehensive understanding of feature importance, with red indicating high feature values and blue representing low values. This analysis not only identifies the most critical predictors, but also reveals the nature of their impact on

the arrival time of the vessel. The results reveal that key features such as “exp_aistime”, “exp_astardistance”, and “Exp_aistime” have the most significant influence, with higher values for these features, such as longer expected travel time or distances, being strongly associated with increased predicted VAT.

Distance-related features, such as “distance_a_star” and “Node_number”, also play a crucial role, indicating the importance of route length in determining arrival time. Physical attributes of the vessel, such as “Width” and “Length”, contribute modestly, likely due to their indirect effects on navigational constraints or port operations. Environmental factors like “Temperature”, “Wind speed”, and “River speed” have a smaller but discernible impact, potentially linked to seasonal or weather-related delays. Meanwhile, traffic flow-related features, demonstrate a noticeable influence, highlighting the role of port congestion and operational delays.

Overall, the analysis highlights the critical role of AIS data and the effectiveness of geometric estimations in predicting VAT. LightGBM and XGBoost stand out as the most reliable models, delivering high accuracy and robustness across various feature conditions. When using traditional random data partitioning, these models significantly reduce VAT prediction errors. However, under scenarios simulating real port operations with data split by time or distance, their performance declines compared to random partitioning. Nevertheless, the predictions remain substantially more accurate than vessel-reported ETAs or VAT estimates based on latitude and lon-

gitude. The findings indicate that integrating advanced ML models with precise geometric calculations and additional features can greatly improve the reliability of VAT predictions, especially in cases where self-reported ETAs are unreliable.

4.3 VTT and VST Prediction

In this subsection, we aim to develope a data-driven approach for forecasting the VTT/VST at the HKP.

4.3.1 Feature engineering

Feature engineering is a crucial aspect of ML, focused on creating more effective data representations from the original dataset to capture the essential elements of a problem, ultimately enhancing a model's predictive performance. In our study, the main steps in feature engineering are feature selection and fusion, which aim to identify and incorporate the most relevant variables into the dataset for prediction. In the context of vessel operations, VTT refers to the total time from a vessel's arrival at the port to its departure, while VST specifically measures the duration from the vessel's arrival at the berth to its departure. Since the two prediction tasks share similar characteristics, we focus on the feature engineering process.

Our method improves the accuracy of VST prediction by integrating four key feature categories that significantly influence VST: generic vessel features, temporal features, operational berth features and generic berth features. The actual remaining VST for each vessel

record serves as the ground truth in our prediction task. The specific description of the related features are as follows:

Vessel generic features

Vessel GT and dimensions, as part of generic vessel features, are crucial for VST prediction at ports. GT indicates a ship's capacity, affecting berth time since larger ships often have more cargo, prolonging loading/unloading. Dimensions like length, beam, and draft influence berth allocation and thus VST. Proper assessment of these features enhances port efficiency, reducing port's service time and boosting operational competitiveness. We integrate external data from the MarineTraffic and World Register of Ships (WRS) to our case dataset (WRS 2023). Using IMO number as an identifier, we compile and integrate vessel generic feature like beam, size, GT, and length from both databases into our dataset. Additionally, vessel type and draft are directly derived from the ATA data. This integrated data enables a thorough analysis of how vessel features influence VST prediction.

Temporal features:

Temporal features are crucial for VST management at port. Accurate EDT aids in optimal resource allocation and berth utilization, while consistent ATA is key to avoiding operational disruptions. Discrepancies between EDT and ATA can lead to inefficiencies, impacting vessel schedules and port productivity. Precise timing ensures effective vessel flow management and enhances overall port efficiency. The study utilizes port call data, incorporating temporal information such as EDT, ADT, and their respective reporting time. Besides the

basic reported time data, we include two calculated temporal features: "Berth time" and "expected VST", associated with a ship's VST at the HKP. "Berth time" indicates the time a vessel has been at berth, while "expected VST" is the expected service time for each vessel. The determination of these two metrics is specifically dependent on whether the vessel is at berth when it reports the EDT data. Specifically, the "Berth time" feature applies to vessels that report their EDT when at berth. This value is calculated by subtracting the ATA from the report time of EDT. If a vessel upload the EDT record before arriving at the HKP, its "Berth time" is set to zero. "expected VST" is calculated differently depending on whether the vessel has reached its berth at the time of reporting EDT. For those reporting EDT before reaching the berth, "expected VST" is the time difference between the EDT and the ETA. For vessels reporting EDT while at berth, it is the difference between the EDT and its reported time. Furthermore, in this project, the target of our prediction, namely the actual remaining service time, is also categorized into two types based on the timing of the EDT report. For vessels that report their EDT before arriving at the berth, the actual remaining service time is calculated as the ADT minus the ATA. Conversely, for vessels that report their EDT when at berth, this value is the difference between the ATD and the reported time of the EDT data.

Berth operational features:

Berth operational characteristics significantly influence VST at ports. The number of vessels currently at port indicates the port's real time workload, where high traffic can cause delays and longer VST due

to waiting time in anchorages. Efficient berths with quick loading/unloading operations reduce VST, enhancing vessel turnover. However, berths with frequent delays can prolong VST. Incorporating these patterns is crucial for precise VST forecast and port operations optimization. This section considers four indicators: the number of vessels at berth, berth delay, vessel visit frequency, and agent performance. The amount of vessels at berth is the count of vessels at the berth when ship reports EDT data. Vessel visit frequency refers to the ship's average number of visits to the HKP per month. If no prior record exists of the vessel arriving at HKP, indicating either our data on vessel visits is truncated or it is the first that the vessel visits the HKP, the value is initially left blank and will later be filled by the average of the training set. Each vessel is represented by an agent, with "agent_delay" signifying the historical VST average delays under a specific agent's management. This value can reflect the agent's performance and is dynamically updated with each vessel arrival . Essentially, it serves as a metric for evaluating the efficiency of different agents in managing their vessels' port stays.

Berth generic features:

Berth characteristics like length, width, and depth are key to efficient port operations, impacting VST and cargo handling. Longer and deeper berths accommodate bigger ships, requiring more time for docking, undocking, and cargo processing. Properly managing these attributes is vital for enhancing berth operational efficiency. The HKP provides comprehensive information in the berth guide (Hong Kong Government 2023). Our study incorporates data on berth length,

and berth depth as berth characteristics. This data inclusion is intended to improve VST prediction and understanding by considering key berth properties. The features used for VST prediction are described in Table 4.11.

Table 4.11: VST prediction feature description

Feature type	Detailed feature	Feature description	Source
Vessel generic features	Beam	Beam of the vessel	WRS
	Size	The size of a vessel	WRS
	Gross tonnage (GT)	Vessel GT	WRS
	Length	Vessel length	WRS
	Max_draft	The maximum distance between between the water surface and the vessel's keel	ATA
	Vessel_type	Type of arrived ship	ATA
Temporal features	ATA_day	Day of the week for ATA	ATA
	ATA_hour	The hour shift of ATA	ATA
	EDT_day	Week day of EDT	EDT
	EDT_hour	Hour shift of EDT	EDT
	EDT_season	Season shift of EDT	EDT
	ATA_season	ATA season shift	ATA
	Berth time	The duration a vessel has already been in berth	EDT
	Expected VST	Vessel EDT minus the report time of the EDT	ATA and EDT
Berth operational features	Inport_number	The count of vessels in port when the ship reports its EDT data.	In port vessels
	ATA_berth_delay	The mean value of historical delay for the berth where the vessel arrives	ATA
	EDT_berth_delay	The mean value of historical delay for the berth where the vessel reports the EDT	EDT
	Vessel_delay	The mean value of vessel historical VST delay	ADT and EDT
	Vessel visit frequency	The average visit time of the vessel in one month	ATA
	Agent_performance	The average VST delay represented by the agent	ADT and EDT
Berth generic features	ATA_berth max draft	The maximum distance between the water surface and the vessel's keel upon arrival at the berth	HKP berth guide
	Overall max length ATA_berth	The maximum length of the berth where the vessel arrives.	HKP berth guide
	ATA_berth length	The length of the berth at which the vessel docks.	HKP berth guide
	EDT_berth max draft	The maximum length of the berth at which the ship is located when reporting the EDT.	HKP berth guide
	Overall max length EDT_berth	The berth overall length at which the ship reports its EDT	HKP berth guide
	EDT_berth max length	The berth length at which the ship reports the EDT	HKP berth guide
Prediction target	Actual VST	Difference between vessel ADT and ATA	ADT and ATA

4.3.2 VTT prediction results

we employ XGBoost, RF, BPNN, LR and ridge regression on the dataset to evaluate their predictive performance in estimating VTT. Subsequently, we compare the predicted VTT by these models against the EDT records provided in the original test set, using MAE, RMSE and R^2 as metrics for comparison. Additionally, to assess the impact of EDT and its related features on the prediction of VTT, we perform an analysis where EDT-related features, specifically "EDT_berth" and "expected VTT", are excluded from the dataset. Furthermore, to demonstrate the superiority of our model and the importance of selecting novel parameters, we also conducted a comparative analysis with the model and parameters used in the paper by Štepec, Martinčič, Klein, Vladušič, and Costa (2020). In this comparison, we applied their parameters and model to our dataset. This means that features like vessel maximum draft, berth generic features, and EDT-related features are excluded from the dataset for training. Following this modification, we then proceed to retrain the RF and XGBoost models, using the revised dataset to conduct further prediction tasks. The prediction results are presented in Table 4.12.

Table 4.12: VTT prediction results by different ML models

Model	RMSE	MAE	MSE	R^2
Test_set	8.00	5.12	64.51	None
Decision_tree	6.64	4.32	44.09	0.761
RF	6.15	4.04	37.81	0.795
XGBoost	**6.06**	**3.94**	**36.80**	**0.804**
BPNN	8.21	5.33	67.41	0.02
Ridge_regression	7.67	4.99	58.83	0.465
LR	7.71	5.11	59.44	0.441
RF_no_edt	7.89	5.08	57.61	0.119
XGBoost_no_edt	7.71	4.83	51.99	0.174
RF_S	7.90	5.08	62.41	0.108
XGBoost_S	7.75	4.88	60.06	0.165

In Table 4.12, the row labeled "Test_set" refers to the error between the original EDT that is reported by individual ships and the real ADT in the test set. We use this data as the basis for comparing the results of our model predictions. "RF_no_edt" indicates the results predicted by the RF model, which does not take into account any EDT related information. Similarly, "XGBoost_no_edt" refers to the results predicted by the XGBoost model, again without considering any EDT related features. "RF_S" and "XGBoost_S" indicate the prediction results obtained by applying the RF and XGBoost models, along with the features proposed by Štepec, Martinčič, Klein, Vladušič, and Costa (2020). to our dataset. The results presented in Table 4.12 reveal that, except for the BPNN, all the other seven VTT prediction models show reasonably more accurate results on the test set, when compared with the initially reported EDT data. In addition, the BPNN performance lags behind, even underperforming in comparison to the original test data. The primary reason for

this discrepancy is that BPNN with one hidden layer, as opposed to tree-based models, do not yield as effective results when applied to tabular data (Grinsztajn, Oyallon, and Varoquaux 2022). Regarding the models that have shown predictive improvements, tree-based models work well in the prediction tasks and XGBoost achieves the best results across all four metrics. The delay error, as measured by the MAE evaluation metric, decreases from 5.12 hours in the original EDT test set to 3.94 hours as predicted by the XGBoost model, marking a decrease of 23%. The RMSE also sees a decrease from 8.0 hours to 6.06 hours with a reduction of 24.3%. The MSE drops from 64.51 h^2 to 36.80 h^2, indicating a substantial decrease of 43%. Furthermore, XGBoost also achieves the best performance in the R^2 evaluation, yielding a score of 0.804.

Table 4.12 also illustrates that if the RF model or XGBoost model do not consider EDT related features during prediction, the effectiveness of the model prediction would significantly diminish. The VTT predictive results from XGBoost and RF without considering EDT related features are only marginally better than the original EDT test set data, with the XGBoost model reducing the delay error by a mere 0.11 hours for RMSE and 0.04 hour for MAE error. The results shown in the RF_no_edt and XGBoost_no_edt rows in Table 4.12 emphasize the significant role that EDT and its related features play in real-world port operation and predicting the actual VTT. Although there may be error in the reported EDT, this feature is still the most crucial basis for VTT evaluation and prediction when ATA is known. As for the comparative experiments with the approach

of Štepec, Martinčič, Klein, Vladušič, and Costa (2020), the results show that the RF_S and XGBoost_S models, which presumably apply the features from Štepec, Martinčič, Klein, Vladušič, and Costa (2020), demonstrate a reduced performance across all metrics compared to our proposed prediction models. Specifically, RF_S and XGBoost_S have higher RMSE values (7.90 and 7.75, respectively) compared to our RF and XGBoost models (6.15 and 6.06, respectively), indicating that our models outperform in terms of RMSE, with the lowest RMSE observed in our XGBoost model. For MAE, which similarly benefits from lower values, RF_S and XGBoost_S again show inferior performance compared to our RF and XGBoost, which have the lowest MAE values among all models (4.04 and 3.94, respectively). The situation is similar for MSE and R^2 metrics. Overall, the experimental results corroborate the effectiveness of our approach. The VTT prediction results can be further analyzed to generate strategic and managerial implications for policymakers and port practitioners.

In addition to evaluating model performance, we also identify the top ten most important features for VTT prediction from the constructed XGBoost model. The feature importance score can be automatically computed using a built-in function, XGBoost.feature_importances_, from the XGBoost Python library (T. Chen, T. He, M. Benesty, Khotilovich, Y. Tang, Cho, K. Chen, Mitchell, Cano, T. Zhou, et al. 2015). Feature importance is presented here in the "total_gain" method, where the sum of all feature importance values equals 1 and it reflects the contribution

of each feature towards improving the model's accuracy. The higher the score, the more significant the feature is in terms of its importance to the model. Feature importance of the top 10 most important features in the XGBoost model constructed is shown in Table 4.13.

