ERASMUS UNIVERSITY ROTTERDAM

Erasmus School Of Economics

Master Thesis Econometrics and Management Science

Operations Research and Quantitative Logistics

Dynamic Routing Policies on the Quay Lane Area: a Simulation Analysis

Author name:
Kristan van Renssen
Student ID number:
502671

Supervisor:
M.A. Mankowski
Second assessor:
T.A.B. Dollevoet

Date final version: 15-09-2022

Abstract

With the continuous development of global trade, the demand for container terminals is ever-increasing. Automated container terminals have appeared to meet this growing demand and contribute to higher efficiency. However, to obtain and maintain a competitive advantage, increasing this efficiency is ever more important. Therefore, in this thesis, we focus on constructing a unidirectional routing policy over the quay lane area. This policy aims to be efficient for both autonomous straddle carriers as well as quay cranes. Therefore, we mathematically formulate the situation as a queueing system. Following this, we introduce two dynamic routing policies to support the decision-making in the quay lane area. Simulation models are used to test the proposed routing policies in different situations and these results are compared to a benchmark routing policy currently adopted. The simulation results show that the dynamic routing policies are able to outperform the benchmark routing policy in the considered situations.

Keywords: Automated Container Terminal, Queueing System, Dynamic Routing, Travel Time Estimation, Simulation

Contents

List of Abbreviations

The following abbreviations are used in this thesis.

ACT - Automated Container Terminal

AGV - Automated Guided Vehicle

ALV - Automated Lifting Vehicle

HRP - Hybrid Routing Policy

JEFQP- Join-the-Estimated-Fastest-Queue Policy

JLOQP- Join-the-Least-Occupied-Queue Policy

PDF - Probability Density Function

QC - Quay Crane

QCSP - Quay Crane Scheduling Problem

RV - Random Variable

SC - Straddle Carrier

TP - Transfer Point

YC - Yard Crane

1 Introduction

Port container terminals provide a service of container handling, more precisely, unloading containers from vessels onto inland vehicles (typically trucks and trains), and vice versa. The main purpose of such operations is to distribute the containers to the end-users (Pjevčević et al., 2011). A terminal can be divided into three components: a landside, a quayside, and a storage area in between. Figure 1 gives a schematic side view of a container terminal. At the landside, containers are unloaded off and/or loaded on inland vehicles. Similarly, at the quayside, containers are unloaded off and/or loaded on the vessel. The storage area in between stores the containers in stacks and facilitates the decoupling of the two external interfaces of the terminal.

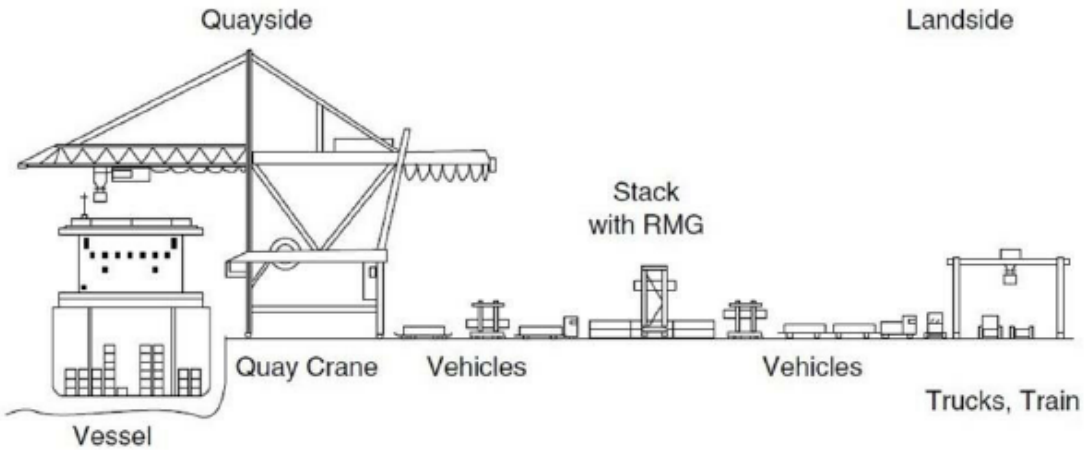


Figure 1: A schematic side view of a container terminal (Steenken et al., 2004).

Due to the continuous development of global trade the demand for container terminals, including unloading and loading operations, and the storage area has been pushed up. Automated Container Terminals (ACTs) have appeared to meet this ever-increasing demand and contribute to the higher efficiency (Yang et al., 2018). Since the establishment of the first ACT in Rotterdam in 1993, the technology for these terminals has been rapidly developing (Yu et al., 2022). ACTs are advanced container terminals equipped with various automated facilities. A large part of equipment found on an ACT is mainly used for container handlings, such as Quay Cranes (QCs) and Yard Cranes (YCs). Automated Guided Vehicles (AGVs) and Automated Lifting Vehicles (ALVs) are the most common unmanned vehicles used for the horizontal transport system of an ACT. Autonomous Straddle Carriers (SCs), belonging to the class of ALVs, can independently lift and set down containers. AGVs, however, must be directly assisted by other pieces of equipment (e.g. QCs or YCs) to perform their tasks.

Within an ACT, many operations can be distinguished in which multiple types of these machines need to work efficiently together. Therefore, operations efficiency is a highly investigated topic in general literature regarding ACTs. Namely, efficiency can provide a competitive advantage over other terminals, as well as increased revenue (Dragović et al., 2017). Hence, in this thesis, we attempt to improve a specific quayside operation at an ACT. More specifically, underneath the QCs, several quay lanes are positioned used for the transportation of the containers. Over these quay lanes, a unidirectional routing policy is searched for in the case of the deployment of autonomous SCs. This routing policy is aimed to be as efficient as possible considering the attributes of both the autonomous SCs as well as the QCs. We explicitly focus on situations where multiple QCs are working closely together, i.e., the autonomous SCs need to go under multiple QCs to leave the quay lane area resulting in a high probability of congestion or deadlock situations among these vehicles. Deadlock situations refer to situations when all vehicles involved are blocked by each other, waiting for the other vehicles to move.

As discussed before, efficiency is a highly sought-after criterion within an ACT. As the goal of an ACT is to minimize the time of berth of each vessel, various operations can be investigated to achieve this goal. The problem investigated in this thesis is such an operation that could positively affect the throughput of an ACT, as congestion and other conflict situations among the vehicles can highly affect the performance. Additionally, a relatively small increase in efficiency could be of significant value in gaining a competitive advantage.

Henceforth, we introduce a queueing system to mathematically formulate the previously denoted problem. Following, two dynamic routing policies to support the decision-making necessary in the quay lane area are formulated. More specifically, the decision for the specific quay lane is based on certain dynamic elements. The first dynamic routing policy merely considers the number of vehicles and, thus, selects the least occupied quay lane. Moreover, the second policy extends on this topic and distinguishes between different vehicles to make a travel time estimation over the relevant quay lane area.

We test our proposed dynamic routing policies on a simulation model of a specific ACT. Different scenarios in terms of QC positions are considered to evaluate the performance of the dynamic routing policies in different circumstances. The performances are measured in terms of relevant performance indicators, which are being compared to a benchmark routing policy currently adopted by TBA. The dynamic routing policies are shown to outperform the benchmark routing policy in the evaluated scenarios for these performance indicators. More specifically, significant improvements are found for the productivity of the QCs, whilst no additional waiting times for the autonomous SCs or the QCs are observed.

The contributions of this thesis are threefold: Firstly, new routing policies are shown for a specific ACT operation. These policies are shown to be robust in different situations which is valuable due to the uncertainty and changing circumstances on an ACT. Secondly, the general topic of dynamic routing is shown to be suitable for a new application. Namely, dynamic routing is a widely used topic, however, it has not yet been applied in a similar situation. Thirdly, in terms of practicality, useful insights are provided for TBA such that an executive decision can be based on known results.

The remainder of this thesis is structured as follows; An analysis of literature related to this problem is given in Section 2. Section 3 contains a more detailed description of the problem itself. In Section 4, the queueing system used to mathematically model the problem is extensively discussed. Next, in Section 5 the current routing policy is discussed in more detail, as well as the two dynamic routing policies are introduced. Further, in Section 6 the experimental design of the thesis is given, after which the obtained results in terms of various performance indicators for the various policies are discussed and compared. Finally, in Section 7 we formulate a conclusion based on the results, as well as suggestions for further research.

2 Related Literature

In recent years, numerous researches have been devoted to the efficiency of the equipment on ACTs. Improving various ACT operations and attaining handling processes excellence is of significant value to gain a competitive advantage (Dragović et al., 2017). The research especially focuses on the efficiency of the QCs. For example, Kaveshgar et al. (2012) construct a mixed-integer problem formulation for the Quay Crane Scheduling Problem (QCSP). The aim of the QCSP is to determine the task sequence for each QC in such a way that the vessel turn time is minimized. A genetic algorithm is proposed to solve the QCSP, as it is shown that this problem is $\mathcal{NP}$-hard (Lee et al., 2008). Other heuristic solution methods proposed for the QCSP involve tabu search algorithms, see e.g. Lee et al. (2011) and Sammarra et al. (2007), lagrangian-relaxation-based heuristics, see e.g. Al-Dhaheri and Diabat (2017), or variable neighborhood search algorithms, see e.g. Expósito-Izquierdo et al. (2011). An extensive survey on the QCSP is provided by Bierwirth and Meisel (2015).

However, the QCSP merely considers the scheduling of the QCs (Yang et al., 2018). Therefore, based on the schedules and routes of the vehicles, the optimal task sequence of the QCs could not be performed as expected. Namely, due to long waiting times for the vehicles, congestion, deadlock situations, and other conflicts could arise, indirectly affecting the performance of all ACT operations.

Therefore, vehicle scheduling and routing is an important process to examine in order to increase ACT efficiency, preferably in combination with QC scheduling. Almost all literature devoted to routing and scheduling of vehicles on container terminals considers the routing of AGVs. Early, static approaches were most commonly examined. Bartlett et al. (2014) state that with static routing approaches the route that a vehicle travels is selected based on static factors, such as distance, maximum travel velocity, and the location of the destination. Examples of such approaches are given by Kim et al. (2007), in which the AGV traveling area within the yard is divided into a large number of grid-blocks. These grid-blocks are reserved in advance by running AGVs to provide a static method of vehicle control. The final routes are determined in a distance-minimizing manner. The advantages of static routing methods are an easy implementation and fast route computation times (Gawrilow et al., 2008). However, common drawbacks of such approaches are that all constructed routes will traverse the same locations regardless of the traffic situation and an increased probability of congestion. Therefore, such methods are not suitable to adapt to often changing circumstances on an ACT.

As briefly discussed in Section 1, ALVs and AGVs differ in terms of operating characteristics. However, almost all literature devoted to the routing and scheduling of ALVs considers a similar static approach. Cai et al. (2012) formulate the Autonomous Straddle Carrier Scheduling Problem, which aims to find a feasible and efficient schedule for the autonomous SCs to finish a static list of container jobs. That is, each autonomous SC should pick up a container and transfer it to a location to set it down. Both the pickup and setdown operations should be executed within a given time window. Cai et al. (2014) extend the problem by incorporating uncertainty in this job list. However, neither approach considers the routes of the autonomous SCs, such that constant travel times are assumed between locations. As a result, congestion and other conflicts can not affect the performance of the vehicles.

Therefore, a more recent method is the dynamic routing approach which can be applied to avoid the drawbacks of a static approach. Bartlett et al. (2014) define dynamic routing as an approach that allows a vehicle to alter its path in response to congestion or other factors by continuously learning about the state of the system and adapting based on the new information. Several dynamic routing methods predicate routing decisions upon estimations of travel times on edges in a graph. Fontanelli et al. (2010) propose a travel time estimation model based on the last traversal of an edge or a linear interpolation between previous and current observations. A more extensive model is proposed by Zhen (2016), incorporating the estimated number of interruptions on an edge, normal cruise speeds, and acceleration/deceleration rates of the vehicles.

Alternatively, the interdependence between traffic flow, vehicle speed, and vehicle density is extensively examined (Greenshields et al., 1935). As a result, queueing models are increasingly more used to model traffic flows and, thus, vehicle speed. Vandaele et al. (2000) develop such queueing models able to predict traffic flow. Van Woensel et al. (2008) extend on this model and use it to estimate travel times. They test their estimations in a dynamic implementation of the vehicle routing problem. Both Van Woensel et al. (2008) and Vandaele et al. (2000) conclude that the queueing models are able to provide a more analytical approach to estimating traffic flow and are useful for sensitivity analysis, forecasts, etc.

Moreover, a prodigious rise in machine learning-based techniques has caused an influx of new travel time estimation methods. The general goal of machine learning is to recognize patterns in large data structures (Carleo et al., 2019). Zhang and Haghani (2015) construct a gradient boosting method to improve travel time predictions. They find a method that performs well in dealing with big data structures and provides considerably better predictions compared with classical statistical approaches. Duan et al. (2016) provide a neural network method able to predict travel times with a relatively small prediction error. However, a common drawback of machine learning methods is parameter optimization as the performance of these methods is often largely influenced by their parameters. The choices of these values are so far guided by trial-and-error and experience of researchers (Carleo et al., 2019).

Subsequently, these varying travel time estimation models, which are continuously updated with current information about the container terminal, can be used to determine the travel time minimizing routes at that point in time for the vehicles. Bartlett et al. (2014) propose such a dynamic method that stores a lookup table at each diverging node containing the next node in the route given a certain destination. These tables are updated periodically over time at a fixed interval. Namely, a one-to-all shortest path calculation using Dijkstra's algorithm is used to update the table. Musolino et al. (2013) resort to using a general optimization model to find the routes that minimize the travel time. Both methods can produce a decrease in travel times and the system is better able to recover from congestion. However, both methods deal with a certain trade-off in optimality and efficiency.

Other dynamic routing approaches incorporate Q-learning methods to base their decisions. Even-Dar et al. (2003) state that Q-learning originates from the field of reinforcement learning. Within reinforcement learning methods, an agent wanders in an unknown environment and tries to maximize its long-term return by performing actions and receiving rewards (Kaelbling et al., 1996). The agent should make its decisions such that it maximizes its long-term reward and can learn by systematic trial-and-error. Jeon et al. (2011) implement a Q-learning-based

dynamic routing method for AGVs in container terminals. They construct the estimated waiting times through the Q-learning technique and by constructing the shortest-time routing matrix. A substantial reduction in travel time is observed for the vehicles compared to a static approach. However, drawbacks of the method are the necessary computational power in combination with the time taken for learning, preventing often a suitable application in practice. Additionally, deadlock situations are difficult to avoid. Other Q-learning methods are given by e.g. Chujo et al. (2020) and Zhou et al. (2021).

Finally, more literature focuses on the combined scheduling of various equipment on an ACT. Meersmans and Wagelmans (2001) propose a first attempt at integrated scheduling of AGVs, QCs, and YCs. They construct a Branch-and-Bound algorithm to minimize the operation time of the QCs. Lau and Zhao (2008) implement a multi-level genetic algorithm to solve a mixed-integer programming model. Chen et al. (2007) propose a tabu search algorithm to solve this integrated planning problem. However, due to the high complexity of such problems, assumptions are needed to solve them. Travel times are often assumed to be deterministic and congestion is ignored within the ACT. Additionally, travel routes for AGVs are limited to a small set of possible routes. As discussed before, these static routing policies can negatively affect the actual throughput of the ACT.

In summary, a large part of existing literature deals with the efficiency of an ACT over various equipment. The routing and scheduling of automated vehicles constitute a large role in the performance of the ACT. However, most of the research focuses on routing within the yard and to the QC in a static or dynamic manner. Relatively little literature incorporates the routing options over the quay lanes in detail, as well as the routing of ALVs including its special characteristics. Henceforth, in this thesis, this routing decision for this specific type of vehicle is investigated further.

3 Problem Description

In this thesis, we focus on constructing a unidirectional routing policy efficient for both the QCs as well as the autonomous SCs when driving over the quay lanes within an ACT. More specifically, we focus on the case in which multiple QCs work closely together, such that the vehicles need to pass under multiple other QCs to leave the quay lane area. The quay lanes are a specific component of the quayside transportation area of an ACT. A schematic overview of the layout of the full area is given in Figure 2.

