Maritime transport applications in an operational efficiency perspective

Gabriel Fuentes

Submitted in partial fulfillment of the
requirements for the degree of
Philosophiae Doctor
at the Norwegian School of Economics

2022

Acknowledgements

I would like to express my deepest appreciation to my advisor Roar Ådland, for his support, and steering me into the path to become a successful researcher in a future. His advice was always taken with respect and admiration.

Also, I appreciate the attention of Stein Wallace and his help with modeling stochastic problems and his valuable input in the second chapter of this thesis. My sincere appreciation goes to François-Charles Wolff and Pierre Cariou for some very enriching discussions and for welcoming me into some joint projects.

I would also like to thank Julio Goez and Mario Guajardo, for their tutoring and for helping my family in having an easier integration to the city. I appreciate the discussions we had on distributed computing, optimization models up to where to take the kids for holidays.

My appreciation to Ishaan Hemnani from Bunker Ex, Semiramis Assimakopolou from The Signal Group, Egil Husby from Western Bulk, and Ronald Jansen from the United Nations Statistics Division for sharing their experiences and resources that partly supported this thesis.

I would like to express my sincere gratitude to my colleagues at the Business and Management Science department that made this a stimulating working environment.

Finally, I am extremely grateful to my wife Rosario for her constant support over these 13 years and her patience over my career shifts, from my years as a seafarer until my new passion as an academic and everything in between (sorrows and triumphs). This path has been accompanied with the joy of welcoming our son Lucas and our daughter Gabriela, to whom I am also grateful.

Bergen, October 11th, 2022
Gabriel Fuentes

Contents

Introduction

«The birds are still ... and the sea is calm; hushed are the winds, and silence broods o'er this narrow firth»

Agamemnon in the Iliad by Homer

With this phrase, Agamemnon referred to the problem he was facing. He had a fleet of 100 ships stranded at the port of Aulis expectant on the winds to follow such as they could invade Troy. If he only had enough information, he would probably have planned differently and choose a more favorable port for starting his invasion. Such an unfortunate port call (he was punished by the goddess Artemis), was later compensated by the sacrifice of his daughter Iphigenia to the goddess on behalf of favorable winds.

At the same time that the Iliad was supposedly written (c. 8th century BC), the Phoenicians dominated the trade by sea at the Mediterranean. The standard way of trading was in the hands of merchant skippers. These merchants, usually also the captains and owners of the vessels, were part of the adventure and would trade cargo as the vessel was visiting ports (Fayle, 2006). They would act with due diligence and plan for an efficient voyage, considering the seasonal wind patterns, the intended trade route and their commercial interests.

The rise of Carthage (a former Phoenician colony) as a maritime power gave rise to years of commercial dominance driven by their exploring skills and long preserved secrets of the discovered routes (Fayle, 2006). With the fall of the Carthaginians to the Romans as the result of the Punic wars (c. 264-146BC), such secrets switched hands and developed, together with the expansion of the empire, into new routes. The Romans would efficiently plan for their trade voyages based on the information collected by their knowledge gained and years of observations. As an example of this, they would export their grain to Alexandria in July and August and arrive in due time for sorting out the complexities of trading at a destination including, in the more

favorable cases, delays arising from corrupt officialdom or bureaucratic concerns (Arnaud, 2018).

Despite the sense of efficiency being part of maritime trade since early civilization, the motives for such efficiency have changed as new technology has emerged (e.g., steam engines, GPS, satellite communications, etc.) or as civilization faces new challenges (e.g., feeding the empire, globalization, climate change, etc.). Whether the intention was to invade a country or be on time at a port according to weather seasonality in the past, or taking fuel with the least delay and reduce vessel emissions based on efficient port calls in the present time; there is an undisputed awareness of the benefits for a maritime venture of being *operationally efficient.* It is the objective of this thesis to explore and measure the resulting benefits (commercially and/or socially) in present maritime transport applications.

Operational efficiency is defined by Lee and Johnson (2013) as the ability to provide services cost-effectively without sacrificing quality. Based on this definition, vessels' operational efficiency in this thesis refers to cost-effective vessel operation and port utilization while keeping to the agreed transport terms and schedule. Such efficient planning is closely related to the vessel time at sea and/or the time in port (Moon and Woo, 2014).

If vessel operational efficiency assumes that there is a tradeoff between time at sea and time in port, there is a direct effect of port efficiency (or, more generally, any waiting time at an intermediate stop such as a bunkering port or a canal) on the efficiency of the voyage. The efficiency of the voyage could be quantified in terms of cost minimization, reduced time at port or emissions reduction.

Port efficiency has been thoroughly researched in the academic literature (see Poulsen and Sampson (2020) and Krmac and Mansouri (2022) for a review). However, many of the studies refer to efficiency as a measure of container port throughput and number of port calls. Few studies refer to the turnaround time or waiting times at port (see Slack et al., 2018 for a review). The main reason, as suggested by Slack et al. (2018), is that time data for measuring port efficiency is seldom available to have a proper comparison. This is the case as it is difficult to collect consistent time data from individual ports and shipping lines, and the use or access to such data might further be restrained by business secrecy.

In order to plan an operationally efficient voyage, the operator needs access to statistics on key variables such as waiting times and servicing times for any potential stops during the intended voyage. Such data and statistics then feed the decision making from an operational efficiency perspective. Along this line, the chapters of this thesis are based on the same overarching concept. First, establishing a framework for generating robust port/canal statistics, including waiting times. Second, to obtain insight from the underlying dynamics of the ports and canals. Third, with the insight gained from the generated statistics, to propose solutions for vessels to be operated efficiently.

In this thesis, we study vessel operational efficiency in the context of two maritime applications. First, in **Chapter 1**[1] and **Chapter 2** the focus is on the bunkering (fuel) management problem and the effect of waiting times in the bunkering port selection. Second, **Chapter 3** investigates the potential GHG emissions impact from vessel speed reduction due to increasing operational efficiency at a maritime chokepoint such as a canal.

In **Chapter 1**, we identify the need for a framework for the extraction of empirical data and statistics on global bunkering operations. Bunkering operations statistics are important for studies on regional port competitiveness and to feed optimization models for the bunkering management problem. For this purpose, the first part of the chapter describes a framework for deriving bunkering statistics, based on Automatic Identification System (AIS) data, spatial analysis and the Density Based Spatial Clustering Application with Noise (DBSCAN) algorithm. The framework can generate bunkering statistics, such as waiting times, service times, number of transits per port and whether a vessel was served at port or at anchorage. The second part of the study uses the framework to generate statistics for bunkering operations at the various ports in the Mediterranean Sea. The results validate earlier results and hypotheses in the literature on bunkering operations, commonly obtained from qualitative methods. More importantly, it presents, for the first time, a comparison metric for the efficiency of bunkering operations in the Mediterranean Sea.

Having an efficiency metric (waiting times) for every port in the Mediterranean, **Chapter 2** addresses the question of *what is the value of efficiency in an optimal bunkering*

[1] The paper is published in *Transportation Research Part E: Logistics and Transportation Review* (2021)

management plan for tramp ships? The research question is explored by a set of experiments simulating the different settings a shipowner would face if trading in the Mediterranean Sea. The experiments are based in a multistage stochastic model of the bunkering management problem with consideration of stochastic waiting times (sourced from Chapter 1 framework), stochastic bunker prices, voyage dynamics and contractual information. The results suggest that when there is large uncertainty in waiting times, a decision solely driven by price could lead to significant losses. As a managerial insight, this study shows that shipowners should evaluate waiting times at potential bunkering ports as part of standard practice.

Finally, **Chapter 3** explores the question: *What are the implications for emissions if vessels reduce speed due to improving operational efficiency at maritime chokepoints such as a canal?* Speed reduction is an operational measure with a high emissions reduction potential (Bouman et al., 2017). A vessel can reduce GHG emissions by reducing speed and spending more time at sea instead of sailing at high speed to wait at anchorage (rush to wait). The findings show that more efficient scheduling at maritime chokepoints, such as the Panama Canal in our case, can have a significant impact on overall emissions. Once again, we develop an algorithmic framework based on AIS data that can generate robust statistics on waiting and transit times as a basis for the empirical evaluation. From a qualitative point of view, our work highlights how many of the well-known contractual barriers to operational efficiency in ports are mitigated in the case of canal operation, making this a suitable avenue for the gradual implementation of policies driven by the need for emission reduction.

References

Arnaud, P. (2018). Reconstituting the maritime routes of the Roman Empire. In C. Ducruet, *Advances in shipping data analysis and modeling.* Routledge.

Bouman, E. A., Lindstad, E., Rialland, A. I., & Strømman, A. H. (2017). State-of-the-art technologies, measures, and potential for reducing GHG emisions from shipping–A review. *Transportation Research Part D: Transport and Environment, 52*, 408-421.

Fayle, C. E. (2006). "Ships of Tarshish". The sea traders of antiquity. In C. E. Fayle, *A short history of the world's shipping industry.* Routledge.

International Maritime Organization. (2020). *Reduction of GHG emissions from ships. Fourth IMO GHG study 2020.* London: International Maritime Organization.

Krmac, E., & Mansouri Kaleibar, M. (2022). A comprehensive review of data envelopment analysis (DEA) methodology in port efficiency evaluation. *Maritime Economics & Logistics*, 1-65.

Lee, C. Y., & Johnson, A. L. (2013). Operational Efficiency. In A. B. Badiru, *Handbook of Industrial and Systems Egineering.* CRC Press.

Moon, D. S., & Woo, J. K. (2014). The impact of port operations on efficient ship operation from both economic and environmental perspectives. *Maritime Policy & Management, 41(5)*, 444-461.

Poulsen, R. T., & Sampson, H. (2020). A swift turnaround? Abating shipping greenhouse gas emissions via port call optimization. *Transportation Research Part D: Transport and Environment, 86*.

Slack, B., Comtois, C., Wiegmans, B., & Witte, P. (2018). Ships time in port. *International Journal of Shipping and Transport Logistics, 10(1)*, 45-62.

Chapter 1

Generating bunkering statistics from AIS data: A machine learning approach*

Gabriel Fuentes [a,b]

[a] Centre for Applied Research at NHH(SNF)
[b] Department of Business and Management Science, NHH

Abstract

In shipping, the optimization of the bunkering location is dependent on price, deviation from the planned route and the cost of delays incurred by the bunkering operation itself. Despite their potential importance, detailed statistics for bunkering operations at the individual port call level (e.g. waiting times, barge capacity, location - anchorage or terminal) are not available. I develop a new method, based on the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm, that can a) identify tanker vessels used as bunkering tankers, b) detect stationary ocean-going vessels at anchorage or alongside terminals and c) automatically recognize bunkering operations as a rendezvous between an ocean-going vessel and a bunkering barge. I find that the high time complexity of the DBSCAN algorithm in this setting can be compensated by adjusting the algorithm to distributed computer settings. In the empirical study, I use the output to describe the relative importance of Mediterranean ports for bunkering and provide statistics on waiting and servicing times. The empirical findings are important for the optimization of the bunkering location decision in shipping and studies on regional port competitiveness.

*This research was partly financed by the Research Council of Norway under the project “Smart digital contracts and commercial management”, project number 280684. The bunker ports position were kindly provided by Bunker-Ex.

Keywords: bunkering operations, machine learning, DBSCAN, spatial recognition, bunkering statistics.

1.1 Introduction

Bunkering information is important to ship operators, ports, and suppliers alike. Ship operators can benefit from knowledge about the servicing and waiting times for bunkering to make informed operational decisions. Examples of such decisions are: whether to deviate for bunkers or bunker during cargo operations, making a route plan, selecting the voyage speed, and deciding the quantity of bunker to load. Bunker suppliers can use the fleet age and capacity of their bunker barges, service times and fuel prices to assess where they stand in the intra-/inter-port competition. This is helpful in decisions about service pricing, fleet renewal, fleet location or fleet maintenance. Port operators can monitor and compare bunkering servicing times and asses how it influences their core business (transferring cargo/passengers) and port competitiveness and, if the service does not already exist, integrate bunker supply as part of their pool of services. Using the port of Antwerp as an example, Aronietis et al. (2017) claim that bunkering operations have no influence on port choice, as bunkering –in the case of liner operations– is optimized on a defined route. Conversely, Lam et al. (2011) argue, based on their comparison of Shanghai and Singapore ports, that the bunkering service is important to ports aiming to be recognized as maritime hubs within existing supply chains. What is clear is that a port-by-port analysis is suitable only when there is access to standardized measures of comparison.

Until now, detailed and comparable information about bunkering operations from individual ports does not exist. Key information such as the number and specifications of actively operating bunker barges, whether bunkering services are performed alongside or as a barge-to-ship operation at anchorage is not standardized and, to my knowledge does not exist as aggregate statistics. The motivation of this paper is to create a general, standardized and automatic detection algorithm that can recognize bunkering operations and generate the data that can serve as a base to rank port performance in terms of bunkering services.

The bunkering statistics presented in this paper is derived from Automatic Identification System (AIS) data from ships. AIS has expanded from its primary use as an aid to

navigation and collision avoidance[1] to several other uses including being a tool for vessel trading pattern recognition. AIS data has been used extensively as part of both commercial activities and research, with more suppliers of data entering the market and benefiting from an ever-increasing network of low-earth orbit satellites and coastal receivers. Higher observational frequency also creates an escalating challenge in terms of algorithm effectiveness for large databases, as is the case for the empirical study in this paper.

The proposed method to extract the bunkering operation information relies on generating an automatic detection algorithm powered by the unsupervised learning method Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Ester et al., 1996). This algorithm solves the problems of classifying the bunker barges, recognizing a stopped vessel at anchorage/berth and identifying a bunker operation from a barge to an oceangoing vessel at close quarters. The first step classifies a bunker barge as a tanker vessel within a particular size range that is positioned within local port waters for an extended period of time. The second step recognizes stopped ocean-going vessels by filtering positions inside port waters and clustering those within a proximity parameter. To model a vessel swinging around an anchor position, the setting uses tidal streams as the proximity parameter. The third and final step identifies a bunkering operation from periods where a bunker barge and an ocean-going vessel are in close proximity. The empirical results are validated against statistics in the bunkering literature.

The paper's contributions are fourfold. First, I develop an algorithm for the recognition of tanker vessels used as bunkering barges, with the classification carried out from the overall list of existing vessels. To be selected, vessels must fulfill the characteristics of a prospective bunker barge, such as the percentage of time spent inside port waters, and the type and size of the vessel. Second, I improve algorithms for the detection of vessels at anchor or alongside berth using an adaptation of the DBSCAN algorithm. The algorithm is capable of separating drifting vessels from anchored vessels using the detection of an anchorage shape, or stationary vessels alongside a berth. Third, I propose an algorithm for the automatic recognition of bunkering operations between a

[1] AIS is mandatory for vessels of 300 gross tonnage and upwards engaged in international voyages and cargo vessels with 500 gross tonnage and upwards not engaged in international voyage as per the International Convention for the Safety of Life at Sea 74/88 (SOLAS).

bunkering vessel and an ocean-going vessel based on the rendezvous between the defined berthed or anchored vessel and prospective bunker barges. Finally, I extract and compare the bunkering statistics across ports in the Mediterranean and Marmara Sea based on the auto-generated results from the algorithm.

The rest of the paper is organized as follows: Section 1.2 contains the literature review, Section 1.3 describes the algorithm for bunkering recognition, and Section 1.4 includes the application of the algorithm, the statistical results of real data from the Mediterranean and Marmara Sea, and provides a sensitivity analysis for parameters testing. Finally, Section 1.5 concludes and identifies prospects for future extensions.

1.2 Literature review

The literature on bunkering includes fuel management strategies, fuel services as a determinant of port competitiveness, financial instruments to hedge against price fluctuations, competition between bunkering ports, or statistical analysis of bunker price dynamics.

Acosta et al. (2011) review the port competitiveness literature and conclude that the primary factor when selecting a port is cost, with the quality and completeness of services (where bunkering is part) being secondary. When focusing on bunkering they argue that the fuel price together with geographical advantage (i.e. proximity to main trading routes) are the major determinants of selecting a port. Following in importance they also identified port tariffs and bunkering supply waiting times. Their conclusions are based on responses to a questionnaire from Gibraltar Strait port operators.

Also following a qualitative approach, Aronietis et al. (2017) interviewed shipping companies calling the port of Antwerp and having offices in Antwerp. They validated that bunkering competitiveness is mainly affected by the fuel price, with the correct delivery of fuel quantity and quality being second in importance. This shows that a care towards service and the perception of a good service is important. Lam et al. (2011) compared the port competitiveness of Shanghai and Singapore and relate this to their position as bunkering hubs. The limited research on the topic drove them to collect information through interviews with regulatory bodies, marine consultants, bunker suppliers/brokers, ship managers and ship operators. Based on their results, they

argue that bunker price ranks third in order of importance following bunker quality and market transparency, with the latter referring to the degree of corruption in the bunkering market. They assert that a bunker market is more transparent when there is strict control from authorities regarding bunkering practices. Better visibility regarding the bunkering activities, such as the micro-level statistics developed in this paper, is an important tool to enhance transparency.

The reliance on qualitative research methods in the literature on bunkering market competition is a consequence of the difficulties in collecting data for conducting analysis. Moreover, the lack of a data standard and unified collection method makes the comparison of ports nearly impossible, effectively limiting the research to small-scale case studies. Shipping as a globally integrated market could benefit from both an accumulation of local data and the analysis along several dimensions at a higher level of aggregation.

The elements assessed in port competitiveness analysis can also be considered as part of bunkering decision support tools within the area of Operations Research (OR). Besbes and Savin (2009) provide a profit optimization model that considers bunkering decisions for both the liner and tramp shipping cases. The liner case models fuel prices as a variable with stochastic dynamics, while the tramp case provides optimal solutions that choose a combination of route and bunkering stops. Both cases use scenarios across ports where the price is constant or variable. Yao et al. (2012) use real data from ships and create empirical approximations of the function that compares fuel consumption rate and ship speed. The empirical models, segmented by ship size, prove to be superior to standard non-data driven functions when used as part of the bunkering optimization model. Together with differences in results, they argue that the traditional standard consumption function - Ronen's (1982) cubic speed model - results in lower cost but is suboptimal as it is not reflecting real scenarios. Wang and Meng (2012) confirm that better estimates can be had from data-driven fuel consumption models in container shipping, but recognize that Ronen's function is a good approximation when historical data is not available. Vilhelmsen et al. (2014) show the bunker consumption and its interaction with the fuel inventory, with ship speed as a decision variable. They provide a heuristic solution to the optimal routing, scheduling and bunkering of a tramp fleet carrying full shiploads. They find that the integration of

bunkering operations in cargo scheduling proves beneficial for the overall solution. Similarly, Sheng et al. (2015) model the multi-stage stochastic version in a liner shipping case. The solution determines the ship-owner's bunkering policy based on a ship's remaining fuel level and the prevailing fuel price when at a scheduled port. They generate the fuel consumption scenarios by sampling average approximation and the fuel prices' percentage difference in a one-stage (one voyage leg) Markovian process.

The problem of bunkering optimization is still open for further improvements. No model has yet included the distributions of bunkering service and waiting times or other measures of service quality within the bunkering location selection problem. This seems relevant, as it may be optimal for a ship-owner to deviate from a port that is "cheap and close" to a more costly and distant port that has a substantially faster service. The former choice might be related to the strict fulfillment of contractual requirements, such as meeting a pick-up window (laycan) for tramp vessels. An inefficient bunkering service implies that the vessels has to speed up to reach a laycan, causing a subsequent increase in consumption and emissions. As a related illustration of these tradeoffs, Aydin et al. (2017) introduce uniformly distributed servicing times (for cargo operations) and recognize its influence in bunkering decisions as part of a discretized dynamic programming model. They found that the uncertainty of waiting times affects the optimal speed decision necessary to complete the voyage on time.

One of the best candidates to recognize bunkering operations is AIS data. Once used strictly for navigational purposes, AIS data has created an interesting niche in the areas of research and commercial management in shipping. It has being used to recognize shipping information related to near-miss ship collisions (Mou et al., 2010; Zhang et al., 2016), ship emissions (Tichavska and Tovar, 2015; Goldsworthy and Goldsworthy, 2015; Johansson et al., 2017), global trade flows (Adland et al., 2017; Jia et al., 2017), the spatial dynamics of fixtures (Prochazka et al., 2019), and underwater exposure of the marine fauna (Merchant et al., 2012), just to mention a few applications.

Recognizing vessel patterns from AIS has gained further attention, as more and better-quality data has become available. An early automatic method to recognize vessel patterns by AIS can be found in Pallotta et al. (2013). The work of Cazzanti and Pallotta (2015) presented one of the first approximations to the classification of areas where vessels are stationary by utilizing the Traffic Route Extraction and Anomaly Detection

algorithm embedded with DBSCAN. They classify anchored vessels by detecting rapid changes in the vessel course over ground while inside an identified cluster. Rather than detecting stoppages, Sheng and Yin (2018) adapted the DBSCAN with a hierarchical density estimate to extract the common shipping patterns in the Port of Tianjin. As an interesting twist, Lei (2015) provided a method of statistical vessel trajectories recognition to detect deviation from common patterns. In related work, Zhang et al. (2019) used vessel patterns to recognize hotspots for high vessel speeds and matching this with the spatial distribution of shipping accidents within the port of Singapore.

The challenges in handling AIS data have themselves provided some interesting topics that have now become new research areas. An important issue addressed by some papers, and of interest in the context of this paper, has to do with the reliability and completeness of the information derived from AIS data. Adland et al. (2017) compared the volume of crude oil exports derived from AIS data with official customs data. They also recognized the challenge of generating micro-level shipment information and the prospects of strengthening detection algorithms. Dobrkovic et al. (2018) documented the existence of incomplete and noisy AIS data, mostly on the high seas and in high-traffic areas, and provides an algorithm to adjust trajectories representing maritime lanes. Related to techniques to recognize patterns, Sheng and Yin (2018) stress the difficulty of vessel traffic analysis because of the large quantity of raw AIS data and proposed a method by a revised DBSCAN algorithm to cluster the information into valid traffic lanes.

Advancing in ways that aim to replicate humans' ways of solving problems, Machine Learning (ML) algorithms have recently proven successful in many settings. In this paper, I adapt the unsupervised learning algorithm DBSCAN due to its range of relevant built-in capabilities. The motivation behind DBSCAN derives from the goal of recognizing clusters inside large databases with arbitrary shapes (Ester et al., 1996). DBSCAN is used extensively on many real-world problems and remains relevant as an important and flexible clustering algorithm, as suggested by Schubert et al. (2017). A testimony of this flexibility can be observed in the many variants of the algorithm summarized by Khan et al. (2014). In shipping, Fuentes and Adland (2020) used DBSCAN to derive access routes to the Suez Canal. In a more general implementation,

Cazzanti and Pallotta (2015) and Sheng and Yin (2018) use the method for vessel pattern recognition, aligning the motivation behind DBSCAN with what researchers are faced with when using AIS data. For this paper, DBSCAN proves to be particularly useful in finding anchorage clusters of different and arbitrary shapes.

From this review I can recognize at least three relevant gaps in the literature. Firstly, there is a need to develop a framework for the standardization and extraction of micro-level empirical data on global bunkering operations. Secondly, optimization models for the bunkering location problem must be expanded to take into account bunkering waiting and servicing times, in addition to the classical fuel price-deviation tradeoff. Thirdly, algorithms for vessel pattern recognition need to be extended to cater for a more granular spatial analysis at the micro level, such as the ability to recognize a rendezvous between two vessels in an offshore barge-to-ship bunkering operation.

1.3 **Methods**

This section describes the components of the algorithm used to generate bunkering statistics. First, the DBSCAN algorithm is reviewed to help in introducing the specific application developed on this paper. Then, the bunkering algorithm is described in three subsegments which I consider to be its natural subdivisions: bunker barge recognition, stopped vessel recognition and bunkering operation recognition. The description of the full bunkering algorithm can be followed as pseudo-code in the Appendix A.

1.3.1 DBSCAN algorithm

The DBSCAN algorithm (Ester et al., 1996) classify positions (points) into clusters C based on parameters selected by a user. In contrast to other machine learning algorithms, DBSCAN does not require predefined input shapes (unsupervised learning), which is a convenient condition for discovering arbitrary shapes such as anchoring patterns.

The algorithm works iteratively one position –from a database *D*– at the time and classify clusters by defining groups of positions from the predefined parameters epsilon distance (ϵ) and minimum of positions (minPts). ϵ works as a threshold for the

detection of neighboring points when compared against distances calculated between every pair of positions $dist(p,q)$ using consistent distance measurements (e.g. Euclidean, Manhattan, etc.). A summary of the algorithm is presented as follow:

Step 1. $Q = \{q \in D : |N(dist(q, \forall p \in D), \epsilon)| \geq minPts, p \neq q\}$

Step 2. $C = \{q, q' : q, q' \in N(q) \cap Q, q \neq q'\}$

Step 3. $C = \{r: dist(r,q) < \epsilon, q \in C, \forall r \in D \backslash Q\} \cup C$

Step 4. $Z = \{z \in D \backslash C\}$

where, Q is a set of core points, C is a set defining clustered points, Z is a set of unclassified points or noise points, N is a set of points within a neighborhood, and D is a database set including all points.

In step 1, a cluster starts by declaring core positions (*q*). They are classified as such if they have a larger or equal number of neighbors than the minPts within the predefined epsilon distance. In step 2, q members of the same neighborhood (N) are merged in the same cluster. The positions not classified as core positions are kept for further classification.

From the clusters built from *q*, an additional refinement (step 3) expands the clusters by detecting border positions (*r*) from those in D not labelled as *q*. These are positions that are within ϵ of already defined *q* inside clusters (directly density-reachable) or positions that connects to *q* via sequence of other *q* (density-reachable). The positions not labelled in either of the classification processes are then classified as noise positions (z) (step 4) and are not part of any cluster.

The algorithm generates clusters that fulfill two main conditions. First, the maximality condition states that if *r* is in a cluster and there is a *q* that is density-reachable from *r,* then *q* also belongs to the cluster. Second, the connectivity condition states that for all *r* and *q* elements of a cluster, *r* is density-connected to *q.* That is, both *r* and *q* are density-reachable to some point *o*.

The bunkering recognition algorithm uses DBSCAN to classify anchoring patterns by modelling the expected movement of a vessel from the forces potentially affecting it

while at anchor. Further description of the parameters selection is covered in the following sections.

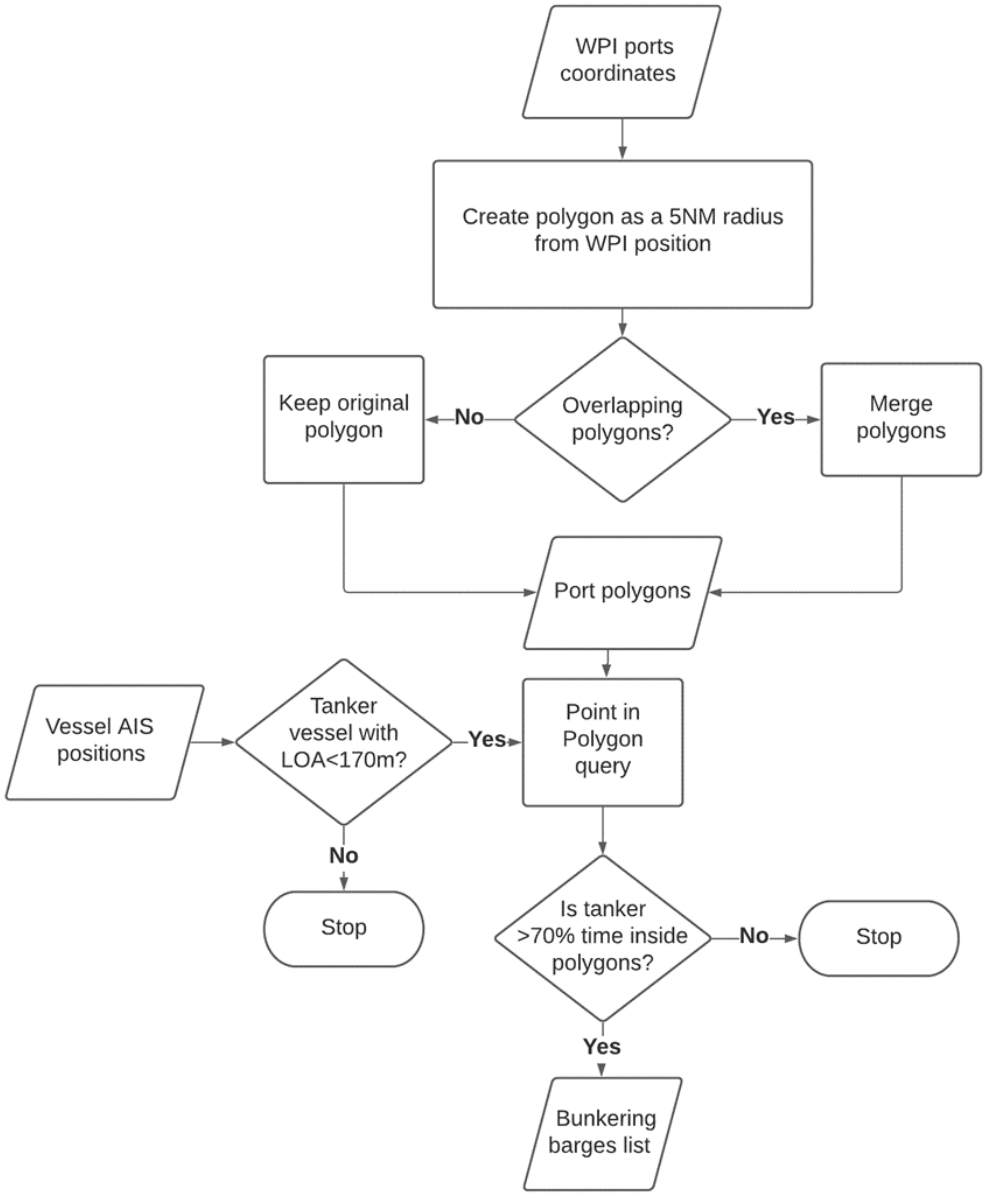
1.3.2 Bunker barges recognition

As bunkering is a marine operation considered as high risk, given its potential impact on the environment and human life should fuel spills occur, the operations are expected to occur within areas where the external natural conditions permits a safe connection between the vessel and it supplier (Ford, 2012). Such conditions are normally fulfilled at places such as anchorages or berths.

The former condition helps in reducing the bunkering operations search space to ships linked within port areas. A procedure to recognize both, the polygons where bunkering operations are carried and a bunker barge is summarized in Figure 1. First, polygons are generated from the World Ports Index (WPI) coordinates. The creation of the “local waters” polygon, where bunker barges are expected to operate, is defined by a circle with a radius of five nautical miles from the port’s position. Any ports where the resulting circles are intersecting are merged and counted as a joint port. Then, a Point-In-Polygon (PIP) query conducts the comparison of the vessels’ positions and the defined anchorages. In spatial analysis, a PIP serves to detect if a polygon has a point inside of it. Note here that it is considered good practice to align both the polygon and the point to the same coordinate standard, which for this case is WGS84. Each ship position, as reported by AIS is pre-filtered via PIP query against larger and fewer polygons to include only those within the region of interest. The valid positions are iteratively tested against the port “local waters” polygons using the PIP query.

Figure 1

Procedure for generating port polygons and recognition of bunker barge prospects



For a tanker vessel to be considered a candidate for use as a bunkering barge, it must have a length overall (LOA) lower than 170 meters and stay inside the local waters of a port or ports for 70 percent of the time. These filters are applied to reduce the complexity of subsequent sections of the algorithm where the position of an ocean-going vessel is tested against a list of prospective barges to recognize the one that provides it with fuel. The maximum LOA criteria is necessary to classify vessels with sufficient maneuvering flexibility that are capable of being moored alongside a bigger vessel with relative ease in restricted navigational areas and is based on the largest LOA (168.83 meters) for vessels classified as “Oil Bunkering Tanker” at the Clarksons’

World Fleet Register database. The minimum stay criteria follows an analysis of the oil bunkering tanker vessels' relative stay at port. From the results, bunker barges shows in average to stay at port waters for 70% of the time as calculated from their AIS positions inside port water relative to the total positions of the vessel. Additionally, barges declared as “Slop Reception Vessel”, “Water Carrier” or “Waste Disposal Carrier” are removed as they could be confused for bunkering operations as they also perform operations alongside bigger vessels.

The application of these filters (line 16 -37 in Appendix A- Algorithm 2) results in a list of bunkering barge candidates in the Mediterranean and Marmara Sea together with their port(s) of service. This list is then iteratively tested for proximity against a vessel waiting for bunkering service. A vessel is recognized as a waiting vessel from the moment it is stopped, an operation described in the next section.

1.3.3 Stopped vessel

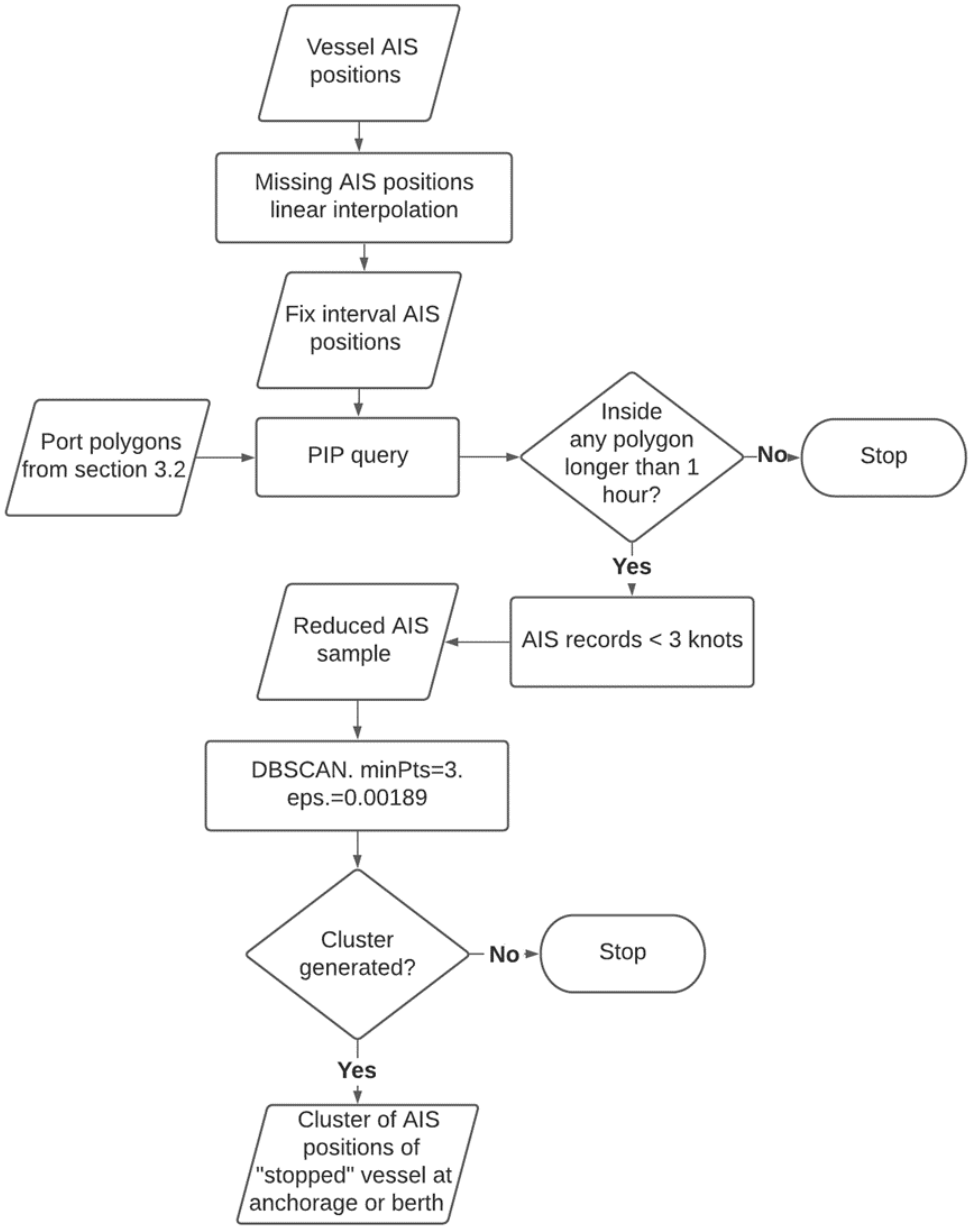
In most settings, a stopped object can be identified simply when it ceases to move, that is, when its speed is zero. AIS data also provides the vessel speed at every timestamp calculated from the distance between coordinates and the time between positions, but it cannot be used directly to find stopped vessels in this framework. A vessel with zero speed is easily recognizable but not observed often, unless the vessel is moored to a berth or kept steady using a dynamic positioning system. Given that most bunkering operations are performed at anchorage, it is expected that a “stopped vessel” will show some movement in a somewhat regular drift around the anchor position—an effect termed “swinging on the anchor”—due to changing wind, tides or sea currents.

To properly account for this, it is necessary to adjust the Cazzanti and Pallotta (2015) stopped vessel recognition algorithm. The algorithm adaptation for recognition of stop vessels under bunker operations is displayed in Figure 2 .The first step is to reduce the sample by recognizing the positions with a higher likelihood of being at anchorage. Such is the case for vessels with speed less than three knots, being positioned inside port waters polygons and staying more than one hour. This filter is important as DBSCAN recognizes position density by comparing distances; therefore, a reduced database helps in simplifying the algorithm while assuring that it does not hit a memory

wall due to bigger distances matrices. The implications of large databases in the DBSCAN algorithm are further analyzed as part of the empirical study on Section 1.4.

Figure 2

Procedure for stopped vessels recognition



Given the proximity of vessels' positions on short timescales, the Euclidean distance calculation method (see Equation 1) is sufficiently accurate to build the distance matrix between all position combinations.

$$\left(\sqrt{(lat.1 - lat.)^2 + (long.1 - long.2)^2}\right) decimal\ degrees\ x \frac{60nm\ x1852m}{decimal\ degreesxnm} \quad (1)$$

The calibration of the DBSCAN parameters follows the characteristics of a vessel anchored with allowance for some movement. For a cluster to be constructed it requires at least three positions and a threshold maximum separation of 210 meters. The chosen setting considers that a vessel may move with a constant upper speed of 0.68 knots for ten minutes, equivalent to 0.11333 nautical miles (210 meters[2]) (see Equation 2), and still classify as a "stopped vessel". The introduction of DBSCAN and recognition of a bunkering stoppage at anchorage is summarized from line 27 to 46 in the Algorithm 3 of Appendix A.

