INTERINSTITUTIONAL POSTGRADUATE PROGRAM IN MARINE SCIENCE AND TECHNOLOGY MANAGEMENT

Dissertation/Thesis

LINER SHIPPING SCHEDULE RELIABILITY: THE ROLE OF BERTHING WINDOWS

Evdokia Bei

Supervisor: Dr. Ioannis Lagoudis

Piraeus
August 2026

DECLARATION OF AUTHENTICITY

The person who is conducting the Thesis bears the entire responsibility of determining the fair use of the material, which is defined based on the following factors: the purpose and character of the use (commercial, non-profit or educational), the nature of the material used (part of the text, tables, figures, images, or maps), the percentage and significance of section, which it uses in relation to the entire copyrighted text, and the possible consequences of this use on the market or the overall value of the copyright text.

Signature: _____

Evdokia Bei

This Diploma Thesis was unanimously approved by the three-member Examination Committee appointed by the Board of Directors of the MSc in accordance with the Regulations of the MSc «Management in Maritime Science and Technology».

The members of the Committee were:

- MEMBER A': Dr. Ioannis Lagoudis
- MEMBER B': Dr. Ioannis Theotokas
- MEMBER C': Dr. Anastasia Christodoulou

The approval of the thesis by the Department of Maritime Studies of the University of Piraeus does not imply acceptance of the author's views.

Preface and Acknowledgements

This thesis was prepared as part of the Interinstitutional Postgraduate Programme «Management in Maritime Science and Technology» of the University of Piraeus and the Hellenic Naval Academy. It brings together a review of the academic and operational literature on berth allocation and schedule reliability with an original discrete-event simulation built on anonymized vessel-call data from a large transshipment container terminal.

I would like to thank my supervisor, Dr. Ioannis Lagoudis, for their guidance and feedback throughout the preparation of this work, and the members of the Examination Committee for their time.

Table of Contents

List of Figures

List of Tables

Abstract

This thesis examines whether an appointment-based ("rendezvous") berth-allocation policy reduces vessel waiting time at a large transshipment container terminal compared with a first-come, first-served (FCFS) regime, and how this effect depends on terminal congestion. The literature review situates the question within research on schedule reliability, the berth allocation problem, berthing windows and berth templates, and appointment and pre-booking systems, drawing on peer-reviewed studies together with the operational procedures of the Piraeus Container Terminal (PCT) as a case context, and identifies a gap: no existing study has tested a PCT-type pro-forma and rendezvous system against pure FCFS under common, stochastic simulated conditions. The empirical chapters build a discrete-event simulation, calibrated on 794 anonymized vessel calls spanning nineteen months and two demand regimes, and compare four policy scenarios — current hybrid practice, pure FCFS, universal appointment, and a cargo-weighted priority rule — across 400 replications each. The central finding is that the appointment system redistributes waiting-time risk from mother vessels to feeders rather than reducing it overall: it functions as insurance rather than speed, although the apparent strength of this protection across congestion regimes depends on whether it is measured in absolute or proportional terms. A cargo-weighted alternative reduces total delay by 10–14% but concentrates the added cost on the smallest consignments. The thesis states that the decision regarding which berth allocation policy to adopt is essentially a distributional one in so far as it determines which cargo has to wait, and this has consequences for the way terminals and carriers negotiate priority.

Keywords: berth allocation; schedule reliability; discrete-event simulation; container terminal; appointment system; queueing

1. Introduction

1.1 Background and problem statement

Container terminals sell two things that cannot be stored: quay length and time. When a vessel arrives, it competes with every other vessel already waiting for one of a fixed number of berths, and the order in which vessels are served has direct consequences for waiting time, turnaround time, berth utilization and, ultimately, the reliability of the liner schedule that connects that port call to the rest of the network (Zhang, Zheng and Teo, 2021). A published liner schedule is not only a timetable for the ocean voyage; it is likewise a commitment to a sequence of port calls, berth services, cargo-handling windows and onward connections. When a vessel arrives late, waits for a berth, or stays in port longer than planned, that disruption propagates to transshipment, inland transport and inventory operations downstream (Karmelić et al., 2025).

One instrument terminals use to manage this competition is the berthing window: a commercial or operational commitment that a vessel arriving within an agreed interval will receive priority, or guaranteed, treatment. Chapter 2 uses the publicly documented operating procedures of the Piraeus Container Terminal (PCT) as an illustrative case context for the literature review, since PCT applies exactly such a system: certain vessels receive a pro-forma window and may request an appointment, while all other traffic is served first-come, first-served (FCFS). This case-study use is confined to Chapter 2; the terminal simulated in Chapters 3 and 4 has been fully anonymized. Whether this kind of hybrid appointment system actually improves outcomes — and for whom — is an empirical question that the existing literature has not answered under controlled, comparable conditions, as Chapter 2 sets out in detail.

1.2 Aim, objectives and research question

The aim of this thesis is to evaluate, through discrete-event simulation, how an appointment-based berth-allocation policy affects vessel waiting time and the distribution of delay relative to pure FCFS, and to interpret the implications for liner schedule reliability under different levels of congestion.

This aim is pursued through three objectives: first, to review and synthesize the literature on schedule reliability, the berth allocation problem, berthing windows and templates, and

FCFS, priority and appointment-based allocation policies; second, to build and calibrate a discrete-event simulation model of a container terminal using anonymized operational data, capable of comparing alternative allocation policies under identical stochastic conditions; and third, to compare the current hybrid policy against pure FCFS, a universal-appointment variant, and a cargo-weighted priority rule, and to interpret the results for both operational and distributional (fairness) implications.

The research question that follows from these objectives is: what effect does a berthing-window policy based on appointments, whether used on its own or in combination with the FCFS rule for unscheduled vessels, have on vessel waiting time and the distribution of delay when compared with pure FCFS under stochastic container-terminal conditions, and how does this effect vary with the level of congestion?

1.3 Methodological approach

The thesis answers this question with a discrete-event simulation (DES) built in SimPy on Python 3.12, driven by 794 anonymized vessel calls recorded at a large transshipment container terminal over nineteen months. The data are split into two demand regimes either side of a statistically verified structural break, and the arrival and service-time processes are estimated separately by regime and vessel class using empirical resampling rather than a fitted theoretical distribution. Competition from other carriers for the same berths, which is not directly observed in the data, is represented explicitly through a calibrated "phantom" arrival stream. Four policy scenarios — the current hybrid policy, pure FCFS, a universal-appointment variant, and a cargo-weighted priority rule — are then compared under identical random conditions across 400 replications each, using waiting-time measures and a cargo-weighted delay measure as outcomes. The full specification, including the model's assumptions and limitations, is given in Chapter 3.

1.4 Structure of the thesis

The remainder of the thesis is organized into five further chapters. Chapter 2 reviews the literature on schedule reliability, the berth allocation problem, berthing windows and templates, and FCFS, priority, appointment and hybrid allocation policies, and situates the Piraeus Container Terminal's operating procedures within that literature as a case context before identifying the research gap addressed by this thesis. Chapter 3 sets out the method: the data and their cleaning, the identification of the two demand regimes, the input analysis,

the simulation model itself, the calibration of the competitor-load parameter, the four policy scenarios, and the model's assumptions and limitations. Chapter 4 presents the results: the empirical waiting-time patterns, the calibration outcome, what actually distinguished the congested period from the normal one, the robustness of the regime split, and the comparison of the four policy scenarios. Chapter 5 draws together the conclusions, discusses the contribution and limitations of the study, and proposes directions for future research.

2. Literature Review

2.1 Introduction

Liner shipping schedule reliability is a service-performance matter which links vessel operations, the capacity of container terminals and broader supply-chain coordination. A published liner schedule functions not only as a timetable for the ocean journey as well as a sequence of intended port calls, berth services, cargo-handling activities and onward connections. If a vessel arrives late, has to wait for a berth or stays in port longer than scheduled, the disruption can be passed on to the next ports as well as to transshipment, inland transport and inventory operations. The connection between schedule reliability and berth allocation is therefore both operational and commercial: the berth decision determines when a vessel can start its services, and the quality of that decision influences whether the planned port-call window can be achieved.

The chapter looks at the existing literature concerning the comparison of first-come, first-served (FCFS), appointment-based, and hybrid berth-allocation policies for container ships. It has four aims. The first is to define schedule reliability, berth allocation, and berthing windows as separate yet interdependent ideas. The second is to classify the main formulations of the berth-allocation problem and to describe the evolution of the priority, pre-booking, stochastic, optimization, and simulation approaches. The third aim is to link berth allocation with port-call synchronization, Just-in-Time (JIT), virtual arrival, vessel-speed decisions, and maritime decarbonization. The fourth is to use the operating procedures of the Piraeus Container Terminal (PCT) as a short case study in order to interpret the academic concepts, at the same time distinguishing this grey literature source from peer-reviewed research.

The relevant literature was found through structured searches in Scopus, Web of Science, ScienceDirect and Google Scholar, together with backward and forward citation tracking of the key papers and by carrying out a specific search of the official sources of the International Maritime Organization and the European Union for up-to-date material on JIT and emissions regulation. The search terms used were schedule reliability, berth allocation, berthing windows, FCFS, rendezvous, pre-booking, stochastic berth allocation, berth templates, port-call synchronization, JIT arrival, virtual arrival, and emissions. Papers were included only if they dealt with berth-allocation policy, berthing windows, or schedule dependability in the area of container shipping and had presented a model, carried out an empirical analysis, or

described an operational system; those focusing solely on landside terminal activities, on yard and gate optimization, or on bulk or tanker terminals were not included. Grey literature was only accepted when it described actual operational practice or regulation and was clearly identified as such. The review is set out as follows: Section 2.2 looks at schedule dependability in the liner shipping sector; Section 2.3 defines the berth allocation problem and categorises its main variants; Section 2.4 deals with berthing windows, berth templates and cyclic planning; Section 2.5 reviews FCFS, priority, appointment and synchronization policies; Section 2.6 examines queueing, stochastic, optimization and simulation approaches; Section 2.7 presents the PCT case context; Section 2.8 identifies the research gap and the research question; and Section 2.9 draws the implications for the thesis.

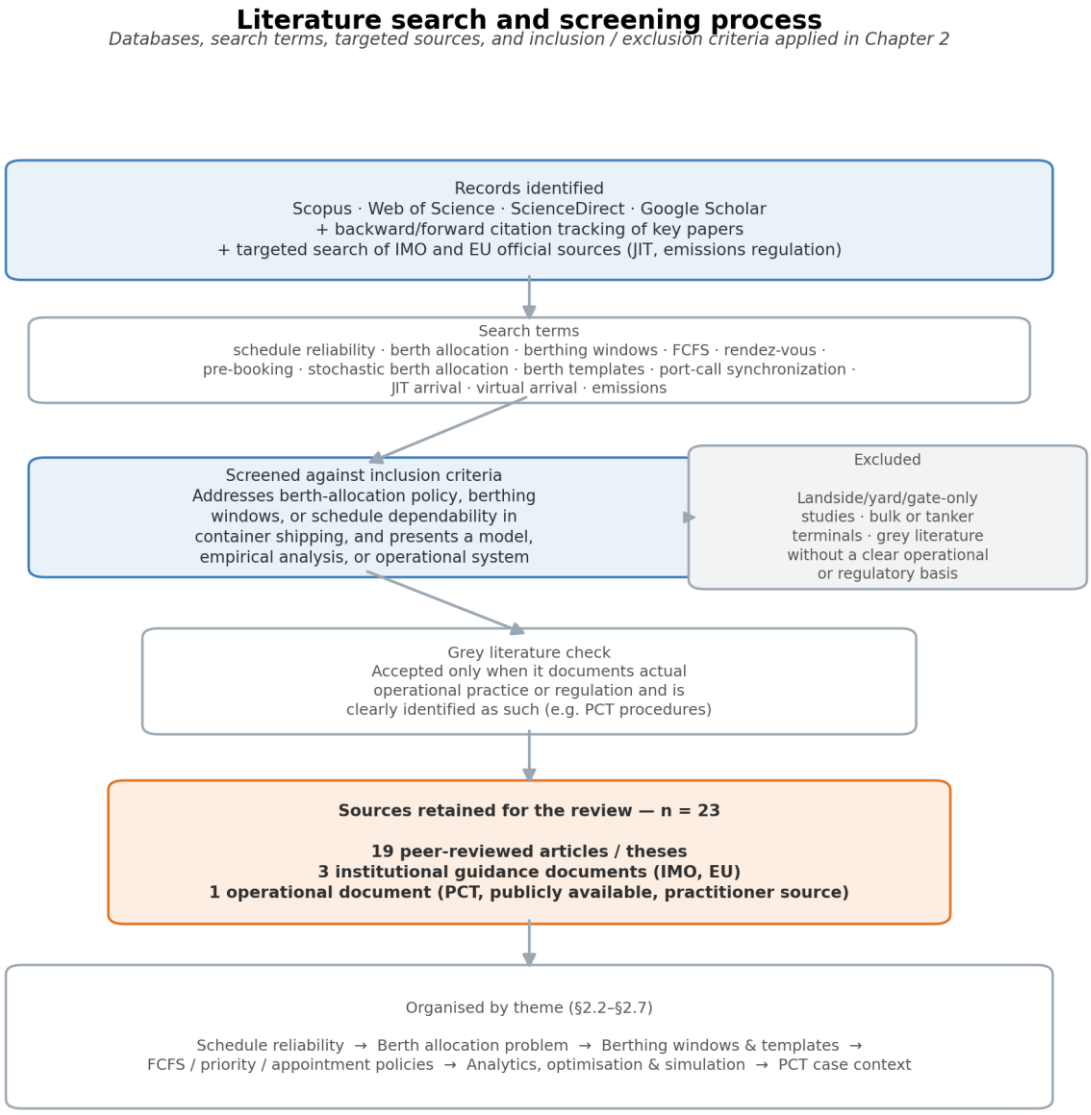

Figure 2.1. Literature search and screening process: databases, search terms, targeted sources, and the inclusion/exclusion criteria applied in this chapter.