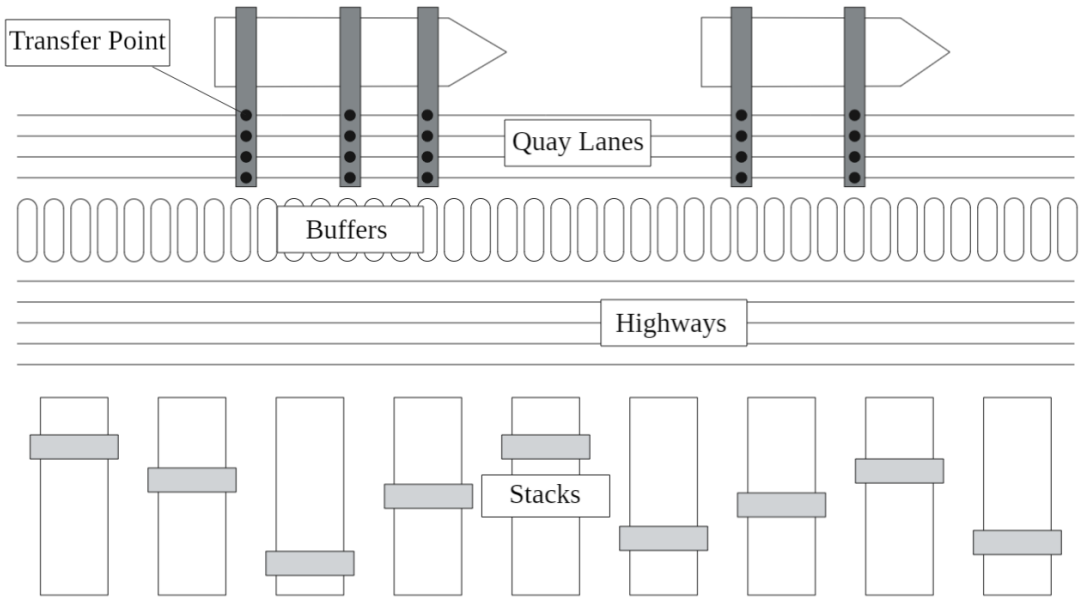


Figure 2: A schematic overview of the transportation area at the quayside of an Automated Container Terminal (ACT).

The layout can be described as follows: In front of the stacks lie several highways. These highways are directed to either the right or the left. Above the highways, a row of perpendicular buffer locations is positioned. These buffers serve as waiting locations for the vehicles. Namely, vehicles can have various reasons to wait, e.g., the vehicle has not been assigned a new task, or the QC can not be reached immediately. Above the buffers, various quay lanes are located. Usually, these quay lanes are chosen to be unidirectional. Finally, underneath the QCs, Transfer Points (TPs) are located depicting the handover points between vehicle and QC. It is possible to position more than one TP underneath each QC.

Underneath the QC, the TPs can be located in two ways. The first option is in between the supporting legs of the gantry. In this case, the QC works in gauge. Secondly, the TPs can be located further inland behind the QC: the QC works in backreach. Minimizing the distance between the vessel and the TP minimizes the unloading and loading times of the containers. Therefore, working in gauge, preferably on the quay lanes closest to the waterside, is generally considered the most efficient for the QC. Additionally, working in gauge takes less space on an ACT, which is preferable as space is scarce and expensive. A downside of working in gauge is the restricted movement space of the vehicles. Namely, if the QCs work closely together the vehicles have to pass underneath several other QCs to leave the quay lanes. This can be inefficient and can cause congestion. When working in backreach, the vehicles can move more freely as the legs of the QCs do not form an obstruction. However, the unloading and loading times are elongated.

Currently, in the case provided by TBA, autonomous SCs are used for the transportation of containers between the QCs and the stacks. In the simulation models, they are routed over the quay lane area using a Hybrid Routing Policy (HRP) depicted in Figure 3. The buffer locations used to enter the quay lane area, referred to as the in-buffers, are depicted in dark gray. Similarly, the buffer locations to leave the area are specified as out-buffers and are given in a light gray color. The remaining buffer locations, shown in white, can be used as general waiting locations. The HRP applies a mostly static decision policy based on a fixed order. If the originally assigned TP is not available at the time it is needed, it dynamically selects a new available TP based on that same order. This order does not take into account the efficiency of the QC or the autonomous SC. The vehicles drive in a unidirectional manner over the quay lane area.

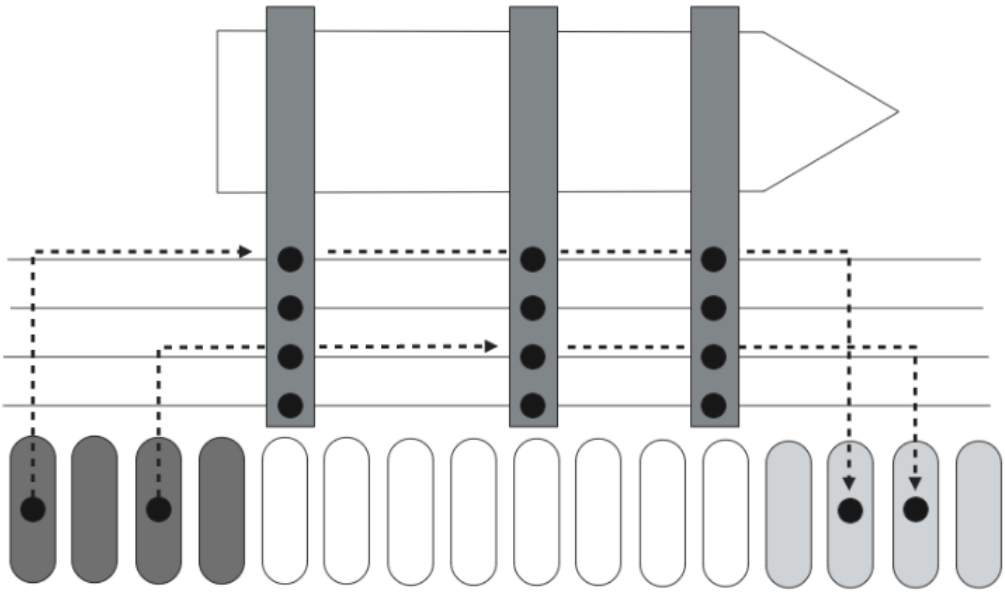


Figure 3: The Hybrid Routing Policy (HRP) for the autonomous Straddle Carriers (SCs) over the quay lanes in the case of a perpendicular quay lane implementation.

However, as discussed before, space on an ACT is often scarce and expensive. Additionally, more container terminals are being replaced with automated variations. For man-operated container terminals, often a different quay lane area implementation is applied in terms of the buffer locations. To prevent man-operated vehicles from entering areas underneath other QCs, these buffer locations are positioned horizontally on the quay lane area as depicted in Figure 4. More specifically, for this implementation, each QC is assigned a specific group of in-buffers and out-buffers positioned both on the same side of the QC. This means that quay lanes are used bidirectionally. Additionally, the outer QCs are chosen to work in gauge, while the inner QCs work in backreach. As there is no space for a full row of perpendicular buffer locations during the

transition, this quay lane implementation is kept for the ACT. However, due to this decision no more than four QCs can be positioned closely together for this quay lane implementation. The decision policy for the HRP can still be applied in this situation, as only the list of possible TPs for each QC is being restricted. However, as the quay lanes are used bidirectionally, the proposed routing policies in this thesis are not suitable. Thus, this quay lane area implementation is used only to compare the performance among the different policies.

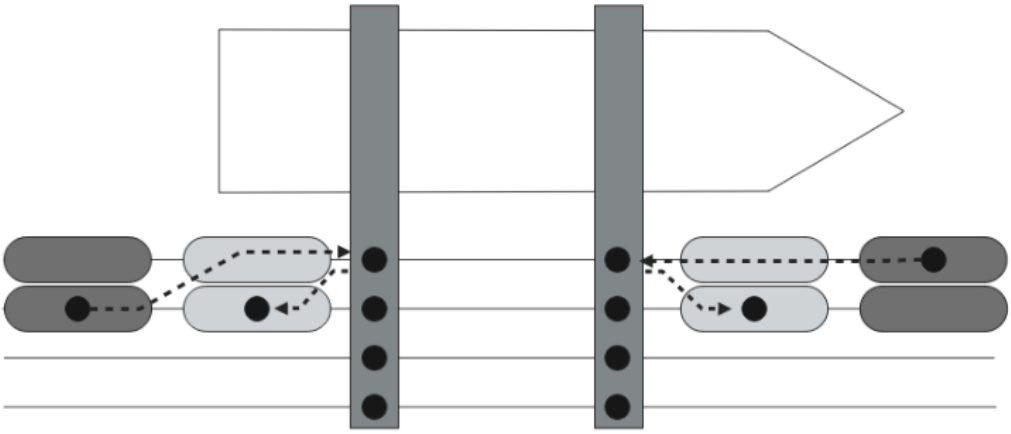


Figure 4: The Hybrid Routing Policy (HRP) for the autonomous Straddle Carriers (SCs) over the quay lanes in the case of a horizontal quay lane implementation.

The objective of an ACT is to maximize its performance. This performance is mostly evaluated as the throughput of containers within a given period. To maximize this throughput, the time of berth of the vessel needs to be minimized. That is, an ACT needs to work as efficiently as possible to unload and/or load a vessel as quickly as possible. Additionally, low times of berth attract more customers, resulting in more profit. To be as efficient as possible, the QCs should maximize their moves per hour. A move is defined as the unloading or loading of a single container. Many components within an ACT contribute to a higher performance of the QCs. Working the QCs mostly in gauge is such a component.

However, as discussed before, working in gauge often causes more congestion on and around the quay lanes. This congestion is caused as most decisions on an ACT are made based on a static policy, i.e., the routes of the vehicles are not adapted to the changing circumstances within a container terminal. As a result, more congestion occurs in specific areas, while in other areas hardly any traffic exists. All events happening on an ACT involve uncertainty to some extent, e.g., unloading and loading times can vary due to unforeseen reasons, or vehicles can break down when driving over the terminal. Dealing with these circumstances dynamically can provide a more efficient operation over the container terminal.

Therefore, the objective of this thesis is to provide an efficient unidirectional routing policy over the quay lane area. The policy needs to allow the QCs to work mostly in gauge, preferably on the quay lanes closest to the waterside, as well as prevent most of the congestion among the autonomous SCs. Using a dynamic method, this routing policy needs to react to the ever-changing conditions on the ACT. Additionally, as the policy is chosen to be unidirectional it is specifically aimed for ACTs with the perpendicular buffer layout. However, it can be compared to the performances of other quay lane area implementations.

4 The Queueing Model

In this section, we present the mathematical formulation to model the problem discussed in Section 3. More specifically, a queueing system is constructed, presented in Section 4.1. Moreover, a detailed description of the job types and the service rates is given in Section 4.2. Finally, the state space of the queueing system is described in detail in Section 4.3.

4.1 The Queueing System

We consider a multi-server continuous-time queueing system to formulate the problem. These systems are shown to be suitable to deal with traffic flow and congestion problems (Vandaele et al., 2000). Generally, queues occur whenever instantaneous demand exceeds service capacity. Queueing theory involves a mathematical approach to these waiting lines, such that various characteristics of these systems can be analyzed, like average waiting time or average length of the queue. Therefore, this queueing system provides an analytical approach to examining the congestion among the vehicles within the quay lane area.

We assume a situation in which K $(1 \leq K < \infty)$ QCs are positioned closely together. These QCs combined are referred to as a QC group and, thus, the following queueing system can be constructed for each QC group in the terminal. Additionally, the quay lane area is assumed to consist of L $(1 \leq L < \infty)$ quay lanes. For convenience, let the sets $\mathcal{K} = \{1, \ldots, K\}$ and $\mathcal{L} = \{1, \ldots, L\}$ include the reference numbers to all QCs and quay lanes in use for the specific QC group, respectively. Generally, we consider QC 1 (K) to be passed first (last) by the vehicles. Similarly, quay lane 1 (L) is positioned closest to (furthest from) the waterside.

Subsequently, we depict the situation as an open queueing network with L parallel queues representing the quay lanes. Each queue has K servers describing each TP underneath the QC. Accordingly, in the queueing system, we distinguish in total $K \times L$ servers. Each server has a separate waiting queue in which vehicles can wait if the server is occupied. The size of each of these separate waiting queues is dependent on the space available between two adjacent QCs.

More specifically, if two QCs are working closely together on the same vessel, there could be only room for a single vehicle or even no vehicles to wait between the TPs of the two different QCs. However, if these cranes are working further apart on the vessel, the space in between could become larger resulting in the possibility of more vehicles waiting in between the TPs. Consequently, the sizes of the separate waiting queues need to be evaluated each time the positions of the QCs are adjusted. A schematic representation of the full queueing system is depicted in detail in Figure 5.

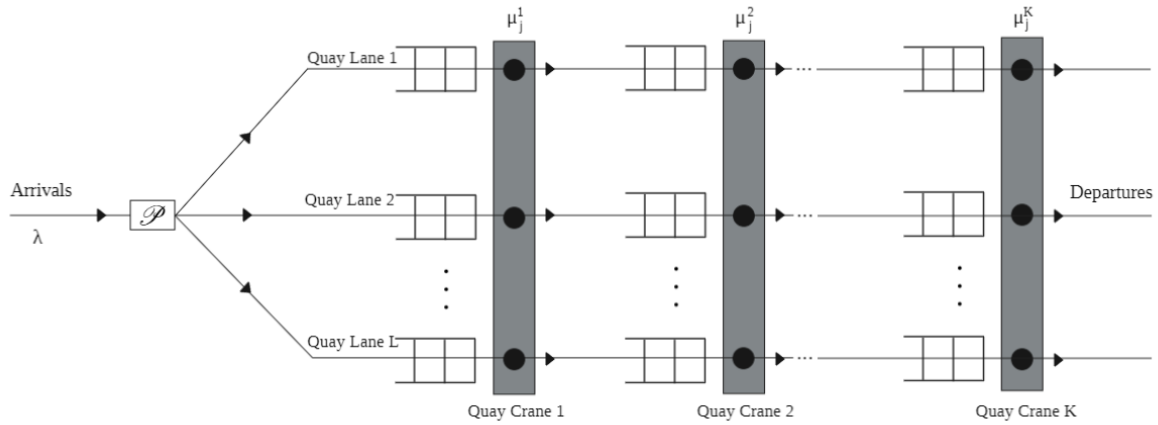


Figure 5: An open queueing network with L parallel queues each with K servers.

4.2 The Job Types and Service Rates

Jobs arrive in the system and represent vehicles arriving for either an unloading or loading operation assigned for one of the QCs in the QC group. As can be seen in Figure 5, we assume that these jobs arrive according to a Poisson arrival process with a mean arrival rate λ, which value can be reevaluated each time a new QC group situation occurs. Following, these arrivals need to be routed over one of the L quay lanes. This is done according to a specific routing policy $\mathscr{P}$. The various routing policies will be discussed in more detail in Section 5. The moment of arrival in the system can be seen as the moment the specific quay lane has to be decided for the job. As it turns out, this moment is slightly different when considering unloading or loading operations. With an unloading operation, the decision is ultimately made when the QC picks the container from the vessel and transports it to the TP. For loading operations, this quay lane decision needs to be made by the vehicle when entering the quay lane area. Therefore, the decision has to be ultimately made in the in-buffer of the quay lane area.

However, in comparison to regular queueing systems, these arriving jobs only need to be served at a single server over the entire quay lane as a container is designated for a single QC. The job can only experience delays at other servers due to blocking. To differentiate between different job destinations, we consider multiple job types, such that the system can be considered a multi-class queueing system. More specifically, by distinguishing between unloading and loading operations for each QC, we can identify $2K$ job types. However, a single QC works only on either unloading or loading operations at once, thus, only K job types are actually in use at one point in time. Let j_k^{unload} denote an unloading job designated for QC $k \in \mathcal{K}$. Similarly, j_k^{load} represents a loading job designated for QC $k \in \mathcal{K}$. For convenience, let $\mathcal{J}$ denote the set of all job types in use in the current system, such that $|\mathcal{J}|$ equals K. Every arriving job is of job type $j \in \mathcal{J}$ with probability p_j, $\sum_{j \in \mathcal{J}} p_j = 1$.

To distinguish between the service times of the different job types, we define the service rates of various servers dependent on this job type. Namely, a server representing a TP at QC $k \in \mathcal{K}$ is assumed to be characterized by its mean service rate μ_j^k for each $j \in \mathcal{J}$. By means of this distinction, the mean service rate for each server differs between various job types. No distinction is made between various quay lanes, as it is assumed that operations on different quay lanes are executed in similar times. The service times and the arrival process are independent of each other.