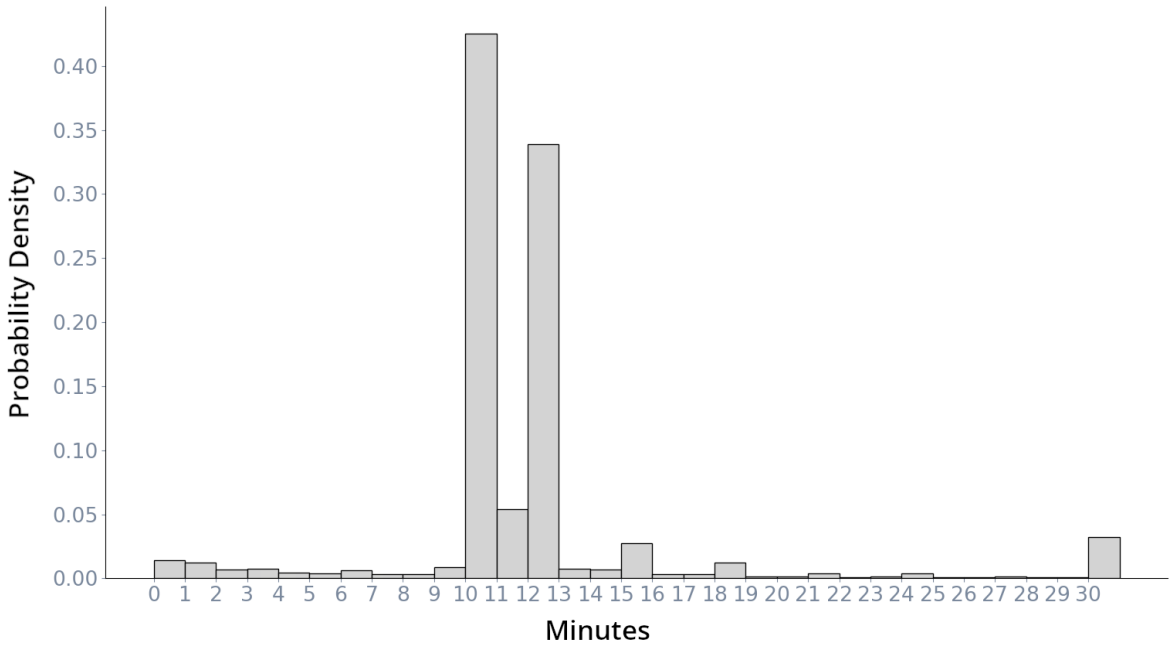
$$\frac{0.68nmx10min}{60min} = \frac{0.11333nmx1852m}{nm} = 209.9m \quad (2)$$

A fix interval is necessary to support the assumption of a vessel movement driven by external forces. As Figure 3 illustrates, from the raw data most of the observations have a frequency of ten minutes. Therefore, this frequency is used to avoid large interpolations and to keep most of the original observations within the data. Any missing positions in the ten minutes interval can, for instance, be generated from the Prochazka and Adland (2019) maritime track interpolation method. Their method uses a prebuilt network of the world oceans constructed from a grid mesh and interpolates the vessel route by finding the closest route from origin to destination. However, as they recognized the model performs better for long distances, and given the closeness of the clustered points, the missing positions are here instead based on linear interpolation of the surrounding positions.

[2] The equivalent of 210m in decimal degrees is 0.00189. This works as a threshold for this level of accuracy only if the AIS positions have at least five decimals.

Figure 3

Frequency between positions at raw database



Note. The rightmost vertical bar includes all frequency observations above 30 minutes.

The stated maximum separation threshold is derived from the phenomena likely to affect the drift of a vessel. The movement of a vessel relative to its anchor position depends on external forces like the wind, tides and currents. The relative direction of seasonal winds and prevailing currents and the closeness of the defined anchorage areas to the coast define the dynamics of vessel movement at the anchoring position, but in here, the focus is on the effect of tidal currents[3]. Using data from 1986 to 2017, collected from surface drifters' installed throughout the Mediterranean, Poulain et al. (2018) reported a maximum tidal current amplitude of 35 cm/s (0.68 knots) along Adventure Bank off western Sicily. This provides an upper bound to the external forces that could cause a rapid change in a vessel's position while at anchorage. This upper bound is the maximum allowed separation between points in the construction of the cluster and, thus, forms the basis for the epsilon parameter of the DBSCAN algorithm. In Section 1.4, a sensitivity analysis compares the performance of this selection on discovering anchoring shapes. To adjust the settings of the algorithm to different

[3] The effect of tidal streams can be replaced or used jointly with other external forces capable of causing rapid vessel drift while at anchor (e.g. Ro-Ro vessels has larger exposed surface to wind, hence, the maximum swing on anchor is likely caused by wind change.)

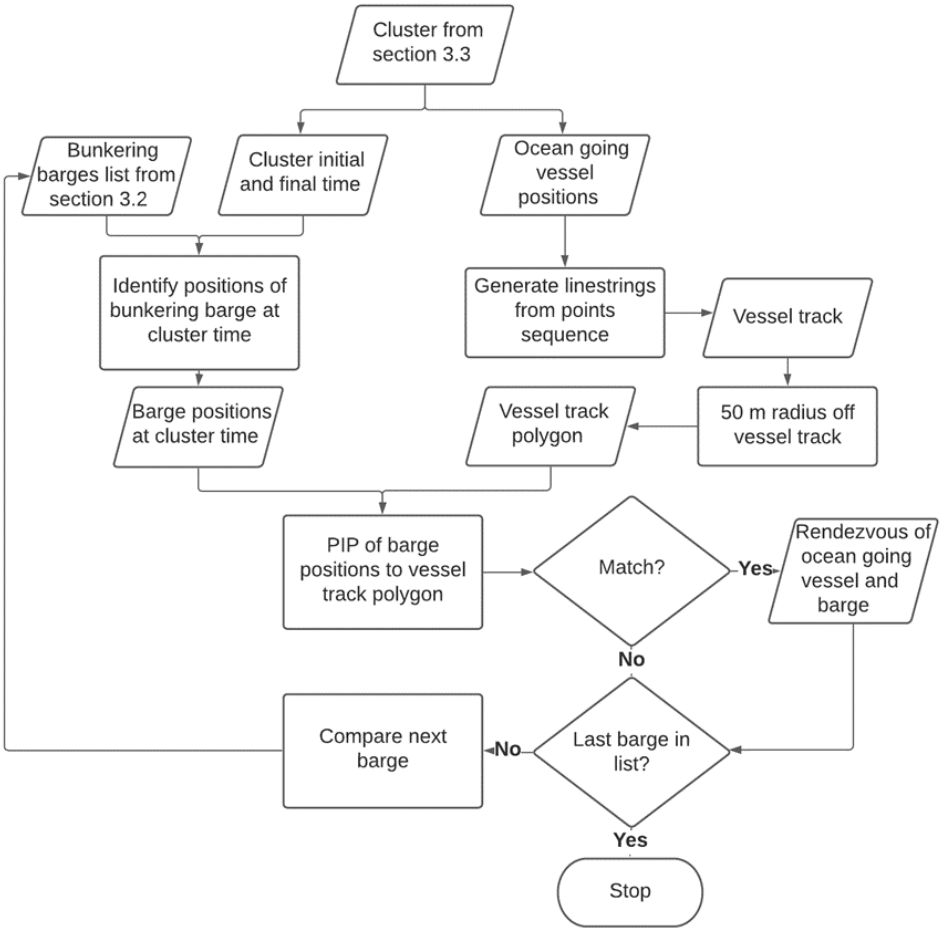
geographical regions, information about tidal streams can be found in Tidal Stream Atlases or Sailing Directions.

1.3.4 Bunkering recognition

Recognizing a stopped vessel and having the list of bunker barge prospects serve as the basis for identifying a bunkering operation. This section consolidates the results generated in previous sections as shown in Figure 4 (e.g. bunker barges list and anchoring clusters) to discover the rendezvous of an ocean going vessel and a bunker barge. Note that the algorithm can only recognize an ocean-going vessel supplied by a bunker barge and not vessels bunkering from shore connections.

Figure 4

Procedure for discovering the rendezvous of an ocean going vessel and a bunkering barge

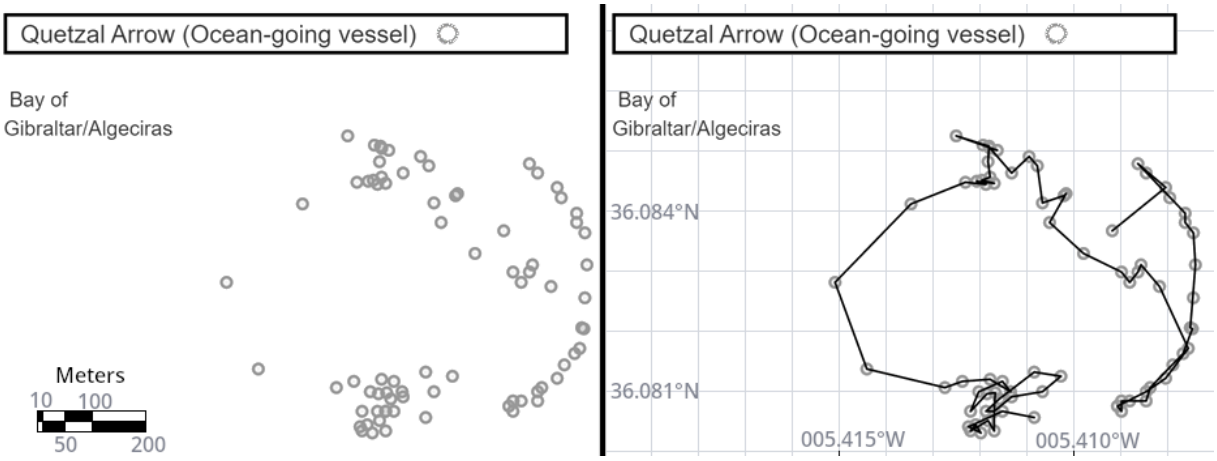


The stopped vessels' AIS information includes positions and timestamps that allow identifying the vessel track. This track represents the drifting pattern of the vessel while

at anchorage and results from the transformation of individual positions (points) to a single line string. The line string approximates the track that the vessel could have followed if positions were observed more frequently than the 10 minutes in the sample data (see Figure 5).

Figure 5

Line string construction from a vessel's positions



The next step is to generate a polygon around the track that represents the close quarters (i.e. immediate vicinity) of the vessel. A bunker barge that is alongside the ocean-going vessel will be positioned only meters away and match the time period of the ocean-going vessel while stopped. This constraint, together with the start and end time of the period at anchor, provides the conditions to test all bunker barges prospects with a PIP query. To avoid detecting barges that are simply passing by, their positions must remain within the "stopped vessel polygon" for at least one continuous hour.

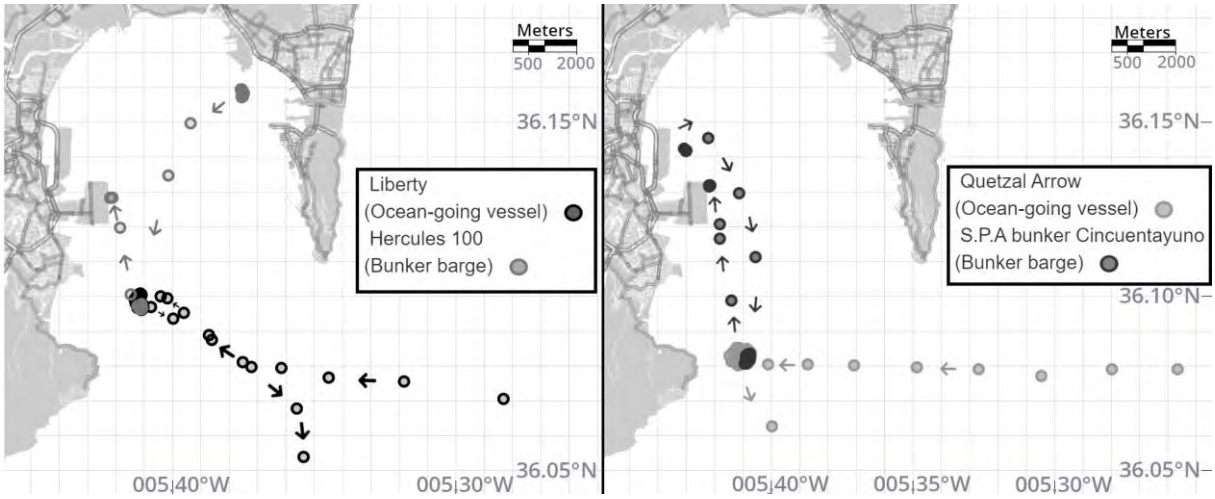
The results from this algorithm generates a database that includes information such as the ocean-going vessel information, the time of arrival and departure at the anchorage/berth, the timestamps for the bunkering service, bunker barge information and the location (port) of the bunkering event.

1.3.5 Algorithm test

I illustrate the algorithm performance by following two ocean-going vessels and visually validate their bunkering operations at the Algeciras port in Spain. The two plots shown in Figure 6 displays the arrival/departure of the vessels 'Quetzal Arrow' and 'Liberty' to the bay of Algeciras/Gibraltar and their servicing by bunker barges 'S.P.A Bunker Cincuentayuno' and 'Hercules 100', respectively.

Figure 6

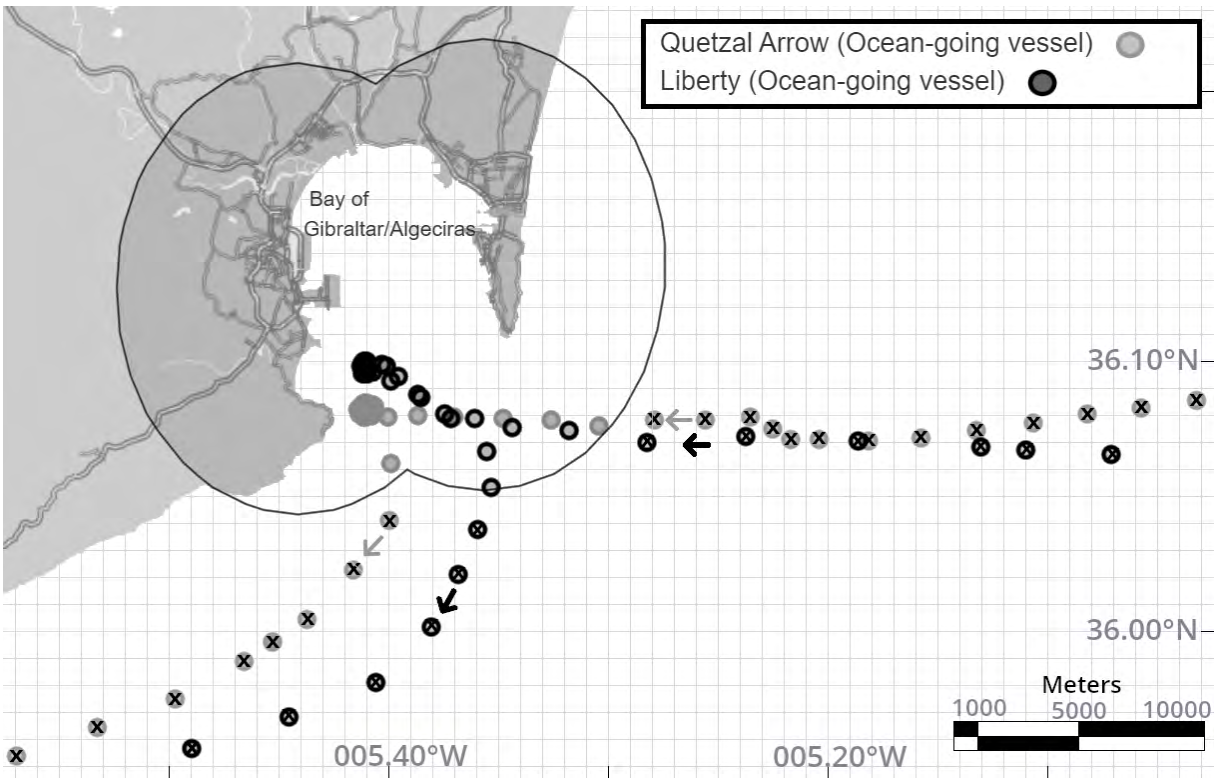
Plots of ocean-going vessels and bunker barges sample in the Bay of Gibraltar/Algeciras



Assuming that the bunker barge list and port/anchorage polygons have already been constructed (from Algorithm 2 of Appendix A), the first filter (see Figure 7) removes the ocean-going vessel's positions that are outside of the port waters and keeps the recorded positions with less than 3 knots speed. The result is a limited search sample that reduces the complexity of implementing the DBSCAN algorithm.

Figure 7

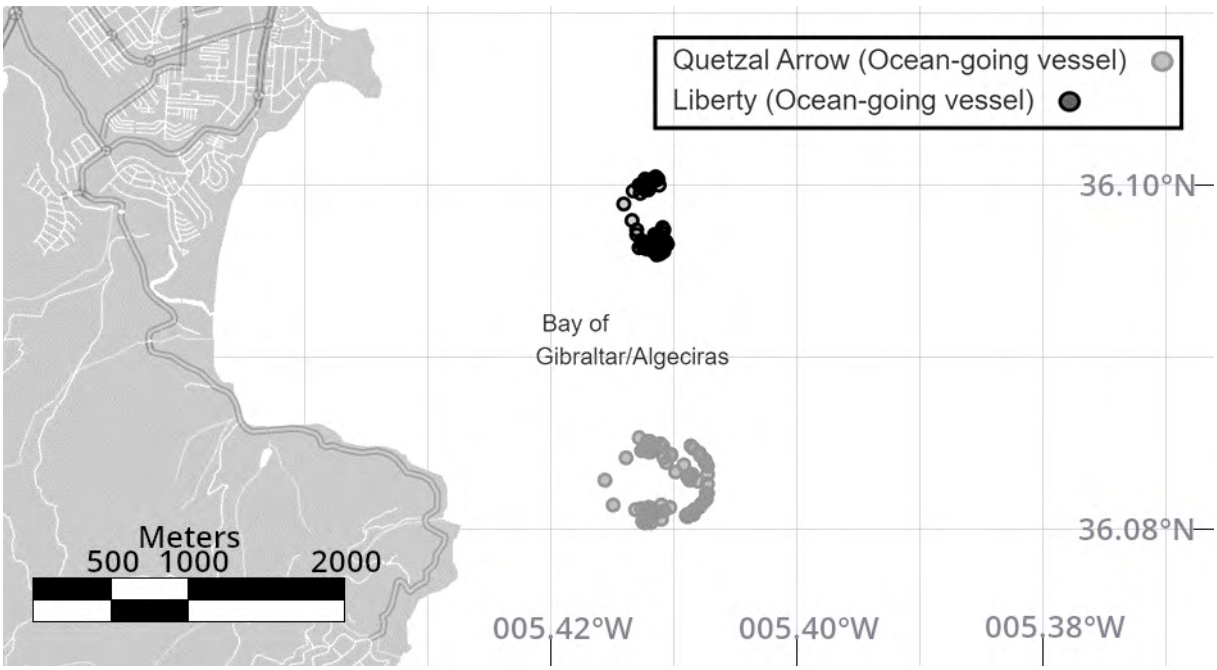
Ocean-going vessels filtered positions inside local waters



The DBSCAN adjusted parameters recognizes clusters by having at least three positions within 210 meters (0.00189 decimal degrees) of each other. As described in Ester et al. (1996), for the minPts parameter, a relatively low value must be used to include all points belonging to the same cluster. As one position cannot build a cluster and two positions can be confused with a vessel slowing down on maneuvers, the setting herein uses three positions as the minimum available option. This setting should remove, in a normal scenario, the positions not belonging to the cluster of anchored positions or swing radius of the vessel. The resulting cluster of anchored positions for the two ocean-going vessels are illustrated in Figure 8. Note that the resulting shape is irregular but most closely resembles the circle implied by a vessel swinging on anchor due to varying tidal streams or wind conditions.

Figure 8

Ocean-going vessels anchoring positions as determined by the DBSCAN algorithm



To compare the positions of barges against those of stopped ocean-going vessels requires the construction of base polygons for conducting a PIP query. The polygon construction starts from connecting the ocean-going vessel positions by lines. The polygon is then generated by extending a radius from the created line strings (lines 5-8 in Algorithm 4 of Appendix A). The resulting polygon represents the potential areas within which a barge is expected when positioned alongside the vessel. Here, a radius of fifty meters represents the maximum distance between the two GPS positions (vessel and barge) for a large beam vessel[4].

The rendezvous of an ocean-going vessel and a bunker barge is recorded in a database with the specifications of both vessels. The entry is complemented with the time that the barge entered the stopped vessel polygon (the service start time) and the time of the last position inside the polygon (the service end time).

To decide whether the barge-to-ship bunkering operation is undertaken at anchorage or at berth, a first investigation performed a PIP query against the high-resolution coastline polygons from the Global Self-consistent, Hierarchical, High-resolution

[4] In theory, a dynamic buffer adjustment can be done with the AIS information for both vessels. This would involve the distance to antennas (dimC and dimD), the vessels' beam and the position of the barge alongside the larger vessel. Such a dynamic radius is not implemented here.

Geography Database from the National Geophysical Data Center of NOAA (2017). As the polygons are built on natural coastline features, some berths appear as artificial extensions of the coastline but are not properly identified as being part of the polygon. To conduct a proper PIP query, I therefore manually construct the berth polygons with the help of WPI port positions and identified the berthed vessels as those inside the polygons.

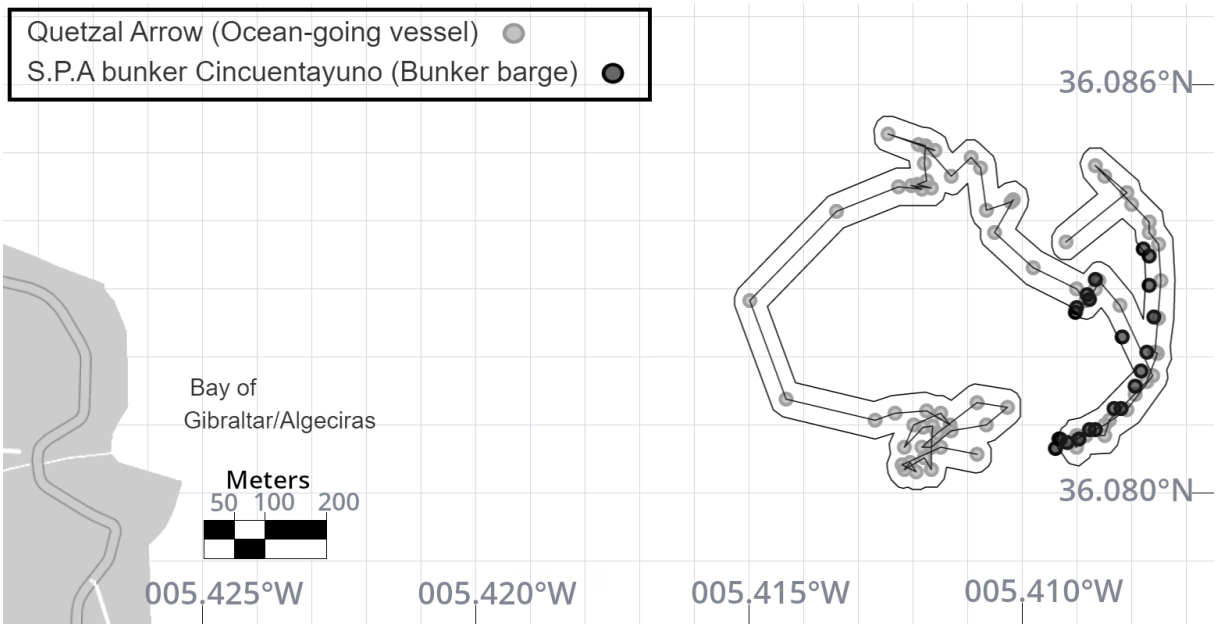
It is important to discuss the concern that might arise with the proximity calculation when distances are measured from GPS positions. GPS is operated and maintained by the US Air Force Space Segment and they commit to broadcasting the GPS signal in space with a global average user range error (URE) or error from the signal in space of $\leq$7.8m (National Coordination Office for Space-Based Positioning, 2020). They have also declared that as from 2016, their performance is URE $\leq$0.715m, 95% of the time. Using earth-based correctors such as the Differential Global Positioning System[5] (DGPS), provided along most of the main shipping routes, might allow users to achieve accuracy of 1 to 3m (Kim, et al., 2017). Given that bunkering operations take place close to shore, typically within the range of DGPS signals, there is reason to believe that the position accuracy is satisfactory.

Figure 9 illustrates how the algorithms identifies Quetzal Arrow as a stopped vessel at anchor, that the vessel is eventually joined by the bunker barge S.P.A Bunker Cincuentayuno, following which both vessels drift together as the bunkering operation takes place.

[5] Differential Global Positioning System is a system that improves location accuracy of GPS satellite signals by a calculation of fixed ground reference and the difference with the GPS positions. It broadcasts the adjustment on radio frequencies.

Figure 9

Ocean-going vessel and bunker barge rendezvous in bunkering position



The most computationally intensive and complex part of the implementation of DBSCAN in this case is the distance matrix calculation for detecting the 'stopped vessels' position cluster. Computational improvements to the distance matrix generation can be achieved with parallel Graphics processing unit (GPU) acceleration (Man et al., 2011), which reduces the overall processing time of the DBSCAN algorithm. It is worth noting here that even though an efficient distance matrix calculation reduces the processing time, data features such as vessels having several visits and the large scale of the ship-barge matching makes executing a sequential algorithm a lengthy process. To reduce the cost of complexity, I assume that a vessel with over 6,500 observations (approx. 45days) inside a port polygon (a single visit) is no longer a vessel waiting for service but a laid-up vessel and the visit is not assessed as a prospective bunkering operation. This criterion is helpful to avoid long layovers, which are difficult to assess based on the assumptions of bunkering service calculation (e.g. waiting time calculation from vessel arrival until vessel served by a bunkering barge). Another convenient condition is that it precludes the algorithm from hitting memory barriers and bottlenecks in the computation that would be caused by large distance matrices.

Moreover, as the algorithm iterations do not need information from other processed vessels, it can be distributed to reduce the overall computational time. The DBSCAN used as the machinery of this algorithm is known to have a complexity of $O(n^2)$ in the worst case (Ali et al., 2010). The prospect of having long processes also works in favor of the application of a distributed setting. Implementing the distributed setting requires separating the whole database in chunks grouped by vessel and assign a vessel per node. To avoid hitting the RAM memory limit of a computer in the cluster, because of the stored distance matrices, and to use the most of the CPU cores I use only the physical cores (12 out of 24) of 10 computers/node for 120 distributed operations. MPI performs the communication between cores and nodes and the assignment of vessels positions defined in an asynchronous process. This setting permits that the results of a core can be written in a central database while other processes in the remaining cores have not finished. This also permits that a new vessel to be processed can use an idle core.

The use of additional cores increases the likelihood of having memory error shutdowns, as synchronous processes will claim resources from the same pool e.g. RAM. Memory error is a critical fault in a distributed process and causes the overall calculation to have an abrupt stop, should that occur. Such scenario might happen if any random allocation of processes (vessels per core) within nodes, demand more RAM than the local node capacity (47.3 GB) can supply. As DBSCAN matrices are the biggest objects in the algorithm[6], having control on its generation works as a safety “relay”. The control is done programmatically by setting an upper bound (6,500 records) within the code, causing a limit in the memory consumption per node. This balance is necessary to assure that the solutions are given in a practical time and removing the risk of a shutdown and loss of information after processing for some time.

1.4 Empirical study

Using the algorithm outlined above, it is possible to build a database with bunkering event information from a dataset of raw AIS data. This section reports the empirical test of the algorithm and the results in the form of descriptive statistics for bunkering

[6] DBSCAN distance matrices of size (nxn) with double precision floating points (8 bytes).

operations. Additionally, a sensitivity analysis with the DBSCAN parameters test the performance of different inputs in the recognition of anchoring shapes.

The algorithm is fed with observations from an AIS database that comprises $1.3x10^9$ positions of vessels in the Mediterranean and Marmara Sea collected from February 2013 to June 2019 and provided by Vesseltracker.com GmbH. I implemented it with Python 3.7 in an asynchronous distributed setting defined with Message Passage Interface (MPI). The distribution was conducted in a cluster of 10 computers with 24 CPU cores (12 physical) each with 2.67 Ghz core frequency and non-shareable RAM memory of 47.3 GB.

1.4.1 Derived statistics and discussion

Running the algorithm results in 49,463 observations of bunkering operations in the Mediterranean and Marmara Seas, with more observations identified from 2015 onwards (see Table 1) given higher accuracy and frequency of AIS positions from improved coverage of satellite and shore AIS receivers.

Table 1

Bunkering observations for the Mediterranean Sea and Marmara Sea per year

Year	Bunker.ops	AIS Vessels positions
2013	105	$68x10^6$
2014	6160	$145x10^6$
2015	9408	$204\ x10^6$
2016	9159	$216\ x10^6$
2017	9997	$227\ x10^6$
2018	9931	$235\ x10^6$
2019	4703	$126\ x10^6$

Note. Observations for 2019 available until June

While there are few benchmark data available, the Gibraltar Port Authority (2019) provides some general bunkering statistics for comparison (see Table 2). The statistics shown in Table 2 suggests that with the better-quality AIS data available from 2015 onwards, the algorithm is able to capture around 80% of reported operations in the Port of Gibraltar.

Table 2

Bunkering operations in Gibraltar per year

Year	Gibraltar statistics	Bunk. algorithm	% of total
2013	5988	101	1.7
2014	5475	2745	47.8
2015	5571	4452	80.0
2016	5720	4261	74.4
2017	6298	5112	81.1
2018	5829	5176	88.8

As seen in Table 3, in order of importance, the most used ports for bunkering are Gibraltar, Algeciras, Istanbul, Ceuta, Piraeus and Valetta. This result is consistent with the observation that the location of a port (being close to main trade routes) is an important factor when selecting the bunkering port, as pointed out by Vilhelmsen et al. (2014). Additionally, Table 3 shows the average waiting time for vessels served for bunkers. This variable can be interpreted as a measure of reliability and punctuality of the ports bunker suppliers. According to Lam et al. (2011), this is an important factor (ranked fourth after price competitiveness) for ship operators when selecting a bunkering port as prolonged bunkering operations could lead to disruption of vessels schedule. Their description suggest that a ship operator with a tight schedule would choose bunkers from the ports with lower waiting times. For instance, a vessel bunkering at anchor would choose Gibraltar (1.68 hours) instead of its close neighbor Algeciras (5.20 hours) when the bunkering decision is based on service efficiency and all else equal (e.g. price, distance, etc.). A more detailed decision model would incorporate the tradeoff of several variables, such as price, route and waiting time.

Table 3

Bunkering operations derived statistics from January 2013 to June 2019

Port	Barges	Service time (hours)	Waiting time-berth (hours)	Waiting time-anchor (hours)	Ops. at berth	Ops. at anchor	Ops.	%
Gibraltar	40	6.63	3.88	1.68	323	23,794	24,117	48.92
Algeciras	46	5.96	3.80	5.20	2,165	7,230	9,395	19.06
Istanbul	23	3.37	6.95	4.02	43	3,797	3,840	7.79
Ceuta	8	5.02	1.45	3.63	17	2,296	2,313	4.69
Piraeus	59	4.90	4.35	5.73	1,145	825	1,970	4.00
Valletta	17	6.32	4.11	2.32	1,246	245	1,491	3.02
Barcelona	5	4.34	4.85	-	1,103	0	1,103	2.24
Limassol	4	3.95	2.72	3.53	196	877	1,073	2.18
Livorno	7	4.44	5.56	-	857	7	864	1.75
Augusta	17	3.78	12.47	5.48	114	739	853	1.73
Genova	10	3.92	6.46	-	665	0	665	1.35
Valencia	5	4.91	4.35	6.05	534	20	554	1.12
Napoli	11	3.18	5.03	7.25	298	25	323	0.66
Syros Isl.	1	3.26	-	2.04	1	314	315	0.64
Tanger	1	5.83	2.92	11.88	0	166	166	0.34
Chioggia	2	3.14	9.75	9.36	150	0	150	0.30
Mersin	11	6.22	3.46	11.69	101	6	107	0.22
Total	267				8,958	40,341	49,299	

Notes. Barges are counted as belonging to the last port where they had a bunkering operation. Table includes ports with more than 100 operations from 2013 to June 2019. Service and waiting times are winsorized at the 10% level.

A comparison between variables suggests weak linear relations, as shown in a correlation matrix in Appendix Table B.1, and a marginally moderate relation (0.30) between operations and number of barges. This could be interpreted as the variables changing based on different underlying dynamics not described in this study. The number of barges deployed in a port could result from a long-term fleet strategy or the fleet plan to fulfill an uncertain expected demand (future operations); the success of meeting that demand and its variance could be suggested by the moderate correlation coefficient (0.30). Waiting times might be driven by the tank capacity of the bunker barges, which result in the number of trips to a restock place, number of vessels waiting for bunkers, or vessel clearance of port requirements such as free pratique. The number of operations could result from a tradeoff in the port's proximity to main trade routes, bunker price and port efficiency (waiting time and service time). The relation

might not be linear, as the weight of importance on the factors can be interpreted as conditional on fuel price and market conditions. Finally, the bunkering service time can be more related to bunker quantity loaded per operation and individual barges pumping rates. This last hypothesis is analyzed later in this chapter.

Segmenting the results by vessel type shows an interesting outcome as seen in Table 4. According to Acosta et al. (2011), the bunkering service provided to liner vessels in the strait of Gibraltar is dominated by the port of Algeciras. They argue that the quick turnaround of these vessels requires them to be served while alongside the container port terminal instead of deviating to the port of Gibraltar and adding the costs of deviation and possible waiting time. They also mention that non-liner vessels, such as bulk carriers or oil tankers prefer to be served at the port of Gibraltar while at anchor as a result of generally lower bunker prices and a lower value of time. The results in Table 4 confirm these qualitative statements with actual data derived from the algorithm.

Table 4

Top 10 preferred bunker ports per vessel type

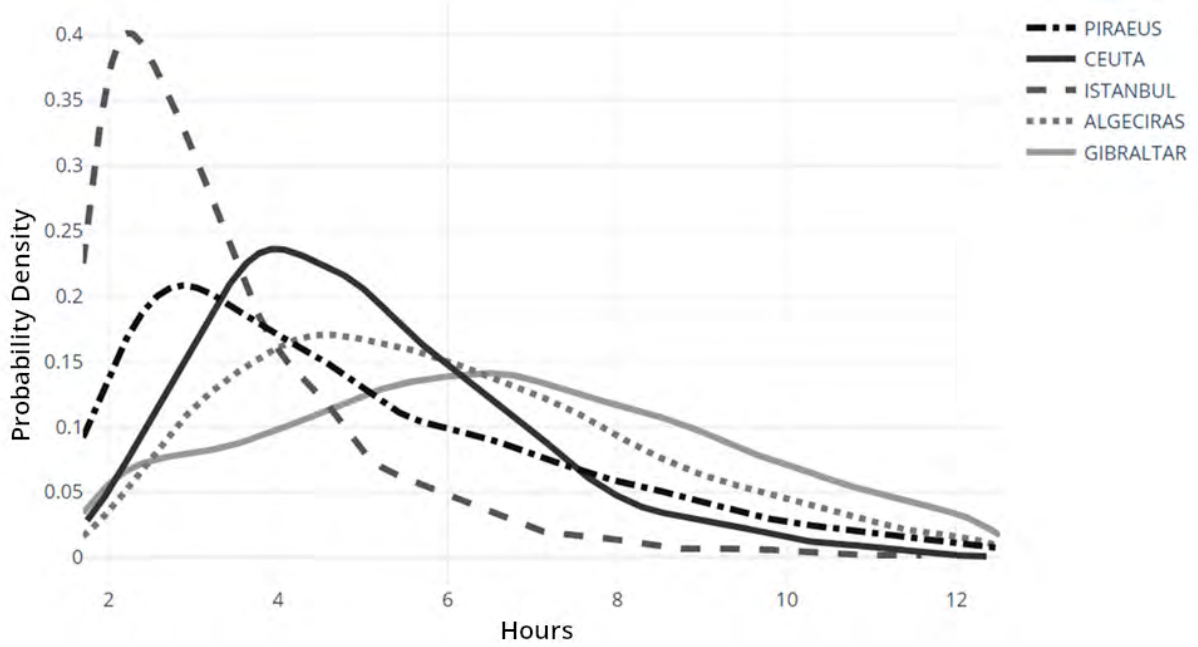
Liner vessels (Containers/Ro-Ro)		**Crude oil tankers**		**Bulk Carriers**	
Port	**Bunk.**	**Port**	**Bunk.**	**Port**	**Bunk.**
Algeciras	2,431	Gibraltar	658	Gibraltar	15,887
Valetta	1,032	Algeciras	144	Algeciras	3,300
Piraeus	973	Genova	23	Istanbul	1,879
Barcelona	532	Augusta	17	Ceuta	1,317
Valencia	503	Piraeus	13	Piraeus	273
Livorno	455	Ceuta	6	Limassol	126
Limassol	295	Limassol	6	Valetta	117
Tanger	260	Valetta	4	Barcelona	92
Genova	234	Livorno	3	Syros Isl.	82
Gibraltar	203	Barcelona	1	Genova	44
At anchorage	1,440	At anchorage	840	At anchorage	22,9371
At berth	5,829	At berth	37	At berth	350

The vessel service time is an important piece of information that helps to assess the effectiveness of the supplying barges in pumping the fuel, conducting the tanks survey and clearing the documentation that will let the vessel resume its voyage. The results shown in Figure 10 derives from the time difference of the first and last positions of the

bunker barge inside the stopped vessel polygon. This effectively assumes that the barge starts the bunkering operation as soon as it enters into the stopped vessel polygon and stops the service as it leaves the polygon, though in practice some time will be lost due to mooring arrangements and hose connection.

