CIRRELT 1971 - 2021

CIRRELT-2022-02

Port Rail Shunting Scheduling Problem

Daniela Ambrosino
Veronica Asta
Teodor Gabriel Crainic

January 2022

Bureau de Montréal
Université de Montréal
C.P. 6 128, succ. Centre-Ville
Montréal (Québec) H3C 3J7
Tél : 1-514-343-7575
Télécopie : 1-514-343-7121

Bureau de Québec
Université Laval,
2325, rue de la Terrasse
Pavillon Palasis-Prince, local 2415
Québec (Québec) G1V 0A6
Tél : 1-418-656-2073
Télécopie : 1-418-656-2624

Port Rail Shunting Scheduling Problem

Daniela Ambrosino[1], Veronica Asta[1], Teodor Gabriel Crainic[2,*]

1. Department of economics and business studies, University of Genoa, and Research center on logistics, transport and infrastructures - CIELI, Genoa, Italy
2. Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT), and Department of Management and Technology, Université du Québec à Montréal

Abstract. This paper deals with the Port Rail Shunting Scheduling Problem (PRSSP), a problema rising in the rail-sea yard where the modal switch between maritime and rail transportation is performed. In particular, the focus is on the trains' transfer operations within the port area, between the maritime terminals, and the railway network. These operations are performed in an area called shunting zone characterized by some infrastructures (i.e., tracks, shunting parks) and some other limited resources. The PRSSP consists in defining the schedule of all activities necessary for transferring export and import trains from the railway network station to the terminals and vice versa, respecting the time limits imposed by the railway network schedule and by the ships one, and the limits due to the finite resources available in the shunting zone. An operations-time-space network representing the rail station and the terminals (either the origin or the destination of the trains) and the operations that might be performed on the trains in each zone of the port is used for modeling the problem. A flow model based on this operations-time-space network is used for solving PRSSP. Extensive experimental tests are reported and show the validation of the model as a useful tool for the shunting managers to either evaluate the maximum number of trains that can be served or define the bottlenecks of the port system.

Keywords: Rail shunting, rail-sea yards, maritime ports, scheduling, operations-time-space network.

Acknowledgements. The authors thank CIRCLE S.P.A. for the contribution to this research. While working on the project, the third author was Adjunct Professor, Department of Computer Science and Operations Research, Université de Montréal. We gratefully acknowledge the support provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), through its Discovery Grant program, and by the Fonds de recherche du Québec through their infrastructure grants.

Results and views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect those of CIRRELT.

Les résultats et opinions contenus dans cette publication ne reflètent pas nécessairement la position du CIRRELT et n'engagent pas sa responsabilité.

* Corresponding author:teodorgabriel.crainic@cirrelt.net

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2022

© Ambrosino, Asta, Crainic, and CIRRELT, 2022

1 Introduction

Goods are transferred every day from origins to final destinations and, generally, shipments involve more than a single mode of transport. The distances between origins and destinations and the importance of multi-modality increases continually with the growth of international trade.

Intermodal transportation is defined in Crainic and Kim (2007) as the transportation of a person or a load from its origin to its destination by a sequence of at least two transportation modes, the transfer from one mode to the next being performed at an intermodal terminal. In particular, as reported in Gronalt et al. (2019), intermodal transport describes the movement of goods in the same loading unit (container, swap body, semitrailer) or vehicle (truck) by successive modes of transport without handling of the goods themselves while changing the modes [rail (trains), road (trucks), inland waterways (vessels/barges)].

Examples of operations research opportunities in intermodal freight transport research are reported in Abbassi et al. (2019), Caris et al. (2013).

Maritime transport should be considered the most relevant mode in the global goods exchange. It is a significant part of the global supply chains, where shippers, public and private navigation companies, port authorities, import-export companies, carriers, consignees, and receivers take part in a collaborative environment to better manage exchanges and trans-shipments (Muñuzuri et al. (2019), Lee and Song (2017), Krämer (2019)).

The activities in maritime ports are important for improving the freight transportation system and they have to be well managed to facilitate the modal switches among maritime transport and both rail and road ones. Containers are generally used in intermodal transports. New planning and control activities in container terminals are described in Kim and Lee (2015). The authors stress the need for better coordination among those crucial activities necessary to connect the seaside to the land-side.

The management of the yard (Carlo et al. (2014b)), as well as the transport operations (e.g., Carlo et al. (2014a)) inside the terminals, have a great impact on the efficiency of unloading and loading operations on both ships and trains/trucks. Anyway, papers integrating seaside, transport, and land-side operations are rare. Yin et al. (2011) propose a model with four agents that negotiate to obtain a berth allocation schedule, vehicle requirements, a storage allocation schedule, and a truck/train schedule for intermodal moves. Schepler et al. (2017) introduce an optimization model of a multi-terminal, multimodal maritime container port. They define the scheduling of ships, trains and trucks on terminals while limiting inter-terminal transport of containers and minimizing weighted turnaround time. They offer a tool for evaluating container port management. The objective is to minimize the weighted turnaround time of ships and trains, which operate on fixed timetables, contrary to trucks. An assumption is the use of a truck appointment system. The model provides the berthing positions and time windows to serve ships, the rail tracks and time windows to serve trains as well as the time windows for the truck appointment. The goal is to improve port performance by coordinating operations between terminals. We think that this is the first attempt of integration.

As far as the road connections are considered, some papers dealing with the truck management system, as a strategy to schedule the arrival of trucks, can be found (see for example Mar-Ortiz et al. (2020)). Few papers deal with collaborative aspects, that seem to be the ideal/necessary process to obtain better results. Phan and Kim (2016) discuss a new appointment process by which trucking companies and terminals collaboratively determine truck operation schedules and truck arrival appointments.

For what concerns the rail connections, the interested actors are rail terminal operators and railway network managers (Heggen et al. (2016)). The terminal operator has to collaborate with the rail network manager for transferring trains from the terminal to the railway network.

To the authors' knowledge, no papers deal with collaboration, and few focus on the organization of rail transport connections (Xie and Song (2018)). This paper is one of the first attempts to investigate in detail a problem arising for connecting seaside and rail land-side.

This work focuses on the modal switch between maritime and rail transportation, thus on the activities performed in the rail-sea yard (Bektas and Crainic (2007)). In particular, the focus is on the trains' transfer operations within the port area, between the maritime terminals and the railway network. These operations are performed in an area called shunting zone characterized by some infrastructures (i.e., tracks, shunting parks) and some other limited resources (Krämer (2019)).

The management of these operations is strongly affected by the schedules of both the trains on the railway network and the ships; also delays and disruption in one of the connected transportation systems impact these activities.

Thus, the scheduling (and the re-scheduling) of the operations for transferring the freight trains within the port's area, is a typical problem faced by the shunting managers.

Note that scheduling and rescheduling are well-known problems (see for example recent surveys in Allahverdi (2016), Janiak et al. (2015)) but, to the authors' knowledge, the scheduling of trains in the port shunting area is an emerging problem.

This work contributes to defining a new scheduling problem (re-scheduling is not included in the current analysis) faced by the shunting managers, and proposes an optimization model framework based on a flow network to solve it.

The paper is organized as follows: Section 2 introduces the Port Rail Shunting

Scheduling Problem (PRSSP) and the related literature; the problem definition and the basic notation are reported in Section 3; Section 4 describes the model approach; Section 5 shows the experimental results, while in Section 6 some conclusions are reported.

2 Port Rail Shunting Scheduling Problem (PRSSP)

2.1 Problem context

The exchange node between maritime and rail transportation, a port, includes three distinct elements: train stations, shunting zones, and maritime terminals. These elements can be combined and connected in different ways Ambrosino et al. (2021).

Each train station is connected with both the railway network and one or more shunting zones. The shunting zone permits the transfer of trains from the railway network to the terminals through the train station, and vice versa, and is composed of tracks for transferring trains and one or more parks where the trains can wait. The maritime terminal receives the train, unloads goods that later will be loaded on ships, then loads the train (with goods waiting in the yard) and, in the end, waits for the train transfer outside the terminal itself.

Note that, in some ports, arriving and departing trains transport goods belonging to different maritime terminals, in other ports all cargo loaded on a train has a single maritime terminal as destination/origin. Therefore, in the first case discussed for example in Rusca et al. (2019), trains arriving at the shunting zone have to be split into cars depending on the terminal of destination (note that the departing trains have to be composed of cars arriving from different terminals). In the second case, proposed in Ambrosino and Asta (2019) the whole train arriving at the shunting zone has to be transferred to the terminal of destination (and vice versa for departing trains).

Our attention focuses on the activities allowing the transfer of the whole trains from the entry (the train station) to the terminals and vice versa, that is on all the activities realized in the shunting zone. These activities, which must be organized in such a way to permit an efficient usage of the available resources, are strongly affected by the schedule of trains on the railway network and also by the schedule of the ships docked at the terminals.

In the export cycle, a train arriving at the station can either wait until it is moved to a shunting park or reach immediately a park. Then, the train can have to wait in the park until the terminal is ready to receive it. It will be finally transferred to the terminal. In some particular cases, the train may be positioned in more than one shunting park before reaching the terminal; in the opposite case, it can be directly transferred from the station to the terminal.

In the import cycle, the train is loaded in the terminal and then leaves for being transferred to a shunting park. After that, it can be transferred either directly to the rail station for its immediate departure in the railway network or in a shunting park waiting for its departure scheduled time. Note that, for import trains, some security and stability checks are usually required. These controls may be realized either in the shunting zone or in the terminals. In this latter case, the train may be directly transferred from the terminal to the rail station.

It is convenient to distinguish all the activities performed in the shunting zone in waiting operations and shunting ones. The shunting operations comprehend the movement of the trains in the port area and, if necessary, stability and security checks. Note that, waiting operations can be on the tracks of either the station or the parks, while, generally, shunting operations are performed on tracks, except for checks. Waiting operations starting and ending times depend on both the times of the shunting ones and the train arrival/departure scheduled time.

The difficulty is scheduling all these activities and managing accurately the limited resources necessary to realize them, that are related to tracks, locomotives, specialized workforce, pilots, etc. Note that the shunting managers plan all the activities to realize in the shunting zone at least 6 months in advance with respect to the time horizon to schedule.

The length of the time horizon is chosen in such a way to be representative of the activities usually realized in the port, i.e., a week. In this way, the weekly schedule can be repeated week by week, unless some changes are required.

In fact, besides this planning scheduling problem, an operative, sometimes real-time, rescheduling problem occurs. The shunting manager, during T, has to reschedule waiting and shunting operations when delays, cancellations, other disruption events occur.

Note that this paper only focuses on the scheduling problem.

Summarising, this problem consists in defining the schedule of all activities necessary for transferring trains from the railway network to the terminals and vice versa, respecting the time limits imposed by the railway network schedule and by the ships one, and the limits due to the finite resources available in the shunting zone.