2.2 The reliability of the schedule in liner shipping

2.2.1 Concept and measurement

The degree to which a liner service can be said to be reliable refers to the consistency of its performance in relation to the published or agreed timetable; it is thus a property that arises from repeated performances, not the punctuality of any single voyage. Zhang et al. (2021) define reliability in terms of the probability of arriving on time and prove that both the time taken to sail and the time spent in port are stochastic. Karmelić et al. (2025) do the same by linking reliability to compliance with the pro forma schedule and point out the causes of disruption at sea, at anchorage, in ports, and during inland transport.

To carry out a study on berth allocation, reliability must be defined for each port call. A ship may arrive within its allotted time but still depart late due to an occupied berth or low cargo handling efficiency. On the other hand, a vessel that arrives late might make up for the delay by berthing immediately or by speeding up the handling process. The analysis should therefore differentiate between arrival reliability, service-start reliability and departure reliability, the latter two being most directly influenced by the berth policy. The main measure should be the proportion of vessels for which both the start and end of service occur within the originally planned operational window, along with waiting and turnaround times.

The difference lies in separating the final section of the schedule from the voyage section. Factors such as weather conditions, currents, delays caused by canals, and decisions regarding sailing speed have an impact on the maritime section, whereas berth availability, the rules governing the queue, handling productivity, and the allocation of equipment affect the port section. Zhang et al. (2021) demonstrate that the design of the timetable entails balancing reliability against fuel consumption and voyage time, and Karmelić et al. (2025) show that port congestion and waiting for a berth can become major causes of unreliability when delays spread through a service string.

2.2.2 Importance with regard to operations and the supply chain

The significance of reliability extends not only to the carrier since a port call is part of terminal operations, transshipment, inland transport, and cargo delivery; a delayed ship might miss a feeder connection, push another vessel off a berth, or shorten the time available at a later port. Thus the costs of unreliability are shared among carriers, terminals, logistics providers, and cargo interests.

According to Böse (2011), who cites Notteboom, a large portion of disruptions to liner schedules is linked to unexpected waiting times before docking and low container-handling efficiency. This result does not imply that berth allocation accounts for all the delay, but it does show that the quay-side operation is a significant point at which action can be taken. A policy that aims to reduce waiting times may improve the current port call and also help future calls, while a poorly designed queue can turn a local capacity shortage into network-wide delays.

Reliability is also connected with decisions regarding speed and buffer times. Although greater speed may save time it will at the same time increase fuel use, while bigger buffers can enhance reliability but will lengthen the planned round voyage (Zhang, Zheng and Teo, 2021). When a reliable berth window is made known early on, the carrier can reduce its speed in order to prevent arriving too early. If there is any uncertainty about berth availability, the carrier may increase its speed as a safeguard against losing its position, which in turn results in an early arrival and the need to wait while anchored.

2.2.3 Reliability as an end result

When assessing schedule reliability, it is necessary to take into account waiting time, turnaround time, berth utilization, throughput, and the effect on non-priority vessels. Depending exclusively on average waiting times can mask long delays affecting a small number of vessels or cause the delays to be transferred from booked calls to non-booked ones. As Yıldırım et al. (2020) and Ursavas and Zhu (2016) show, allocation objectives must be made explicit since improving one performance measure may lead to a deterioration of another. The detailed simulation measures and the model assumptions should be included in the methods chapter.

2.3 The problem of allocating berths

2.3.1 Definition and decision points

The problem involved in berth allocation (BAP) is the assignment of arriving ships to the available quay space and their service times. When making a decision the terminal has to determine both when a vessel will start its service and where it will be placed, possibly taking into account issues such as sequencing, the length of service, the suitability of the berth, clearance, and the cranes and other resources. Qin et al. (2009) describe the problem in terms of time-window constraints relating to vessel service and berth availability, while Gkolias

(2007) views it as a time-space scheduling problem in which contractual service windows and various service objectives can be expressed.

The BAP goes beyond what a queue-order rule provides. While FCFS specifies which vessel is chosen first under certain conditions, it does not determine the feasible berth position, the duration of service, or the interaction with the cranes. The allocation problem involves deciding both when and where service takes place, and in the case of integrated models, how quay-crane capacity is assigned. The relevant decisions comprise the berth position, the start and completion times, the vessel sequence, quay space, the handling of early and late arrivals, whether service is interrupted or shifted, and resource management. The manner in which these options are made has an impact on waiting times, turnaround time, utilization, throughput, vessel connections, and schedule reliability.

2.3.2 The classification of BAP formulations

The literature is usually divided according to three different criteria. In terms of space, a discrete BAP splits the quay into a finite number of berths, a continuous BAP permits a vessel to take up any possible position along the quay, and a hybrid BAP combines fixed areas with flexible positioning. Li et al. (2023) regard discrete, continuous and hybrid formulations as a fundamental organizing principle. With regard to time, a static BAP sets out a fixed collection of vessels and parameters for a given period, while a dynamic BAP makes changes to its decisions as vessels arrive, leave or new information becomes available. In relation to uncertainty, deterministic models make use of fixed arrival times, handling times and capacity, whereas stochastic, robust and distributionally robust models show the variation by means of random variables, scenarios or uncertainty sets.

The separations can be combined; a model can be continuous and dynamic together with being stochastic, or on the other hand it can be discrete and static and deterministic. This is important since they indicate which operating conditions are being represented. While a static and deterministic comparison can show nominal efficiency, it is not able to determine the effect of missed appointments, varying handling times or revised arrivals on reliability. Although the simulation implemented in this thesis makes use of a simplified spatial representation, the thesis itself is conceptually concerned with a dynamic and stochastic operational environment.

Classification of berth allocation problem formulations – Table 2.1

Table 2.1. Classification of berth-allocation problem formulations

Dimension	Categories	Relevance to this thesis
Spatial	Discrete, continuous, hybrid quay representation	A discrete berth representation is sufficient for comparing queue policies
Temporal	Static (fixed vessel set) versus dynamic (decisions revised as information arrives)	The comparison requires a dynamic setting in which late arrivals are resolved as they occur
Uncertainty	Deterministic, stochastic, robust or distributionally robust	Stochastic arrivals and handling times are essential to test whether reservations survive variability

2.3.3 Queueing and capacity

BAP may also be regarded as an example of a queueing problem: ships arrive at a fixed number of parallel service locations and have to wait until a suitable berth and the necessary resources are available. The arrival patterns can be regular or disrupted, and the service time is determined by the container movements, the size of the vessel, the productivity of the cranes and any interruptions. When demand nears capacity, waiting becomes very sensitive to small changes in the arrival rate or in the length of service.

Legato and Mazza (2001) treat the activities involving arrival, berthing and departure as a queueing network featuring time-varying priorities and interacting resources. Since analytical methods become difficult when the service stations and the allocation rules are non-standard, they apply discrete-event simulation in order to carry out what-if planning for berths. Ursavas and Zhu (2016) do the same in respect to queueing approaches and demonstrate how stochastic arrivals and handling times affect policy design. Queueing theory is thus useful for explaining capacity and congestion, while simulation is useful whenever there is interaction among multiple resources and operational rules.

The view from the queueing standpoint also makes it clear what the cost of reserving vessels is. Although an appointment system may reduce uncertainty for certain vessels it can result in idle capacity if a booked vessel is late or fails to turn up. Evaluation should therefore take into account the waiting times for both booked and non-booked vessels, the length of the queue and berth utilization, and not merely the performance of the booked vessels.

2.4 Berthing windows, templates and cyclic planning

2.4.1 Berthing windows as contractual and operational entities

A berthing window refers to a time period connected with a vessel's scheduled arrival, berthing, servicing or departure. It may be established commercially as a result of an agreement between a carrier and a terminal, or it can be determined operationally on the basis of the berth plan taking into account the expected size of the vessel, the container movements, the capacity of the berth and the working shifts. It is important to make the distinction between a window and a point estimate. An estimated time of arrival is a forecast, while a berthing window can act as a service commitment, a priority condition or a constraint in the allocation model.

The PCT procedures show this operational meaning by associating pro-forma windows with an estimated number of container lifts and with a required day and time of the week (Piraeus Container Terminal Single Member S.A., n.d.). In a similar way, the academic literature regards time windows as constraints which link vessel service requirements with berth availability (Qin et al., 2009). A window can therefore contribute to reliability if it is flexible enough to take account of normal variations but specific enough to aid in berth and resource planning. A window that is too narrow may result in frequent breaches and in the need for emergency rescheduling, while a window that is too wide may offer little protection for the carrier or little guidance for the terminal.

Berthing windows also establish a kind of governance relationship between the carrier and the terminal. The terminal has to decide whether a vessel that falls within its window is given priority, guaranteed berthing, or merely preferential treatment. The carrier is required to give a credible arrival time and to provide details regarding cargo volume, the equipment needed and the connections required. The effectiveness of the window therefore depends on the quality of the information and on the rule that is applied when the vessel arrives early, late or outside the agreed time interval.

2.4.2 Berth templates and the design of the cyclic berthing window

In the shipping industry it is common for a large number of ship calls to take place in accordance with weekly or other cyclic service patterns. It is therefore possible to prepare a berth template, which is defined as a recurring assignment of berth space and time to regular services over a cycle. Unlike daily reactive allocation, template design consists in setting up a tactical pattern to assist with future calls. Moorthy and Teo (2006) treat berth management as a template design problem, while Hendriks et al. (2010) look at robust cyclic berth planning in the face of uncertain vessel delays. The aforementioned studies indicate that

cyclic planning can stabilize service patterns, but it has to include some slack or robustness since a delay in one cycle can have an impact on later calls.

A berth template can contribute to schedule reliability in two ways: by setting aside recurring capacity for strategic services, which in turn reduces the necessity of having to renegotiate each call, and by offering a standard against which deviations can be measured. At the same time, the template may become inflexible. In the event that a service is delayed, a reserved slot might stay empty or could cause the terminal to have to move another vessel. Because of this, a recovery mechanism is needed, for example in the form of controlled shifting, temporary changes in priority or dynamic re-optimization.

Charisis et al. (2018) go beyond the terminal stage by incorporating port time windows and coordinated arrivals into the design of liner routes and schedules. Their model indicates that the limited availability of berths at ports results in the need to coordinate the arrival times of vessels going to the same port. This represents an important link between network schedule design and terminal allocation in that the berthing window is not merely a constraint imposed by the terminal on the carrier, but also a factor which can influence decisions regarding routing, speed and service frequency.

2.5 From FCFS to policies involving coordinated arrivals

2.5.1 FCFS as the basic rule

The FCFS method chooses the vessel which arrived first or that joined the queue first, provided that it is suitable for the berth and for the operations. The approach is attractive because of its transparency, the ease with which it can be put into practice and the general sense of fairness it inspires. However, its disadvantage is that it does not automatically take into account factors such as vessel size, handling time, contractual time windows, transshipment connections, the urgency of the service, as well as the environmental cost of arriving early and of waiting while anchored.

FCFS may lead to a rush-to-wait phenomenon when carriers think that arriving early will improve their position in the queue. Kontovas and Psaraftis (2011) connect this behavior to fuel consumption and emissions, and Yıldırım et al. (2020) give the reason why flexible priority policies are looked at as alternatives. Although FCFS remains a useful benchmark, it should not be regarded as a neutral optimal solution.

2.5.2 Priority and hybrid rules

Priority rules alter the FCFS system by giving different weights or service orders to vessels according to their features. Examples are the large-vessel-first policy, the earliest-due-date approach, serving vessels with the lowest container load first, customer-specific service levels, and priorities based on transshipment connections. The Earliest Due Date is connected with reliability since it sets urgency by means of a completion deadline which can take into account arrival time, the expected handling duration and a buffer. It does, however, rely on reliable handling estimates and on an agreed-upon definition of lateness.