Figure 10

Service times approximation of the top five bunkering ports

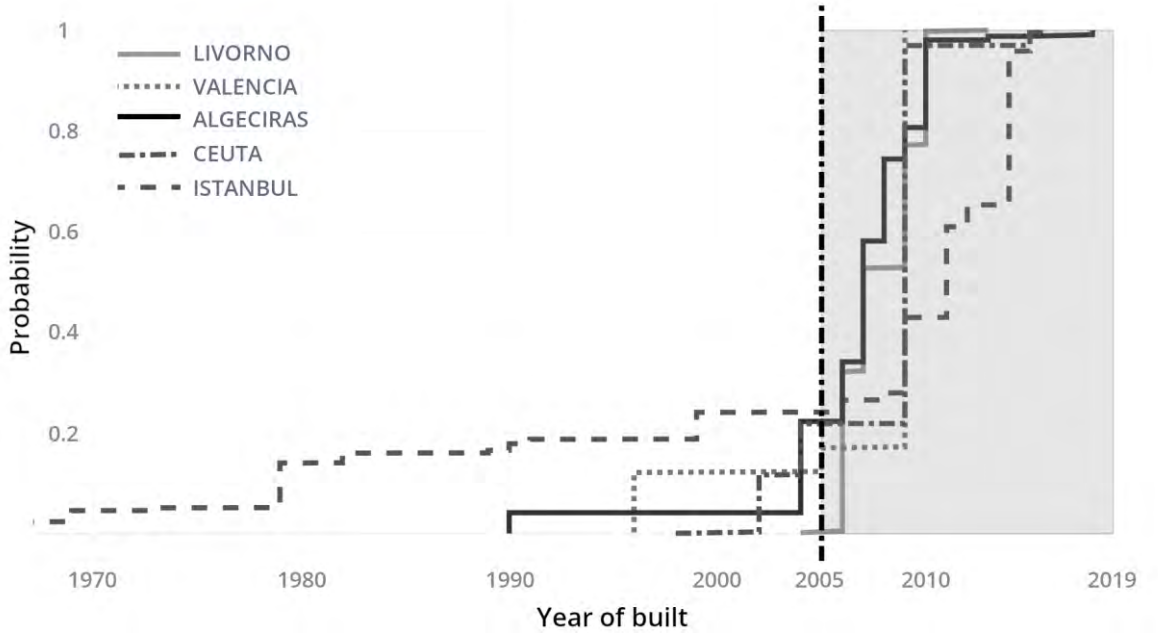


The service time is dependent on the pumping rate of the bunker barge and the quantity of fuel requested by a vessel. The first could be thought as being related to the barge's age. Older bunker barges will tend to have lower pumping capacity, due to more sediments in their tanks, older lines and older pumps. Newer barges can also indicate suppliers' willingness to invest in offering a better service, and so improving the overall port bunker servicing time. Figure 11 shows the difference in bunker barge fleet age distribution for the top five bunkering ports. The shadow region display part of the cumulative distribution where vessels are less than 15 years old (2005 laid barges or newer) with Livorno with 99%, Valencia with 83%, Algeciras and Ceuta with 78%, and Istanbul with 76%. The ranking signals the need to offer faster services based on the short calls of the vessels calling their ports. Specifically, the reason for calling at these ports has to do with waiting for Bosporus Strait transit (Istanbul), servicing passenger vessels (Ceuta and Livorno), fast turnaround services e.g. Ro/Ro vessels at Livorno, and loading/unloading containers at the port of Algeciras and

Valencia. Other drivers have to do with the individual behavior, expertise and sophistication of the bunker buyer. An approximation of fuel requirements could be estimated with knowledge of factors such as the duration of the next voyage, its expected speed, the total fuel oil tank capacity of the vessel, expected weather conditions, and draft restrictions. I assume these the very basic considerations that are made when requesting the necessary fuel for a vessel.

Figure 11

Empirical cumulative distribution function for bunker barges year of built at the top five bunkering ports



Notes. Ports with five or more bunker barges.

It is also of interest to investigate the factors influencing the bunkering servicing times. In a transparent market, all that would be needed is the quantity loaded and the pumping rate used. However, given that such data is unobservable due to the lack of publicly available bunker delivery notes, it is necessary to estimate the relationship using proxies. Taking bulk carriers as an example (the vessel type with the most records in the AIS data), the criteria chosen were (with expected sign in parenthesis): fuel tanks maximum capacity (+), next two legs voyage distance (+), average speed of the next two legs (+), age of barge (+) at the time of the bunkering operation and a dummy variable denoting whether the vessel is laden or in ballast (-). The maximum capacity of the fuel tanks in the ocean-going vessel ought to be positively correlated with the volume of fuel taken onboard, though there is no information available on the

amount of fuel remaining onboard. Similarly, the distance and speed of the next planned voyage serve as a proxy for expected consumption. It is worth noting that both variables will be known by the ship's officers (or can be estimated with some certainty) at the time of bunkering based on their voyage plan. The greater speed, the higher the consumption per nautical mile (nm) of distance (Ronen, 1982). The age of the barge is likely to influence the pumping rate, with higher age typically implying lower pumping rates and therefore longer servicing times. The impact of the variables was estimated in a fixed effects regression and the results are shown in Table 5.

Table 5

Estimates of the service time

Variables	(1) Service Time (hours)	(2) Service Time (hours)	(3) Service Time (hours)
Bunker capacity (m^3)(/100)	0.060*** (0.005)	0.055*** (0.004)	0.056*** (0.005)
Next voyage distance (nm)(/100)	0.008*** (0.001)	0.007*** (0.001)	0.007*** (0.001)
Next voyage average speed (knots)	-0.002 (0.012)	0.009 (0.011)	0.013 (0.011)
Barge age (years)	0.003 (0.006)	0.005 (0.006)	0.001 (0.006)
Laden vessel	-0.280*** (0.073)	-0.302*** (0.071)	-0.320*** (0.072)
Constant	5.745*** (0.214)	5.758*** (0.201)	5.749*** (0.202)
Barges Fixed Effects	No	Yes	Yes
Time Fixed Effects	No	No	Month
Number of observations	5176	5176	5176
R^2	0.067	0.067	0.067

Notes. Robust standard errors. Significance levels are 1% (***), 5% (**) and 10% (*). Service time and next voyage average speed windsorized at the 10% level. Regression carried on bulk carriers' records (23,294) with greater than or equal to two legs after the bunker stop (11,692) and non-empty values (6,392).

While the explanatory power is low (R^2 = 0.067), and could be interpreted as the independent variables can explain 7% of the service time variance, it is more important in the context of this analysis that the individual variables are highly significant and that the coefficient have a plausible sign, as a "common sense" validation of the variation

in bunkering servicing times generated by the algorithm. This is acceptable in this study, considering that the tested variables are proxies of the underlying effect.

Table 5 suggests that the fuel tank capacity of the ocean-going vessel and the upcoming voyage distance are significant at the 1% level. Their coefficients imply that 100 m^3 additional bunker tank capacity on a vessel leads on average to 0.056 hours (3.4 minutes) change to the service time or 100 nautical miles for the next voyage, an increase of 0.007 hours (0.4 minutes). The age of barges turns out to be non significant at a 95% confidence interval. This could be explained by some barge-owners having a rigorous maintenance plan that can include overhauling the pumping arrangement of underperforming barges, regardless of age. The laden vessel dummy variable has a negative coefficient; the reason is that laden vessels will typically be close to their maximum permitted sailing draft and therefore have less allowance to load fuel, compared to ballasting vessels of the same size. The laden vessel coefficient shows that laden vessels will have, on average, 0.32 hours (19.2 minutes) less service time than ballast vessels. There is no improvement from including barge fixed effects. Such effects could include the respective Captains'/Chief Engineers' pumping rate preferences.

It is worth noting that certain robustness tests were also performed. First, the model is iterated with 100 random samples of 1000 records each (approx. 20% of total sample) to check that the significance of the variables is not a result of the sample size. By averaging the coefficients, standard deviations, and p-values, the results shows that the variables bunker capacity, next voyage distance and laden vessel remains to be significant and the sign coefficients are consistent with the interpretation of the relations to the service time. Second, the Breusch-Pagan test for residual homoscedasticity shows a value of 44.74 and p-value < 0.05. This implies the rejection of the null hypothesis of homoscedasticity and supports the use of robust standard errors as reported. Third, the independent variables do not suffer from multicollinearity as the Variance Inflation Factor is lower than five in all cases. Fourth, the independent variables show exogeneity when compared via Pearson's r correlation to the model's residuals. Finally, the expected residual value is zero but the distribution is non-Normal. However, violating this assumption causes no significant problems as a comparison of

the kernel density approximation and the normal distribution suggests the difference is marginal.

1.4.2 Sensitivity analysis on anchoring cluster detection

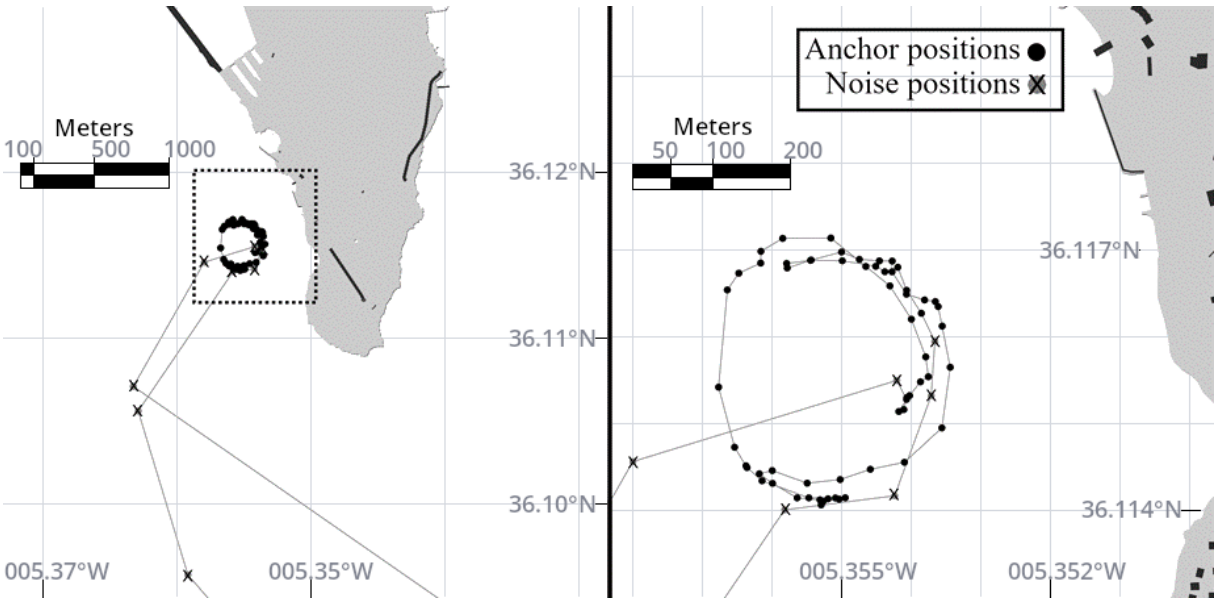
A good selection of ϵ and minPts is critical for detecting anchoring shapes from DBSCAN clustering. Hou et al. (2016) suggest that a priori knowledge of the phenomena to be clustered can be advantageous for selecting parameters, and Karami and Johansson (2014) argues that obtaining prior knowledge of the density shapes is seldom achieved. For the phenomena of ships at anchor, such prior knowledge is reachable from environmental information, alike the estimated speed of an anchored vessel approximated by the external factors affecting it.

Working with a single ϵ and a fix minPts helps in simplifying the parameter selection while maintaining a relatively stable performance. In order to prove such, this section includes a sensitivity analysis to compare among input changes of parameters on DBSCAN for the comparison of results to a synthetic database of anchoring shapes.

The synthetic database is produced by selecting at random, five AIS positions samples per port with registered bunkering operations at anchor. These samples are aligned with the timestamps of records resulting from the bunkering algorithm. Also, the records are extended six hours before and after the service time in order to test the parameters fitness to detect noise points. For producing the anchor labels, the records at the synthetic database are manually classified as either anchor or noise positions as illustrated in Figure 12. Note that the berthing records are removed, as the parameters that detects anchoring patterns, in theory, should also detect berth patterns where vessels are expected to move less.

Figure 12

Selection of anchoring positions for performance testing



The clustering performance is thereafter assessed by the Rand Index (RI). This measure appraises the accuracy of clusters matching records, ranging from zero when the clusters have no similarity and one when they are identical. In the case of the bunkering algorithm, such test is carried between the result on the DBSCAN labels -triggered by ϵ and minPts- and the synthetic data labels. RI is described in (Rand, 1971) as: given N points, $X_1, \dots, X_k$, and two clustering of them $A = \{A_1, \dots, A_k\}$ and $B = \{B_1, \dots, B_k\}$, then:

$$R(A,B) = \sum_{i<j}^{N} \frac{\gamma_{ij}}{\binom{n}{2}} \tag{3}$$

where,

$$\gamma_{ij} = \begin{cases} 1\ if\ there\ exists\ k\ and\ k^{'}\ such\ that\ both\ X_i\ and\ X_j\ are\ in\ both\ A_k\ and\ B_k \\ 1\ if\ there\ exists\ k\ and\ k^{'} such\ that\ X_i\ is\ in\ both\ A_k\ and\ B_k \\ while\ X_j\ is\ in\ neither\ A_k\ or\ B_k \\ 0\ o.w \end{cases}$$

$\binom{n}{2}$ *is the binomial coefficient*

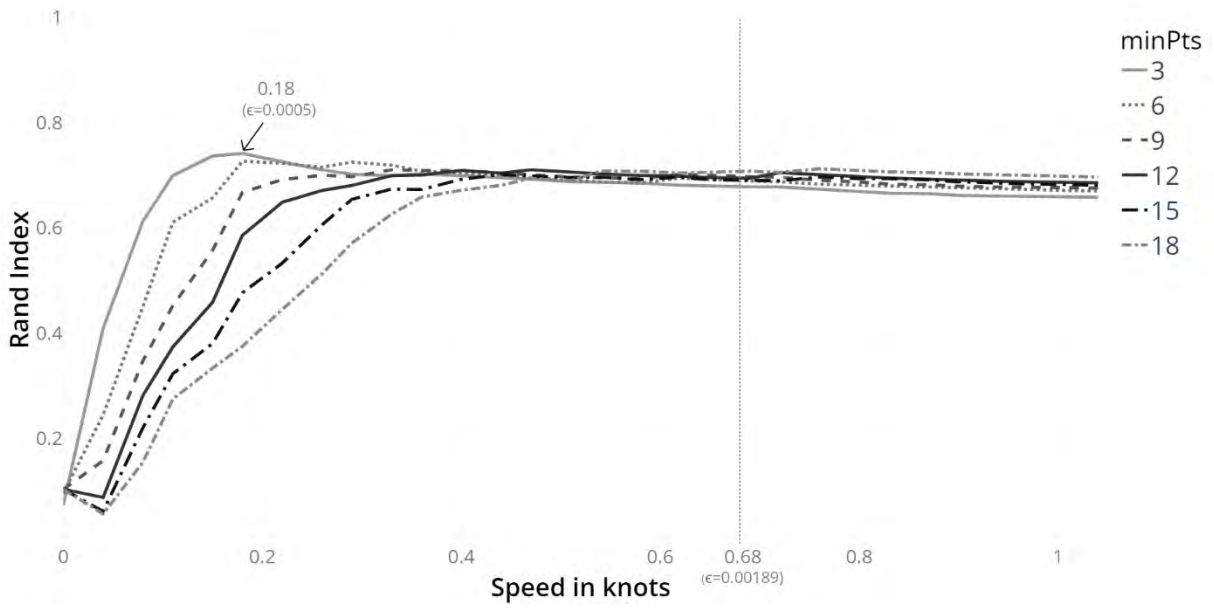
Running the sensitivity test in all ports results in an RI of 0.68 for the fix parameters used at the bunkering algorithm (0.00189 decimal degrees and three minPts). If compared to the maximum RI (0.74) at ϵ =0.00050, the difference as illustrated in

Figure 13 suggest that the fix parameter is within the best ϵ neighborhood. This is more evident when compared with higher values. For instance, running an ϵ of 0.0194, which is equivalent to seven knots, has an RI of 0.50 and the performance keeps reducing as higher ϵ are used.

Figure 13 also illustrates the influence of minPts in the anchoring cluster detection. Higher minPts shift the best area to higher ϵ .This is likely the case as the DBSCAN should search more distant points to form a cluster. As the input data is discretized to have a ten minutes frequency between positions, the minPts can be interpreted as the minimum sample necessary to cluster an anchorage e.g. six points is a one-hour sample. Using higher minPts works as a lower bound constraint as it will condition the service times to be higher than the minPts sample in order to be detected as a bunkering operation. This is relevant as some ports might consistently show fast services that would go undetected on large minPts.

Figure 13

Performance on clustering matching from DBSCAN parameters changes



Using fix parameters also helps in avoiding undershooting the cluster shape. Too low ϵ can rapidly make the DBSCAN to underperform as displayed in Figure 13. For that reason, it is more efficient to choose an above the best parameter in a close stable neighborhood. As seen in Table 6, the ϵ related to the max RI is different for every port, suggesting that using a very low fixed ϵ could end of being adverse on detecting

anchoring clusters. By deriving an upper bound with external factors information, the bunkering algorithm ϵ will consistently be above the best ϵ within a close distance. The RI difference in Table 6 shows the distance in RI values of the fix ϵ used for the bunkering algorithm and the best ϵ for every port.

Table 6

Performance on clustering matching segmented by bunkering port on three minPts

Port	RI (0.68kts)	Max RI	RI difference	ϵ for Max RI
Augusta	0.94	0.96	0.02	0.00050 (0.18 kts)
Ceuta	0.81	0.86	0.05	0.00120 (0.44 kts)
Civitavecchia	0.6	0.66	0.04	0.00050 (0.18 kts)
Gibraltar/Algeciras	0.78	0.85	0.07	0.00069 (0.25 kts)
Istanbul	0.60	0.72	0.13	0.00042 (0.15 kts)
Kalimnos	0.65	0.72	0.07	0.00030 (0.11 kts)
Limassol	0.73	0.83	0.11	0.00030 (0.11 kts)
Livorno	0.70	0.75	0.05	0.00042 (0.15 kts)
Napoli	0.67	0.79	0.12	0.00050 (0.18 kts)
Piraeus	0.84	0.88	0.04	0.00100 (0.36 kts)
Rijeka	0.50	0.61	0.11	0.00061 (0.22 kts)
Syros	0.72	0.84	0.11	0.00022 (0.08 kts)
Tanger	0.47	0.46	0.03	0.00080 (0.29 kts)
Valencia	0.94	0.98	0.03	0.00050 (0.18 kts)
Valetta	0.67	0.83	0.15	0.00022 (0.08 kts)

1.5 **Conclusions**

This paper implements the DBSCAN algorithm to automatically recognize barge-to-ship bunkering operations using high-frequency spatial data on ship positions from AIS. The algorithm is shown to reliably identify stopped ocean-going vessels, tanker vessels operating in local port waters as bunker barges, and a rendezvous between a bunker barge and a larger vessel. The algorithm is capable of detecting the irregular shape of vessel movement at anchor and can separate between bunkering performed at anchor and alongside a terminal during cargo operations. The framework is used to generate micro-level statistics for bunkering operations. The results include information such as the ranking of bunkering ports by activity and vessel type, bunker barge age distribution, waiting times and servicing times. A comparison with public data on the number of bunkering operations in the port of Gibraltar shows that the algorithm can

detect a high percentage of operations. The quality and accuracy of the results are highly dependent on the quality of the input data and port polygons.

Shipping practitioners can benefit from the algorithm, as it is customizable to filter bunkering operations by specific vessels, vessel segments, geographical regions, or individual ports. Moreover, the results in this paper provide bunker suppliers with the framework to assess their position in intra- and inter-port competition. Port operators can judge if entering the bunker market and become a full-service maritime hub is beneficial versus focusing on cargo handling only. Finally, ship-owners can make more informed bunkering decisions by accounting for both waiting and servicing times as well as fuel prices and deviation.

The results generated under this paper framework are aligned to the conclusions presented by Acosta et al. (2011) and Vilhelmsen et al. (2014), regarding bunkering competition and bunkering choice determinants. More importantly, the method helps to overcome the difficulty of generating bunkering information via qualitative methods in a standardized way that could be extended to different geographical regions.

The services and waiting times, as generated in this paper, will be important input in the bunkering location optimization problems. Services times introduced from arbitrary distributions such as in the model of Aydin et al. (2017) proved to be a major factor in the selection of optimal bunkering ports for liner ships. Therefore, new research could extend the findings to the tramp case by exploring the importance of real service and waiting times on their selection of bunkering ports.

Future research should extend this analysis of bunkering operations to a global scale, and investigate the interaction between bunkering competition and fuel prices. Importantly, the method described in this paper can help in discovering new activities via AIS, such as transit through Canals from the detection of anchoring patterns before transits. There is also room for improvement in the calibration of the settings of the algorithm, such as the use of real weather conditions to assess suitable distance thresholds for ships at anchor and the dynamic selection of parameters per bunkering port.

References

Acosta, M., Coronado, D., & Cerban, M. D. M. (2011). Bunkering competition and competitiveness at the ports of the Gibraltar Strait. *Journal of Transport Geography, 19*(4), 911-916. https://doi.org/10.1016/j.jtrangeo.2010.11.008

Adland, R., Jia, H., & Strandenes, S. P. (2017). Are AIS-based trade volume estimates reliable? The case of crude oil exports. *Maritime Policy & Management, 44*(5), 657-665. https://doi.org/10.1080/03088839.2017.1309470

Ali, T., Ashgar, S., & Sajid, N. A. (2010). Critical analysis of DBSCAN variations. *2010 International Conference on Information and Emerging Technologies*, (pp. 1-6). IEEE. https://doi.org/10.1109/ICIET.2010.5625720

Aronietis, R., Sys, C., Van Hassel, E., & Vanelslander, T. (2017). Investigating the bunkering choice determinants: the case of the port of Antwerp. *Journal of Shipping and Trade,2*(1),8. https://doi.org/10.1186/s41072-017-0025-7

Aydin, N., Lee, H., & Mansouri, S. A. (2017). Speed optimization and bunkering in liner shipping in the presence of uncertain service times and time windows at ports. *European Journal of Operational Research, 259*(1), 143-154. https://doi.org/10.1016/j.ejor.2016.10.002

Besbes, O., & Savin, S. (2009). Going bunkers: The joint route selection and refueling problem. *Manufacturing & Service Operations Management, 11*(4), 694-711. https://doi.org/10.1287/msom.1080.0249

Cazzanti, L., & Pallotta, G. (2015). Mining Maritime Vessel Traffic: Promises, challenges, techniques. *OCEANS 2015-Genova*, (pp. 1-6). IEEE. https://doi.org/10.1109/oceans-genova.2015.7271555

Dobrkovic, A., Iacob, M. E., & van Hillegersberg, J. (2018). Maritime pattern extraction and route reconstruction from incomplete AIS data. *International Journal of Data science and Analytics, 5*, 111-136. https://doi.org/10.1007/s41060-017-0092-8

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Proceedings of the*

Second International Conference on Knowledge Discovery and Data Mining, 226-231.

Ford, M. C. (2012). *A Masters Guide to: Using Fuel Oil Onboard Ships*. Retrieved from Charles Taylor &Co Limited, The Standard and the American Bureau of Shipping: https://www.standard-club.com/media/24163/AMastersGuidetoUsingFuelOilOnboardships.pdf

Fuentes, G., & Adland, R. (2020). A Spatial Framework for Extracting Suez Canal Transit Information from AIS. *2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 586-590. IEEE. https://doi.org/10.1109/ieem45057.2020.9309882

Gibraltar Port Authority. (2019, March 10). *Gibraltar Port Authority activity 2013-2019*. Retrieved from HM Government of Gibraltar Port Statistics: https://www.gibraltar.gov.gi/statistics/statistics-topic-area/2019/port

Goldsworthy, L., & Goldsworthy, B. (2015). Modelling of ship engine exhaust emissions in ports and extensive coastal waters based on terrestrial AIS data - An Australian case study. *Environmental Modelling & Software, 63*, 45-60. https://doi.org/10.1016/j.envsoft.2014.09.009

Hou, J. Gao, H., & Li, X. (2016). DSets-DBSCAN: A parameter-free clustering algorithm. *IEEE Transactions on Image Processing, 25(7),* 3182-3193. IEEE. https://doi.org/ 10.1109/TIP.2016.2559803

Jia, H., Lampe, O. D., Solteszova, V., & Strandenes, S. P. (2017). An automatic algorithm for generating seaborne transport pattern maps based on AIS. *Maritime Economics & Logistics, 19*(4), 619-630. https://doi.org/10.1057/s41278-017-0075-7

Johansson, L., Jalkanen, J. P., & Kukkonen, J. (2017). Global assesment of shipping emissions in 2015 on a high spatial and temporal resolution. *Atmospheric Environment, 167*, 403-415. https://doi.org/10.1016/j.atmosenv.2017.08.042

Karami, A., Johansson, R. (2014). Choosing DBSCAN parameters automatically using differential evolution. *International Journal of Computer Applications, 91(7),* 1-11. https://doi.org/10.5120/15890-5059

Khan, K., Rehman, S. U., Aziz, K., Fong, S., & Sarasvady, S. (2014). DBSCAN: Past, present and future. *The fifth international conference on the applications of digital information and web technologies (ICADIWT 2014)*, 232-238. IEEE. https://doi.org/10.1109/icadiwt.2014.6814687

Kim, J., Song, J., No, H., Han, D., Kim, D., Park, B., & Kee, C. (2017). Accuracy improvement of DGPS for low-cost single-frequency receiver using modified Flächen Korretur parameter correction. *International Journal of Geo-Information, 6*(7), 222. https://doi.org/10.3390/ijgi6070222

Lam, J. S. L., Chen, D., Cheng, F., & Wong, K. (2011). Assessment of the competitiveness of ports as bunkering hubs: Empirical studies of Singapore and Shanghai. *Transportation journal, 50*(2), 176-203. https://doi.org/10.5325/transportationj.50.2.0176

Lei, P. R. (2015). A framework for anomaly detection in maritime trajectory behavior. *Knowledge and Information Systems*, *47*(1), 189-214. https://doi.org/10.1007/s10115-015-0845-4

Man, D., Uda, K., Ito, Y., & Nakano, K. (2011). A GPU implementation of computing euclidean distance map with efficient memory access. *2011 Second International Conference on Networking and Computing*, (pp. 68-76). IEEE. https://doi.org/10.1109/ICNC.2011.19

Merchant, N. D., Witt, M. J., Blondel, P., Godley, B. J., & Smith, G. H. (2012). Assesing sound exposure from shipping in coastal waters using a single hydrophone and Automatic Identification System (AIS) data. *Marine Pollution Bulletin, 64*(7), 1320-1329. https://doi.org/10.1016/j.marpolbul.2012.05.004

Mou, J. M., Van Der Tak, C., & Ligteringen, H. (2010). Study on collision avoidance in busy waterways by using AIS data. *Ocean Engineering, 37*(5-6), 483-490. https://doi.org/10.1016/j.oceaneng.2010.01.012

National Centers for Environmental Information. National Oceanic and Atmospheric Administration. (2017). *Global Self-consistent, Hierarchical, High-resolution Geography Database*. Retrieved January 31, 2020, from https://www.ngdc.noaa.gov/mgg/shorelines/data/gshhg/latest/

National Coordination Office for Space-Based Positioning. (2020, 01 31). *GPS.gov*. Retrieved from GPS Accuracy: https://www.gps.gov/systems/gps/performance/accuracy/

Pallotta, G., Vespe, M., & Bryan, K. (2013). Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction. *Entropy, 15*(6), 2218-2245. https://doi.org/10.3390/e15062218

Poulain, P. M., Menna, M., & Gerin, R. (2018). Mapping Mediterranean tidal currents with surface drifters. *Deep Sea Research Part I: Oceanographic Research Papers, 138*, 22-33. https://doi.org/10.1016/j.dsr.2018.07.011

Prochazka, V., & Adland, R. (2019). Ocean mesh grid: Applications in shipping modeling. *2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 330-334. IEEE. https://doi.org/10.1109/IEEM44572.2019.8978936

Prochazka, V., Adland, R., & Wolff, F. C. (2019). Contracting decisions in the crude oil transportation market: Evidence from fixtures matched with AIS data. *Transportation Research Part A: Policy and Practice, 130*, 37-53. https://doi.org/10.1016/j.tra.2019.09.009

Rand, W. M. (1971). Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association, 66(336),* 846-850. https://doi.org/10.2307/2284239

Ronen, D. (1982). The effect of oil price on the optimal speed of ships. *Journal of the Operational Research Society, 33*(11), 1025-1040. https://doi.org/10.2307/2581518

Schubert, E., Sander, J., Ester, M., Kriegel, H. P., & Xu, X. (2017). DBSCAN revisited, revisited: why and how you should (still) use DBSCAN. *ACM Transactions on Database Systems (TODS)*, *42*(3), 1-21. https://doi.org/10.1145/3068335

Sheng, P., & Yin, J. (2018). Extracting shipping route patterns by trajectory clustering model based on Automatic Identification System data. *Sustainability, 10*(7), 2327. https://doi.org/10.3390/su10072327

Sheng, X., Chew, E. P., & Lee, L. H. (2015). (s, S) policy model for liner shipping refueling and sailing speed optimization problem. *Transportation Research Part E: Logistics and Transportation Review, 76*, 76-92. https://doi.org/10.1016/j.tre.2014.12.001

Tichavska, M., & Tovar, B. (2015). Port-city exhaust emission model: An application to cruise and ferry operations in Las Palmas port. *Transportation Research Part A: Policy and Practice, 78*, 347-360. https://doi.org/10.1016/j.tra.2015.05.021

Vilhelmsen, C., Lusby, R., & Larsen, J. (2014). Tramp ship routing and scheduling with integrated bunker optimization. *EURO Journal on Transportation and Logistics, 3*(2), 143-175. https://doi.org/10.1007/s13676-013-0039-8

Wang, S., & Meng, Q. (2012). Sailing speed optimization for container ships in a liner shipping network. *Transportation Research Part E: Logistics and Transportation Review, 48*(3), 701-712. https://doi.org/10.1016/j.tre.2011.12.003

Yao, Z., Ng, S. H., & Lee, L. H. (2012). A study on bunker fuel management for the shipping liner services. *Computers & Operations Research, 39*(5), 1160-1172. https://doi.org/10.1016/j.cor.2011.07.012

Zhang, L., Meng, Q., & Fwa, T. F. (2019). Big AIS data based spatial-temporal analyses of ship traffic in Singapore port waters. *Transportation Research Part E: Logistics and Transportation Review, 129*, 287-304. https://doi.org/10.1016/j.tre.2017.07.011

Zhang, W., Goerlandt, F., Kujala, P., & Wang, Y. (2016). An advanced method for detecting possible near miss ship collisions from AIS data. *Ocean Engineering, 124*, 141-156. https://doi.org/10.1016/j.oceaneng.2016.07.059

Appendices

Appendix A

Algorithms for bunkering operation recognition

Algorithm 1 Functions for Algorithm

function POLYGONRADIUS(coordinate, distance)
    Generate Polygon out of coordinate point
    return polygon with radius of distance
end function
function TIMERANGE(timestamp1,timestamp2)
    Calculate the difference between timestamps
    return hours range
end function
function SEQUENCETEST(iterator)
    Detect separate calls in same port. Test if sequence exist and assign unique number to sequence of True values between False values
    return set of unique sequence group values
end function
function TIME-POSITIONINTERPOLATION(Dataframe)
    Detects discontinuity in the expected sequence of timestamps and interpolates the empty positions from the previous and next available positions
    Parameters:
    timerange←10. Minutes, starting from the first position in the dataframe
    return adjusted dataframe with interpolated positions
end function
function DBSCAN(positions iterator)
    Creates clusters based on DBSCAN. -1 is assigned to noise positions and a unique number to every cluster
    Parameters:
    minPts←3.Minimum number of positions to generate a cluster
    ϵ←0.00195. Decimal degrees maximum distance. From 0.68 kts upper bound or 0.117NM allowance for 10 minutes movement.
    return set of either unique clusterID or noise value
end function
function VESSELPOLYGONGENERATOR(positions iterator)
    Generates linestrings from the sequence of ordered positions (lon,lat). Then creates a polygon as the radius buffered from the linestring
    Parameters:
    radius←0.00045. Decimal degrees, Radius of 50 meters away from linestring
    return ocean-going vessel polygon
end function
function MERGESERVICE(dataframe)
    Shift method that calculates the time difference between the last timestamp of a service ID and the first timestamp of the next service ID. If the difference is less than a threshold then the services merges and the first group service ID is absorbed by the following group
    Parameters:
    time←24. Hours, considers a barge refueling cargo in the middle of the bunkering operation. The refueling of the barge keeps counting as service as the vessel has to wait.
    return dataframe with merged services
end function
function MERGEDATAFRAMES(dataframe1,dataframe2,leftIndex,rightIndex)
    Merges dataframes by leftIndex=rigtIndex
    Parameters:
    return merged dataframe with algorithm results and vessels characteristics. Named as dataframe1.
end function

Algorithm 2 Bunker barge prospects recognition

Inputs:
A dataframe with columns [portName,coordinate,berthPolygon], Ports;
A dataframe with columns [vesselsID,LengthOverAll,TypeOfVessel], VesselsSpecs;
A dataframe with columns [vesselID,positions,time]; AIS
An empty dataframe with columns [bargeID,portOfService]; BargesInfo
Default Parameters:
loa ←150, meters of length overall of ship
localThreshold ←0.70, minimum stay at local waters
minimumObs ←100, minimum observations to have a valid sample
set polygonGeometry as new column in Ports
for coordinate in Ports **do**
    polygon ← **call** POLYGONRADIUS(coordinate, polygon)
    store polygon to Ports.polygonGeometry
end for
for polygon in Ports.polygonGeometry **do**
    for testPolygon in Ports.polygonGeometry **do**
        if polygon overlaps testPolygon **then**
            newPolygon ← **merge** polygon and testPolygon
            remove polygon and testPolygon
            store newPolygon to Ports.polygonGeometry
        end if
    end for
end for
set inLocalWaters as new column in AIS
for vesselID in VesselsSpecs **do**
    if vesselID in VesselsSpecs.LengthOverAll $\leq$ loa AND VesselsSpecs.TypeOfVessel=tanker **then**
        filter position of vesselID in AIS
        for polygon in Ports **do**
            for position in AIS **do**
                bool,polygonName ← **call** POINTINPOLYGON(position,polygonGeometry)
                if bool = TRUE **then**
                    store [TRUE to AIS.inLocalWaters ; polygonName to AIS.portVisited]
                else
                    store [FALSE to AIS.inLocalWaters ; NULL to AIS.portVisited]
                end if
            end for
            totalRows ← **count** total rows in AIS
            filter positions in AIS WHERE AIS.inLocalWaters = TRUE
            totalFilteredRows ← **count** total filtered rows in AIS
            if totalRows$\geq$ minimumObs AND totalFilteredRows/totalRows $\geq$ localThreshold **then**
                store [vesselID in BargeInfo.bargeID ; AIS.portVisited in BargeInfo.portOfService]
            end if
        end for
    end if
end for
return BargesInfo

Algorithm 3 Stopped vessel

Inputs:
A dataframe with columns [portName,coordinate,polygonGeometry,berthPolygon], Ports;
A dataframe with columns [vesselID,positions,time,speed], AIS;
An empty dataframe with columns[vesselID, initialStopTime, finalStopTime, atBerth, callID, clusterID, stopPosition, portVisited, stopID], Stops;
A set of
Default Parameters:
maxSpeed ← 3, knots of vessel entering or leaving cluster
minStay ← 1, hour. Minimum stay to be classified as stopped.
laidUp ← 1084, hours. Laid-up vessel with more than 45 days in port
noisePts ← -1. Assigment by DBSCAN to noise positions(not clustered positions)
listOfVessels← set of unique values of AIS.vesselID
listOfCallID ← empty set
listOfClusterID ← empty set
set inLocalWaters, callID, clusterID, portVisited as new columns in AIS
set counter to 0
for uniqueVesselID in listOfVessels **do**
  filter AIS WHERE AIS.vesselID=uniqueVesselID. Sort ascending by AIS.time
  for polygonGeometry in Ports **do**
    for positions in AIS **do**
      bool, polygonName ← **call** POINTINPOLYGON(positions,polygonGeometry)
      if bool=TRUE **then**
        store [TRUE in AIS.inLocalWaters ; polygonName in AIS.portVisited]
      else
        store [FALSE in AIS.inLocalWaters ; NULL in AIS.portVisited]
      end if
    end for
  end for
  store [**call** SEQUENCEGENERATOR(AIS.inLocalWaters)] in AIS.callID
  filter AIS WHERE AIS.inLocalWaters=TRUE AND AIS.speed≤maxSpeed
  store unique values of AIS.callID in listOfCallID
  for uniqueCallID in listOfCallID **do**
    filter AIS WHERE AIS.callID=uniqueCallID
    timeDifference ← **call** TIMERANGE(AIS.time[-1], AIS.time[0])
    if timeDifference≥minStay AND timeDifference≤laidUp **then**
      call TIME-POSITIONINTERPOLATION(AIS)
      store [**call** DBSCAN(AIS.positions)] in AIS.clusterID
      filter AIS WHERE AIS.clusterID≠noisePts
      store unique values of AIS.clusterID in listOfClusterID
      for uniqueClusterID in listOfClusterID **do**
        add 1 to counter
        filter AIS WHERE AIS.clusterID=uniqueClusterID
        store [uniqueVesselID in Stops.vesselID ; AIS.time[0] in Stops.initialStopTime ; AIS.time[-1] in Stops.finalStopTime ; uniqueClusterID in Stops.clusterID; uniqueCallID in Stops.callID ; AIS.portVisited[0] in Stops.portVisited ; uniqueCallID in Stops.callID ; AIS.positions[0] in Stops.stopPosition, counter in Stops.stopID]
        for berthPolygon in Ports **do**
          bool, ← **call** POINTINPOLYGON(AIS.positions[0],berthPolygon)
          if bool=TRUE **then**
            store TRUE in Stops.atBerth
          else
            store FALSE in Stops.atBerth
          end if
        end for
      end for
    end if
  end for
end for
return Stops

Algorithm 4 Bunkering recognition

Inputs:
A dataframe with columns [vesselID,positions,time], AIS;
A dataframe with columns[vesselID, initialStopTime, finalStopTime, atBerth, callID, clusterID, stopPosition, portVisited, stopID], Stops;
A dataframe with columns[bargeID,portOfService]; BargesInfo;
A dataframe with columns [vesselID,grossTonnage,TypeOfVessel,yearOfBuilt], VesselsSpecs;
An empty dataframe with columns [vesselID,positions,time], VesselAIS;
An empty dataframe with columns [vesselID,positions,time, alongsideVessel, serviceID], BargeAIS;
An empty dataframe with columns [bargeID, initialServiceTime, finalServiceTime, vesselID, initialStopTime, finalStopTime, atBerth, callID, clusterID, stopPosition, portVisited, stopID], Results;
Default Parameters:
minStay ← 1, hour. Minimum stay of barge inside ocean-going vessel polygon
listOfStops ← set of values of AIS.stopID
listOfBarges ← empty set
listOfServices ← empty set
for uniqueStop in listOfStops **do**
  filter Stops.stopID=uniqueStop
  VesselAIS←**filter** AIS.vesselID=Stops.vesselID AND AIS.time≥Stops.initialStopTime AND AIS.time≤Stops.finalStopTime
  vesselPolygon ← **call** VESSELPOLYGONGENERATOR(VesselAIS.positions)
  filter BargesInfo.portOfService=Stops.portVisited
  store unique values of BargesInfo.bargeID in listOfBarges
  for uniqueBargeID in listOfBarges **do**
    BargeAIS←**filter** AIS.vesselID=uniqueBargeID AND AIS.time≥Stops.initialStopTime AND AIS.time≤Stops.finalStopTime
    for positions in BargeAIS.positions **do**
      bool, ← **call** POINTINPOLYGON(positions,vesselPolygon)
      if bool=TRUE **then**
        store TRUE in BargeAIS.alongsideVessel
      else
        store FALSE in BargeAIS.alongsideVessel
      end if
    end for
    store [**call** SEQUENCEGENERATOR(BargeAIS.alongsideVessel)] in BargeAIS.serviceID
    filter BargeAIS WHERE BargeAIS.alongsideVessel=TRUE
    call MERGESERVICE(BargeAIS)
    store unique values of BargeAIS.serviceID in listOfServices
    for service in listOfServices **do**
      filter BargeAIS.serviceID=service
      timeDifference ← **call** TIMERANGE(BargeAIS.time[-1], BargeAIS.time[0])
      if timeDifference≥minStay AND BargeAIS.vesselID≠Stops.vesselID **then**
        store [BargeAIS.vesselID in Results.bargeID ; BargeAIS.time[0] in Results.initialServiceTime ; Barge AIS.time[-1] in Results.finalServiceTime ; Stops.vesselID in Results.vesselID ; Stops.initialStopTime in Results.initialStopTime ; Stops.finalStopTime in Results.finalStopTime ; Stops.atBerth in Results.atBerth ; Stops.callID in Results.callID ; Stops.clusterID in Results.clusterID ; Stops.stopPosition in Results.stopPosition ; Stops.portVisited in Results.portVisited ; uniqueStop in Results.stopID]
      end if
    end for
  end for
end for
call MERGEDATAFRAME(Results,VesselsSpecs,vesselID,vesselID)
call MERGEDATAFRAME(Results,VesselsSpecs,bargeID,vesselID)
return Results

Appendix B

Table B.1

Correlation matrix for algorithm derived variables

	Service time (hours)	Operations	Waiting time anchor (hours)	Waiting time berth (hours)	Available barges
Service time (hours)	1	0.09	-0.01	0,05	0.08
Operations	0.09	1	-0.09	0.02	0.30
Waiting time anchor (hours)	-0.01	-0.09	1	-0.01	0.07
Waiting time berth (hours)	0.05	-0.02	-0.01	1	-0.02
Available barges	0.08	0.30	0.07	-0.02	1

Chapter 2

The effects of waiting times on the bunkering decision for tramp ships*

Gabriel Fuentes [a,b], Roar Adland [b], Stein W. Wallace [b]

[a] Centre for Applied Research at NHH(SNF)
[b] Department of Business and Management Science, NHH

Abstract

This study explores the influence of waiting times, both deterministic and stochastic, together with uncertain fuel price in the selection of bunker fuel stops for a ship-owner engaged in tramp shipping. We introduce stochastic bunkering waiting times in the bunkering optimization problem using scenarios constructed from vessels' Automatic Information System (AIS) records and test their relevance to the optimal decision. We demonstrate that choosing a bunkering port solely based on the level of fuel prices can lead to suboptimal decisions. From our empirical results based on a case study of the Mediterranean Sea, Gibraltar and Valetta are the preferred bunkering options based on their relatively low fuel prices and short waiting times.