4.3 The State Space of the Queueing System

The queueing system, as previously described, can be considered a dynamic system as it evolves dynamically over time. At each point in time, the system can be described by its state and changes based on the state transitions. These transitions are dependent on the arrival rates and service rates previously discussed. The set of all possible states of the system is referred to as the state space and is assumed to be finite or countably infinite.

The state of this queueing system can be considered as follows: Denote by $Q^{k,l}(t)$ the total queue length (including the job at service) at server k on quay lane l at moment t. Then, for each $k \in \mathcal{K}$ and $l \in \mathcal{L}$, $t \geq 0$, the vectors

$$S^{k,l}(t) = (s_1^{k,l}(t), \ldots, s_{Q^{k,l}(t)}^{k,l}(t), \varsigma^{k,l}(t)), \quad k \in \mathcal{K},\ l \in \mathcal{L},$$

describe the states of each server at time t, where $s_i^{k,l}(t)$ denotes the job type of the ith job at server k on quay lane l. For notation, introduce $k_i^{k,l}(t)$ to be the reference number of the designated QC of the ith job at server k on quay lane l. Moreover, $\varsigma^{k,l}(t)$ depicts the residual service time of the first job in this queue. If $Q^{k,l}(t)$ is equal to zero, we put $\varsigma^{k,l}(t)$ also equal to

zero. Let $K^{k,l}$ be the set of finitely terminated sequences dependent on the size of the queue of server k on quay lane l. Then $S^{k,l}(t)$ takes values in the state space $K^{k,l} \times \mathbb{R}^+$.

Additionally, the vector $S(t) = (S^{1,1}(t), \ldots, S^{K,L}(t))$ represents the state of the full system at time t. In particular, $S(0)$ denotes the initial state for which we assume $Q^{k,l}(t)$ is set to zero for all $k \in \mathcal{K}$ and $l \in \mathcal{L}$. That means all queues are assumed to be empty at the beginning. $S(t)$ takes values in the state space $\mathcal{X} = \prod_{k \in \mathcal{K}} \prod_{l \in \mathcal{L}} (K^{k,l} \times \mathbb{R}^+)$.

5 The Routing Policies

This section presents the routing policies proposed to efficiently direct the various jobs over the quay lane area. In Section 5.1, the quay lane decision strategy for the HRP, as introduced in Section 3, is discussed in more detail. Namely, the performance of this HRP is used throughout this thesis as a benchmark. Thereafter, in Section 5.2, two dynamic routing policies are introduced.

5.1 The Hybrid Routing Policy

In the case provided by the TBA based on a specific simulation model, the routing decisions are based on the HRP. It was shortly introduced in Section 3, however, the particular decision features of the policy have not been discussed. The HRP can be applied in both quay lane area implementations with perpendicular buffer locations or horizontal buffer locations.

For the first step of the HRP, the allowed TPs are determined for each QC. This is where the situation with the horizontal buffer locations deviates from that of the perpendicular buffer locations. Namely, in the case of the perpendicular buffer locations, all QCs are allowed to work on all TPs, both in gauge as well as in backreach. However, as said before, in the case of horizontal buffer locations the outer QCs are chosen to work in gauge, whilst the inner QCs are working in backreach. Therefore, each QC is assigned these respective TPs on which it is allowed to perform its handling operations. In no situation, the QC can deviate from the assigned TPs.

Thereafter, the HRP is similar for both quay lane area implementations. Namely, in a static manner, a quay lane is assigned to each arriving job. More specifically, each arriving job is assigned a TP among the allowed TPs based on a fixed order. Assume a QC is allowed to work on all TPs on all quay lanes, such that $\mathcal{L} = \{1, \ldots, L\}$ are the allowed quay lanes. Then, the fixed order is generally based on an ascending or descending arrangement. That is, if an ascending arrangement is chosen, the TP on quay lane 1 is assigned to the first arriving job, followed by the TP on quay lane 2 for the second job, etc. If the highest allowed TP on quay lane L is assigned to an arriving job, the next arriving job gets again assigned the TP on quay lane 1. The opposite holds if a descending arrangement is selected.

However, the HRP also incorporates a slight dynamic element in the decision policy. Namely, at the ultimate decision moment, it is checked if the originally assigned TP is available. If not, due to a not yet handled container or a stationary vehicle, an alternative quay lane is selected. This alternative is decided in the same ascending or descending arrangement as discussed before. However, only the available TPs are considered at that moment in time. We denote the set of available TPs as $\mathcal{L}^{\mathrm{a}}$. Consequently, if the original TP is blocked, the alternative TP is the first TP in the set $\mathcal{L}^{\mathrm{a}}$ in the selected order. If an alternative TP needs to be selected, it becomes more likely for the following jobs to also encounter unavailable TPs. Therefore, the assigned TPs for all known jobs at that point in time are recalculated based on the fixed order.

The HRP as discussed here will be treated as a benchmark model. That is, the performance of the dynamic routing policies is compared to the performance of the HRP for both quay lane area implementations. Based on this comparison, a detailed understanding of the performance of the proposed routing policies can be created.

5.2 The Dynamic Routing Policies

Dynamic routing is shown to be suitable to deal with congestion problems in vehicle networks. As discussed before, Bartlett et al. (2014) state that dynamic routing allows a vehicle to alter its path in response to the state of the system. The efficacy of dynamic routing can be accounted to the direct relation between flow (q), density (m), and speed (v). Namely, Greenshields et al. (1935) provide seminal work on this topic and captured these relationship results in speed-flow-density diagrams. The general relationship between these diagrams is shown in Figure 6.

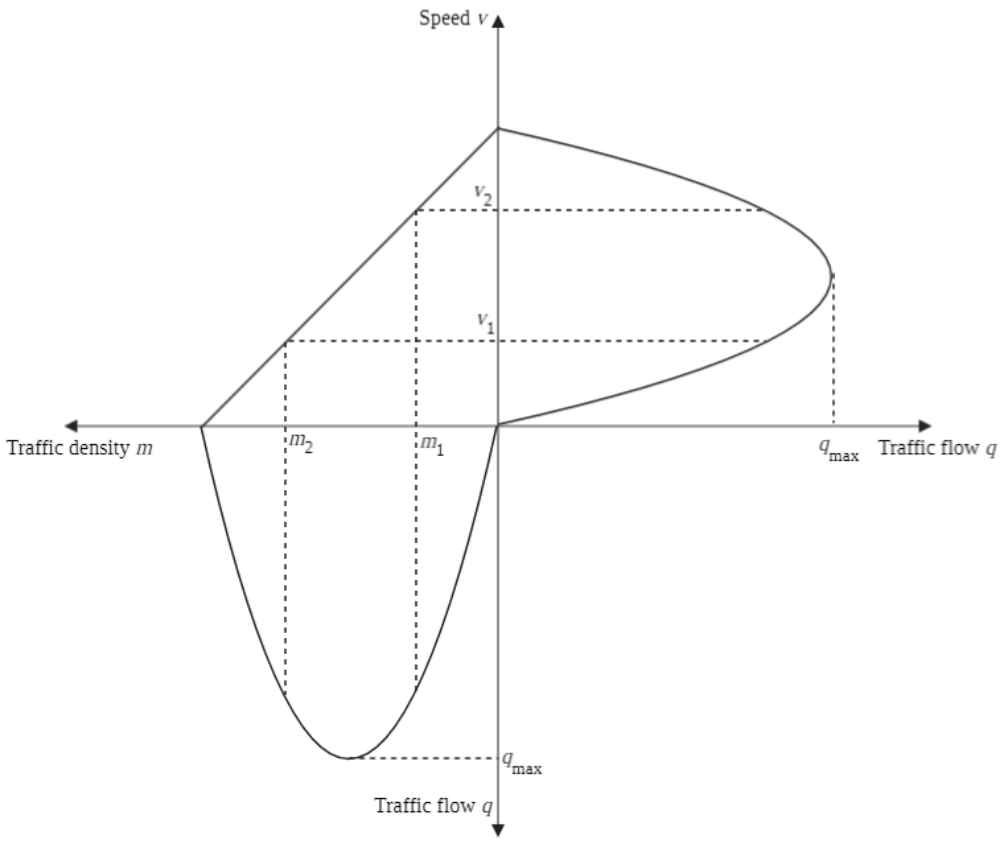


Figure 6: The relations between the speed-flow, the speed-density, and the flow-density diagrams.

Figure 6 shows that every traffic flow q does not correspond to a single traffic speed v. There are two speeds for every traffic flow: a lower half (v_1) where speed increases as flow increases and an upper half (v_2) where speed decreases as flow increases. Daganzo (1997) explains as the flow moves from 0 to $q_{\max}$, congestion increases but the flow rises because the decline in speed is over-compensated by the higher traffic density. However, if traffic grows past $q_{\max}$ the traffic flow falls again because the decline in speed more than offsets the additional number of vehicles. The flow-density and speed-density diagrams are similar representations and can be interpreted in the same way.

As a result, the dynamic routing policies proposed in this thesis are aimed to provide a mechanism of load-balancing over the quay lanes. By means of this, the vehicle density is regulated and, thus, the vehicle speed and flow. Using the dynamic element, the policy is more flexible to react to unforeseen changes. Section 5.2.1 provides a dynamic routing policy disregarding the different job types in the queueing system and merely focusing on the overall density of the quay lanes. Section 5.2.2 extends this policy and aims to make a travel time estimation based on the current state of each quay lane incorporating the different job types.

5.2.1 Join-the-Least-Occupied-Queue Policy

The first dynamic routing policy is referred to as the Join-the-Least-Occupied-Queue Policy (JLOQP). This policy aims to regulate traffic density in order to guarantee vehicle speed and flow over the quay lane area. Therefore, the JLOQP is formulated to send an arriving job to the quay lane with the least vehicles present. To do this, assume the arriving job arrives at moment $t \geq 0$ and the system is in state $S(t)$. Recall that with $Q^{k,l}(t)$ we denote the number of jobs present at server k on quay lane l at moment t. Then, the total number of jobs on the entire quay lane $l \in \mathcal{L}$ can be calculated as $\sum_{k \in \mathcal{K}} Q^{k,l}(t)$. This term can be compared between each quay lane and the best option can be chosen.

However, the proposed policy also needs to consider the efficiency of the QC. As discussed before, working in gauge is preferred for a QC, ideally working on the quay lanes closest to the waterside. However, it is unwanted that a QC waits due to a chosen TP being unavailable, especially during unloading. As such, we make a decision only considering the available TPs. Moreover, a weighting term can be added to the score of each quay lane to establish an additional preference for the quay lanes closest to the waterside. As a result, we add the $l \times \delta$ to the scores of each quay lane. This total value can be seen as additional dummy vehicles. Higher values of δ correspond to a stronger preference for the quay lanes closer to the waterside. The policy can formally be described as

Assume the job arrives at moment $t \geq 0$ and the system is in state $S(t)$.

1) The arriving job is sent to quay lane $i \in \mathcal{L}^{\mathrm{a}}$ if the sum of queue lengths of all servers on this quay lane $\sum_{k \in \mathcal{K}} Q^{k,i}(t) + (i \times \delta) = \min \left\{ \sum_{k \in \mathcal{K}} Q^{k,l}(t) + (l \times \delta) \mid l \in \mathcal{L}^{\mathrm{a}} \right\}$.

2) If this number is not unique, the routing policy selects from among the ties the quay lane closest to the waterside.

As can be seen, the policy only considers the length of the queues on each quay lane multiplied by the weighting term and disregards the specific types of jobs present in the system. Condition 2) is one of many possibilities to solve a tie between multiple quay lanes and it will be used for all dynamic routing policies. This condition is chosen as the quay lane closest to the waterside is preferred as this quay lane minimizes the movement of the QC.

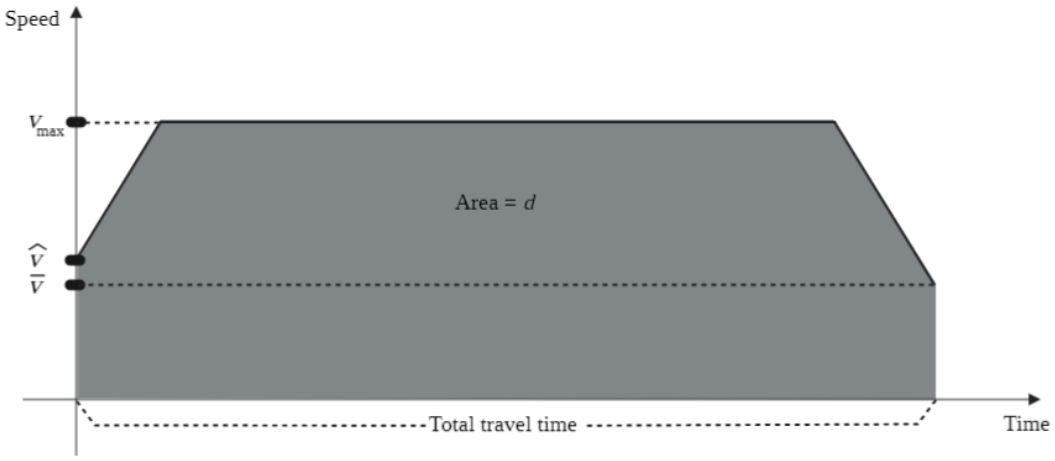
5.2.2 Join-the-Estimated-Fastest-Queue Policy

The second dynamic routing policy we formulate is the Join-the-Estimated-Fastest-Queue Policy (JEFQP). As the previous policy merely considers the total number of jobs and does not distinguish between the different types of jobs, we extend this second policy to incorporate such components. More specifically, for each quay lane, a travel time estimation is made based on the expected duration of interruptions due to other jobs being already present in the queueing system. Similar to the JLOQP, this estimation can be multiplied with a similar weighting term in order to partially bias the decision towards the quay lanes closest to the waterside.

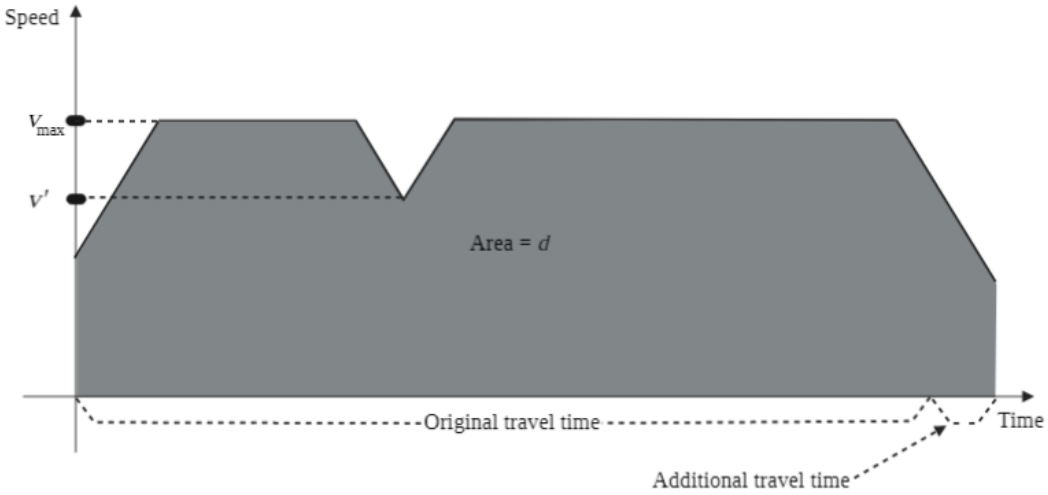
Firstly, we discuss the travel time estimation model. Initially, we focus on the situation in the case of no interruptions. The travel time estimation model is based on a model introduced by Zhen (2016). They propose an extensive model that incorporates maximum travel speed and acceleration rates. For notation, we use d to represent the distance of the travel region over the quay lane. Let $v_{\max}$ denote the maximum travel speed of the autonomous SCs. Additionally, $\hat{v}$ and $\bar{v}$ are the speeds of the autonomous SCs at the start and the end of the quay lane, respectively. Finally, we assume the acceleration rate and deceleration rate to be equal for the vehicle and to be constant over time. We denote this acceleration/deceleration rate of the vehicle by a. According to the chart given in Figure 7a, the total travel time of the autonomous SC without interruptions, given as t_0, can be calculated as

$$t_0 = \frac{d}{v_{\max}} + \frac{\hat{v}^2 + \bar{v}^2}{2av_{\max}} + \frac{v_{\max} - \hat{v} - \bar{v}}{a}. \tag{1}$$

A formal proof of the validity of Equation 1 is presented in Appendix A, as this is not given by Zhen (2016).



(a) No interruptions for the vehicle.



(b) One interruption for the vehicle.

Figure 7: Charts representing the speed of a vehicle over time in two different scenarios.