Yıldırım et al. (2020) examine single-queue, multiple-queue and hybrid queue-priority arrangements. Their findings indicate that flexible priorities could lead to shorter waiting times and better utilization and throughput, although the results will vary according to the type of queue structure and the way in which larger vessels are treated. A hybrid rule is able to safeguard urgent or booked vessels while still providing a queue for those without special status, but it also has to establish fairness and deal with the problem of repeated displacement.

2.5.3 Systems for appointment and rendezvous

A system for appointments or rendezvous enables a carrier to ask for a service period in advance and to obtain either planned or guaranteed berth treatment, provided that the necessary conditions are met. The terminal can make use of the booking to even out demand and to reserve both berth and crane capacity. Mubder (2024) examines the shift from FCFS arrivals to pre-booking and claims that a guaranteed berth window can facilitate JIT arrivals. The study also points out that information quality, port-community systems, governance, external nautical services and trust are necessary for implementation.

A rendezvous rule has to state what is to occur when the vessel comes early, when it is late, or when it is not there at all. As an example of how it might be applied, the PCT shows that vessels arriving within their pro-forma window are given priority, non-rendezvous feeders are attended on a first-come, first-served basis and late-coming vessels may lose their appointment (Piraeus Container Terminal Single Member S.A., n.d.). This results in a hybrid system rather than a full replacement of FCFS. The rule has the advantage of linking the carrier's decision regarding arrival with terminal capacity, but it can only be regarded as credible if the terminal is able to keep the reservation and pass on any changes.

2.5.4 Optimizing and synchronizing port calls

By optimizing port calls, berth reservations are extended to take into account the coordinated availability of berths, pilots, tugboats, fairways, mooring services and cargo-handling resources. Ceder et al. (2018) make use of Port Collaborative Decision Making (PortCDM) in order to represent the various port actors and arrive at a recommended time of arrival as a result of information exchange and through successive agreement. The vessel's target time of arrival can then be adjusted until the relevant parties are able to accommodate the call.

According to the IMO JIT Arrival Guide, a vessel is said to have a virtual arrival when it is asked to arrive later and is permitted to reduce its speed so that it reaches the port at the time when the berth and the related services are ready (International Maritime Organization, 2020). This approach shifts some of the coordination duties from the period when the vessel is anchored to the planning stage before arrival. However, the advantages of this method depend on the availability of reliable information and external services; Mubder (2024) and Ceder et al. (2018) thus advocate a careful interpretation of the matter, namely that synchronization may enhance reliability and environmental performance, but it does not guarantee arrival at the berth universally.

A comparison of the berth-allocation policies found in the literature, as shown in Table 2.2

Table 2.2. Comparison of berth-allocation policies in the literature

Policy	Basis for ordering	Advantages	Disadvantages
FCFS	Order of arrival at the port or queue	Transparent, simple, perceived as fair	Ignores vessel size, connections and urgency; encourages rush-to-wait
Priority rules	Vessel attribute (size, due date, load, contract)	Reduces waiting for critical calls; improves utilization	Requires credible estimates; can displace smaller vessels repeatedly
Pro-forma window	Contractually agreed recurring interval	Stabilises planning; gives a measurable reliability reference	Rigid under delay; reserved slots may stand idle
Appointment / rendezvous	Advance booking subject to compliance	Enables JIT arrival and speed reduction	Depends on information quality; needs a rule for missed slots
Hybrid	Appointment or window priority plus FCFS for the remainder	Protects strategic services while retaining a queue	Fairness and displacement rules must be defined explicitly

2.6 Analytical methods, optimization, simulation, and environmental integration

2.6.1 Stochastic and optimization approaches

The research carried out by BAP makes use of mixed-integer programming, dynamic programming, heuristics, metaheuristics, stochastic programming and robust optimization; these methods aim at finding allocations which minimize waiting time, service time, lateness, cost or emissions within the given constraints. Qin et al. (2009) use a tabu-search heuristic to formulate the time-window constraints; Ursavas and Zhu (2016) develop policies in the case of uncertain arrival times and handling times; and Charisis et al. (2018) link port time windows and coordinated arrivals with the design of liner routes and schedules.

Optimization offers formal insight and good schedules for a given instance, but it may be based on information that is not actually available in practice. Both network schedule design and terminal allocation deal with different levels of decision-making. While Zhang et al. (2021) optimize both reliability and bunker consumption over a liner route, BAP studies usually concentrate on quay-side allocation; the first establishes the planned arrival patterns and buffers, while the second decides whether these patterns can be turned into timely service.

2.6.2 The use of simulation and queueing methods for evaluation

Simulation is suitable in cases where the system involves queues, limited resources, random arrivals, varying handling times, and policy rules that influence each other over time. Henesey et al. (2004) make use of a berth allocation management system in order to compare berth policies under different quay configurations, different spacing requirements and various arrival sequences. Abdel Hafez et al. (2013) link berth allocation with quay-crane assignment and stress the importance of turnaround time. These studies show that simulation can serve as a useful method for carrying out experiments when comparing policies under common conditions.

A universal optimum cannot be determined by simulation; instead, it sets up a controlled environment in which the same vessels, arrivals, capacity constraints and handling assumptions are evaluated under different rules. This approach is suitable for comparing FCFS, appointment-based and hybrid policies. The detailed specification of the entities, resources, events, scenarios, outputs and validation is left for the methods chapter.

The outcome of the simulation is influenced by the arrival process, the assumptions regarding service times, the berth constraints, the priority rules and the replication design; a

policy might work well under conditions of low demand but break down when congestion occurs, or although it may reduce average waiting it could at the same time increase maximum waiting. Carvalho et al. (2026) show that it is possible to assess waiting time, emissions and resource utilization together when JIT and speed optimization are taken into account.

2.6.3 Optimizing speed, reducing emissions and complying with regulation

The connection with the environment can be explained by the speed at which vessels travel. When the availability of a berth is uncertain, ships may go faster in order to keep their position in the queue or to arrive early and then wait by anchoring. If a reliable requested arrival time is given, slower sailing becomes possible and excess fuel consumption can be reduced. Zhang et al. (2021) show the trade-off between reliability and bunker consumption, whereas Kontovas and Psaraftis (2011) associate coordinated arrivals and shorter waiting times with lower emissions.

JIT does not eliminate uncertainty; instead, it transfers the need for coordination to the stage before the vessel arrives. A vessel can only reduce its speed when it expects the berth, the fairway, and the nautical services to be available at the agreed time. The EU regulation makes this relationship more important. The EU ETS has been applied to maritime transport since 2024 for large ships visiting ports in the EEA, featuring phased surrender obligations and, from 2026, the inclusion of methane and nitrous oxide in the relevant carbon dioxide equivalent framework (European Maritime Safety Agency, 2026). FuelEU Maritime comes into full effect from 2025, progressively lowers the permitted level of greenhouse gas intensity of the energy on board, and introduces mandatory zero-emission-at-berth requirements for the relevant container and passenger ships in EU ports from 2030 (European Commission, 2025).

The measures do not include a rule regarding berths, but rather alter the context in which such a rule is assessed. A policy which decreases waiting time for anchoring and allows for slower sailing might result in lower fuel consumption and reduced emissions; on the other hand, a policy that sets aside capacity but leads to idle berths or frequent re-planning may produce weaker benefits. Reliability, capacity and emissions should therefore be considered as closely related outcomes, even if the emissions are calculated as a secondary figure.

2.7 Piraeus Container Terminal as a case context

In this thesis the Piraeus Container Terminal Operations Procedures are employed as a specific example of an operational setting and are not used in place of peer-reviewed literature. The document mentions that when preparing berth plans the PCT takes into account pro forma berthing windows, vessel berthing appointments, arrival announcements, berth availability, vessel connections, and cargo plan information (Piraeus Container Terminal Single Member S.A., n.d.). The fact that such a list includes contractual scheduling, advance information, present capacity and network dependencies shows that practical berth planning involves all these elements.

The procedures set out pro-forma windows for certain main-trade vessels, together with the negotiated arrangements concerning the expected container lifts and the day and time of the week required. Vessels arriving within this pro-forma window are given priority over those arriving outside the window or making incidental calls, provided that special circumstances are not involved and taking into account the impact on the other schedules and the normal terminal workflow. The procedures thus give a clear example of a service-window policy which is more structured than one based on pure FCFS.

The document also includes a system for arranging appointments or rendezvous for eligible liner vessels, this involving prior applications and a series of steps relating to the information provided before arrival. This shows the kind of information and compliance that is required in the case of pre-booking. It is further stated that non-rendezvous feeder services are handled on a first-come, first-served basis and that a vessel which does not arrive on time will lose its appointment. All of these rules thus form a mixed allocation system consisting of appointment or pro-forma priority for certain services, together with FCFS treatment being applied to all other vessels and to those that miss their appointments.

The PCT context is useful since it converts the general literature into specific rules which can be included in a simulation. However, it does not mean that the same standards or procedures apply to other terminals. Procedures relating to customs, dangerous goods, reefers, cargo documents and safety are not covered by the literature review since they only serve to define the berth-policy comparison when they are specifically incorporated into the method.

It should also be made clear that the Port of Piraeus and PCT are used in this chapter solely as an illustrative, publicly documented case study of a real-world berthing-window policy, in order to ground the concepts discussed above in an operational setting. This use is

confined to the literature review: the terminal modeled in the discrete-event simulation of Chapters 3 and 4 has been fully anonymized (§3.1), and no claim is made, here or elsewhere in this thesis, that it is PCT specifically.

2.8 The research gap and the research question

It cannot plausibly be said that there is a lack of research into FCFS. FCFS is widely studied as a practical baseline, and its limitations have been identified: it is capable of ignoring the effects specific to each vessel, it tends to encourage early arrivals, and it performs badly when there is heterogeneity in handling times and connections (Yıldırım, Aydın and Gökkuş, 2020; Kontovas and Psaraftis, 2011).

Another group looks at priority rules and hybrid rules. Yıldırım et al. (2020) compare different queue structures and flexible priorities, whereas Böse (2011) looks at FCFS and the earliest due date. Tighilt et al. (2025) view the move towards berthing-window systems as a reaction to congestion. Although these studies show the potential of precedence rules, they do not explain how a container terminal should combine contractual windows, rendezvous calls and an FCFS queue for unbooked vessels when the experimental conditions are the same.

Another group focuses on pre-booking, synchronization and JIT. According to Mubder (2024), the conditions needed for implementation are information sharing, trust, governance and access to external resources. Ceder et al. (2018) have developed a proof of concept for PortCDM relating to recommended arrival times, whereas Kontovas and Psaraftis (2011) explain the environmental reasons for rendezvous. Although these studies show why appointment systems are attractive, they do not include a controlled comparison of the outcomes of container-terminal queues under pure FCFS, appointment priority and a hybrid rule.

A fourth group is concerned with formal optimization. Qin et al. (2009) formulate the time-window constraints; Charisis et al. (2018) link the windows to route and schedule design; Ursavas and Zhu (2016) model stochastic arrivals and handling times; and Li et al. (2023) carry out a classification of the wider BAP literature. Although these studies offer rigorous formulations, their objectives and assumptions differ network schedule design does not directly assess the terminal queue discipline, whereas stochastic optimization generally aims to find a policy for a specific mathematical setting rather than to test implementable rules using common simulated inputs.

The simulation studies carried out by Henesey et al. (2004), Abdel Hafez et al. (2013) and Carvalho et al. (2026) show that simulation is suitable for comparing policies, examining resource interactions and assessing JIT-related indicators. The problem yet to be resolved is one that is integrative and evaluative: although the literature includes the various components, it fails to adequately link a PCT-type pro forma and rendezvous structure with stochastic arrivals and handling times in a single controlled container-terminal comparison of policy scenarios.

The question that arises is this: what effect does a berthing-window policy based on appointments, whether used on its own or in combination with the FCFS rule for unbooked vessels, have on vessel waiting time and the distribution of delay when compared with pure FCFS under stochastic container-terminal conditions?

2.9 Closing synthesis

The literature treats schedule reliability as a performance outcome and separates berth allocation as a time-space decision process, the two being linked by berthing windows which convert a commercial schedule into an operational commitment. The BAP may be discrete, continuous or hybrid; static or dynamic; and deterministic, stochastic or robust. Queueing theory is used to explain capacity and congestion, optimization offers formal methods for assignment, and simulation enables implementable rules to be compared when the same stochastic conditions are applied. In all these areas the consistent conclusion is that no one rule performs best in all situations: the effectiveness of the FCFS, priority, appointment and hybrid policies varies according to the relationship between demand and capacity, depending on the variability of the arrival and handling times, and also on the way in which vessels falling outside their windows are handled.