Keywords: bunker management, tramp shipping, stochastic programming, bunker port efficiency

2.1 Introduction

In this study, we investigate the bunker (fuel) management problem for tramp shipping in an optimization model that integrates both waiting times (deterministic and

* This research was partly financed by the Research Council of Norway under the project "Smart digital contracts and commercial management", project number 280684.

stochastic) and stochastic fuel prices. We focus on a vessel trading in a tramp shipping setting under a voyage charter (contract). This implies that the shipowner pays the voyage costs, including the bunker costs. Therefore, our study focuses on the shipowner's position.

For tramp shipping, reducing bunker costs is crucial as it might represent up to 70% of the ship's operating cost (Rehmatulla & Smith, 2015). Better bunker management is one way for ship operators to maximize their voyage profit. According to Yao et al. (2012), bunker management refers to three interrelated components: Selecting a bunkering port, deciding on the bunker volume to order, and selecting the vessel's speed. Based on these elements, the vessel's officers can prepare an efficient voyage plan. A voyage plan provides all details of the safest and most efficient route, including any deviation to a bunkering port, to fulfill the contracted voyage instructions.

For a tramp vessel under a voyage contract, choosing where to bunker is a complex problem as the decision must balance the tradeoff between the local fuel price, deviation from the shortest route and the time spent in each potential bunkering port. Based on the wording in standard chartering contracts, the vessel is also legally limited to take on bunkers only on the ballast (empty) voyage leg. This tradeoff weighs the value of savings from cheaper bunkering ports versus the value of time resulting from the deviation to the bunkering port, the time spent there, the route taken, and the speed selection. In addition, voyage and vessel specific factors (e.g. vessel draft, consumption, speed, and remaining voyage distance) must be considered.

The assessment of waiting times is relevant, as an owner might take on bunkers in a relatively faster bunkering port at the expense of paying a higher fuel price. Thus, the bunkering management decision could help to avoid cheap but inefficient ports that would subsequently cause the vessel to speed up in order to pick up the cargo within the contracted window (laycan[1]). Such an increase in speed would increase fuel consumption and emissions. In other words, the cost of the additional fuel consumption could outweigh the apparent savings of buying fuel in a cheap but inefficient port. In

[1] Laycan is short for Lay Days Cancellation Days. It is a time window agreed by the ship owner and the charterer at which the vessel must arrive and be ready to carry cargo operations in a port. Failure to reach this window gives the charterer the right to cancel the contract.

our study, we refer to an inefficient bunker port as one where large deviations, long waiting times and/or slow bunker barge pumping rates are observed.

The bunker management problem can be considered using either an operational and/or a tactical approach. The tactical approach is about medium- and long-term decisions that help in controlling uncertainties (e.g., bunker contracts and fleet deployment). Conversely, the operational strategy uses short-term decisions based on expected outcomes or the observed voyage evolution (e.g., speed and route adjustment). Gu et al. (2019) showed that a higher performance can be obtained using a mix of both strategies in the liner case. Here, we follow a similar approach where an owner can choose to bunker in a port using either a forward (tactical) or spot (operational) contract or a combination of both.

Pursuing a pure operational strategy implies that bunkering decisions are based on the spot fuel market. A complicating factor in this setting is that fuel prices demonstrate large differences across ports within a relatively close distance [see Yao et al. (2012) for a comparison]. Furthermore, as fuel oil is derived from crude oil, it inherits its global price volatility and is, additionally, exposed to local supply and demand dynamics. As the expected fuel price for a voyage is considered during freight rate negotiations, a realized large difference from the expected price could cause a substantial loss for a spot market strategy.

To reduce such risks, owners can implement operational strategies to reduce fuel consumption, such as slow steaming, or reduce fuel price risk using financial hedging instruments such as oil futures or bunker swaps. A cautious owner would request all bunkers on a forward basis, even though she would also lose the opportunity of benefiting if lower prices materialize during the voyage. A halfway measure would be to decide that some of the required volume of bunkers, and the port in which to bunker, need to be nominated at the time of entering a new freight contract. This leaves the possibility of benefiting from future price drops and flexibility to deal with operational delays.

We formulate the problem as a three-stage stochastic optimization model. In the first stage, the shipowner is negotiating a voyage and is also deciding where to bunker, the speed and route to take based on expected weather, and how much to hedge via forward contracts with known prices. This decision must acknowledge the second

stage, where the price of spot fuel prices is revealed and a spot bunker contract is secured. In stage three, after the waiting time is revealed at the bunkering port, an owner can have a recourse on the second stage decision and decide how much spot bunker to take onboard, paying a penalty for any unused volume and/or a penalty for late arrival in the ultimate loading ports. The late arrival penalty reflects the market state where, in a low freight market, time is perceived as less valuable than in a high freight market.

In this study, we introduce a framework to solve the bunkering management problem in a tramp shipping environment. The framework is used in a case study which incorporates real data from fuel prices and waiting time estimations from Mediterranean bunkering ports. Our results and framework are relevant to ships operating under a voyage contract, as it offers the potential for more efficient voyage plans and emissions reduction. By using this model, we also analyze the effect of ports' waiting times in the bunkering decision.

The remainder of this paper is structured as follows. Section 2.2 contains the literature review. Section 2.3 describes our interpretation of the bunkering management problem, and Section 2.4 introduces our mathematical model and related assumptions. Section 2.5 introduces the details of our empirical case study. Section 2.6 presents the results and discusses their implications. Finally, Section 2.7 concludes and identifies prospects for future extensions.

2.2 Literature review

Bunker management strategies have been proposed in the literature as part of an increasing number of maritime operations research papers. The main division is between liner vessel which operate with fixed schedules and routes, and tramp vessels operating on an on-demand basis.

Most of the literature proposes solutions to liner services, either for a single vessel or a fleet of vessels. Although the goal of reducing bunker costs is shared, these studies enrich the literature by a myriad of methods, modeling strategies or small objective additions (e.g. emissions reductions), some of which also apply to the tramp case. In a single liner vessel context, Yao et al. (2012) introduced empirical fuel consumption functions and concluded that the optimal fuel management decisions are strongly

influenced by the assumed relationship between speed and consumption. Kim (2014) introduced a carbon tax on vessel emissions, thus including environmental objectives as part of a bunkering problem. Sheng et al. (2014) used a multistage scenario tree to model the problem with stochastic consumption and prices. Additionally, they adopted a rolling horizon strategy to reduce the size of the scenario tree. Sheng et al. (2015) gave a comprehensive review of the general inventory management problem and compared it to the bunker management problem. Moreover, they modeled the bunker prices as a one-step Markovian process and the weather affecting the consumption with a normal truncated distribution, finding that the weather effect on consumption is significant in the bunkering decision. Ghosh et al. (2015) proposed a framework where vessels are supplied with fuel from both long-term contracts and the spot market. Aydin et al. (2017) modeled uncertain cargo handling times for liner ships in terminals and demonstrated that speed decisions are influenced by the time in port, a feature that is comparable to the bunkering waiting time in the tramp case. Wang. et al. (2018) modeled stochastic fuel prices with a distribution-free method, using the mean, variance, upper and lower bound of historical fuel prices. Gu et al. (2019) modeled risk preferences by introducing a Conditional Value at Risk (CVaR) constraint.

The fleet case merges the tactical decision of vessel deployment with the bunkering policies of vessels belonging to a fleet liner service. Wang and Meng (2015) developed such a case by modeling a robust bunker management strategy under the worst-case bunker consumption scenario and introduced uncertain speed. Wang and Chen (2017) modeled the problem as non-linear, without the piecewise linearization of the consumption function as is commonly observed in other papers. However, the problem was simplified by avoiding price stochasticity. Wang et al. (2018) deviated from the traditional setting and develop a tactical model of bunker contract swaps as a two-stage stochastic problem. The scenario tree to model the fuel prices was generated by an adjusted moment matching method from a distribution built from an mGARCH model. Liu et al. (2020) used a two-stage stochastic problem with recourse and included the uncertain container demand of a liner fleet. They approached the solution by both a Sample Average Approximation method (SAA) and an L-Shaped method, with the L-Shaped method outperforming the SAA.

Liner vessels have the advantage–modeling wise–of having voyage plans defined in advance, removing the added complexity of deviations, and vastly reducing the

number of bunkering port options. The assumption taken by most liner vessel papers is that they can only be bunkered while in a cargo port that is a part of their schedule. Hence, as long as there are no unusual delays, we can surmise that the bunkering service time is not important in their case. Having to deviate for bunkers reinforces the importance of bunkering waiting times as a feature that influences the vessel speed and fuel consumption.

Given that the bunker consumption is linked to the total voyage distance and voyage speed, it is natural to model the routing and bunkering problems together (Meng, et al., 2015). Besbes and Savin (2009) modeled the bunker management problem and route selection for both the liner and tramp cases, recognizing the higher complexity in the tramp case (also noted by Fan et al. (2019)). They formulated the problem to be solved by heuristics of stochastic dynamic programming, with bunker prices generated by a combination of the forecast of crude oil prices based on a first-order autoregressive process, a global price adjustment factor, and a local correction factor. The results for the liner case suggest a “buy up to” policy where bunker quantity is requested based on the remaining fuel on board and the fuel price while stopped at scheduled ports. The tramp setting considered the bunker refueling decision and its interaction with the routing selection. They assumed an instantaneous bunkering time, fixed speed, and fixed consumption to avoid their influence in the model and to focus on assessing the influence of stochasticity in bunker prices on the selection of an optimal routing. They proved that, even in the absence of price dispersion, a modest degree of stochasticity in fuel prices causes a substantial degree of complexity in the optimal routing decisions.

Oh and Karimi (2010) proposed speed as a decision variable in a mixed integer programming model for bunkering and routing decisions of multi-parcel tankers. The speed decision is affected by the consumption and the relation was modeled with discrete consumption rates for the different operations of a tanker vessel (e.g., steaming, cleaning tanks, loading and unloading cargoes, and a base consumption when it is not moving). In their case, the way multiparcel tankers operate permits modelling the tramp case as if it was a liner case with fixed routes, and no consideration is given to the operational features of the voyage (e.g., water depth, weather). To model the fuel prices, they used arbitrary equally probable scenarios and prices derived from historical data.

Vilhelmsen (2014) proposed a joint model with the scheduling, routing, and bunkering optimization of tramp vessels. The author introduced reformulation of bunker consumption based on cargo load and speed. To model the fuel prices, she made use of forward curves to have a full overview of the prices in subsequent ports. Here, the fuel consumption rate is not associated with a vessel speed function but depends simply on the vessel either being laden or in ballast. She also assumed a fixed bunkering time at every port. Hence, the service level's influence within the model is not analyzed.

While Vilhelmsen allowed deviation to bunker ports, Meng, et al. (2015) only considered bunkering at ports scheduled with cargo. However, the empirical work of Fuentes (2021) shows that deviations to perform pure bunkering operations are commonly observed in tramp shipping.

From this review, we can recognize at least four gaps in the literature. Firstly, no model in the tramp case has considered the influence of bunkering waiting times on the bunker management problem, despite the demonstrated influence of cargo ports' servicing time on the speed and consumption of a liner vessel. Secondly, important features of the voyage such as draft restrictions, the influence of weather on fuel consumption, and the relationship between the ordered fuel quantity and servicing times and waiting times have not been explicitly modeled. Thirdly, the literature has not analyzed the problem subject to the full set of standard contractual constraints for the tramp case (e.g., voyage charter party clauses regarding laycan and cargo intake optionality; bunker contracts; loading area, etc.). Finally, there are no studies which evaluate risk measures, such as CVaR, in a multistage tramp vessel bunkering model.

To fill these gaps, the contribution of this paper is fourfold. Firstly, we model the bunkering management problem which incorporates stochastic bunkering waiting times and stochastic fuel prices. The waiting time scenarios are modeled from observed empirical distributions in bunkering ports in the Mediterranean Sea. Secondly, we introduce draft dynamics as a novel feature of the bunkering management problem and estimate the vessel consumption for the expected weather. Draft changes throughout the voyage because of fuel being gradually consumed are modeled by the tons per centimeter of immersion (TPC) function of a real vessel. Thirdly, we take assumptions regarding appropriate constraints from a standard

voyage charter. Finally, we evaluate the effect of risk aversion per stage in the shipowner's decision.

2.3 Problem description

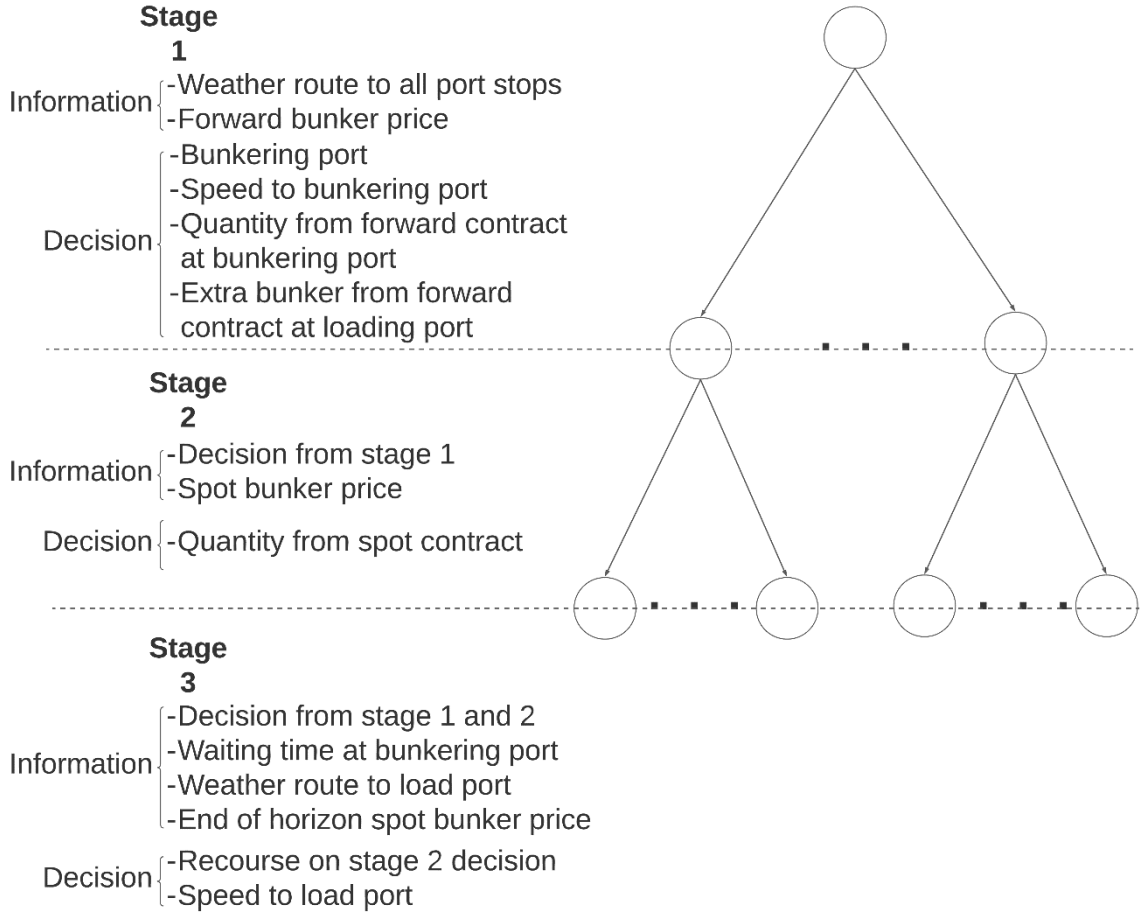
We consider the bunkering management problem of a tramp vessel under a voyage charter from the position of the shipowner. Under such freight contracts, the owner is responsible for purchasing bunkers and covering all related voyage costs. Additionally, the vessel's captain (subject to the ship-owner's instructions) must decide on the voyage route and the speed to reach the laycan as agreed in the charter party.

The problem is formulated in a three-stage stochastic programming model as described in Figure 1. At stage one, the owner in liaison with the vessel's captain, is to decide where to bunker, instruct on the vessel speed to the bunkering stop, and nominate the quantity of fuel to fix on a forward contract. The decision is taken subject to uncertain spot prices (that are revealed as the vessel approaches the bunker port) and uncertain waiting times for the bunkering operation. The vessel has the option to request fuel from a forward contract either at an intermediate bunkering port, directly from the loading port or as a combination of both. Fuel from the forward contract shall be received in full. In stage two, based on the consequences of the first stage decision, the owner must nominate how much spot bunkers to take on at the selected port given the revealed spot price. In stage three, the waiting time at the bunkering port is revealed and the owner can adjust on the quantity to receive from the spot bunker contract nomination, thus adjusting the service time for loading bunkers, and pay a penalty for the fuel not purchased. This recourse represents an adjustment to the waiting time in the bunkering port, leaving a buffer to potentially reduce the subsequent speed and still make the loading port laycan. Finally, the captain decides whether to take additional fuel at the loading port, while considering cargo quantity optionality[2] and vessel draft limitations. Note that the end of horizon (EOH) price at stage 3 has no consequence for the stage 3 decision but is introduced to condition the first and second stage orders.

[2] In a dry bulk voyage charter, it is common to include the option for owners to nominate more or less than the agreed cargo quantity for the voyage (usually +/- 5% or 10%). This gives flexibility to the captain to adjust the cargo intake based on the vessel's navigational or stability considerations. As the freight revenue is paid as a function of cargo loaded, owners have the incentive to load as much as possible.

Figure 1

Problem interpretation



2.4 Modeling

2.4.1 Assumptions

We assume that every fuel purchase contract is drafted following BIMCO's bunkering terms 2018. This means that the quotation, once accepted, is binding; the owner must arrive at the agreed delivery time and the supplier must supply the bunker quantity as agreed, irrespective of market conditions after the agreement is made. Refusing to fulfil the contractual terms could lead to a hefty fine–usually higher than the cost of the requested fuel. More importantly, it would entail the expenses of a Court of Arbitration and claims of loss/damages, not to mention the cost of a damaged reputation in a tightly woven industry. This underscores the importance of honouring first stage

commitments if a spot price lower than the forward price is revealed at the bunkering port while underway.

Another important assumption is related to the permitted bunker stops. While Oh and Karimi (2010), and Vilhelmsen (2014) permit only one bunkering stop, we allow the vessel to take on bunkers both at an intermediate bunkering stop, time permitting, and to top up during cargo operations at the loading port.

The following assumptions also apply:

- Once a decision is made on stage 1, no deviation is permitted to other bunkering ports while sailing.
- The vessel has access to forward bunkering contracts for physical supply at the time of voyage negotiations (Stage 1).
- All potential bunkering ports have enough fuel to supply the vessel.
- The vessel can choose any bunkering port and plan for a voyage deviation or steam directly to the loading port and refuel while loading cargo.
- Only Heavy Fuel Oil is considered.
- The value of the remaining fuel at the EOH is assumed to be the average spot price for bunkering ports within the nominated discharging cargo area at the expected time of discharging
- The vessel is only allowed to load bunkers when empty (in ballast). This is industry practice as standard voyage contracts customarily include a clause that the vessel must steam with ‘utmost dispatch’ on the laden (loaded) leg.
- New forward contracts are not available while the vessel is underway (Stage 2).
- The vessel is not cancelled if it cannot meet the laycan window. Any delay is priced based on a Time Charter Equivalent (TCE) that is consistent with the agreed voyage freight rate and the average voyage time (including port stays). A good review of the TCE calculation can be found in Panayides (2018).

2.4.2 Mathematical formulation

Before introducing the mathematical formulation, we list the notation as follows:

Sets

I	Bunkering ports
L	Bunkering ports except for loading port
V	Discrete speed options for vessel on leg 1 (Departure port to bunkering port)
Z	Discrete speed options for vessel on leg 2 (bunkering port to load port)
S	Scenarios with spot price, EOH price and waiting time

Decision and derived variables

	Stage 1
z_{ivs}^{F}	Agreed amount from forward contract at port i and speed v under scenario s
z_{es}^{F}	Extra amount agreed from forward contract at loading port under scenario s
δ_{ivs}	Binary variable. 1 if selected port i and speed v, 0 otherwise under scenario s
F_{ivs}	Fuel remaining on board at bunkering port i when speed v is selected under scenario s
	Stage 2
z_{ivs}^{γ}	Agreed amount from spot market at port i and speed v under scenario s
${m_s}^{F}$	Received amount from forward contract under scenario s
u_s^F	Unused amount from forward contract under scenario s
	Stage 3
m_s^{γ}	Amount received from spot contract under scenario s
u_s^{γ}	Unused amount from spot contract under scenario s
δ_{ivzs}	Binary variable. 1 if port i and speed v is selected at leg 1, speed z for leg 2 under scenario s , 0 otherwise.
w_{ivzs}	Cargo requested at loading port after bunkering at port i with speed v at leg 1, speed z at leg 2 under scenario s

φ_s	Hours after laycan under scenario s
F_{es}	Fuel on board at arrival to load port under scenario s
d_{ivzs}	Draught at loading port after loading cargo and bunkering at port i with speed v at leg 1 and speed z at leg 2 under scenario s
G_s	Quantity remaining on board at the end of problem horizon under scenario s
	Risk variables
η_s^τ	Value-at-risk (VaR) at stage τ under scenario s
σ_s^τ	Linearization variable for the risk constraint at stage τ under scenario s

Parameters

α^τ	Confidence level for CVaR at stage τ; $\tau \in \{2,3\}$
λ	CVaR relative weight
K_{iv}	Fuel consumption to port i with speed v
t_{iv}	Time to port i with speed v
F_{max}	Maximum fuel tanks capacity
F_{min}	Minimum fuel on board at any time
F_o	Fuel on board at departure port
π	Freight rate ($/tonnes)
$C(\pi, R, \mathbb{E}(\mathrm{c}_{\mathrm{is}}^{\gamma}(\xi_2))$	Time charter equivalent (prorated to hours)
$\mathrm{c}_{\mathrm{i}}^{\mathrm{F}}$	Forward contract price at port i ($/tonnes)
$\mathrm{c}_{\mathrm{is}}^{\gamma}(\xi_2)$	Spot contract price at port i under scenario s
$\mathrm{t}_{\mathrm{is}}^{\mathrm{w}}(\xi_3)$	Waiting time at port i under scenario s
$\mathrm{c}_{\mathrm{i}}^{\mathrm{P}}$	Penalty for amount not received from spot contract fuel
K_{ivz}	Fuel consumption to port i with speed v at leg 1 and speed z at leg 2
$\mathbb{E}(K_{min})$	Expected fuel consumption at leg 3 (loading port to discharging port)
t_{ivz}	Time from port i to loading port after speed v at leg 1 and speed z at leg 2
R	Time to end of laycan from departure port
H	Tonnes per centimeter of immersion (TPC)

B	Tonnes of ballast on board
d_{max}	Max draught
w_{max}	Max weight (10% additional to cargo nomination)
β	Pumping rate of barge $\frac{tonnes}{hour}$
$c_s^e(\xi_3)$	Spot fuel price in the proximity of the discharge port at the expected arrival time

The problem can be formulated as:

$$max \sum_{s\in S} p_s \left(\sum_{i\in I}\sum_{v\in V}[-c_{iv}^F z_{ivs}^F] - c_e^F z_{es}^F - \sum_{i\in I}\sum_{v\in V} c_{is}^\gamma z_{ivs}^\gamma + \sum_{i\in L}\sum_{v\in V}\left[\left(c_{is}^\gamma - c_i^P\right)u_s^\gamma\right] + \sum_{i\in I}\sum_{v\in V}\sum_{z\in Z}[\pi w_{ivzs}] + \sum_{j\in I\setminus L}\sum_{v\in V}[\pi w_{jvs}] + c^e G_s - \varphi_s C + \lambda\left(\eta_s^2 - \frac{1}{1-\alpha^2}\sigma_s^2 + \eta_s^3 - \frac{1}{1-\alpha^3}\sigma_s^3\right)\right) \quad (1)$$

subject to:

CVaR Constraints:

$$\eta_s^2 - \left(\sum_{i\in I}\sum_{v\in V}[-c_{iv}^F z_{ivs}^F] - c_e^F z_{es}^F - \sum_{i\in I}\sum_{v\in V} c_{is}^\gamma z_{ivs}^\gamma\right) \le \sigma_s^2 \qquad \forall s \in S \quad (2)$$

$$\eta_s^3 - \left(\sum_{i\in I}\sum_{v\in V}[-c_{iv}^F z_{ivs}^F] - c_e^F z_{es}^F - \sum_{i\in I}\sum_{v\in V} c_{is}^\gamma z_{ivs}^\gamma + \sum_{i\in L}\sum_{v\in V}\left[\left(c_{is}^\gamma - c_i^P\right)u_s^\gamma\right] + \sum_{i\in I}\sum_{v\in V}\sum_{z\in Z}[\pi w_{ivzs}] + \sum_{j\in I\setminus L}\sum_{v\in V}[\pi w_{jvs}] + c^e G_s - \varphi_s C(\pi, R, \mathbb{E}(c^e))\right) \le \sigma_s^3 \qquad \forall s \in S \quad (3)$$

$$F_{ivs} = \delta_{ivs}(F_o - K_{iv}) \qquad \forall i \in I, \forall v \in V, \forall s \in S \quad (4)$$

$$F_{ivs} \ge \delta_{ivs} F_{min} \qquad \forall i \in I, \forall v \in V, \forall s \in S \quad (5)$$

$$\sum_{i\in I}\sum_{v\in V}\delta_{ivs} = 1 \qquad \forall s \in S \quad (6)$$

$$z_{ivs}^F + z_{ivs}^\gamma \le \delta_{ivs}(F_{max} - F_o + K_{iv}) \qquad \forall i \in I, v \in V, \forall s \in S \quad (7)$$

$$\delta_{lvs} - \sum_{z\in Z}\delta_{lvzs} = 0 \qquad \forall l \in L, \forall v \in V, \forall s \in S \quad (8)$$

$$u_s^\gamma = \sum_{l\in L}\sum_{v\in V}\left[z_{lvs}^\gamma\right] - m_s^\gamma \qquad \forall s \in S \quad (9)$$

$$u_s^F = \sum_{l\in L}\sum_{v\in V}[z_{lvs}^F] - m_s^F \qquad \forall s \in S \quad (10)$$

$$\left(m_s^F + m_s^\gamma\right)\frac{1}{\beta} \leq \sum_{l\in L}\sum_{v\in V}\sum_{z\in Z}\left[\delta_{lvzs}\left(R - t_{lv} - t_{lvz} - \mathrm{t}_{l_s}^{\mathrm{w}}\right)\right] + \varphi_s \quad \forall s \in S \quad (11)$$

$$\delta_{jvs}t_{jv} \leq R \quad \forall j \in I\backslash L\,, \forall v \in V, \forall s \in S \quad (12)$$

$$F_{es} = F_o - \sum_{i\in I}\sum_{v\in V}\delta_{ivs}K_{iv} - \sum_{l\in L}\sum_{v\in V}\sum_{z\in Z}\delta_{lvzs}K_{lvz} + \sum_{j\in I\backslash L}\sum_{v\in V}[z_{jvs}^F + z_{jvs}^\gamma] + m_s^F + m_s^\gamma \quad \forall s \in S \quad (13)$$

$$z_{es}^F \leq F_{max} - F_{es} \quad \forall s \in S \quad (14)$$

$$F_{es} - E(K_{min}) + z_{es}^F \geq G_s \quad \forall s \in S \quad (15)$$

$$F_{es} \geq F_{min} \quad \forall s \in S \quad (16)$$

$$d_{lvzs} = \left(d_o + \frac{F_{es} - F_o - B + w_{lvzs} + z_{es}^F}{100H}\right)\delta_{lvzs} \quad \forall l \in L, \forall v \in V, \forall z \in Z, \forall s \in S \quad (17)$$

$$d_{jvs} = \left(d_o + \frac{F_{es} - F_o - B + w_{jvs} + z_{es}^F}{100H}\right)\delta_{jvs} \quad \forall j \in I\backslash L, \forall v \in V, \forall s \in S \quad (18)$$

$$d_{lvzs} \leq d_{max} \quad \forall l \in L, \forall v \in V, \forall z \in Z, \forall s \in S \quad (19)$$

$$d_{jvs} \leq d_{max} \quad \forall j \in I\backslash L, \forall v \in V, \forall s \in S \quad (20)$$

$$w_{lvzs} \leq \delta_{lvzs}w_{max} \quad \forall l \in L, \forall v \in V, \forall z \in Z, \forall s \in S \quad (21)$$

$$w_{jvs} \leq \delta_{jvs}w_{max} \quad \forall j \in I\backslash L, \forall v \in V, \forall s \in S \quad (22)$$

$$z_{ivs}^F = z_{ivs'}^F, z_{ivs}^\gamma = z_{ivs'}^\gamma, \delta_{ivs} = \delta_{ivs'}, m_s^F = m_{s'}^F, u_s^F = u_{s'}^F, z_{es}^F = z_{es'}^F \quad \forall s = s' \in S\; s.t.\; \xi_s^t = \xi_{s'}^t, \forall \tau \in \{1,2\} \quad (23)$$

$$\sigma_s^\tau = \sigma_{s'}^\tau \quad \forall s = s' \in S\; s.t.\; \xi_s^t = \xi_{s'}^t, \forall \tau \in \{2,3\} \quad (24)$$

$$\delta_{ivs} \in \{0,1\} \quad \forall i \in I, \forall v \in V, \forall s \in S \quad (25)$$

$$\delta_{lvzs} \in \{0,1\} \quad \forall l \in L, \forall v \in V, \forall z \in Z, \forall s \in S \quad (26)$$

$$z_{ivs}^F, z_{ivs}^\gamma \geq 0 \quad \forall i \in I, \forall v \in V, \forall s \in S \quad (27)$$

$$w_{ivzs}, d_{ivzs} \geq 0 \quad \forall i \in I, \forall v \in V, \forall z \in Z,\ \forall s \in S \quad (28)$$

$$w_{jvs}, d_{jvs} \geq 0 \quad \forall j \in I\backslash L, \forall s \in S \quad (29)$$

$$z_{es}^F, u_s^\gamma, u_s^F, G_s, m_s^\gamma, m_s^F, \varphi_s \geq 0 \quad \forall s \in S \quad (30)$$

$$\sigma_s^\tau \geq 0\,, \tau \in \{2,3\} \quad \forall s \in S \quad (31)$$

The objective function (1) maximizes a multistage CVaR time consistent function as presented by Homem-de-Mello and Pagnoncelli (2016) and adapted for the maximization case by Bushaj et al. (2022). For our case, the model minimizes the cost of forward contracted fuels; the spot contracted fuels and penalties for not loading the agreed spot fuel[3]. The EOH fuel value and the voyage income generated by the cargo

[3] Note that forward fuel has no salvage value and the unused part of a forward fuel order is lost in full. This forces the use of forward fuel before using spot fuel.

intake are added to the fuel costs. A penalty for arriving late to the laycan proxies the influence of the value of time under different market conditions. The CVaR importance to the objective function is controlled by λ and can be adjusted to compare the relative importance of CVaR against the expected value. The risk neutral problem equivalent is derived by λ being zero.

Constraints (2) and (3) represents the risk attitude of the shipowner via CVaR constraints for the second and third stages. They are linearizations (together with σ^{τ} at the objective function) of the optimization formula for CVaR as per Rockafellar and Uryasev (2000).

Constraint (4) is a bookkeeping constraint of fuel remaining on board at the bunkering port and at the loading port. Constraint (5) ensures that the vessel has always more fuel than a safe margin. Constraint (6) imposes that only one bunkering option is selected. For the resupply path (i.e. when the fuel is partly loaded at a bunkering port and topped up at the cargo loading port), this constraint could be interpreted as selecting the bunkering port where the first bunker order is taken.

Constraint (7) limits any fuel order to the maximum capacity of the vessel's fuel tanks. Constraint (8) links the port and speed decision on the first leg to the second leg speed decision. Constraints (9) and (10) derive the unused fuel from the spot and forward contracts, respectively. Constraint (11) reduces the pumped fuel quantity if the ordered bunker is so large that the vessel ends up being largely penalized for late arrival at the loading port. We introduce this soft constraint as a mechanism to assess the market conditions on the bunkering decision. By using this version instead of its hard constraint equivalent (i.e. where the laycan is enforced in all scenarios), we permit late arrivals to be penalized according to the market state. A hard constraint limits the options to ports with low waiting time and a fast vessel speed if, from their waiting time distribution, no scenario leads to infeasibility. The fast speed would not be a response to the uncertainty, but a way to avoid infeasibility with large extremes.

Constraint (12) limits the vessel to arrive before the laycan when there is no deviation for bunkers. Constraint (13) is a bookkeeping constraint for fuel remaining on board at the loading port. Constraint (14) limits fuel orders to the bunkering tank max capacity, only at the loading port. Constraint (15) forces the EOH fuel quantity to be higher than the sum of the expected fuel consumption for the third leg (loading port to discharging

port), the fuel from the discharging port to the closest bunkering port, and a safety margin for the minimum fuel remaining. Constraint (16) ensures that the fuel remaining on board at the loading port is higher than the minimum fuel allowance. Constraint (17) defines the vessel's draught at the loading port after any loaded bunker and cargo. Constraint (18) defines the vessel's draught if the vessel travels directly to the loading port with no bunkering deviation. Constraints (19) and (20) limit the vessel draught to the maximum permitted draught. Constraints (21) and (22) limit the loaded cargo to less than the nominated quantity plus ten percent. Constraints (23) and (24) defines the non-anticipativity constraints. Constraints (25-31) define the domains of the decision variables.

2.5 Empirical case study

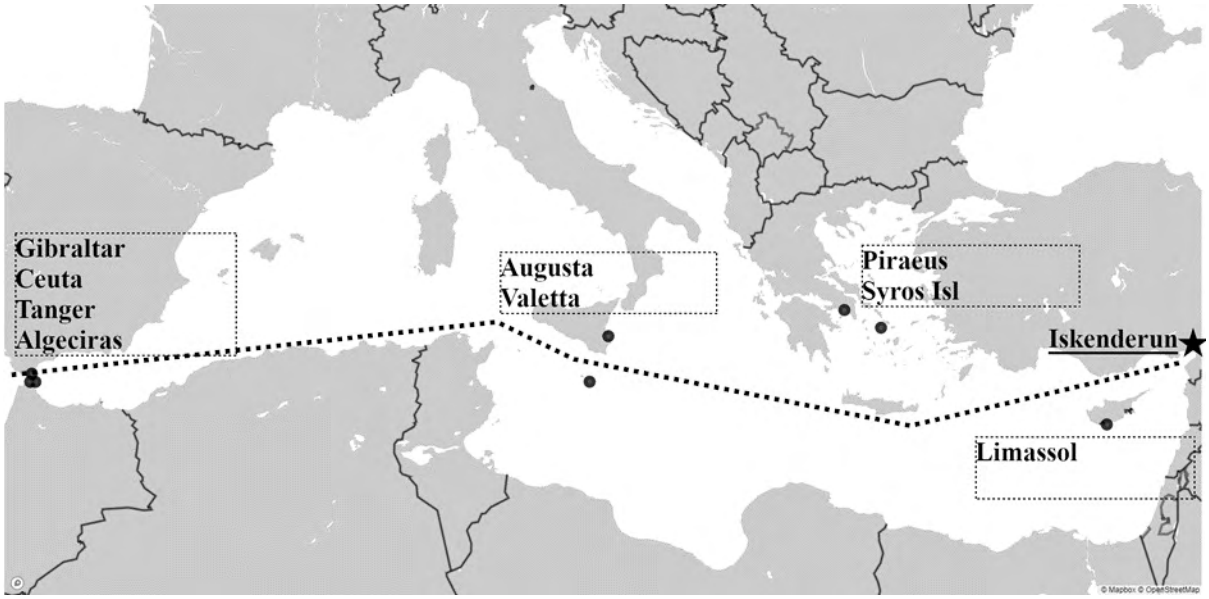
2.5.1 Case particulars

We test the proposed mathematical formulation by introducing the case of a Supramax[4] bulk carrier on a voyage contract. The owner has agreed the freight rate and voyage contract terms for a cargo of scrap from Boston, USA, to Iskenderun, Turkey. The contract establishes that the vessel will be discharging in the Iskenderun area, without mentioning to what port. The vessel starts from Iskenderun, such that this is in effect a round trip voyage with a ballast leg from Iskenderun to Boston followed by a laden leg from Boston to Iskenderun. We assume that at the first stage the vessel has remaining fuel from a previous voyage and has been instructed on the contractual requirements (i.e. laycan). With this information, the decisions to be made at the first stage are: to what port to deviate for bunkers, the sailing speed, and the quantity of fuel to fix on a forward contract at the bunkering port. Figure 2 shows the available bunkering port options for the intended voyage and the direct route if no deviation for bunkering is considered.