Table 4.13: The top 10 most importance features and their importance scores of the constructed XGBoost model

Rank	Feature	Importance score
1	Expected VTT	0.5651
2	Length	0.1329
3	Beam	0.1004
4	GT	0.0812
5	Max_draft	0.0542
6	ATA_inport	0.0358
7	EDT_berth_dealy	0.0310
8	Vessel_delay	0.0219
9	ATA_shift_2	0.0181
10	ATA_berth_delay	0.0108

The feature importance analysis presented in Table 4.13 indicates that several key factors primarily determine the performance of VTT. These factors include the expected VTT (calculated as the vessel's EDT minus its ATA), vessel generic features (such as length, beam and GT), the number of vessels in port at the ATA time and the average historical turnaround time delay data for both the berth and the vessel. Specifically, the expected VTT is the most critical determinant among the features, as it is reported by the vessel itself and the port uses it as a reference for berth scheduling planning. The vessel's length and beam, which rank second and third, respectively,

directly influence the allocation of berth (J. Yu, G. Tang, Song, X. Yu, Qi, D. Li, and Y. Zhang 2018). Additionally, the vessel's GT and the maximum draft during navigation, ranking fourth and fifth, as they reflect the carrying capacity of the container vessels, and thus affect the turnaround time. As for the number of vessels in port, it reflects the busyness of the port at the time when the ship arrives at the port. When there are too many vessels in port, the port becomes busy and may be unable to effectively schedule berth allocations for the incoming vessel, which in turn can impact the actual turnaround time of the vessels. Besides, the historical delay data of the berth where the vessel's EDT data is reported, as well as the vessel's own historical delay data, have significant impacts on the vessel final turnaround time. Low historical VTT delay data of the berth reflects the efficiency of that berth, making the ADT of the vessel closer to the EDT. This, in turn, leads to more accurate results in predictive modeling. Besides, vessels with a history of less delay in VTT tend to provide more trustworthy EDT information and reliable EDT information ultimately enhances the accuracy of model predictions.

Additionally, based on the results of the feature importance analysis, we select the ten most important parameters from Table 4.13 and actual VTT for feature correction analysis and visualization. The results of the feature correlation matrix visualization are shown in Figure 4.17.

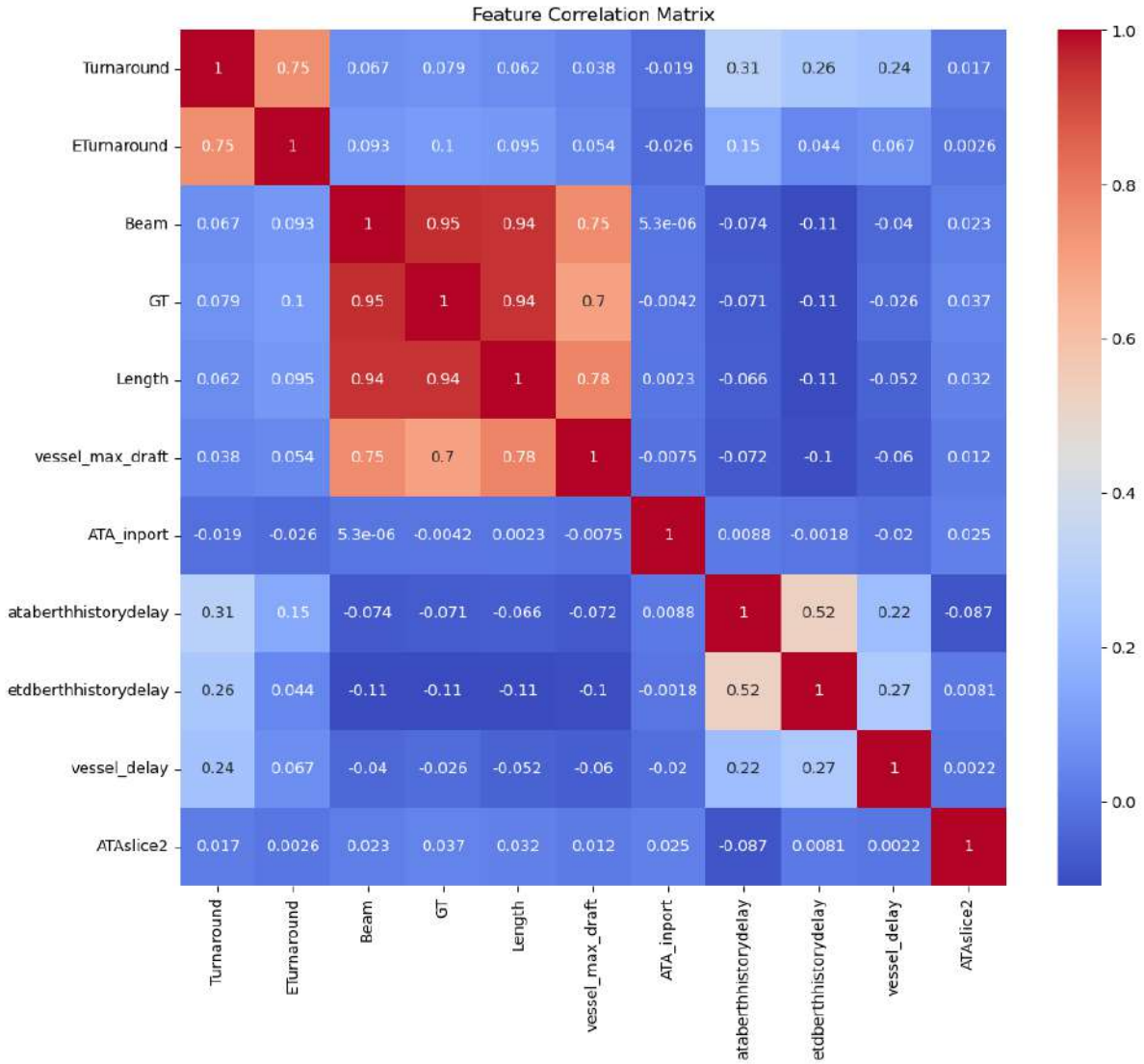


Figure 4.17: Feature correlation matrix visualization results

In correlation matrix analysis, each cell in the matrix represents the correlation coefficient between the variables on the corresponding row and column. A correlation coefficient can range from -1 to 1, where: 1 indicates a perfect positive correlation (as one variable increases, the other also increases). -1 indicates a perfect negative correlation (as one variable increases, the other decreases). In

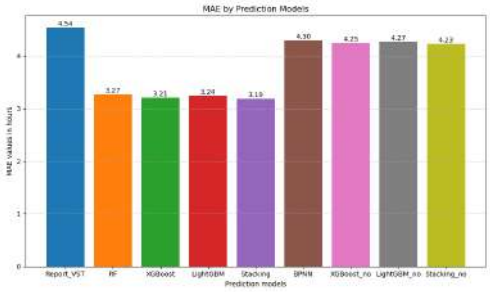
this part of the analysis, we focus on the actual VTT, represented as "Turnaround," and its related parameters. The correlation coefficient of 0.75 between "Turnaround" and the expected VTT ("ETurnaround") suggests a substantial positive relationship, indicating that longer expected VTT is often associated with longer actual VTT. The analysis reveals that the expected VTT as a significant indicator of the actual VTT which is similar to the findings from the feature importance analysis. Besides, the vessel's physical attributes such as beam, GT, length, and maximum draft show a week linear relationships with "Turnaround" feature. Port operational features including berth and vessel historical delay records, also exhibit very weak correlations with "Turnaround." This indicates that these elements potentially affect the VTT.

4.3.3 VST prediction results

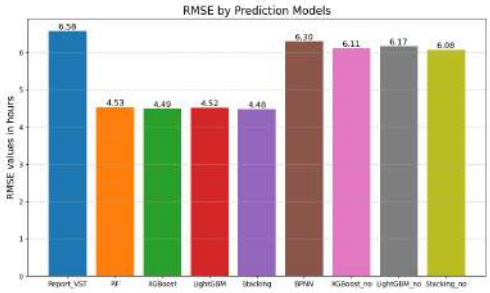
We implement tree-based stacking model and other models on the dataset to assess their efficacy in predicting VST. Furthermore, we compare the VST predicted by these models to the expected VST provided by the vessel and port, with RMSE, MSE, MAE and R^2 as evaluation metrics. Besides, to underscore the importance of the expected VST features in VST prediction, we removed this parameter and conducted a new round of training. We then compared these new predictions with the original results. The prediction outcomes are presented in Table 4.14 and the visualization of RMSE, MSE, MAE, and R^2 results can be found in Figure 4.18.

Table 4.14: Comparison of VST prediction results

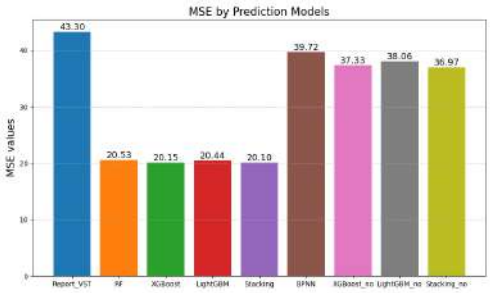
Model	RMSE	MAE	MSE	MAAPE	R^2
Report_VST[1]	6.58	4.54	43.30	0.18	0.06
RF	4.53	3.27	20.53	0.15	0.58
XGBoost	4.49	3.21	20.15	**0.13**	0.79
LightGBM	4.52	3.24	20.44	**0.13**	0.78
Stacking	**4.48**	**3.19**	**20.10**	**0.13**	**0.80**
BPNN	6.30	4.30	39.72	0.18	0.11
XGBoost_no	6.11	4.25	37.33	0.17	0.22
LightGBM_no	6.17	4.27	38.06	0.17	0.21
Stacking_no	6.08	4.23	36.97	0.17	0.22



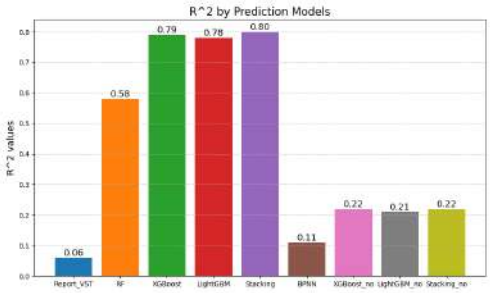
(a) MAE results by prediction models



(b) RMSE results by prediction models



(c) MSE results by prediction models



(d) R^2 results by prediction models

Figure 4.18: Comparison of different prediction models

Table 4.14 and Figure 4.18 compare various ML models based

[1]Recall that "Report_VST" represents the VST reported by the vessel at berth and it is calculated by EDT minus ETA. Alternatively, if the vessel is at berth when reporting the EDT data, Expected_VST is calculated by subtracting the report time of this EDT records from the EDT itself.

on their performance in VST prediction. "Report_VST" denotes the VST errors as reported by the vessel in the test set. Rows labeled "XGBoost_no", "LightGBM_no", and "Stacking_no" show the VST prediction results of the proposed XGBoost, LightGBM, and stacking models without incorporating the "Expect VST" feature. The "Report_VST" row data shows that the original VST reported by the vessel has an RMSE of 6.58 hours, an MAE of 4.54 hours, an MSE of 43.30 squared hours, an MAAPE of 0.18. Moreover, if we regard the reported VST by the vessels themselves as a prediction of VST,the R^2 value is 0.06, which indicates that the original VST reported by vessels is unreliable and can provide little valid information for port operations management. These values provide a benchmark for comparing prediction accuracy in our study. Table 4.14 indicates that the tree-based stacking model with the lowest RMSE outperforms the others, which decreases the RMSE from 6.58 hours in the original reported VST data to 4.48 hours in the prediction results, and the reduction rate is 31.9%. The MAE also decreases from 4.54 to 3.19 hours, marking a 29.8% improvement, and the MSE is reduced from 43.3 to 20.10 hours squared, a significant reduction of 53.6%. Additionally, the stacking model yields the highest R^2 value of 0.80 and the lowest MAAPE of 0.13, demonstrating the best fit. XGBoost and LightGBM show comparable results with slightly higher errors and R^2 values of 0.79 and 0.78, respectively. The RF model has a moderate performance, while the BPNN model exhibits the highest errors in RMSE, MAE, MSE, MAAPE metrics and the lowest R^2 value (0.11), suggesting it is the least ac-

curate model for predicting VST in our problem. The cause is that BPNN, unlike other tree models, is not good at dealing with tabular data.

Table 4.14 also shows that the effectiveness of the tree-based model predictions substantially decreases if the expected VST feature is not used for prediction. The predictive results of the proposed XGBoost, LightGBM, and stacking models, excluding the expected VST feature, show only minor improvements over the baseline values in the test set. Among these models, the stacking model maintains the best performance, despite a slight reduction in delay error: 6.11 hours for RMSE and 4.25 hour for MAE. XGBoost_no, LightGBM_no and Stacking_no rows in Table 4.14 highlight the critical role of feature the expected VST play in VST prediction. Despite potential inaccuracies in the reported and expected VST values, this feature is still the most crucial for evaluating and forecasting VST.

In addition to assessing model performance, we identify the ten most significant features for VST prediction with the XGBoost model. Feature importance is evaluated using the "XGBoost.feature_importances function" in the library (T. Chen, T. He, M. Benesty, Khotilovich, Y. Tang, Cho, K. Chen, Mitchell, Cano, T. Zhou, et al. 2015). This function calculates feature importance using the "total_gain" method, which indicates each feature's contribution to model accuracy. A higher score signifies a greater impact of the feature on the prediction model. The importance values for the top ten features in the VST prediction model are presented in Table 4.15.

Table 4.15: The analysis of top ten most important features

Ranking	Variable	Score
1	Expected_VST	0.5576
2	Berthing_time	0.083
3	Max_draft	0.055
4	GT	0.053
5	Beam	0.047
6	Length	0.046
7	Size	0.042
8	Visit_frequency	0.037
9	In_port_number	0.030
10	Agent_delay	0.029

Table 4.15 reveals various crucial factors are primary determinants of VST prediction performance. These factors include the expected VST, "Berth time", generic vessel features (max_draft, GT, beam, length, and size), the number of vessels in port, frequency of vessel visits to HKP, and their agents' historical performance. Notably, the expected VST is the most critical feature, as it is reported by the vessel and consistently used by the port authority for berth scheduling planning. "Berth time", reflecting the duration a vessel has been at the berth, ranks second in importance. The longer a vessel has been in the berth, the more operations may have been completed, potentially indicating a shorter remaining VST. Additionally, this shorter period may allow for more definitive planning and reduced uncertainty. The vessel's max_draft, GT, beam, length and size, ranking third and seventh respectively, directly impact berth allocation. The feature "Visit_frequency" ranks eighth in importance among the features. This value represents the frequency at which a vessel visits HKP each month, highlighting the significance of HKP in the

vessel's voyage routine. The "In_port_number" feature reflects the number of vessels in the port at the time a vessel reports its EDT data. This value can indicate the port's activity level or congestion at the time of the vessel's EDT report. Additionally, "agent_delay" represents the historical average of VST delays managed by a specific agent for a vessel, which is also crucial for predicting VST. Such value acts as a metric to evaluate the efficiency of various agents in managing their vessels' port stays.