The existing literature has not yet included a controlled comparison in which a pro-forma and rendezvous system of the type documented at PCT—this covering the FCFS treatment of non-booked feeders—is tested against a pure FCFS system using the same vessels, arrivals, and handling assumptions. It is for this reason that the present chapter lays down both the conceptual foundation for such a comparison and the criterion that reliability and waiting times should be evaluated together rather than relying on a single averaged measure. The model which puts this comparison into practice is given in Chapter 3.

3. Methodology

The method used in this chapter is that of a discrete-event simulation (DES) study, which assesses the berth-allocation policy at a large transshipment container terminal. The research question is whether a vessel berthing appointment (rendez-vous) system leads to shorter waiting times than a first-come, first-served (FCFS) system, and whether this difference depends on the level of congestion at the terminal.

The terminal, the carrier and the port are not named in this thesis. The terminal is described only as a port in the Mediterranean, and is referred to throughout as “the terminal”, the carrier as “the carrier”. This differs from Chapter 2, where the Piraeus Container Terminal (PCT) is named and its published operating procedures are cited by URL as a publicly documented, illustrative case context for the literature review (§2.7); no claim is made anywhere in this thesis that the terminal modeled here and in Chapter 4 is PCT. The vessel-level operating data are commercially sensitive and have been anonymized. Vessel names and voyage numbers are replaced by codes, service codes by labels, and infrastructure figures are rounded (to approximately 2,800 metres of total quay, about 30 quay cranes). The operational rules are taken from the terminal's own published operational procedures.

3.1 Data

The study initially comprised 796 vessel calls made by the carrier at the terminal. The study period is set as 01/01/2025 to 05/08/2026, a duration of nineteen months, calculated according to the berthing date. The raw file starts on 26/12/2024 since four of the calls arrived during the last week of December 2024 and berthed during the early days of January 2025; as the period is defined based on berthing, these four calls are included within it and therefore retained. Each call includes the time records for the service cycle (arrival at the port limits, pilot boarding, berthing, start and completion of operations, and sailing), the service code, the service type, the number of container moves, and the auxiliary productivity figures.

Treatment variable. The variable in question is the type of service, which defines two classes of vessel: mother vessels, large ocean-going vessels given priority treatment when an appointment is made, and feeder vessels, smaller vessels handled on a first-come, first-served basis. Although mother vessels make up 15% of the calls, they account for 40% of

all moves (the median number of moves per call being 3,554 as compared to 1,101 for the feeder vessels), and this fact offers a sound economic reason for giving them priority.

Cleaning. Cleaning was carried out before the analysis; for each case, it was assumed that the error affected just one record. The extra spaces in a service code were removed, an incorrect year (changed from 2016 to 2026) and an incorrect day in the sailing times that came before berthing were corrected, two errors involving hours and years in the start-of-operations and sailing times were fixed, and one mistake relating to a month in the completion-of-operations time that had distorted the length of the operations was amended. After the cleaning process, the number of records (796), the date range and the distribution by class (679 feeders and 117 mother vessels) were checked automatically. Two further calls were then excluded before analysis (see below); the reproduction of the observed waiting statistics was checked automatically against the resulting 794-call sample.

Two calls, both feeders, were excluded before analysis. Each recorded a waiting time in excess of 14 days (827.9 hours and 402.8 hours respectively), far beyond any other observation. These are best read as schedule changes, a rebooking or a canceled-then-reinstated call, rather than genuine queueing waits, and their inclusion would distort the feeder waiting-time distribution on which the calibration and the reported statistics depend. After this exclusion, 794 calls remain (677 feeders and 117 mother vessels); this is the sample used throughout the rest of the thesis. Unless stated otherwise, ‘796’ below refers to the cleaned record before this exclusion and ‘794’ to the final analysis sample.

Derived variables.

- waiting time: wait = berthing time − arrival time
- service time: ops = completion of operations − start of operations
- berth occupancy: berth_occupancy = unberthing − berthing

3.2 The structural break and the definition of the regimes

There is a structural break on 1 June 2025, a result verified by a Chow test ($F = 75.7$, $p = 8.3 \times 10^{-31}$) applied to log(1 + wait), including an intercept and a class dummy, both allowed to shift. Examination of each possible week shows that the most significant break occurs on 9 June 2025, eight days away, so the date is well supported. After this break, the calls become smaller and more frequent. The median number of cargo moves per mother-vessel call drops from 6,199 to 3,378, while weekly arrivals increase from 8.3 to 10.0. Determining the cause

of this change is, however, beyond the scope of the present study. The data are divided into two regimes according to the berthing date:

- Regime A "congested": before 1/6/2025 (n = 179)
- Regime B "normal": from 1/6/2025 (n = 615)

The analysis that follows is carried out separately by regime and vessel class, because both factors have a real effect on both waiting times and service times.

3.3 Input analysis

The model is driven by two random processes: the arrival process and the service-time process, both estimated separately by regime and class.

Arrival process. Two criteria were used: (i) testing whether the times between arrivals could be fitted with an exponential distribution using a Kolmogorov–Smirnov test, and (ii) examining the variance-to-mean (dispersion) ratio of the weekly arrival counts. The exponential distribution is seldom rejected when testing inter-arrival times, since the combination of numerous fixed weekly schedules produces intervals that closely resemble an exponential distribution (the Palm–Khintchine theorem). The criterion that actually discriminates is the dispersion of the weekly counts, which is below one in every subgroup (ranging from 0.51 to 0.91) and notably under-dispersed in the largest subgroup (feeders, regime B; $p = 0.001$). Since the arrivals are therefore scheduled rather than following a Poisson process, the inter-arrival times in the model are generated by resampling with replacement from the observed gaps. It is better to use empirical resampling than to fit a parametric distribution since none of the theoretical distributions that were tested was able to capture the under-dispersion and structure present in the actual arrival pattern; empirical resampling retains both the rate and the variability of real-world arrivals without imposing a predetermined shape. To be clear, this is resampling with replacement, not trace replay. In each case, a gap is randomly selected with replacement from the set of observed gaps for the corresponding regime and class, and the recorded sequence of arrivals is never reproduced in the order in which it was originally recorded. As a result, the number of arrivals in each replication is stochastic and not fixed.

Service time. Service time was analyzed using the lognormal, gamma and Weibull distributions, fitted by maximum likelihood. None of these fits all of the subsets adequately (the largest subset is rejected by the KS test, $p = 0.028$). Service time also varies considerably

by regime (Mann–Whitney U, $p < 0.001$). During the congested regime vessels work considerably longer (mother-vessel median 85 h versus 37 h; feeder median 45 h versus 30 h). Service time is therefore also obtained by empirical resampling, separately by regime and class, so as to retain the true shape and tail of the distribution.

3.4 Simulation model

The model was built as a discrete-event simulation using SimPy 4.1 (Python 3.12).

Resource. The resource consists of four identical berth slots, a realistic rather than a fitted number. Once a vessel has come alongside, it occupies one slot until its service is complete.

Unobserved competition: the phantom stream. The data record only the carrier's own ships, even though other lines call at the same berths and compete for the same slots. An earlier version absorbed this competition into the berth count itself, making c an "effective" capacity. That approach is not adopted here, for the reason given in §3.5. The quantity being estimated is not an integer, so no whole number of berths could correspond to the data.

Instead, the competition is modeled explicitly. An independent arrival stream of phantom vessels stands in for the ships of other carriers. A phantom vessel queues and occupies a slot exactly as any other vessel does, is never given an appointment, and is excluded from every reported statistic. Its only function is to consume capacity. Phantom arrivals follow a Poisson distribution with rate

$$\lambda_{phantom} = (\rho \cdot c) / \text{mean\_phantom\_service\_time}$$

so that competitors are expected to use a share ρ of total quay capacity. ρ is the only free parameter in the model and is calibrated in §3.5. Phantom service times are drawn from the empirical service-time distribution of the feeder class in the matching regime; this rests on the assumption that other carriers' vessels are broadly feeder-like in size, which should be read as an assumption rather than a factual observation. The results depend on this assumption. If competitor service times were in reality considerably longer or shorter than those of feeders, for example if a considerable share consists of larger, mother-vessel-like ships or of smaller, faster-turnaround vessels, the amount of quay capacity these competitors occupy would change. Longer competitor service times would mean the same ρ translates into heavier quay use, and potentially longer waits for every class; shorter competitor service times would mean the model overstates the congestion competitor traffic actually causes.

Since there is no direct data on competitor call patterns, the sensitivity of the conclusions to this assumption is a limitation of the study, and the results should be read with that in mind.

Service time. Quay cranes are not modeled explicitly. This is a deliberate simplification, since the research question concerns queue priority, which vessel berths first, rather than the working rate once a vessel has berthed. It is a claim supported by evidence rather than a bare assertion. The median gang allocation for mother vessels is 6 in both regimes, so crane assignment does not differ between the periods in a way that could confound the comparison. The berth-holding time equals the empirical service time ops (§3.3).

It is still possible that omitting the explicit assignment of cranes could lead to bias, for example if conflicts between vessels regarding cranes cause delays which are not included in the total service time, or if the varying way cranes are allocated systematically disadvantages a certain class. In cases where crane bottlenecks are important from an operational point of view or where crane priority is linked to the berth allocation policy, the lack of explicit crane modeling might either underestimate or overestimate the waiting times for particular classes. The fact that these possible effects are present is recognized as the model's boundary.

Arrival process. Two independent arrival streams are generated, one per class, by bootstrap sampling of the inter-arrival times of the respective regime. Service time and container moves are drawn together from the same observed call, preserving the link between cargo size and handling time. This is necessary because scenario S3 prioritizes on cargo, and cargo and service time are the same underlying quantity observed twice.

Pluggable priority policy. The policy determines which classes are eligible for an appointment and which key orders the queue, and so realizes the four scenarios of §3.6. Priority is a key under the convention that a lower value is served first:

Table 3.1. Priority-key definitions by queue-ordering rule

Vessel	Key
holding an appointment	(0, slot_time, arrival_time)
arrival order (FCFS)	(1, arrival_time, 0)
cargo order (S3)	(1, −container_moves, arrival_time)

All keys share a three-element shape so that vessels under different rules, and phantom vessels, remain comparable element by element. Phantom vessels take their place in the

queue under whatever rule is in force. Under S3, a competitor with a large load outranks a small one, exactly as a carrier vessel would.

Appointment mechanics. Appointments are only allocated at shift boundaries (07:00 / 15:00 / 23:00, that is, every 8 hours). Each appointment vessel is assigned the nearest shift boundary as its scheduled time; if it arrives before that time it waits, but if it arrives within the one-shift tolerance (8 hours) it moors and retains its priority. Because the slot is defined as the nearest boundary, all arrivals must fall within the tolerance by definition; it follows that no appointment is lost in the current model. The simulation of schedule non-adherence and therefore appointment loss is left as a limitation.

Berthing lead time. The observed waiting time includes an unavoidable component of approach, pilotage and mooring that is present even without a queue. This was estimated from the 10th percentile of the observed waiting time in the normal regime (about 1.7 hours for mother vessels and 1.3 hours for feeders, figures independent of regime) and is added as a constant offset to the waiting time produced by the model, which otherwise yields only the pure queueing wait.

Run parameters. For each replication the time horizon is 104 weeks, with the first 4 weeks being discarded as a warm-up period, and there are 400 independent replications in each scenario. The reasons for these two choices are given empirically in §3.7 and not just stated.

Common random numbers. The seed is based solely on the replication number, never on the scenario, on c, or on rho. As a result, every scenario within a replication experiences identical arrivals and service times, so scenario-to-scenario comparisons are paired and their confidence intervals narrower. This was confirmed directly by comparing the vessel-by-vessel traces of S0, S1, S2 and S3 at a common seed. Arrival and service times agree exactly across all four, on 841 vessels in the congested regime and 974 in the normal one.

Summary of assumptions. (i) The cranes are not based on a model; instead, the service time is determined on the basis of empirical data. (ii) The slots are regarded as being the same; no account is taken of the spatial arrangement or of the length of the vessel. (iii) It is not assumed that any appointment is lost. (iv) The berthing lead time is fixed according to the class of vessel. (v) It is assumed that the competitor vessels have a service-time profile similar to that of feeders.

3.5 Calibration of the competitor load

3.5.1 Why the berth count is no longer the fitted parameter

The previous formulation searched over integer berth counts c = 1,…,6, using the weighted score

Score(c) = 2·|p90_sim − p90_obs|/p90_obs + 1·|median_sim − median_obs|/median_obs

assessed on the feeder class. This search is retained as a diagnostic; its loss surface is the reason it was abandoned. Reported by regime and class, it shows the two classes select different integers:

Table 3.2. Feeder- and mother-vessel loss score by candidate berth count c, by regime

Regime	Feeders select	Loss	Mother vessels select	Loss
A congested	c = 3	0.67	c = 2	1.38
B normal	c = 3	0.90	c = 4	0.25

In the congested regime, the mother-vessel loss never falls below 1.38 at any value of c. In neither period can a whole number of berths be found to suit both classes, because the quantity the search is really estimating, the capacity left over once competition has taken its share, is not an integer. Choosing c = 3 was therefore an elimination among unsatisfactory alternatives rather than a genuine fit.