[4] A Supramax is a class of bulkcarrier vessel ranging from 50,000 to 60,000 DWT

Figure 2

Bunkering port options



Note: The Atlantic leg (after Gibraltar strait) is assumed a rhumb line route to Boston

The first stage decisions should account for the possibility that the spot price might change while underway and an opportunity might open for buying bunkers at a lower price. Additionally, they should consider that the waiting time before loading bunkers might force the vessel to either reduce the bunker quantity loaded, paying a penalty, or arriving late to the loading port. In reality, a late arrival gives the right (but not the obligation) to the charterer to cancel the voyage contract, regardless of how far the vessel is from the port. For illustration, we assume that a late arrival is penalized by an hourly prorated Time Charter Equivalent (TCE), internally consistent with the agreed freight rate and the expected spot fuel price across the route This is one way of reflecting the value of time under different market conditions and the opportunity cost of not being available for a next voyage. It is also a proxy for the daily rate at which the vessel could be fixed again if the contract is cancelled. Overall, the hypothesis is that in strong markets, it would be more expensive to bunker at ports with long waiting times. Additional relevant parameters are summarized in Table 1.

Table 1

Parameters for the case study problem

Parameter	Value	Comments
F_{max}	3,780 tonnes	
F_{min}	50 tonnes	
F_o	160 tonnes	Enough fuel to reach Gibraltar at 12.5 knots + safe margin
c_i^P	$20.00 \frac{\$}{\text{tonnes}}$	
π	$20.5 \frac{\$}{\text{tonnes}}$	
$\mathbb{E}(K_{min})$	470 tonnes	Based on $24 \frac{MT}{day} x 17 days(12.5 kts) + 15\%$
R	420 hours	Voyage time. Equivalent to a direct route from Iskenderun to Boston at 12.5 kts.
w_{max}	48,000 tonnes	
$\mathbb{E}(c_{is}^{\gamma}(\xi_2))$	$400.9 \frac{\$}{\text{tonnes}}$	Expected spot fuel prices across route options $c_{is}^{\gamma}(\xi_2)$
$C(\pi, R, \mathbb{E}(c_{is}^{\gamma}(\xi_2)))$	$1{,}397.98 \frac{\$}{\text{hour}}$	$C = \frac{w_{max}\pi - 2\mathbb{E}(K_{min})\mathbb{E}(c_{is}^{\gamma}(\xi_2)) - port\ cost}{R}$ (32)
H	$55 \frac{\text{cm}}{\text{tonnes}}$	
B	25,000 tonnes	
d_0	6.9 m	
d_{max}	11.5 m	Maximum draught at Schnitzer Scrap Terminal, Boston.
β	$\beta = 300 \frac{\text{tonnes}}{\text{hour}}$	
α_2, α_3	0.05	Expected return on the worst 5% scenarios
λ	1	Value set to 0 is equivalent to the risk neutral problem

2.5.2 Consumption estimation

The bunker consumption is an important factor for the bunkering decision, as demonstrated by Yao et al. (2012). Some studies include consumption into the problem as discrete values, conditional on vessel activity (Oh and Karimi, 2010) or load condition (Vilhelmsen, 2014). Another approach would be to incorporate fuel consumption functions as part of the optimization model, such as with the empirical functions for container vessels introduced by Yao et al. (2012), or linear piecewise approximations (Wang and Meng, 2015; Gu et al., 2019) from traditional quadratic consumption functions (e.g., Ronen, 1982).

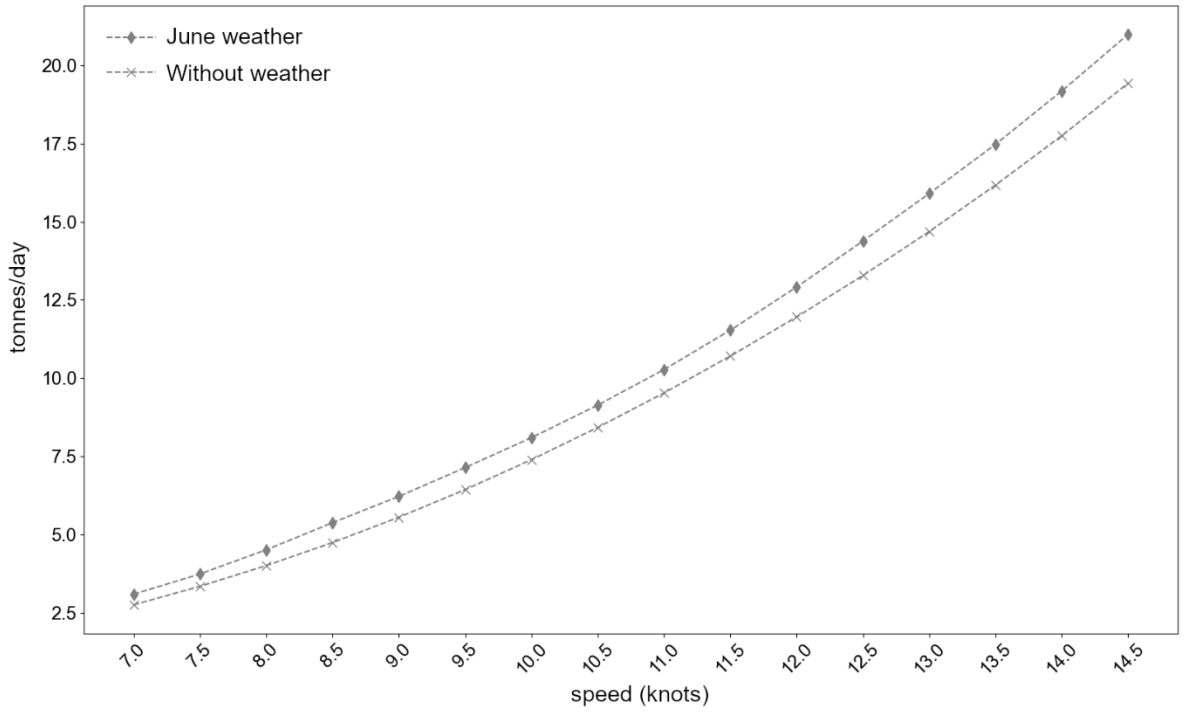
The fuel consumption is estimated as the product of the power demand for the vessel's route and the Specific Fuel Oil Consumption as derived from the Fourth IMO Green

House Gas Study (2020). This approach is motivated by the vast literature on a vessel's resistance calculation, nicely summarized by Zis et al. (2020), which incorporates the interaction of the vessel specific factors e.g., speed, displacement and draught, with the environmental effects, such as wind, currents and waves. As our objective is to simply a realistic physical environment for the route, our selection is based on the average weather and oceanographic conditions of the month of June between 2018 and 2021, this being the month with the least variance for the Mediterranean according to Barbariol, et al. (2021). A description of the estimation method used for this study can be found in the Appendix. The fuel consumption curves in June are 10% higher than when weather is not considered and represents an average weather margin for the month. We leave the evaluation of the impact of stochastic weather on the bunkering decision for future research.

The fuel consumption is calculated on leg 1 from Iskenderun to every bunker port option and on leg 2 from every bunker port to the loading port (Boston). The time of departure for leg 2 is estimated as the arrival time at leg 1 resulting from the speed decision at V plus the expected waiting time at the bunkering port. An example for a trip from Iskenderun to Boston is shown in Figure 3.

Figure 3

Fuel consumption from Iskenderun to Boston for the case study Supramax vessel in ballast condition



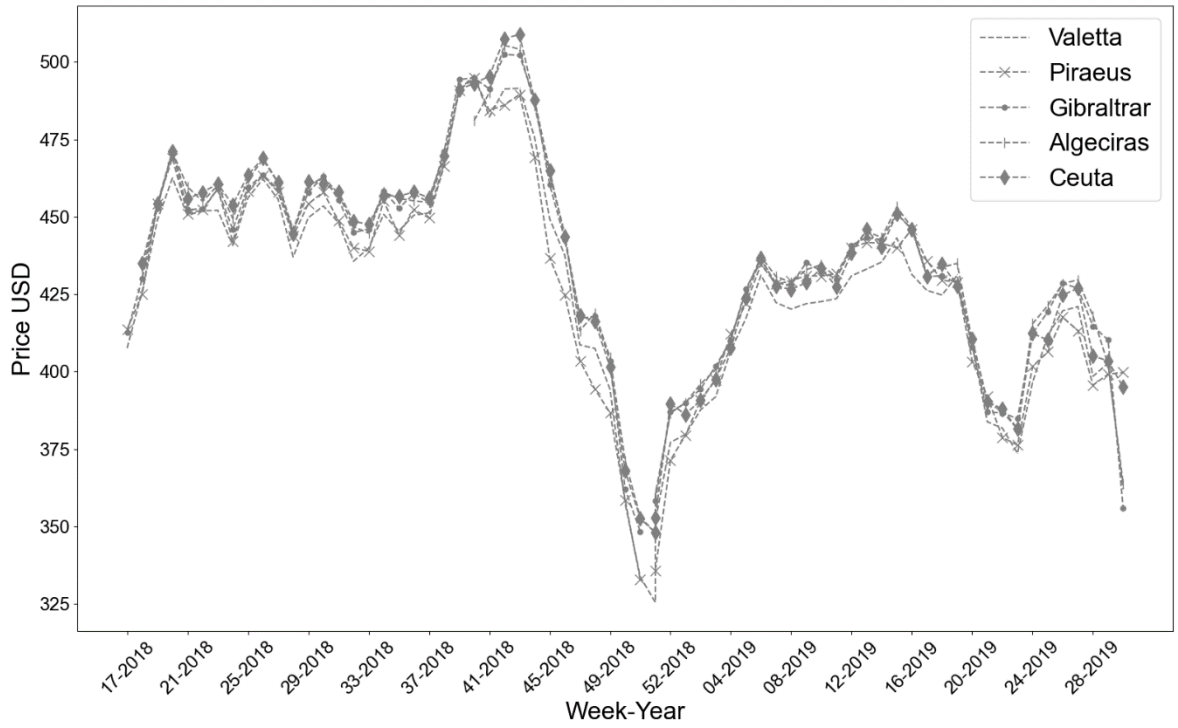
Note: Authors' own calculations

2.5.3 Scenario tree generation

The model incorporates three stochastic parameters, the spot fuel oil price, the spot fuel price at the EOH, and the bunkering ports' waiting times. The spot fuel oil price is sourced from historical records (April 2018 to July 2019) from daily HFO 380 bunker prices kindly provided by Bunker-Ex. A sample with the five cheapest ports at the Mediterranean is presented in Figure 4. Valetta is, on average, the port with the lowest fuel price, followed by Piraeus and Gibraltar. The tight relationship between Gibraltar and Algeciras prices suggests fierce competition between the two geographically close ports. The high volatility of fuel oil prices is derived from oil market dynamics and the local balance of supply and demand (Alizadeh et al., 2004). Gu et al. (2019) suggest that the fuel price development can be considered as a Lévy process with independent increments, referencing the crude oil models of Gencer and Unal (2012).

Figure 4

HFO 380 fuel price for the five (on average) cheapest ports in the Mediterranean



Note: Authors' graph information sourced from Bunker Ex historical records

Similar to the strategy employed by Gu et al. (2019) we generate distributions for fuel price increments from the difference between fuel price of port i at day n_i and days $n_i + \frac{dist_i}{24v}, \forall v \in V, \forall i \in I$. Thereafter, the results are averaged per port and compiled.

As the vessel's sailing days to the bunkering port varies based on speed and distance, the time and space features of the fuel prices in the bunkering ports also show across the calculated increments. For instance, we expect that the ports closest to the departing position should have, on average, lower increments than those further away. This can be seen from the descriptive statistics in Table 2, as the lowest "increment" is observed in Limassol (the closest port), while the highest is in Boston (farthest away). However, this relationship is not consistent in all cases (e.g., Valetta and Augusta) as some ports' prices might be more affected by local and regional dynamics and time. From Table 2 it is important to highlight the extreme values as potential scenarios to be observed in a voyage. The resulting joint empirical distribution and their underlying dynamics in our model are discretized by generating equiprobable scenarios via the Copulas scenario generation method (Kaut & Wallace, 2011). The mean and standard

deviations of the resultant spot prices after discretization are shown in the right section of Table 2.

Table 2

Descriptive statistics of empirical bunker price increments at Mediterranean ports and spot prices in stage 2 for the case study

	Bunker price increments				Spot prices	
Port	**Mean**	**Std.Dev**	**Min**	**Max**	**Mean**	**Std.Dev**
Boston	-1.280	13.943	-36.00	30.14	396.43	13.50
Valetta	-0.918	15.170	-46.94	53.45	385.84	14.23
Algeciras	-0.853	14.842	-37.61	43.30	390.29	14.11
Gibraltar	-0.711	15.795	-53.35	57.5	386.53	15.38
Ceuta	-0.697	14.562	-38.98	46.15	388.83	14.02
Syros Island	-0.621	10.330	-32.44	33.79	429.79	9.92
Augusta	-0.559	15.569	-46.38	67.83	412.75	14.30
Piraeus	-0.260	14.507	-56.05	60.53	386.62	13.16
Limassol	-0.115	16.764	-83.98	50.23	430.96	14.62

A similar strategy, incorporating time and space features, is used for generating the forward price c_{iv}^{F} parameter. To maintain consistency, we here assume that the forward price is a function of the expected spot price, such that: $c_{iv}^{F} = E[c_i^{\gamma}]e^{\left(0.05\frac{t_{iv}}{8760}\right)}$. The selection of 0.05 is arbitrary and reflects the holding cost and the risk-free rate of the contract. We assume that the cost of carry principle holds, resulting in a forward curve that is in contango such that the futures price is higher than the spot price. Considering that bunker prices inherit the characteristics of the crude oil market, we acknowledge that backwardation (spot price higher than the future price) is also possible and would suggest the incentive to buy forward instead of spot. However, to avoid further complexity, we here consider only the more common contango situation.

For modelling the waiting times at port, we replicate the method for generating bunkering statistics from AIS data by Fuentes (2021). In our case, we compile the waiting times for all vessels that had bunker operations at anchorage from January

2015 to January 2018. From Table 3, Gibraltar is shown as the bunkering port with the lowest waiting times, on average, followed by Syros Island and Valetta. It is important to highlight that looking only at the average waiting time does not ensure the shortest stay, as a port with lower waiting times might still have scenarios with higher extremes (i.e., Syros vs. Valetta). We explore the effects of the waiting time distribution by discretizing it in the scenario tree via the Copulas method.

Table 3

Waiting time (hours) per bunkering port from January 2015 to January 2018

Port	Mean	St.Dev	Median	Max
Gibraltar	1.68	0.41	1.62	3.22
Syros Island	1.94	2.42	0.76	13.28
Valetta	2.33	2.17	1.74	10.05
Limassol	3.14	3.39	2.33	25.79
Ceuta	3.36	2.94	2.71	20.32
Augusta	5.56	5.36	3.46	25.83
Algeciras	5.79	2.53	5.47	15.87
Piraeus	6.42	6.15	3.90	26.23

Our motivation for using copulas is threefold. First, the generation method is compatible with the empirical distributions used as input. In the absence of a parametric distribution that could fit our data, this is an important feature. Second, the copulas scenario generation method matches the first and second moments of our data empirical distribution, regardless of the tree size. Third, our scenarios can convey the relationship of price increments along the voyage and across bunkering options.

A natural interpretation of the copula theory in this setting is that the copula function mapping the multivariate joint distribution–i.e. spot fuel prices and waiting times across ports–contains the dependence information of all ports' variables and their marginal distributions. An additional gain, as suggested by Kaut (2014), the copulas generation method is a better scenarios generation method compared to other methods such as sampling and moment matching. This argument is based on the fewer scenarios needed to achieve stable solutions. This is an important feature in our context

considering that multistage models can become intractable if the scenario tree is too large.

We propose an in-sample stability test (Kaut & Wallace, 2007) to verify that the results are not affected by the tree size. A stable result should have similar objective values across different trees constructed with the same generation method (Kaut & Wallace, 2007). As our tree is generated with the Copulas method, which is based on an optimal assignment of scenarios as seen in Kaut (2014), we compare 11 different trees ranging from 50 to 100 scenarios. These trees are based on combinations of 10 to 15 scenarios at the second stage and 5 to 10 scenarios at the third stage. A summary of the tested scenario trees is introduced in Table 4.

Table 4

Objective values for scenario trees tested for the in sample stability analysis

2nd stage scenarios	3rd stage scenarios	Objective value	Stage compared: max (diff to min, diff to max)		
			3rd stage	2nd stage	All
10	**5**	695,258.8	0.8%	-	1.3%
10	**6**	693,256.2	0.5%	-	1.0%
10	**7**	689,788.9	0.8%	-	0.8%
10	**8**	691,997.3	0.4%	-	0.8%
10	**9**	691,805.2	0.3%	-	1.3%
10	**10**	694,831.3	0.7%	1.3%	0.9%
11	5	688,748.3	-	0.9%	1.2%
12	5	687,075.0	-	1.1%	1.2%
13	5	686,986.7	-	1.1%	1.0%
14	5	686,039.0	-	1.3%	1.3%
15	5	688,585.9	-	0.9%	0.9%

From the results in Table 4, stability can be observed at the third stage with as few as five scenarios. This is based on slight changes in the objective values (within 0.8%) when the third stage is tested with 5 to 10 scenarios. These results are stable despite

the random assignment of EOH prices at the third stage. It also shows that the model is stable regardless of the EOH price and waiting times match in a scenario. Similarly, the second stage is tested by five trees from 10 to 15 scenarios, with a difference of 1.3% between the trees' objective values. Overall, all trees included have a difference within 1.3%, suggesting that a tree with 10 scenarios at the second stage and 5 scenarios at the third stage is stable for our model in the current setting.

2.6 Computational experiments and discussion

2.6.1 Experiments description

Our chosen experiments are based on distinct situations that a shipowner could face when choosing her bunkering plan. For all situations, the vessel decides on its next voyage while completing a cargo discharge at Iskenderun. For this next voyage, the owner takes a stance on risk aversion and such attitude is modeled by the CVaR measure.

Our motivation for using CVaR is based on the observation that in a risk neutral setting with our contango assumption, a shipowner would always be better off (in expectation) buying bunkers from the spot market. This strategy seems obvious based on expectations, yet it has the risk of large losses when an adverse (higher) spot price is revealed. Introducing risk aversion in this setting helps to hedge against those adverse scenarios by reducing the expected solution variance. We test the effect on the solution when risk aversion is modeled for stage 2 and risk neutrality on stage 3, and when risk aversion is modeled for both stages. Based on our formulation, a shipowner could isolate risk-aversion levels per decision stage or transform the model back into a risk neutral setting.

Situation 1 refers to our benchmark. In *situation 2*, we remove the waiting times and explore the effect on the proposed solution.

Situation 3 explores the decision if instead of valuing the EOH fuel in a high price area (Iskenderun); we value the EOH fuel in the more price competitive Gibraltar Strait area (c.f. Table 2). Disclosing the discharge area is a customary practice for voyage charters. For this experiment, we explore the effect of using an EOH price different to that of our discharge area. This could be interpreted as the assessment a shipowner

would have to make if she were considering repositioning her vessel after the current voyage.

In *situation 4,* the voyage laycan is reduced ($R = 380\ hours$) to model the effect of a shortened laycan on the speed decision, the effect of the exponential fuel consumption function on depleting the fuel inventory and the ordered bunker quantity. Conversely, in *situation 5*, we refer to a pressing laycan ($R = 800\ hours$) and compare the results to *situation 4*.

For *situation 6 and situation 7,* we explore the effect of freight and fuel market conditions on the bunkering decisions. This model evaluates the influence on changes to the TCE and the relative cost of time. For *situation 6* we assume a high freight market ($\pi = 30.9 \frac{\$}{\text{tonnes}}$) and a low freight market ($\pi = 11.5 \frac{\$}{\text{tonnes}}$) for *situation 7*.

2.6.2 Numerical results and discussion

In this section we present and discuss the results for our experiments as run by modeling the problem in mpisppy (Knueven, et al., 2020). This module works as an extension for solving stochastic Pyomo (Hart, et al., 2017) formulations. The numerical experiments were conducted on a single computer with 16GB RAM memory and 12 virtual processors with 2.60 Ghz core frequency.

2.6.2.1 Benchmark (Situation 1 and Situation 2)

The results of situation 1 and situation 2 in Table 5 reflect the close competition of Valetta and Gibraltar. A review of their descriptive statistics in Table 2 (spot prices) and Table 3 (waiting times) reveals that Valetta is the cheapest port with a relatively short average waiting time (ranked 3rd) while Gibraltar has the shortest waiting times and the second cheapest price. Simplifying their profiles in this manner helps in judging whether the proposed decisions were driven by price (Valetta) or waiting time (Gibraltar). An interesting feature of our case study is that given that Valetta and Gibraltar are close to the direct route to the loading port, they cause similar deviations. Therefore, we could simplify their profiles without the concern of other parameters affecting the solution. Additional parameters that could cause different optimal decisions are the laycan, the penalty for fuel not taken from the order and barge pumping rates per port. We set them to be identical for all port options.

Table 5

Numerical results in situation 1 and situation 2 under differerent risk level assumptions

	Situation 1	Situation 2	Situation 1	Situation 2
Risk-aversion (stage2, stage3)	**(0.95,0)**		**(0.95,0.95)**	
First stage				
Port - speed	Valetta 13.3	Valetta 13.2	Gibraltar 13.2	Gibraltar 12.3*
Forward fuel contracted	630.9	869.5	568.7	563.6
Second stage				
Average spot fuel contracted	1,973.9	1,880.5	1,025.9	1,023.1
Average unused forward fuel	0	0	0	0
Third stage				
Average voyage consumption	272.8	270.8	265.1	263.3
Average unused spot fuel	198.3	228.8	38.9	38.2
Average fuel at EOH	1,523.0	1,640.3	680.6	675.0
Expected total profit	695,258.80	698,761.30	673,257.37	673,643.79
Total cost standard deviation	67,659.15	66,699.79	52,251.85	56,741.06

** Speed reduction compared to situation 1 is caused by the option to "adjust" speed on the third leg and discretization of third leg speeds to every 0.5 knots.*

From Table 5, we observe that the cheapest port (Valetta) is selected when risk neutrality is assumed for stage 3. Here, the shipowner takes the following two actions. First, she speculates on the spot fuel as the EOH price of fuel (stage 3) is higher in the EOH area, in expectation. Second, they accept a longer stay in port (waiting time plus service time) caused by that "large" bunkering order. A large bunkering order results in a longer service time at a port that, together with long waiting times, could develop in too large penalties for late arrival. The waiting time affects the solution by limiting the bunker order to that necessary to reduce the penalties for late arrival at the load port. The cost of the expected "extra" profit from the EOH fuel outweighs the loss of the penalty of the expected delays caused by the waiting times.

For a strategy that models risk aversion separately for stage 2 and stage 3, the decision is to take bunkers at the port with the smaller waiting time distribution (Gibraltar). The

strategy also includes a smaller order from the forward market and the spot market. The forward fuel order is in place to hedge against spot price fluctuations, while the port selection and a smaller spot fuel order have two effects on stage 3 uncertainty. First, a smaller fuel order reduces the service times; therefore, the time penalty is on scenarios with larger waiting times. The larger waiting times are also reduced by choosing a port with a compact distribution of low waiting times. Second, the smaller order reduces the loss caused on scenarios where the EOH price was lower than the spot price. In this setting, the waiting times defines the port to choose, based on tighter distributions.

Results for situation 2 (no waiting time used), reflects a stronger effect (when comparing the expected profit value) in Valetta than in Gibraltar (where almost no effect is observed). The reason is the lower average and more compact distribution for Gibraltar waiting times. For this situation, either Valetta or Gibraltar could be selected from a flat objective function caused by equal waiting times (0) and tight fuel prices differences. These differences are within the margins of our in sample stability test presented in Table 4.

Based on these results, we can conclude that decisions based on prices alone will not cause a significant difference in profits if the cheap port has low waiting times. In our setting, the selection of one port or the other is mainly based on whether the shipowner is assumed to be risk averse or risk neutral on stage 3. Assuming risk neutrality on stage 3 results in a speculative strategy where the shipowner chooses the cheapest port, aiming for profit on the EOH fuel value while just assuming the effect of the expected waiting time on that strategy. Conversely, a strategy with risk aversion on stage 3 selects the port with lower waiting times and reduces the fuel order to hedge scenarios with longer waiting times and/or lower EOH prices.

Special case (Situation 2 –no waiting time and Valetta is not available)

In addition to the results in Table 5, we extend our experiments to assess a special case of Situation 2 where Valetta is not available, for instance due to fuel availability constraints. The optimal decision, based on this setting, is to bunker in Piraeus and sail there at 11.2 knots when risk neutrality is assumed in stage 3. The selection of Piraeus is based on three factors. First, as risk neutrality is assumed in stage three, the appetite for speculation on bunker orders promotes buying from the cheapest port.

Piraeus has the third lowest price on average and a lower standard deviation of prices compared to Gibraltar (see Table 2). Second, the port selection does not consider waiting times, thus, the lower speed is permitted. The inclusion of waiting times limits the selection of Piraeus otherwise. Finally, despite Piraeus having a larger deviation than Gibraltar, the difference in consumption is smaller due to the lower speed and fuel consumption being an exponential function of speed.

Importantly, if the decision is evaluated in a tree that incorporates waiting times, this causes a loss of 4.1% in expected profits compared to the optimal decision's expected profits. This highlights what can happen when a bunker option is selected based on price alone in an inefficient port. We note that sailing speed also proves to be important to compensate for large waiting times, with the large impact on fuel cost that this can have.

We also evaluate the case of not having Valetta as an option for situation 2 when risk aversion is modeled separately on stage 2 and stage 3. Here, the decision would be to bunker at Gibraltar. However, the difference to Piraeus is smaller than in the stage 3 risk neutrality case (1.0%). This is the case as there is less expected profit from the EOH fuel value. Therefore, the slight difference in spot price would have a lower effect than in the stage 3 risk neutrality case.

Hedging on stage 3 has an additional implication for this case. From the results, the best option is to steam to Piraeus at 11.4 knots. If we plug this decision into a tree with waiting times, it results in a loss of 0.8% compared to situation 1 with risk aversion on stage 2 and stage 3 (c.f. Table 5). This result is consistent with our interpretation of smaller orders, reducing the cost of larger stays and reducing the losses on scenarios with an EOH price lower than the spot fuel price.

From this special case, the results suggest that a decision based on prices alone could result in a significant loss if the selected port has long waiting times. More importantly, the waiting time has an effect on which port to choose. Speed is also an important decision to reduce the effect of long waiting times on the voyage time progression. Our experiment suggests that hedging for third stage decisions reduces the potential large losses observed when no risk measures are assumed in the third stage. Therefore, a strategy (midway solution), when waiting times are not available, is to model the problem so that decisions hedge against third stage risk.

2.6.2.2 EOH effect (Situation 3)

By comparing expected values, we can infer that having a higher EOH price, such as in the Iskenderun area, would prompt the shipowner to buy fuel in the cheapest bunkering port so as to take maximum advantage of the price difference to the high price area. Such a case was observed in the form of a large spot order for situation 1 when risk neutrality is assumed for stage 3.

Here, we observe in Table 6 that the decision is to bunker at Valetta or Gibraltar when assuming risk neutrality or risk aversion on stage 3, respectively. The decisions, including speed, are consistent with our results for situation 1 and situation 2[5]. Specifically, we observe a larger order of forward fuel but a small order from the spot market. This suggest that without the incentive to speculate, there is no case for buying fuel in the spot market and the required bunker for the voyage is taken in full from a forward contract. As the order is limited to cover the current voyage, the waiting time has an effect on the port selection only if assumed risk aversion is assumed. It turns out that the Valetta option, despite not having the lowest waiting times, still offers a good tradeoff of price vs. waiting time.

[5] A slower speed decision on the first stage is compensated by a faster speed on the third stage. The opposite is observed on situation 1 and situation 2. A way to prove such occurences is by comparing the average voyage consumption reported in Table 5 and Table 6.

Table 6

Numerical results in situation 3 under differerent risk level assumptions

Risk-aversion (stage2, stage3)	Situation 3 (0.95,0)	Situation 3 (0.95,0.95)
First stage		
Port - speed	Valetta 12.2	Gibraltar 12.5
Forward fuel contracted	876.0	873.9
Second stage		
Average spot fuel contracted	61.8	7.4
Average unused forward fuel	0	0
Third stage		
Average voyage consumption	266.0	263.9
Average unused spot fuel	12.4	1.5
Average fuel at the EOH	49.4	5.9
Expected total profit	645,752.4	645,748.3
Total cost standard deviation	3,452.42	346.60

One could argue that this setting is more realistic given that a cargo-carrying vessel's main business is not speculating on bunkers. However, it is important also in practice to assign an appropriate price to the fuel remaining on board at the end of the current voyage, and one way of doing so is use price information at the place where the vessel is likely open for a new contract.

Taking situation 1 as an example and comparing it to situation 3, we can reason that if we do not consider the gain in value at the EOH, then an optimal decision would be to bunker the minimum required order. Such a decision is then reduced to a search space with the goal of finding a cheap bunker port with no excessive waiting times. There is, however, a win for the decision maker if the next voyage order has no such cheap option (e.g. transiting the Suez Canal to the Red Sea) and they have some additional fuel to counter the situation. However, this is beyond the end of our horizon.

2.6.2.3 Laycan and voyage speed effect (Situation 4 and Situation 5)

The laycan is an important information for chartering negotiations. For a voyage to be accepted, a shipowner will verify whether the vessel can realistically reach the load port on time and the fuel quantity they would need to do so. The exponential speed-consumption relationship (c.f. Figure 3) implies that a pressing laycan will have important implications for the voyage order and the bunkering plan. First, a need for larger fuel orders to compensate for the higher consumption caused by the required higher speeds. Second, the resulting longer servicing times in the bunkering port caused by such larger orders. Third, the possibility that the vessel is unable to reach the farthest bunkering ports at the required high speeds with the initial fuel allocation onboard.

For the case of a pressing laycan $(R = 380\ \text{hours})$ in situation 4, the results in Table 7 reflects the need for speeding up and the decision to reach the bunkering port offering the lowest waiting time.

Table 7

Numerical results in situation 4 and situation 5 under different risk level assumptions

	Situation 4 $\psi = 23.7$	**Situation 5**	**Situation 4** $\psi = 22.4$	**Situation 5**
Risk-aversion (stage2, stage3)	**(0.95,0)**		**(0.95,0.95)**	
First stage				
Port - speed	Gibraltar 14.4	Valetta 7.0	Gibraltar 14.3	Piraeus 7.0
Forward fuel contracted	187.7	292.3	611.1	237.0
Second stage				
Average spot fuel contracted	1,899.9	2,121.6	1,024.8	1,298.2
Average unused forward fuel	0	0	0	0
Third stage				
Average voyage consumption	311.0	99.7	308.7	103.1
Average unused spot fuel	109.4	180.8	39.1	62.2
Average fuel at EOH	1,057.2	1,523.4	678.2	763.6
Expected total profit	668,168.51	763,030.72	655,903.49	739,826.78
Total cost standard deviation	68,835.90	71,850.61	56,244.95	55,537.70

The decision to go Gibraltar is only possible if we permit to top up additional first stage fuel ψ at the current spot price c^{ψ} at the port of origin. By adding fuel, the vessel can sail at higher speeds to the more distant (cheaper) bunkering ports. This extension to our model could be interpreted as the shipowner requesting bunkers to be pumped while discharging its previous cargo at the time of negotiating a next voyage.

If no topping up is allowed at Iskenderun or too much fuel is needed, then the decision will be to go to Valetta at 14.1 knots and maintain safe tank levels as per Constraints (4) and (5). Note that it is feasible to reach Valetta at such higher speeds as it is closer than Gibraltar (c.f. Figure 2).

When risk aversion is assumed on stage 3, the spot order for situation 4 is reduced in the second stage and a larger order is secured from a forward contract. This result suggests that with due consideration of the third stage EOH price and waiting time risk,

the vessel is better off by securing more fuel from the forward contract than speculating on spot fuel. This is the case as the spot fuel prices at the EOH is likely going to be lower than the spot price for the purchased fuel, representing a loss according to our assumptions. Additionally, as a smaller order is secured, then the vessel can marginally save some fuel by reducing its speed.

The case of a relaxed laycan $(R = 800\ hours)$ in situation 5 highlights the opposite situation. With less time pressure, the decision is to bunker in cheap ports at the slowest permitted speed. An additional feature in this setting is that the slow speed results in small differences in bunker consumptions between Piraeus and Valetta, despite the difference in deviation.

The results in these experiments suggest that the waiting time at a bunkering port has a stronger effect when there is a shorter time remaining to reach the laycan. While this may be an obvious result, a less obvious interpretation is that at higher speeds, the bunkering options are reduced to ports that are closer to the point of origin. Additionally, as the distance to a port increase in the first leg, the higher the speed and consumption is needed to reach it within our model constraints.

2.6.2.4 Freight rate effect (Situation 6 and Situation 7)

For situation 6 and situation 7, we observe in Table 8 that the decision in all cases is to bunker in Gibraltar. For situation 6 (high freight rate), bunkering at Gibraltar results from the relative value of time. From our TCE calculation (Equation 32), we note that a higher freight rate increases the TCE; therefore, late arrivals are penalized more strongly. A higher penalty for delays promotes short stays at port, so the effect of waiting times is stronger in high freight markets. A low standard deviation when compared to the expected profit, suggests that in strong freight markets a strategy that reduces bunkering costs is not as important as compared to a low freight market.

Table 8

Numerical results in situation 6 and situation 7 under differerent risk level assumptions

	Situation 6	Situation 7	Situation 6	Situation 7
		$*$ **MIP Gap** $\mathbf{5.5e^{-3}}$		
Risk-aversion (stage2, stage3)	**(0.95,0)**		**(0.95,0.95)**	
First stage				
Port - speed	Gibraltar 12.4	Gibraltar 9.5	Gibraltar 13.2	Gibraltar 9.5
Forward fuel contracted	225.8	0	606.2	505.2
Second stage				
Average spot fuel contracted	1,777.6	2,058.3	967.5	964.6
Average unused forward fuel	0	0	0	0
Third stage				
Average voyage consumption	270.2	159.4	267.6	158.9
Average unused spot fuel	102.5	80.4	59.1	24.8
Average fuel at EOH	1,790.7	1,208.6	637.0	676.2
Average time passes laycan (φ)	**0.6**	**162.4**	**1.0**	**161.8**
Expected total profit	1,574,834.74	273,076.80	1,557,252.43	255,985.40
Total cost standard deviation	62,810.04	77,255.16	52,251.85	60,066.36

For situation 7 (low freight rate), the decision is to bunker at Gibraltar, though there are only slight differences between Valetta[6] and Gibraltar in terms of expected total profit. This result is similar to situation 2 when no waiting time is considered. This suggest that the waiting time effect is negligible in low freight markets as the value of time is low

If risk neutrality is assumed on stage 3, we observe that speculating with a large spot order and not taking fuel from a more expensive (in expectation) forward contract compensates for an expected low profit in a low freight market. Assuming risk aversion on stage 3 results in a lower standard deviation, signaling a better control of scenarios

[6] Valetta 9.1 knots objective value of 254,822.20 for Situation 7 (0.95,0.95)

with lower EOH fuel prices by increasing the forward bunkers order. Also, we observe a sharp decrease in the selected speed. This speed decrease results in longer voyage times, signaling the low value assigned to time.

2.7 Conclusions

In this study, we modeled the bunkering management problem for a tramp vessel in a voyage charter with due consideration of stochastic waiting times and stochastic fuel prices. We included voyage features, such as draft restrictions and weather conditions, and made use of contractual information as part of the problem. Additionally, we introduced risk measures to our multistage problem with the use of time consistent CVaR constraints. From this framework, we assessed the effect that stochastic waiting times have on the bunkering decision.

Our results suggest that for the route examined in our case study, it does not make much difference to bunker in Valetta or Gibraltar given their “ideal” conditions for bunkering, i.e. low prices and short waiting times. Nonetheless, the various situations considered in our empirical results illustrate that when there is large uncertainty in waiting times, a decision driven solely by price could lead to significant losses. This is perhaps best highlighted by the observation that a traditional bunkering port such as Piraeus, which is an attractive option based on its price alone, can lead to losses due to adverse waiting times. With the benefit of hindsight, we acknowledge that an empirical case study based on a different geographical area with lower spatial integration of prices might have yielded a better illustration of the power of our model.

We identified that the price of fuel and the deviation from the shortest route are the main factors for selecting a bunkering port. This is the same logic that would have resulted from a model with no waiting time. However, the waiting times add a new dimension to the problem, where there is a stronger effect for a bunkering port with low price if it is disproportionally less efficient than its competitors. We also show that with a tight laycan, the level of initial fuel on board become important. With a tight laycan, bunkering options are limited to nearby ports, as the high speed needed to compensate for the time at port and, hence, the high fuel consumption, does not allow for longer distances. Finally, we highlight the importance of a hedging strategy that accounts for uncertainty which is manifesting itself in different stages (time consistency). In our

case, neglecting the time consistency is equivalent to not acknowledging the uncertain waiting times and EOH prices as individual risks from the spot fuel order. This results in a speculative strategy where the second stage decision is such as to expect a large profit from the EOH fuel price.

