Strategic allocation of cyclically calling vessels for multi-terminal container operators

M. P. M. Hendriks · D. Armbruster · M. Laumanns · E. Lefeber · J. T. Udding

Abstract We consider a terminal operator who provides container handling services at multiple terminals within the same port. In this setting, the well-known berth allocation problem can no longer be considered for each terminal in isolation since vessel calls should be spread over the various terminals to avoid peaks and troughs in quay crane utilization, and an allocation of two connecting vessels to different terminals will generate inter-terminal container transport. In this paper, we address the problem of spreading a set of cyclically calling vessels over the various terminals and allocating a berthing and departure time to each of them. The objectives are (1) to balance the quay crane workload over the terminals and over time and (2) to minimize the amount of inter-terminal container transport. We develop a solution approach based on mixed-integer programming that allows to solve real-life instances of the problem within satisfactory time. Additionally, a practical case study is presented based on data from the terminal operator PSA Antwerp who operates multiple terminals in the port of Antwerp, Belgium. The computational results show the cost of the currently agreed schedules, and that

M. P. M. Hendriks · D. Armbruster · E. Lefeber (✉) · J. T. Udding
Department of Mechanical Engineering, Eindhoven University of Technology,
PO Box 513, 5600 MB Eindhoven, The Netherlands
e-mail: a.a.j.lefeber@tue.nl

D. Armbruster
School of Mathematical and Statistical Sciences, Arizona State University,
Tempe, AZ 85287-1804, USA

M. Laumanns
IBM Research-Zurich, 8803 Rueschlikon, Switzerland

J. T. Udding
PSA Antwerp, Napelsstraat 79, 2000 Antwerp, Belgium

relatively small modifications can significantly reduce the required crane capacities and inter-terminal transport.

1 Introduction

The importance of efficiently managing the scarce and expensive resources at a modern container terminal is well recognized. For an overview including problem classifications as well as solution methods, the reader is referred to the survey papers (Bierwirth and Meisel 2010; Stahlbock and Voß 2008; Steenken et al. 2004; Vis and de Koster 2003; Rashidi and Tsang 2006) and references therein. These studies classify the so-called berth allocation problem (BAP) as one of the key issues in a container terminal. Container shipping is characterized by loops. A loop is a set of vessels visiting a sequence of terminals according to a more or less fixed schedule, similar to a timetable for trains. For a terminal this results in a cyclic schedule of calling vessels. The BAP concerns the allocation of quay space and quay crane capacity to these loops in time minimizing an objective function.

This problem can be considered at three different levels: strategic, tactical, and operational. The timetable itself, i.e. the arrival and departure times of the various vessels, is fixed for a long period of time, usually in the order of years. It is subject to change only when contracts come up for renewal, or when new contracts are negotiated. At this strategic level, it is of paramount importance to determine the most suitable schedule for a terminal operator, and try to agree with the vessel operators upon the most favorable schedule. Due to its nature, optimization algorithms to solve this problem can be allowed to run for a number of hours, if not days. Given such a timetable of arrivals and departures at a terminal, we can consider the same problem at a tactical level. At this level tactical decisions and schedules are made. This results in (optimal, but still proforma) schedules for weekly quay crane allocations and yard reservations. At an operational level a terminal operator needs to cope with the day-to-day disturbances and adjust such a proforma schedule to reflect actual arrivals and departures of vessels, break-down of equipment, repair of yard pavement, etc. The optimization problems at this operational level need instant solutions.

An extensive number of studies on the BAP can be found in the literature, for instance (Giallombardo et al. 2010; Guan and Cheung 2004; Monaco and Sammarra 2007; Imai et al. 2005; Cordeau et al. 2005; Kim and Moon 2003; Hansen et al. 2008; Wang and Lim 2007; Moorthy and Teo 2006). They all have in common that they consider a single terminal at a tactical or operational level. Usually the various terminals in a port are not related to one another since each terminal is run by a different terminal operator who compete for the container flows into and out of the port. In those cases, it makes sense to consider the BAP for each terminal individually.