This problem is here denoted Port Rail Shunting Scheduling Problem (PRSSP).

In Section 3 the detailed problem definition with the basic notation is presented. Here below a brief literature overview on works that present some aspects similar to the PRSSP is reported.

2.2 Literary overview

The main aim of the problem under inspection is to schedule a given number of activities necessary for transferring trains through the port area. Most of the literature on rail scheduling focuses on the whole railway network instead of on a node of the network, as the present work. Moreover, few literature related to schedule in a single node (i.e. a rail yard) faces problems that are quite different from PRSSP.

The rail yard is a railway depot used to perform the operations on arriving and departing trains. A depot consists of a set of tracks and, depending on the type of tracks, it is called shunting yard or marshalling yard. In the first case, all the tracks can be approached from only one side, in a marshalling yard each track can be approached from both sides (Di Stefano and Koči (2004)). Shunting operations consist of the split of the inbound trains into single cars or blocks and into build new outbound trains ready for their departure. Often the schedule regards either to a portion of the train or to a wagon/car, and the related more popular problems are known as train unit scheduling problem (TUSP) (Lin and Kwan (2014), Lin and Kwan (2016)) and train unit assignment problem (TUAP) (Cacchiani et al. (2012)).

Boysen et al. (2012) reviews the literature on the operational processes at shunting yards over the last 40 years.

Cordeau et al. (1998) reviews analytical yard models and underlines that yard policies concern the specification of the activities to be performed in the yards of a rail network. More precisely, there are a lot of models indicating how trains entering each yard should be inspected and disassembled, and how cars should be sorted and reassembled into blocks that will form new outbound trains. Among these, the Classification Problem (CP), specifically addressed also in a dedicated chapter of the recent Borndörfer et al. (2018), consists in modeling the train formation and the train sorting.

The rail process through the shunting zone of a port area has rarely been addressed and, as consequence, the related literature is few. It has been recently described in Krämer (2019) and Ambrosino et al. (2021). In Ambrosino and Asta (2019) the shunting operations of freight trains either within the port area or the intermodal inland terminal are introduced. Moreover, the authors present a discussion on possible approaches for solving the scheduling of shunting operations.

Some works present simulation approaches for evaluating transit capacity through its various compartments of port shunting yard, as Rusca et al. (2018) and Rusca et al. (2019). Other studies with the aim of sizing the port rail networks and planning shunting operations for container terminals are Caballini et al. (2016) and Fioribello et al. (2016).

Tomii et al. (1999) and Tomii and Zhou (2000) deal with the shunting scheduling problems set in a rail station, called depot. The problems include the schedule of both tasks, shunting (that is movement of trains between tracks) and workforce assignment.

These works on general scheduling and re-scheduling problems in the rail field address a problem similar to the port rail shunting optimization one. The authors deal with the shunting scheduling problems set in a rail station. The problems include the schedule of both tasks, shunting, and workforce assignment. In this problem, the movement of trains between tracks is called shunting. Shunting is necessary when either a train which arrives at a track and is scheduled to depart from a different track, or a train which arrives and departs from the same track, but another train is planned to use the track in the same time interval, or a train on a side track if other train uses it. Tasks to be scheduled are related to several types of inspections, maintenance, cleaning of rolling stock, and the shunting of rolling stock. In the schedule of tasks, the type, the track where the task is performed, start and end times are prescribed. In the scheduling of shunting, origin and destination tracks and their execution timings have to be decided. Workforce assignment includes the assignment of the workforce for each task together with the time for rest. To summarize, the aim of the depot shunting scheduling problem is to decide for trains that need shunting, side tracks to be assigned and shunting times. Arrival and departure times of trains and tracks are prescribed by train schedules. That is exactly as in port rail shunting optimization problems.

Unfortunately, no more literature exists on these scheduling and re-scheduling problems for the port rail shunting activities, to the authors' knowledge. Only some papers related to the trains management in the hinterland terminals can be cited; a recent and interesting review is Borndörfer et al. (2018).

3 Problem definition and basic notation

In the present paper, the port (i.e., the modal exchange node) under consideration includes one railway station, a shunting zone, and several maritime terminals. The railway station has a given number of tracks connected with the shunting zone. The shunting zone is composed of a park with several tracks and some tracks that join the station to the shunting park, the shunting park to the terminals, and the station directly to the terminals. The considered layout is shown in Fig. 1, where, just to let be the figure more readable, it is assumed that two tracks are present in the railway station and the shunting park, and two maritime terminals can be reached. Let us describe the specific elements in more detail.

- z^0 represents the whole railway network; it is included in this sketch just for completeness since it represents the rail transport system imposing some time constraints to the shunting manager;
- z^1 represents the railway station with its two tracks here denoted (z_1^1 and z_2^1) where the waiting operations for trains arriving from either the railway network or the shunting area can be performed;

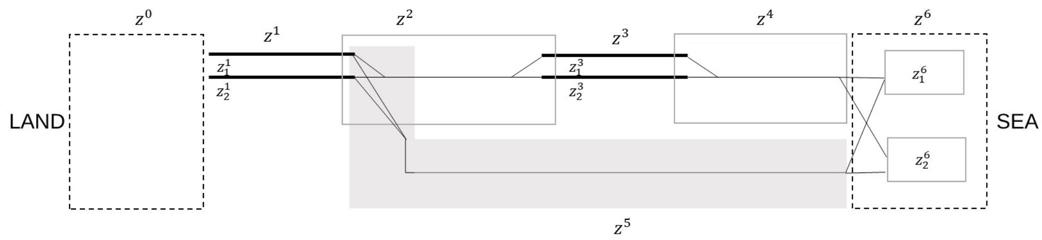

Figure 1: Port modal exchange node - Physical layout

- z^2 represents the connection between the station (z^1) and the park (z^3). This zone, inside the shunting area, is here called primary area. The connection track in the primary area is used to transfer trains from the railway station to the shunting park and viceversa;
- z^3 represents the shunting park, in the following just called park, with its two tracks here denoted (z_1^3 and z_2^3) where the waiting operations for trains arriving from either the railway station or the terminals can be performed;
- z^4 represents the connection between the park (z^3) and the terminals (z^6). This zone, inside the shunting area, is here called secondary area. The connection track in the secondary area is used to transfer trains from the shunting park to the terminals and viceversa;
- z^5 represents the connection between the station (z^1) and the terminals (z^6). This zone, inside the shunting area, is here called unique area. The connection track in the unique area is used for direct transfer of trains from the railway station to the terminals and viceversa;
- z^6 represents the seaside exchange node, here represented by two terminals (z_1^6, z_2^6), that are either the origin or the destination of trains passing through the port.

Let us indicate as I the set of import trains, those having the terminals as origin, and E the set of export trains, those having the terminals as destination. Thus, J, the union of sets E and I, represents the trains that have to be managed during the time horizon T.

Time horizon T is discretized; T is split into τ equal time intervals $T = \{1, 2, ...\tau\}$. Each time interval t in T is here denoted $[t; t+1)$ and when an operation is related to time t (i.e., has index t) means that it happens in the interval $[t; t+1)$.

An export train j arrives at the railway station in a given time instant e_j (i.e., at the beginning of time interval $[e_j; e_j + 1)$), and has to be at destination, i.e., inside its destination terminal (p_j), within a given time window $[e_j^{min}; e_j^{max}]$ that is a subset of T.

An import train j' is ready inside its origin terminal $p_{j'}$ for being picked up within a given time window $[e_{j'}^{min}; e_{j'}^{max}]$ and has to depart from the railway station in a given time instant $e_{j'}$.

Each train has to perform specific operations in order to be transferred within the port area, that is it has to pass through some of the different zones in the shunting area (see Fig.1).

Let us indicate Z the set of zones of the port area in which shunting and waiting operations are preformed and Z_j the set of zones that must be visited by train j.

An export train may have to realize one of the following paths:

- it arrives at the railway station z^1 in its fixed arrival time, enters the shunting zone for reaching the unique area z^5 and finally enters in its destination terminal in z^6;
- it arrives at the railway station z^1 in its fixed arrival time, enters the shunting zone for reaching the primary area z^2, enters the park z^3, enters the secondary area z^4 and finally enters in its destination terminal in z^6.

For an import train the possible paths are the following:

- it leaves its origin terminal in z^6, enters the shunting zone for reaching the unique area z^5 and finally arrives at the railway station z^1 from which it will enter the national railway network in its fixed departure time;
- it leaves its origin terminal in z^6, enters the shunting zone for reaching the secondary area z^4, enters the park z^3, enters the primary area z^2 and finally arrives at the railway station z^1 from which it will enter the national railway network in its fixed departure time.

The time necessary to execute the operations in each zone on each train is known too, and is here denoted $d_{i,j}$ $\forall i \in Z, \forall j \in J$.

Note that, there is a capacity for the number of operations that can be realized in each time interval in each zone, depending on the specific layout of these zones. Let u_t^i be the maximum number of operations that, in time interval t, can be realized in zone i, $\forall i \in Z$, $\forall t \in T$.

Moreover, there is a limited number of teams for realizing all the operations in the different zones. Thus, let be k_t the maximum number of shunting operations that can be executed in t in the shunting zones z^2, z^4 and z^5.

In the present problem, in each zone only one type of operation can be executed on the trains; thus, in the following, we will identify the operations with the zones.

The decisions to take are related to the time instant in which to perform the required operations on each train, in such a way to respect time constraints (the time windows for entering and leaving terminals, the arrival/departure time at the railway station) and capacity constraints related to the resources needed to perform the shunting operations.

Due to the above-mentioned limited resources, a train has often to wait when passing from one zone to another in the port area. A train can wait in the railway station and the park. In these cases, the train is performing a waiting operation. Thus, the time spent by a train to go through its path depends on the duration of both the shunting operations and the waiting ones. The main aim is to perform all the required operations on trains minimizing their total waiting time.

4 Model approach

The model used for solving PRSSP is based on an operations-time-space network representing the operations that might be performed on the import and export trains in each zone of the port (see Fig. 1) and the terminals that are either the origin or the destination of the trains. This network has been derived from the one proposed in Ambrosino and Asta (2021).

The nodes of the network, representing the zones and at the same time the operations to execute on trains, are replicated for each time interval of the schedule horizon T. Arcs are events representing the starting time of each operation, more precisely vertical arcs represent the end of a given operation in the time instant t and therefore the simultaneous begin of the following required operation, and thus the transfer of a train from one zone to another one. The horizontal arcs represent the temporal advancement of the operations i.e., from t to $t+1$, that is the time spent by a train in a given zone. Note that in some zones it is necessary to distinguish the available resources such as the tracks, while in other zones this is not required.