3.5.2 Calibrating rho

With the number of berths fixed at four and the competition modeled explicitly (§3.4), the only free parameter is rho, the share of quay capacity used by other carriers. It is calibrated separately per regime against the feeder waiting distribution, feeders serving as the congestion signal since they carry no priority protection, make up 85% of all calls, and their tail carries the finding. The same weighted score is reused with rho in place of c:

Score(rho) = 2·|p90_sim − p90_obs|/p90_obs + 1·|median_sim − median_obs|/median_obs

The sweep goes from rho = 0.00 to 0.50 in 0.05 increments, and then it adjusts with steps of 0.01 over the range chosen in the initial run, carrying out 400 replications at each point; no value had been specified in advance.

Result. Both regimes calibrate to approximately 0.25. That is, a quarter of the quay is occupied by other carriers. The estimate is deliberately reported to that precision. Repeating the calibration under a different family of random seeds, with everything else unchanged,

moves it by ±0.02 to ±0.03, a range wider than the 0.01 grid on which the minimum is located; quoting two decimal places would therefore suggest a precision the data do not support.

Finding. Competitor pressure on the quay is the same in both periods. The longer waiting times are therefore not due to the carrier facing more competition for berths during the busy period. Section 4.3 of the Results chapter identifies what actually did change.

3.6 Scenarios and experimental design

Four policy scenarios are compared, each under both regimes:

Table 3.3. Description of the four policy scenarios (S0-S3)

Scenario	Description
S0	Hybrid policy, as operated: mother vessels by appointment, feeders by FCFS — the baseline validation scenario
S1	Pure FCFS: all vessels in a single queue, in arrival order
S2	Universal appointment: every vessel receives a shift-boundary slot
S3	Cargo-weighted priority: the vessel with the most container moves is served first, whatever its class

Each scenario is run at the calibrated (c = 4, rho), with 400 independent replications; for each vessel class, the mean and 95% confidence interval of the median, mean and p90 of the waiting time are reported.

A metric beyond waiting time. Comparing policies on waiting time alone answers who waits, not what the waiting costs. A fifth measure is therefore reported for all four scenarios, the cargo-weighted delay, the sum over vessels of waiting hours multiplied by container moves, in container-hours:

cargo_weighted_delay = Σ (waiting_hours × container_moves)

What the statement does is to change a descriptive observation, namely that appointments merely redistribute waiting rather than eliminate it, into a normative one by asking which rule will achieve that redistribution at the lowest cost to the cargo that is actually being carried.

The standing of S2. Under the present queue discipline, a vessel's eligibility time is max(arrival, appointment slot). When every vessel is assigned a slot shortly after arriving,

eligibility rises steadily with arrival time, and S2's queue order collapses onto arrival order, that is, onto S1's. This was tested directly rather than assumed. Running S1 and S2 from a common seed and comparing the order in which vessels actually berthed gives an order agreement of 0.998 in the congested regime and 0.994 in the normal one, with a Kendall rank correlation of exactly +1.000 in both. Not one pair of the carrier's vessels is ever reversed between the two policies.

S2 is therefore reported as an internal consistency check on the model, not as an independent policy alternative. A further caution applies to any S2 figure quoted in isolation. Under S2 the carrier's whole fleet holds appointments and so occupies priority tier 0, ahead of the phantom stream; under S1 the same vessels sit in tier 1 and compete with other carriers on equal terms. Removing the competitors (rho = 0) makes S1 and S2 indistinguishable, 25.0 h against 25.8 h on the feeder p90, while at the calibrated load they differ by more than a factor of two. That difference reflects the carrier's fleet collectively overtaking its competitors, not an effect of the appointment mechanism, which is why S2 is excluded from the cargo-weighted-delay ranking.

3.7 Run parameters, justified

Warm-up. Because the quay starts out empty, the early vessels have to wait for less time than they do in the steady state. If you plot the average waiting time of the n-th vessel to arrive over 200 replications (Figure 3.1), you can see the transient effect clearly. In the congested situation the average waiting time increases from 17.8 hours for the first 25 vessels to a steady value of 42.6 hours, which is reached by about the 50th vessel, while the four-week warm-up period eliminates 33 hours. On its own, this might indicate that the warm-up is too short. However, the direct test shows that this is not the case. Here is the reported feeder p90 in relation to the length of the warm-up:

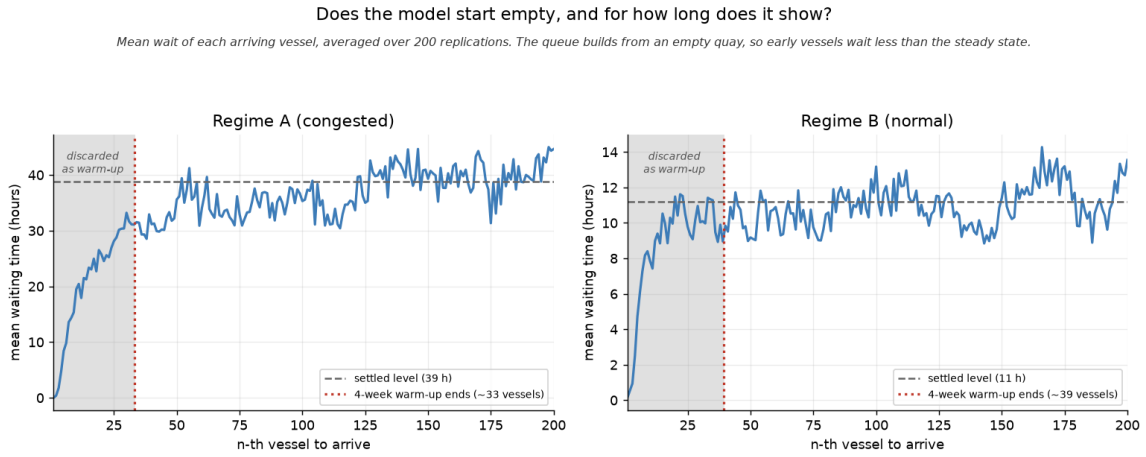

Figure 3.1. Warm-up transient. Mean waiting time of the n-th arriving vessel, averaged over 200 replications, showing the settling of the queue from an empty initial state to its steady-state level.

Table 3.4. Reported feeder p90 waiting time by warm-up period length and regime (hours)

Warm-up	Regime A	Regime B
4 weeks	134.9	41.0
8 weeks	134.5	41.3
12 weeks	134.9	41.6
16 weeks	135.2	41.8

Quadrupling the warm-up shifts the reported figure by 0.3 h in the congested regime and 0.8 h in the normal one, both well inside the confidence interval, because the residual transient spans roughly 20 vessels out of some 840 in a 104-week run. Four weeks is adequate.

The absolute level in this table should not be compared directly with the calibrated congested-regime feeder p90 of 129.6 h reported in Chapter 4 (Table 4.4). This warm-up check was run to test the sensitivity of the reported figure to warm-up length alone, holding the other run settings used at that stage fixed; Chapter 4 reports the figure produced under the final calibrated settings (§3.5). The two are not directly comparable figures for that reason, but the conclusion drawn from this table, that four weeks of warm-up is adequate, rests on the differences between its rows, not on their absolute level, and is unaffected by which settings were current at the time.

Replication count. The 95% confidence-interval half-width of the p90:

Table 3.5. 95% confidence-interval half-width of the p90 waiting time, by replication count, regime and class

Regime	Class	30 reps	100 reps	400 reps
A congested	mother	±0.5 h (2.4%)	±0.4 h (1.9%)	±0.2 h (1.0%)
A congested	feeder	±11.4 h (9.7%)	±10.9 h (8.8%)	±5.7 h (4.4%)

B normal	mother	±0.6 h (4.0%)	±0.3 h (2.0%)	±0.2 h (1.2%)
B normal	feeder	±4.4 h (10.4%)	±2.1 h (5.1%)	±1.1 h (2.6%)

Three of the four cells behave conventionally. The feeder p90 in the congested regime does not. Its half-width barely moves between 30 and 100 replications, and the point estimate rises. This is due to the shape of the across-replication distribution, which over 400 replications has a mean of 129.6 h, a median of 114.3 h, a standard deviation of 57.9 h and a skewness of +2.36, with a maximum of 535 h. A small sample under-represents that right tail, so the mean and the variance estimate grow together and the two effects cancel out in the half-width. The replication count was therefore fixed at 400 for all reported results, at which the estimate settles to about ±4.4%. This is itself a finding about the model. With a quarter of the quay taken by competitors and only four berths available, the congested regime sits close enough to saturation that some replications produce very long queues.

3.8 Software and reproducibility

The analysis was carried out using Python 3.12 and the following libraries: SimPy 4.1, NumPy, pandas, SciPy and statsmodels. The workflow is divided into separate stages, each a single script that saves its intermediate outputs: data cleaning and anonymization, input analysis, the model itself, calibration, scenario runs, and the diagnostic and stability checks. Seeding follows the single shared rule described in §3.4, so every reported figure can be reproduced from the raw file.

At the cleaning stage the identifiers are replaced. Vessel calls become V001…V794 in order of arrival, and services become S01…S19. A separate file holding the mapping between real and anonymized identifiers is created and is never committed or shared.

3.9 Statement regarding the use of an AI tool

What the tool was used for. Claude Code (Anthropic), an AI coding assistant, was used throughout the implementation of the Python 3.12 codebase behind this chapter and the next: for the data-cleaning and anonymization scripts (§3.1), the input-distribution fitting (§3.3), the SimPy 4.1 discrete-event simulation model itself (§3.4), the calibration sweep (§3.5), and the scenario, diagnostic and soundness code (§3.6–§3.7). It was used to write, debug and refactor Python code to a specification set by the author, and, in a limited number of places, to draft descriptive text from tabulated numerical output, which the author then reviewed

and edited. It was used for the same reason a statistical package is used under expert supervision. It sped up the mechanical, syntactic parts of the implementation, for what is, at this stage, an educational simulation exercise.

What it was not used for. The formulation of the research question, the definition of the two regimes, the phantom-vessel formulation and its calibration target, the specification of the four policy scenarios, the metrics reported, including the cargo-weighted delay, and the interpretation of every result are the author's own work. No numerical result in this chapter or the next was accepted without independent verification against the underlying dataset. Reported figures were checked against the raw data, and the automated checks described in §3.1 (record count, class distribution, reproduction of the observed waiting statistics) were re-run after every substantive change to the codebase.

Responsibility. Research design, the reading of the results, and responsibility for their correctness rest with the author, not with the tool.

3.10 Limitations

1. Competitor traffic is modeled, not observed. Rather than the earlier device of absorbing competitors into the berth count, a phantom stream is introduced, which removes the integer problem but is still a model of something that cannot be measured directly. Its service times are assumed to follow the feeder distribution. This is an assumption about other carriers' fleet profile, not a measurement.

2. The mother-vessel queue is not reproduced. The calibration process addresses the feeder distribution and matches it very closely (the model gives a p90 of 129.6 hours compared to the observed value of 130.0 hours). However, the mother-vessel wait is not well reproduced. When there is congestion, the model produces a p90 of 22.6 hours, whereas the observed figure is 96.8 hours. It is impossible for a single shared resource to at the same time support both a small protected queue and a heavy unprotected one; although the direction of the findings is reliable, the model exaggerates the extent of mother-vessel protection. Refinements could improve the model's fidelity here: separate resource pools for priority and non-priority vessels, explicit class-based berth access, variable protection levels, or a hybrid resource structure reserving certain berths or time periods for mother vessels. Such changes would require more detailed data than is currently available, but they represent clear routes for future development.

3. The observed mother-vessel figure is itself uncertain. It rests on 27 calls. Bootstrapping the observed p90 with 10,000 resamples gives a 95% confidence interval of 66.3 to 252.7 hours around the point estimate of 96.8 hours. This interval is almost twice the size of the estimate. No conclusion should rest on its exact value.

4. The congested period is not internally uniform. The monthly median values vary between 27.6 and 71.0 hours, and for March 2025 the p90 is 193.2 hours, as opposed to 36.9 hours for the entire period. The figures for 'the congested regime' are averages calculated over months that differ by a factor of two.

5. Omission of cranes. Service time is determined empirically rather than being modeled; crane assignment and working rate are not modeled. The stability of the gang count across regimes (§4.3) reduces, but does not eliminate, the risk this poses.

6. No spatial dimension. Slots are treated as identical; neither vessel length nor position along the quay is taken into account.