Recently, in an increasing number of ports (e.g. Singapore, Antwerp, and Rotterdam), one terminal operator is or will become responsible for multiple

terminals in that port. The BAP in these ports should then no longer be considered for individual terminals. By exploiting the fact that multiple terminals are under the control of one terminal operator, it is possible to avoid peaks and troughs in quay and quay crane utilization and to spread vessels evenly over the various terminals. Moreover, transshipment containers, i.e. containers that arrive but also leave by vessel generate inter-terminal traffic if the two vessels involved berth at different terminals. The cost of this inter-terminal traffic should also be taken into account and minimized (see Fig. 1).

The overall problem becomes to allocate (1) a terminal to each vessel (or actually the loop to which a number of vessels belong), (2) a time interval for berthing to each vessel, (3) a suitable berth position to each vessel in a terminal, and (4) a number of quay cranes to a vessel during the time that it is berthed (Hendriks 2009). Each of these subproblems can be nicely formulated mathematically, and is practically relevant and theoretically interesting in its own right. To the best of our knowledge and as stated in Ottjes et al. (2006), the BAP for multiple interrelated terminals has not yet been considered.

Multiple terminals in the same port are considered in Imai et al. (2008a) but the setting is entirely different from ours. In their formulation, an additional terminal comes into play only if the main terminal can no longer handle the expected number of vessels at some point in time. A number of vessels is then rerouted to a nearby terminal in order to minimize the total vessel service time. Compared to our research, this is a problem at an operational level, and the extra costs incurred for the haulage and reshuffling of containers to be transported between the two terminals is noted, but does not play a role in the resulting minimization problem.

In addition to the fact that almost all studies consider the BAP for a single terminal, also the inherent cyclic nature of the system, i.e. vessels in a loop call at regular (usually weekly) intervals, is hardly ever taken into account. Most of the existing studies of the BAP consider a set of vessels within a certain time horizon.
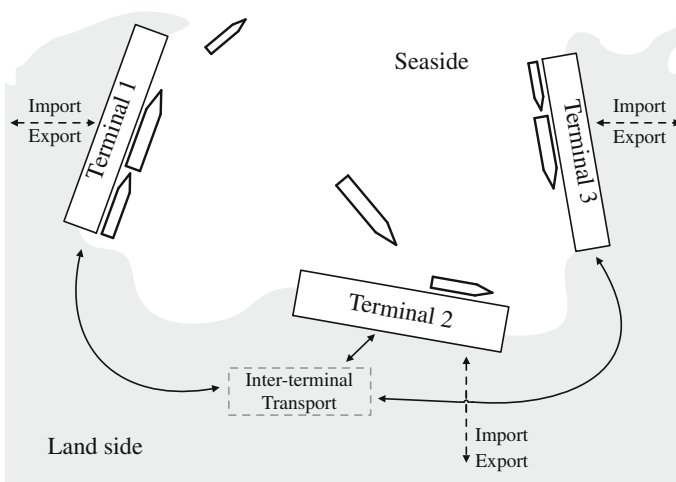

Fig. 1 Container flows in a cluster of multiple terminals

The corresponding objective in these studies often reduces to the well-known packing problem of fitting all vessels within a time horizon and minimizing the total weighted handling time for all vessels. In practice vessels may arrive at the end of this time period (cycle) and leave at the beginning of the same period (next cycle). Relating this to the packing problem implies that rectangles (vessels) may have to be cut into two pieces, where one piece is placed at the end of the time horizon and the other piece at the beginning. Another way of phrasing this is to consider this a packing problem on a cylinder. The authors in Moorthy and Teo (2006) take this phenomenon into consideration for a single terminal BAP in a continuous time setting. In this paper, we develop a formulation that takes this cyclic property into account for a multi-terminal container port in a discrete time setting.