However, it is highly unlikely an autonomous SC can traverse the horizontal area over the quay lanes without any interruptions. Figure 7b depicts the speed over time in the case a vehicle incurs an interruption during its travel. Let us denote the decelerated speed as v'. Then the additional travel time t_{add} of the vehicle is given by

$$t_{\text{add}} = \frac{(v_{\max} - v')^2}{a v_{\max}}, \tag{2}$$

and can be added to the original travel time that would be achieved if no interruptions are encountered. Again, a formal proof of the validity of Equation 2 is given in Appendix A.

Generally, an autonomous SC can incur multiple reasons to be interrupted over the quay lane area. Firstly, the vehicle has to come to a full stop to perform its container operation, i.e., v' can be set to zero. However, in addition to the additional time t_{add} that is added because of the deceleration and acceleration of the vehicle, we have to account for the stop time t_{stop} that is used to put down or pick up the container. As a result, the minimal travel time $t_{\min}$ for any autonomous SC over the quay lane area is at least equal to

$$t_{\min} = \frac{d}{v_{\max}} + \frac{\hat{v}^2 + \bar{v}^2}{2a v_{\max}} + \frac{v_{\max} - \hat{v} - \bar{v}}{a} + \frac{(v_{\max} - 0)^2}{a v_{\max}} + t_{\text{stop}}. \tag{3}$$

Secondly, additional travel time can be incurred due to other vehicles being in the way. In general, we can distinguish three different kinds of progress among the jobs that can be present in the queueing system, namely:

1. A job can be present that still needs to do its container operation at an upcoming QC $k \in \mathcal{K}$. The newly arriving job could be blocked by this job in the case the sum of all remaining service times until QC k of the present job is more than the sum of service times until QC k of the new job;
2. A job can be present that is performing its container operation at the current QC k. This could result in a blockage if the remaining service time of this job at QC k is larger than the sum of service times of the new job until QC k;
3. A job could have performed its container operation at a previous QC k. The presence of this job can generally never result in a blockage of the new job. The new job could become stationary behind this job, however, the standstill can then be accounted to another job in front on the quay lane performing its own container operation.

Therefore, we estimate the expected additional travel time due to blockages dependent on the three different kinds of progress among the vehicles. To do this, we assume that the new vehicle corresponding to the new job has to come to a full stop behind a vehicle that has to do or is doing its job at QC k if the service times of the new job to QC k is at least equal to the service time of the present job. Additional stopping time is added if the remaining service time of the present job exceeds the sum of service times of the new job. As the service times are considered Random Variables (RVs), we can use probability theory to calculate the probability of such a blockage.

Namely, this probability can be calculated as follows: Let X_j^k be an exponential RV with the respective mean service rate μ_j^k, $j \in \mathcal{J},\ k \in \mathcal{K}$, written as $X_j^k \sim \text{Exp}(\mu_j^k)$. This RV denotes the service time of a job of job type j at server k. Then, we define the RV

$$\hat{X}_j^{m,n} = \sum_{i=m}^{n} X_j^i, \quad \forall\, m, n \in \mathcal{K},\ m \leq n,\ \forall\, j \in \mathcal{J},$$

to be the Hypoexponential RV with parameters $\mu_j^i,\ i = m, \ldots, n$, written as

$$\hat{X}_j^{m,n} \sim \text{Hypoexp}(\mu_j^m, \ldots, \mu_j^n).$$

This RV denotes the sum of service times from server $m \in \mathcal{K}$ until server $n \in \mathcal{K}$. Assuming distinct parameters for the Hypoexponential RV, we can denote the Probability Density Function (PDF) as

$$f_{\hat{X}_j^{m,n}}(x) = \sum_{i=m}^{n} \mu_j^i e^{-\mu_j^i x} \left(\prod_{h=m, h\neq i}^{n} \frac{\mu_j^h}{\mu_j^h - \mu_j^i} \right) = \sum_{i=m}^{n} \ell_i(0) \mu_j^i e^{-\mu_j^i x},$$

with $\ell_m(x), ..., \ell_n(x)$ the *Lagrange basis polynomials* associated with the points $\mu_j^m, \ldots, \mu_j^n$ (Nadarajah, 2008). Both the sum of service times of the present job and the newly arriving job can be denoted with such Hypoexponential RVs. Thus, let $\hat{X}_j^{m,n}$ and $\hat{Y}_j^{1,n}$ be the Hypoexponential

RVs representing the sum of service times of a present job and a newly arriving job, respectively. However, to evaluate the probability of a positive waiting time, the PDF of the difference between the two Hypoexponential RVs is needed. As such, we define the RV $Z^{m,n}_{j_1,j_2} = \hat{X}^{m,n}_{j_1} - \hat{Y}^{1,n}_{j_2}$ with the PDF formulated as

$$f_{Z^{m,n}_{j_1,j_2}}(z) = \sum_{\alpha=m}^{n}\sum_{\beta=1}^{n} \frac{\ell_\alpha(0)\ell_\beta(0)\mu^\alpha_{j_1}\mu^\beta_{j_2}}{\mu^\alpha_{j_1}+\mu^\beta_{j_2}} \left(e^{\mu^\beta_{j_2} z}\mathbb{1}_{\{z<0\}} + e^{-\mu^\alpha_{j_1} z}\mathbb{1}_{\{z\geq 0\}} \right), \tag{4}$$

with $\mathbb{1}_A$ an indicator function representing the value 1 if property A holds, 0 otherwise. A formal proof of the derivation of Equation 4 using the convolution formula for probability theory is presented in Appendix B.

Next, we can calculate the expected additional travel time due to each present job by multiplying each of the possible outcomes by the probability each outcome will occur and summing over these values. Then, let us denote $w^{k,l}_i$ as the expected additional travel time due to the job in the ith position in the waiting queue of server $k \in \mathcal{K}$ on quay lane $l \in \mathcal{L}$. For notation, let $W^l_{M,N}$ be the summed value of all additional travel times due to vehicles up until the Nth job in the queue for the Mth server on quay lane l, such that

$$W^l_{M,N}(t) = \sum_{k=1}^{M-1}\sum_{i=1}^{Q^{k,l}(t)} w^{k,l}_{Q^{k,l}(t)-i+1}(t) + \sum_{i=1}^{N} w^{M,l}_{Q^{k,l}(t)-i+1}(t).$$

As the incurred travel time of previously encountered jobs needs to be taken into account for the next job, we use a recurrent formula to determine the additional travel time due to each present job. Then, the recurrent formula denoting the additional travel time for a new job of job type j_{new} is equal to

$$w^{k,l}_i(t) = \begin{cases} \sum_{j=1}^{\infty} f_{Z^{k,k^{k,l}_i(t)}_{x^{k,l}_i(t),j_{\text{new}}}}(W^l_{k,i-1}(t)-j) \times \left(\frac{(v_{\max}-0)^2}{a v_{\max}} + j\right) \\ \qquad , \quad \text{if } k \geq 1, i > 1, k^{k,l}_i(t) \geq k \\ \sum_{j=1}^{\infty} f_{Z^{k+1,k^{k,l}_i(t)}_{x^{k,l}_i(t),j_{\text{new}}}}(W^l_{k,i-1}(t)-j-\varsigma^{k,l}(t)) \times \left(\frac{(v_{\max}-0)^2}{a v_{\max}} + j\right) \\ \qquad , \quad \text{if } k \geq 1, i = 1, k^{k,l}_i(t) \geq k \\ 0 \quad , \quad \text{otherwise} \end{cases} , \; k \in \mathcal{K},\, l \in \mathcal{L},\, 1 \leq i \leq Q^{k,l}(t).$$

As a result, we find the expected total travel time over quay lane l to be equal to $t_{\min} + W^l_{K,Q^{K,l}}(t)$. Based on this value, we can select the appropriate quay lane to send the new job onto. As said before, we can use the weighting term to bias the decision towards the quay lanes closer to the waterside. As the score of each quay lane is denoted in time, the added weighting term can be considered as dummy time. The policy can be formally described as

Assume the job arrives at moment $t \geq 0$ and the system is in state $S(t)$.

1) The arriving job is sent to quay lane $i \in \mathcal{L}^{\mathrm{a}}$ if the total travel time on this quay lane $t_{\min} + W^{i}_{K,Q^{K,i}}(t) + (i \times \delta) = \min \left\{ t_{\min} + W^{l}_{K,Q^{K,l}}(t) + (l \times \delta) \mid l \in \mathcal{L}^{\mathrm{a}} \right\}$.

2) If this number is not unique, the routing policy selects from among the ties the quay lane closest to the waterside.

Similarly as for the JLOQP, Condition 2) is used to solve a tie in the case multiple quay lanes have the same value.

6 Experimental Results

In this section, simulated terminal performance metrics are analyzed and compared for the various routing policies. Section 6.1 briefly discusses the simulation model as well as the validation of the simulation models. Section 6.2 gives the layout design settings and parameter values studied in this thesis. In Section 6.3 various performance indicators are presented and compared for the various routing policies. Finally, Section 6.4 conducts a sensitivity analysis on certain model parameters.

6.1 Simulation

The simulation model used to analyze various operations on container terminals is referred to as TIMESquare (Terminal In-depth Models Evaluation Studies). This fully developed simulation model is created by using the commercial software package eM-Plant. eM-Plant applies discrete event simulation. These events are generated using methods stored in the *Eventlist*. These methods in the *Eventlist* are called chronologically. Once a method is called, it executes the code it contains. Generally, this code consists of instructions for equipment or events that take place. As a result, TIMESquare can model the behavior of QCs, YCs, autonomous SCs, trucks, and trains on an ACT in detail.

A single simulation run is chosen to correspond to eight hours of work and processing time on the ACT. To achieve steady simulation results for each experiment, the first hour is treated as a warm-up period, and corresponding results are not included for performance evaluations. After this period, model outputs become steady. Consequently, one simulation run produces seven hours of useful data. Generally, in TIMESquare it is assumed that at least 100 hours of simulation data is necessary to draw conclusions that are not based on stochastic deviations. Therefore, all the results in this thesis are obtained using 175 hours of simulation data, equivalent to 25 simulation runs per experiment.

6.2 Terminal Layout Settings

In this thesis, a specific ACT is used to compare terminal performance for the different combinations of the concerned routing policies and terminal layouts. The used layout of the ACT is proceeding from a generic autonomous SC terminal with a perpendicular stack layout. Two rail-mounted gantry cranes are used as YCs per stack. We consider a total of six QCs positioned on the main quay. This number of QCs rules out the possibility of the stack capacity being the bottleneck in the productivity of the terminal. This is preferred as QC productivity is mostly considered. Each QC is assigned a total of six TPs of which four are positioned in gauge. The remaining two TPs are located in backreach. The terminal as implemented in TIMESquare is presented in Figure 8. As can be seen, for the simulation experiments, we evaluate a complete ACT. The operational implementation of additional ACT operations necessary for the full simulation are shortly discussed in Appendix C.

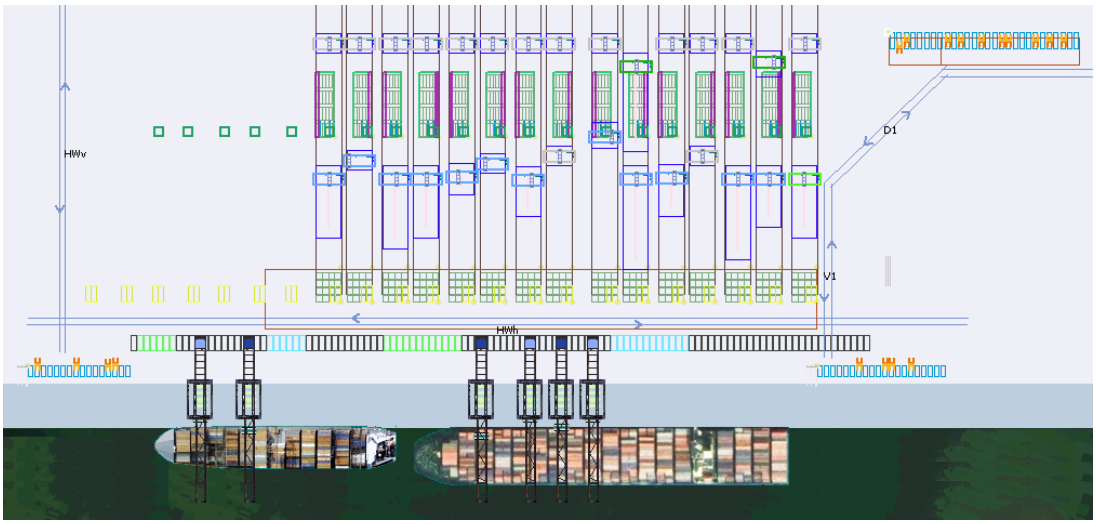


Figure 8: The layout of the Automated Container Terminal (ACT) as depicted in TIMESquare.

To evaluate the performance of the routing policies in different situations we consider three scenarios in which the positions of the six QCs are adjusted. For Scenario 1, the QCs are split up into two QC groups. The first group consists of two QCs and the second group consists of the remaining four QCs. This specific scenario is depicted in Figure 8. Similarly, Scenario 2 considers two QC groups. However, these groups both consist of three QCs. Finally, in Scenario 3 all QCs are positioned closely together forming a single QC group. As discussed before, for the horizontal quay lane implementation no more than four QCs can be positioned closely together. As such, this scenario is excluded for this quay lane implementation. An overview of coordinate positions of the six QCs in the various scenarios is given in Table 1. The coordinate scale used

in the model is expressed in meters. As can be seen, in a QC group no more than 50 meters is left between the different QCs. This is considered to be the distance when vehicles are not allowed to travel between the gantry legs of the cranes to leave or enter the quay lane area.

Table 1: An overview of the coordinate positions of the Quay Cranes (QCs) in the different scenarios.

	QC 1		QC 2		QC 3		QC 4		QC 5		QC 6	
Scenario type	X-coord.	Y-coord.	X-coord.	Y-coord.	X-coord.	Y-coord.	X-coord.	Y-coord.	X-coord.	Y-coord.	X-coord.	Y-coord.
Scenario 1	190.00	20.00	235.00	20.00	450.00	20.00	495.00	20.00	524.00	20.00	553.00	20.00
Scenario 2	210.00	20.00	255.00	20.00	298.00	20.00	495.00	20.00	524.00	20.00	553.00	20.00
Scenario 3	360.00	20.00	405.00	20.00	450.00	20.00	495.00	20.00	524.00	20.00	553.00	20.00

Additionally, important for the simulation experiments is the distribution of the QC move types. This distribution determines the proportion of the various QC move types throughout the simulation run. In the model, QCs are allowed to perform either a single or a twin move. In the case of a single move, the QC (un)loads a single container. A twin move (un)loads two containers short sides adjacent, such that they can be positioned on the same TP. As a result, roughly 63% of container operations correspond to single moves. Complementary, approximately 37% of the containers are twin-lifted by the QC. At the start of each simulation, a random schedule consisting of bays based on these proportions is generated. A bay corresponds to a specific number of containers for a specific QC move. The same schedule is reused for each experiment with a different scenario or routing policy. This ensures good comparability for the different simulation experiments.

Moreover, 24 autonomous SCs are used for the transportation of the containers between the stack and the quay lane area. Again, using this number of vehicles we ensure the vehicle productivity to not be the bottleneck in the productivity of the terminal. These vehicles are considered to be bidirectional and symmetrical, i.e., they can move forwards and backwards in an identical fashion. As discussed before, these vehicles can pick up and set down containers independently. For the travel time estimation model, the following values are assigned to the necessary parameters:

- $v_{\max}$:= 8.3 m/s,
- $\hat{v}$:= 3 m/s,
- $\bar{v}$:= 3 m/s,
- a := 0.6 m/s^2,
- t_{stop} := 13 sec.

The values for $\hat{v}$ and $\bar{v}$ are chosen based on the turning speeds of the autonomous SCs due to the turning necessary at the start and end of the horizontal travel region over the quay lane area. The other parameters d and t' are dependent on the scenario and the specific QC group.

Finally, we discuss the estimation of the service rates of the queueing model. These parameters are needed for calculations in the JEFQP. As no data is available for such information, we estimate the parameter values using travel time estimations similarly as presented in Section 5.2.2. Namely, the service times can be interpreted as the time elapsed between two QCs. Consequently, the service rate is given as one over the service times. As an estimation of the service times, we use the travel time estimations between the QCs in the case of no interruptions and including the stopping time if a container operation is needed. As such, the service rates used for the JEFQP are estimated as one over the estimated service times. An overview of the service rate estimations for the different scenarios is given in Table 2. As can be seen, in the job types no distinction is made between unloading and loading operations as travel time estimations for these are similar. Additionally, no values on any row are equal to comply with the assumption made in Section 5.2.2.