7. No appointment loss. Schedule non-adherence, and the resulting loss of an appointment, are not modeled. This applies to S0, S2 and S3 alike.

8. Constant berthing lead time. The approach and pilotage component is treated as constant per class.

4. Results

The empirical findings from the data, the calibration result, and the comparison of the four policy scenarios are given in this chapter. Times are stated in hours unless otherwise noted. "p90" refers to the 90th percentile of waiting time, the wait exceeded only by the worst vessel in ten, meaning it relates to tail risk rather than the typical call.

The figures which have been simulated are based on four hundred replications for each scenario, using the same random numbers in each replication for the various scenarios.

4.1 Empirical waiting-time patterns

Observed waiting times to berth, by regime and class (n = 794):

Table 4.1. Observed waiting time to berth, by regime and class (n = 794)

Regime	Class	n	Median	Mean	p90
A congested	mother (appointment)	27	31.0	50.0	96.8
A congested	feeder (FCFS)	152	38.9	54.3	130.0
B normal	mother (appointment)	90	3.0	7.2	6.9
B normal	feeder (FCFS)	525	3.1	15.2	47.1

The result is located in the tail, not in the middle. Within each regime the medians are almost the same, 3.0 against 3.1 hours in the normal period, although the p90 values differ several-fold. In the normal period the feeder p90 is 6.8 times the mother-vessel p90 at the chosen break date. This ratio depends on where the two periods are divided. It is 8.2 with a break at 1 July and 3.0 with a break at 1 May (§4.5). Whatever choice is made, it stays more than threefold. The interpretation does not rely on the exact multiple. The appointment system does not make the typical call any faster; it removes the risk of a very bad one. It provides insurance, not speed.

The mother-vessel figures for the busy period rest on 27 calls and are correspondingly uncertain. Bootstrapping the observed p90 with 10,000 resamples gives 96.8 hours with a 95% confidence interval of 66.3 to 252.7 hours. This interval is given wherever that figure is referred to below.

Fairness and financial rationale. Priority for mother vessels rests on economic grounds, not exclusively commercial ones. Although they make up 117 of the 794 calls (14.7%), they are

responsible for 40.2% of container moves (523,528 of 1,303,473), with a median of 3,554 moves per call against 1,101 for feeders.

4.2 Input characteristics

Arrivals. The dispersion index of weekly arrivals is below one in every subset (0.51 to 0.91) and significantly under-dispersed for feeders in the normal regime ($p = 0.001$). Arrivals are therefore scheduled, not Poisson. In the model they are generated by bootstrap sampling of the observed gaps, drawn at random with replacement and never reproduced in the order in which they were recorded.

Service time. Service time differs markedly by regime (Mann–Whitney, $p < 0.001$). Vessels work substantially longer during the congested period (mother-vessel median 85.0 h versus 37.2 h; feeder median 44.8 h versus 30.5 h).

4.3 What actually distinguished the congested period

The longer service times of the congested period could reflect a slower terminal, a larger cargo per call, or both; the question is answered directly rather than inferred, and the answer differs by class.

Call size and productivity, by class and regime (medians):

Table 4.2. Call size and productivity by class and regime (medians)

Class	Regime	n	Moves	Gangs	Service h	Moves/hour
mother	A	27	6,199	6	85.0	73.7
mother	B	90	3,378	6	37.2	88.1
feeder	A	152	1,222	3	44.8	27.1
feeder	B	525	1,045	4	30.5	33.9

Did the terminal slow down? Yes, and noticeably. Median gang productivity, moves per crane-hour, the rate at which an individual gang works, fell from 18.8 to 16.1 for mother vessels (a decrease of 14.4%) and from 17.1 to 13.6 for feeders (a decrease of 20.5%). Regressing the logarithm of the terminal's own reported gang productivity on regime and class shows the congested period to be 26.9% less productive per crane-hour (95% CI 24.1–29.7%, $p < 10^{-77}$); adding controls for cargo size and gang count raises this to 28.2%. This is a genuine finding, not a null result, and it is reported as one.

But for mother vessels, cargo grew faster than the terminal's overall handling rate fell. This second decomposition uses a different measure from the gang productivity above. It is the terminal's aggregate moves-per-hour rate for the call, moves divided by service time, which reflects the number of gangs deployed as well as their individual speed, not the per-gang (crane-hour) productivity just described. Separating the rise in service time multiplicatively, since service time is cargo divided by this moves-per-hour rate:

Table 4.3. Decomposition of the service-time increase into cargo growth and the decline in terminal moves-per-hour (aggregate handling rate, not gang productivity), by class

Class	Service time	Share from larger cargo	Share from lower moves-per-hour
mother	×2.28	73.5%	21.6%
feeder	×1.47	40.7%	58.2%

For mother vessels the two shares do not sum to 100% (73.5% + 21.6%). About 4.9% of the increase is not attributed to either channel. This residual comes from the multiplicative interaction between cargo growth and the moves-per-hour decline, not from a third factor.

For mother vessels, roughly three quarters of the longer service time is simply more cargo. Median moves per call rose from 3,378 to 6,199, an increase of 84%. For feeders the balance reverses. Cargo rose only 17%, and close to 60% of their service-time increase comes from the decline in terminal moves-per-hour. A single statement covering both classes would be wrong for one of them.

It was not a busier port. The congested period received 8.29 calls per week; the normal period received 9.97, a fifth more, and waited far less. This rules out increased arrival volume as an explanation, a point the productivity argument alone cannot make.

Gang allocation is stable. The median gang count for mother vessels is 6 in both regimes. The mean and upper quartile are in fact higher during congestion (7.48 and 10.5, against 6.19 and 8.0). The terminal put more cranes on the larger calls, not fewer.

A secondary check. Regressing service time on cargo, gangs, class and regime gives 0.0075 hours per additional container move, or 7.5 hours per thousand moves, significant at $p < 10^{-64}$. The regime term in that model remains large, retaining 88% of the raw 16.45-hour gap between periods, and this holds across specifications (57–90% depending on interactions and functional form). The reason is that a linear cargo term absorbs a multiplicative quantity poorly. Its implied marginal rate of 133 moves per hour is far above the observed median productivity of 34. The decomposition set out above, not this regression, is the evidence for

the conclusion; the regression is reported for completeness, and the residual regime effect is not claimed to be small.

4.4 Calibration

With the quay fixed at four berths and competitor traffic modeled as an explicit phantom stream, the calibrated competitor load is approximately 0.25 in both regimes. This is about a quarter of the quay occupied by other carriers. Competitor pressure is therefore indistinguishable between the two periods, consistent with §4.3. What changed was cargo size and working rate, not the contest for berth space.

Validation of the baseline against the collected data:

Table 4.4. Validation of the simulated baseline (S0) against the observed data

Class	Regime	Metric	Observed	Simulated (95% CI)	Difference
mother	A	median	31.0	7.0 ± 0.1	−77.6%
mother	A	p90	96.8 (CI 66.3–252.7)	22.6 ± 0.2	−76.6%
feeder	A	p90	130.0	129.6 ± 5.7	−0.3%
mother	B	p90	6.9	14.6 ± 0.2	+111.5%
feeder	B	p90	47.1	42.4 ± 1.1	−10.0%

The feeder distribution, the calibration target and the class that carries the finding, is reproduced closely, to within 0.3% on the congested-period tail. The mother-vessel queue is not. Under congestion the model gives 22.6 h against an observed 96.8 h, and even against the lower bound of the bootstrap interval (66.3 h) the gap remains large, so this is a genuine limitation of the model rather than an artifact of the small sample. Under normal conditions, the model overstates the mother-vessel wait instead. A single shared resource cannot reproduce both a small protected queue and a heavy unprotected one at the same time; the qualitative redistribution finding is robust, but the magnitude of mother-vessel protection is not.

4.5 Robustness of the regime split

The interruption on 1 June 2025 was initially selected by means of an examination of the monthly medians and has withstood the tests.

When a Chow test is carried out on log(1 + wait) on that date with both the intercept and the class dummy allowed to change, the result is F = 75.7 and p = 8.3 × 10^{-31}. Among all the possible weeks examined, the most significant break occurs on 9 June 2025 (with an F value of 88.9), which is eight days away; furthermore, the five most prominent candidates are all within a month of the selected date (Figure 4.1).

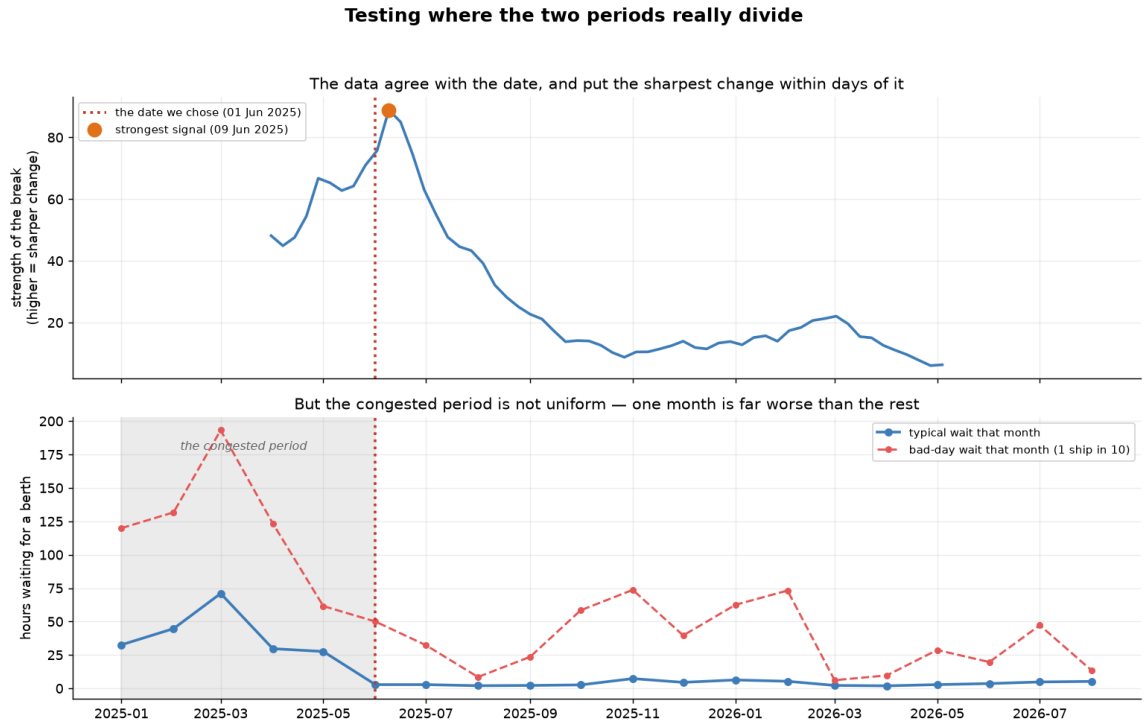

Figure 4.1. Chow-test F-statistic for every candidate break week, identifying the structural break in the data.

Re-running the descriptive split at alternative dates confirms the substantive conclusions but not any single headline number:

Table 4.5. Sensitivity of key waiting-time statistics to the choice of regime break date

Break	mother p90, A	feeder p90, A	mother p90, B	feeder p90, B
1 May 2025	100.6	144.0	15.7	47.7
1 June 2025	96.8	130.0	6.9	47.1
1 July 2025	99.1	123.7	5.6	46.1

The ordering holds under every choice. The ratio between the feeder and mother-vessel p90 in the normal period does not. It ranges from 3.0 to 8.2, which is why §4.1 reports it as a range.

The congested period is not internally uniform. Its monthly medians run 32.6, 44.7, 71.0, 29.7 and 27.6 hours from January to May 2025, against 36.9 hours for the period as a whole.

March 2025 is the outlier in both the center and the tail, with a median of 71.0 hours and a p90 of 193.2 hours. Statistics reported for "the congested regime" are averages over months that differ by a factor of two, and March contributes more to the picture than its share of calls would suggest.

4.6 Policy-scenario comparison

Each scenario was run at the calibrated capacity with 400 replications. The results are shown in Figure 4.2, Figure 4.3, Figure 4.4 and Figure 4.5.