The operations-time-space network

Starting from the layout shown in Fig.1, let us define the operations time space network $G = (N, A)$, where N is the set of the nodes and A the set of the arcs. In the next, we will use only the term network to refer to this operations time space network. Note that the set of zones (operations) Z that we want to represent as nodes of G includes zones in which we have to distinguish the available resources used for performing operations on trains; for this purpose let us define:

Z^+ the set of operations for which we have to distinguish the available resources (i.e., the available tracks);

$Z^- = Z - Z^+$ the set of operations for which we have not to distinguish the available resources (in the considered layout, see Fig.1, $Z^- = \{z^2, z^4, z^5\}$);

$Z^C = Z^C(E) \cup Z^C(I)$ the set of couple of operations that can be executed in sequence, both on export trains ($Z^C(E)$) and on import ones ($Z^C(I)$), necessary to define transfer arcs (that permit the trains to pass from one operation to another, i.e. from one zone to another one, and to enter and leave the network).

$\mathcal{R}_\rangle$ the set of resources available for executing the operation i, $i \in Z^+$.

$\mathcal{N} = \cup_{\rangle \in \mathcal{Z}} \mathcal{N}^\rangle$ where :

$\mathcal{N}^\rangle = \{\setminus^\rangle_\sqcup | \sqcup \in \mathcal{T}\}$, $\forall i \in Z^-$ the set of nodes representing the operation i, in each t of the time horizon;

$\mathcal{N}^\rangle = \cup_{\triangledown \in \mathcal{R}_\rangle} \mathcal{N}^\rangle_\triangledown$, $\forall i \in Z^+$ with

$\mathcal{N}^\rangle_\triangledown = \{\setminus^\rangle_{\triangledown,\sqcup} | \sqcup \in \mathcal{T}\}$ the set of nodes representing the resource r available for operation i, $i \in Z^+$, $r \in \mathcal{R}_\rangle$, in each t of the time horizon.

$\mathcal{A} = (\cup_{\rangle \in \mathcal{Z}} \mathcal{A}^{\mathcal{N}^\rangle}) \cup (\cup_{\rangle,| \in \mathcal{Z}^\mathcal{C}} \mathcal{A}^{\mathcal{N}^\rangle,\mathcal{N}^|})$ the set of arcs of the network, given by the union of horizontal arcs ($\mathcal{A}^{\mathcal{N}^\rangle}$) and vertical ones ($\mathcal{A}^{\mathcal{N}^\rangle,\mathcal{N}^|}$), defined as in the following:

$\mathcal{A}^{\mathcal{N}^\rangle}$ the set of arcs $\{(n^i_t, n^i_{t+1}), \forall t \in T\}$, $\forall i \in Z^-$

$\mathcal{A}^{\mathcal{N}^\rangle} = \cup_{\triangledown \in \mathcal{R}_\rangle} \mathcal{A}^{\mathcal{N}^\rangle_\triangledown}$, $\forall i \in Z^+ - Z^6$ with

$\mathcal{A}^{\mathcal{N}^\rangle_\triangledown}$ the set of arcs $\{(n^i_{r,t}, n^i_{r,t+1}), \forall t \in T\}$, $\forall i \in Z^+ - Z^6, \forall r \in \mathcal{R}_\rangle$

Since terminals represent either origin or destination for the trains and we are not interested at representing the operations on trains inside the terminals, horizontal arcs connecting the terminals are not necessary.

For what concerns vertical arcs, they link couple of operations that can be executed in sequence and different definitions are required, depending on the kind of operations they refer to. In general form, let us denote:

$\mathcal{A}^{\mathcal{N}^\rangle,\mathcal{N}^|}$ the set of vertical arcs $\forall i, j \in Z^C$

Note that $\mathcal{A}^{\mathcal{N}^\rangle}_\sqcup$ and $\mathcal{A}^{\mathcal{N}^\rangle,\mathcal{N}^|}_\sqcup$ refer to the set of horizontal and vertical arcs related to time period t, $\forall t \in T$;

$\mathcal{A}^{\mathcal{N}^\rangle,+}_\sqcup$ and $\mathcal{A}^{-,\mathcal{N}^\rangle}_\sqcup$ refer to the subset of vertical outbound/inbound arcs of node n^i_t in t and depend on the compatible operations;

In next section we introduce the flow formulation used to solve PRSSP.

4.1 The network flow model

Each export train that has to be transferred from the rail station to the maritime terminal represents a unit of flow that enter in a given time instant and has to reach its destination terminal within a given time window. The vice-versa is for import trains. Note that while managing trains some capacity constraints must be satisfied.

Let us now introduce the additional useful notation for the flow formulation. The model based on the network described above has the flow decision variables and auxiliary variables for computing the train's waiting time in the network (to be minimized); let be:

$x_{a,j} \in \{0,1\}, \forall j \in J, \forall a \in A$, $x_{a,j} = 1$ if arc a is used for train j;

$y_j^i \geq 0, \forall j \in J, \forall i \in \{z^1, z^3\}$, define the total time spent by a train j in the rail station or in the shunting park, waiting either for its departure or for starting the shunting operations.

The resulting model is the following:

$$MIN \sum_{j \in J} \sum_{i \in \{z^1, z^3\}} y_j^i \tag{1}$$

subject to:

$$\sum_{a \in \mathcal{A}^{\mathcal{N}\rangle}} x_{a,j} = y_j^i \quad \forall j \in J, \forall i \in \{z^1, z^3\} \tag{2}$$

$$\sum_{a \in \mathcal{A}_{\rceil|}^{\mathcal{N}\ddagger', \mathcal{N}\ddagger\infty}} x_{a,j} = 1 \quad \forall j \in E \tag{3}$$

$$\sum_{a \in \mathcal{A}_{\rceil|}^{\mathcal{N}\ddagger\infty, \mathcal{N}\ddagger'}} x_{a,j} = 1 \quad \forall j \in I \tag{4}$$

$$\sum_{a \in (\cup_{e_j^{min} \leq t \leq e_j^{max}} \mathcal{A}_{\sqcup}^{-, \mathcal{N}\ddagger/_{\nabla}})} x_{a,j} = 1 \quad \forall j \in E, r = p_j \tag{5}$$

$$\sum_{a \in (\cup_{e_j^{min} \leq t \leq e_j^{max}} \mathcal{A}_{\sqcup}^{\mathcal{N}\ddagger/_{\nabla}, +})} x_{a,j} = 1 \quad \forall j \in I, r = p_j \tag{6}$$

$$\sum_{j\in J}\sum_{a\in\mathcal{A}_{\sqcup}^{-,\mathcal{N}_{\nabla}^{\ddagger/}}\cup\mathcal{A}_{\sqcup}^{\mathcal{N}_{\nabla}^{\ddagger/},+}} x_{a,j}\le 1 \quad \forall t\in T,\forall r\in\mathcal{R}_{\ddagger/} \tag{7}$$

$$\sum_{a\in\mathcal{A}_{\sqcup-\infty}^{\mathcal{N}^{\rangle}}\cup\mathcal{A}_{\sqcup}^{-,\mathcal{N}^{\rangle}}} x_{a,j} = \sum_{a\in\mathcal{A}_{\sqcup}^{\mathcal{N}^{\rangle}}\cup\mathcal{A}_{\sqcup}^{\mathcal{N}^{\rangle},+}} x_{a,j} \quad \forall j\in J,\forall i\in Z_j,\forall t\in T \tag{8}$$

$$\sum_{j\in J}\sum_{a\in\mathcal{A}_{\sqcup}^{\mathcal{N}_{\nabla}^{\rangle}}\cup\mathcal{A}_{\sqcup}^{\mathcal{N}_{\nabla}^{\rangle},+}} x_{a,j}\le 1 \quad \forall t\in T,\forall i\in\{z^1,z^3\},\forall r\in\mathcal{R}_{\rangle} \tag{9}$$

$$\sum_{j\in J}\sum_{a\in\mathcal{A}_{\sqcup}^{\mathcal{N}^{\rangle}}\cup\mathcal{A}_{\sqcup}^{\mathcal{N}^{\rangle},+}} x_{a,j}\le u_t^i \quad \forall t\in T,\forall i\in\ Z^- \tag{10}$$

$$\sum_{j\in J}\sum_{a\in\cup_{i\in Z^-}(\mathcal{A}_{\sqcup}^{\mathcal{N}^{\rangle}}\cup\mathcal{A}_{\sqcup}^{\mathcal{N}^{\rangle},+})} x_{a,j}\le k_t \quad \forall t\in T \tag{11}$$

$$\sum_{a\in\mathcal{A}^{\mathcal{N}^{\rangle}}} x_{a,j}=d_{ij} \quad \forall j\in J,\forall i\in Z_j \tag{12}$$

The objective function (1) minimizes the sum of the trains' waiting time both in the shunting park and in the rail station. Constraints (2) define the auxiliary variables $y_j^{z^1}$ and $y_j^{z^3}$. Constraints (3) and (4) impose that for each export/import train there must be one variable representing the entrance/exit of the train in the network equal to 1 exactly at time $t=e_j$.

Constraints (5) and (6) are related to both the entrance and the exit of the trains in/from the their destination/origin terminals; they permit to satisfy the entry/exit terminals time windows ($e_j^{min}\le t\le e_j^{max}$).

Constraints (7) are used to impose that there must be at maximum one shunting operation (considering both entry and exit operations) for each time instant involving the same terminal.

The flow conservation constraints (8) operate for each unit of flow on the network, i.e. for each train j, for each time instant t verifying each node visited by the train ($i\in Z_j$).

Constraints (9),(10) and (11) impose the capacity restrictions. In particular, the capacity of tracks is considered and thanks to constraints (9) it is imposed that at most one train can wait on each station's track and on each shunting park's tracks in each time instant t. Constraints (10) limit the number of shunting operations that can be executed simultaneously in each zone of the port when resources are not explicitly considered ($i \in Z^-$). Finally, the maximum number of simultaneous shunting operations in the whole port area (more precisely in zone z^2, z^4 and z^5) is limited to the number of available shunting teams in each time instant t as imposed by constraints (11).

Constraints (12) fix the processing time of each shunting operation that each train has to execute to the required time d_{ij}.