Table 2: The estimated values of the service rate parameters for the different scenarios.

		Estimated value of					
Scenario type	Job type j	μ_j^1	μ_j^2	μ_j^3	μ_j^4	μ_j^5	μ_j^6
Scenario 1	(QC) 1	0.036289	0.065969	0.107223	0.184444	0.286206	0.158372
	(QC) 2	0.130897	0.039466	0.107223	0.184444	0.286206	0.158372
	(QC) 3	0.130897	0.121330	0.034196	0.081048	0.286206	0.158372
	(QC) 4	0.130897	0.121330	0.107223	0.039466	0.096056	0.158372
	(QC) 5	0.130897	0.121330	0.107223	0.184444	0.042716	0.075581
	(QC) 6	0.130897	0.121330	0.107223	0.184444	0.286206	0.042716
Scenario 2	(QC) 1	0.035513	0.081048	0.124984	0.116233	0.286206	0.158372
	(QC) 2	0.121330	0.039466	0.067034	0.116233	0.286206	0.158372
	(QC) 3	0.121330	0.184444	0.039845	0.116233	0.286206	0.158372
	(QC) 4	0.121330	0.184444	0.124984	0.035063	0.096056	0.158372
	(QC) 5	0.121330	0.184444	0.124984	0.116233	0.042716	0.075581
	(QC) 6	0.121330	0.184444	0.124984	0.116233	0.286206	0.042716
Scenario 3	(QC) 1	0.031473	0.081048	0.184450	0.184455	0.286207	0.158372
	(QC) 2	0.084342	0.039466	0.081048	0.184455	0.286207	0.158372
	(QC) 3	0.084342	0.184444	0.039466	0.081048	0.286207	0.158372
	(QC) 4	0.084342	0.184444	0.184450	0.039466	0.096056	0.158372
	(QC) 5	0.084342	0.184444	0.184450	0.184455	0.042716	0.075581
	(QC) 6	0.084342	0.184444	0.184450	0.184455	0.286207	0.042716

To evaluate the performance of the various routing policies, we consider several performance indicators. These performance indicators are calculated using the results of the simulation model. The main performance indicators are:

- Quay Crane productivity: average moves per hour that a QC performs;
- Transfer Point utilization: percentage of container operations positioned on each TP considered over the entire quay lane area;
- Quay Crane utilization: percentage of time each QC is productive complementary to the time the QC is idle or waiting for an autonomous SC;
- Autonomous Straddle Carrier productivity: average containers per hour an autonomous SC handles, as well as the average time it takes per container.

Per routing policies, these performance indicators are presented and discussed. As a result, we can evaluate and compare the performances of the various routing policies.

6.3 Practical Performance

After obtaining the results of the simulation models for all scenarios, we can compare the discussed performance indicators for the various routing policies. Section 6.3.1 compares the QC productivity for the three routing policies. Section 6.3.2 discusses the distribution for the container operations over the different TPs. In Section 6.3.3 the percentage of time each QC is active is presented. Finally, Section 6.3.4 compares the autonomous SC productivity. The results of the various performance indicators are presented per scenario.

6.3.1 Quay Crane Productivity

This section discusses the QC productivity for each of the different routing policies. As discussed before, the QC productivity depicts the average moves per hour of a QC. QC productivity is generally assumed to be the most important performance indicator of an ACT, as it corresponds with the main objective of a container terminal to minimize the time of berth of a vessel. In order to test the significance of the different routing policies, a one-sided alternative test is imposed. This test is chosen as only an improvement is accepted. Hence, the test is whether the alternative model is better or not. The test statistic is based on the following t-statistic given as

$$t = \frac{\bar{X}_A - \bar{X}_B}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}},$$

with $\bar{X}_A$ and $\bar{X}_B$ the sample variances of both selected models. Additionally, n and m are the sample variances of model A and model B, respectively. Finally, s_p is considered as the pooled sample standard error, calculated as the square root of

$$s_p^2 = \frac{(n-1)S_A^2 + (m-1)S_B^2}{n+m-2},$$

with S_A^2 and S_B^2 the sample variances of both models. The t-statistic follows a t-distribution with $n+m-2$ degrees of freedom. Table 3 lists the results of the QC productivity for all three scenarios. The final four columns depict if the specific policy performs significantly better than the policy it is compared to. For Scenario 3, no comparisons are included to the HRP for the horizontal quay lane implementation as this scenario does not allow for such a QC situation.

Table 3: The results of the Quay Crane (QC) productivity for the different scenarios. Policies designated with '++', '+' are significantly better than the policy it is compared to at level $\alpha = 0.05$, $\alpha = 0.10$, respectively, and policies with a '-' are not significantly better based on level $\alpha = 0.10$.

					Compared to			
Scenario type	Policy	n	$\bar{X}$	S^2	HRP (Hori.)	HRP (Perp.)	JLOQP	JEFQP
Scenario 1	HRP (Hori.)	25	36.3	1.64	-	-	-	-
	HRP (Perp.)	25	36.4	1.43	-	-	-	-
	JLOQP	25	37.1	2.20	++	++	-	-
	JEFQP	25	37.3	1.72	++	++	-	-
Scenario 2	HRP (Hori.)	25	35.5	1.77	-	-	-	-
	HRP (Perp.)	25	36.6	1.82	++	-	-	-
	JLOQP	25	37.2	1.61	++	+	-	-
	JEFQP	25	37.2	1.97	++	+	-	-
Scenario 3	HRP (Perp.)	25	34.2	1.60	N.A.	-	-	-
	JLOQP	25	34.7	1.14	N.A.	+	-	-
	JEFQP	25	35.0	1.67	N.A.	++	-	-

As can be seen, the QC productivity for the dynamic routing policies is higher compared to that of the HRP for either quay lane implementation. However, we are mostly interested to find these improvements to be significant. We find the improvements of the dynamic routing policies to be significant at a 10% significance level for all ten comparisons to the HRP of either quay lane implementation. Moreover, seven of these ten comparisons are significant on a 5% significance level.

Moreover, we can compare the performances of the dynamic routing policies with each other. As such, we find the JEFQP to not be able to significantly outperform the JLOQP for any scenario, despite the additional complexity used in the decision-making. This suggests most improvements are made when transitioning from a static routing policy to a dynamic routing policy. Additional complexity does not significantly add to productivity.

Comparing the QC productivity of the different routing policies over the different scenarios, we find the dynamic routing policies to produce significant improvements when compared to the benchmark policies for all scenarios. However, both dynamic routing policies perform relatively similarly as no significant differences can be found between the results.

6.3.2 Transfer Point Utilization

Secondly, the TP utilization is assessed. With the dynamic routing methods, we aimed to work more consistently on TPs in gauge, preferably on those closest to the waterside. Therefore, we present the percentage distribution of the number of container operations to be executed on each specific TP over the quay lane area.

Firstly, this distribution is presented in Figure 9 for the HRP for the different routing policies for Scenario 1. As can be seen, an equal percentage distribution is observed for the applications of the HRP of both quay lane implementations. This is to be expected as for both quay lane implementations a similar fixed order is used for the TP selection. Moreover, for each distribution, each TP is used approximately the same. Similarly, due to the fixed order, each TP is assigned equally to the arriving container operations resulting in the presented results. Considering the JLOQP we find 84.4% of container operations to be executed in gauge, which is substantially more compared to the HRP. Moreover, the JLOQP performs 90.8% of container operations in gauge.

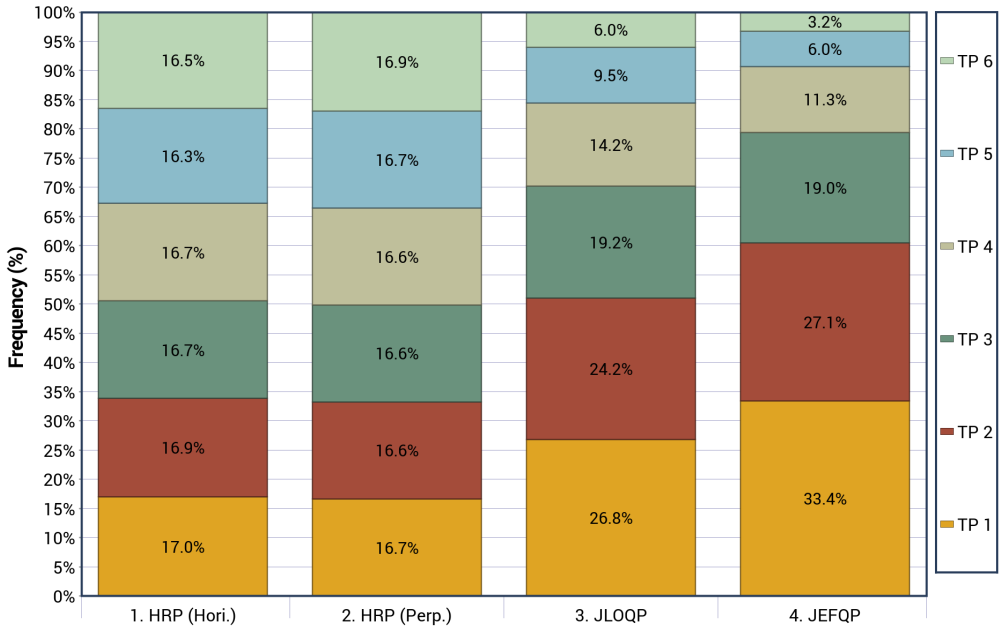


Figure 9: The Transfer Point (TP) utilization for Scenario 1.

Secondly, Figure 10 indicates this percentage distribution for Scenario 2. Generally, the results are relatively similar compared to Scenario 1. Namely, the HRP indicates similar percentage distributions for each TP complying with the remarks made for Scenario 1. The JLOQP and JEFQP perform 87.0% and 91.6% of container operations in gauge, respectively.

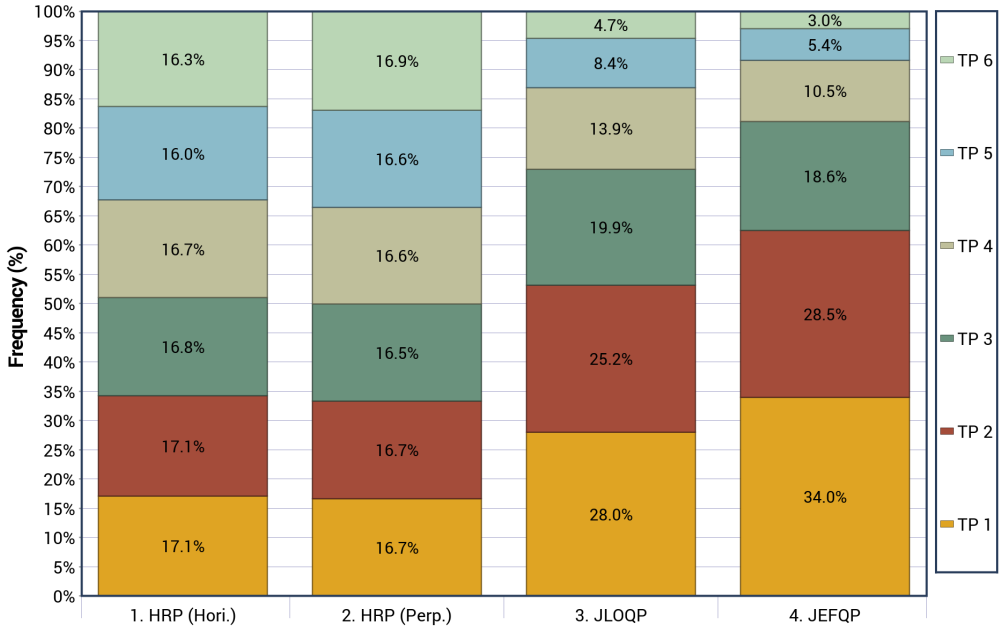


Figure 10: The Transfer Point (TP) utilization for Scenario 2.

Finally, Figure 11 presents the percentage distribution for Scenario 3. The results for HRP comply with the results previously seen for Scenario 1 and Scenario 2. However, both dynamic routing policies perform substantially fewer container operations in gauge. Namely, the JLOQP performs 72.6% of container operations in gauge, whilst the JEFQP does this for 79.4%. This can be accounted to the fact that for Scenario 1 and Scenario 2 the six QCs are divided over two QC groups. As a result, vehicles on the quay lane area for the one QC group are not being accounted for when making routing decisions for the other QC group. As in Scenario 3, we only consider a single QC group, all vehicles on the quay lane area are being taken into account, resulting in a larger probability of a vehicle being counted in the selection process.

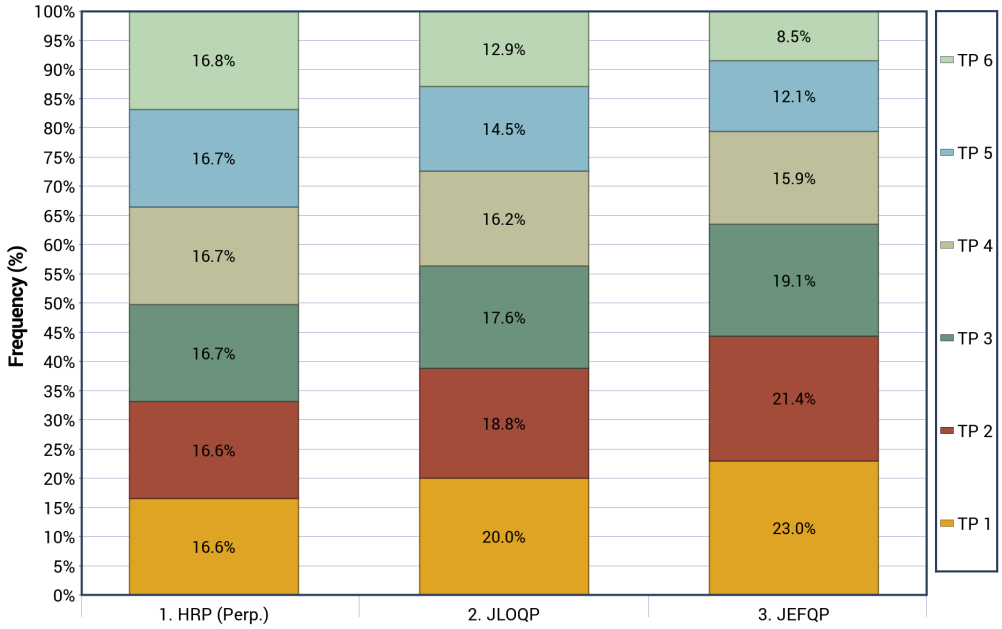


Figure 11: The Transfer Point (TP) utilization for Scenario 3.

As a result, we find that the dynamic routing policies are able to select the TP in a smarter way compared to the HRP. As expected, due to the additional components, the JEFQP outperforms the JLOQP in this performance indicator. However, most of the increase in QC productivity, discussed in Section 6.3.1, is already achieved with the initial improvement of the JLOQP.

6.3.3 Quay Crane Utilization

Thirdly, we assess the QC utilization for each routing policy. As we have shown that we execute more of the container operations in gauge, it is important to know if this result does not negatively affect the QC utilization. This performance indicator is evaluated as the percentage of time the QC is productive. Complementary, the QC can have various reasons to be inactive. Namely, a QC can be idle performing no container operations for a specific amount of time. Additionally, the QC can be changing its specific type of container operations. As said, the plan for the QC consists of bays representing each a specific container operation. As such, in the simulation model, a short amount of time is incorporated when the QC moves to the next bay of the plan. Finally, a QC can be generally waiting. We identify three different reasons for a QC to be waiting. Firstly, the selected TP is occupied by a vehicle such that the QC waits for the TP claim. Secondly, the QC can be waiting during loading. This incurs when the crane moves faster than the autonomous SCs can arrive with the next loading containers. Finally, the QC can be waiting during unloading. In this case, the QC moves faster than the autonomous SCs can arrive to pick up the unloading containers.

Firstly, we present the QC utilization of the different routing policies for Scenario 1. This is depicted in Figure 12. As can be seen, the HRP performs approximately similarly for both quay lane implementations. Moreover, the dynamic routing policies indicate slightly less productive time over the simulation period. This is mainly due to additional waiting time during loading, i.e., the crane has to wait for the autonomous SCs to arrive with new loading container operations. As such, we find that the QC indeed works faster when more containers are positioned as close as possible to the waterside. No additional waiting time during unloading is observed.