Shipping practitioners can make use of our model to evaluate bunkering options and introduce the waiting time as a parameter in their bunkering procurement planning. Therefore, we propose that evaluating waiting times at bunkering ports must become part of standard practice. For implementation purposes, we propose that shipowners collect the underlying waiting time distribution from agents' information or specialized software, and then base their decision on the tradeoff between the price differences and the waiting time differences valued by the prevailing TCE for the vessel. A midway solution, when no information of the underlying waiting time is available, is to hedge against potential large losses (from stage 3) via forward fuel contracts and make smaller orders on spot contracts.

The proposed model has some limitations, mainly related to the underlying assumptions. First, we recognize that a linear penalty for surpassing the laycan could be refined by a more elaborate function that penalizes late arrivals as a non-linear function. This would better represent the real-life case, where minor infractions pass without penalty and larger delays disproportionally increases the chances of contract cancellation. Second, we assume a constant bunker barge pumping rate. This could be refined by introducing pumping rates per port, as the service provider in a port could compensate for longer waiting times with faster pumping rates (servicing times). There is also the argument that faster pumping rates, and therefore quicker turnarounds per vessel, would lead to lower waiting times. Third, the TCE could be modeled as a function of several decision variables (e.g. cargo quantity, fuel quantity, fuel port, etc.) as part of stage 3. Note that such a transformation results in a non-linear model.

Our case study results are limited to a single route and a single vessel type. In general, the bunkering decision is conditional on the deviation distance from the main route and the exposure of the vessel type to the weather elements (e.g., a vehicle carrier is more exposed to wind than it is to waves and currents). An additional observation is that waiting time distributions could reflect more recent events or seasonal events. We use the historical waiting times estimation from January 2015 to January 2018 and it could

be that an initial inefficient service was later compensated by a better service not captured in our data. We are therefore cautious to generalize the results of this study, which are based on specific conditions.

Future research should expand this study to consider stochastic fuel consumption based on weather. A potential way forward is expanding our model to a 4-stage problem where, instead of knowing beforehand the bunker volume to order, the shipowner could use some weather expertise and plan on the basis of seasonal and forecasted weather. Such a problem implies larger trees and potentially the use of heuristic approaches, such as the Progressive Hedging algorithm as proposed by Rockafellar and Wets (1991).

Additionally, the relation of the forward/spot contracts is a consequence of our contango futures market assumption. A future study could evaluate the bunkering decision on different forward-spot contracts margins with real forward price information.

References

Alizadeh, A. H., Kavussanos, M. G., & Menachof, D. A. (2004). Hedging against bunker price fluctuations using petroleum futures contracts: constant versus time-varying hedge ratios. *Applied Economics, 36(12)*, 1337-1353.

Aydin, N., Lee, H., & Mansouri, S. A. (2017). Speed optimization and bunkering in liner shipping in the presence of uncertain service times and time windows at ports. *European Journal of Operational Research, 259(1)*, 143-154.

Barbariol, F., Davison, S., Falcieri, F. M., Ferretti, R., Ricchi, A., Sclavo, M., & Benetazzo, A. (2021). Wind Waves in the Mediterranean Sea: An ERA5 Reanalysis Wind-Based Climatology. *Frontiers in Marine Science.*

Besbes, O., & Savin, S. (2009). Going Bunkers: The Joint Route Selection and Refueling Problem. *Manufacturing & Service Operations Management 11(4)*, 694-711.

Besse, P. C., Guillouet, B., Loubes, J. M., & Royer, F. (2016). Review and perspective for distance-based clustering of vehicle trajectories. *IEEE Transactions on Intelligent Transportation Systems, 17(11)*, 3306-3317.

Bushaj, S., Büyüktahtakın, İ. E., & Haight, R. G. (2022). Risk-averse multi-stage stochastic optimization for surveillance and operations planning of a forest insect infestation. *European Journal of Operational Research, 299(3)*, 1094-1110.

Conn, J., & Ferguson, A. (1968). Results obtained with a series of geometrically similar models. *Transactions of the Royal Institution of Naval Architechts, 93*, 255-230.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische mathematik, 1(1)*, 269-271.

Fan, H., Yu, J., & Liu, X. (2019). Tramp Ship Routing and Scheduling with Speed Optimization Considering Carbon Emissions. *Sustainability, 11(22)*, 6367.

Fuentes, G. (2021). Generating bunkering statistics from AIS data: A machine learning approach. *Transportation Research Part E: Logistics and Transportation Review, 155*, 102495.

Fuentes, G., & Adland, R. (2020). A spatial framework for extracting Suez Canal transit information from AIS. *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 586-590.

Gencer, M., & Unal, G. (2012). Crude Oil price modelling with levy process. *International Journal of Economics and Finance Studies, 4(1)*, 139-148.

Ghosh, S., Lee, L. H., & Ng, S. H. (2015). Bunkering decisions for a shipping liner in an uncertain environment with service contract. *European Journal of Operational Research, 244(3)*, 792-802.

Gu, Y., Wallace, S. W., & Wang, X. (2019). Integrated maritime fuel management with stochastic fuel prices and new emission regulations. *Journal of the Operational Research Society, 70(5)*, 707-725.

Hart, W. E., Laird, C. D., Watson, J. P., Woodruff, D. L., Hackebeil, G. A., Nicholson, B. L., & Siirola, J. D. (2017). *Pyomo-optimization modeling in python (Vol. 67).* Berlin: Springer.

Harvald, S. (1983). *Resistance and Propulsion of Ships.* Wiley.

Homem-de-Mello, T., & Pagnoncelli, B. K. (2016). Risk aversion in multistage stochastic programming: A modeling and algorithmic perspective. *European Journal of Operational Research, 249(1)*, 188-199.

IMO. (2020). *Fourth IMO Greenhouse gas study.* London: International Maritime Organization.

ITTC. (2011). *Fresh water and seawater properties.* Zurich: International Towing Tank Conference.

ITTC. (2014). *Analysis of Speed/Power Trial Data.* Zurich: International Towing Tank Conference.

ITTC. (2017). *Resistance Test.* Zurich: International Towing Tank Conference.

ITTC. (2021). *General guideline for uncertainty analysis in resistance test.* International Towing Tank Conference.

Kaut, M. (2014). A copula-based heuristic for scenario generation. *Computational Management Science, 11(4)*, 503-516.

Kaut, M., & Wallace, S. W. (2007). Evaluation of scenario-generation methods for stochastic programming. *Pacific Journal of Optimization*.

Kaut, M., & Wallace, S. W. (2011). Shape-based scenario generation using copulas. *Computational Management Science, 8(1)*, 181-199.

Kim, H. J. (2014). A lagrangian heuristic for determining the speed and bunkering port of a ship. *Journal of the Operational Research Society, 65(5)*, 747-754.

Knueven, B., Mildebrath, D., Muir, C., Siirola, J. D., Watson, J. P., & Woodruff, D. L. (2020). A Parallel Hub-and-Spoke System for Large-Scale Scenario-Based Optimization Under Uncertainty. *Preprint. Optimization Online*.

Kristensen, H. O., & Lützen, M. (2013). Prediction of resistance and propulsion power of ships. *Clean Shipping Currents, 1(6)*, 1-52.

Liu, M., Liu, X., Chu, F., Zhu, M., & Zheng, F. (2020). Liner ship bunkering and sailing speed planning with uncertain demand. *Computational and Applied Mathematics, 39(1)*, 22.

Meng, Q., Wang, S., & Lee, C.-Y. (2015). A tailored branch-and-price approach for a joint tramp ship routing and bunkering problem. *Transportation Research Part B: Methodological, 72(1)*, 1-19.

Oh, H., & Karimi, I. (2010). Operation planning of multiparcel tankers under fuel price uncertainty. *Industrial & Engineering Chemistry Research, 49(13)*, 6104-6114.

Panayides, P. (2018). *Principles of Chartering (pp. 174-180).*

Prochazka, V., & Adland, R. (2019). Ocean Mesh Grid: Applications in Shipping Modeling. *2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 330-334.

Rehmatulla, N., & Smith, T. (2015). Barriers to energy efficiency in shipping: A triangulated approach to investigate the principal agent problem. *Energy Policy, 84*, 44-57.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of risk, 2*, 21-42.

Rockafellar, R. T., & Wets, R. J. (1991). Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of operations research, 16(1)*, 119-147.

Ronen, D. (1982). The effect of oil price on the optimal speed of ships. *Journal of the Operational Research Society, 33(11)*, 1035-1040.

Sheng, X., Chew, E. P., & Lee, L. H. (2015). (s,S) policy model for liner shipping refueling and sailing speed optimization problem. *Transportation Research Part E: Logistics and Transportation Review, 76*, 76-92.

Sheng, X., Lee, L. H., & Chew, E. P. (2014). Dynamic determination of vessel speed and selection of bunkering ports for liner shipping under stochastic environment. *OR Spectrum, 36(2)*, 455-480.

Vilhelmsen, C. (2014). Tramp ship routing and scheduling with integrated bunker optimization. *EURO Journal on Transportation and Logistics, 3(2)*, 143-175.

Wang, C., & Chen, J. (2017). Strategies of refueling, sailing speed and ship deployment of containerships in the low-carbon background. *Computers & Industrial Engineering, 114*, 142-150.

Wang, S., & Meng, Q. (2015). Robust bunker management for liner shipping networks. *European Journal of Operational Research 243(3)*, 789-797.

Wang, Y., Meng, Q., & Kuang, H. (2018). Jointly optimizing ship sailing speed and bunker purchase in liner shipping with distribution-free stochastic bunker prices. *Transportation Research Part C: Emerging Technologies 89*, 35-52.

Wang, Y., Meng, Q., & Tan, Z. (2018). Short-term liner shipping bunker procurement with swap contracts. *Maritime Policy & Management*, 211-238.

Worldscale Association (London) Limited & Worldscale Association (NYC) Inc. (2021, 06 05). *Basis of Calculation.* Retrieved from Worldscale: https://www.worldscale.co.uk/help#t5

Yao, Z., Ng, S. H., & Lee, L. H. (2012). A study on bunker fuel managemnent for the shipping liner services. *Computers & Operations Research 39(5)*, 1160-1172.

Zis, T. P., Psaraftis, H. N., & Ding, L. (2020). Ship weather routing: A taxonomy and survey. *Ocean Engineering, 213*, 107697.

Appendix

Estimating weather fuel consumption

Our estimations are based on the resistance calculations of a Supramax vessel, with characteristics (refer to Table A.1) derived from a general arrangement and capacity plan of a vessel, kindly supplied by a shipping operator. In this section, the notation **is not** related to the bunkering model presented in the main text of this study. We assume the total resistance is the sum of calm water resistance and appended resistance to wind, currents and waves.

Table A.1

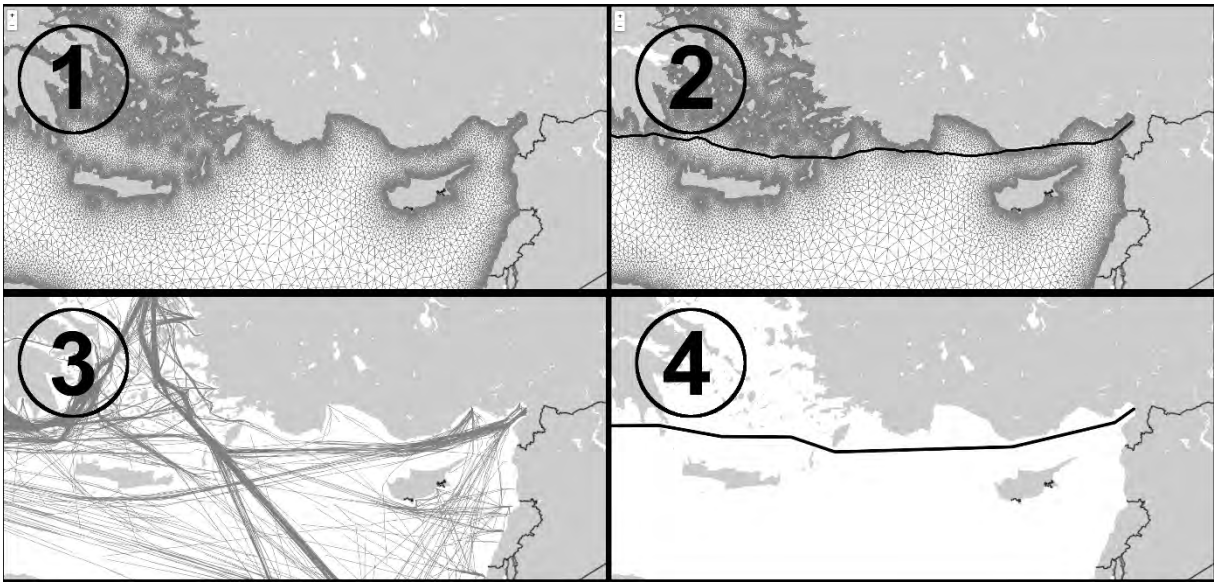
Case study Supramax characteristics

Notation	Description	Value
L	Length between perpendiculars	183.3 m
Δ	Displacement	$\Delta_o = 65595$ tonnes
D	Depth	18.5 m
P	Beam	30.4 m
K	Kilowatt-hour of vessel at design speed (16.7 knots) 100% $Kw\ MCR$	$8{,}562$ mKw
H^l	Tonnes per centimeter of immersion (laden)	57
H^b	Tonnes per centimeter of immersion (ballast)	52
A^S	Traverse superstructure area	272 m^2
d	Vessel draught (Salt water $1025.9 kg/m^3$)	$d_0 = 6.9$ m

To calculate the effects of the underlying environmental forces on the vessel, we assume that the vessel positions along the route can be derived based on the speed and distance relationships described below. The procedure can be visually represented in Figure A.1. In step 1, the navigable area is populated with nodes and edges feeding a graph. This is motivated by the ocean mesh concept described by Prochazka and Adland (2019). Step 2 retrieves the shortest route from a sequence of nodes via the Dijkstra shortest path algorithm (Dijkstra, 1959). The resulting shortest path is thereafter smoothed in step 3 by comparing it to historical routes (i.e. sequences of AIS positions) of vessels of the same type with a similar draught and retrieving the closest route based on their Symmetrized Segmented Path Distance (step 4). The algorithm to calculate this distance is introduced by Besse et al. (2016) and implemented by Fuentes and Adland (2020) for the interpolation of shipping routes.

Figure A.1

Shortest route generation method



Having pre-defined the shortest route, it is possible to interpolate positions every hour based on different speed selections. The route is interpreted as $n_t \in N(z)$, where n_t is a node defined by latitude and longitude, with a subscript t indexing its hourly increment from a set T, and $N(z)$ is a set of ordered nodes defining the shortest route with speed $z \in \{7.0 \ldots, 14.5\}$[7].

The route is then iteratively compared to grids of wind and current data and filtered with the grids traversed by a ray $n_t \rightarrow n_{t+1}$. The grid values (wind and current vectors represented by u and v components) associated with $n_t \rightarrow n_{t+1}$ are assumed[8] to happen at t. Information on currents and wind on a six-hour frequency are collected from the 0.083° grid Operational Mercator Global Ocean Analysis database and the 0.25° grid Global Ocean Wind L4 database from the Copernicus Marine Environment Monitoring Service.

With all necessary input on hand, new vessel variables are transformed for every $n_t \rightarrow n_{t+1}$, as:

[7] Speed z is assumed to be fix. It could be intrepreted as RPM adjustments to mantain speed regardless of weather.
[8] The u and v components are averaged if more than one grid is crossed -i.e., due to routes interpolated by higher speed.

$$h_t(n_t, n_{t+1}) = atan2(X, Y)\left(\frac{180°}{\pi rad}\right) \quad \text{(A.1)}$$

where h_t is the vessel heading and:

$$\mathrm{X} = \cos lat_t \ sin|lon_{t+1} - lon_t| \quad \text{(A.2)}$$

$$\mathrm{Y} = \cos lat_t \, sinlat_{t+1} - \ sinlat_t \ \cos lat_{t+1} \ cos|lon_{t+1} - lon_t| \quad \text{(A.3)}$$

Wind (w) and current (x) variables are transformed as:

$$Z_{it} = 1.94384 \frac{\mathrm{kts}}{\mathrm{m/s}} \sqrt{u_{it}^2 + v_{it}^2}, \forall i \in \{w, x\}; \forall t \in T \quad \text{(A.4)}$$

$$\phi_{it} = atan2(u_{it}, v_{it})\left(\frac{180°}{\pi rad}\right), \forall i \in \{w, x\}; \forall t \in T \quad \text{(A.5)}$$

where Z is the true speed in $knots$, ϕ is the direction of the ocean current c or wind w, u is the zonal velocity and v is the meridional velocity.

The wind, currents and waves (q) relation with the vessel's speed and heading are described as:

$$B_{it} = \begin{cases} \phi_{it} - h_t & , \phi_{it} - h_t \geq 0 \\ 360 + \phi_{it} - h_t & , o.w \end{cases}; \forall i \in \{w, x, q\}; \forall t \in T \quad \text{(A.6)}$$

$$Z_{it}^A = \sqrt{Z_{it} + z + 2Z_{it} z \ cosB_{it}}; \forall i \in \{w, x\}; \forall t \in T \quad \text{(A.7)}$$

$$\phi_{it}^A = acos\left(\frac{Z_{it} \cos B_{it} + z}{Z_{it}^A}\right); \forall i \in \{w, x\}; \forall t \in T \quad \text{(A.8)}$$

where B is the angle between the true wind/current/wave direction and the vessel's heading, Z^A is the apparent speed, ϕ^A is the apparent direction and z is the vessel's speed in $knots$.

The vessel's calm water resistance R^c is estimated as:

$$R_t^c = \frac{C^c \rho_{sw} z^2 W_t^x}{2} \ ; \forall t \in T \quad \text{(A.9)}$$

where C^c is the total resistance coefficient, ρ^{sw} is the salt water density $\left(1025.9 \frac{\mathrm{kg}}{\mathrm{m}^3}\right)$, and W^x is the wetted hull area estimated from the Denny-Mumford equation (Kristensen & Lützen, 2013):

$$W_t^x = \frac{\Delta_t}{d_t} + 1.7(d_t L)\ ; \forall t \in T \tag{A.10}$$

The function to estimate C_t according to Kristensen and Lützen (2013) is:

$$C^c = C^f + C^{aa} + C^r + C^{as} \tag{A.11}$$

where C^f is the frictional coefficient, C^r is the residual drag coefficient, C^{aa} is the air resistance coefficient, C^{as} is the correction for the steering resistance coefficient. C^f is a function of Re the Reynold's number and γ is the coefficient of kinematic viscosity. The coefficient and relevant parameters are summarized in Table A.2.

Table A.2

Coefficients used for the vessel's consumption estimation

Coefficient	**Reference**
$C^r = \begin{cases} 0.0940, & \text{ballast} \\ 0.0016, & \text{laden} \end{cases}$	(Harvald, 1983)
$C^f = \frac{0.075}{(\log_{10} Re - 2)^2}$	(ITTC, 2021)
$C^{as} = 0.041e^{-3}$	(Harvald, 1983)
$C^{aa} = 0.071e^{-3}$; for Handymax	(Kristensen & Lützen, 2013)
$Re = \frac{zL}{\gamma}$	(ITTC, 2021)
$\gamma = 1.191e^{-6} m^2 s^{-1}$; for saltwater at 15°C	(ITTC, 2011)

Appended resistance from wind R^w is estimated from:

$$R_t^w = \cos(\phi_{tc}^A)\, C^{aa} \frac{\rho^a (Z_{wt}^A)^2}{2} W_t^a; \forall t \in T \tag{A.12}$$

where ρ^w is the air density $\left(1.225 \frac{\text{kg}}{\text{m}^3}\right)$, W^w is the vessel area exposed to air and can be estimated as:

$$W_t^w = A^S + P(D - d_t); \forall t \in T \tag{A.13}$$

Appended resistance from current R^x is estimated from:

$$R_t^x = \cos(\phi_{ct}^A)\, C^v \frac{\rho^{sw}(Z_{ct}^A)^2}{2} W_t^x; \forall t \in T \tag{A.14}$$

C^v is the coefficient of viscous resistance and can be estimated as per ITCC (2017) as:

$$C^v = C^f(1+k) \tag{A.15}$$

and k is the form factor estimated as per Conn and Ferguson (1968):

$$k_t = 18.7\left(C_b \frac{d_t}{L}\right)^2 \tag{A.16}$$

Appended resistance due to waves R_t^q is based on the STAwave-2 empirical method (ITTC, 2014) which main function is summarized as follows:

$$R_t^q = 2\int_0^\infty \frac{R_{AWM}(\omega,\zeta_t) + R_{AWR}(\omega,\zeta_t,z)}{\zeta_t^2} S(\omega)d\omega; \forall t \in T \tag{A.17}$$

where R_{AWM} is the vessel motion induced resistance, R_{AWR} is the increase due to wave reflection, ω is the circular wave frequency, ζ is the wave amplitude and S is the frequency spectrum for ocean waves.

Thereafter, required power Pe is estimated as:

$$Pe_t = \frac{z(R_t^c + R_t^x + R_t^q)}{1000}; \forall t \in T \tag{A.18}$$

Engine load j is:

$$j_t = \frac{\mathrm{Pe_t}}{K}; \forall t \in T \tag{A.19}$$

Specific Fuel Oil Consumption $SFOC^l$ as a function of load (IMO,2020) is estimated as:

$$SFOC_t^l = SFOC^b(0.455j_t^2 - 0.71j_t + 1.28)\ ; \forall t \in T \tag{A.20}$$

and SFOC base value as derived by IMO GHG 4th (2020) is $175\,\frac{\mathrm{g}}{\mathrm{kwH}}$

Fuel consumption F_t can then be estimated as:

$$F_t = Pe_t SFOC_t^l\ ; \forall t \in T \tag{A.21}$$

Chapter 3

Greenhouse gas mitigation at maritime chokepoints: The case of the Panama Canal*

Gabriel Fuentes [a,b], Roar Adland [b]

[a] Centre for Applied Research at NHH(SNF)
[b] Department of Business and Management Science, NHH

Abstract

In this study, we investigate the impact on shipping emissions from improving operational efficiency in a maritime chokepoint such as a canal. We consider several scheduling proposals that allow for different levels of speed reduction for incoming vessels and estimate the resulting emission reduction compared to a benchmark established from ship position data. For our case study of the Panama Canal, we estimate that the canal could have removed up to 1.8 million tonnes of CO_2e per year for the period of 2019 to 2021. Our findings suggest that emission reduction can be easier to obtain at intermediate points as many of the contractual barriers to improving operational efficiency do not apply.

Keywords: climate action, GHG mitigation, maritime chokepoints, chokepoints efficiency

* This research was partly financed by the Research Council of Norway under the project “Smart digital contracts and commercial management”, project number 280684. The AIS data and big data handling resources were kindly provided by the United Nations Global Platform.

3.1 Introduction

According to the 4th International Maritime Organization (IMO) greenhouse gas (GHG) study (2020), shipping contributed to about 2.89% of global anthropogenic GHG emissions in 2018, equivalent to 1,056 million tonnes. Emissions generated by ships burning fossil fuels can have both a global effect on climate change, caused by Carbon Dioxide (CO^2), Methane (CH^4) and Nitrous Oxide (N_2O), and local effects on human health and the environment caused by Sulphur Oxides (SOx), Particulate Matter (PM) and Nitrogen Oxides (NOx).

In 2018, the IMO as the international shipping regulator proposed an initial strategy for reducing GHG emissions from international shipping by at least 50% by 2050, compared to the 2008 level (MEPC 72/17, 2018). To reach this goal, they acknowledge that the reduction can be based on both technical and operational energy efficiency measures. Technical measures focus on aspects relevant to the vessel's design and construction, e.g., hull design, use of alternative fuels, and improved propulsion. (Bouman et al., 2017). Operational measures improve energy efficiency by operational changes such as speed reduction, weather routing, and improved fleet management.

Speed reduction is an operational measure with a high emissions reduction potential (Bouman et al., 2017). An important feature of its implementation is that a reduction of waiting time in port can be used (in part or in full) to reduce the sailing speed of arriving vessels without increasing the duration of the voyage. Despite its theoretical simplicity and its high potential for emission reduction, its implementation at the vessel or voyage level is often hindered by barriers such as misaligned incentives and contractual constraints (Rehmatulla and Smith, 2015) leading to unproductive waiting times at port (Johnson and Styhre, 2015; Jia et al., 2017).

Until now, the challenges and benefits in implementing the concept of speed reduction due to operational efficiency has only been analyzed in the context of cargo-handling ports, that is, the end points of ships' voyages. However, unproductive waiting time may also occur at intermediate points during a voyage, such as during bunkering stops and at maritime chokepoints such as canals and straits. Such waiting times could be caused by factors such as dynamic demand, lack of resources to serve the vessel, weather closures and restrictions on throughput etc. As any unproductive waiting can be considered an opportunity for improving operational efficiency and implementing

speed reduction measures, thereby reducing fuel consumption and emissions, this focus on cargo-handling ports creates a gap in the literature. Arguably, some maritime chokepoints are so central to global seaborne trade that they are as important to consider as very large ports and key trade lanes.

A chokepoint refers to locations that limit the capacity of throughput and cannot be easily bypassed, if at all (Rodrigue, 2004). The primary contribution of this study is to expand the literature on emission mitigation from speed reduction due to improving efficiency to the case of maritime chokepoints, specifically the Panama Canal. There are at least two additional reasons why canals and straits should be considered separately from the literature on speed reduction and port efficiency. Firstly, as in the case of the Panama Canal, scheduling policies for transits (the equivalent of berthing during a port call) can be substantially more complex, involving not only a chronological queue of vessel arrivals but also Canal Authority discretion, pre-booked slots and rules determining priority by vessel type. Secondly, the centralization of authority matters for the ability to implement such measures successfully, compared to port operations where several stakeholders related to cargo, vessel and port interests determine the outcome. As a secondary contribution, we develop new algorithms, based on the processing of AIS vessel position data, to generate more accurate statistics for canal waiting times and transits by vessel type, size and direction.

The rest of the paper is organized as follows: Section 3.2 presents our literature review. Section 3.3 introduces our strategies of implementation with due consideration of the Panama Canal Authority (PCA) transit scheduling rules. Section 3.4 includes the methods for estimating the transits and emission information. In section 3.5, we present the results derived from the algorithm. In section 3.6 we discuss the implications of speed reduction implemented by a chokepoint and report the limitations of the proposals. Finally, Section 3.7 concludes and proposes extensions for future research.

3.2 Literature review

Air pollution from ships is regulated by the International Maritime Organization (IMO) via the International Convention for the Prevention of Pollution from Ships (MARPOL) Annex VI. Annex VI was the last of the annexes to enter into force (May 2005) from

MARPOL and initially focused on reducing SOx via Emissions Control Areas and NOx by implementing marine engine emission standards.

The literature on maritime emission reduction can broadly be divided into technical and operational measures. In a comprehensive review, Bouman et al. (2017) suggest that the highest individual abatement potential is observed for biofuels (technical) and speed optimization (operational). However, the highest overall emissions abatement effect requires combinations of individual reduction measures. They also acknowledge that there is low agreement among studies reporting on the abatement potential due to speed reduction, as evidenced by a larger observed variance compared to other measures.

The potential of speed related measures in the emissions abatement is a popular topic in the academic literature as evidenced by the 26 studies summarized by Bouman et al. 2017, in part because the measure has a low burden of implementation (Lindstad et al. 2011). Additionally, it is found to be a cost-effective measure as the emission reductions can be achieved with zero or negative marginal abatement cost (Eide et al., 2009; Corbett et al., 2009; Eide et al., 2011; Lindstad et al., 2011; Schwartz et al., 2020). However, Corbett et al. (2009) suggest that the real marginal abatement cost is higher than reported in some studies as more vessels are needed to maintain a certain service frequency or level of trade, therefore increasing the cost for operating a fleet. Regardless of the interpretation of cost effectiveness, there is consensus in the literature of the high emissions reduction potential from speed reduction measures.

Speed reduction could be enabled by an increased fleet size (Corbett et al., 2009; Cepeda et al., 2017) to compensate the loss in cargo transported or, alternatively, by reducing unproductive waiting times at port (Lavon and Shneerson, 1981; Psaraftis et al., 2009; Johnson and Styhre, 2015; Andersson and Ivehammar, 2017, Jia et al., 2017). Our focus is on the latter concept, named by Johnson and Styhre (2015) as *speed reduction due to port efficiency.* The large emission reduction potential from implementing this concept was presented by Jia et al. (2017) in the case of Very Large Crude Carriers. As extensions to their study, they encouraged the introduction of algorithms to recognize waiting times with higher accuracy and expanding studies to other vessel sizes and types. Slack et al (2018) also recognized the absence of time

related metrics in port operations and related this to the difficulties of gathering consistent data from different stakeholders.

Port efficiency has been extensively studied in the academic literature (see, Krmac and Mansouri, 2022, for a review) but mainly in the context of container port infrastructure (e.g. container throughput and vessel port calls). However, in our context, studies dealing with the barriers to increasing efficiency and the causes of inefficiency are more relevant. In the case of bulk carriers, Johnson and Styhre (2015) find that the main factors, in order of importance, are the port opening hours, early arrivals, congestion and clearance procedure, unspecific reasons, and waiting for pilot. Slack et al. (2018) suggest that the time in port could be attributed to local or regional factors, pointing out the role of port officials in the port waiting time. Poulsen and Sampson (2020) argue in a study of tankers port calls that dealing with a large group of stakeholders influences the causes of inefficiencies (e.g., waiting for berth, availability of cargo, waiting for surveyors, etc.) and the corresponding waiting times. Accordingly, Poulsen and Sampson (2020) observe that "best practices might be present when the same company owns terminals and cargoes and/or ships and ensures port stakeholder coordination".

There are several contractual barriers to the implementation of speed reduction in parallel with increasing port efficiency. Firstly, a vessel on a voyage charter currently has an incentive to tender the Notice of Readiness (NOR)[1] to load as early as possible irrespective of the port availability to serve the vessel (Rehmatulla and Smith, 2015), both due to the First-in-first-out scheduling policy operated by most ports and to reduce the risk of contract cancellation due to late arrival (not meting the agreed laycan[2]). Poulsen and Sampson (2019) suggest that this rationale to "rush to wait" could also be attributed to the owners of ships engaged on voyage charters aiming for the counting of laytime to start as early as possible[3]. In theory, this increases the probability of claiming demurrage, yet the outcome is uncertain as the charterer can be so efficient

[1] NOR is a statement where the vessel declares it has arrived and is ready to commence cargo operations at the charterer's disposal. It is an important declaration as it commences the laytime counting and is relevant to demurrage calculations.
[2] Laycan is short for Laydays/Cancelling and is an agreed range of days where the vessel must be at port.
[3] Laytime in voyage charters is the time allowed to charterers for loading and/or unloading the vessel. The laytime starts counting from the NOR been tendered. If the laytime is passed, the owner is entitled to claim for demurrage (damages per delay).

with cargo handling that demurrage is not payable. For a time-chartered vessel, they argue that the behaviour could be due to the perceived reduction in uncertainty of having a buffer of vessels ready in port. Note that this is the only barrier specified for time-chartered vessels.

A second contractual barrier is the 'utmost dispatch' clause for voyage charters (Rehmatulla and Smith, 2015; Jia et al., 2017; Global Industry Alliance, 2020) that instructs the vessel to proceed to a destination port without delay[4]. Consequently, an owner is not generally free to reduce the sailing speed on the laden voyage in the interest of reducing emissions or fuel costs, unless an exemption clause exists to that effect.

The third barrier is the split incentives problem arising from the principal-agent relationship in a voyage charter (Rehmatulla and Smith, 2015). Under such a contractual arrangement, the shipowner bears the cost of fuel and, therefore, has an incentive to increase operational efficiency. However, the charterer has no other concern for the vessel speed than the vessel arriving on time. We note that there is no split incentive in the time charter case, as the charterer is the one benefiting from the speed reduction as they cover the fuel cost.

To our knowledge, there is not yet any literature on the impact and implementation of slow steaming as a consequence of increasing efficiency in maritime chokepoints such as canals and straits.

3.3 Strategies for speed reduction on voyages to the Panama Canal

Our proposals for implementing speed reduction measures are based on suggested changes to the existing Panama Canal scheduling rules. Therefore, in this section, we review the relevant scheduling rules from the PCA and the general modeling assumptions applied to our proposals. Thereafter, we present our four proposals for the speed reduction of inbound vessels

[4] Dockray (2013) review the most notorious cases where the utmost dispatch clause was disputed.

3.3.1 PCA transit scheduling rules

The scheduling system for the Panama Canal transits is based on vessels assignment within two categories. 1) *Non-booked transits* and 2) *Pre-booked transits*. Vessels in the first category are scheduled at the PCA discretion, conditional on the number of pre-booked vessels and various operational factors (e.g. vessel type mix, daylight transits restrictions). The vessels in the second group (pre-booked vessels) pay an extra fee for ensuring a slot on a particular transit date, with their transit time on the day subject to operational discretion. Note that the last booking period closes two days before a transit date.

An important feature of the reservation system is that vessels with booked transits must arrive at the canal before an arrival cutoff. Vessels not arriving at the cutoff time must pay a penalty, which is a function of how late the vessel is and equivalent to up to 100% of the original booking fee. More importantly, they may not be able to re-book the same slot with the risk of waiting in a queue as a regular transit vessel.

According to the PCA reservation system rules (2021b), the time limit for arrival is at 2200 hours of the day before transiting for Neo-Panamax vessels except LNG's, 0200 hours for LNG Neo-Panamax and Super/Regular[5] with restrictions and 1400 hours of the same day for Regular vessels. Also, in the transit reservation rules, a vessel becomes an arrived vessel for a booked transit when the vessel is physically sighted or is within eight nautical miles from either of the canal entrances. No reference is made to regular transit vessels; therefore, we assume that such "physical" arrival also is their trigger for being included in a day's transit schedule.

The scheduling rules make reference to priority based on vessel types or a customer ranking. For example, cruise vessels can book a reservation up to two years in advance while other vessel types can book at the earliest, one year in advance (Panama Canal Authority, 2021b). This implies that if policies need to be implemented in steps, then the PCA could easily segment their strategy (e.g., by market, vessel type or customer ranking).

[5] Vessels transiting the Panamax locks are subdivided for scheduling purposes in Supers higher or equal to 27.74m beam) and Regulars (lower than 27.74m beam). Neo-Panamax are vessels that can only transit through the Neo-Panamax locks or has a beam of more than 32.62m

3.3.2 General modeling assumptions

There are two important modeling assumptions that are common across all of our strategies. First, we assume that the speed reduction is only implemented on the voyage leg before the canal and that the historically observed transit time for a vessel does not change. This is simply due to the huge complexity of generating alternative scenarios for the dynamic fleet movements through the canal that would remain feasible given the real-life constraints that the PCA operates under (e.g., vessels mix, daylight limitation, no simultaneous crossing of Neopanamax vessels in opposite directions, daily availability of pilots and tugs, etc.). We acknowledge that this limits the emission reduction potential compared to an optimization model that had more degrees of freedom, including the rescheduling of vessel transits in time.

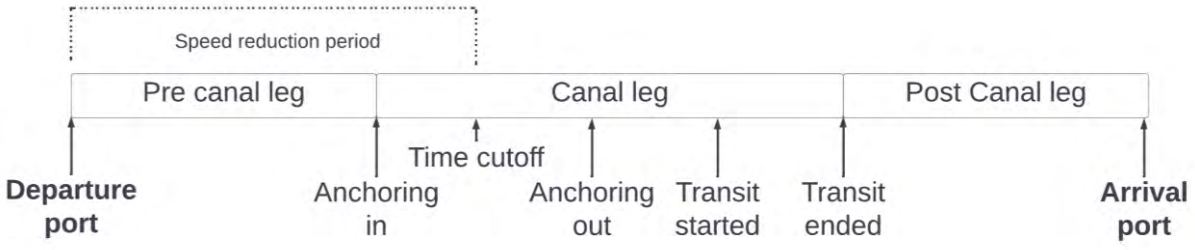
Second, we assume that a canal leg is composed of the interval of time that the vessel is within canal waters, starting with its wait at anchorage up to the time the vessel exits the canal locks at the opposing entrance. The canal waters are defined by the PCA (2019) and include anchorage, the access lanes and the transit lanes.

3.3.3 Proposal 1: Coordinated voyage

A coordinated voyage, as illustrated in Figure 1, implies that speed reduction is implemented for the full duration of the leg from the time vessels depart the last port until the arrival cutoff for the Panama canal transit as declared in the PCA transit reservation rules (2021b).

Figure 1

Coordinated voyage strategy principle



Arguably, the strategy is unrealistic, as the canal would be unable to assign transits for vessels given the large uncertainty in ETAs for long voyages. However, our motive for calculating the effects of this proposal is to investigate the limits of the speed reduction

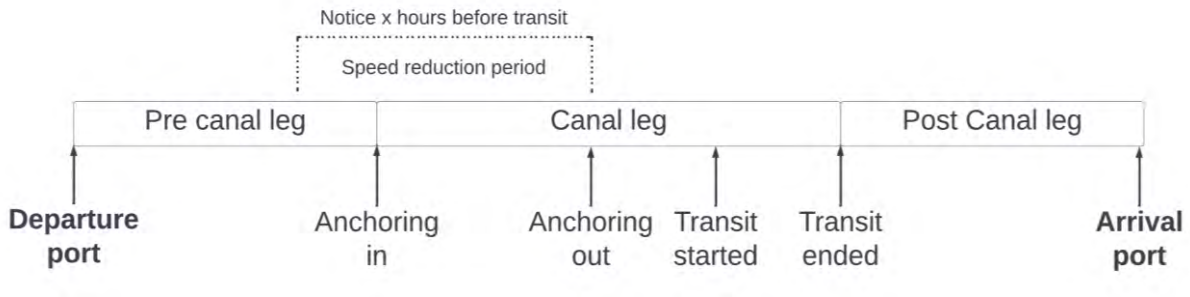
implementation and to estimate the theoretical upper bound based on our assumptions.

3.3.4 Proposal 2: Just-in-time (JIT) speed orders 48 hours before arrival

For a strategy based on JIT, a vessel is accepted as an arrived vessel without being sighted and instructed to reduce speed to arrive at a Required Time of Arrival (RTA). For the purpose of estimating the effect of this proposal, as illustrated in Figure 2, we assume the RTA is the time when the vessel leaves the anchorage prior to transit. We also assume there is no distinction between non booked and booked transits, and that vessels with less than 48 hours sailing reduce the speed upon leaving the last port.

Figure 2

JIT strategy principle



3.3.5 Proposal 3: Green slots

This proposal is based on the existing rules that the Panama Canal has for JIT arrivals. To avoid confusion with proposal 2, we refer to this proposal as ‘green slots’.

In the current scheduling system, the PCA has a limited number of green slots to assign per day. These slots are only assigned by request from a booked vessel that is capable of transiting immediately upon arrival[6]; therefore, it is natural that some of them remain unused. According to PCA rules, they can assign up to four green slots for *Supers* (max two per direction) and two for *Regulars* without restriction (max one per direction).