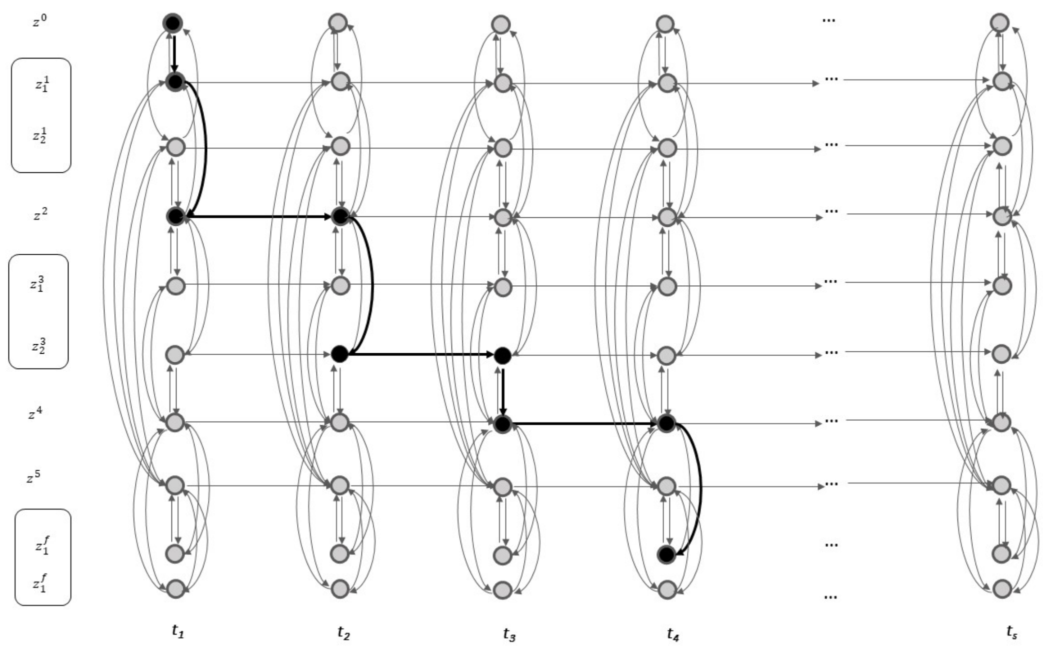

Figure 2: The network

Fig. 2 shows the network used in the present work for PRSSP in which only two tracks and two terminals are represented. In black is reported the path of an export train arriving in the first track of the station, starting immediately the primary operation, reaching and waiting in the track number two of the park, performing the secondary operation and arriving at its destination terminal number one.

4.2 Reducing the number of variables

Note that it is possible to reduce the number of flow variables by taking into consideration the time limitations imposed to trains. As far as the export cycle is considered, it is possible to use $e_j \leq t \leq e_j^{max}$ to generate the related variables only between the time of arrival of train j and the maximum time in which it has to be at destination. In particular, $t \leq e_j^{max}$ only for variables representing the end of either the secondary or unique operations, bringing trains to their destination and $t < e_j^{max}$ for the variables representing the start of the operations.

The same reasoning can be done for the import cycle.

4.3 Model extension: tight constraints vs soft constraints

The model presented above can support the shunting manager in planning the schedule of import and export trains. Moreover, for the shunting manager, it should be useful to discover bottlenecks of the port area to the expansion in terms of the number of trains to manage.

Strong constraints to this growth are due to the national railway network and its schedules, together with the layout of the port area (i.e., the number of tracks in the station and the shunting park); other constraints are related to the organization of the port area and fix a limit to the number of operations that can be executed in each period of the time horizon; the time windows imposed by the terminals could represent another limitation to the growth of the managed trains.

The model presented above can be modified by replacing strong time window constraints and capacity constraints with soft ones. Moreover, it is possible to permit not to serve a train. Note that the penalization of the deviations must be included in the objective function. In the following, the new constraints are described together with the new objective function.

Time Windows constraints vs time window deviation

Constraints (5) and (6) are related to both the entrance and the exit of the trains in/from the their destination/origin terminals within the given terminals time windows. If it is permitted to an export train j to enter the terminal either before e_j^{min} or later than e_j^{max} (an import train j to leave the terminal before e_j^{min} or later than e_j^{max}), it is necessary to compute the distance of the entry(exit) from its required time window, in such a way to be able to minimize all these deviations. Let be f_j the deviation of the train j with respect to its original time window.

Constraints (5) and (6) are re-written considering larger time windows and the following new constraints are required for computing the deviation for each train.

$$\sum_{a \in (\cup_{en_j^{min} \leq t \leq en_j^{max}} \mathcal{A}_{\sqcup}^{-,\mathcal{N}^{\ddagger/}_{\nabla}})} tx_{a,j} - e_j^{max} \leq f_j \quad \forall j \in E, r = p_j \tag{13}$$

$$e_j^{min} - \sum_{a \in (\cup_{en_j^{min} \leq t \leq en_j^{max}} \mathcal{A}_{\sqcup}^{-,\mathcal{N}_{\nabla}^{\ddagger/}})} t x_{a,j} \leq f_j \quad \forall j \in E, r = p_j \tag{14}$$

$$\sum_{a \in (\cup_{en_j^{min} \leq t \leq en_j^{max}} \mathcal{A}_{\sqcup}^{\mathcal{N}_{\nabla}^{\ddagger/},+})} t x_{a,j} - e_j^{max} \leq f_j \quad \forall j \in I, r = p_j \tag{15}$$

$$e_j^{min} - \sum_{a \in (\cup_{en_j^{min} \leq t \leq en_j^{max}} \mathcal{A}_{\sqcup}^{\mathcal{N}_{\nabla}^{\ddagger/},+})} t x_{a,j} \leq f_j \quad \forall j \in I, r = p_j \tag{16}$$

Strong capacity constraints vs soft ones

Capacity constraints (10) and (11) can be modified in such a way to permit a larger use of the limited resources. For each time instant the number of additional required shunting teams is computed in (17) thanks gk_t. For each time instant and for each operation the number of additional resource used is computed in (18) thanks variables gu_t^i.

The new constraints replacing (10) and (11) are the following:

$$\sum_{j \in J} \sum_{a \in \cup_{i \in Z^-} (\mathcal{A}_{\sqcup}^{\mathcal{N}\rangle} \cup \mathcal{A}_{\sqcup}^{\mathcal{N}\rangle,+})} x_{a,j} \leq k_t + gk_t \quad \forall t \in T \tag{17}$$

$$\sum_{j \in J} \sum_{a \in \mathcal{A}_{\sqcup}^{\mathcal{N}\rangle} \cup \mathcal{A}_{\sqcup}^{\mathcal{N}\rangle,+}} x_{a,j} \leq u_t^i + gu_t^i \quad \forall t \in T, \forall i \in \ Z^- \tag{18}$$

Possible exclusion of trains

Constraints (3) and (4) impose to each export/import train to enter/leave the network at time $t = e_j$. If it is possible not to serve a train, it is enough to modify these constraints by adding a binary variable g_j assuming value one when a train is not served (see constraints (19) and (20)).

Note that also constraints (12) must be replaced by the following (21) which fix the processing time of every shunting operation of each served train to the required time d_{ij}.

$$\sum_{a \in \mathcal{A}_{\rceil|}^{\mathcal{N}\ddagger',\mathcal{N}\ddagger\infty}} x_{a,j} = 1 - g_j \quad \forall j \in E \tag{19}$$

$$\sum_{a \in \mathcal{A}^{\mathcal{N}\ddagger^{\infty},\mathcal{N}\ddagger'}_{\rceil|}} x_{a,j} = 1 - g_j \quad \forall j \in I \tag{20}$$

$$\sum_{a \in \mathcal{A}^{\mathcal{N}\rangle}} x_{a,j} = d_{ij}(1 - g_j) \quad \forall j \in J, \forall i \in Z_j \tag{21}$$

The new objective function that includes the penalisation of the above defined deviations is the following:

$$MIN \sum_{j \in J} \sum_{i \in \{z^1, z^3\}} y^i_j + \alpha_1 \sum_{j \in J} f_j + \alpha_2 \sum_{t \in T} gk_t + \alpha_3 \sum_{t \in T} \sum_{i \in Z^-} gu^i_t + \alpha_4 \sum_{j \in J} g_j \tag{22}$$

where $\alpha_1, \alpha_2, ...\alpha_4$ are the penalties used for the deviations.

5 Numerical experiments and analyzes

In this section, some computational tests for validating both the proposed model and the quality of the obtained solutions are presented. The validation regards model (1)-(12) modified as described in Section 4.3. In the following we will refer to it as Mod1.

Mod1 has been implemented in Python 3.7, and solved by commercial solver Gurobi 8.1.0, on a machine intel (R), i5, 7200U CPU, 2.5 GHz, 8.00 GB RAM.

All the tests are based on a port system characterized by a rail station, a shunting park, and four terminals. Two terminals are directly connected to the station, while the others are connected to the shunting park. The rail station and the shunting park have ten tracks each. The generated number of trains to manage ranges from 50 to 100 and is here equally distributed between import and export trains. The time horizon is fixed to six days, discretized into time intervals of ten minutes.

The random instances have been generated varying the arrival distribution for the trains along the considered time horizon. Two types of arrival distribution are analyzed: a) a homogeneous distribution in which the generated number of trains is constant for specific time intervals within the time horizon; b) a compact distribution in which the generated number of trains is distributed in a specific time interval by splitting it into four slots and assigning the 50% of the total number of trains to the first slot, the 25% in the second and the third.

For the homogeneous distribution, the following time intervals have been considered: i) two days (i.e. 3-time intervals of two days in the six days time horizon, called **H2**);

ii) one day (i.e. 6-time intervals of one day in the six days time horizon, called **H1**); iii) one working shift (i.e. three shifts per day, thus 18-time intervals in the six days time horizon, called **Hs**). The time interval considered for the compact distribution is two days (called **Co**), thus the first slot with the greatest number of trains is composed of shifts 1 and 2 of day 1.

The arrival and departure time for each train within the time intervals is randomly assigned while respecting the distributions explained above.

Each train has a terminal as origin/destination and is characterized by the types of shunting operations to perform. Different scenarios concerning these elements characterizing the trains have been considered.

Concerning the shunting zones (i.e., the shunting operations) the distribution of trains can be balanced (B), unbalanced (U), or strongly unbalanced (S_U). A distribution is balanced (B) if the partition of trains between the operations ranges from 45-55 (unique operation)/55-45 (others operations); it is unbalanced (U) if the ranges are 25-35/75-65 and strongly unbalanced (S_U) with ranges as 15-20/85-80.

The distribution of trains among the terminals regards the trains having the same types of operations to execute, and can be either balanced (B) or strongly unbalanced (S_U). For instance, the strongly unbalanced distribution among the terminals means that the distribution of the trains having to perform the unique operations is unbalanced between the two terminals that can be reached by the station (passing through the unique zone). A distribution is balanced (B) if the split of trains between the terminals ranges from 40-60/60-40, while in the case of strongly unbalanced (S_U) the range varies between 15-25/85-75.

Note that, the combination of these three kinds of distributions for operations (B, U, S_U) and two kinds of distributions for terminals (B, S_U), provides six different sets of instances. Table 1 shows the obtained six sets with their trains distribution between both the shunting operations (Unique and Others) and the terminals (Ter 1/2/3/4).