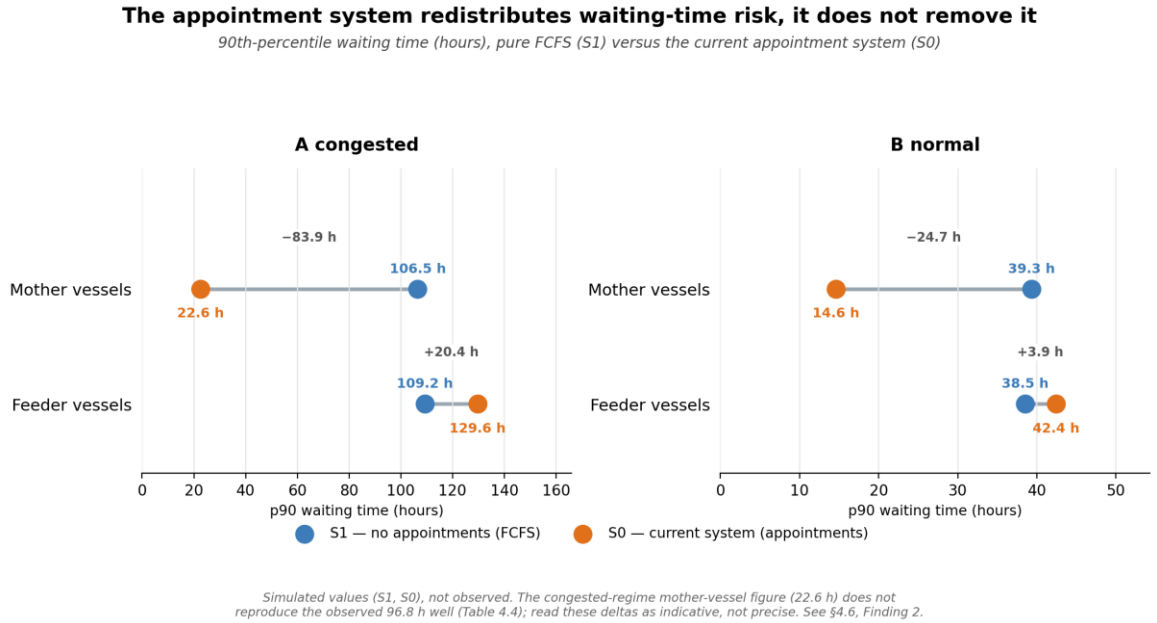

Figure 4.2. Redistribution of waiting-time risk between mother vessels and feeders under the appointment system (S0) relative to pure FCFS (S1).

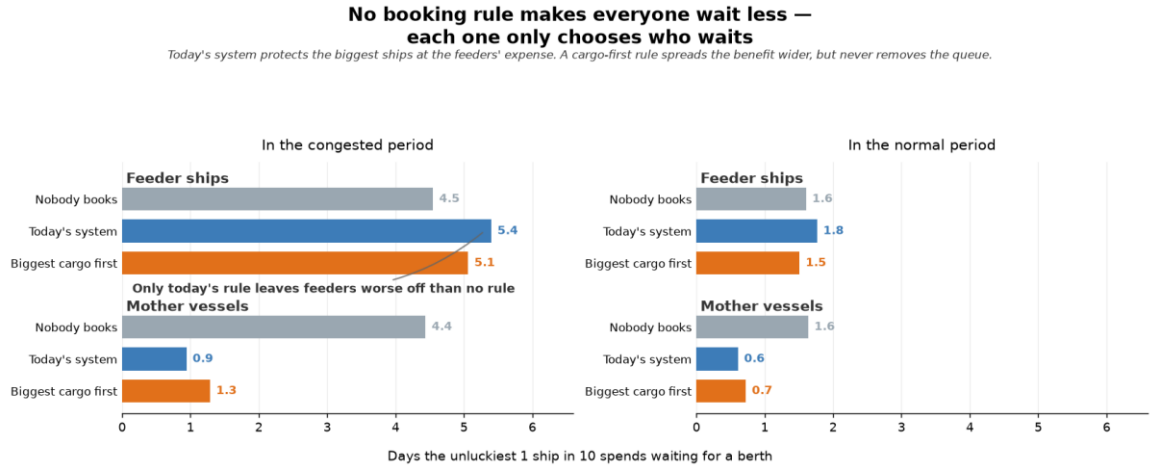

Figure 4.3. Median and 90th-percentile waiting time by scenario (S0–S3), regime and vessel class.

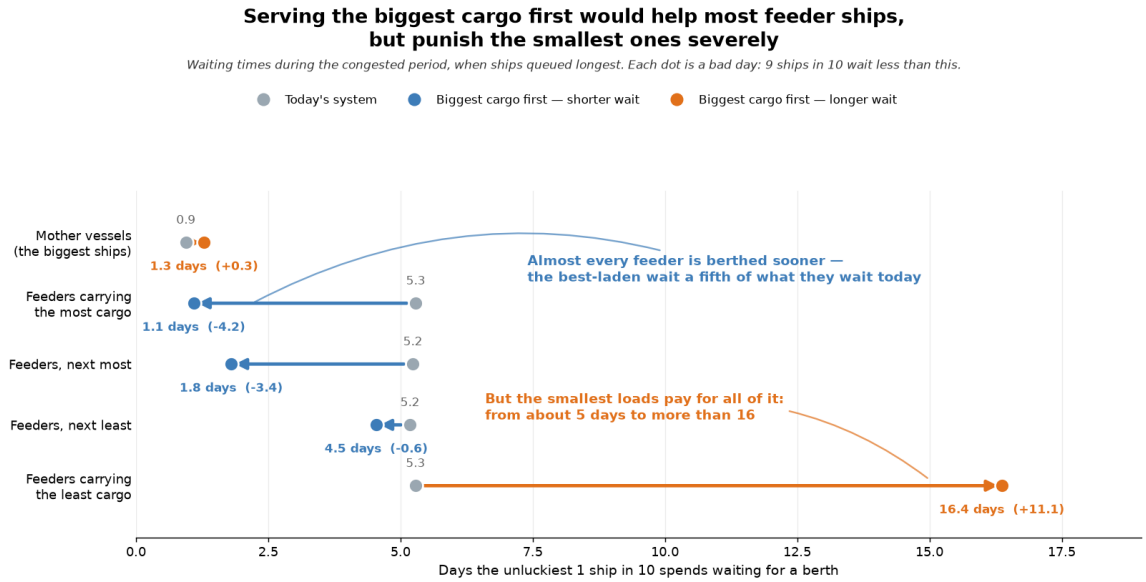

Figure 4.4. Effect of the cargo-weighted priority rule (S3) on the feeder waiting-time distribution, by cargo quartile.

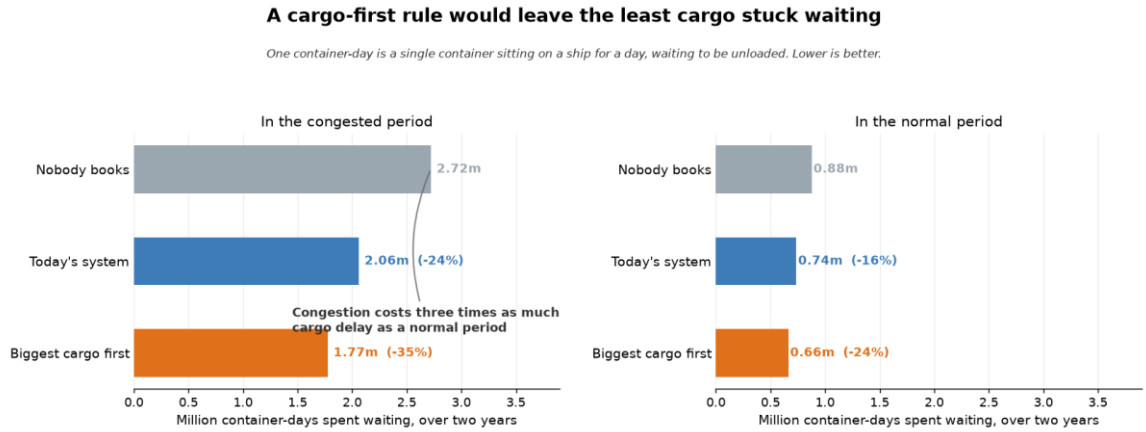

Figure 4.5. Total cargo-weighted delay (container-hours) by scenario and regime.

Waiting time in hours, mean over 400 replications:

Table 4.6. Mean waiting time (hours) by scenario (S0-S3), regime and vessel class

Regime	Policy	Mother: median	Mother: p90	Feeder: median	Feeder: p90
A congested	S1 no appointments	24.2	106.5	25.4	109.2
A congested	S0 mother booked	7.0	22.6	27.8	129.6
A congested	S2 all booked (check only)	11.4	40.5	12.2	43.1
A congested	S3 biggest cargo first	6.4	30.9	10.4	121.5
B normal	S1 no appointments	4.1	39.3	3.3	38.5

B normal	S0 mother booked	4.5	14.6	3.4	42.4
B normal	S2 all booked (check only)	5.1	21.2	4.6	21.2
B normal	S3 biggest cargo first	2.5	17.3	2.3	36.2

Finding 1. Redistribution, not reduction. Comparing S0 with S1 in the same regime, the mother-vessel p90 falls from 106.5 to 22.6 h under congestion and from 39.3 to 14.6 h in normal conditions, while the feeder p90 rises from 109.2 to 129.6 h and from 38.5 to 42.4 h respectively. The appointment system transfers tail risk from mother vessels to feeders; total waiting time is nearly unchanged. It is priority, not speed.

Note also that in the normal period the mother-vessel median rises slightly under appointments, from 4.1 to 4.5 h. Mother vessels pay a little more typical waiting to remove the risk of a very bad call. This is the shape of an insurance premium.

Finding 2. The picture depends on whether protection is measured in hours or as a ratio. Observed p90 waiting times are 96.8 h for mother vessels against 130.0 h for feeders in the congested regime, and 6.9 h against 47.1 h in the normal regime (§4.1). Comparing the simulated S1 and S0 scenarios, the mother-vessel p90 saving from having an appointment is about 84 h under congestion and about 25 h in normal conditions, so in absolute hours the appointment system appears to matter most when the terminal is under pressure. This absolute comparison should be read as indicative rather than as a measurement. It rests on the simulated congested-regime mother-vessel queue, which the calibrated model does not reproduce well. The model returns a p90 of 22.6 h for that cell against an observed 96.8 h (§4.4, Table 4.4), so the 84 h figure is likely overstated, and no conclusion should rest on its exact size.

Expressed as a ratio rather than in hours, the relationship reverses. The observed feeder p90 is 6.8 times the mother-vessel p90 in the normal regime, but only 1.3 times in the congested regime (§4.1). Proportionally, appointments buy mother vessels more protection when the terminal is quiet, the opposite of what the absolute-hour comparison above suggests. This tracks the direction of the model's own mismatch reported in §4.4. The model understates the mother-vessel wait under congestion (22.6 h simulated against 96.8 h observed) and overstates it under normal conditions (14.6 h simulated against 6.9 h observed), errors running in opposite directions rather than as noise, which points to a specific simplification.

The model treats the four berths as a single, fully shared and interchangeable pool, whereas mother vessels in practice tend to need one dedicated or size-appropriate berth. Under congestion the pooling assumption lets a simulated mother vessel take whichever of the four berths frees first, whereas in practice it must wait for the one berth long enough to accommodate it. The simulated queue therefore clears far faster than the real one, which is why the model understates mother-vessel waiting in that regime. Under normal conditions, the same pooling assumption still exposes mother vessels to some competition for the shared pool that a dedicated berth would remove, making the model too pessimistic there. Neither framing should be read as the more accurate one without this caveat; the qualitative conclusion that appointments redistribute protection is robust, but the magnitude and its ordering across congestion regimes should not be treated as precise.

Finding 3. A universal appointment changes nothing internally. Under S2 the order in which the carrier's own vessels berth is identical to S1. Kendall rank correlation is +1.000 in both regimes, with no pair ever inverted. When everyone has priority, no one does. S2's apparently short waits in the table above arise because the carrier's whole fleet then outranks the competitor traffic collectively, an effect of tier assignment against third parties rather than of the appointment mechanism, and S2 is therefore reported as a consistency check on the model rather than as a policy alternative.

Finding 4. Cargo-weighted delay. Total container-hours of delay, the measure of what the waiting costs rather than who bears it:

Table 4.7. Total cargo-weighted delay (container-hours) by scenario and regime

Regime	S1	S0 current	S3 cargo-first
A congested	65.3M	49.4M (−24.3%)	42.6M (−34.8%)
B normal	21.0M	17.7M (−16.0%)	15.9M (−24.4%)

Among the policies that can be meaningfully compared, S3 is the cheapest in both regimes, 13.8% below the current policy under congestion and 10.0% below it in normal conditions. The terminal's class-based priority rule therefore leaves value on the table relative to a pure cargo-volume rule.