In principle our problem can be considered as a multiple-job-on-one-processor scheduling problem, where each processor represents a terminal and each vessel a job. However, for reasons of computational complexity we split this problem into two. In the first step, discussed in this paper, we allocate a terminal and a berthing time interval to a set of cyclically calling vessels. At this stage, we only guarantee that the total lengths of all vessels berthed at a terminal at any point in time does not exceed that terminal's quay length. In addition we guarantee that a terminal's quay crane capacity to process the vessels berthed, assuming their nominal load compositions, will not be exceeded. To this end we assume that a vessel's processing time is inversely proportional to the crane capacity allocated to it. These simplifications make the problem computationally tractable, although there is clearly no guarantee that a feasible solution will result to do the actual berth and quay crane allocations at a tactical level. However, for typical berth and quay crane utilizations, this problem never surfaced. For very high utilization ratios this may no longer be the case. While we assume that each vessel has a preferred terminal and a preferred time interval, as is the case in practice, we phrase the problem in such a way that we can relax these restrictions to any desired level. We can impose bounds on the deviations from preferred berthing terminals and time slots. The objective is to balance the workload over the terminals and over time, so as to minimize the maximally required crane capacity per terminal, and at the same time to minimize the amount of container transport between the various terminals.

In a subsequent step (see Hendriks 2009; Hendriks et al. 2011) a joint berth position and container stacking problem can be solved for a given terminal allocation and a given timetable, in order to actually position the vessels along the quay, and the containers in blocks in the yard. At the same time the actual assignment of an integer number of quay cranes to the berthed vessels can be determined, taking into account the non-crossing constraint of quay cranes (Lim et al. 2004). A limited number of studies address the joint problem of berth allocation and crane scheduling in a two-phase approach (Imai et al. 2008b; Park and Kim 2003; Meisel and Bierwirth 2009). The authors in Imai et al. (2008b) assume that the process time of a vessel depends on its berth position. They construct a mixed-integer linear program (MILP) for the joint problem of berth allocation and crane scheduling and solve it using a genetic algorithm. The authors in Park and Kim (2003) take into account that a vessel's process time is inversely proportional to the number of cranes assigned to it and solve the problem in two

phases. The first phase determines the vessels' positions and berth times and an integer number of quay cranes in each time segment for each vessel. A sub-gradient optimization technique is applied to obtain a near-optimal solution of the first phase. The second phase constructs a schedule for each individual quay crane guaranteeing non-crossing of cranes. The main difference of our approach to Park and Kim (2003) is that we consider the multi-terminal BAP (rather than the single-terminal BAP) and face the additional problem of allocating each of the vessels to a terminal. Techniques to make such an allocation robust to disturbances at an operational level, such as deviations in arrival times or load compositions are explained in Hendriks et al. (2010).

The BAP can be modeled with a discrete or a continuous quay. In the discrete case, the quay is divided into segments with specific lengths, or even points when the quay and vessels' lengths are ignored. The problem can then be modeled as a parallel machine scheduling problem (Pinedo 1995; Lim 1998) where each vessel is a job and each berth a machine. Large segments result in a poor space utilization, whereas small segments increase the computational complexity of the problem. The continuous approach adopted in this paper is still complex from a computational point of view. However, with the proposed split of the problem, it can still be solved for real-world instances.

Existing studies on the single-terminal BAP consider either a static or dynamic arrival of vessels. The static case assumes that all vessels are in port waiting for a berth allocation. The discrete-quay-static-arrival BAP (Imai et al. 1997) is an assignment problem and is solvable in polynomial time with the Hungarian method (Papadimitriou and Steiglitz 1982). This method assigns jobs to machines by sequentially computing shortest paths until each job is assigned to a machine. In the dynamic case, vessels arrive while work is in progress, in which case there may be idle times between successive vessels. The dynamic BAP is NP-hard for both the continuous and the discrete quay case (Cordeau et al. 2005). Although the dynamic BAP is NP-hard, we can still solve real-life instances of our problem within satisfactory time, since we abstract from the position and quay crane allocations at the strategic level.

In order to show the feasibility of the model we have performed a case study, based on data from terminal operator PSA Antwerp, Belgium. The results show that this approach can yield significant reductions in required crane capacities and inter-terminal transport. The practical contribution of this paper is therefore twofold. First of all, it allows a set of terminals to balance their workload, so as not to exceed the available labor and/or equipment pool at any point in time. Moreover, given the enormous lead times for purchasing new cranes or expanding quays, it allows a terminal to search for (possibly temporary) measures that can be taken to accommodate a new customer on a short notice, or to deal with berths or quay cranes that will be out of commission for a longer period of time. Whether changes to the current schedules are commercially feasible is beyond the scope of this paper, but our study shows the cost of the current schedules and the potential for cost reduction.