Chapter 5

Optimizing Berth Allocation through Vessel Arrival Time Prediction

This chapter explores the benefits of incorporating quantitative VAT predictions into the following BAP operation. To achieve this, we propose a two-stage prediction-then-optimization framework. In the first stage, we develop a VAT prediction model to enhance the accuracy of vessel arrival estimates. In the second stage, the predicted VAT is integrated into the BAP model to optimize berth scheduling decisions. To assess the effectiveness of VAT predictions, we compare the performance of a BAP model using predicted VAT against one that relies on vessel-reported ETA across both discrete and continuous berth settings. The results demonstrate that VAT-based scheduling significantly improves berth allocation efficiency, reduces vessel waiting times, and enhances overall resource utilization. These findings underscore the advantages of data-driven forecasting in optimizing berth operations across different configurations.

5.1 BAP Model Formulation

BAP focuses on assigning available berths to incoming container vessels that require loading or unloading at a port. Given the high volume of vessels arriving daily and the limited number of berths, the primary objective of the BAP is to determine the optimal berth assignment for each vessel and the precise starting time for berth operations. Effective berth allocation plays a critical role in increasing port profitability, which can be achieved by either boosting revenue or reducing operational costs. Moreover, it is also vital to improve the port's good reputation and its customers' satisfaction. Since the primary source of revenue for operational berths comes from the handling of cargo containers, efficient management of vessel arrivals, departures, and berth resources is essential. By improving the efficiency of berth operations, the port can process more vessels within the same timeframe, leading to higher throughput and increased revenue.

Currently, BAP decisions at port are made based on the ETA reported by incoming vessels to berth. However, there is often a significant discrepancy between the reported ETA and the ATA of the vessel. This section explores the potential benefits of incorporating predicted ship arrival time, as derived in the previous section, into daily berth allocation strategies. We aim to evaluate whether VAT prediction offers advantages to BAP when compared to solely relying on the vessel's reported ETA.

In the context of BAP, there are generally two types of berths: discrete and continuous. In the discrete BAP, the quay line is di-

vided into specific berths, each of which is designated for a single vessel, limiting the berthing to the predefined locations. In contrast, in the continuous BAP, the vessels can dock at any position along the quay line without fixed restrictions, allowing flexibility in berthing locations. The illustrations of the two different types of berths are shown in Figure 5.1: In our study, we explore BAP optimization in

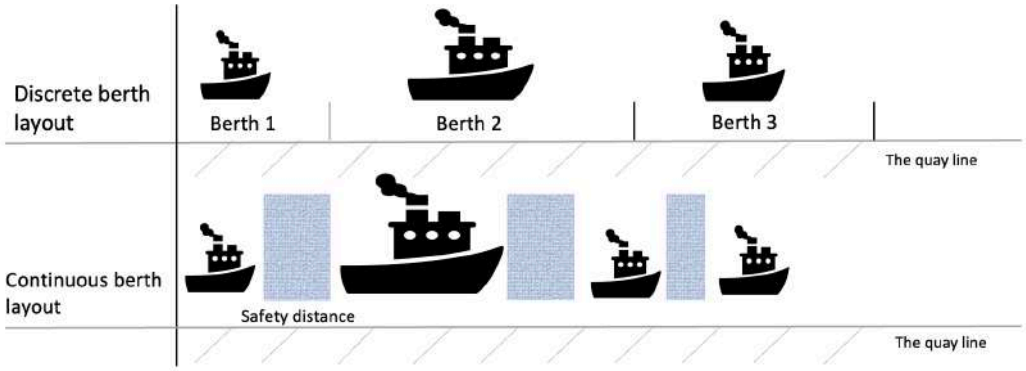


Figure 5.1: The layout of discrete and continuous berths

both discrete and continuous cases with VAT predicted by ML models as input. First, we consider the type of berth in HKP as discrete. The set of available berths is denoted by B and the set of vessels to be handled by V. According to the Hong Kong Marine Department's requirement that vessels report their ETA 36 hours prior to arrival (Hong Kong Maritime And Port Board 2024), this study focuses on the BAP for vessels scheduled to arrive within the upcoming 36-hour time frame. Traditional BAP typically assumes a planning horizon of one week (168 hours). However, the ETA time planning horizon in data-driven berth allocation frameworks poses challenges to this assumption. Consequently, the remaining time beyond this 36-hour window in the one-week planning horizon lacks reliable

data, creating a gap in the planning process. Before addressing the model in detail, the underlying assumptions are clarified as follows:

1. Every vessel reports the ETA data before arriving at the berth;
2. When a vessel arrives at the port and there is no available berth, it will anchor and wait for the berth;
3. The move time of container vessels from anchor to the berth is neglected;
4. Vessel will leave the port as soon it has finished container handling;
5. The time used for container handling is determined for each vessel;

For the model parameters, the notation of the sets, parameters, and decision variables used in the discrete BAP is summarized in Table 5.1. Let V denote the set of vessels under consideration. For each vessel $i \in V$, the expected arrival time is denoted by a_i, the requested departure time by d_i, and the vessel's length by L_i. In the discrete BAP, the maximum allowable vessel length at berth j, represented by ϕ_j, determines the docking capacity of that berth. Moreover, the total handling time for each vessel depends on its assigned berth; accordingly, we denote the handling time of vessel i at berth j ($j \in B$) by h_{ij}. From the perspective of the terminal operator, the departure delay cost coefficient for vessel i is p_i^{δ} with unit \$ per hour. Discrete BAP seeks to minimize the total cost associated with vessel berthing services, including cargo operations and departure tardiness. The output of the discrete BAP model includes the following.

Table 5.1: Notation of sets, parameters, and decision variables in the discrete BAP

Notation	Explanation
Sets	
V	The set of all vessels.
B	The set of all berths.
Indices	
i, h	The indices of vessels i and h, $i, h \in V$.
j	The index of berth j, $j \in B$.
Parameters	
a_i	The expected arrival time of vessel i, given in hours.
d_i	The requested departure time of vessel i, given in hours.
L_i	The length of vessel i, given in meters.
ϕ_j	The maximum allowable vessel length for docking at the berth j, given in meters.
h_{ij}	The handling time of vessel i at berth j, given in hours.
p_i^{δ}	The cost coefficient of departure delay for vessel i, given in \$ per hour.
M	A sufficient large constant.
Decision variables	
s_i	The berth handling starting time in the planning horizon for vessel i, which is a positive integer.
e_i	The berth handling end time in the planning horizon for vessel i, which is a positive integer.
x_{ij}	Binary, equal to 1 if vessel i is assigned to berth j, and 0 otherwise.
y_{ih}	Binary, equal to 1 if vessel h is operated at the same berth as vessel i next successor, and 0 otherwise.
f_i	Binary, equal to 1 if vessel i is the first vessel operated at the assigned berth, and 0 otherwise.
l_i	Binary, equal to 1 if vessel i is the last vessel operated at the assigned berth, and 0 otherwise.

- The berth assigned to each incoming vessel for cargo handling;
- The starting time of handling operations for each vessel;
- The time at which each vessel leaves the berth.

In terms of decision variables, once the vessel i arrives at the berth, the time at which the berth operations begin should be decided and is indicated by s_i. Similarly, the time when the berthing operations are completed should be determined and is denoted by e_i, representing the berthing end time for the vessel i. Four decision binary variables are also introduced to formulate the BAP model: $x_{ij} = 1$ if vessel i is operated at berth j and 0, otherwise; $y_{ih} = 1$ if vessel h $(h \in V, h \neq i)$ is operated at the same berth as vessel i immediately after vessel i finishes all the berth operations and 0, otherwise; $f_i = 1$ if vessel i is

the first vessel in the planning horizon that is assigned to be operated at the berth, and 0 otherwise; $l_i = 1$ if vessel i is the last vessel in the planning horizon to be operated at the assigned berth and 0, otherwise. The above discrete BAP is formulated as a mixed-integer linear programming (MILP) model as follows:

$$\min \quad \sum_{i \in V} p_i^{\delta}(e_i - d_i)^+ \tag{5.1}$$

Subject to:

$$\sum_{j \in B} x_{ij} = 1, \quad \forall\, i \in V, \tag{5.2}$$

$$s_i + \sum_{j \in B} h_{ij} x_{ij} = e_i, \quad \forall\, i \in V, \tag{5.3}$$

$$f_i + \sum_{h \in V \setminus \{i\}} y_{hi} = 1, \quad \forall\, i \in V, \tag{5.4}$$

$$l_i + \sum_{h \in V \setminus \{i\}} y_{ih} = 1, \quad \forall\, i \in V, \tag{5.5}$$

$$f_i + f_h \leq 3 - x_{ij} - x_{hj}, \quad \forall\, j \in B, \forall\, i, h \in V, i \neq h, \tag{5.6}$$

$$l_i + l_h \leq 3 - x_{ij} - x_{hj}, \quad \forall\, j \in B, \forall\, i, h \in V, i \neq h, \tag{5.7}$$

$$y_{ih} - 1 \leq x_{ij} - x_{hj} \leq 1 - y_{ih}, \quad \forall\, j \in B, \forall\, i, h \in V, i \neq h, \tag{5.8}$$

$$s_i + \sum_{j \in B} h_{ij} x_{ij} - M\,(1 - y_{ih}) \leq s_h, \quad \forall\, i, h \in V, i \neq h, \tag{5.9}$$

$$a_i \leq s_i, \quad \forall\, i \in V, \tag{5.10}$$

$$L_i x_{ij} \leq \phi_j, \ \forall\, i \in V, \forall\, j \in B, \tag{5.11}$$

$$x_{ij}, y_{ih}, f_i, l_i \in \{0, 1\}, \quad \forall\, j \in B, \forall\, i, h \in V, \tag{5.12}$$

$$s_i, e_i \in \mathbb{Z}^+, \forall\, i \in V. \tag{5.13}$$

The objective function (5.1) minimizes the total cost the cost incurred by delays in the departure of the vessel for the terminal operator. In particular, $(\alpha)^+ = \max\{\alpha, 0\}$ captures the non-negative component of the delay. Constraints (5.2) guarantee that each vessel is assigned one and only one berth. Constraints (5.3) present the vessel berthing end time. Constraints 5.4 guarantee that each vessel is either served first or follows another vessel. Constraints (5.5) specify that each vessel is served last or is followed by another vessel. Constraints (5.6) and (5.7) specify that only one vessel can be scheduled as the first and last to be served in each berth, respectively. Constraints (5.8) require that a vessel can follow another vessel in service only if both are assigned to the same berth. Constraints (5.9) define the berthing start time of each vessel and M is a sufficiently large positive number. Constraints (5.10) guarantee that the start time of the berth cannot be earlier than the arrival time of the vessel. Constraints (5.11) ensure that the length of the vessel assigned to a berth does not exceed the maximum allowed vessel length. Constraints (5.12) specify that the four decision variables are binary. Constraints (5.13) indicate decision variables s_i and e_i are positive integers. Compared to the discrete BAP, the continuous BAP scenario typically assumes a single continuous berth that can accommodate multiple vessels simultaneously, allowing vessels to dock at any position along the wharf and enabling more efficient utilization of the wharf. In addition to the description provided in Table 5.1, the supplementary notation for the sets, parameters, and

decision variables used in the formulation of the continuous BAP is summarized in Table 5.2.

Table 5.2: Notation of sets, parameters, and decision variables in the continuous BAP

Notation	Explanation
Parameters	
LB	The length of the continuous berth wharf, given in meters.
h_i	The handling time of vessel i at the berth, given in hours.
P_i	The pre-defined low-cost berthing position of vessel i, given in meters.
C_i	The berth position deviation cost coefficient for vessel i, given in \$ per meter.
M	A sufficient large constant.
Decision variables	
s_i	Positive integer, the berth staring handling time for vessel i.
x_i	Positive integer, the berth position of vessel i.
θ_{ih}^X	Binary, equal to 1 if vessel i is berthed to the left of vessel h.
θ_{ih}^S	Binary, equal to 1 if the start time of berthing for vessel h is no earlier than the departure time of vessel i.

In the continuous BAP case, the total length of the continuous berth is denoted as LB. Each vessel i is characterized by a length L_i, an expected arrival time a_i, a requested departure time d_i, and an expected handling time h_i. Before starting berthing operations, each vessel is assigned a low-cost berthing position P_i. This value is pre-defined on the basis of the operational conditions of the berth and terminal, serving as a reference to minimize overall operational costs. The cost coefficient for the deviation between the actual berthing position of the vessel i and the predefined optimal position is represented by C_i (Rodrigues and Agra 2022). The objective of continuous BAP is to minimize two key cost components:

- The deviation cost incurred by the vessels due to their berthing positions deviating from their optimal low-cost positions (Rodrigues and Agra 2022).

- The cost of departure delay associated with ships that exceed their requested departure time.