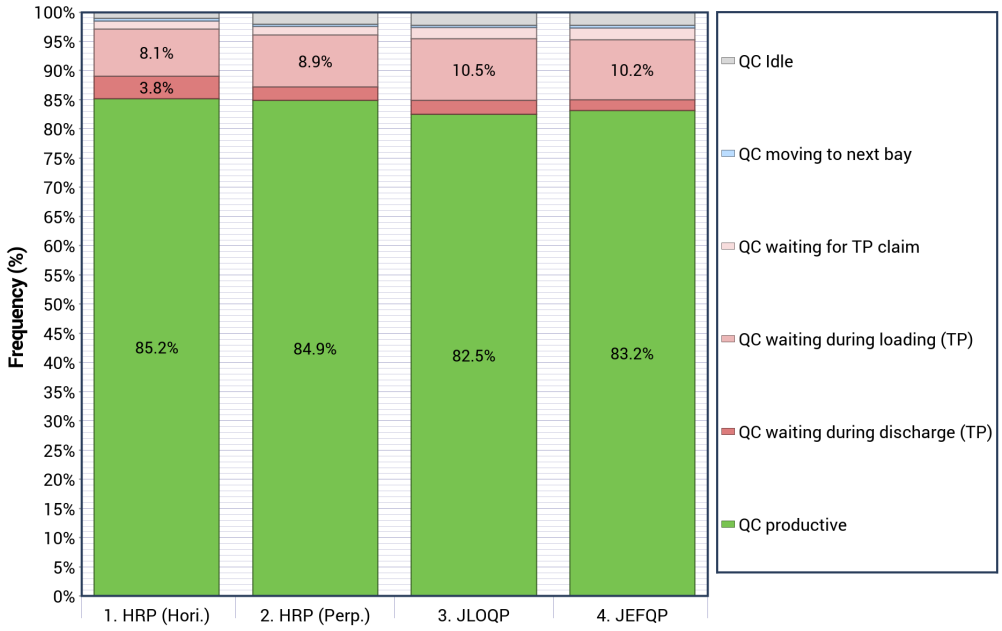


Figure 12: The Quay Crane (QC) utilization for Scenario 1.

Secondly, the results for the QC utilization are shown for Scenario 2 in Figure 13. This time we observe relatively similar results for all routing policies, i.e., no substantial differences are to be seen. Moreover, the results in terms of percentage distribution are similar as seen for Scenario 1 as both scenarios are relatively similar.

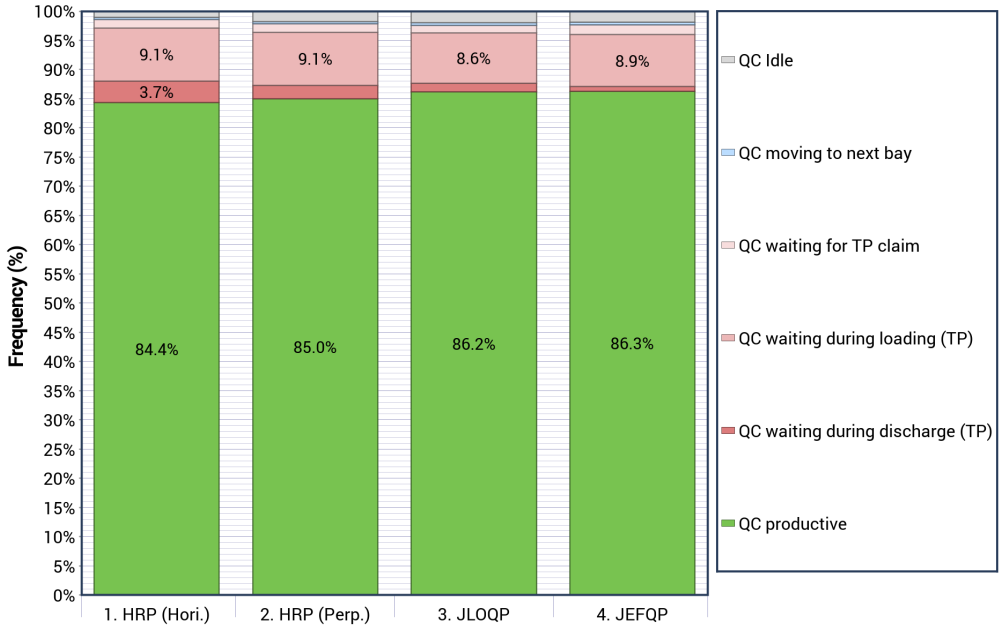


Figure 13: The Quay Crane (QC) utilization for Scenario 2.

Finally, Figure 14 indicates the QC utilization for the different routing policies for Scenario 3. Comparing the results to the previous scenarios, we observe substantially lower performance for this particular scenario. Generally, more waiting time during unloading is observed for the QCs. This can be accounted to the fact that due to the large QC group, the TPs are harder to reach for the autonomous SC due to the additional QCs that needs to be passed. Moreover, no substantial differences are observed between the different routing policies.

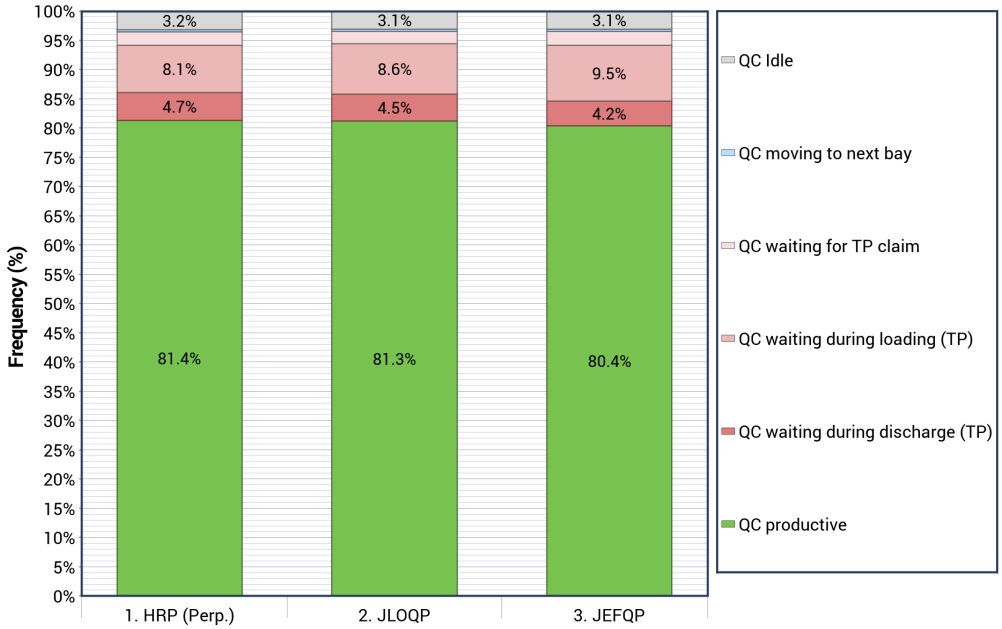


Figure 14: The Quay Crane (QC) utilization for Scenario 3.

Comparing the three routing policies, we observe no changes in the QC utilization due to the increased selection of the TPs in gauge. As a result, we can identify that the QCs are able to work at the same utilization rates, regardless of the fact it increasingly works on TPs closer to the waterside. Thus, this confirms the cause of the increase in QC productivity we have seen in Section 6.3.1.

6.3.4 Autonomous Straddle Carrier Productivity

Finally, the autonomous SC productivity is considered. With the increased use of TPs in gauge, it is unwanted that this results in more congestion or longer waiting times for the autonomous SCs. Therefore, we compare the autonomous SC productivity expressed in the average number of containers handled per hour. Moreover, we compare the average time spent per container for an autonomous SC, as well as the buildup of this accumulated time. Namely, within this time different components can be identified. Firstly, an autonomous SC can be waiting for its approach to the QC in the in-buffer, or it can wait for the grounding location to bring the container to the out-buffer. Moreover, it could wait for the determination of its route, however, as this calculation is relatively straightforward this should not accumulate to a large component of total travel time. Additionally, the time at the TPs of the QCs and YCs should be taken into consideration. Finally, an autonomous SC can be driving either empty or laden.

Firstly, we assess the results of the autonomous SC productivity for Scenario 1. Figure 15 presents these results. On the left axis of the graph, the average time per box in minutes is shown, as well as the distribution of the components of this travel time. On the right axis, the average productivity in boxes per hour is given. As seen with the previous performance indicator, the autonomous SC productivity is relatively similar to the different routing policies. The horizontal quay lane implementation requires slightly more waiting time for the QC approach as each QC has less TPs to work on. Additionally, the general time at the QC is substantially higher affecting the productivity of the autonomous SC. However, for all routing policies, most of the productive time is spent on general driving.

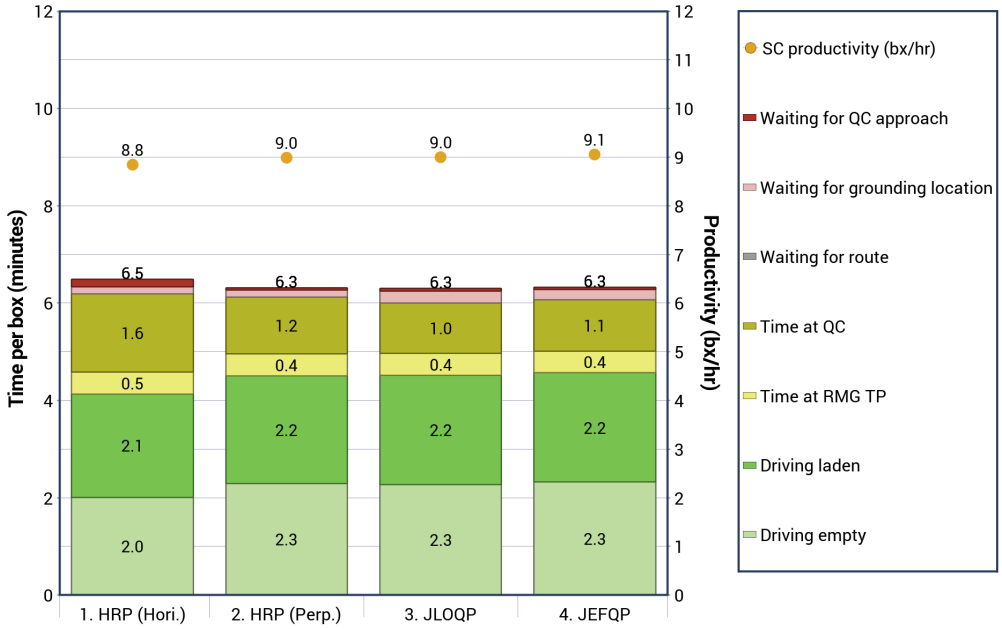


Figure 15: The autonomous Straddle Carrier (SC) productivity for Scenario 1.

Secondly, Figure 16 indicates the autonomous SC productivity for Scenario 2. Generally, the results are approximately similar as seen for Scenario 1. Similarly, no significant differences are observed for the different routing policies.

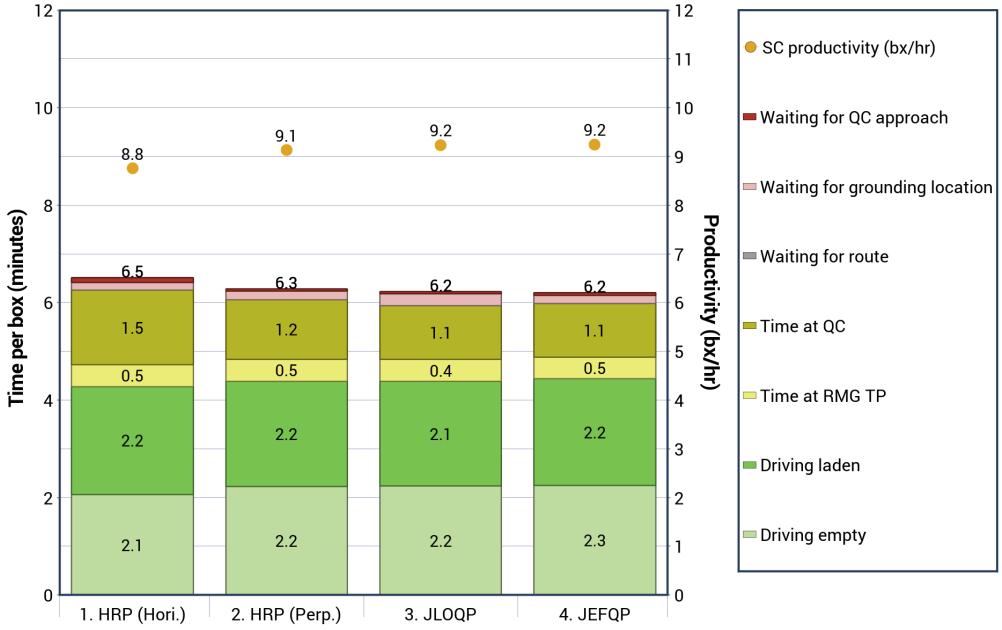


Figure 16: The autonomous Straddle Carrier (SC) productivity for Scenario 2.

Finally, the autonomous SC productivity is shown for Scenario 3 in Figure 17. Comparing these results to the results of the previous scenarios, we observe substantial increases in the time per container, complying with the decreased QC productivity previously. Additionally, the longer average travel times of the autonomous SC confirms the suggested longer travel distance over the quay lane area as reasoning for the decreased QC productivity. Moreover, no significant differences can be observed between the different routing policies.

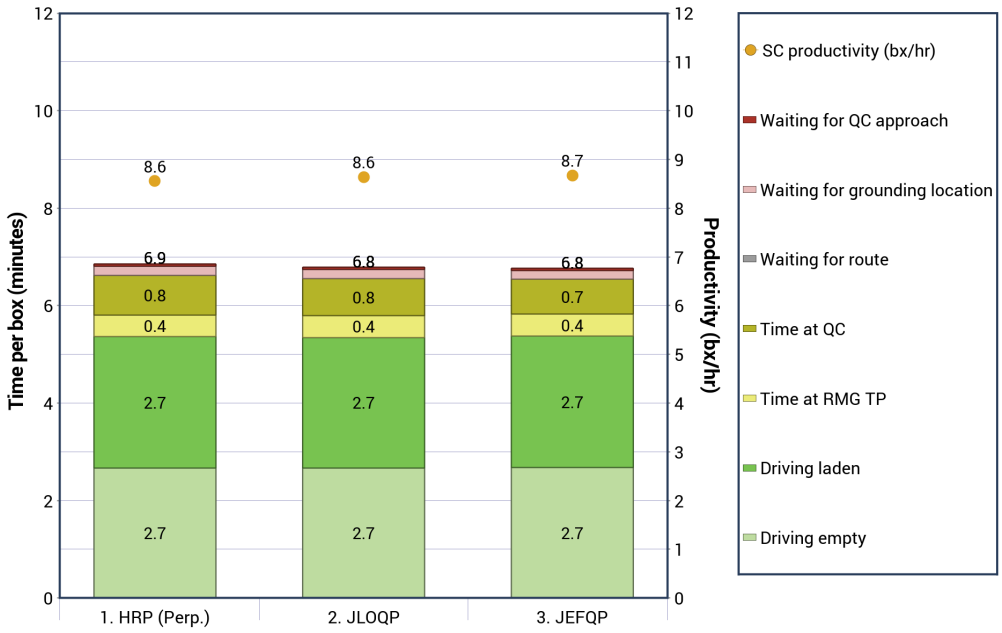


Figure 17: The autonomous Straddle Carrier (SC) productivity for Scenario 3.

Comparing the results of both dynamic routing policies and the results of the HRP, we observe no significant changes in the autonomous SC productivity. As a result, we can conclude that the dynamic routing policies effectively select a suitable TP without increasing congestion and the duration times of the autonomous SCs.

6.4 Sensitivity Analysis

In this section, we evaluate the effects of certain layout decisions or parameters included in the dynamic routing policies. Firstly, we consider the effects of the number of TPs beneath each QC, which is presented in Section 6.4.1. Moreover, Section 6.4.2 discusses the effects of the preference parameter.

6.4.1 The Number of Transfer Points

This section assesses the effects of a decreased number of TPs on the QC productivity when applying the dynamic routing policies. As discussed in Section 3, a horizontal quay lane implementation is chosen due to restricted space when transitioning from a man-operated container terminal to an ACT. However, another method to handle restricted space is to sacrifice several TPs inland in order to facilitate the space needed for a full row of perpendicular buffer locations. As such, the results of the horizontal quay lane implementation shown in Section 6.3 are more suitable for comparison. For this implementation, the six TPs are kept as this situation is already representative of a compact ACTs.