The outline of the paper is as follows: In the next section, we formulate the problem and derive an appropriate MILP which can be adjusted to exhibit the right

amount of flexibility to allocate terminals and berthing intervals to vessels. In Sect. 3, we consider a real-life case, impose bounds on the deviations from preferred terminals and time intervals that express relevant commercial considerations, and solve the problem in two sequential steps. We end with conclusions and recommendations in Sect. 4.

2 Mathematical model

2.1 Problem description

For the rest of this paper we postulate the following set-up. We are given a cluster of terminals $\{1, 2, \ldots, T\}$, with typical elements s and t, a set of vessels $\{1, 2, \ldots, V\}$, with typical elements v and w, and a set of container destinations $\{0, 1, 2, \ldots, V\}$, with typical element z. A container destination can be any vessel, viz. the vessel that a container is to be loaded on. A hinterland destination, to which a container is transported by train or truck, is labeled destination 0. We assume vessels call cyclically, where each vessel in the set arrives exactly once in each cycle. We assume the cycle to be a sequence of K time slots, with k a typical time slot. Addition and subtraction in the domain of time slots is done by adding a proper multiple of K, so as to end up between 1 and K.

In the cluster of terminals, the container vessels have to be discharged and loaded. Each vessel v brings a pre-determined (nominal) number of inbound containers $I_{vz} \in \mathbb{N}$ with destinations z, where $v \neq z$. Besides the containers brought in by vessels, a certain amount of containers H_v with destination $v \neq 0$ is brought into the terminals by trucks and trains during the cycle. Each container is to be delivered to one of the terminals. Usually, such a container is delivered to the terminal where the destination vessel berths. However, if at the time of arrival, the yard capacity of that particular terminal is fully used, containers have to be temporarily stacked in a different terminal, resulting in inter-terminal transport of containers, which is established by trucks. Furthermore, each vessel v exports a number of outbound containers $O_v \in \mathbb{N}$ in each cycle.

Terminal t has a restricted quay length $L_t \in \mathbb{R}^+$ and a maximum quay crane capacity $N_t \in \mathbb{N}$. One aspect of the problem is to determine whether all of the quay cranes are necessary. Once berthed, vessel v occupies a certain amount of quay meters M_v. In addition, this length M_v determines the maximum number of quay cranes $S_v \in \mathbb{N}$ that can process vessel v in parallel, and the efficiency $\eta_v \in [0, 1]$ of the quay cranes on vessel v. In practice, quay cranes with different processing rates are present in the terminals. We do not take the specific allocation of quay cranes to vessels into account in this strategic planning stage, but assume the average processing rate $\lambda_t \in \mathbb{N}$ to be the processing rate of each quay crane in terminal t. Therefore the handling time of vessel v in terminal t depends on (1) the mean processing rate λ_t in terminal t, (2) the efficiency η_v of quay cranes operating vessel v, (3) the number of quay cranes processing vessel v and (4) the total number of inbound and outbound containers I_{vz} and O_v of vessel v. We assume the processing

time of vessel v to be inversely proportional to the first three of these items and proportional to the latter. Furthermore, the number of quay cranes processing vessel v may change from one time slot to another. After the discharging and before the loading, containers can temporarily be stored in the yard of terminal t up to the yard's capacity W_t. The number of time slots that it takes to transport a container from terminal s to terminal t is defined as $\Delta_{st} \in \mathbb{N}$. Furthermore, the time that a vessel is berthed is bounded by the length of the calling cycle. In addition, we assume that vessels arrive at the beginning of a time slot and depart at the end of a time slot.

Our goal is to minimize the total costs of the system, which consist of two components. First, costs are associated with each operating quay crane. We define c_t to be the linear cost factor for the required crane capacity at terminal t. Second, a transportation cost c_{st} is associated with each container that is transported from terminal s to terminal t.