The primary decision variables in the continuous BAP model are the berthing position x_i and the start time of the berthing operations s_i for each vessel i. To assist in the formulation of the model, two binary decision variables are also introduced: θ_{ih}^X, which equals 1 if the vessel i is berthed to the left of the vessel h and 0, otherwise; θ_{ih}^S, which equals 1 if the start time of berthing for the vessel h is not earlier than the departure time of vessel i and 0, otherwise. In this way, the continuous BAP is formulated as follows

$$\min \sum_{i \in V} \left(C_i \left| x_i - P_i \right| + p_i^{\delta} \left(s_i + h_i - d_i \right)^+ \right. \tag{5.14}$$

Subject to:

$$x_i + L_i \leq LB, \quad \forall\, i \in V, \tag{5.15}$$

$$x_i + L_i \leq x_h + M \left(1 - \theta_{ih}^X \right), \quad \forall\, i, h \in V, i \neq h, \tag{5.16}$$

$$s_i + h_i \leq s_h + M \left(1 - \theta_{ih}^y \right), \quad \forall\, i, h \in V, i \neq h, \tag{5.17}$$

$$\theta_{ih}^X + \theta_{hi}^X + \theta_{ih}^S + \theta_{hi}^S \geq 1, \quad \forall\, i, h\ \in V, i \neq h, \tag{5.18}$$

$$s_i \geq a_i, \quad \forall\, i \in V, \tag{5.19}$$

$$x_i, s_i \geq 0, \forall\, i \in V, \tag{5.20}$$

$$\theta_{ih}^x, \theta_{ih}^y \in \{0, 1\}, \forall\, i, h \in V, i \neq h, \tag{5.21}$$

$$s_i, x_i \in \mathbb{Z}^+, \forall\, i \in V, \tag{5.22}$$

The objective function (5.14) aims to minimize the total cost associated with deviations in vessel berthing positions and vessel departure

delays. The first term in the objective function (5.14) captures the total cost associated with the deviation between the actual berthing position and the predefined optimal position, while the second term accounts for the total cost of the delay in departure of the vessel. Constraints (5.15) guarantee that all vessels are positioned within the limits of the berth wharf. Constraints (5.16) ensure that the vessels are spatially separated within the distance space of the wharf when the vessel i is placed to the left of the vessel h. Similarly, constraints (5.17) guarantee that vessels do not overlap in the time domain when vessel h arrives after vessel i has left. Constraints (5.18) eliminate an extreme case where $\theta_{ih}^X = \theta_{hi}^X = \theta_{ih}^S = \theta_{hi}^S = 0$, which would result in vessels i and h overlapping in both the distance domain and the time domain. Constraints (5.19) ensure that a vessel can start operating on its berth only after its arrival in the port. Constraints (5.20) define the vessel berth position p_i and the start berthing time b_i as nonnegative integer variables. Constraints (5.21) define θ_{ih}^X and θ_{ih}^S as binary variables. Constraints (5.22) indicate decision variables s_i and x_i are positive integers.

5.1.1 Linearization of the objective function for the BAP

The second term of the objective function (5.1), in the discrete BAP model, is non-linear due to the positive part function $(e_i - d_i)^+$. Such nonlinear terms increase the problem's complexity and make it incompatible with linear programming solvers. To address this, an auxiliary variable z_i is introduced for each $i \in V$, where z_i represents the positive deviation of $e_i - d_i$. This linearization introduces

the following two constraints:

$$z_i \geq e_i - d_i, \quad \forall\, i \in V, \tag{5.23}$$

$$z_i \geq 0, \quad \forall\, i \in V. \tag{5.24}$$

With these constraints, the objective function (5.1) is reformulated as:

$$\min \sum_{i \in V} p_i^{\omega} \left(s_i + \sum_{j \in B} h_{ij} x_{ij} - a_i \right) + \sum_{i \in V} p_i^{\delta} z_i, \tag{5.25}$$

subject to constraints (5.2)–(5.13), (5.23), and (5.24).

For the continuous BAP model, the objective function (5.14) includes two non-linear components: $|x_i - P_i|$, the absolute value term and $(a_i + h_i - d_i)^+$, the positive part function. Two auxiliary variables u_i and w_i are introduced to linearize the non-linear components in the objective function (5.14), with the following constraints:

$$u_i \geq x_i - P_i, \quad \forall\, i \in V, \tag{5.26}$$

$$u_i \geq -(x_i - P_i), \quad \forall\, i \in V, \tag{5.27}$$

$$z_i \geq a_i + h_i - d_i, \quad \forall\, i \in V, \tag{5.28}$$

$$z_i \geq 0, \quad \forall\, i \in V. \tag{5.29}$$

The linearized form of the objective function (5.14) is:

$$\min \sum_{i \in V} \left(C_i u_i + C_i^d w_i \right), \tag{5.30}$$

subject to constraints (5.15)−(5.22) and (5.26)−(5.29).

5.2 BAP Optimization with VAT Prediction

In the previous chapter, we presented a predictive model for VAT, leveraging vessel port call data and AIS data in an innovative manner. Building upon these predictions, this section focuses on optimizing the subsequent BAP. Before presenting the results and analysis, we first introduce the scenarios considered in this study, followed by the parameter settings for BAP and the validation steps. Additionally, to assess the robustness of the proposed approach under varying operational conditions, we conduct a sensitivity analysis by adjusting the number of arriving vessels per unit time. By analyzing the impact of different traffic loads, we aim to evaluate the effectiveness and adaptability of the berth allocation process.

To demonstrate the value of VAT prediction for BAP, three scenarios are considered:

- **Ideal case**: Assume that the ETA of the vessel a_i in Tables 5.1 and 5.2 corresponds to the ATA vessel. This scenario represents an ideal extreme situation where the terminal operator has prior knowledge of the exact ATA of the vessel. Berth allocation strategy in this case is optimal, and the resulting objective function value serves as a optimal value for comparison.

- **Vessel reported ETA case**: In this scenario, the arrival time of the vessel a_i is based on the reported ETA of the vessel. The terminal operator makes the berth allocation using this reported

ETA, and the objective function is evaluated based on the actual ATA for validation.

- **Predicted VAT case**: In this scenario, the ETA of the vessel is determined on the basis of the predicted VAT provided by the prediction model. The subsequent berth allocation is made according to this predicted VAT, and the objective function is then validated using the actual ATA of the vessel. The validated objective function value is compared with both the ideal case and the reported ETA case to evaluate the benefit of VAT prediction on berth allocation.

5.2.1 Parameters setting and validation steps

For the discrete BAP case, we consider three types of berths: small, medium, and large. Correspondingly, vessels are classified into three categories based on their length. Small vessels are those with a length of 200 meters or less, medium vessels have a length between 200 and 300 meters, and large vessels exceed 300 meters in length. A small berth can accommodate only small vessels, a medium berth can accommodate both small and medium vessels, and a large berth can accommodate vessels of all three categories. As stated in the previous section, each discrete berth can accommodate only one vessel at a time.

To enhance the realism of the port operation simulation, we extract VAT and vessel length from the HKP real-world VAT dataset. Specifically, we select vessel arrival data where the ETA report time falls within the time window of January 1, 2021, between 12:00 PM

and 1:00 PM, representing vessels that are expected to arrive within the next 36 hours. Each data record represents a vessel that reported an ETA within this period and is paired with an ATA, a predicted VAT, vessel length. Furthermore, each vessel is assigned a quay crane (QC) working hour value based on its length, ranging from 20 to 50 hours. QCs are essential for port operations, handling the loading and unloading of containers. The assigned QC working hours represent the total time required for a single QC to complete all container operations for a vessel.

For the discrete BAP, we consider three berths: "small", "medium", and "large". Correspondingly, vessels are classified into three categories based on their length. "Small" vessels are those with a length of 200 meters or less, "medium" vessels have a length between 200 and 300 meters, and "large" vessels exceed 300 meters in length. The maximum allowable vessel length for the "small" berth is 200 meters, while for the "medium" berth, it is 300 meters. A "small" berth can accommodate only "small" vessels, a "medium" berth can accommodate both "small" and "medium" vessels, and a "large" berth can accommodate vessels of all three categories, with no specific maximum vessel length restriction. As stated in the previous section, each discrete berth can accommodate only one vessel at a time.

The allocation of QCs is determined based on vessel size and berth assignment. According to statistical data, the average turnaround time for container vessels in HKP is approximately 10 hours, which includes the waiting time in the anchorage area. Based on this statistical value, we make a reasonable assumption for QC allocation:

"small" vessels are assigned 3, 4, and 5 QCs when berthed at "small", "medium", and "large" berths, respectively. "Medium" vessels are assigned 5 and 6 QCs when berthed at "medium" and "large" berths, respectively. "Large" vessels, which can only be berthed at the "large" berth, are allocated 7 QCs. Given the previously established QC working hours, this allocation ensures that the total operation time for each vessel at its most suitable berth, defined as the smallest berth that can accommodate the vessel based on its length constraints, combined with potential waiting time, is approximately 10 hours.

As for the vessel's requested departure time, we first determine its operational time at its most suitable berth. This operational time is then multiplied by a uniform distribution factor randomly drawn from the range $[1, 1.5]$. Finally, by adding the vessel's ETA a_i, vessel requested departure time d_i is obtained. The vessel departure delay is then defined as the positive part of the difference between the berth handling end time e_i and the requested departure time d_i. The cost coefficient of vessel departure delay is also set according to vessel length, with small vessels incurring a cost of \$1000/h, medium vessels \$1500/h, and large vessels \$2000/h.

Next, we consider the continuous BAP case, where the continuous BAP considers only a single berth, which can simultaneously accommodate multiple vessels for consecutive berth operations. In our study, we set the continuous berth length LB to 400 meters, which exceeds the length of any vessel in our dataset. The handling time h_i of vessel i at the continuous berth remains fixed and corresponds to the handling time of the vessel at its most suitable berth in the

discrete case. The optimal low-cost berthing position P_i for vessel i is influenced by various factors onshore, including real-time terminal and dock infrastructure operation. In our study, we first ignore the position deviation cost term $C_i\,|x_i - P_i|$ in the continuous BAP objective function 5.14 and perform optimization using the ATA values to determine an initial set of optimal low-cost berthing positions. These positions are then fixed as the optimal low-cost berthing positions in the subsequent formulation of the BAP problem. The berth position deviation cost coefficient is set equal to the vessel departure delay cost coefficient: \$1000/h for small vessels, \$1500/h for medium vessels, and \$2000/h for large vessels.

In evaluating the impact of VAT prediction on BAP performance, it is irrational to directly compare the objective function values obtained from optimizing BAP with different VAT values (ETA, ATA, and predicted VAT). This is because berth operations are scheduled based on these VAT values, while vessel ATA may deviate from ETA and predicted VAT. If ETA or predicted VAT is earlier than ATA, the berth may be reserved for the target vessel while it has yet to arrive. Conversely, if ETA or predicted VAT is later than ATA, the vessel may already be at the port as per the schedule requirements, while the berth is still occupied by another vessel, leading to additional waiting time and inefficient use of port resources. To properly validate the impact of VAT prediction on BAP, an additional verification step is required.

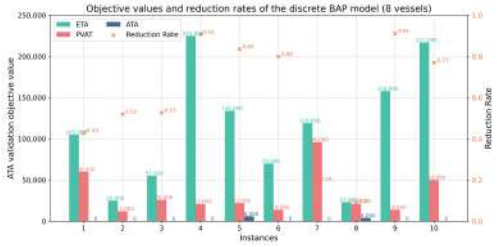
For the discrete berth case, we first optimize BAP using predicted VAT or ETA to determine the specific berth assignment for each ves-

sel and the vessel sequence at each berth. Then we fix the decision variables x_{ij} and y_{ih} from optimization in the discrete BAP model, which correspond to the berth assignment and the vessel sequence schedule, ensuring that the vessels follow the predefined berth allocation and entry order. We then re-run the optimization using ATA as the vessel arrival data to obtain the final results, reflecting a realistic operational scenario. For the continuous berth case, we first determine each vessel's starting handling time s_i through optimization and use this information to derive the vessel sequence at the continuous berth. Since the continuous berth can accommodate multiple vessels simultaneously, vessels with the same beginning service time are prioritized based on their length, with longer vessels receiving higher service priority. Based on this sequence, we introduce additional constraints on the berth starting handling time for each vessel and re-run the optimization model using ATA values to obtain the final BAP results under different VAT planning scenarios in the continuous BAP.

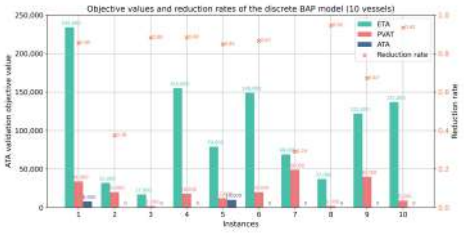
5.2.2 Discrete BAP results analysis

In this section, we first evaluate the results of the discrete BAP by analyzing scenarios in which 8, 10, 12, 14, 16, and 18 vessels arrive at the port within a 36-hour time frame. The evaluation compares the vessel-reported ETA, serving as the benchmark, with the ATA, representing the optimal value, and the predicted VAT (PVAT) results. To minimize randomness, each experiment randomly selects different vessels, and each scenario is repeated ten times. The

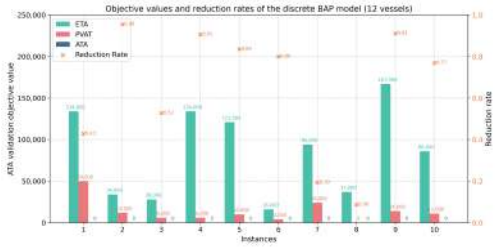
analysis focuses on the objective value of the discrete BAP model and the reduction rate achieved by the predicted VAT, calculated as $(ETA - PVAT)/ETA$. Additionally, we examine two key metrics: vessel average turnaround time, defined as the average duration from the commencement to the completion of berth operations across all vessels, and total vessel waiting time, which represents the cumulative waiting time for all vessels from their arrival at the port to the start of their respective berth operations. The impact of predicted VAT on these metrics is further assessed by computing the respective reduction rates, providing insights into the potential efficiency improvements introduced by the proposed VAT prediction model and the following BAP evaluation framework. The results are visualized in Figures 5.2, 5.3 and 5.4.