	Unique	Others	Ter 1	Ter 2	Ter 3	Ter 4
$S_1(B-B)$	48.00	52.00	50.00	50.00	46.15	53.85
$S_2(B-S_U)$	48.00	52.00	20.83	79.17	18.85	81.15
$S_3(U-B)$	30.20	69.80	47.01	52.99	48.42	51.58
$S_4(U-S_U)$	29.40	70.60	21.88	78.12	20.96	79.04
$S_5(S_U-B)$	15.40	84.60	41.69	58.31	49.18	50.82
$S_6(S_U-S_U)$	15.00	85.00	24.08	75.92	23.05	76.95

Table 1: Trains distribution (in %) among operations & terminals

The following subsections report three different experimental campaigns, which aims are here below introduced:

- first campaign: evaluation of the difficulty in solving instances characterized by a given number of trains and different arrival distribution of the trains during the time horizon. Moreover, the evaluation also regards different distribution of the trains with respect to the operations to perform and their origin/destination terminals;

- second campaign: evaluation of the capability of model Mod1 to solve bigger instances, in terms number of trains to manage, and evaluation of the capacity of the port to manage a higher number of trains stressing possible bottlenecks;

-third campaign: showing the capability of the proposed model to be an useful tool for both scheduling a given number of trains in an operative fix contest, and evaluating how to change the operative contest (i.e., in terms of required resources and new agreements to define).

5.1 First campaign

The first campaign aims at evaluating the impact of different trains distributions concerning their arrival during the time horizon (that is, H2, H1, Hs and Co), and their origin/destination terminals together with their operations to perform (that is, S_1, S_2,...,S_6). For this experimental tests instances used are characterized by 50 trains (50% are import trains and 50% export ones).

The results obtained by solving Mod1 are reported in the following tables; note that, each row gives the average of five solved instances. Each set of the generated instances reported in Table 2, is characterized by low standard deviations, as shown in Table 8 in the Appendix.

Table 2 focuses on the size of the solved model and shows the number of variables (**#V**), the number of constraints (**#C**), the CPU times expressed in seconds (**CPU**), the objective function values (**Obj**) and the optimality gap in percentage (**Gap**).

Mod1 has been solved with a time limit of 1 hour. After a tuning phase for choosing the parameters for the penalties, the following weights have been used in (22): 10 for the time window deviation, 100 for the deviation from both the operations in the zone, and the maximum number of teams and 500 for the trains exclusion.

Looking at Table 2 we can note that all instances have been solved up to optimality, and the CPU time is always smaller than half an hour. The number of variables ranges between 306591 and 504285, while the number of constraints is between 151358 and 229418.

Table 3 shows the characteristics of the obtained solutions and reports some data related to the usage of the port resources: the total waiting time of the trains (**Tot Wait**) together with the percentage of the time spent in waiting operations (**% Wa**)

Distr	Set	#V	#C	CPU	Obj	Gap
	S_1	390606	188024	789.92	2932.00	0.00
	S_2	387736	186646	846.15	3059.00	0.00
H1	S_3	446129	207423	1258.38	2861.00	0.00
	S_4	444271	206766	901.05	2954.20	0.00
	S_5	486868	222193	1412.93	3138.40	0.00
	S_6	504285	229418	1465.21	3347.80	0.00
	S_1	354611	171955	804.08	1712.40	0.00
	S_2	359878	173882	789.77	1763.80	0.00
H1	S_3	403805	189302	844.68	1681.20	0.00
	S_4	405832	190112	795.91	1713.60	0.00
	S_5	447526	205152	1272.23	1738.80	0.00
	S_6	449839	205771	899.57	1709.20	0.00
s	S_1	308125	151924	82.93	370.40	0.00
	S_2	306501	151358	139.15	378.00	0.00
H	S_3	354071	168285	245.93	380.60	0.00
	S_4	356634	169203	188.39	377.60	0.00
	S_5	392839	182081	303.95	381.80	0.00
	S_6	395627	183082	402.69	383.20	0.00
o	S_1	358724	173781	873.89	2347.00	0.00
	S_2	361196	175154	867.58	2461.80	0.00
C	S_3	410611	192784	760.10	2516.00	0.00
	S_4	414603	194173	1196.64	2463.00	0.00
	S_5	462528	211471	913.33	2535.80	0.00
	S_6	444398	204006	1263.60	2314.00	0.00

Table 2: Computational results obtained by solving Mod1

and the relative occupancy of tracks (**% Oc**) in the station (St) and in the park (Pa). Moreover, for each train is reported the average time spent in waiting operations (**Avg/t**) in the station (St) and in the park (Pa); the last four columns refer to the deviation from the limited shunting teams (**Te**) and resources (**Zo**), from the not-served trains (**Ex**) and from the time windows (**TW**).

Remember that, in Mod1 the objective function (22) penalizes both the over capacity requests (the number of shunting teams and the number of operations to execute in the different zones) and the violated time windows and the not served trains.

The waiting time of the trains is generally spent for the most part in the track of the station. Anyway, the % of occupancy of the tracks in the rail station is slow, and range from 0.3 to 33.9%. The percentages of occupancy are lower for the tracks of the shunting park (from 0.1 to 17.9%).

Note that only in few cases the obtained solutions present a deviation with respect to

Distr	**Set**	**Tot Wait**	**%Wa** (St)	**%Wa** (Pa)	**%Oc** (St)	**%Oc** (Pa)	**Avg/t** (St)	**Avg/t** (Pa)	**Te** dev	**Zo** dev	**Ex** dev	**TW** dev
H2	S_1	2585.40	76.03	23.97	33.27	10.91	393.16	123.92	0.00	0.00	0	2.00
	S_2	2710.60	72.15	27.85	33.19	12.86	391.16	150.96	0.00	0.00	0	4.00
	S_3	2516.60	63.93	36.07	27.38	15.20	321.76	181.56	0.00	0.00	0	2.00
	S_4	2612.40	65.68	34.32	29.20	15.43	343.16	179.32	0.00	0.00	0	0.00
	S_5	2797.80	63.44	36.56	30.49	17.88	355.00	204.56	0.00	0.00	0	0.00
	S_6	3006.00	65.48	34.52	33.89	17.73	393.64	207.56	0.00	0.00	0	0.00
H1	S_1	1367.80	70.79	29.21	16.36	6.83	193.64	79.92	0.00	0.00	0	2.00
	S_2	1418.40	67.41	32.59	16.02	7.97	191.24	92.44	0.00	0.00	0	2.00
	S_3	1336.60	56.58	43.42	12.81	9.80	151.24	116.08	0.00	0.00	0	2.00
	S_4	1371.40	63.96	36.04	14.49	8.47	175.44	98.84	0.00	0.00	0	0.00
	S_5	1396.00	51.00	49.00	12.13	11.49	142.40	136.80	0.00	0.00	0	0.00
	S_6	1367.20	49.03	50.97	11.30	11.90	134.08	139.36	0.00	0.00	0	0.00
Hs	S_1	26.80	70.90	29.10	0.32	0.12	3.80	1.56	0.00	0.00	0	0.00
	S_2	35.20	76.14	23.86	0.42	0.13	5.36	1.68	0.00	0.00	0	0.00
	S_3	38.00	86.32	13.68	0.54	0.08	6.56	1.04	0.00	0.00	0	0.00
	S_4	35.00	80.00	20.00	0.45	0.11	5.60	1.40	0.00	0.00	0	0.00
	S_5	40.80	85.78	14.22	0.56	0.10	7.00	1.16	0.00	0.00	0	0.00
	S_6	41.20	81.55	18.45	0.56	0.13	6.72	1.52	0.00	0.00	0	0.00
Co	S_1	2002.00	69.91	30.09	23.89	10.20	279.92	120.48	0.00	0.00	0	0.00
	S_2	2116.80	74.57	25.43	26.15	8.86	315.68	107.68	0.00	0.00	0	0.00
	S_3	2170.80	69.52	30.48	25.66	11.48	301.84	132.32	0.00	0.00	0	2.00
	S_4	2119.20	61.63	38.37	22.17	13.81	261.20	162.64	0.00	0.00	0	0.00
	S_5	2193.00	65.18	34.82	23.90	12.77	285.88	152.72	0.00	0.00	0	0.00
	S_6	1969.80	62.60	37.40	21.06	12.43	246.60	147.36	0.00	0.00	0	0.00

Table 3: Computational results – characteristics of the solutions obtained by solving Mod1

the time windows; anyway, the deviation is limited to few minutes. No other deviations are present.

Looking at Tables 2 and 3 it is clear that the arrival distributions have a great impact on both the model size, the CPU time and the quality of the obtained solutions. In fact, for instances characterized by the arrival distribution H2, the total waiting time spent by the trains in the tracks of the port is the highest one, there is a little violation of the time windows and also the corresponding CPU time is the highest one (1112 seconds on average). The opposite case is represented by the set with Hs as arrival distribution, in which trains are homogeneously distributed in each shift of each working day. The total waiting time is negligible.

The following graphs permit to better understand these relations.

In Fig. 3 the number of variables and constrains, CPU times and objective function values are reported. In case of instances characterized by H2 (Homo2day) scheduling all the activities is more complex. In fact, the highest values reported in the graph are related to H2, while the lowest are related to Hs (Homo1shift). The gap percentages between the highest and the lowest values of variables and constraints is around 20%, while for CPU and objective function is around 85%.

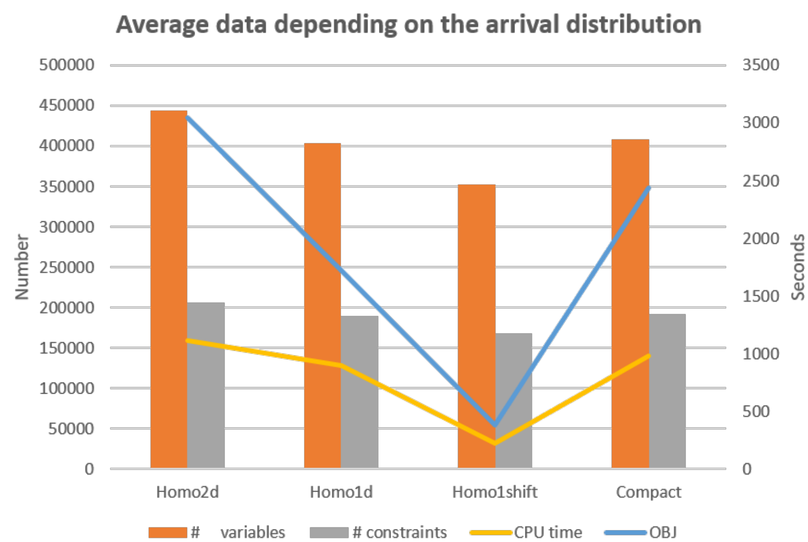

Figure 3: Variables, constrains, CPU time and objective function values for different arrival distributions

The graph in Fig.4 reports the average percentages of both waiting times and tracks occupancy in the rail station and in the shunting park. Looking to the tracks occupancy, the worst case is again for H2 (Homo2d), while the best results have been obtained for Hs (Homo1shift).

Generally, the time passed in the rail station is higher than the time spent in the shunting park for all distributions.