Finding 5. But S3 concentrates the cost on the smallest consignments. The paired comparison of S3 against S0, differenced replication by replication:

Table 4.8. Paired difference between S3 and S0 in waiting-time statistics, by regime

Regime	Median	Mean	p90
A congested	−17.5 ± 1.2 h	+6.0 ± 0.6 h	−8.1 ± 3.6 h
B normal	−1.1 ± 0.1 h	+1.8 ± 0.1 h	−6.2 ± 0.6 h

All six differences are significant. The median falls sharply and the p90 falls as well, yet the mean rises, which places the damage above the 90th percentile. The feeder distribution under the two rules, pooled over 400 replications in the congested regime, shows exactly that:

Table 4.9. Feeder waiting-time distribution under S0 and S3 by percentile (congested regime)

Percentile	S0	S3	Difference
50th	25.6	10.3	−15.3
75th	69.0	32.7	−36.3
90th	127.2	115.3	−11.9
95th	172.8	242.0	+69.2
99th	285.3	781.6	+496.3

S3 improves every part of the distribution up to the 90th percentile and concentrates the entire cost in the last decile, where it is severe. The worst 1% of feeder calls goes from about 12 days of waiting to about 33.

Splitting feeders by cargo quartile identifies who pays (congested regime):

Table 4.10. Feeder waiting-time statistics under S0 and S3 by cargo quartile (congested regime)

Quartile	S0 median	S0 mean	S0 p90	S3 median	S3 mean	S3 p90
Q1 smallest	24.2	46.8	126.8	24.9	141.5	392.8
Q2	24.0	46.1	124.3	11.4	39.2	109.0
Q3	24.2	46.6	125.6	7.1	17.0	43.2
Q4 largest	24.7	47.2	126.9	5.3	10.0	26.2

Under the current rule the four quartiles are almost indistinguishable. The class-based rule is blind to size within the feeder fleet. Under S3 they separate sharply and the smallest quartile carries the whole cost. Its mean triples and its p90 reaches 392.8 h, more than two weeks. The mechanism is clear. Median mother-vessel cargo in the congested period is 6,199 moves against 1,222 for feeders, and 85% of mother-vessel calls exceed the feeder median, so a small feeder sits behind nearly the entire fleet at all times.

4.7 Synthesis

The reason for having an appointment system is not that it speeds up the average call, but rather that it provides protection for strategically important vessels during busy times; although mother vessels make up only 15% of the calls, they account for 40% of container movements, which brings about a corresponding increase in the costs for feeder vessels. It is like insurance for some, financed by others, and how large that premium looks depends on whether it is measured in absolute hours or as a proportion of the feeder baseline, since the two framings do not point the same way across congestion regimes (§4.6). From a policy point of view, the findings indicate that terminals and carriers can use appointment systems not to reduce overall waiting times, but to safeguard valuable or strategic cargo flows from the most severe delays when capacity is limited. In order to ensure reliability for their main services, carriers may request priority arrangements, while terminals have to take into account the effect of the priority rule on the way waiting costs are distributed among users, weighing commercial incentives against broader considerations of service fairness.

Extending that priority to the whole fleet (S2) changes nothing about who is served first among the carrier's own vessels. Replacing class with cargo volume (S3) does change it, lowering the total container-hours of delay by 10–14% against the current rule. The terminal's class-based priority is not the cheapest arrangement available. That gain is not free. It is financed almost entirely by the smallest consignments, whose worst-case waits roughly triple. The choice between the two rules is therefore not a technical one regarding efficiency, but a distributional one about which cargo the terminal is willing to leave waiting.

5. Conclusions

5.1 Summary of findings

This thesis set out to test, through discrete-event simulation, how an appointment-based berth-allocation policy affects vessel waiting time and the distribution of delay relative to pure FCFS, and what these effects imply for schedule reliability under different levels of congestion. Calibrated against 794 anonymized vessel calls across two demand regimes, the model shows that the terminal's current hybrid policy, mother vessels served by appointment and feeders by FCFS, does not reduce total waiting time. Instead, it redistributes tail risk. Mother-vessel p90 waiting time falls sharply under appointments, while feeder p90 waiting time rises by a comparable amount, and the apparent size of this transfer differs by congestion regime, although its ordering depends on whether protection is measured in absolute hours or proportional terms. The appointment system functions as insurance for strategically important vessels, financed by the vessels without that protection, rather than as a device for making the terminal faster overall.

A universal-appointment variant, in which every vessel receives a slot, changes nothing about the order in which the carrier's own vessels are served relative to pure FCFS. When everyone has priority, no one effectively does. A cargo-weighted priority rule, by contrast, does change outcomes. It lowers total cargo-weighted delay by 10–14% relative to the current policy, but it does so by concentrating additional waiting on the smallest consignments, whose worst-case waits roughly triple. The choice between these rules is therefore not primarily a question of technical efficiency, but a distributional one, about which cargo the terminal and its customers are willing to leave waiting when capacity is under pressure.

5.2 Contribution

The literature on berth allocation includes formal optimization studies, queueing and simulation analyses, and descriptions of appointment and pre-booking practice, but no existing study has tested a pro-forma and rendezvous system of the type documented at PCT (§2.7), including FCFS treatment of non-booked feeders, against a pure FCFS system using the same vessels, arrivals and handling assumptions, evaluated jointly on waiting time and a cargo-weighted delay measure. This thesis provides that controlled comparison, grounded in real operational data and an operational rule set documented at a live container terminal,

and shows that the value of an appointment system is best understood as risk redistribution rather than average-case speed. This conclusion is consistent with, but not directly demonstrated by, the wider literature reviewed in Chapter 2.

5.3 Limitations

The findings should be read alongside the limitations already detailed in Section 3.10. In summary: competitor traffic at the quay is modeled through a calibrated phantom stream rather than observed directly, and its service-time profile is assumed, not measured; the mother-vessel queue under congestion is not well reproduced by the calibrated model, so the direction of the findings is more reliable than their exact magnitude; the observed mother-vessel waiting statistics themselves rest on a small number of calls and carry wide confidence intervals; the congested period is not internally uniform, and one month drives much of the reported effect; quay cranes, vessel length and spatial berth position are not modeled explicitly; and the model does not allow for a booked vessel to miss its appointment. None of these simplifications is fatal to the central conclusion, that appointment systems redistribute rather than eliminate waiting, but they bound the precision with which any single number in Chapter 4 should be interpreted.

5.4 Recommendations for future research

Three extensions follow directly from the limitations above. First, richer data on competitor vessel calls, even aggregate counts or size classes, would allow the phantom-stream assumption to be tested rather than assumed. Second, separate or partially reserved berth resources for priority and non-priority traffic could be modeled explicitly, which the current shared-resource formulation cannot represent, and might close the gap between the model's mother-vessel waiting time and the value observed under congestion. Third, incorporating schedule non-adherence, the possibility that a booked vessel misses its appointment, and an explicit spatial or crane-capacity dimension would move the model closer to full operational realism. Beyond the model itself, applying the same comparative framework to other terminals and trades, and linking the cargo-weighted delay measure to commercial penalty or compensation structures, would help translate the distributional findings of this thesis into concrete policy guidance for terminals and carriers negotiating berthing-priority arrangements.

5.5 Concluding remarks

Berth-allocation policy is often discussed as if it were purely a technical efficiency problem, namely which rule minimises average waiting time. The evidence in this thesis suggests a different framing. Under realistic, stochastic terminal conditions, no policy examined here eliminates queueing; each one chooses, in effect, who bears the risk of a very bad day. Recognizing that choice explicitly, rather than treating today's practice as neutral, is the main practical contribution this thesis offers to terminals and carriers negotiating berthing-priority arrangements.

References

Abdel Hafez, M.A., Hubbard, N.J., Tipi, N.S. and Eltawil, A.B. (2013) 'A simulation based model for the berth allocation and quay crane assignment problem', paper presented at the LRN Annual Conference and PhD Workshop, Birmingham, 4–6 September. Available at: http://eprints.hud.ac.uk/id/eprint/19208/

Böse, J.W. (ed.) (2011) Handbook of Terminal Planning. New York: Springer.

Carvalho, C.C., Santos, R., Marques, C.M. and Pinho de Sousa, J. (2026) 'Just-in-time strategies for sustainable operations in seaports: a simulation approach', Transportation Research Procedia. doi: 10.1016/j.trpro.2026.02.072.

Ceder, S., Gustafsson, N., Nilsson, H., von Nordenskjöld, A., Stjernström, P. and Torbjörnsson, J. (2018) 'Port call synchronization', Bachelor's thesis. Chalmers University of Technology and University of Gothenburg, Gothenburg.

Charisis, A., Mitrović, N. and Kaisar, E. (2018) 'Container shipping route and schedule design with port time windows and coordinated arrivals', in Proceedings of the 21st International Conference on Intelligent Transportation Systems. Piscataway, NJ: IEEE, pp. 2538–2543. doi: 10.1109/ITSC.2018.8569562.

European Commission (2025) Decarbonizing maritime transport: FuelEU Maritime. Available at: https://transport.ec.europa.eu/transport-modes/maritime/decarbonising-maritime-transport-fueleu-maritime_en

European Maritime Safety Agency (2026) Reducing GHG emissions: ETS extension to maritime. Available at: https://www.emsa.europa.eu/reducing-emissions/extension-ets.html

Gkolias, M.D. (2007) The discrete and continuous berth allocation problem: Models and algorithms. PhD dissertation. Rutgers, The State University of New Jersey, New Brunswick, NJ.

Henesey, L., Davidsson, P. and Persson, J.A. (2004) 'Using simulation in evaluating berth allocation at a container terminal', in Proceedings of the 3rd International Conference on Computer Applications and Information Technology in the Maritime Industries (COMPIT'04). Parador Sigüenza, Spain.

Hendriks, M., Laumanns, M., Lefeber, E. and Udding, J.T. (2010) 'Robust cyclic berth planning of container vessels', OR Spectrum, 32(3), pp. 501–517.

International Maritime Organization (2020) Just In Time Arrival Guide: Barriers and Potential Solutions. Prepared by the GEF-UNDP-IMO GloMEEP Project and the members of the Global Industry Alliance. London: IMO. Available at: https://wwwcdn.imo.org/localresources/en/OurWork/PartnershipsProjects/Documents/GIA-just-in-time-hires.pdf (Accessed: 2 August 2026).

Karmelić, J., Mihanović, M.J., Hadžić, A.P. and Brčić, D. (2025) 'Liner schedule reliability problem: An empirical analysis of disruptions and recovery measures in container shipping', Logistics, 9(4), article 149. doi: 10.3390/logistics9040149.

Kontovas, C.A. and Psaraftis, H.N. (2011) 'Reduction of emissions along the maritime intermodal container chain: Operational models and policies', Maritime Policy & Management, 38(4), pp. 451–469.

Legato, P. and Mazza, R.M. (2001) 'Berth planning and resources optimization at a container terminal using discrete event simulation', European Journal of Operational Research, 133(3), pp. 537–547. doi: 10.1016/S0377-2217(00)00200-9.

Li, B., Elmi, Z., Manske, A., Jacobs, E., Lau, Y.-y., Chen, Q. and Dulebenets, M.A. (2023) 'Berth allocation and scheduling at marine container terminals: A state-of-the-art review of solution approaches and relevant scheduling attributes', Journal of Computational Design and Engineering. doi: 10.1093/jcde/qwad075.

Moorthy, R. and Teo, C.-P. (2006) 'Berth management in container terminal: The template design problem', OR Spectrum, 28(4), pp. 587–610. doi: 10.1007/s00291-006-0036-5.

Mubder, A.A.A.M. (2024) 'The implementation of berth allocation policies that enable Just-in-Time arrival in port calls', International Journal of Physical Distribution & Logistics Management, 54(6), pp. 610–630. doi: 10.1108/IJPDLM-11-2023-0442.

Piraeus Container Terminal Single Member S.A. (n.d.) Piraeus Container Terminal operations procedures at:https://www.pct.com.gr/sites/default/files/PCT%20Operational%20Procedures%20SOPs.pdf (Accessed: 18 August 2026).

Qin, J., Miao, L., Shi, F. and Chen, C. (2009) 'Model and algorithm for the berth allocation problem with time windows', in Proceedings of the 2009 Chinese Control and Decision Conference, pp. 4194–4199. doi: 10.1109/CCDC.2009.5194918.

Tighilt, F., Bouchellal, Y. and Daddi Addoun, N. (2025) 'Congestion of Algerian seaports issue: A transition from the first-come, first-served (FCFS) system to the berthing window system (BWS) at marine container terminals', Economic Researcher Review, 13(2), pp. 250–262.

Ursavas, E. and Zhu, S.X. (2016) 'Optimal policies for the berth allocation problem under stochastic nature', European Journal of Operational Research, 255(2), pp. 380–387. doi: 10.1016/j.ejor.2016.04.029.

Yıldırım, M.S., Aydın, M.M. and Gökkuş, Ü. (2020) 'Simulation optimization of the berth allocation in a container terminal with flexible vessel priority management', Maritime Policy & Management, 47(6), pp. 833–848. doi: 10.1080/03088839.2020.1730994.

Zhang, A., Zheng, Z. and Teo, C.-P. (2021) 'Schedule reliability in liner shipping timetable design: A convex programming approach', Transportation Research Part B: Methodological, 155, pp. 499–525.