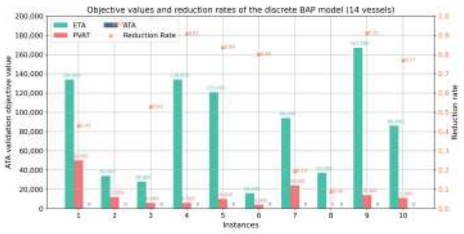
(a) Objective value and reduction rate for 8 vessels



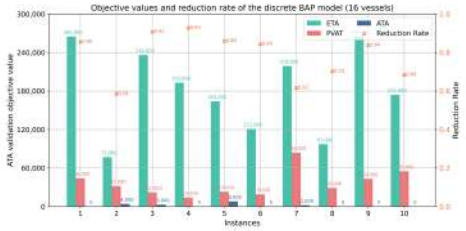
(b) Objective value and reduction rate for 10 vessels



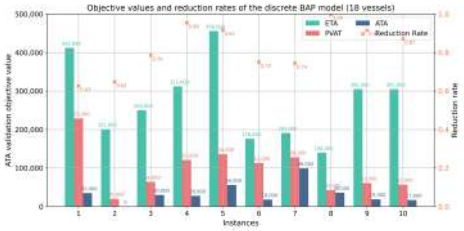
(c) Objective value and reduction rate for 12 vessels



(d) Objective value and reduction rate for 14 vessels



(e) Objective value and reduction rate for 16 vessels



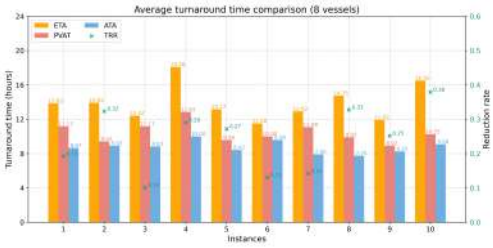
(f) Objective value and reduction rate for 18 vessels

Figure 5.2: Comparison of cost and reduction rate for different vessels in the discrete BAP

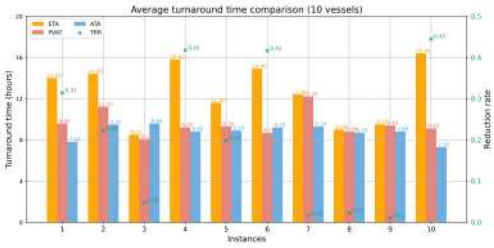
The discrete BAP objective value visualization results in Figure 5.2 show that under six different fleet sizes (8, 10, 12, 14, 16, and 18 vessels), using only ETA-based scheduling generates the highest vessel departure delay costs for the berth operator, and these costs intensify as the vessel number grows. In the 8 vessel scenario, the "ETA" case cost typically ranges from around 50,000 to 90,000$, whereas using "PVAT" reduces the related cost to roughly half or

less. For example, Instances 1 and 5 show a reduction of more than 50 percent when comparing "PVAT" to "ETA". As the number of vessels increases to 10 or 12, the "ETA" case costs can climb to approximately 90,000 or 150,000, yet PVAT still produces savings of about 50 to 80 percent. In the 14 or 16 vessel scenarios, "ETA" case values often exceed 170,000 or even 200,000, while "PVAT" remains 10,000 lower, yielding delay cost reductions between 40 and 70 percent.

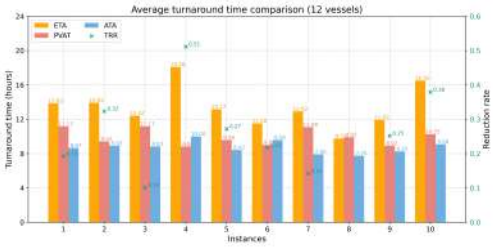
In the 18 vessel case, the contrast is especially pronounced: "ETA" case costs can exceed 300,000 or even 400,000, whereas PVAT generally stays in the 120,000 to 180,000 range, corresponding to reduction rates of 70 to 90 percent in many instances. Meanwhile, the "ATA" case, which assumes perfect prior knowledge of vessel ATA, consistently drives costs down to nearly zero, such as in the 8 vessel Instances 1 and 2, the 10 vessel Instance 2, 3. In these cases, suppose the port staff knows the vessel ATA beforehand, they can develop an optimal schedule that ensures each vessel departs on time, thereby eliminating additional operational costs. Overall, the results underscore that inaccurate or misreported arrival times impose substantial cost penalties on berth operators, whereas more precise vessel arrival time predictions can deliver significant reductions in departure delay costs. Following the objective value discussion, Figure 5.3 presents the average turnaround time under different VAT conditions (ETA, PVAT, and ATA). As might be expected, improving the accuracy of arrival predictions (switching from ETA to PVAT) generally reduces the vessel average turnaround time for most vessel numbers and in-



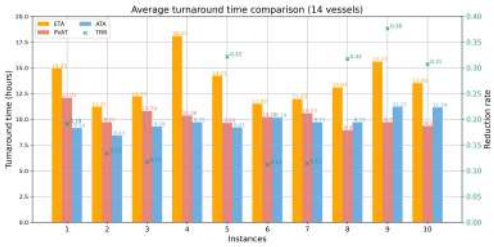
(a) Average turnaround time and reduction rate for 8 vessels



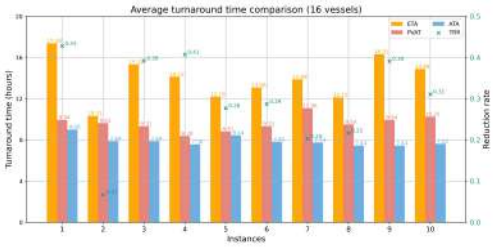
(b) Average turnaround time and reduction rate for 10 vessels



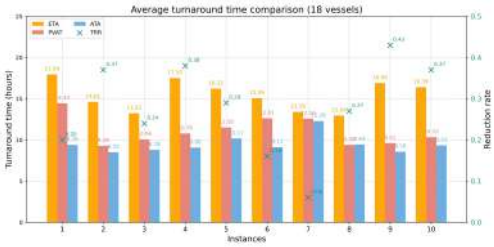
(c) Average turnaround time and reduction rate for 12 vessels



(d) Average turnaround time and reduction rate for 14 vessels



(e) Average turnaround time and reduction rate for 16 vessels

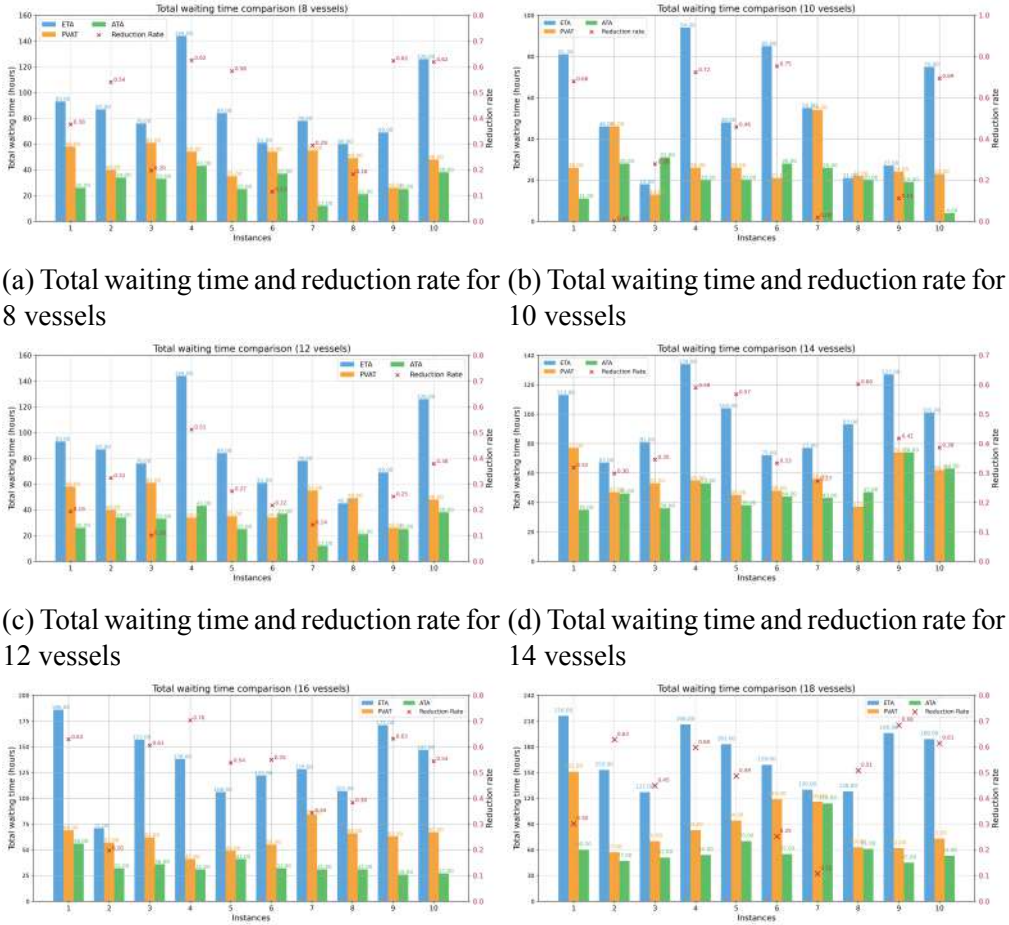


(f) Average turnaround time and reduction rate for 18 vessels

Figure 5.3: Comparison of average turnaround time and reduction rate for different vessels

stances. For example, in the 8 and 10 vessel cases in Figures 5.3a and 5.3b, the "ETA" case often lie several hours above the "PVAT" case, illustrating how ports can more efficiently plan berth assignments when equipped with better VAT estimates.

However, there are a few notable instances where the "PVAT" based solution does not outperform ATA. For example, in the 10 vessel scenario illustrated in Figure 5.3b, Instance 3 shows the average turnaround time value under PVAT that is higher than ATA by about 0.70 to 1.00 hours. A similar pattern emerges in the 12 vessel setting in Figure 5.3c, especially in Instances 4 and 6, where PVAT's TAT exceeds ATA by roughly 0.40 to 0.60 hours. This outcome stems from the fact that the BAP objective function emphasizes minimizing vessel departure delay cost, not strictly the turnaround time. Although improved VAT predictions typically allow port authorities to reduce both departure delays and turnaround time of vessels, certain edge cases can arise. In these cases, the scheduling algorithm may opt to delay a smaller vessel, despite its early arrival, so that a larger vessel can be assigned a berth slot immediately. This tactical choice lowers the overall departure delay cost, because ensuring punctual departure of a large vessel (which incurs a higher cost if delayed) yields greater savings for the port operator. Figure 5.4 presents the total waiting time of vessels under the three VAT conditions, showing that increased VAT accuracy (from "ETA" to "PVAT") generally decreases vessel total waiting time at the port, with ATA often yielding the lowest totals across most instances. Nonetheless, several cases diverge from this trend. In the 10 vessel scenario in Figure 5.4b, In-



(a) Total waiting time and reduction rate for 8 vessels

(b) Total waiting time and reduction rate for 10 vessels

(c) Total waiting time and reduction rate for 12 vessels

(d) Total waiting time and reduction rate for 14 vessels

(e) Total waiting time and reduction rate for 16 vessels

(f) Total waiting time and reduction rate for 18 vessels

Figure 5.4: Comparison of vessel total waiting time in the discrete BAP

stance 3 shows that the total waiting time for both ETA and PVAT is almost identical (e.g., 46.0 hours for both “ETA” and “PVAT” cases), “ATA” case achieves a modestly lower total waiting time of about 28.0 hours, whereas in Instance 4, both “ETA” and “PVAT” values decrease and lie below that of ATA.

An even more unexpected result is seen in Instance 8, where the ETA-based schedule yields a slightly lower total waiting time than

the PVAT-based one. As with the total waiting time, These deviations can be attributed to the objective function of the BAP primary focus on reducing departure delay costs, which is similar to what was observed for vessel average turnaround time situations. By assigning a berth immediately to a larger, high-priority vessel and temporarily delaying a smaller one, even if the smaller vessel arrives first, the port operator can avoid substantial penalty charges tied to the late departure of high-priority vessels. Although this tactic inflates the total waiting time (and consequently the turnaround time) of the delayed vessel, it ultimately delivers greater cost savings at the system level by ensuring on-time departures for larger ships.

Table 5.3: Comparative analysis of reduction rates for different vessels and metrics in the discrete BAP

Metric	**Stat**	**Vessel number**					
		8	**10**	**12**	**14**	**16**	**18**
Objective value	Max	0.91	0.95	0.96	1.00	0.93	0.90
	Min	0.09	0.29	0.09	0.63	0.58	0.33
	Mean	0.60	0.76	0.64	0.82	0.78	0.64
Average turnaround time	Max	0.38	0.45	0.51	0.43	0.43	0.33
	Min	0.10	0.01	−0.02	0.11	0.07	0.22
	Mean	0.24	0.21	0.24	0.24	0.30	0.26
Total waiting time	Max	0.63	0.75	0.76	0.60	0.70	0.68
	Min	0.11	−0.05	−0.09	0.27	0.20	0.11
	Mean	0.42	0.37	0.44	0.41	0.51	0.46

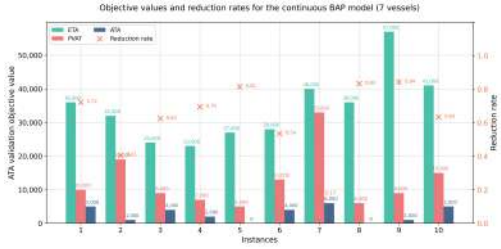
Table 5.3 presents a comparative analysis of the reduction rates for different vessel numbers and key performance metrics in the discrete BAP scenarios. Across the tested scenarios of 8, 10, 12, 14, 16, and 18 vessels, the maximum, minimum, and mean reduction

rates consistently show that the VAT-based predictions generally offer substantial improvements over schedules derived purely from vessel-reported ETAs. In particular, the mean reduction rates for all three metrics commonly range between 20% and 80%, with certain instances, such as the objective value at 14 vessels, even reaching 100%. However, negative reduction rates appear in a few cases (highlighted in red), most notably the average turnaround time for 12 vessels (-0.02) and the total waiting time for 10 and 12 vessels (-0.05 and -0.09, respectively), indicating that in these specific instances, the ETA-based approach slightly outperforms the VAT-based one. Such divergences can arise when a stronger emphasis on minimizing delay costs within the BAP optimization lead to marginal trade-offs in other metrics.