We propose that the PCA assigns the unused green slots based on a ranking of available vessels by the highest emission reduction. This highlights the benefit from involving fewer stakeholders (PCA and the vessel) compared to the challenge of

[6] This requires that the operational conditions are in place to transit the vessel directly on arrival (e.g. the vessel has free pratique, customs clearance, etc.).

implementing JIT in a port setting (Global Industry Alliance, 2020). We note that the difference between this and proposal 2 is that the green slots are allocated only to some vessels while the rest follow business as usual (i.e., arrive before an arrival cutoff or wait at anchorage for its transit assignment).

3.3.6 Proposal 4: Blend of JIT 24 hours before transit and green slots

This proposal is an alternative if the potential limitations in previous proposals (i.e. uncertainty of vessels' ability to comply with the RTA for JIT with 48 hour notice or not being within the rules for JIT assignment), hinder their implementation.

In this blended strategy, we propose that green slots be assigned as per proposal 3 and that the remaining vessels reduce their speed from 24 hours before arriving at the canal.

3.4 Methods

3.4.1 Input data

Our model relies on three sources of information: maritime geofences, including the Panama Canal and coastal lines; vessels specifications and AIS data.

Geofences[7] are virtual geographic boundaries that define an area and reduces the search space for recognizing a vessel operational phase (e.g. anchoring, at berth, maneuvering, slow transiting or normal cruising). For the Panama Canal estimations, we define the geofences including locks, ports and anchoring areas. Separately, for estimating emissions of the legs before and after the Canal, we make use of worldwide port positions from the World Port Index (WPI) (2019) and coastlines from the National Oceanic and Atmospheric Administration (2017).

Detailed vessel specifications are important in order to segregate our Panama Canal statistics by vessel type and transit type, to filter out non-transiting vessels such as Panama Canal tugboats, and to estimate emissions using a bottom-up method. Similarly to the 4th IMO GHG study (2020) we use the IHS Markit database, which includes over 100,000 records of vessel, including scrapped vessels.

[7] The term geofences and polygons are used interchangeably in this study

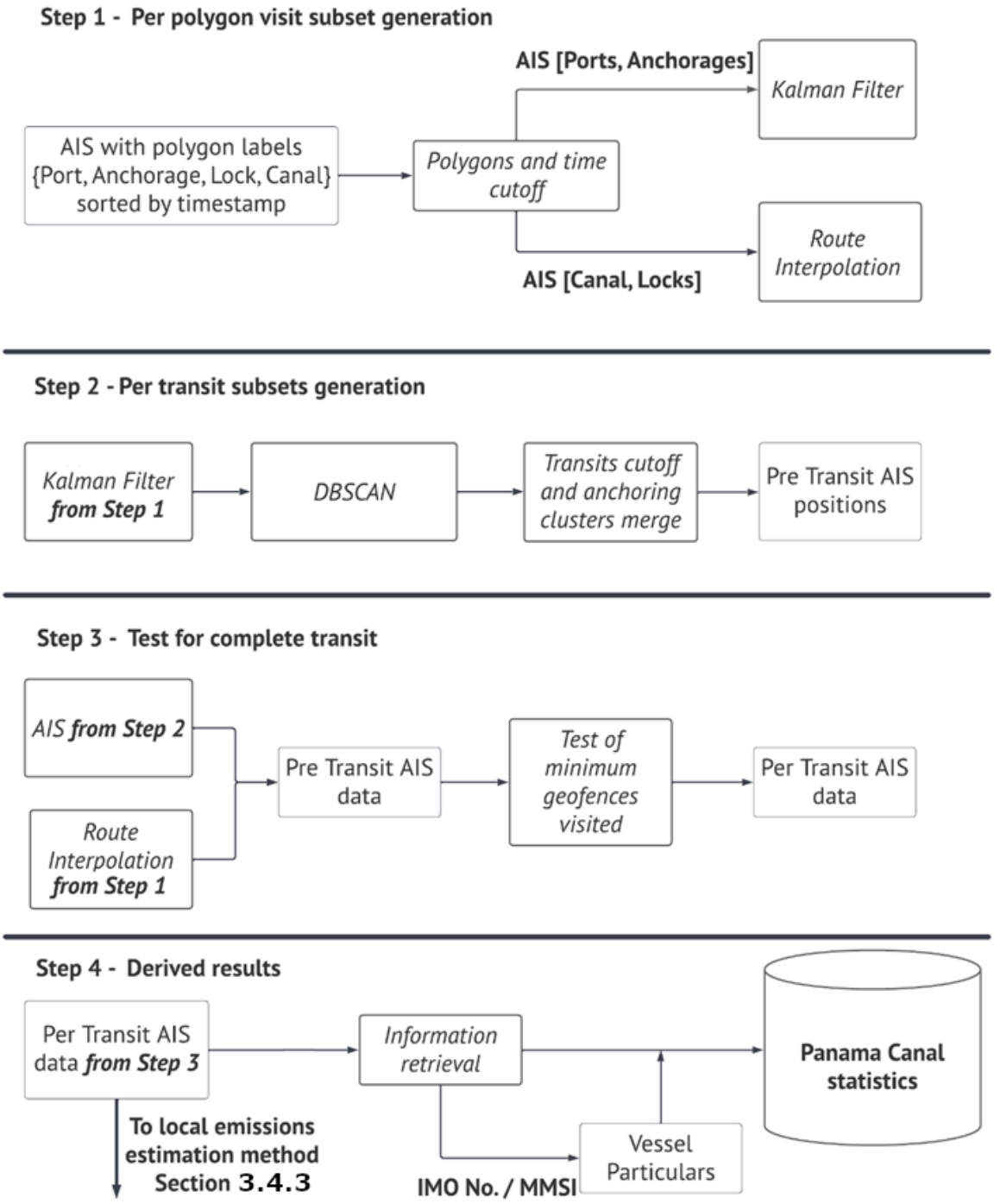
We use satellite and terrestrial AIS data from the UN Global Platform encompassing worldwide vessel records (3.6 TB) from January 2019 to December 2021. The data for the Panama Canal statistics is filtered by querying the records within our manually constructed geofences for the Panama Canal. Data records for emissions calculation before and after transiting are filtered with geofences based on the WPI records. Our filtered data includes approximately 8.5 billion records of dynamic AIS reports (i.e. latitude, longitude, timestamp, course over ground (COG), speed over ground (SOG), heading, and draught) and static AIS records (i.e. IMO Number, vessel type, length and width). The processing of such a big data set is made possible by an Apache Spark machinery and Apache Sedona.

3.4.2 Generating Panama Canal statistics

The algorithm used in this study is adapted from Fuentes and Adland's (2020) algorithm for recognizing transits in the Suez Canal. We estimate the waiting times by recognizing anchoring patterns before transits, and the transit times and time at ports are derived from the vessel track through the Canal and visits to the adjacent ports. We introduce our version of the algorithm for generating Panama Canal statistics in Figure 3.

Figure 3

Framework for generating Panama Canal statistics

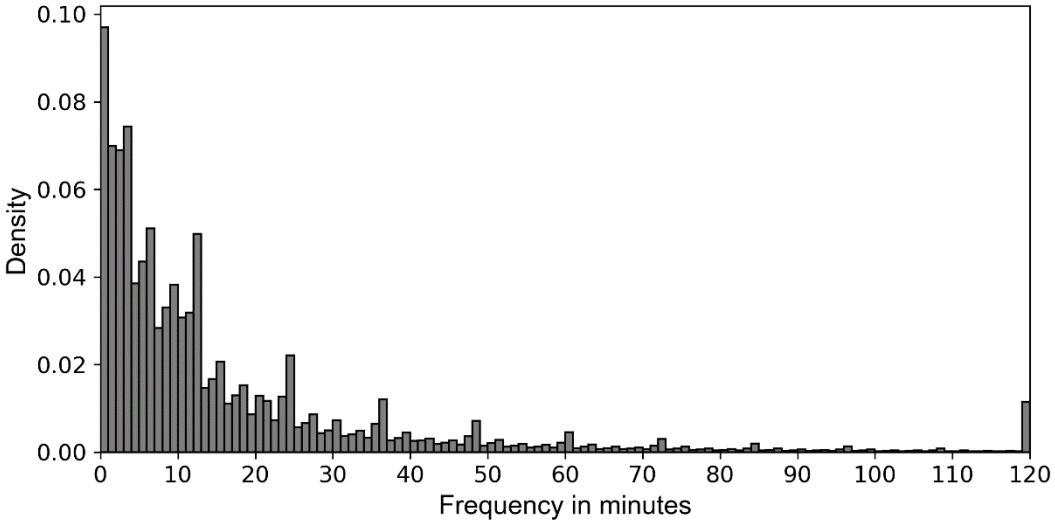


The first step (*polygons and time cutoff*) is to recognize when a vessel leaves and enters a polygon. A detailed description of the polygons and time cutoff test is presented in Algorithm 1 of Appendix A.

Also in step one; data is interpolated into fixed intervals of ten minutes. As shown in Figure 4, the frequency of positions, calculated as the difference between subsequent timestamps for the same vessel, is not fixed, causing some assumptions used in the algorithm, such as those for vessels swinging on the anchor, to be invalid if not tuned. A potential solution inside anchoring and port polygons is to do linear interpolation between observations; however, to make the best use of the data, we implement a Kalman filter corrected by positions, SOG and COG. The Kalman filter is an algorithm that estimates unknown variables given the measurements observed over time (Youngjoo & Hyochoong, 2018). For a detailed description, we refer to Kalman (1960).

Figure 4

Histogram of AIS time frequency between positions



Note. The rightmost vertical bar includes all frequency observations above 120 minutes.

Another challenge is to recognize the time of entering the Canal based on a vessel crossing the locks. A similar approach with Kalman filter has the risk of generating spurious estimations if the latency between concurrent positions is too large to detect fast turnarounds, as is commonly observed in canals. For interpolating the vessel records inside the canal and its locks, we take advantage of the limited navigable options inside a canal and use historical transit routes, a method described in Fuentes and Adland (2020).

After the frequency interpolation of anchorage records, in ***step two*** we derive the waiting times by recognizing the vessel at anchor using the Density-based Spatial

Clustering of Applications with Noise (DBSCAN) algorithm. The DBSCAN hyper parameter ϵ is calibrated from the assumption that external natural forces are producing the movement of the vessel swinging on the anchor. For simplicity, we derive for each anchoring area, the maximum values of the sum of historical tidal streams and ocean current speed over the last three years using the Global Ocean Analysis database from the Copernicus Marine Environment Monitoring Service. Our use of ocean currents as a proxy for vessel movements while at anchorage is based on the discussion in Fuentes (2021). For port records (i.e. when the vessel is deemed to be alongside a quay) the DBSCAN cluster is defined by an arbitrary small ϵ ($1x10^{-6}$) to reflect small changes in the vessel position caused by slight inaccuracies in the AIS-reported position.

Thereafter, we create the *transit cutoff* by completing two tests. The first test, described in Algorithm 2 of the Appendix A, refines the anchorage clusters where consecutive visits could result from: 1) The *time cutoff*[8] or 2) the *polygon cutoff* from the first step. The second test -described in Algorithm 3 of the Appendix A - depends on refinements from the first test and compares whether a consecutive visit to anchorage polygons on opposing sides of the Canals could be considered as a complete transit. The goal of the transit cutoff is to create a mapping of positions to recognize whether a vessel has separate transits through the canal.

In **step three**, we use the interpolated position records from the first step and the mapping of transits from the second step to generate a full sequence of all positions that are part of a canal transit. A final refinement test in step three checks whether a minimum set of polygons (locks, transit and anchorage polygons) is visited by each vessel. This is important to ensure that the results are robust interpretations of the canal visit and reduce the risk of faulty records.

Finally, in **step four**, we derive and record the required information per vessel transit: the waiting time before transiting, the transit time, transit direction, whether the vessel used the neo or regular locks, and any stop that the vessel had in between. The waiting time is the difference in time between the last and first DBSCAN clustered positions at the anchorage before transiting. The canal transit time is estimated as the difference between the time the vessel enters the first lock and the time when it exits the final lock

[8] Not to be confused with the arrival cutoff rule from the PCA.

on the opposite side of the canal. Transit direction and the type of transit (neo or regular) is derived from what lock is visited first. In addition, consistent with PCA statistics, we estimated the Canal Water Time (CWT) as the time from when the vessel anchors until it completes transit less any stoppage at port. To test the performance of our algorithm, we compare in Appendix B the resulting CWT statistics to the official statistics of the PCA.

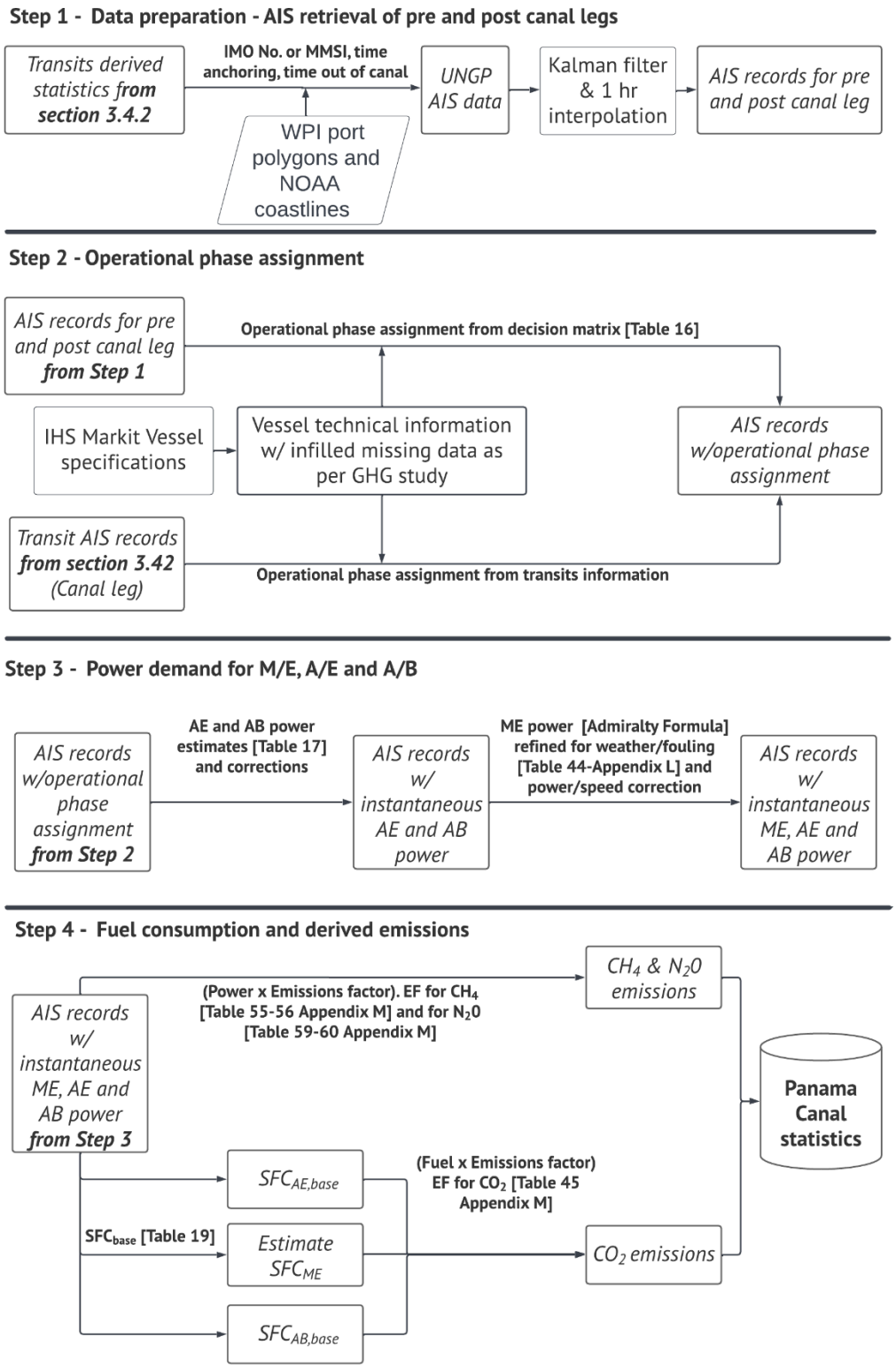
3.4.3 Estimating emissions

In this study, we estimate the canal-related emissions based on the assumptions of the bottom-up method used in the 4th IMO GHG study (2020). The canal-related emissions, in our context, refer to emissions during a vessel's canal transit and on the legs before and after a transit. The method derives emissions at the individual vessel level using position and speed data from AIS and the vessel's technical specifications, enabling aggregation into various statistics as desired (e.g. per geographical region, time, vessel type, etc.). The bottom-up method has been extensively used in the literature for generating emissions inventory (see Andersson and Ivehammar, 2017; Chang et al., 2013; Chen et al., 2017; Coello et al., 2015; Jia et al., 2017; Toscano et al., 2021, Tran et al., 2021). The assumptions of the 4th IMO GHG study have been subject to quality control, have known limitations and enable comparison of results.

We introduce the algorithm for generating vessel emissions in Figure 5. From **step one,** we retrieve vessel records six months before and after a vessel has transited the canal. The transit reference times are the time the vessel was anchored awaiting a transit and the time the vessel cleared the canal transit. These reference times are the first and last recorded positions of our transits recognition algorithm. Thereafter, we keep the records from the vessel's last port call before a transit until the first port call after the transit. Such port visits are assumed to occur when a vessel maintains a speed below 1 knot inside a port polygon for at least 6 hours, an assumption consistent with the IMO GHG study. Using a Kalman filter in this step helps to reduce the noise from faulty GPS readings in the data, unlike linear interpolation which can be more sensitive to such noise.

Figure 5

Framework for estimating Panama Canal related emissions



Note: Information in square brackets refers to the 4th IMO GHG study. ME stands for Main Engine, AE is for Auxiliary Engines and AB is Auxiliary Boilers

In **step two**, for positions outside the canal, we assign the operational phase of the vessel (i.e. at berth, at anchor, maneuvering, in slow transit or normal cruising) per the vessel's position record. Note that records inside the canal are labelled as maneuvering, except for periods where, from our derived Panama Canal statistics, the vessel is at anchorage or is stopped at a local port. In these cases, the records are assigned to at anchorage and at berth, respectively. For the remainder of the voyage (i.e. the legs before and after the canal), we assign the operational phases according to the decision matrix described in the IMO GHG study (2020, pp. 80-81).

Finally, in **step three** and **step four,** we replicate the estimation of the vessels' engine power and the related emissions from the method used in the IMO GHG study.

3.4.4 Coordinated voyage and JIT emissions reduction

For generating the emissions reduction information, we assume every vessel is a booked vessel such that the date of transit is known. A different interpretation would be that a vessel is assigned a slot as it departs from the last pre-canal port, so no distinction is made between booked transits or regular transits.

We assume that a vessel will only reduce speed if four conditions are fulfilled. First, the proposed RTA (i.e. the arrival cutoff for a coordinated voyage or heaving up anchor time for the JIT strategy), must be after our benchmark arrival time but no later than one hour before a transit begins. Second, we assume that vessels will not be instructed to reduce speed below the equivalent of 55% main engine load – a typical limit to slow-steaming (Dere et al., 2022). This implies that very low speeds are not feasible and we avoid suboptimal loads that are potentially harmful to the main engine. Third, a vessel would have reduced emissions if the new emissions for the leg before the canal is lower than the benchmark emissions. It could be the case that a vessel speed reduction would end up with a higher consumption, particularly around the vessel optimal speed. Fourth, vessels that have no waiting time already arrive JIT and are excluded as they have no scope for emission reduction.

3.4.5 Green slots emissions reduction

Our proposal for green slots allocation is based on hindsight from the Panama Canal statistics, as presented in this study. In this case, we allocate the remaining (non-requested) green slots for vessels in the 48 to 96 hours interval prior to transit.

Our motive for using 96 hours as the start of our planning window is based on the PCA (2019) requirement for vessels to submit their intention of transiting at least 96 hours before arriving to the Canal waters. Vessels with a voyage time less than 96 hours, can submit their information as they depart from the last port preceding its transit. By using 48 hours as the end of the planning window, it follows that the last green slots are assigned on the final day of the booking period (2 days before transiting).

The result for our proposed strategy is modeled using the following four steps. In **step one**, we estimate the emissions reduction for every vessel at time $t \in \{96,72,48\}$. The emissions reduction per vessel is the difference between the pre-canal leg plus the canal leg emissions, estimated as of time t, and the benchmark estimates for the same legs. For **step two**, we estimate the non-available (used) green slots by counting those transits with lower waiting times than 30 minutes. In doing so, we assume that a vessel that did not stop before transit was allocated a green slot upon request. From our anchoring recognition method, 30 minutes is the minimum detected cluster based on our assumptions. Finally, we calculate the per day and per direction green slots transits for Super, Regular and Neo-Panamax vessels and estimate the number of remaining slots according to the PCA rules. Note that according to the rules, Neo-Panamax vessels are not assigned green slots. Therefore, we assume that the Neo-Panamax JIT transits take the position of Supers' green slots, transiting in the same direction.

In **step three**, the slot assignment model matches candidates and available green slots based on the following mathematical model:

Sets

D	Transit dates
B	Transit direction $\{North, South\}$
V	Transit type $\{Super, Regular, Neo\}$
I	Transit candidates
L	Transit candidates per day $L \subset I$
F	Transit candidates per day per vessel type $F \subset L$
S	Transit candidates per day per vessel type per direction $S \subset F$

Parameters

e_i	CO_2e reduced on pre-canal leg plus canal leg from transit i
R_{dbv}	Max green slots on day d, direction b for transit type v
P_d	Max green slot for all transits on day d

Decision variable

δ_i	Binary variable. 1 if transit i selected for a green slot, 0 otherwise

$$max \sum_{i \in I} \delta_i e_i \quad (3)$$

s.t:

$$\sum_{i \in S} \delta_i \leq R_{dbv} \; ; \forall d \in D, \forall b \in B \; , \forall v \in V \quad (4)$$

$$\sum_{i \in L} \delta_i \leq P_d \; ; \forall d \in D \quad (5)$$

$$\sum_{i \in F} \delta_i \leq M_{dv} \; ; \forall d \in D \; ; \; \forall v \in V \quad (6)$$

$$\delta_i \in \{0,1\} \quad (7)$$

The objective function (3) maximizes the emission reduction from the green slot assignments. Constraint (4) limits the transits per day, direction, and time according to the PCA JIT assignment rules. Constraint (5) ensures that the maximum number of available slots per day is not surpassed. Constraint (6) limits the green slots per type per day. Constraint (7) defines the decision variable as a binary variable. Note that this model only assigns the vessels that may arrive JIT without being penalized (i.e. after the arrival cutoff). Therefore, the transit assignment is not changed from what was observed in our statistics.

In **step four**, the resulting assigned green slots and related emissions reductions are stored. Thereafter, in a rolling horizon approach, the input parameters for candidates and options are updated, and the model is rerun. The iteration stops when no more slots are available, there are no more candidates, or after a solution is obtained for $t = 48$.

3.5 Results

3.5.1 Panama Canal transits derived statistics

Our algorithm generates 36,812 records of Panama Canal transits between January 2019 and December 2021. Table 1 highlights the differences in waiting and transit times based on transit direction and lock size. As shown in Table 1, 26,853 records are vessels transiting the original Panamax locks and 9,959 are Neo-Panamax

vessels. If subdivided by transit direction, 18,870 records were northbound[9] transits (Gulf of Panama to Caribbean Sea) and 16,371 records were southbound transits (Caribbean Sea to the Gulf of Panama). The greater number of northbound vessels is also reflected in the waiting times, as southbound vessels average 33.6 hours waiting compared to 43.0 hours for northbound vessels, both with large variations. The variation in waiting times is in part driven by the booking system, where different types of vessels have different transit priorities and prevalence of paying a premium for faster transits via pre-booking or auctions.

Table 1

Derived statistics of Panama Canal transits

	Mean	St.Dev	Median	Min	Max	Sample
Waiting Northbound (hours)	43.0	46.9	27.0	0	351.6	16,042
Waiting Southbound (hours)	33.6	43.2	19.5	0	351.1	12,623
Waiting Panamax (hours)	37.9	42.4	23.0	0	351.6	22,150
Waiting Neo (hours)	42.2	54.2	20.1	0	351.1	6,875
Transit time Northbound (hours)	9.8	5.3	8.1	5.5	47.9	18,870
Transit time Southbound (hours)	11.8	6.2	9.0	5.6	49.3	16,371
Transit time Panamax (hours)	9.4	4.2	8.0	5.6	47.9	26,016
Transit time Neo (hours)	14.6	6.9	12.0	5.5	49.3	9,225

The AIS-generated statistics has the advantage, compared to official statistics, that it can be customized based on selected criteria, e.g., vessel type, vessel size, transit direction, etc. As an example of this flexibility, we present a count of transits per vessel type (defined consistently with the 4th IMO GHG study) in Table 2. We can observe that bulkers have the most transits, mainly using the Panamax locks. This reflects how the canal is frequently used by medium-size bulk carriers on the US East Coast to Asia route, this being one of the most important trade routes connected by the canal (Panama Canal Authority, 2022). Neo-Panamax locks are mostly used by large containerships, followed by LPG and LNG tankers.

[9] Some readers might relate to a westbound route being a south transit and eastbound route been a north transit

Table 2

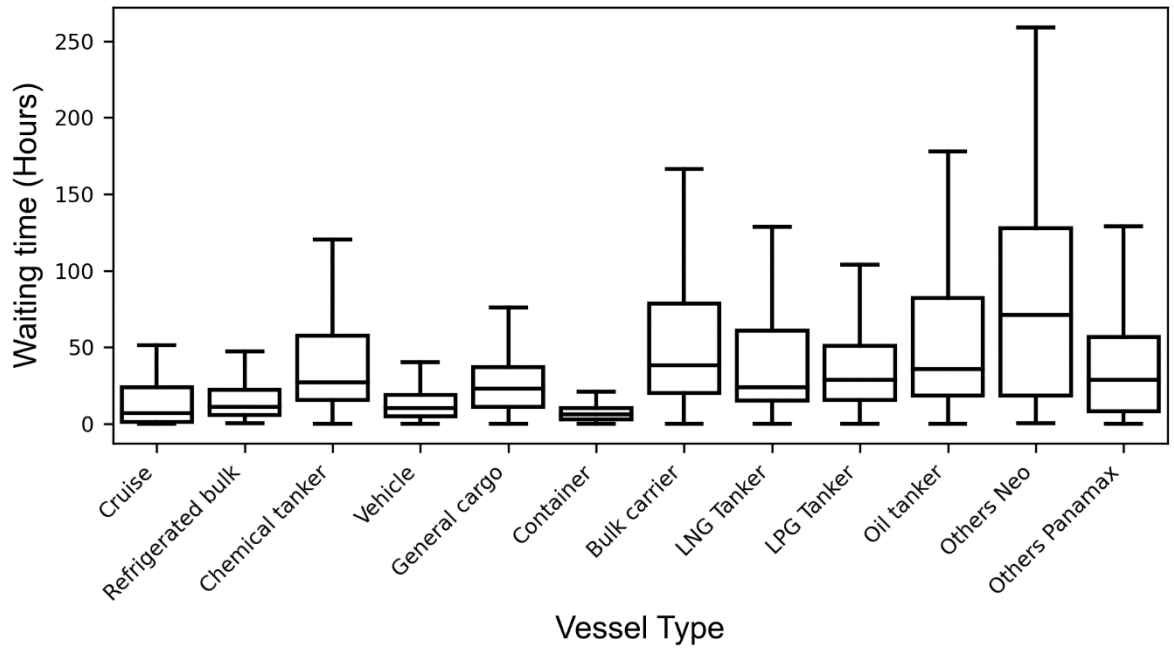
Transits per vessel type

Type of Vessel	Transits	Regulars	Supers	Neo-Panamax
Bulk carriers	7,966	356	6,348	1,262
Containers	7,176	562	2,554	4,060
Chemical tankers	5,727	1,144	4,511	72
LPG	3,934	441	1,138	2,355
Others	2,859	1,008	1,568	283
General cargo	2,137	1,424	705	8
Vehicle	2,065	0	1,925	140
Oil tanker	1,774	133	1,145	496
Refrigerated bulk	1,641	1,577	64	0
LNG	1,256	0	0	1,256
Cruise ships	277	123	127	27

The boxplot in Figure 6 shows there is a difference in waiting times depending on vessel types. For container vessels, 75% of values are within 10.8 hours, showing a compact distribution in expected waiting times. Cruise vessels have the lowest median with 7.0 hours, followed by container vessels with 7.1 hours and vehicle carriers with 10.3 hours. Conversely, crude tankers and bulk carriers have the highest third quartile, with 79.7 and 77.6 hours respectively, as well as a larger standard deviation. The larger variance could result from the mix of strategies for requesting a transit slot: awaiting a turn to be assigned by the PCA, requesting a booked transit, participating in slot auctions, or requesting just-in-time slots. Other factors affecting the assignment of transit slots could be the mix of scheduled vessels under consideration and the customer ranking. A comprehensive description of the transit slot allocation process can be found in the rules for the Panama Canal transit reservation system (2021b).

Figure 6

Waiting time per vessel type



By comparing both the waiting times and the transits made per vessel type, we can infer that a policy based on green slots (limited to Supers and Regulars) could have a significant effect on emissions reduction from LPG tankers, chemical tankers, oil tankers and bulk carriers – the vessel types transiting mainly through the old locks.

LNG tankers are also facing prolonged waiting times before transiting. While a strategy for green slots is not applicable in their case per the current canal rules, their inclusion could lead to important emission reductions if allowed to arrive JIT. We also note here that many use LNG as a fuel, and so there would not only be a reduction of CO_2 emissions but also methane (CH_4), which is a powerful GHG as highlighted by the Global Warming Potential (GWP) of 25. The GWP can be interpreted as how much energy a gas will absorb over a period compared to one tonne of CO_2 (United States Environmental Protection Agency, 2022).

3.5.2 Emissions from existing canal scheduling

We derive the canal related emissions for 36,112 transits (98.1% of the reported transits) of our generated Panama Canal statistics[10]. The results are used to describe the current emissions generated and as a benchmark to compare our proposed strategies.

The estimated emissions per leg and per year are summarized in Table 3. For the period under consideration (2019 - 2021), the voyages of vessels transiting the Panama Canal emitted approximately 101.9 million tonnes of CO_2 equivalent (CO_2e), of which 97.0% are emissions of CO_2 and the rest are CO_{2e} transformations[11] of CH_4 and N_2O. The share of CO_2 to the total is consistent with the global estimates of the 4th IMO GHG study. The total, 31 million tonnes of CO_{2e} for the year 2019, represents 2.9% of shipping emissions in 2018 per the 4th IMO GHG study.

Table 3

Estimated emissions

	CO_2 $x10^3$ (tonnes)			**CH_4 $x10^3$ (tonnes)**			**N_2O $x10^3$ (tonnes)**		
	Before	**Canal**	**After**	**Before**	**Canal**	**After**	**Before**	**Canal**	**After**
2019	15,081.6	375.70	14,713.5	6.387	0.261	7.693	0.863	0.022	0.840
2020	16,446.2	468.81	15,543.0	8.155	0.324	8.251	0.942	0.027	0.888
2021	17,585.7	689.19	18,008.2	9.811	0.449	11.265	1.003	0.040	1.026
Total	49,113.6	1,533.7	48,264.7	24.353	1.034	27.209	2.808	0.089	2.755

In Table 3, it can also be observed how emissions are roughly even between the voyage legs before and after the canal transit, with slightly higher emissions occurring on the leg before.

Emissions segregated by vessel types are presented in Table 4. The higher emitters are container vessels followed by LPG tankers, both transiting through the Neo-Panamax locks. Container vessels emit approximately one third of the total, despite having the lowest waiting time before transiting (as seen in Figure 6). This is of course

[10] Some of our transits are not considered for the emissions inventory if, based on our assumptions, no port was identified either before or after the canal transit.
[11] Global Warming Potential on 100 year basis assumes a transformation factor of 25 for CH_4 and 298 for N_2O

mainly a result of their larger average vessel size, as indicated by the number of transits through the Neo-Panamax locks, larger main engines, and a higher average sailing speed (13.6 knots) than other vessel types such as bulk carriers (10.5 knots) and oil tankers (9.9 knots). For LPG tankers, their overall emissions are driven by their high speed (12.9 knots), number of transits and higher power demand from auxiliary engines from activities such as cargo liquefaction. The highest emitter for the Panamax locks is the bulk carrier segment, with the large number of transits being the main reason.

Table 4

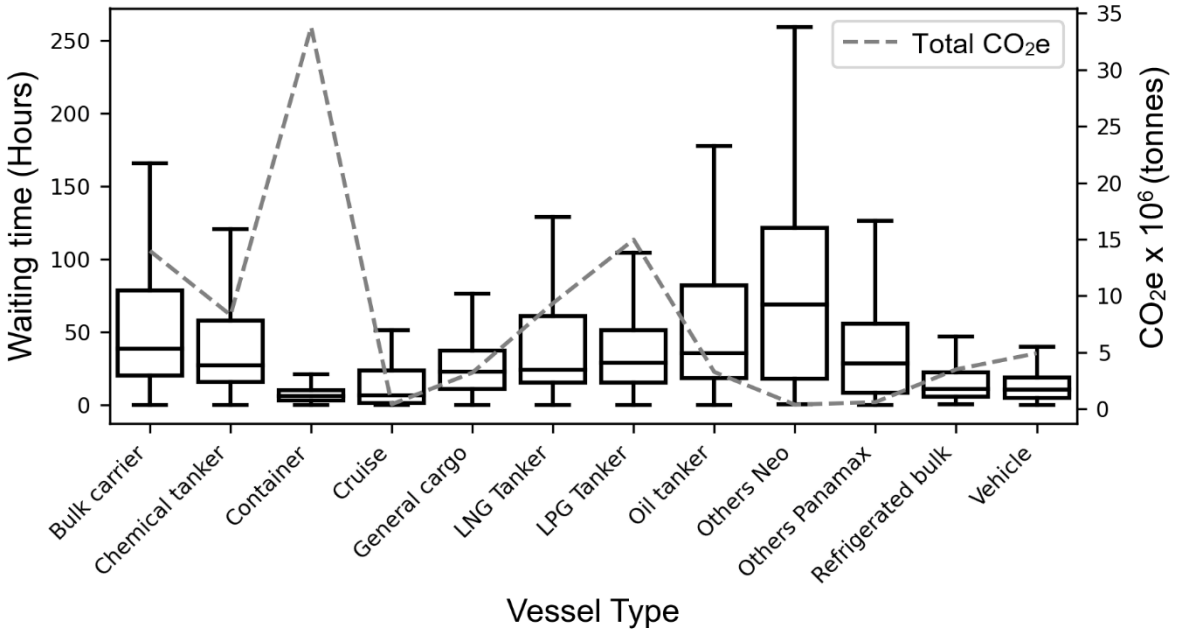
Total estimated emissions per vessel and transit type (Top 10)

	Transit type	CO_{2e} $x10^6$ (tonnes)			
		2019	2020	2021	Total
Container	Neo-Panamax	7.94	9.10	10.80	27.84
Liquefied gas tanker (LPG)	Neo-Panamax	3.43	3.77	4.47	11.67
Bulk carrier	Super	2.99	3.63	3.85	10.47
Liquefied gas tanker (LNG)	Neo-Panamax	2.82	3.27	3.93	10.02
Chemical tanker	Super	2.49	2.36	2.39	7.24
Container	Super	2.50	2.35	2.17	7.02
Vehicle	Super	1.74	1.35	1.66	4.75
Bulk carrier	Neo-Panamax	0.90	1.44	1.49	3.83
Refrigerated bulk	Regular	1.23	1.17	1.09	3.49
Liquefied gas tanker (LPG)	Super	0.70	0.90	1.50	3.10

If we compare the emissions per vessel type to the waiting times presented in Figure 7, there is seemingly a significant opportunity for reducing emissions from LPG and LNG via JIT and for bulk carriers and chemical tankers transiting the Panama locks via JIT and green slots. The potential mitigating effect on emissions from implementing arrival policies by vessel type is presented in the next section.

Figure 7

Waiting time and total CO_2e emissions per vessel type



If the implementation of a speed reduction strategy for vessels approaching the Panama Canal is limited to a subset of the relevant fleet for operational reasons, then an alternative PCA strategy could be to focus on vessels trading on the specific routes (or origin-destination port-pairs) that would have the highest impact on emissions. The results in Table 5 suggest that the highest-emitting routes are dominated by container routes from Asia to the East Coast of the USA, with the Busan-NY/NJ route producing the most emissions. If based on individual ports, an initial strategy should consider Busan port and the Port of Houston, covering four and two of the top routes, respectively. The idea of a per route focus is inspired in the proposal of green corridors presented by the Getting to Zero coalition (2021). Green corridors are “specific trade routes between major port hubs where zero-emission solutions have been demonstrated and are supported”- (Getting to Zero Coalition, 2021). Despite their criteria for selection of green corridors being based on more factors than just the emissions reduction potential (e.g., share of global trade, carbon intensity on route, national policy incentives and regulations, etc.), we recognize that a similar effect is achievable by the PCA in terms of leveraging favorable conditions for accelerated action and creating a spill-over effect to other maritime chokepoints and routes.