Therefore, for each scenario, the number of TPs beneath each QC is decreased to either three or four TPs. Consequently, the two TPs in backreach are abandoned to simulate the implementation of a full row of perpendicular buffer locations in a more compact ACT. For the case with three TPs, we even simulate a situation with a smaller QC allowing for less TPs in gauge. The results of the dynamic routing policies are compared to the results of the HRP for horizontal quay lane implementation. Table 4 lists the results of the QC productivity for different scenarios with variations in the number of TPs. Significance tests are included to see if the dynamic routing policies perform significantly better than the HRP. To test significance, a similar test statistic is used as presented in Section 6.3.1.

Table 4: The results of the Quay Crane (QC) productivity for different numbers of Transfer Points (TPs) for the different scenarios. Policies designated with '++', '+' are significantly better than the policy it is compared to at level $\alpha = 0.05$, $\alpha = 0.10$, respectively, and policies with a '-' are not significantly better based on level $\alpha = 0.10$.

Scenario type	Policy	Num. of TPs	n	$\bar{X}$	S^2	Compared to HRP (Hori.)
Scenario 1	HRP (Hori.)	6	25	36.3	1.64	-
	JLOQP	3	25	36.0	1.72	-
	JEFQP	3	25	35.9	1.81	-
	JLOQP	4	25	36.3	2.18	-
	JEFQP	4	25	36.4	1.86	-
Scenario 2	HRP (Hori.)	6	25	35.5	1.77	-
	JLOQP	3	25	36.6	1.69	++
	JEFQP	3	25	36.7	1.62	++
	JLOQP	4	25	36.9	2.02	++
	JEFQP	4	25	37.2	1.53	++
Scenario 3	JLOQP	3	25	33.5	1.20	N.A.
	JEFQP	3	25	33.2	1.13	N.A.
	JLOQP	4	25	34.0	1.47	N.A.
	JEFQP	4	25	34.6	1.28	N.A.

As can be seen, over the different scenarios we observe different outcomes. For Scenario 1, the QC productivity for the various cases of the dynamic routing policies is never significantly different from HRP. However, for Scenario 2, the differences are significant at a 5% significance level. Finally, for Scenario 3, the largest improvement is made in absolute terms, as this scenario can not be executed for the horizontal quay lane implementation.

Concluding from the presented results, we find the dynamic routing policies to be relatively robust. Namely, in different situations with different parameters, improvements for the dynamic routing policies can be found. Moreover, sacrificing several TPs in backreach to allow for a full row of perpendicular buffers is shown to provide at least similar QC productivity for various scenarios. Additionally, more arrangements in terms of QC positions are possible as with the perpendicular quay lane implementation no limitations on the size of the QC group are imposed.

6.4.2 The Preference Parameter

In Section 5.2, an additional preference parameter was introduced for the decision-making of the dynamic routing policies. This preference parameter allows us to balance the decisions of the policy even more towards the waterside, as technically a dummy score is being added weighted by the reference numbers of the various quay lanes. A sensitivity analysis of the results of this parameter is executed on both dynamic routing policies. In terms of the scenario, we only consider Scenario 3 as the largest improvement can be made in this scenario.

Firstly, the TP utilization is shown for both dynamic routing policies. Figure 18 and 19 present the TP utilization for different values of the preference parameter δ for the JLOQP and the JEFQP, respectively. Generally, the results for both dynamic routing policies are relatively similar. Namely, as expected, we observe the percentage distribution to shift closer to the waterside. However, as the value of δ increases, the percentage increase for the use of the TPs closest to the waterside decreases. This suggests that for larger values of the preference parameters the availability of the TPs increasingly limits the decision options. This availability is mostly limited by the working speed of the autonomous SCs and the QCs.

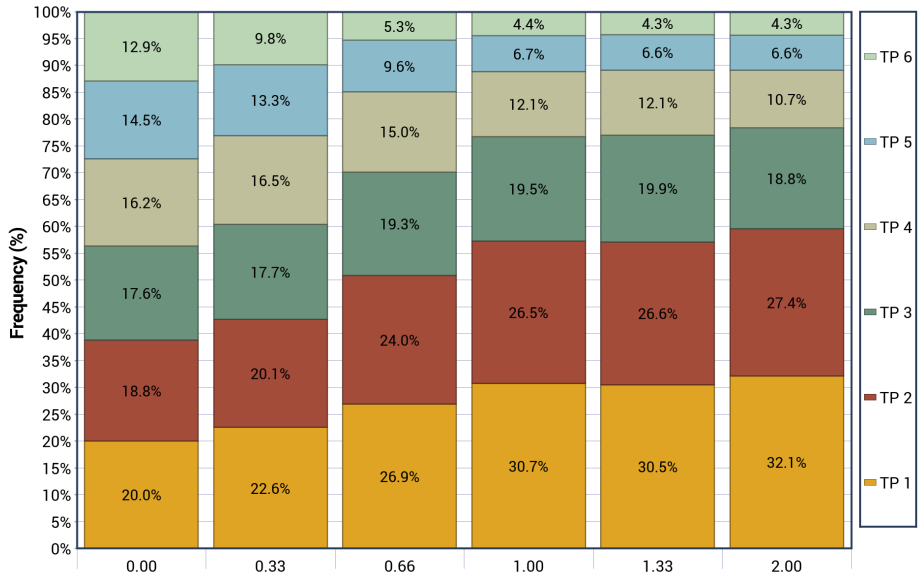


Figure 18: The Transfer Point (TP) utilization for different values of δ for the Join-the-Least-Occupied-Queue Policy (JLOQP).

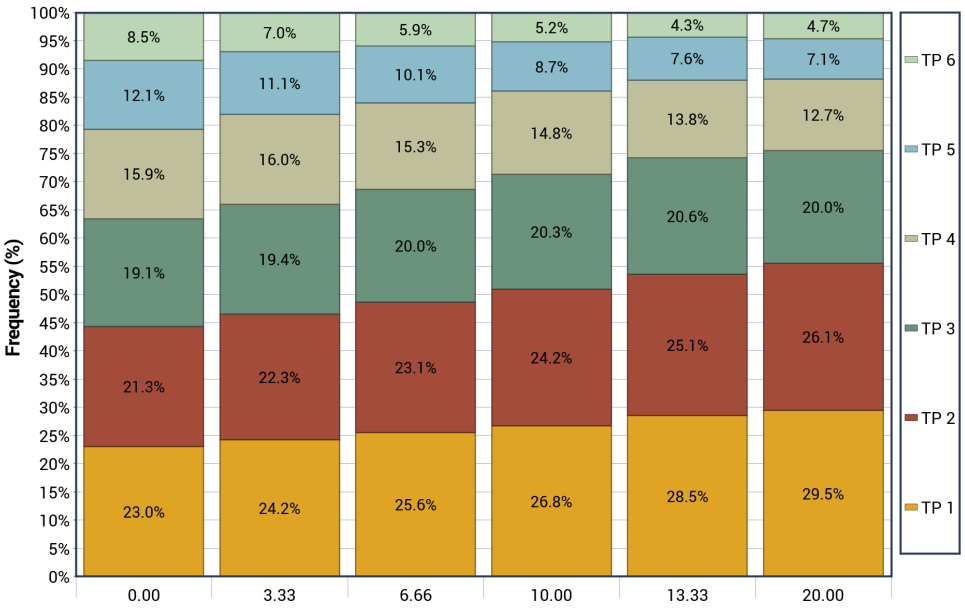


Figure 19: The Transfer Point (TP) utilization for different values of δ for the Join-the-Estimated-Fastest-Queue Policy (JEFQP).

Secondly, we can compare the QC productivity for each corresponding value of δ to see if additional preference has any effect on the general productivity. Table 5 lists the results on the QC productivity for the different values of δ for both dynamic routing policies. For the JLOQP, the differences between the different values are generally small and insignificant. Only for substantial increases of this preference parameter, significant differences are observed. However, for the evaluated values, it is clear that the policy with lower values for δ performs better.

For the JEFQP, a stronger downwards trend can be observed, suggesting an actual negative influence of the additional preference of the quay lanes closest to the waterside. However, as seen before, relatively few differences are significant. Only substantial increases of δ result in significant performance differences. This complies with the results previously seen in Section 6.3 that most improvements are made when the transition is made from a static routing policy to a dynamic routing policy. Differences due to additional preference or complexity are often insignificant.

Table 5: The results of the Quay Crane (QC) productivity for the different values of δ for both dynamic routing policies.

Policy	Value of δ	n	$\bar{X}$	S^2
JLOQP	0.00	25	34.7	1.14
	0.33	25	34.4	2.36
	0.66	25	34.4	1.97
	1.00	25	34.9	2.08
	1.33	25	34.3	1.78
	2.00	25	34.1	1.79
JEFQP	0.00	25	35.0	1.67
	3.33	25	34.8	1.26
	6.66	25	34.4	1.63
	10.00	25	34.0	2.02
	13.33	25	33.9	2.11
	20.00	25	33.7	1.94

Concluding from these results, we find that with the preference parameter δ we are able to perform even more container operations in gauge. However, similar as seen for the results in Section 6.3, we do not observe great differences in this relatively small additional shift towards the waterside of the quay lane area. Most improvements in terms of QC productivity are already made with the shift from a static routing policy to a dynamic routing policy.

7 Discussion

In this thesis, we focused on constructing a unidirectional routing policy over the quay lane area. This routing policy is aimed to be both efficient for the QCs as well as the autonomous SCs. Therefore, we introduced two dynamic routing policies to support the decision-making in the quay lane area. The Join-the-Least-Occupied-Queue Policy (JLOQP) considers merely the number of vehicles on each quay lane to select the suitable TP. The more advanced Join-the-Estimated-Fastest-Queue Policy (JEFQP) incorporates an additional distinction between the different vehicles present in the quay lane area and based the decision on a travel time estimation.

These routing policies were compared to the Hybrid Routing Policy (HRP) which is the adopted routing policy within the experimental case provided by TBA. Simulation models were used to obtain the results incorporating a total of six QCs. Additionally, three scenarios in terms of QC positions were considered to evaluate the performance in different situations. More specifically, for two scenarios the six QCs were split up into two QC groups: a group of two

QCs and a group of four QCs, or two groups of three QC. For the final scenario, we considered a single group of six QCs. Various performance indicators were used for the different routing policies to differentiate the results.

We found the dynamic routing policies to outperform the HRP in terms of QC productivity in the different scenarios considered in this thesis. This is considered to be the most crucial performance indicator on an ACT as this corresponds with its objective to minimize the time of berth of a vessel. However, the improvement of QC productivity is relatively modest. This can be accounted to the fact that the quay lane decision-making considered in this thesis is a relatively minor decision over the full ACT. The increase in QC productivity can be mainly assigned to the efficient approach to the TP selection. Both dynamic routing policies are able to make substantially more use of the TP in gauge, without resulting in any increased congestion or waiting times for the QCs or autonomous SCs.

Moreover, the effects of certain layout decisions as well as model parameters were examined. Most notably, we found that the dynamic routing policies were able to perform at least as well as the horizontal quay lane implementation when simulating a restricted scenario by means of decreasing the number of TPs. This suggests that sacrificing several TPs in backreach to facilitate a full row of perpendicular buffers is profitable in the long-term. Moreover, this allows for more possibilities in terms of QC positions as no restrictions are imposed for the size of a QC group when implementing the perpendicular quay lane implementation. Additionally, additional preference towards the quay lanes closest to the waterside does not provide additional improvements in terms of QC productivity. Insignificant differences are observed for small values of the preference parameter, while larger values resulted in significant decreases in QC productivity.

However, the additional complexity of the JEFQP does not result in an equally further increase in QC productivity. We even find the JEFQP to not be able to perform significantly better than the JLOQP in any of the evaluated scenarios. This suggests a relatively straightforward dynamic routing policy, such as the JLOQP, to provide most of the productivity improvement already. As a result, in an application, it could be left to the company to decide if the additional complexity is worthwhile.

7.1 Limitations and Extensions

Some limitations of the proposed methods are discussed and suggestions are made for further research. Firstly, most limitations arise from the assumptions made for the travel time prediction model. General assumptions on the distribution of the variables used for this model were

necessary, such that we adhered to the general distributions used in queueing systems. Other distributions as well as additional variations in the parameter values could be applied to indicate the effects of these decisions.

Moreover, assumptions were made on the blocked time of the vehicle, i.e., a vehicle always comes to a full stop behind a vehicle if it is at the same location. In practical situations, this is not necessary in every case. Therefore, with the prodigious rise in machine learning methods and their capability to predict values given certain states of a system, these models seem very applicable for this type of application. Additionally, with the known improvements due to dynamic routing policies, it could become imminent what additional improvements machine learning methods are able to produce given their additional complexity.

Additionally, the question still remains whether any additional improvements in terms of QC productivity are possible, as no information is known on the optimality of the problem. An optimization problem could be formulated for the problem formulation to get an indication of the optimality gap of the current policies.

Finally, this thesis has considered various container terminal parameter values as given. Working speeds of QCs and autonomous SCs, distributions on container operations, etc., are kept as given in the original model. In real applications, variations among these parameter values could occur, such that the effects of the variations could be further analyzed. This could allow for a more confident insight to conclude if the dynamic routing policies are indeed robust enough to deal with the variations and uncertainty present on an ACT.

Bibliography

Al-Dhaheri, N., & Diabat, A. (2017). A lagrangian relaxation-based heuristic for the multi-ship quay crane scheduling problem with ship stability constraints. *Annals of Operations Research*, *248*(1), 1–24.

Bartlett, K., Lee, J., Ahmed, S., Nemhauser, G., Sokol, J., & Na, B. (2014). Congestion-aware dynamic routing in automated material handling systems. *Computers & Industrial Engineering*, *70*, 176–182.

Bierwirth, C., & Meisel, F. (2015). A follow-up survey of berth allocation and quay crane scheduling problems in container terminals. *European Journal of Operational Research*, *244*(3), 675–689.

Cai, B., Huang, S., Liu, D., & Dissanayake, G. (2014). Rescheduling policies for large-scale task allocation of autonomous straddle carriers under uncertainty at automated container terminals. *Robotics and Autonomous Systems*, *62*(4), 506–514.

Cai, B., Huang, S., Liu, D., Yuan, S., Dissanayake, G., Lau, H., & Pagac, D. (2012). Multiobjective optimization for autonomous straddle carrier scheduling at automated container terminals. *IEEE transactions on automation science and engineering*, *10*(3), 711–725.

Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., Vogt-Maranto, L., & Zdeborová, L. (2019). Machine learning and the physical sciences. *Reviews of Modern Physics*, *91*(4), 045002.

Chen, L., Bostel, N., Dejax, P., Cai, J., & Xi, L. (2007). A tabu search algorithm for the integrated scheduling problem of container handling systems in a maritime terminal. *European Journal of Operational Research*, *181*(1), 40–58.

Chujo, T., Nishida, K., & Nishi, T. (2020). A conflict-free routing method for automated guided vehicles using reinforcement learning. *International Symposium on Flexible Automation*, *83617*, V001T04A001.

Daganzo, C. (1997). *Fundamentals of transportation and traffic operations* (Vol. 30). Pergamon Oxford.

Dragović, B., Tzannatos, E., & Park, N. K. (2017). Simulation modelling in ports and container terminals: Literature overview and analysis by research field, application area and tool. *Flexible Services and Manufacturing Journal*, *29*(1), 4–34.

Duan, Y., Yisheng, L., & Wang, F.-Y. (2016). Travel time prediction with lstm neural network. *2016 IEEE 19th international conference on intelligent transportation systems (ITSC)*, 1053–1058.

Even-Dar, E., Mansour, Y., & Bartlett, P. (2003). Learning rates for q-learning. *Journal of machine learning Research*, *5*(1).

Expósito-Izquierdo, C., Melián-Batista, B., & Moreno-Vega, J. M. (2011). Variable neighbourhood search for the quay crane scheduling problem. *2011 11th International Conference on Intelligent Systems Design and Applications*, 463–468.

Fontanelli, S., Bini, E., & Santi, P. (2010). Dynamic route planning in vehicular networks based on future travel estimation. *2010 IEEE Vehicular Networking Conference*, 126–133.

Gawrilow, E., Köhler, E., Möhring, R. H., & Stenzel, B. (2008). Dynamic routing of automated guided vehicles in real-time. *Mathematics–key technology for the future* (pp. 165–177). Springer.