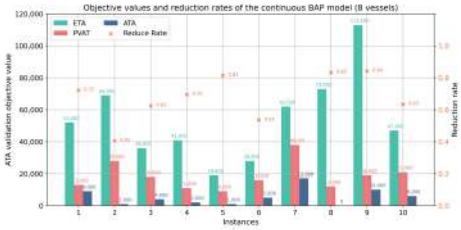
5.2.3 Continuous BAP results analysis

For the continuous BAP, the primary concern in the objective value in Function 5.14 lies in the total cost associated with deviations in vessel berthing positions and delays in vessel departure times. In our analysis, we first evaluate the objective value and its corresponding reduction rate. Similar to the discrete BAP, we then assess the total waiting time of the vessels. Lastly, given the nature of the continuous BAP, we introduce a metric called berth space utilization, which is defined as the sum of the product of each vessel's length and the time it occupies the berth, divided by the product of the total berth length LB and the time span from when the first vessel begins its berthing operation to when the last vessel finishes. The im-

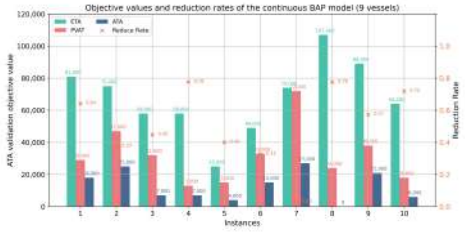
provement rate for berth space utilization predicted by VAT is then assessed using the fraction $(PVAT - ETA)/ETA$. To ensure robustness and reduce randomness, scenarios involving 7, 8, 9, 10, 11, and 12 vessels arriving at the berth are tested, and each scenario is repeated ten times. Figure 5.5 presents the objective values and



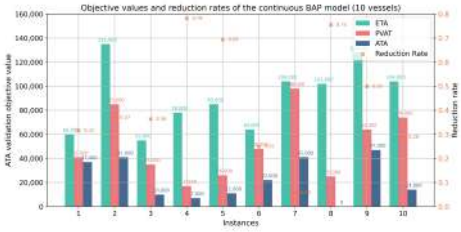
(a) Objective function value and reduction rate for 7 vessels



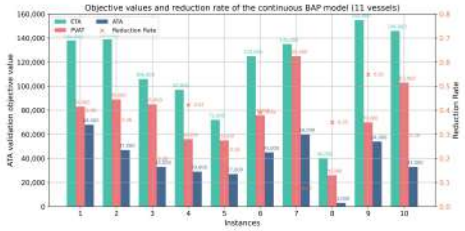
(b) Objective function value and reduction rate for 8 vessels



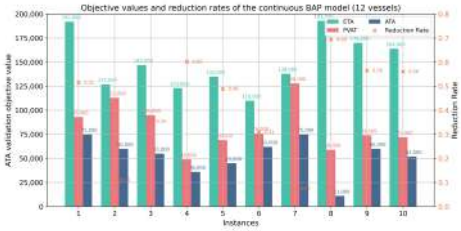
(c) Objective function value and reduction rate for 9 vessels



(d) Objective function value and reduction rate for 10 vessels



(e) Objective function value and reduction rate for 11 vessels

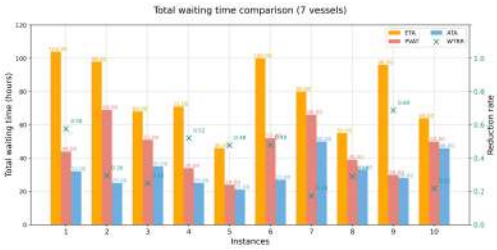


(f) Objective function value and reduction rate for 12 vessels

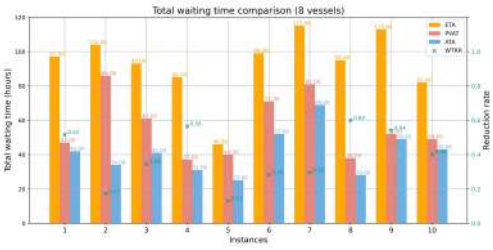
Figure 5.5: Comparison of objective function values for the continuous BAP model

reduction rates for the continuous BAP under scenarios with seven to twelve vessels, illustrating a consistent pattern in which "PVAT"

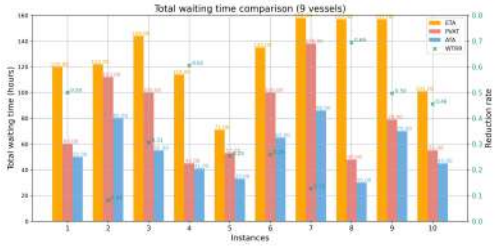
substantially reduces the objective value cost compared to "ETA". For instance, in Figure 5.5a (7 vessels case), the objective values of "ETA" cases range from about 23,000 to 57,000, whereas "PVAT" cases stay between 5,000 to 33,000, achieving reductions of roughly 17% to 84%. As the number of vessels grows from seven to twelve, the objective value for "ETA" may climb above 100,000 (and even exceed 150,000), yet "PVAT" value continues to cut costs significantly, frequently to around half of the value in "ETA" and registers most reductions between 30% and 70%, peaking at 75% or more in certain cases. In general, these findings demonstrate that VAT-based predictions provide cost savings in every case considered, clearly highlighting their efficacy in reducing objective values for continuous BAP. Figure 5.6 illustrates the total waiting time for continuous BAP in scenarios with seven to twelve vessels. In most cases, the solution of "PVAT" significantly reduces the total vessel waiting time compared to "ETA". For example, with 7 vessels, "ETA" sometimes pushes the total waiting beyond 100 hours, while "PVAT" typically keeps it close to 40-60 hours, resulting in a considerable reduction in queueing delays. Similar trends appear when the problem scale increases to 8 or 9 vessels: vessel total waiting times in "ETA" case may exceed 90–120 hours, while "PVAT" generally cuts tens of hours, lowering overall waiting closer to the 40–80 hours range. Nonetheless, a few special instances emerge, particularly with 11 vessels, where instance 7 shows no improvement (PVAT waiting time remains unchanged from ETA) and instance 8 even exhibits higher waiting under "PVAT" case. Similar



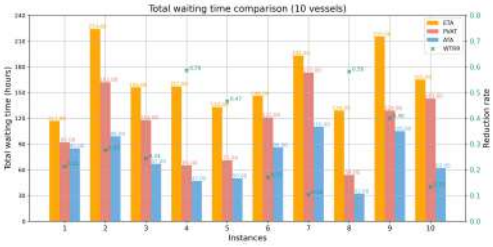
(a) Total waiting time and reduction rate for 7 vessels



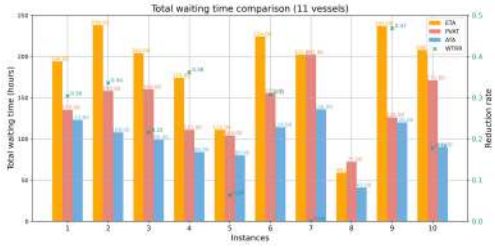
(b) Total waiting time and reduction rate for 8 vessels



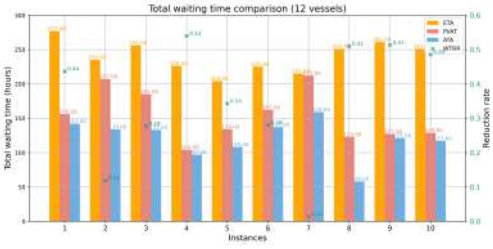
(c) Total waiting time and reduction rate for 9 vessels



(d) Total waiting time and reduction rate for 10 vessels



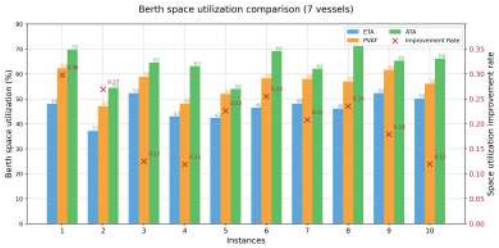
(e) Total waiting time and reduction rate for 11 vessels



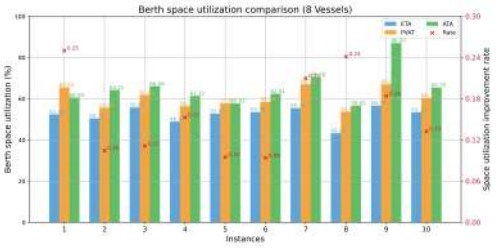
(f) Total waiting time and reduction rate for 12 vessels

Figure 5.6: Comparison of vessel total waiting time for the continuous BAP model

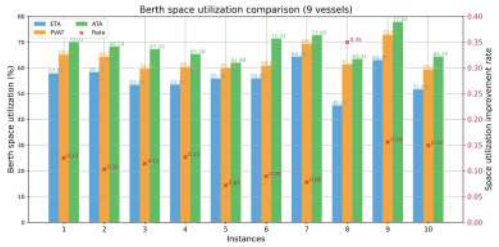
to the discrete BAP, the underlying scheduling may assign specific segments of the continuous berth to larger vessels first, compelling smaller vessels that arrive earlier to remain at anchor longer. While this increases overall waiting time, it can ultimately reduce the total cost objective by prioritizing the berthing positions and time for vessels whose deviations and delays incur higher penalties. Consequently, although "PVAT" generally decreases vessel waiting times, these cases demonstrate a trade off between minimizing queuing delays and achieving broader cost savings related to berthing position deviations and departure delays in the continuous BAP. Figure 5.7 displays berth space utilization rates under different VAT values and vessels. In general, "PVAT" yields higher berth space utilization compared to "ETA", and "ATA" remains the optimal cases among the three scenarios. With seven vessels, for example, "ETA" cases usage might hover around 48–60%, while "PVAT" raises it by 10–15 percentage points in several instances, giving improvement rates of about 0.12 to 0.27; similar patterns are visible in the 8 and 9 vessel cases, where utilization commonly shifts from the mid-50% range under "ETA" to 60% range under "PVAT", leading to improvement rates near 0.10–0.25. A notable exception appears in the 10 vessel cases: while the improvement rates vary mainly between 0.07 and 0.16, instance 6 shows a rate of zero because the last vessel's departure time remains unchanged between "ETA" and "PVAT". In continuous BAP, the operational time of each vessel in the berth is fixed; if the final vessel completes its service at the same time in both schedules, the overall utilization ratio of the berth remains constant, thus



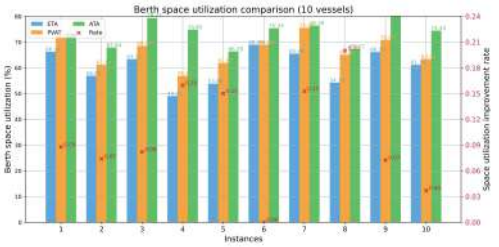
(a) Berth space utilization rate for 7 vessels



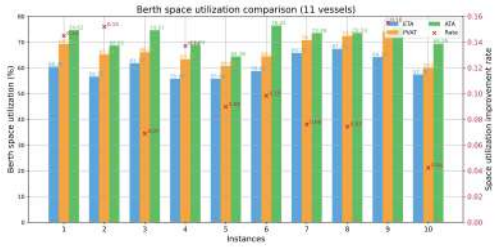
(b) Berth space utilization rate for 8 vessels



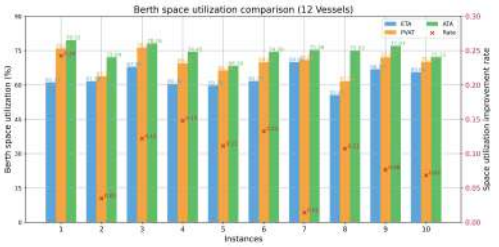
(c) Berth space utilization rate for 9 vessels



(d) Berth space utilization rate for 10 vessels



(e) Berth space utilization rate for 11 vessels



(f) Berth space utilization rate for 12 vessels

Figure 5.7: Comparison of berth space utilization rate for the continuous BAP model

producing no improvement. However, in most cases, across seven to twelve vessels, PVAT consistently improves the utilization of berth space over ETA, demonstrating the added value of VAT predictions in maximizing the efficiency of use of continuous berth.

Table 5.4: Comparative analysis of change rates for different vessels and metrics in the continuous BAP

Metric	**Stat**	**Vessel number**					
		7	**8**	**9**	**10**	**11**	**12**
Objective value	Max	0.84	0.84	0.78	0.78	0.55	0.69
	Min	0.17	0.41	0.03	0.06	0.07	0.07
	Mean	0.63	0.68	0.50	0.44	0.33	0.43
Total waiting time	Max	0.69	0.60	0.69	0.59	0.47	0.54
	Min	0.18	0.13	0.08	0.10	-0.22	0.01
	Mean	0.40	0.38	0.38	0.32	0.20	0.35
Berth space utilization	Max	0.24	0.25	0.35	0.20	0.16	0.24
	Min	0.01	0.09	0.07	0.00	0.04	0.01
	Mean	0.11	0.16	0.14	0.10	0.10	0.11

Table 5.4 summarizes the change rates in three key metrics (objective value, total waiting time, and berth space utilization) when comparing the predicted VAT approach (PVAT) with the baseline ETA under various vessel counts. A positive figure in the objective value column indicates the percentage reduction (i.e., improvement) achieved by the predicted VAT relative to the reported ETA. The maximum reduction can reach 0.84 (84 percent) for 7 and 8 vessels, and remains robust at higher fleet sizes, as suggested by mean values that remain above 0.40. The utilization of the berth space exhibits positive improvement rates, with maxima ranging from about 0.20 to 0.35, showing that "PVAT" generally increases the efficiency

of berth use compared to "ETA".

Two special cases appear. With 11 vessels, the minimum total waiting time improvement is -0.22, indicating that "PVAT" can lead to higher waiting times in extreme situations. This aligns with earlier observations that prioritizing certain berthing sequences may reduce overall costs but prolong the waiting of smaller vessels. For 10 vessels, the minimum improvement in berth space utilization is 0.00, which means the final vessel's departure time was the same under both PVAT and ETA, so the overall utilization rate did not change.

Overall, considering both discrete and continuous BAPs, the results demonstrate that predicted VAT (PVAT) consistently outperforms schedules based on vessel-reported ETA. In particular, PVAT significantly lowers the objective value across a wide range of instances. By strategically anticipating vessel arrivals, the proposed predictive approach effectively allocates berth resources and mitigates costly idling and misalignment. Consequently, the findings highlight that adopting VAT-based predictions can bring substantial benefits to berth operators, including reduced overall costs and more efficient berth utilization in both discrete and continuous berth allocation settings.

Chapter 6

Conclusion and Further research

6.1 Contribution and conclusions

In this thesis, we explore the application of ML in port operations through a structured framework consisting of evaluation, prediction, and optimization. By first assessing the current state of vessel arrivals and port operations and identifying key performance indicators, we establish a solid foundation for developing the following predictive models. These models enabled accurate forecasting of critical variables, such as ocean-going and inland waterway VAT, VTT, and VST. The insights gained from VAT predictions are then leveraged to optimize the following BAP, ensuring a more efficient assignment of berths, reducing vessel waiting time, and improving overall port throughput.