Table 5

Top 10 estimated emissions per origin destination pair from Jan. 2019 to Dec. 2021

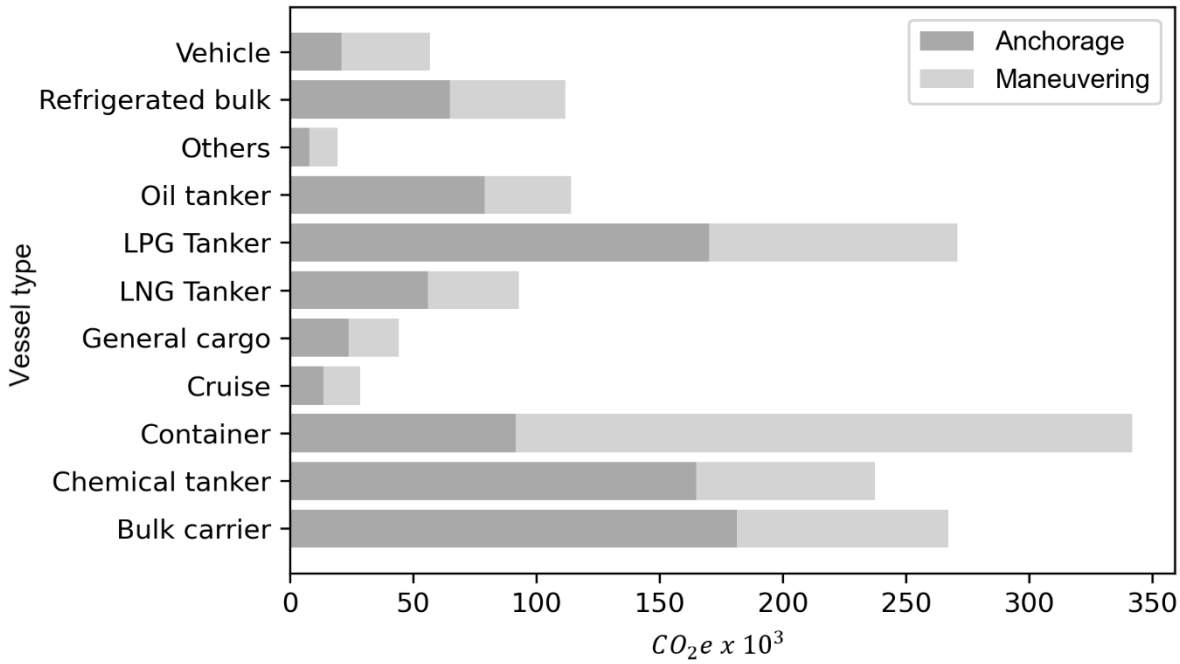
Origin	Destination	CO_{2e} $x10^6$ (tonnes)	Main vessel type
Busan port	Port of New York and New Jersey	2.23	Container
Busan port	Port of Houston	1.98	Container
Xiamen	Port of New York and New Jersey	1.35	Container
Hong Kong/ Shekou/ Yantian	Port of Savannah	1.12	Container
Busan port	Port of Savannah	1.05	Container
Port of Houston	Tokyo Bay ports	0.97	LPG/Chemical
Busan port	Kingston	0.94	Container
Hong Kong/ Shekou/ Yantian	Port of Houston	0.82	Container
Port of Houston	Ulsan	0.80	LPG/Chemical
Ulsan	Port of Houston	0.73	LPG/Chemical

Note: Estimates from vessel transiting the canal with no intermediate ports (i.e., Panama ports inside canal waters). Port of Houston incl. Galveston Bay, Trinity Bay and Houston Channel ports

Reducing emissions at the canal anchorages is an additional benefit of the lower waiting times resulting from our proposals. As these anchoring areas are close to populated areas (Panama City and Colon), the benefit extends to the reduction of emissions related to health hazard such as Particulate Matter and Nitrogen Oxides. This is illustrated in Figure 8, showing the estimated emissions per operational phase and vessel type inside canal waters. While all vessels produce some emissions from auxiliary engines when awaiting transit, certain vessel types (i.e. LNG, LPG and refrigerated bulks) have higher auxiliary power demand related to cargo handling activities such as liquefaction for LNG and LPGs, and cooling for refrigerated cargo. The still significant emissions for vessel types with lower power demand from auxiliary engines (i.e., bulk carriers, chemical tankers, and oil tankers) are mainly driven by their longer anchorage stays before transiting and higher number of vessels.

Figure 8

Estimated emissions at the Panama Canal per operational phase from Jan. 2019 to Dec. 2021



Note: Berth CO_2e emissions for container vessels only (21.4 tonnes)

3.5.3 Emissions reduction from the speed reduction strategies

Under proposal 1 (coordinated voyage), vessels could have saved 5.2 million tonnes of CO_2e in total (5.2% reduction from current emissions) and an average of 1.8 million tonnes per year. (January 2019 to December 2021). The results are limited by our conservative assumptions of vessel not reducing speed below the 55% ME load and vessel arriving to their arrival cutoff before transiting. Yet despite the assumptions, 17,132 transiting vessels (47% of transiting vessels) could have reduced their speed subject to coordination between the canal and the vessels since their departure from the last port.

Details of the emission reduction per vessel segment are presented in Table 6. The biggest effect is observed from speed reduction for container vessels (1.16 million CO_2e tonnes), despite its relatively low waiting time reduction. The effect is stronger for container vessels due to a larger speed reduction over a longer distance. LPG transits also benefit from the speed reduction, as evidenced by their 1 million CO_2e tonnes savings. Notably, they reduce their waiting times on average by 26 hours.

Similarly, bulk carriers, LNG vessels and chemical tankers benefit from a reduction of relatively long average waiting times, affecting both average sailing speeds and removing emissions while at anchor.

Table 6

Emissions reduction (tonnes x ,000) and waiting time (hours) reduction per proposal

Vessel type	Coordinated voyage plan		JIT 48 hours		Green slots		Green slots and JIT 24 hours	
	AWTR*	CO_2e	AWTR	CO_2e	AWTR	CO_2e	AWTR	CO_2e
Container	2.0	1,163.2	2.5	291.8	0.1	9.8	2.0	219.0
LPG	26.1	1,037.7	18.5	323.3	1.7	31.4	11.8	200.0
Bulk carrier	21.7	965.3	16.2	393.3	3.5	91.0	12.1	289.4
LNG	33.0	725.8	17.0	130.6	0	0	7.8	69.1
Chemical tanker	18.0	554.4	15.0	274.4	3.7	78.0	11.5	216.4
Oil tanker	21.9	217.8	15.1	87.9	2.2	13.7	10.6	60.4
General cargo	14.4	210.9	12.5	74.5	3.2	24.2	9.6	58.0
Ref. bulk	6.8	123.6	6.4	61.3	1.2	14.2	5.6	50.8
Vehicle	3.2	102.2	5.7	66.3	0.8	10.4	4.2	47.4
Other Neo vessels	50.6	39.1	16.6	7.9	0	0	8.3	4.6
Other Panamax	7.8	19.6	5.1	9.4	0.7	2.4	3.5	7.2
Cruise vessels	1.3	7.8	2.4	4.5	0	0	2.6	5.8
Total CO_2e		5,167.4		1,725.2		275.1		1,228.1
(%)		(5.2%)		(1.7%)		(0.3%)		(1.2%)
2019 CO_2e		1,423.4		544.4		79.3		364.6
(%)		(4.6%)		(1.8%)		(0.3%)		(1.2%)
2020 CO_2e		1,860.0		570.1		92.8		392.4
(%)		(5.6%)		(1.7%)		(0.3%)		(1.2%)
2021 CO_2e		2,051.2		644.6		101.7		445.9
(%)		(5.5%)		(1.7%)		(0.3%)		(1.2%)

* Average Waiting Time Reduction (hours)

From the results, we also highlight that for the two highest-emitting vessel types (LPG tankers and container vessels), implementing an incentive policy based on this strategy could be customized for vessel departing from specific origin ports.

A strategy based on proposal 2 (JIT 48 hours before arrival) could have saved 1.7 million tonnes of CO_2e (1.7% reduction as compared to the benchmark) from 21,952 transits and an average of 586 thousand tonnes of CO_2e per year. As presented in Table 6, the largest reduction is caused by bulk carriers, followed by LPGs and container vessels. When compared to the coordinated voyage strategy, we note that the impact for container vessels is reduced, despite a higher waiting time reduction. This reaffirms that for some vessel types (i.e., container vessels and LNG) a significant effect is obtained by reducing speed over a longer distance.

We also estimate that proposal 3 (green slots) could have saved 275,090 tonnes of CO_2e (a 0.3% reduction as compared to the benchmark) an average of 91.3 thousand tonnes of CO_2e per year, a lower reduction compared with the previous proposals.

However, this reduction results from just 1,924 new assignments of green slots. From Table 6, the largest reduction would arise from bulk carriers, chemical tankers, and LPGs; vessels with estimated long waiting times and more transits through the Panamax locks. A lower reduction is observed for container vessels, caused by the restriction that JIT slots are assigned just to vessel transiting the old Panamax locks.

Finally, a strategy based on proposal 4 (blend of JIT 24 hours before transit and green slots) would save approximately 1.2 million tonnes of CO_2e (1.2% reduction as compared to the benchmark) from 22,718 transits and an average of 401 thousand tonnes per year. To put this in perspective, this strategy saves more emissions than the highest emitting vessel type in the coordinated voyage strategy, a strategy that is perceived to be more difficult to implement. Importantly, the emission reduction is obtained with less uncertainty (i.e. whether the vessel would arrive on time) than when instructing a vessel to reduce speed over a long distance.

3.6 Discussion and limitations

From the analysis we observed that there is a tradeoff between the perceived complexity of implementation and the potential for emission reduction. Complexity in this context can be thought of as a measure of the number of vessels reducing speed and how early the speed reduction is enforced, leading to uncertainty regarding vessel arrival versus the proposed RTA. The most complex strategy is to implement a coordinated voyage plan, followed by JIT 48 hours before transit for all vessels, the blend of JIT 24 hours before transit and green slots, and the simplest strategy - a green slot assignment under the existing assignment rules. The reduction potential also follows this ranking.

The complexity of implementation is in large part driven by the limitations of the strategies presented in this study. An important limitation is that of uncertainty of vessels arrival versus their RTA. The risk of late arrival becomes more probable the longer a speed reduction is enforced, such as in the coordinate voyage case. A too-early instruction to reduce speed would in reality lead to higher sailing speeds than what is optimal in a deterministic setting to ensure that a vessel does not miss the assigned transit slot. To mitigate this limitation we propose that the implementation is instructed to vessels closer to the canal such as in the JIT 48 hours or green slots

strategies. Another way to mitigate the uncertainty is to focus on specific market segments, vessels coming from ports where the emissions effect is the highest, or perhaps even based on past vessel reliability. This will give the PCA a buffer to maintain business as usual for the vessels that are not part of the proposed speed reduction strategy. A per-route policy could be developed in parallel to the Panama Canal becoming part of green corridors.

An additional limitation is that a vessel needs to be “within sight” to be considered an arrived vessel and that it needs to be physically arrived before a cutoff time. This, of course, are effectively the same barriers to increasing efficiency that we have in ports, that is, the requirement to be able to tender NOR and meeting a laycan. We recognize that these regulations are enforced to maintain control of vessels' arrivals and a buffer for the daily transit schedule. Nonetheless, most of our proposals assume that these requirements are lifted.

Clearly, scheduling strategies that imply speed reduction for only a small subset of vessels have a lower impact on emission reduction compared to other strategies with more flexibility. Such is the case for the green slot strategy which is limited to the assignment of any unused slots. A more balanced measure such as the blended strategy adds flexibility to the green slot assignment by instructing speed reduction to all vessels 24 hours before arrival. As the speed reduction is enforced closer to the time of transit, the Canal could provide accurate transit assignments based on updated information on their resource availability (e.g., pilots, tugboats, canal slots capacity due to weather, etc.).

An additional observation from Table 4 is that many vessel types have a higher potential emission reduction in 2020 and 2021, despite having fewer overall transits as per the official statistics of the PCA (2021c). This suggests that a strategy for reducing emissions prior to a chokepoint has a stronger effect when there are inefficiencies caused by longer waiting times. It follows that for chokepoints with less waiting time uncertainty, such as the Suez Canal where the assignment is based on the time of arrival in a convoy system, there is less opportunity for reducing emissions. There may also be restrictions on reducing vessel speeds in neighboring areas for security reasons, such as piracy threats in the Gulf of Aden.

We recognize that there are additional barriers to improve efficiency in canal operations that are related to the broader discussion of contractual barriers to operational efficiency in shipping. Firstly, under the simplifying assumption that the time of transit does not change, our scheduling proposals – and the accompanying speed and emission reduction - would not affect the tendering of NOR in any subsequent loading port on the post-canal leg. However, we acknowledge that this restriction may not produce the overall societal optimal solution.

Secondly, the utmost dispatch barrier in a voyage charter (Rehmatulla and Smith, 2015; Jia et al., 2017; Global Industry Alliance, 2020) remains to be solved as a matter of contractual wording, typically including an exemption clause. However, the PCA could promote a fairer (to the charterer) exemption clause enforcement. It might be the case that the charterer is the carrier and held liable to the shipper for the cargo on board (Hague Visby Rules, 1979) and an unreasonable deviation, caused by speed reduction, would make them lose rights and liability limitations if an accident occurs. If the PCA enforces the rule that a vessel may reduce speed only under PCA instruction, then it gives the charterers assurance that the “deviation” is not an arbitrary decision to suit the shipowner’s commercial interests, but a matter of operational efficiency. On the charter party side, a new exemption clause could be drafted in the spirit of the Baltic and International Maritime Council Just in Time clause (2022).

Thirdly, regarding the split incentives problem on a voyage charter (Rehmatulla and Smith, 2015), the charterer would remain unaffected as the speed reduction takes place before an expected stop and not on the final leg to the destination port. This assumes that the charterer has agreed to waive the utmost dispatch clause (if applicable), for the period the vessel is instructed to reduce speed. The same argument holds for the rush-to-wait behavior (Poulsen and Sampson, 2019) in the destination port, which is not affected by changes on the pre-canal leg.

3.7 Conclusions

In this study, we explore the implications of implementing speed reduction measures driven by increasing operational efficiency at chokepoints such as the Panama Canal. We find that the barriers to implementation discussed in the literature, such as the rush-to-wait behavior caused by first-come-first-served port policies and the potential for

earning demurrage, and split incentives for fuel cost savings, are broadly mitigated in the case of intermediate stops such as a canal transit. This highlights how changes to scheduling policies in the Panama Canal can be an important source of emission reduction in shipping that has hitherto not been considered in the literature.

An additional argument supporting this finding is that canal authorities are usually in control of most decisions affecting the transit of a vessel, reducing or removing the inefficiencies caused by the interaction among several stakeholders (Global Industry Alliance, 2020; Poulsen and Sampson, 2020; Slack et al., 2018).

In our Panama Canal case study, we measured the effect of implementing speed reduction measures from increasing operational efficiency in the canal. Our results show potential reductions from 275,090 tonnes to 5.2 million tonnes of CO_2e emissions from January 2019 to December 2021 (an average of 91,000 to 1.8 million tonnes of CO_2e per year) by promoting speed reduction measures based on four separate proposals with due consideration of their scheduling rules. Our study also highlights the tradeoff between effect and perceived complexity of implementation, and shows how proposals could be implemented gradually by market segmentation.

As a methodological contribution, we introduced a framework for estimating the Panama Canal operational statistics and derived emissions from AIS data. We also adapted the bottom-up method in the 4th IMO GHG study to fit our case study.

This study is limited to emissions reduction on the pre-canal leg of a voyage. We acknowledge that in order to also reduce post-canal emissions, a proposal should also include a scheduling model for assigning transits. The logic is that, as a community planner, the canal could assign, within their operational constraints, the best transit mix with due consideration to the emissions savings both before and after a transit. As we do not have access to the details of such operational constraints, this is not pursued here.

Future research should make a similar assessment for other maritime chokepoints, such as the Suez Canal and Bosphorus Strait, and consider the impact of changes to their existing scheduling rules. Ideally, a more advanced model would also consider optimization of the entire voyage, as well as arrival uncertainty caused by weather and operational constraints within the canal. With the help of our empirical methodology,

future studies could also consider the strengths and weakness of the Canal for being part of a green corridor and propose a strategy for improving such weaknesses.

References

Andersson, P., & Ivehammar, P. (2017). Green approaches at sea–The benefits of adjusting speed instead of anchoring. *Transportation Research Part D: Transport and Environment, 51*, 240-249.

Baltic and International Maritime Council. (2022, August). *Just in Time arrival clause for voyage charter parties 2021*. Retrieved from https://www.bimco.org/contracts-and-clauses/bimco-clauses/current/just-in-time-arrival-clause-for-voyage-charter-parties-2021

Bouman, E. A., Lindstad, E., Rialland, A. I., & Strømman, A. H. (2017). State-of-the-art technologies, measures, and potential for reducing GHG emissions from shipping–A review. *Transportation Research Part D: Transport and Environment, 52*, 408-421.

Cepeda, M. A., Assis, L. F., Marujo, L. G., & Caprace, J. D. (2017). Effects of slow steaming strategies on a ship fleet. *Marine Systems & Ocean Technology*, *12*(3), 178-186.

Chang, Y. T., Song, Y., & Roh, Y. (2013). Assessing greenhouse gas emissions from port vessel operations at the Port of Incheon. *Transportation Research Part D: Transport and Environment, 25,*, 1-4.

Chen, D., Wang, X., Li, Y., Lang, J., Zhou, Y., Guo, X., & Zhao, Y. (2017). High-spatiotemporal-resolution ship emission inventory of China based on AIS data in 2014. *Science of the Total Environment*, *609*, 776-787.

Coello, J., Williams, I., Hudson, D. A., & Kemp, S. (2015). An AIS-based approach to calculate atmospheric emissions from the UK fishing fleet. *Atmospheric Environment*, *114*, 1-7.

Corbett, J. J., Wang, H., & Winebrake, J. J. (2009). The effectiveness and costs of speed reductions on emissions from international shipping. *Transportation Research Part D: Transport and Environment*, *14*(8), 593-598.

Dere, C., Zincir, B., Inal, O. B., & Deniz, C. (2022). Investigation of the adverse effects of slow steaming operations for ships. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*.

Dockray, M. (2013). *Cases and Materials on the Carriage of Goods by Sea.* Routledge.

Eide, M. S., Endresen, Ø., Skjong, R., Longva, T., & Alvik, S. (2009). Cost-effectiveness assessment of CO2 reducing measures in shipping. *Maritime Policy & Management, 36(4)*, 367-384.

Eide, M. S., Longva, T., Hoffmann, P., Endresen, Ø., & Dalsøren, S. B. (2011). Future cost scenarios for reduction of ship CO2 emissions. *Maritime Policy & Management*, *38*(1), 11-37.

Fuentes, G. (2021). Generating bunkering statistics from AIS data: A machine learning approach. *Transportation Research Part E: Logistics and Transportation Review, 155*, 102495.

Fuentes, G., & Adland, R. (2020). A spatial framework for extracting Suez Canal transit information from AIS. *2020 IEEE International Conference on Industrial Engineering Management (IEEM)*, 586-590.

Getting to Zero Coalition. (2021). *The next wave: Green corridors.*

Global Industry Alliance. (2020). *Just In Time Arrival Guide.* London: International Maritime Organization.

International Maritime Organization. (2020). *Reduction of GHG emissions from ships. Fourth IMO GHG Study 2020 – Final report. MEPC 75/7/15.* London: International Maritime Organization.

Jia, H., Adland, R., Prakash, V., & Smith, T. (2017). Energy efficiency with the application of virtual arrival policy. 54. *Transportation Research Part D: Transport and Environment. 54* , 50-60.

Johnson, H., & Styhre, L. (2015). Increased energy efficiency in short sea shipping through decreased time in port. *Transportation Research Part A: Policy and Practice*, *71*, 167-178.

Kalman, R. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 35-45.

Krmac, E., & Mansouri Kaleibar, M. (2022). A comprehensive review of data envelopment analysis (DEA) methodology in port efficiency evaluation. *Maritime Economics & Logistics*, 1-65.

Lavon, B., & Shneerson, D. (1981). Optimal speed of operating ships. *Maritime Policy & Management*, *8*(1), 31-39.

Lindstad, H., Asbjørnslett, B. E., & Strømman, A. H. (2011). Reductions in greenhouse gas emissions and cost by shipping at lower speeds. *Energy policy*, *39*(6), 3456-3464.

National Centers for Environmental Information. National Oceanic and Atmospheric Administration. (2017). Global Self-consistent, Hierarchical, High-resolution Geography Database. Retrieved January 10, 2022, from https://www.ngdc.noaa.gov/mgg/shorelines/data/gshhg/latest/.

Panama Canal Authority. (2019, December). *Regulation on Navigation in Panama Canal Waters.* Retrieved from https://www.pancanal.com/eng/legal/reglamentos/Acuerdo-360.pdf

Panama Canal Authority. (2021a). *Advisory to Shipping No. A-24-2021 (Monthly Canal Operations Summary - April 2021).*

Panama Canal Authority. (2021b, July). *OP Notice to Shipping No. N-7-2021*. Retrieved from Panama Canal Transit Reservation System: https://www.pancanal.com/eng/op/notices/2021/N07-2021.pdf

Panama Canal Authority. (2021c, July). *Panama Canal Traffic by Fiscal Year*. Retrieved from https://www.pancanal.com/eng/op/transit-stats/2020/Table-01.pdf

Panama Canal Authority. (2022, July). *Panama Canal Traffic Along Principal Trade Routes.* Retrieved from https://pancanal.com/wp-content/uploads/2022/03/Table-00.pdf

Poulsen, R. T., & Sampson, H. (2019). 'Swinging on the anchor': The difficulties in achieving greenhouse gas abatement in shipping via virtual arrival. *Transportation Research Part D: Transport and Environment, 73*, 230-244.

Poulsen, R. T., & Sampson, H. (2020). A swift turnaround? Abating shipping greenhouse gas emissions via port call optimization. *Transportation Research Part D: Transport and Environment*, *86*, 102460.

International Convention for the Unification of Certain Rules of Law relating to Bills of Lading (Hague rules), 25 August 1924, as amended by the Protocol of 23 February 1968 (Hague-viby rules) and by the Protocol of 21 December 1979 (SDR protocol).

Psaraftis, H. N., Kontovas, C. A., & Kakalis, N. M. (2009, October). Speed reduction as an emissions reduction measure for fast ships. In *10th International Conference on Fast Sea Transportation FAST* (pp. 1-125).

Rehmatulla, N., & Smith, T. (2015). Barriers to energy efficiency in shipping: A triangulated approach to investigate the principal agent problem. *Energy Policy, 84*, 44-57.

Rodrigue, J. P. (2004). Straits, passages and chokepoints: a maritime geostrategy of petroleum distribution. *Cahiers de géographie du Québec*, *48*(135), 357-374.

Schwartz, H., Gustafsson, M., & Spohr, J. (2020). Emission abatement in shipping–is it possible to reduce carbon dioxide emissions profitably? *Journal of Cleaner Production, 254*, 120069.

Slack, B., Comtois, C., Wiegmans, B., & Witte, P. (2018). Ships time in port. *International Journal of Shipping and Transport Logistics*, *10*(1), 45-62.

Tran, N. K., Lam, J. S., Jia, H., & Adland, R. (2021). Emissions from container vessels in the port of Singapore. *Maritime Policy & Management*, 1-17.

Toscano, D., Murena, F., Quaranta, F., & Mocerino, L. (2021). Assessment of the impact of ship emissions on air quality based on a complete annual emission inventory using AIS data for the port of Naples. *Ocean Engineering*, *232*, 109166.

United States Environmental Protection Agency. (2022). *Greenhouse Gas Emissions*. Retrieved from Understanding Global Warming Potentials: https://www.epa.gov/ghgemissions/understanding-global-warming-potentials

WPI. (2019). *World Port Index.* Springfield: National Geospatial-Intelligence Agency.

Youngjoo, K., & Hyochoong, B. (2018). Introduction to Kalman filter and its applications. In *Introduction and Implementations of the Kalman Filter* (pp. 1-16).

Appendices

Appendix A

Tests for generating Panama Canal statistics

Algorithm 1 Time and polygon cutoff

Inputs:
$Z_s \leftarrow$ [i-j]$_{Ix5}$ ({i|:i $\in \{1, ..., I\} \wedge$ i $\in \mathbb{Z}$}; j $\in$ {polygon,time,cutoff,group,groupPosition})
s$\in$S; S is a set of vessels
A is a set of anchorage polygons
Default Parameter:
timeCutoff $\leftarrow$ 48, hours between ordered records to declare a time cutoff
for s $\in$ S **do**
  set $Y_s \leftarrow 0_{Mx5}$
  $z_{ij} \leftarrow Z_s$
  $y_{mj} \leftarrow Y_s$
  m$\leftarrow$0
  for i $\in \{1, ..., I\}$ **do**
    $z_{i,groupPosition} \leftarrow 0$
    if $z_{i,time}$-$z_{i-1,time}$ > timeCutoff $\vee$ $z_{i,polygon} \neq z_{i-1,polygon}$ **then**
      $z_{i,cutoff} \leftarrow 1$
      $z_{i,groupPosition} \leftarrow 1$
      $z_{i-1,groupPosition} \leftarrow 2$
    else
      $z_{i,cutoff} \leftarrow 0$
    end if
  end for
  for i $\in \{1, ..., I\}$ **do**
    $z_{i,group} \leftarrow \sum_{t=1}^{i} z_{t,cutoff}$
    if $z_{i,polygon} \in$ A $\wedge$ ($z_{i,groupPosition}$=1 $\vee$ $z_{i,groupPosition}$=2) **then**
      m$\leftarrow$m+1
      $y_{m,:} \leftarrow z_{i,:}$
    end if
  end for
end for

Algorithm 2 Transits cutoff - Test 1 (Merge of anchor visits)

Inputs:
$Y_s \leftarrow$[m-j]$_{Mx5}$ ({m|m $\in \{1, ..., M\} \wedge$m $\in \mathbb{Z}$}; j $\in$ {polygon,time,cutoff,group,groupPosition}) ; from Algorithm 1
$Z_s \leftarrow$[i-j]$_{Nx5}$ ({i|:i $\in \{1, ..., N\} \wedge$i $\in \mathbb{Z}$}; j $\in$ {polygon,time,cutoff,group,groupPosition})
s$\in$S; S is a set of vessels
Default Parameter:
timeCutoff $\leftarrow$ 48, hours between consecutive visits to a same polygon
for s $\in$ S **do**
  set $X_s \leftarrow 0_{Vx5}$
  $y_{mj} \leftarrow Y_s$
  $x_{vj} \leftarrow X_s$
  v$\leftarrow$0
  for m $\in \{1, ..., M\}$ **do**
    if $y_{m,polygon} \neq y_{m-1,polygon}$ **then**
      $y_{m,cutoff} \leftarrow 1$
    else
      $y_{m,cutoff} \leftarrow 0$
    end if
  end for
  for m $\in \{1, ..., M\}$ **do**
    $y_{m,group} \leftarrow \sum_{t=1}^{m} y_{t,cutoff}$
  end for
  for m $\in \{1, ..., M\}$ **do**
    if $y_{m,group}$=$y_{m-1,group} \wedge y_{m,groupPosition}$=1 **then**
      if $y_{m,time}$-$y_{m-1,time} \leq$timeCutoff **then**
        if call PORTINBETWEEN($y_{m-1,time}$,$y_{m,time}$,Z_s)=TRUE **then**
          $y_{m-1,cutoff} \leftarrow$-1
          $y_{m-2,cutoff} \leftarrow$-1
        else
          $y_{m-2,cutoff} \leftarrow 1$
          $y_{m-1,cutoff} \leftarrow$-1
          $y_{m,cutoff} \leftarrow$-2
        end if
      end if
    end if
  end for
  for m $\in \{1, ..., M\}$ **do**
    if $y_{m,cutoff} \geq 0$ **then**
      v$\leftarrow$v+1
      $x_{v,:} \leftarrow y_{m,:}$
    end if
  end for
end for

Algorithm 3 Transits cutoff - Test 2 (Transits recognition)

Inputs:
$X_s \leftarrow$[v-j]$_{Mx5}$ ($\{v|v \in \{1,...,V\} \wedge v \in \mathbb{Z}\}$; j $\in$ {polygon,time,cutoff,group,groupPosition}) ; from Algorithm 2
$Z_s \leftarrow$[i-j]$_{Nx5}$ ($\{i|:i \in \{1,...,N\} \wedge i \in \mathbb{Z}\}$; j $\in$ {polygon,time,cutoff,group,groupPosition})
s$\in$S; S is a set of vessels
for s $\in$ S **do**
  set $G_s \leftarrow 0_{Px5}$
  $x_{vj} \leftarrow X_s$
  $g_{pj} \leftarrow G_s$
  p$\leftarrow$0
  for v $\in \{1,...,V\}$ **do**
    if $x_{m,polygon}$=$x_{v-1,polygon}$ $\wedge x_{v,groupPosition}$=1 **then**
      $x_{v,cutoff} \leftarrow 1$
    else
      $x_{v,cutoff} \leftarrow 0$
    end if
  end for
  for v $\in \{1,...,V\}$ **do**
    $x_{v,group} \leftarrow \sum_{t=1}^{v} x_{t,cutoff}$
  end for
  row$\leftarrow$0
  for v $\in \{1,...,V\}$ **do**
    if $x_{v,cutoff}$=1 **then**
      row$\leftarrow$1
    else
      row$\leftarrow$row+1
      if row>4$\wedge x_{v,groupPosition}$=1 **then**
        if ($x_{v,time}$-$x_{v-1,time}$)<($x_{v-2,time}$-$x_{v-3,time}$) **then**
          $x_{v-2,cutoff} \leftarrow 1$
        else
          $x_{v,cutoff} \leftarrow 1$
        end if
      end if
    end if
  end for
  for v $\in \{1,...,V\}$ **do**
    $x_{v,group} \leftarrow \sum_{t=1}^{v} x_{t,cutoff}$
  end for
  row$\leftarrow$0
  for v $\in \{1,...,V\}$ **do**
    if $x_{v,cutoff}$=1 **then**
      row$\leftarrow$1
    else
      row$\leftarrow$row+1
      if row=4 **then**
        p$\leftarrow$p+1
        $g_{p,:} \leftarrow x_{v-3,:}$
        p$\leftarrow$p+1
        $g_{p,:} \leftarrow x_{v,:}$
      end if
    end if
  end for
end for

Appendix B

Performance testing for the algorithm that generates the Panama Canal transit statistics

The PCA (2021a) publishes a summary of their operations in monthly reports as part of their Advisories to Shipping. The performance of our algorithm relate mainly to the ability to generate timely statistics, yet it is also important that the generated statistics are robust as a correct representation of the operations. To verify the robustness, we test our results against official statistics from the PCA for canal transits.

In Figure B.1, we observe the comparison of the number of transits as calculated by our algorithm to monthly transits as reported by the PCA. Overall, our algorithm discovers fewer transits than officially reported. The difference between the values is on average 33.3 transits per month and the algorithm accuracy ranges from 88.9% to 99.8%. As the algorithm aims for accuracy to build a representative sample of variables, such as waiting times and service times, the difference is expected and is caused by the algorithm not reporting vessels with incomplete sequences. Vessels with incomplete sequences are potentially caused by the height of the AIS antennas on smaller vessels or environmental factors affecting the AIS radio signal.

Figure B.1

Monthly Panama Canal transits from the algorithm and official statistics

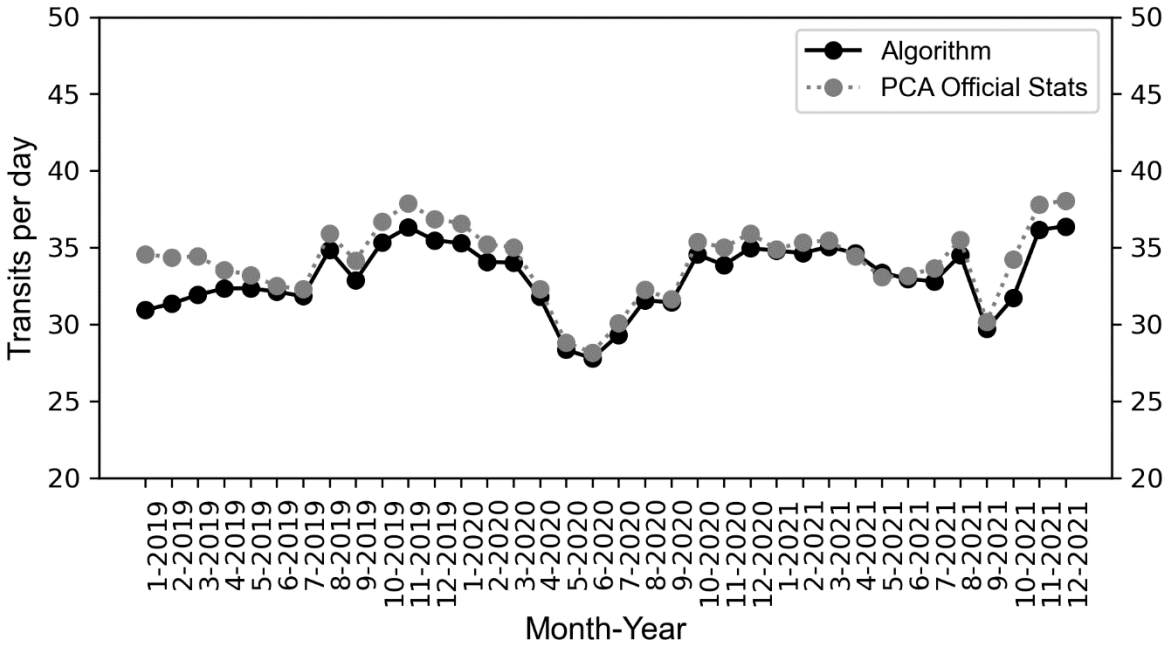
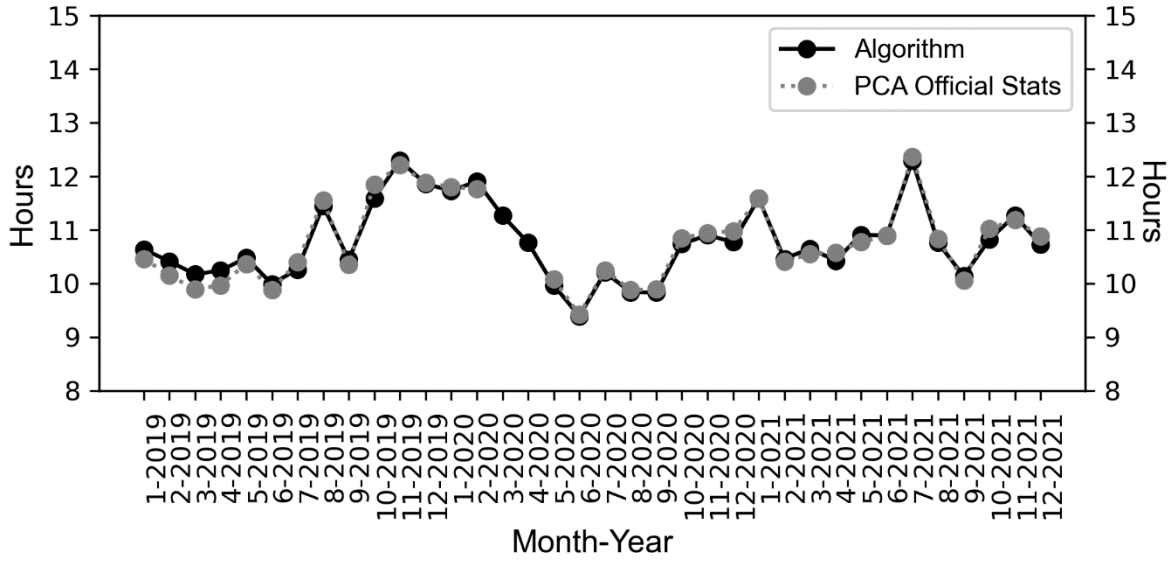


Figure B.2 shows the relation of statistics for the monthly in transit time (ITT). ITT is the time lapsed from the moment a vessel enters the first lock until it exits the last lock (Panama Canal Authority, 2021b). Our algorithm has a close fit to the true values, as the algorithm results are within 97.1% and 99.9% from the official statistics records. Importantly, the algorithm generated information for "blind" months (March and April 2020), where official statistics were not reported.

Figure B.2

Monthly Panama Canal In transit time (ITT) from the algorithm and official statistics



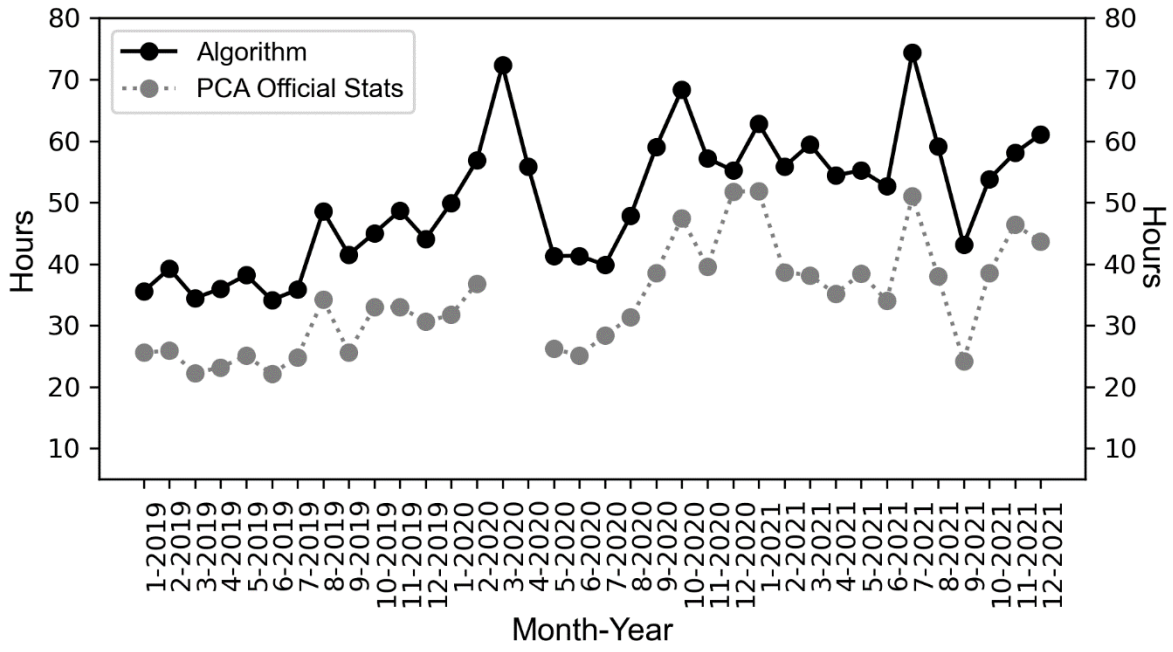
Note: PCA missing values as months not reported in their Monthly Operations Summary.

Similar to ITT, the PCA reports CWT as a measure of service performance. CWT is the time lapsed from the moment a vessel is declared "ready" for transit until it completes transit. Therefore, it is a measure of waiting time before transit, plus ITT. This measure encompasses the whole service where the PCA is responsible, and can be considered a proxy of service quality. As the interpretation of a ready vessel is not publicly defined by the PCA, we assume the period starts at the beginning of the anchoring period, as recognized by our algorithm, and ends when the vessel clears the last lock. It might be the case that PCA is reporting that a vessel is ready after being cleared for transit, such as after finishing inspection by PCA boarding officers. Figure B.3 shows that our algorithm is compatible with the official records. The fit between both values is of an R^2 of 0.84 and the difference is on average 15.52 hours, which is

likely the average time taken to clear a vessel for transit after it has arrived at anchorage.

Figure B.3

Monthly Panama Canal water times from the algorithm and official statistics



Note: PCA missing values as months not reported in their Monthly Operations Summary.

From this evaluation, we can conclude that our generated statistics are suitable proxies of the underlying operations at the Panama Canal.