Finally, the impact of the distribution of the trains inside the port area is analysed. The graph in Fig. 5 shows the average number of variables and constraints, CPU time and objective function values for each set of Table 1.

Passing from a balanced situation (S_1) to strongly unbalanced one (S_6) the values of the objective function present a slow increasing trend, while the other values increase on average of about 25%. This means that more unbalanced is the distribution of trains, more difficult is the trains management and more congestioned is the system. Anyway, the developed model is able to manage all the trains in these different cases.

Thanks to graph reported in Fig. 6, it is possible to note that from S_1 to S_6, the average percentages of tracks occupancy increase for the shunting park (33% between S_1 and S_6) while decrease for the rail station (12% between S_2 and S_3). In all cases the time

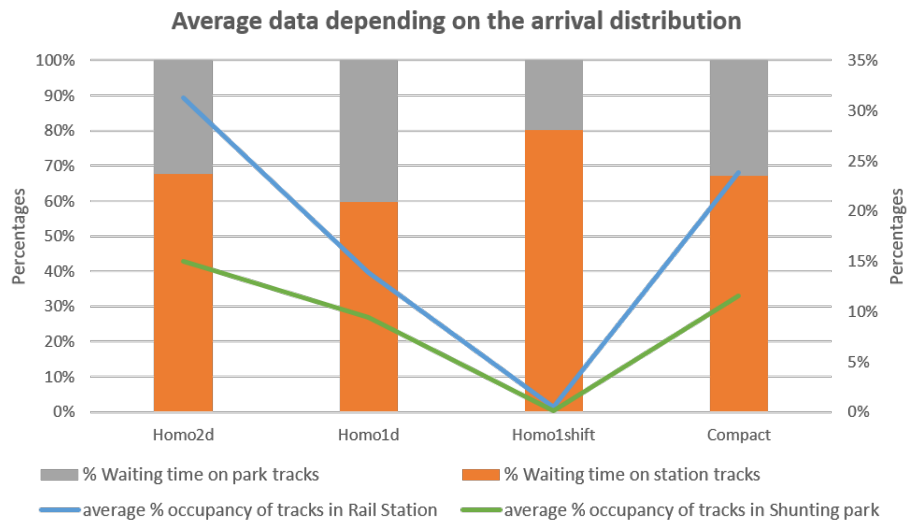

Figure 4: % of waiting times and tracks occupancy for different arrival distributions

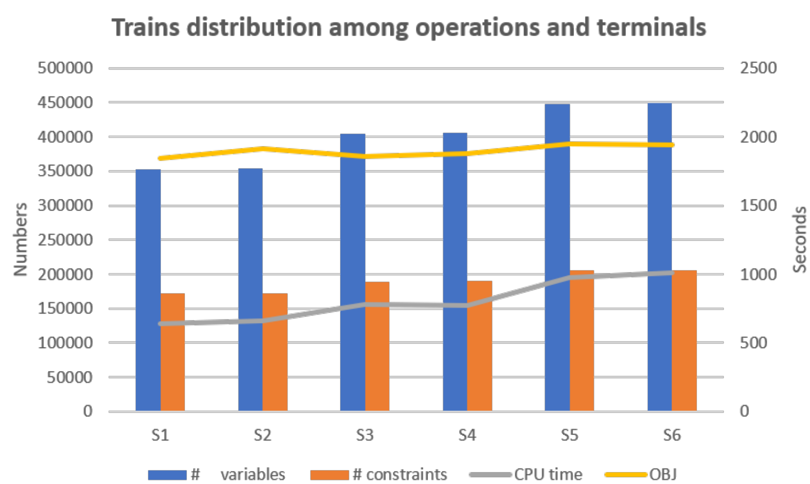

Figure 5: Variables, constrains, CPU time and objective function for different arrival distributions

spent in the railway station is higher than the time spent in the park.

The last graph (see Fig. 7) shows that the highest total waiting time is for sets S_5 and S_6 (i.e. the strongly unbalanced sets), while the lowest is for the balanced set S_1 and the balanced at the terminals S_3. The gap between the highest and the lowest values is of about 7%. The average time spent by a train in the station is almost the same in all sets, excepted for S1 and S2 that present higher values for the station than for the park. The time spent in the shunting park increases passing from S_1 to S_6 of about 34%.

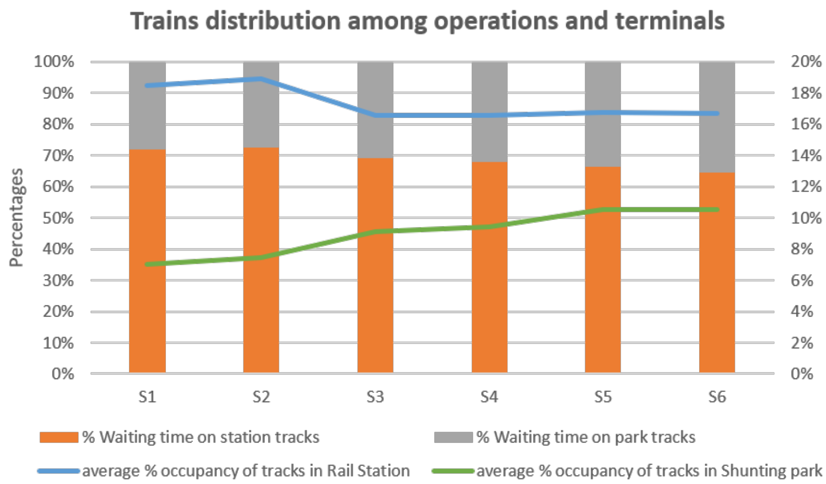

Figure 6: % of waiting times and tracks occupancy for different arrival distributions

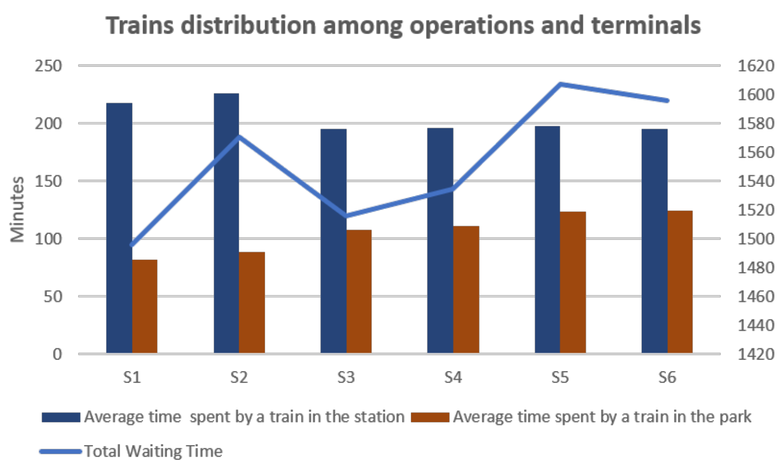

Figure 7: Waiting times analysis for different arrival distributions

5.2 Second campaign

The second experimental campaign focuses on the evaluation of the capability of Mod1 to solve bigger instances, in terms of number of trains to manage, that grows from 50 to 100.

The tests are based on instances characterized by the homogeneous distribution among each day (H1), that is the more realistic one in the contest we are involved with, and inside the port area two scenarios are considered. Trains may be: i) homogeneously distributed among both the shunting zones and the terminals (S_1); ii) unbalanced with respect to the shunting zones and strongly unbalanced with respect to the terminals (S_4).

The obtained results are reported in Tables 4 and 5. As before, each row gives the average of five solved instances and the standard deviations of the sets of instances reported in Table 4 are shown in Table 9 in the Appendix.

Table 4, as Table 2, focuses on the size of the solved model and shows the number of variables (**#V**), the number of constraints (**#C**), the CPU times expressed in seconds (**CPU**), the objective function values (**Obj**) and the optimality gap in percentage (**Gap**).

The number of variables ranges between 354611 to 759987, while the number of constraints from 171955 to 357939. The CPU time is about 13 minutes for instances with 50 trains, while the time limit of 1 hour is reached when solving 70-trains instances.

Set	**#Trains**	**#V**	**#C**	**CPU**	**Obj**	**Gap**
S_1	50	354611	171955	804.08	1712.40	0.00
S_1	60	422632	202275	2114.09	2140.60	0.04
S_1	70	492659	233076	3600.57	3033.60	0.23
S_1	80	566292	265676	3601.07	4063.20	0.30
S_1	90	644110	299473	3600.56	6840.40	0.54
S_1	100	706089	326557	3600.81	11424.40	0.66
Avg	75.00	531065	249835	2886.87	4869.10	0.30
S_4	50	405832	190112	795.91	1713.60	0.00
S_4	60	483612	223949	2278.98	2155.40	0.04
S_4	70	565643	259326	3601.15	3294.00	0.22
S_4	80	641940	292360	3600.96	4020.80	0.32
S_4	90	729953	329874	3600.37	6211.40	0.47
S_4	100	795987	357939	3600.27	12650.00	0.72
Avg	75.00	603828	275593	2912.94	5007.53	0.29

Table 4: Results obtained by solving Mod1

The quality of the solutions can be analyzed by looking at Table 5 that reports the same information showed in Table 3.

The trains spend more time in the station than in the park. In case of S1, on average about 68% of the waiting time is related to the station and about 32% to the shunting park, with an average tracks' occupancy of 22% and 10.6% respectively for the station's tracks and park's tracks. The situation is similar but more balanced for sets S4. Sets S4 are characterized by higher average time spent by each train in the park and lower average time spent in the station than sets S1.

Finally, the obtained solutions present some small deviations for what concerns the number of needed shunting teams and no deviation for the zones. When increasing of the number of trains, for both sets (S1 and S4), it is easy to note an increasing number of excluded trains, from 0 to around 19, and an increasing TW deviation, that on average

Distr	Set	Tot Wait	%Wa (St)	%Wa (Pa)	%Oc (St)	%Oc (Pa)	Avg/t (St)	Avg/t (Pa)	Te dev	Zo dev	Ex dev	TW dev
S_1	50	1367.80	70.79	29.21	16.36	6.83	193.64	79.92	0.00	0.00	0.00	2.00
S_1	60	1599.40	74.87	25.13	19.98	6.59	199.57	67.00	0.00	0.00	0.20	0.00
S_1	70	1809.20	69.12	30.88	21.02	9.25	178.66	79.80	0.00	0.00	1.40	16.00
S_1	80	2148.60	66.35	33.65	24.42	12.08	178.20	90.38	0.00	0.00	2.60	42.00
S_1	90	2280.80	64.71	35.29	25.05	13.73	163.98	89.44	0.00	0.00	7.20	344.00
S_1	100	2355.20	63.13	36.87	25.36	14.89	148.68	86.84	2.00	0.00	16.40	228.00
Avg	75.0	1926.83	68.16	31.84	22.03	10.56	177.12	82.23	0.33	0.00	4.63	105.33
S_4	50	1371.40	63.96	36.04	14.49	8.47	175.44	98.84	0.00	0.00	0.00	0.00
S_4	60	1603.60	60.16	39.84	15.97	11.03	160.80	106.47	0.00	0.00	0.20	0.00
S_4	70	1880.00	56.06	43.94	17.67	13.69	150.57	118.00	0.00	0.00	1.60	98.00
S_4	80	2102.80	53.57	46.43	18.86	16.27	140.80	122.05	0.00	0.00	2.40	132.00
S_4	90	2267.20	55.07	44.93	20.94	17.35	138.73	113.18	0.00	0.00	6.20	210.00
S_4	100	2107.60	50.68	49.32	18.24	17.23	106.82	103.94	6.00	0.00	19.40	174.00
Avg	75.00	1888.77	56.59	43.41	17.70	14.01	145.53	110.41	1.00	0.00	4.97	102.33

Table 5: Analysis of the quality of the solutions

is about 103 minutes.