Specifically, Chapter 3 presents two novel approaches to data fusion, with the first integrating vessel-reported ETA, vessel ATA, and corresponding AIS data to quantify vessel arrival delays. The analysis demonstrates that as the vessels approach their destination port,

their reported ETA becomes increasingly accurate in both spatial and temporal dimensions, both for the oceangoing vessels and the inland waterways. The second approach combines vessel-reported EDT, vessel ADT with the timestamps of vessel entry and departure from the berth to quantify VTT and VST. In addition, the chapter conducts a quantitative analysis of COVID-19's impact on port operations, using HKP as a case study. The findings confirm that COVID-19 and its associated restrictions have led to fewer vessel calls in various vessel types, resulting in a decrease in overall port throughput at HKP. Furthermore, pandemic-related restrictions have exacerbated vessel arrival delays and prolonged turnaround time, further diminishing port operational efficiency. The study also discloses that throughout the COVID period, a positive correlation is observed between vessel arrival delay, container turnaround time at HKP, and the global freight rate.

Chapter 4 focuses on time prediction for VAT, VTT, and VST. Based on the evaluation framework, relevant datasets are constructed to facilitate time prediction, utilizing tree-based models. For oceangoing vessels, this study marks the first to simultaneously incorporate vessel-reported ETA and AIS data for VAT prediction. To better reflect real-world port operations, the dataset is partitioned chronologically. Compared to vessel-reported ETA, our approach reduces MAE from 6.84 hours to 3.11 hours, achieving a 54.53% reduction. In inland waterway shipping, we integrate the unique characteristics of river transport by incorporating vessel traffic flow data and utilizing the A-Star algorithm to estimate the remaining sailing dis-

tance for VAT prediction. Our results demonstrate a significant improvement, reducing MAE from 17.06 hours to 3.49 hours, a 79.54% reduction. For VTT prediction, our model improves accuracy by reducing MAE from 5.12 hours to 3.94 hours, a 23.05% decrease compared to vessel-reported values. Similarly, for VST prediction, the MAE is reduced from 4.54 hours to 3.19 hours, achieving a 29.74% reduction.

Chapter 5 explores the benefits of incorporating quantitative VAT predictions into the following BAP operations. Based on the predicted VAT obtained from the previous chapter, we propose a two-stage prediction-then-optimization framework. In the first stage, a VAT prediction model is developed to enhance the accuracy of vessel arrival estimates. In the second stage, the predicted VAT is integrated into the BAP model to optimize berth scheduling decisions. To evaluate the effectiveness of VAT-based scheduling, we compare the performance of a BAP model utilizing predicted VAT against one that relies on vessel-reported ETA in both discrete and continuous berth settings. The results demonstrate that, in the discrete berth scenario with 12 vessel arrivals, VAT-based BAP scheduling reduces additional BAP costs by 64%, decreases average VTT by 21%, and lowers total vessel waiting time by 73% compared to ETA-based scheduling. In the continuous berth setting, VAT-based scheduling achieves a 43% reduction in additional BAP costs, a 35% decrease in total vessel waiting time, and an 11% improvement in berth space utilization. These findings highlight that VAT-based scheduling significantly enhances berth allocation efficiency, reduces vessel wait-

ing time, and optimizes overall resource utilization.

These studies highlight the potential of improving port operations through a structured framework for evaluation, prediction, and optimization driven by ML. By systematically incorporating data-driven insights into decision-making, our study demonstrates the substantial potential of AI-powered port management in optimizing daily port operations. Real-world case studies have demonstrated that leveraging ML enhances efficiency, mitigates operational uncertainties, and optimizes resource allocation. As ports face growing demands and increasing complexities, adopting intelligent, data-driven methodologies will be essential for ensuring sustainable and adaptive operational improvements.

6.2 Future Research Directions

Further research on machine learning applications in port operations can be systematically extended across three key dimensions: evaluation, prediction, and optimization. In the evaluation component, the current study primarily relies on correlation-based feature importance analysis for VAT, VTT, and VST prediction. However, correlation does not imply causation and may obscure the true influence of input variables. Future research could incorporate causal inference methods, such as instrumental variable approaches or frameworks like Do-Why, to more accurately identify the causal impact of individual features on VAT predictions. Compared to traditional correlation analysis, causal inference allows for a more robust understanding of feature relationships under counterfactual scenarios,

which is particularly valuable when informing downstream operational decisions. By disentangling direct and indirect effects, these methods can enhance the interpretability, stability, and transferability of predictive models, especially when applied across different ports or evolving operational environments.

Currently, VAT prediction models leverage existing forecasting techniques, but future research can explore the development of specialized models tailored specifically to VAT prediction. By utilizing historical vessel arrival records, port congestion patterns, and maritime traffic data, a domain-specific VAT forecasting framework can be established to improve accuracy and adaptability.

In the prediction component, one potential enhancement lies in the integration of unstructured and external data sources to improve model adaptability and robustness. For instance, real-time meteorological data, such as wind speed, temperature, and other atmospheric conditions, can be fused with AIS records based on spatiotemporal alignment using latitude, longitude, and timestamp information. This enables the model to better account for environmental factors that influence vessel arrival times. In addition, incorporating port-related news and social media content may offer early signals of unexpected disruptions, such as severe weather events or port congestion. Furthermore, considering dynamic traffic flow conditions, such as the density of nearby vessels or trajectory-level movement patterns, can help capture congestion-induced delays. By combining these heterogeneous data sources, future VAT prediction models can become more resilient to real-world variability and better suited for

real-time operational deployment.

Another promising direction is to incorporate port resilience into the model, which captures a port's ability to respond to and recover from disruptions. Factors such as recovery time (how quickly the port resumes normal operations after a disturbance) and adaptive capacity (how effectively it manages variable throughput under stress) can be quantified through a Port Resilience Index (PRI). This index can be derived from historical records of operational disruptions, delays, and recovery patterns—for example, by measuring the slope of throughput restoration curves or the variance in berth productivity post-disruption. Once constructed, the PRI can be temporally fused into the dataset by aligning its values with AIS and port call records based on timestamp. As a time-series feature, the PRI allows the model to contextualize current vessel movements within the dynamic operational status of the port, improving robustness under uncertainty and better reflecting real-time resilience conditions.

To further improve the generalizability and practical value of the proposed framework, future studies may consider extending the analysis from a two-month evaluation period to a longer time span of one to two years. A more extensive dataset would enable deeper insights into seasonal variations, long-term behavioral trends, and the impact of recurring or rare disruptions on vessel arrivals. Moreover, validating the framework across multiple ports with varying geographic, infrastructural, and operational characteristics would offer a comprehensive assessment of its adaptability and robustness in heterogeneous port environments. This cross-port evaluation could help

demonstrate the model's transferability and support its application in diverse real-world contexts.

Beyond improving prediction accuracy, future research should also focus on seamless VAT integration into port operations. By leveraging high-precision VAT predictions, ports can implement just-in-time (JIT) docking strategies that synchronize vessel arrivals with berth availability. This reduces unnecessary anchorage time, minimizes fuel consumption, and cuts down on associated emissions. Additionally, integrating VAT forecasts into onboard navigation systems can enable vessels to dynamically adjust speed, optimize routes, and avoid inefficient holding patterns—leading to lower fuel costs and a reduced environmental footprint.

Building upon VAT prediction, the next stage is to enhance BAP by integrating predictive data into scheduling models. The current study employs a basic BAP model where each vessel is assigned a fixed operational time at the berth. Future research could expand this by incorporating quay crane allocation, where the number of cranes per vessel varies dynamically, directly impacting handling time. By integrating VAT predictions into a joint berth and quay crane scheduling framework, ports can transition toward a data-driven optimization approach that accounts for operational uncertainties.

Another area of improvement is the berth scheduling horizon. The present model limits scheduling to vessels arriving within the next 36 hours, reflecting the typical reporting window of ETA data. A more dynamic approach would involve extending this horizon, implementing a rolling scheduling procedure every 24 hours, and lever-

aging VAT predictions to create an adaptive berth allocation strategy covering an extended period, such as one week. This would provide greater flexibility in managing berth availability while responding to fluctuations in vessel arrivals. Building on this idea, further research could transform the current framework from a one-time static scheduling model into a dynamic BAP that allows for real-time adjustments as new information becomes available. By continuously updating arrival predictions and system states, the berth plan can be revised in an online manner to reflect the evolving operational environment. To efficiently solve this more complex and computationally intensive problem, deep reinforcement learning (DRL) techniques could be employed to learn effective scheduling policies that balance short-term operational efficiency with long-term optimization objectives. This would enable ports to proactively adapt to real-time uncertainties and enhance overall berth utilization.

Lastly, the prediction-then-optimization framework used in this study currently treats VAT forecasting and BAP scheduling as two independent stages, where the predicted VAT is directly input into the BAP model without feedback loops. However, the most accurate VAT predictions do not necessarily result in the most efficient berth schedules. A promising direction for future research is the development of an end-to-end contextual learning framework, where the VAT prediction model incorporates BAP optimization errors into its loss function. By integrating predictive and prescriptive analytics, this approach would ensure that VAT predictions are optimized not just for accuracy but also for their impact on berth

scheduling efficiency. This shift from a traditional predictive model to a data-driven prescriptive framework would significantly enhance decision-making in berth allocation, leading to improved operational performance and resource utilization.

References

Abreu, L. R., I. S. Maciel, J. S. Alves, L. C. Braga, and H. L. Pontes (2023). "A decision tree model for the prediction of the stay time of ships in Brazilian ports". In: *Engineering Applications of Artificial Intelligence* 117, p. 105634.

Akyurek, E. and P. Bolat (2020). "Port state control at European Union under pandemic outbreak". In: *European Transport Research Review* 12.1, pp. 1–13.

Alessandrini, A., F. Mazzarella, and M. Vespe (2018). "Estimated time of arrival using historical vessel tracking data". In: *IEEE Transactions on Intelligent Transportation Systems* 20.1, pp. 7–15.

Authority, P. P. (2023). *Definition of Terms*. URL: `https://www.ppa.com.ph/sites/default/files/Definition%20of%20Terms.pdf` (visited on 03/19/2023).

Bai, X., H. Jia, and M. Xu (2022). "Identifying port congestion and evaluating its impact on maritime logistics". In: *Maritime Policy & Management*, pp. 1–18.

Bai, X., M. Xu, T. Han, and D. Yang (2022). "Quantifying the impact of pandemic lockdown policies on global port calls". In: *Transportation Research Part A: Policy and Practice* 164, pp. 224–241.

Barua, L., B. Zou, and Y. Zhou (2020). "Machine learning for international freight transportation management: a comprehensive review". In: *Research in Transportation Business & Management* 34, p. 100453.

Benesty, J., J. Chen, Y. Huang, and I. Cohen (2009). "Pearson Correlation Coefficient". In: *Noise Reduction in Speech Processing*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–4. ISBN: 978-3-642-00296-0. DOI: `10.1007/978-3-642-00296-0_5`. URL: `https://doi.org/10.1007/978-3-642-00296-0_5`.

Branch, A. (2012). *Elements of port operation and management*. Springer Science & Business Media.

Breiman, L. (2001). "Random forests". In: *Machine Learning* 45, pp. 5–32.

Brouer, B. D., C. V. Karsten, and D. Pisinger (2016). "Big Data Optimization in Maritime Logistics". In: *Big Data Optimization: Recent Developments and Challenges*. Ed. by A. Emrouznejad. Cham: Springer International Publishing, pp. 319–344.

Bureau, H. K. D. (2022). *Latest Situation of Coronavirus Disease (COVID-19) in Hong Kong*. URL: `https://chp-dashboard.geodata.gov.hk/covid-19/en.html` (visited on 12/31/2022).

Chen, Q., Y.-E. Ge, Y.-y. Lau, M. A. Dulebenets, X. Sun, T. Kawasaki, A. Mellalou, and X. Tao (2022). "Effects of COVID-19 on passenger shipping activities and emissions: Empirical analysis of passenger ships in Danish waters". In: *Maritime Policy & Management*, pp. 1–21.

Chen, T., T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, K. Chen, R. Mitchell, I. Cano, T. Zhou, et al. (2015). "XGBoost: extreme gradient boosting". In: *R package version 0.4-2* 1.4, pp. 1–4.

Cullinane, K. and H. Haralambides (2021). *Global trends in maritime and port economics: the COVID-19 pandemic and beyond*.

Du, Y., Q. Chen, J. S. L. Lam, Y. Xu, and J. X. Cao (2015). "Modeling the impacts of tides and the virtual arrival policy in berth allocation". In: *Transportation Science* 49.4, pp. 939–956.

Ducruet, C. and O. Merk (2013). "Examining container vessel turnaround times across the world". In: *Port Technology International* 59.

El Mekkaoui, S., L. Benabbou, and A. Berrado (2023). "Deep learning models for vessel's ETA prediction: bulk ports perspective". In: *Flexible Services and Manufacturing Journal* 35.1, pp. 5–28.

Filom, S., A. M. Amiri, and S. Razavi (2022). "Applications of machine learning methods in port operations–A systematic literature review". In: *Transportation Research Part E: Logistics and Transportation Review* 161, p. 102722.

Golias, M. M., G. K. Saharidis, M. Boile, S. Theofanis, and M. G. Ierapetritou (2009). "The berth allocation problem: Optimizing vessel arrival time". In: *Maritime Economics & Logistics* 11, pp. 358–377.

Grinsztajn, L., E. Oyallon, and G. Varoquaux (2022). "Why do tree-based models still outperform deep learning on typical tabular data?" In: *Thirty-sixth Conference on Neural Information Processing Systems Datasets and Benchmarks Track*. New Orleans.

Grinter, M. (2023a). *2022 container throughput at Port of Hong Kong down 6.9%*. URL: http://www.hongkongmaritimehub.com/2022-container-throughput-at-port-of-hong-kong-down-6-9/ (visited on 02/24/2023).

Grinter, M. (2023b). *China's coronavirus lockdowns jam ports hong kong and shenzhen*. URL: https://www.scmp.com/economy/china-economy/article/3171870/chinas-coronavirus-lockdowns-jam-ports-hong-kong-and-shenzhen (visited on 02/24/2023).

Grinter, M. (2023c). *Hong Kong Port imposes further measures to restrict COVID-19 cluster*. URL: http://www.hongkongmaritimehub.com/hong-kong-port-imposes-further-measures-to-restrict-covid-19-cluster/ (visited on 02/24/2023).