In Table 1 we give a summary of all parameters introduced in this section. These parameters are not entirely independent. Due to the cyclic property of the considered system, we assume conservation of containers, i.e. everything brought into a terminal by vessel or from the hinterland to be loaded onto a vessel will leave with that vessel, i.e.,

Table 1 Model parameters

Parameter	Definition
T	Number of terminals in the cluster
V	Number of vessels in the set
K	Number of discrete time slots within the cycle
L_t	Quay length (m) of terminal t
M_v	Quay length required for vessel v (m)
I_{vz}	Number of inbound containers to be discharged from vessel v with destination z
O_v	Number of outbound containers to be loaded onto vessel v
H_v	Number of containers with destination v arriving from the hinterland
N_t	Maximum quay crane capacity in terminal t
S_v	Maximum number of quay cranes that can process vessel v in parallel
λ_t	Mean processing rate of quay cranes in terminal t (containers/time slot)
η_v	Vessel efficiency with respect to quay crane rate
Δ_{st}	Number of time slots needed to transport a container from terminal s to t
W_t	Maximum number of containers that can be stored in terminal t
c_{st}	Transportation costs of a container from terminal s to t
c_t	Quay crane costs in terminal t

Table 2 Overview of decision variables

Variable	Definition
$\mathcal{X}_{tv}$	1, if vessel v calls at terminal t, 0 otherwise
$\mathcal{A}_v$	The arrival time slot of vessel v
$\mathcal{D}_v$	The first time slot after vessel v's departure
$\mathcal{M}_{tv}(k)$	The amount of quay meters consumed in terminal t by vessel v during time slot k
$\mathcal{Q}_{tv}(k)$	The amount of quay cranes that process vessel v in terminal t during time slot k
$\mathcal{H}_{tv}(k)$	The number of containers from the hinterland arriving into terminal t with destination v during time slot k
$\mathcal{F}_{stv}(k)$	The number of containers transported from terminal s to terminal t with destination v during time slot k, $s \neq t$
$\mathcal{W}_{tv}(k)$	WIP in terminal t with destination v at the end of time slot k
$\mathcal{N}_t$	The amount of quay cranes required in terminal t
$\mathcal{B}_v(k)$	Indicator whether vessel v is berthed during time slot k
$\mathcal{E}_v$	Indicator whether $\mathcal{A}_v \geq \mathcal{D}_v$

$$\sum_{w=1}^{V} I_{wv} + H_v = O_v \quad \forall v. \tag{1}$$

Another way of reading this equation is as a definition of O_v.

2.2 Formulation as a mixed-integer linear program

In this section we introduce a number of decision variables, for which we shall use curly letters to distinguish them clearly from the variables and parameters introduced in the previous section. The problem is to assign to each vessel a terminal and a time interval during which that vessel is berthed at that terminal. Therefore, we introduce a binary decision variable $\mathcal{X}_{tv}$, which is set to 1 if vessel v calls at terminal t and integer decision variables $\mathcal{A}_v$, the arrival time slot of vessel v, and $\mathcal{D}_v$, the first time slot after vessel v's departure. Hence, the processing of vessel v starts in time slot $\mathcal{A}_v$ and ends in time slot $\mathcal{D}_v - 1$.

These decision variables need to satisfy a number of constraints. In order to express these constraints more easily we will introduce a set of auxiliary decision variables. All decision variables are summarized in Table 2.

2.2.1 Terminal and time allocation

Each vessel should berth at exactly one terminal, i.e.

$$\sum_{t=1}^{T} \mathcal{X}_{tv} = 1 \quad \forall v. \tag{2}$$

Berthing and departure times of each vessel are within the cycle (although not necessarily the same cycle):

$$1 \le \mathcal{A}_v, \mathcal{D}_v \le K \quad \forall v. \tag{3}$$

A vessel can only be processed when it is berthed, which is exactly between its arrival and departure times. To express whether a vessel is berthed during a certain time slot, we introduce a binary variable $\mathcal{B}_v(k)$

$$\mathcal{B}_v(k) = \begin{cases} 1 & \text{if vessel } v \text{ is berthed during times lot } k, \\ 0 & \text{otherwise.} \end{cases} \quad \forall v, k.$$