Greenshields, B., Bibbins, J., Channing, W., & Miller, H. (1935). A study of traffic capacity. *Highway research board proceedings*, *1935*.

Jeon, S. M., Kim, K. H., & Kopfer, H. (2011). Routing automated guided vehicles in container terminals through the q-learning technique. *Logistics Research*, *3*(1), 19–27.

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of artificial intelligence research*, *4*, 237–285.

Kaveshgar, N., Huynh, N., & Rahimian, S. K. (2012). An efficient genetic algorithm for solving the quay crane scheduling problem. *Expert Systems with Applications*, *39*(18), 13108–13117.

Kim, K. H., Jeon, S. M., & Ryu, K. R. (2007). Deadlock prevention for automated guided vehicles in automated container terminals. *Container terminals and cargo systems* (pp. 243–263). Springer.

Lau, H. Y., & Zhao, Y. (2008). Integrated scheduling of handling equipment at automated container terminals. *International journal of production economics*, *112*(2), 665–682.

Lee, D.-H., Chen, J. H., & Cao, J. X. (2011). Quay crane scheduling for an indented berth. *Engineering Optimization*, *43*(9), 985–998.

Lee, D.-H., Wang, H. Q., & Miao, L. (2008). Quay crane scheduling with non-interference constraints in port container terminals. *Transportation Research Part E: Logistics and Transportation Review*, *44*(1), 124–135.

Meersmans, P. J., & Wagelmans, A. P. (2001). Effective algorithms for integrated scheduling of handling equipment at automated container terminals.

Musolino, G., Polimeni, A., Rindone, C., & Vitetta, A. (2013). Travel time forecasting and dynamic routes design for emergency vehicles. *Procedia-Social and Behavioral Sciences*, *87*, 193–202.

Nadarajah, S. (2008). A review of results on sums of random variables. *Acta Applicandae Mathematicae*, *103*(2), 131–140.

Pjevčević, D., Vladisavljević, I., Vukadinović, K., & Teodorović, D. (2011). Application of dea to the analysis of agv fleet operations in a port container terminal. *Procedia-Social and Behavioral Sciences*, *20*, 816–825.

Sammarra, M., Cordeau, J.-F., Laporte, G., & Monaco, M. F. (2007). A tabu search heuristic for the quay crane scheduling problem. *Journal of Scheduling*, *10*(4), 327–336.

Steenken, D., Voß, S., & Stahlbock, R. (2004). Container terminal operation and operations research-a classification and literature review. *OR spectrum*, *26*(1), 3–49.

Van Woensel, T., Kerbache, L., Peremans, H., & Vandaele, N. (2008). Vehicle routing with dynamic travel times: A queueing approach. *European journal of operational research*, *186*(3), 990–1007.

Vandaele, N., Van Woensel, T., & Verbruggen, A. (2000). A queueing based traffic flow model. *Transportation Research Part D: Transport and Environment*, *5*(2), 121–135.

Yang, Y., Zhong, M., Dessouky, Y., & Postolache, O. (2018). An integrated scheduling method for agv routing in automated container terminals. *Computers & Industrial Engineering*, *126*, 482–493.

Yu, H., Deng, Y., Zhang, L., Xiao, X., & Tan, C. (2022). Yard operations and management in automated container terminals: A review. *Sustainability*, *14*(6), 3419.

Zhang, Y., & Haghani, A. (2015). A gradient boosting method to improve travel time prediction. *Transportation Research Part C: Emerging Technologies*, *58*, 308–324.

Zhen, L. (2016). Modeling of yard congestion and optimization of yard template in container ports. *Transportation Research Part B: Methodological*, *90*, 83–104.

Zhou, P., Lin, L., & Kim, K. H. (2021). Anisotropic q-learning and waiting estimation based real-time routing for automated guided vehicles at container terminals. *Journal of Heuristics*, 1–22.

Appendix A Proof of the Travel Time Estimation Model

In this appendix, we formally proof the travel time estimation equations given in Equation 1 and Equation 2. Recall the notation used:

d : the distance to travel for the autonomous SC,

$v_{\max}$: the maximum travel speed of the autonomous SC,

$\hat{v}$: the speed at the start of the quay lane area of the autonomous SC,

$\bar{v}$: the speed at the end of the quay lane area of the autonomous SC,

a : the acceleration/deceleration rate of the autonomous SC,

t_0 : the travel time over the quay lane area without interruptions,

v' : the decelerated speed of the autonomous SC due to an interruption,

t_{add} : the additional travel time due to an interruption.

Proof of Equation 1. According to Figure 7a, the total distance traveled can be seen as the area underneath the chart. This area can be calculated as

$$\begin{aligned} d &= t_0 v_{\max} - \frac{1}{2}\left((v_{\max} - \hat{v})\,\frac{v_{\max} - \hat{v}}{a}\right) - \frac{1}{2}\left((v_{\max} - \bar{v})\,\frac{v_{\max} - \bar{v}}{a}\right) \\ &= t_0 v_{\max} - \frac{v_{\max}^2 - 2v_{\max}\hat{v} + \hat{v}^2}{2a} - \frac{v_{\max}^2 - 2v_{\max}\bar{v} + \bar{v}^2}{2a}. \end{aligned}$$

However, as we want to calculate t_0, we can rewrite this equation as

$$\begin{aligned} t_0 &= \frac{d}{v_{\max}} + \frac{v_{\max}^2 - 2v_{\max}\hat{v} + \hat{v}^2}{2av_{\max}} + \frac{v_{\max}^2 - 2v_{\max}\bar{v} + \bar{v}^2}{2av_{\max}} \\ &= \frac{d}{v_{\max}} + \frac{\hat{v}^2 + \bar{v}^2}{2av_{\max}} + \frac{2v_{\max}^2 - 2v_{\max}\hat{v} - 2v_{\max}\bar{v}}{2av_{\max}} \\ &= \frac{d}{v_{\max}} + \frac{\hat{v}^2 + \bar{v}^2}{2av_{\max}} + \frac{v_{\max} - \hat{v} - \bar{v}}{a}. \end{aligned}$$

□

Proof of Equation 2. Similarly as for the previous proof, we start with calculating the area underneath the chart in Figure 7b. Let the total travel time, i.e., t_0 combined with t_{add}, be denoted with t_{total}. Then, the area can be calculated as follows

$$\begin{aligned} d &= t_{\text{total}} v_{\max} - \frac{1}{2}\left((v_{\max} - \hat{v})\,\frac{v_{\max} - \hat{v}}{a}\right) - \frac{1}{2}\left((v_{\max} - \bar{v})\,\frac{v_{\max} - \bar{v}}{a}\right) - \frac{1}{2}\left((v_{\max} - v')\,\frac{2v_{\max} - 2v'}{a}\right) \\ &= t_{\text{total}} v_{\max} - \frac{v_{\max}^2 - 2v_{\max}\hat{v} + \hat{v}^2}{2a} - \frac{v_{\max}^2 - 2v_{\max}\bar{v} + \bar{v}^2}{2a} - \frac{2v_{\max}^2 - 4v_{\max}v' + 2v'^2}{2a}. \end{aligned}$$

Again, we can rewrite this equation as

$$\begin{aligned} t_{\text{total}} &= \frac{d}{v_{\max}} + \frac{v_{\max}^2 - 2v_{\max}\hat{v} + \hat{v}^2}{2av_{\max}} + \frac{v_{\max}^2 - 2v_{\max}\bar{v} + \bar{v}^2}{2av_{\max}} + \frac{2v_{\max}^2 - 4v_{\max}v' + 2v'^2}{2av_{\max}} \\ &= \frac{d}{v_{\max}} + \frac{\hat{v}^2 + \bar{v}^2}{2av_{\max}} + \frac{v_{\max} - \hat{v} - \bar{v}}{a} + \frac{v_{\max}^2 - 2v_{\max}v' + v'^2}{av_{\max}} \\ &= t_0 + \frac{(v_{\max} - v')^2}{av_{\max}}. \end{aligned}$$

Thus, we find that $t_{\text{add}} = \frac{(v_{\max}-v')^2}{av_{\max}}$. □

Appendix B Proof of the PDF of the Difference of Two Hypoexponential RVs

In this appendix, we proof the PDF of a RV representing the difference of two Hypoexponential RVs as given in Equation 4.

Proof of Equation 4. Assume X and Y to be Hypoexponential RVs with parameters $\mu_i,\ i = 1,\dots,n$ and $\lambda_i,\ i = 1,\dots,m$, respectively. As we assumed all parameters to be distinct, the PDFs of both RVs can be denoted as

$$f_X(x) = \sum_{k=1}^{n} \ell_k(0)\mu_k e^{-\mu_k x},\ f_Y(y) = \sum_{s=1}^{m} \ell_s(0)\lambda_s e^{-\lambda_s y}$$

Let the difference between the Hypoexponential RVs be denoted as $Z = X - Y$. Using the convolution theorem, we can denote the PDF of Z as

$$f_Z(z) = \int_{\mathbb{R}} f_X(x) f_Y(x - z)dx$$

However, as both X and Y have a support region equal to $[0, \infty)$, we have to account for values of z positive or negative. Namely, when $z < 0$ the integral is evaluated in $(0, \infty)$, while for $z \geq 0$ in (z, ∞). Therefore, we evaluate the integral for two different cases.

Case 1: $z < 0$

$$\begin{aligned}
f_Z(z) &= \int_0^{\infty} f_X(x) f_Y(x - z)dx \\
&= \int_0^{\infty} \sum_{k=1}^{n} \ell_k(0)\mu_k e^{-\mu_k x} \sum_{s=1}^{m} \ell_s(0)\lambda_s e^{-\lambda_s(x-z)} \\
&= \sum_{k=1}^{n}\sum_{s=1}^{m} \ell_k(0)\ell_s(0)\mu_k\lambda_s e^{\lambda_s z} \int_0^{\infty} e^{-(\mu_k+\lambda_s)x}dx \\
&= \sum_{k=1}^{n}\sum_{s=1}^{m} \frac{\ell_k(0)\ell_s(0)\mu_k\lambda_s}{\mu_k + \lambda_s} e^{\lambda_s z}, \qquad\qquad z < 0.
\end{aligned}$$

The last equality holds as it is know that $\mu_k + \lambda_s > 0,\ \forall\, k, s$, as the parameter values of a Hypoexponential RV are always assumed to be positive.

Case 2: $z \geq 0$

$$\begin{aligned}
f_Z(z) &= \int_z^\infty f_X(x) f_Y(x-z) dx \\
&= \int_z^\infty \sum_{k=1}^{n} \ell_k(0) \mu_k e^{-\mu_k x} \sum_{s=1}^{m} \ell_s(0) \lambda_s e^{-\lambda_s (x-z)} \\
&= \sum_{k=1}^{n} \sum_{s=1}^{m} \ell_k(0) \ell_s(0) \mu_k \lambda_s e^{\lambda_s z} \int_z^\infty e^{-(\mu_k + \lambda_s)x} dx \\
&= \sum_{k=1}^{n} \sum_{s=1}^{m} \frac{\ell_k(0) \ell_s(0) \mu_k \lambda_s}{\mu_k + \lambda_s} e^{\lambda_s z} e^{-(\mu_k + \lambda_s) z} \\
&= \sum_{k=1}^{n} \sum_{s=1}^{m} \frac{\ell_k(0) \ell_s(0) \mu_k \lambda_s}{\mu_k + \lambda_s} e^{-\mu_k z}, \qquad z \geq 0.
\end{aligned}$$

As a result, using indicator functions, the PDF can be written as

$$f_Z(z) = \sum_{k=1}^{n} \sum_{s=1}^{m} \frac{\ell_k(0) \ell_s(0) \mu_k \lambda_s}{\mu_k + \lambda_s} \left(e^{\lambda_s z} \mathbb{1}_{\{z<0\}} + e^{-\mu_k z} \mathbb{1}_{\{z \geq 0\}} \right).$$

□

Appendix C Simulation Modeling

This thesis focuses on a specific operation within an ACT. However, other operations are necessary for the execution of container jobs. These operations are required in order to run simulation experiments and, thus, are briefly discussed in this appendix. More specifically, we consider three particular methods. In Section C.1, the methods regarding dispatching are discussed. In Section C.2, the routing methods for the vehicles are considered. Finally, in Section C.3, the area claiming of the vehicles is handled.

C.1 Dispatching

Dispatching refers to the methods used to assign a container job to a specific vehicle. Dispatching is handled by a central planning system that uses a scoring mechanism for the assignment process. At first, it identifies all available vehicles to execute a job. These available vehicles have been idle since they finished their last job. The central planner can select among a large group of vehicles as in ACTs a pooling strategy is used. This pooling strategy allows a group of vehicles to work for a group of destinations, e.g., vehicles 1 to 30 working for all QCs, or vehicles 1 to 15 working on QCs 1 to 3, and vehicles 16 to 30 working on QCs 4 to 6.

After obtaining a list of all available vehicles, it needs to be checked which jobs can be executed without causing deadlock situations and within a certain time window. In order to prevent high-density areas of vehicles, destinations may have different maximum allowed vehicles to work towards it dependent on the type of operation. If this maximum number of vehicles is already reached, the current jobs for this destination are set on hold.

Finally, if all previous criteria are verified, an arriving job is assigned to a vehicle based on a certain score. This scoring is done based on the following parameters:

- Travel distance from the current location of the vehicle to the origin location of the job. Shorter distances are preferred in order to minimize empty traveling of vehicles;
- Urgency of the job. This is measured by the difference in slack time between the actual arrival time of the vehicle at the origin location of the job and the latest allowed arrival time of the vehicle in order to arrive at the job destination in time. More slack time is preferred in order to deal with unforeseen conflicts along the route of the vehicle;
- Number of jobs already directed to the corresponding locations of the arriving job. It is preferred to send a vehicle to a location with few other vehicles underway to prevent heavy congestion and deadlock situations.

The final score is calculated based on a weighted average among these parameters. The scoring weights can be tuned for different concepts in order to obtain good dispatching which balances low driving distances as well as job urgency.

C.2 Routing

After a job is assigned, the routing method allows for an efficient route to travel the container from its origin to its destination. The part of the route that travels the vehicle over the quay lane area is covered in this thesis. However, the vehicle also has to travel between the buffer area to the stacks in the storage yard and vice versa. These specific routing decisions are discussed in this section.

The routing method within these areas is mainly considered as an organization strategy of lanes. Namely, lanes positioned closely together and horizontally to each other are considered a lane group. All lanes within this lane group are selected to serve a particular purpose. If we would consider a highway lane area with six lanes, an organization could be: lanes one and two are reserved for vehicles traveling from one QC to another QC, lanes three and four are used for vehicles traveling from a stack module to a QC, and lanes five and six are used for vehicles traveling from one stack module to another stack module. Generally, both lanes for a particular purpose are chosen to work in opposite directions.

The organization of the quay lane area can be done in two ways as seen in Section 3. In the case of the perpendicular buffer locations, the quay lane area is used in only one direction. Adjusting the situation to the horizontal buffer locations, vehicles can drive bidirectionally over the quay lane area. However, in both situations, the vehicle has to enter the quay lane area through the in-buffers, and, consequently, leave through the out-buffers. This allows for a straightforward organization of the quay lane area for both situations.

With the organization of the lane groups in place, the route determination is relatively simple. The shortest route can be calculated, for example using Dijkstra's algorithm. However, the decision options are restricted and the decision generally boils down to picking an appropriate lane complying with implemented lane organization and the specifications of the origin and destination. Moreover, the entry and exit locations of the lanes need to be determined. Similarly, these last decisions are highly guided as the locations of the buffers, loading points, and unloading points are generally positioned directly next to the various lane areas.

C.3 Claiming

Finally, claiming methods are used to guarantee conflict-free routing of the vehicles over the transportation area of the ACT. Claiming refers to the action of a vehicle in which it exclusively reserves a specific area on an ACT. If the area is successfully claimed, only this vehicle can enter the area and traverse it without conflict.

The method for claiming is relatively straightforward. As the route of the vehicle is determined, it knows its next destination. Therefore, it constantly claims a part of this route which it will traverse next. If another vehicle is traversing this area or another vehicle has just claimed it, the claiming attempt of the vehicle is not granted. In this case, the vehicle has to wait until the area is available again. Once the claim is granted when the area is available again the vehicle continues along its route.

Turning points and crossings are crucial areas prone to conflicts and deadlock situations. Therefore, if the vehicle needs to make a turn, the full turning area needs to be claimed before the turn is made. These restrictions for the vehicles in combination with the organization strategy for the lanes over the quayside area prevent most deadlock situations on the ACT.