Grinter, M. (2023d). *Hong Kong Port shrugs off COVID-19 effect in March*. URL: http://www.hongkongmaritimehub.com/hong-kong-port-shrugs-off-covid-19-effect-in-march/ (visited on 02/24/2023).

Grinter, M. (2023e). *Hong Kong praised for prompt action in facilitating cargo flow*. URL: http://www.hongkongmaritimehub.com/hong-kong-praised-for-prompt-action-in-facilitating-cargo-flow/ (visited on 02/24/2023).

Grinter, M. (2023f). *Port of Hong Kong feels the pain of diminishing trade*. URL: http://www.hongkongmaritimehub.com/port-of-hong-kong-feels-the-pain-of-diminishing-trade/ (visited on 02/24/2023).

Grinter, M. (2023g). *Relaxed transit arrangements for seafarers through HK extended to passenger ship crews*. URL: http://www.hongkongmaritimehub.com/relaxed-transit-arrangements-for-seafarers-through-hk-extended-to-passenger-ship-crews/ (visited on 02/24/2023).

Guo, L., J. Wang, and J. Zheng (2021). "Berth allocation problem with uncertain vessel handling times considering weather conditions". In: *Computers & Industrial Engineering* 158, p. 107417.

Hansen, P., C. Oğuz, and N. Mladenović (2008). "Variable neighborhood search for minimum cost berth allocation". In: *European journal of operational research* 191.3, pp. 636–649.

Havenbedrijf Rotterdam B.V. (2023). *Highlights Annual Report 2023*. URL: https://reporting.portofrotterdam.com/FbContent.ashx/pub_1018/downloads/v240305112534/PoR_AR_2023_Annual_Report_Highlights.pdf (visited on 01/07/2025).

HKSAR (2023). *Enhancing Hong Kong's position as a maritime centre*. URL: https://www.legco.gov.hk/research-publications/english/2022rt03-enhancing-hong-kongs-position-as-a-maritime-centre-20220519-e.pdf (visited on 03/19/2023).

Hong Kong Government (2022). *Vessel Traffic Management System Report*. URL: https://data.gov.hk/en-data/dataset/hk-md-mardep-vessel-traffic-management-system-report. (visited on 05/23/2022).

Hong Kong Government (2023). *The Complete Berthing Guidelines for Port of Hong Kong*. https://www.mardep.gov.hk/en/pub_services/pdf/berthguide.pdf. Accessed: 2023-11-29.

Hong Kong Maritime And Port Board (2024). *Port of Hong Kong introduction*. https://www.hkmpb.gov.hk/en/port.html. Accessed: 2024-05-09.

Humphreys, R. M., A. Dumitrescu, N. O. Biju, and Y. Y. Lam (2020). *COVID-19 and the maritime and logistics sector in Africa*. URL: https://openknowledge.worldbank.org/server/api/core/bitstreams/ec0cf9ad-efa5-5a4e-963d-d88a7341554d/content (visited on 04/01/2023).

Imai, A., E. Nishimura, and S. Papadimitriou (2001). "The dynamic berth allocation problem for a container port". In: *Transportation Research Part B: Methodological* 35.4, pp. 401–417.

Jasmina, Ovcina Mandra (2024). *Global liner performance hits lowest point since September 2022*. https://www.worldcargonews.com/news/2024/03/global-liner-performance-hits-lowest-point-since-september-2022/?gdpr=accept. Accessed: 2024-05-05.

Jin, L., J. Chen, Z. Chen, X. Sun, and B. Yu (2022). "Impact of COVID-19 on China's international liner shipping network based on AIS data". In: *Transport Policy* 121, pp. 90–99.

Ke, G., Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu (2017). "LightGBM: A Highly Efficient Gradient Boosting Decision Tree". In: *Advances in Neural Information Processing Systems*. Vol. 30. Curran Associates, Inc.

Kim, K. H. and K. C. Moon (2003). "Berth scheduling by simulated annealing". In: *Transportation Research Part B: Methodological* 37.6, pp. 541–560.

Kolley, L., N. Rückert, M. Kastner, C. Jahn, and K. Fischer (2023). "Robust berth scheduling using machine learning for vessel arrival time prediction". In: *Flexible services and manufacturing journal* 35.1, pp. 29–69.

Lewis, R. J. (2000). "An introduction to classification and regression tree (CART) analysis". In: *Annual meeting of the society for academic emergency medicine in San Francisco, California*. Vol. 14. Citeseer.

Li, B. and Y. He (2020). "Container terminal liner berthing time prediction with computational logistics and deep learning". In: *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, pp. 2417–2424.

Li, H., H. Jiao, and Z. Yang (2023). "AIS data-driven ship trajectory prediction modelling and analysis based on machine learning and deep learning methods". In: *Transportation Research Part E: Logistics and Transportation Review* 175, p. 103152. DOI: 10.1016/j.tre.2023.103152.

List, L. (2023a). *Hong Kong port operators close in on approval for joint venture*. URL: https://lloydslist.maritimeintelligence.informa.com/LL1133486/Hong-Kong-port-operators-close-in-on-approval-for-joint-venture (visited on 02/24/2023).

List, L. (2023b). *Weekly briefing: Steep fall in tanker tonne-mile demand; Maersk reinstates full-year guidance*. URL: https://lloydslist.maritimeintelligence.informa.com/LL1133580/Weekly-briefing-Steep-fall-in-tanker-tonnemile-demand-Maersk-reinstates-fullyear-guidance (visited on 02/24/2023).

Millefiori, L. M., P. Braca, D. Zissis, G. Spiliopoulos, S. Marano, P. K. Willett, and S. Carniel (2021). "COVID-19 impact on global maritime mobility". In: *Scientific Reports* 11.1, p. 18039.

Mokhtar, K. and M. Z. Shah (2006). "A regression model for vessel turnaround time". In: *Tokyo Academic, Industry & Cultural Integration Tour*, pp. 1–15.

Narasimha, P. T., P. R. Jena, and R. Majhi (2021). "Impact of COVID-19 on the Indian seaport transportation and maritime supply chain". In: *Transport Policy* 110, pp. 191–203.

Notteboom, T., T. Pallis, and J.-P. Rodrigue (2021). "Disruptions and resilience in global container shipping and ports: the COVID-19 pandemic versus the 2008–2009 financial crisis". In: *Maritime Economics & Logistics* 23.2, pp. 179–210.

Pani, C., T. Vanelslander, G. Fancello, M. Cannas, et al. (2015). "Prediction of late/early arrivals in container terminals-a qualitative approach". In: *European Journal of Transport and Infrastructure Research* 15.4, pp. 536–550.

Park, K., S. Sim, and H. Bae (2021). "Vessel estimated time of arrival prediction system based on a path-finding algorithm". In: *Maritime Transport Research* 2, p. 100012.

Park, Y.-M. and K. H. Kim (2003). "A scheduling method for berth and quay cranes". In: *OR spectrum* 25.1, pp. 1–23.

Placek, M. (2022). *Global container freight rate index from January 2019 to December 2022*. URL: `https://www.statista.com/statistics/1250636/global-container-freight-index/` (visited on 12/31/2022).

Post, S. C. M. (2022). *China coronavirus: Hong Kong scrambles to roll out containment plan stopping short of total closure, with cuts on cross-border travel and reduced transport services with mainland*. URL: `https://www.scmp.com/news/hong-kong/health-environment/article/3047907/china-coronavirus-hong-kong-government-deny-entry` (visited on 12/31/2022).

Rodrigues, F. and A. Agra (2022). "Berth allocation and quay crane assignment/scheduling problem under uncertainty: A survey". In: *European Journal of Operational Research* 303.2, pp. 501–524.

Rødseth, K. L., P. B. Wangsness, and H. Schøyen (2018). "How do economies of density in container handling operations affect ships' time and emissions in port? Evidence from Norwegian container terminals". In: *Transportation Research Part D: Transport and Environment* 59, pp. 385–399.

Saeed, N., S. Nguyen, K. Cullinane, V. Gekara, and P. Chhetri (2023). "Forecasting container freight rates using the Prophet forecasting method". In: *Transport Policy* 133, pp. 86–107.

Scott, M., L. Su-In, et al. (2017). "A unified approach to interpreting model predictions". In: *Advances in neural information processing systems* 30, pp. 4765–4774.

Slack, B. and E. Gouvernal (2011). "Container freight rates and the role of surcharges". In: *Journal of transport geography* 19.6, pp. 1482–1489.

Smith, D. (2021). "Big data insights into container vessel dwell times". In: *Transportation Research Record* 2675.10, pp. 1222–1235.

Statista (2023a). *Sales of liquefied petroleum gas (LPG) in Hong Kong in 2021, by type of users*. URL: `https://www.statista.com/statistics/1085341/hong-kong-sales-of-lpg-by-type-of-users/` (visited on 02/24/2023).

Statista (2023b). *Sales of oil products for local consumption in Hong Kong in 2021, by oil product*. URL: `https://www.statista.com/statistics/1085377/hong-kong-sales-of-oil-products-for-local-consumption-by-product/` (visited on 02/24/2023).

Štepec, D., T. Martinčič, F. Klein, D. Vladušič, and J. P. Costa (2020). "Machine learning based system for vessel turnaround time prediction". In: *2020 21st IEEE International Conference on Mobile Data Management (MDM)*. IEEE, pp. 258–263.

Tai, Z., J. Guo, Y. Guan, and Q. Shi (2021). "Impact of COVID-19 on port production and operation based on system dynamics: A case study of Shanghai port in China". In: *Journal of Advanced Transportation* 2021, pp. 1–13.

Tchang, G. S. (2020). "The impact of ship size on ports' nautical costs". In: *Maritime Policy & Management* 47.1, pp. 27–42.

Tsui, K. W. H., X. Fu, T. Chen, Z. Lei, and H. Wu (2021). "Analyzing Hong Kong's inbound tourism: The impact of the COVID-19 pandemic". In: *IATSS Research* 45.4, pp. 440–450.

UNCTAD (2025). *The Review of Maritime Transport 2024*. `https://unctad.org/system/files/official-document/wir2024_en.pdf`. Accessed: 2025-03-13.

Veenstra, A. and R. Harmelink (2021). "On the quality of ship arrival predictions". In: *Maritime Economics & Logistics*, pp. 1–19.

Wang, X., Z. Liu, R. Yan, H. Wang, and M. Zhang (2022). "Quantitative analysis of the impact of COVID-19 on ship visiting behaviors to ports-A framework and a case study". In: *Ocean & coastal management* 230, p. 106377.

Wenzel, P., R. Jovanovic, and F. Schulte (2023). "A neural network approach for ETA prediction in inland waterway transport". In: *International Conference on Computational Logistics*. Springer, pp. 219–232.

WRS (2023). *World Shipping Register*. `https://world-ships.com/`. Accessed: 2024-10-13.

Xu, L., S. Yang, J. Chen, and J. Shi (2021). "The effect of COVID-19 pandemic on port performance: Evidence from China". In: *Ocean & Coastal Management* 209, p. 105660.

Yan, R., H. Mo, X. Guo, Y. Yang, and S. Wang (2022). "Is port state control influenced by the COVID-19? Evidence from inspection data". In: *Transport Policy* 123, pp. 82–103.

Yan, R., S. Wang, L. Zhen, and G. Laporte (2021). "Emerging approaches applied to maritime transport research: Past and future". In: *Communications in Transportation Research* 1, p. 100011.

Yang, D., L. Wu, S. Wang, H. Jia, and K. X. Li (2019). "How big data enriches maritime research–a critical review of Automatic Identification System (AIS) data applications". In: *Transport Reviews* 39.6, pp. 755–773.

Yang, Y., Y. Liu, G. Li, Z. Zhang, and Y. Liu (2024). "Harnessing the power of Machine learning for AIS Data-Driven maritime Research: A comprehensive review". In: *Transportation Research Part E: Logistics and Transportation Review* 183, p. 103426.

Yang, Y., R. Yan, and S. Wang (2024). "An efficient ranking-based data-driven model for ship inspection optimization". In: *Transportation Research Part C: Emerging Technologies* 165, p. 104731.

Yoon, J.-H., D.-H. Kim, S.-W. Yun, H.-J. Kim, and S. Kim (2023). "Enhancing container vessel arrival time prediction through past voyage route modeling: A Case Study of Busan New Port". In: *Journal of Marine Science and Engineering* 11.6, p. 1234.

Yu, J., G. Tang, X. Song, X. Yu, Y. Qi, D. Li, and Y. Zhang (2018). "Ship arrival prediction and its value on daily container terminal operation". In: *Ocean Engineering* 157, pp. 73–86.

Yu, J., G. Tang, S. Voß, and X. Song (2023). "Berth allocation and quay crane assignment considering the adoption of different green technologies". In: *Transportation Research Part E: Logistics and Transportation Review* 176, p. 103185.

Zhai, D., X. Fu, X. F. Yin, H. Xu, and W. Zhang (2022). "Predicting berth stay for tanker terminals: a systematic and dynamic approach". In: *arXiv preprint arXiv:2204.04085*.

Zhang, T., J. Yin, X. Wang, and J. Min (2023). "Prediction of container port congestion status and its impact on ship's time in port based on AIS data". In: *Maritime Policy & Management*, pp. 1–29.

Zhang, X., X. Fu, Z. Xiao, H. Xu, X. Wei, J. Koh, D. Ogawa, and Z. Qin (2024). "Prediction of Vessel Arrival Time to Pilotage Area Using Multi-Data Fusion and Deep Learning". In: *arXiv preprint arXiv:2403.09969*.

Zhao, H.-M., H.-D. He, K.-F. Lu, X.-L. Han, Y. Ding, and Z.-R. Peng (2022). "Measuring the impact of an exogenous factor: An exponential smoothing model of the response of shipping to COVID-19". In: *Transport Policy* 118, pp. 91–100.

Zhen, L., L. H. Lee, and E. P. Chew (2011). "A decision model for berth allocation under uncertainty". In: *European Journal of Operational Research* 212.1, pp. 54–68.

Zhou, C., S. Zhu, M. G. Bell, L. H. Lee, and E. P. Chew (2022). "Emerging technology and management research in the container terminals: Trends and the COVID-19 pandemic impacts". In: *Ocean & Coastal Management* 230, p. 106318.

Zhou, Z.-H. (2012). *Ensemble Methods: Foundations and Algorithms*. CRC press.

Zhou, Z.-H. (2021). *Machine Learning*. Singapore: Springer Nature.