Thanks to the following graphs it is possible to understand how increasing the number of trains, the difficulty in solving Mod1 increases and the management of the scheduling of the trains too.

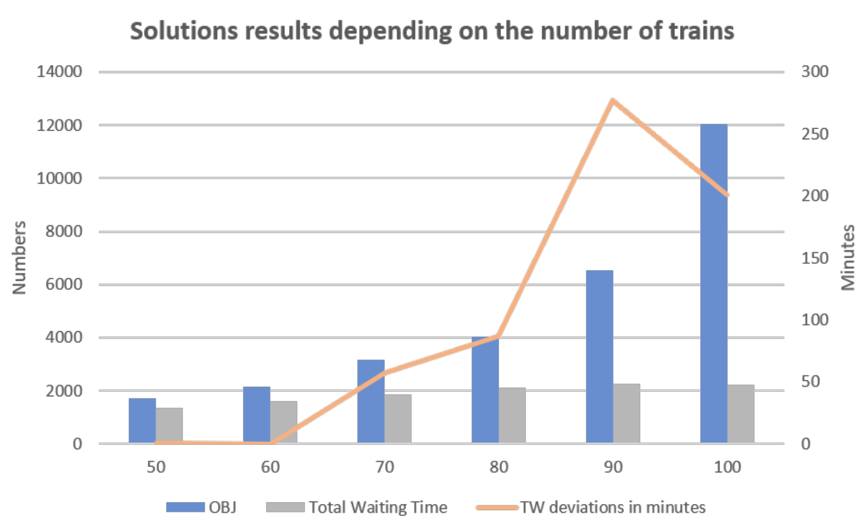

Figure 8: Objective function values, waiting times and TW deviations varying the number of trains to manage

From the graph in Fig. 8, it is possible to note that the enormous growth in the

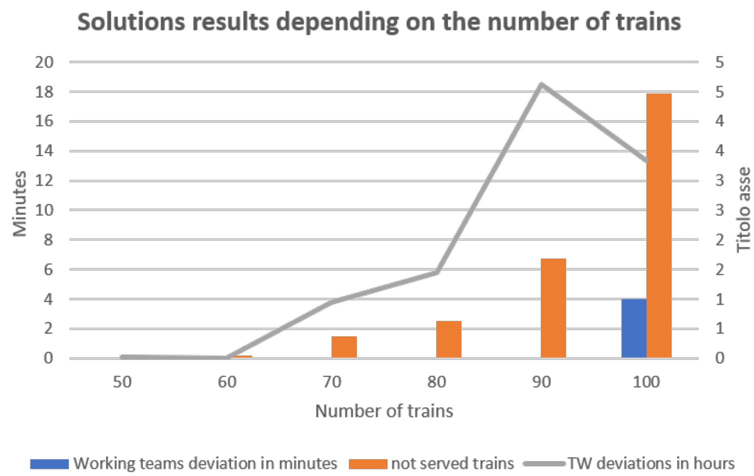

Figure 9: Working teams deviations, excluded trains and TW deviations varying the number of trains to manage

objective function value (86%) is due to the deviations that are weighted, while the total waiting time of served trains remain more or less the same. In fact, the time window deviations in minutes increases with a gap of 57% between 80 and 100 trains.

Thanks to the graph in Fig. 9 it is possible to analyze the obtained solutions from a management point of view. In fact we can use the model for tow different evaluations: the maximum number of trains that we are able to serve with the available resources and, vice versa, the required resources for serving a given number of trains.

The layout structure and the available human resources seem to be not enough for serving more than 80 trains. In fact, the number of not served trains grows between 70 and 100 trains. In this latter case, there is also a major request in terms of human resources. The time windows deviation (same trend as in the previous graph) does not represent a real problem to highlight because time windows have been randomly generated, and thus they could be not perfectly realistic generated and some deviations has expected.

Finally, thanks to the proposed model, this campaign also reveals the capacity of the port to manage a higher number of trains, thanks to the station and the shunting park capacity expressed in term of number of tracks (ten tracks each).

5.3 Third campaign

This last campaign aims at showing the capability of the proposed model to be an useful tool for the shunting manager. The experiments focus on instances of the group H1-S1-80 trains. Note that, this analysis has also been used to fix the weights in the objective function (22).

The time limit in the previous experimental campaign has been fixed to one hour but in the preliminary tests also three hours as time limit has been used.

Moreover, a lot of tests have been made in order to fix the weights for the penalization described in Section 4.3. The weights used in the previous campaigns, as already said, are $\alpha_1 = 10, \alpha_2 = \alpha_3 = 100, \alpha_4 = 500$. Some extreme cases, useful for the aim of this campaign, are reported in this section. In particular, the following Tables report results related to the penalization listed below:

1a) $\alpha_1 = 0$, i.e., time window deviations are not penalized;

1b) $\alpha_1 = \alpha_2 = 0$, i.e., the number of the used shunting teams that exceed the number of the available and the time window deviations are not penalized;

1c) $\alpha_1 = \alpha_2 = \alpha_3 = 0$, i.e., only the train exclusion is penalized.

All the obtained results have been reported in Tables 6 and 7, in such a way to compare the obtained solutions in terms of CPU time and quality.

In particular, Table 6 includes the following information: test name, the CPU time, the objective function value and the optimality gap in percentage. The three following columns are related to the waiting time of the trains, and in particular, the total waiting time and percentage of waiting time in both the rail station and the shunting park are reported. The last four columns are related to the tracks occupancy: the percentages of occupancy at the station and at the park is shown together with the average waiting time per train (at the station and at the park).

Table 7 includes the deviations obtained in each test, as in Tables 3 and 5, but more details related to the time windows are reported due to their high deviation values. Thus, the last four columns show the time window deviations both in minutes and in hours, and the time window deviation per train, again in minutes and hours.

Test	**CPU**	**Obj**	**Gap** %	**Tot Wait**	**%Wa** (St)	**%Wa** (Pa)	**%Oc** (St)	**%Oc** (Pa)	**Avg/t** (St)	**Avg/t** (Pa)
Mod1-1h	3601.07	4063.20	0.30	2148.60	66.35	33.65	24.42	12.08	178.20	90.38
Mod1-3h	10800.59	3300.40	0.13	2206.80	67.48	32.52	25.19	12.54	186.15	89.70
Mod1a	3259.97	677.40	0.04	87.00	74.48	25.52	1.08	0.37	8.10	2.78
Mod1b	539.97	30.00	0.00	30.00	95.33	4.67	0.50	0.02	3.58	0.18
Mod1c	337.38	24.00	0.00	24.00	85.83	14.17	0.33	0.06	2.58	0.43

Table 6: Analysis of the solutions - Instances H1-S1-80trains

From Table 6 it is possible to note that the solutions obtained by solving the model with a time limit of 1 hour have an average gap of 30.5%. From a quality point of view, these solutions are similar to those obtained with a time limit of three hours; the only

Test	Te dev	Zo dev	Ex dev	TW dev (m)	TW dev (h)	TW-dev/t (m)	TW-dev/t (h)
Mod1-1h	0.00	0.00	2.60	42.00	0.70	0.53	0.01
Mod1-3h	0.00	0.00	1.00	10.00	0.17	0.13	0.00
Mod1a	0.00	0.00	0.00	19778.00	329.63	247.23	4.12
Mod1b	528.00	0.00	0.00	20404.00	340.07	255.05	4.25
Mod1c	636.00	184.00	0.00	20596.00	343.27	257.45	4.29

Table 7: Analysis of the quality of the solutions (deviations) - Instances H1-S1-80trains

fact to stress is the difference in the trains exclusion (EX) that passes from 2.6 to 1 train.

Considering the different objective functions tested, i.e., those explained above (1a), 1b) and 1c)), results are reported in the rows Mod1a, Mod1b, Mod1c. We can note an important reduction in the CPU time when having less to penalize in the objective function: the model is able to furnish optimal solutions in few seconds; these solutions are characterized by very low tracks occupancy percentages and average waiting time per train, as shown in the last four columns of Table 6. This is due to the fact that, in these last three cases, it is possible to violate the time windows, thus permitting the trains to immediately enter the terminals instead of waiting in the station or in the park.

Looking at Table 7, we can note that time window deviations for the last three cases are very high. In particular, in the last case Mod1c all the trains have been served with an high deviation for all the reported terms (Teams, Zones and Time windows), while in case of Mod1a only the time window deviations are present. Note that, in these last cases the manager could stipulate new agreements with the terminals for performing the proposed schedules.

Summarizing, Mod1 solved with a time limit of one hour seems to be useful to support shunting managers in the scheduling of the trains planned to approach the terminal in a given time horizon, while both Mod1a and Mod1b can be used by the shunting managers for evaluating the required resources to manage a given number of trains.

6 Conclusions

In the present paper, a new problem arising in the port area has been described and a model for solving it has been proposed. The port area is characterized by a layout and the management has a given organization for operating each day a given number of import and export trains.

A deep analysis has been conducted for validating the model and for tuning some parameters (i.e., the weights used in the objective function of the proposed model) in such a way to have a model useful for both scheduling a given number of trains in an operative fix contest, and evaluating the resources necessary to manage a given number of trains when the shunting manager can re-organize the resources and can stipulate new agreements with the terminals.

Different scenarios concerning the number of trains to manage, the arrival distribution of trains during the time horizon, and distribution of the trains inside the the port zone have been analyzed.

The arrival distribution of the trains has a great impact on the waiting time in both the station and the park, even if the considered shunting area has a number of tracks enough for the management of the trains in all the analyzed situations. The station represents a very useful buffer for permitting the schedule of the trains constrained by the schedule of both the railway network and the ships. The number of tracks never represents a bottleneck for the system: even with one hundred trains, the percentage of tracks' occupancy is not greater than 25 %.

Moreover, the distribution of trains with respect to their paths inside the port area (that is, with respect to their origin/destination terminals and the operations they have to perform) has not a great impact on the trains' schedules. The only remarkable difference is related to the number of shunting teams required for managing one hundred trains: more unbalanced instances, as for those in set S4, required more shunting teams for managing trains in more congestion areas.

Finally, it is really interesting to note that Mod1 with different objective functions can be a useful support for evaluating the network layout, the available human resources, and the organization, within the strong constraints imposed by the ships and the national railway network.

As last remark, port systems characterized by a different layout can be easily evaluated thanks to the innovative operations-time-space network used to tackle this problem.

Acknowledgments

The authors thank CIRCLE S.P.A. for the contribution to this research. While working on the project, the third author was Adjunct Professor, Department of Computer Science and Operations Research, Université de Montréal. We gratefully acknowledge the support provided by the Natural Sciences and Engineering Council of Canada (NSERC), through its Discovery Grant program, and by the Fonds de recherche du Québec through their infrastructure grants.

References

A. Abbassi, A. E. h. Alaoui, and J. Boukachour. Robust optimisation of the intermodal freight transport problem: Modeling and solving with an efficient hybrid approach. *Journal of Computational Science*, 30:127–142, 2019.

A. Allahverdi. A survey of scheduling problems with no-wait in process. *European Journal of Operational Research*, 255(3):665–686, 2016.

D. Ambrosino and V. Asta. Intermodality and rail transport: Focus on port rail shunting operations. In *Advances in Optimization and Decision Science for Society, Services and Enterprises*, pages 351–366. Springer, 2019.

D. Ambrosino and V. Asta. An innovative operation-time-space network for solving different logistic problems with capacity and time constraints. *Networks*, 78(3):350–367, 2021.

D. Ambrosino, V. Asta, and T. G. Crainic. Optimization challenges and literature overview in the intermodal rail-sea terminal. *Transportation Research Procedia*, 52: 163–170, 2021.

T. Bektas and T. Crainic. *A brief overview of intermodal transportation*. CIRRELT, 2007.

R. Borndörfer, T. Klug, L. Lamorgese, C. Mannino, M. Reuther, and T. Schlechte. *Handbook of Optimization in the Railway Industry*, volume 268. Springer, 2018.

N. Boysen, M. Fliedner, F. Jaehn, and E. Pesch. Shunting yard operations: Theoretical aspects and applications. *European Journal of Operational Research*, 220(1):1–14, 2012.

C. Caballini, S. Fioribello, S. Sacone, and S. Siri. An milp optimization problem for sizing port rail networks and planning shunting operations in container terminals. *IEEE Transactions on Automation Science and Engineering*, 13(4):1492–1503, 2016.

V. Cacchiani, A. Caprara, and P. Toth. A fast heuristic algorithm for the train unit assignment problem. In *12th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2012.

A. Caris, C. Macharis, and G. K. Janssens. Decision support in intermodal transport: A new research agenda. *Computers in Industry*, 64(2):105–112, 2013.

H. J. Carlo, I. F. Vis, and K. J. Roodbergen. Transport operations in container terminals: Literature overview, trends, research directions and classification scheme. *European Journal of Operational Research*, 236(1):1–13, 2014a. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2013.11.023. URL `https://www.sciencedirect.com/science/article/pii/S0377221713009405`.

H. J. Carlo, I. F. Vis, and K. J. Roodbergen. Storage yard operations in container terminals: Literature overview, trends, and research directions. *European journal of operational research*, 235(2):412–430, 2014b.

J.-F. Cordeau, P. Toth, and D. Vigo. A survey of optimization models for train routing and scheduling. *Transportation science*, 32(4):380–404, 1998.

T. G. Crainic and K. H. Kim. Intermodal transportation. *Handbooks in operations research and management science*, 14:467–537, 2007.

G. Di Stefano and M. L. Koči. A graph theoretical approach to the shunting problem. *Electronic Notes in Theoretical Computer Science*, 92:16–33, 2004.

S. Fioribello, C. Caballini, S. Sacone, and S. Siri. A planning approach for sizing the capacity of a port rail system: scenario analysis applied to la spezia port network. *IFAC-PapersOnLine*, 49(3):371–376, 2016.

M. Gronalt, R.-C. Schultze, and M. Posset. Intermodal transport—basics, structure, and planning approaches. In *Sustainable Transportation and Smart Logistics*, pages 123–149. Elsevier, 2019.

H. Heggen, K. Braekers, and A. Caris. Optimizing train load planning: review and decision support for train planners. In *International Conference on Computational Logistics*, pages 193–208. Springer, 2016.

A. Janiak, W. A. Janiak, T. Krysiak, and T. Kwiatkowski. A survey on scheduling problems with due windows. *European Journal of Operational Research*, 242(2):347–357, 2015.

I. Krämer. Shunt-e 4.0—autonomous zero emission shunting processes in port and hinterland railway operations. *Journal of Traffic and Transportation Engineering*, 7:157–164, 2019.

C.-Y. Lee and D.-P. Song. Ocean container transport in global supply chains: Overview and research opportunities. *Transportation Research Part B: Methodological*, 95:442–474, 2017.

Z. Lin and R. S. Kwan. A two-phase approach for real-world train unit scheduling. *Public Transport*, 6(1-2):35–65, 2014.

Z. Lin and R. S. Kwan. A branch-and-price approach for solving the train unit scheduling problem. *Transportation Research Part B: Methodological*, 94:97–120, 2016.

J. Mar-Ortiz, N. Castillo-García, and M. D. Gracia. A decision support system for a capacity management problem at a container terminal. *International Journal of Production Economics*, 222:107502, 2020. ISSN 0925-5273. doi: https://doi.org/10.1016/j.ijpe.2019.09.023. URL `https://www.sciencedirect.com/science/article/pii/S0925527319303226`.

J. Muñuzuri, L. Onieva, P. Cortés, and J. Guadix. Using iot data and applications to improve port-based intermodal supply chains. *Computers & Industrial Engineering*, 2019.

M.-H. Phan and K. H. Kim. Collaborative truck scheduling and appointments for trucking companies and container terminals. *Transportation Research Part B: Methodological*, 86:37–50, 2016. ISSN 0191-2615.

A. Rusca, F. Rusca, E. Rosca, V. Dragu, and M. Rosca. Improving capacity of port shunting yard. In *Progress in Maritime Technology and Engineering*, pages 35–42. CRC Press, 2018.

F. Rusca, M. Popa, E. Rosca, A. Rusca, M. Rosca, and O. Dinu. Assessing the transit capacity of port shunting yards through discrete simulation. *Transport Problems: an International Scientific Journal*, 14(4), 2019.

X. Schepler, S. Balev, S. Michel, and Éric Sanlaville. Global planning in a multi-terminal and multi-modal maritime container port. *Transportation Research Part E: Logistics and Transportation Review*, 100:38–62, 2017. ISSN 1366-5545. doi: https://doi.org/10.1016/j.tre.2016.12.002. URL `https://www.sciencedirect.com/science/article/pii/S1366554516301697`.

N. Tomii and L. J. Zhou. Depot shunting scheduling using combined genetic algorithm and pert. *WIT Transactions on The Built Environment*, 50, 2000.

N. Tomii, L. J. Zhou, and N. Fukumura. An algorithm for station shunting scheduling problems combining probabilistic local search and pert. In *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, pages 788–797. Springer, 1999.

Y. Xie and D.-P. Song. Optimal planning for container prestaging, discharging, and loading processes at seaport rail terminals with uncertainty. *Transportation Research Part E: Logistics and Transportation Review*, 119:88–109, 2018.

X. F. Yin, L. P. Khoo, and C.-H. Chen. A distributed agent system for port planning and scheduling. *Advanced Engineering Informatics*, 25(3):403–412, 2011. URL `https://www.sciencedirect.com/science/article/pii/S1474034610001023`. Special Section: Engineering informatics in port operations and logistics.

7 Appendix

In the following Tables are reported the standard deviations for each set of generated instances used in the experimental tests described in Section 5.

Distr	**Set**	Unique	Others	Ter 1	Ter 2	Ter 3	Ter 4
2	S_1	0.00	0.00	0.00	0.00	0.00	0.00
	S_2	0.00	0.00	1.10	1.10	1.10	1.10
H	S_3	1.10	1.10	1.10	0.00	1.10	0.00
	S_4	1.10	1.10	1.10	2.00	2.00	2.28
	S_5	2.19	2.19	1.10	1.10	1.10	1.10
	S_6	2.00	2.00	0.89	2.19	1.67	2.61
	S_1	0.00	0.00	0.00	0.00	0.00	0.00
	S_2	0.00	0.00	1.10	1.10	0.89	0.89
H1	S_3	1.10	1.10	1.10	0.00	1.10	0.00
	S_4	0.00	0.00	0.00	0.00	1.67	1.67
	S_5	1.67	1.67	1.10	0.89	1.10	0.89
	S_6	1.10	1.10	0.89	1.79	1.67	2.68
s	S_1	0.00	0.00	0.00	0.00	0.00	0.00
	S_2	0.00	0.00	1.10	1.10	1.10	1.10
H	S_3	1.10	1.10	1.10	0.00	1.10	0.00
	S_4	1.10	1.10	1.10	1.67	0.89	1.10
	S_5	1.67	1.67	1.10	0.89	1.10	0.89
	S_6	1.79	1.79	0.00	1.79	1.67	2.28
o	S_1	0.00	0.00	0.00	0.00	0.00	0.00
	S_2	0.00	0.00	1.10	1.10	1.10	1.10
C	S_3	1.10	1.10	1.10	0.00	1.10	0.00
	S_4	1.10	1.10	1.10	1.67	2.19	2.28
	S_5	1.10	1.10	1.10	0.00	1.10	0.00
	S_6	1.41	1.41	0.00	1.41	0.89	0.89

Table 8: Standard deviations of generated instances - first campaign

Set	**#Trains**	Unique	Others	Ter 1	Ter 2	Ter 3	Ter 4
S_1	50	0.00	0.00	0.00	0.00	0.00	0.00
S_1	60	0.00	0.00	0.00	0.00	0.00	0.00
S_1	70	0.00	0.00	0.00	0.00	0.00	0.00
S_1	80	0.00	0.00	0.00	0.00	0.00	0.00
S_1	90	0.00	0.00	0.00	0.00	0.00	0.00
S_1	100	0.00	0.00	0.00	0.00	0.00	0.00
S_4	50	0.00	0.00	0.00	0.00	1.67	1.67
S_4	60	1.10	1.10	0.89	1.10	1.41	1.79
S_4	70	1.67	1.67	1.10	1.67	1.41	2.19
S_4	80	1.79	1.79	0.89	1.67	2.28	2.61
S_4	90	1.10	1.10	1.10	1.41	2.00	2.28
S_4	100	1.67	1.67	0.89	1.10	1.79	2.97

Table 9: Standard deviations of generated instances - second campaign