Since we study a cyclic setting the arrival time slot of a vessel does not necessarily precede the departure time slot. A vessel may arrive in a time slot in one cycle and leave in an earlier time slot in the next. Since each vessel can be entirely processed in at most K time slots, the amount of time that a vessel is berthed spans at most 2 cycles. Thus, if $\mathcal{A}_v < \mathcal{D}_v$ then the vessel is berthed between time slots $\mathcal{A}_v$ and $\mathcal{D}_v$ in the same cycle, and if $\mathcal{A}_v \ge \mathcal{D}_v$ then the vessel is berthed during the time slots that are at least $\mathcal{A}_v$ or less than $\mathcal{D}_v$. Hence the case $\mathcal{A}_v = \mathcal{D}_v$ is interpreted as a vessel v that is continuously berthed for a full cycle.

In order to properly phrase this relationship between $\mathcal{A}$, $\mathcal{D}$, and $\mathcal{B}$ we distinguish 2 cases: $\mathcal{A}_v < \mathcal{D}_v$, and $\mathcal{A}_v \ge \mathcal{D}_v$. In each of these cases we first look at a numerical example and then derive the general constraints for each of the cases. Finally, we combine them into one again by introducing an auxiliary binary decision variable $\mathcal{E}_v$.

1. $\mathcal{A}_v < \mathcal{D}_v$. Here vessel v arrives and departs in the same cycle, and the time slots during which a vessel is berthed are contiguous in a cycle. As an example, assume $K = 10$, $\mathcal{A}_v = 3$ and $\mathcal{D}_v = 7$. Since $\mathcal{B}_v(k)$ is 1 between $\mathcal{A}_v$ and $\mathcal{D}_v - 1$, and 0 everywhere else, we find the table below, where we also list the values of $k \cdot \mathcal{B}_v(k)$.

			$\mathcal{A}_v$				$\mathcal{D}_v$			
k	1	2	3	4	5	6	7	8	9	10
$\mathcal{B}_v(k)$	0	0	1	1	1	1	0	0	0	0
$k \cdot \mathcal{B}_v(k)$	0	0	3	4	5	6	0	0	0	0

This table can be fully characterized by 3 constraints: First, after its departure a vessel is not berthed anymore leading to

$$k \cdot \mathcal{B}_v(k) \le \mathcal{D}_v - 1. \tag{4}$$

Second, before its arrival time slot, a vessel is not berthed yet leading to

$$(K - k) \cdot \mathcal{B}_v(k) \le K - \mathcal{A}_v. \tag{5}$$

Note that the two constraints above set all non-berthing elements of $\mathcal{B}_v$ to 0. The number of time slots k for which $\mathcal{B}_v(k) = 1$ is equal to the number of time slots between v's arrival and departure, and is given by

$$\sum_{k=1}^{K} \mathcal{B}_v(k) = \mathcal{D}_v - \mathcal{A}_v. \tag{6}$$

2. $\mathcal{A}_v \geq \mathcal{D}_v$. In this case the time slots when a vessel is *not* berthed are contiguous in a cycle. As an example, assume $K = 10$, $\mathcal{A}_v = 8$ and $\mathcal{D}_v = 4$. Hence $\mathcal{B}_v(k)$ should be 0 between $\mathcal{D}_v$ and $\mathcal{A}_v - 1$, and 1 everywhere else. This is shown in the table below, where we also list the values of $k \cdot (1 - \mathcal{B}_v(k))$.

				$\mathcal{D}_v$				$\mathcal{A}_v$		
k	1	2	3	4	5	6	7	8	9	10
$\mathcal{B}_v(k)$	1	1	1	0	0	0	0	1	1	1
$k \cdot (1 - \mathcal{B}_v(k))$	0	0	0	4	5	6	7	0	0	0

The case is similar to case 1 with the roles of $\mathcal{A}_v$ and $\mathcal{D}_v$ interchanged and $\mathcal{B}_v(k)$ replaced by $1 - \mathcal{B}_v(k)$. Therefore, the three constraints for this case are

$$k \cdot (1 - \mathcal{B}_v(k)) \leq \mathcal{A}_v - 1 \tag{7}$$

$$(K - k) \cdot (1 - \mathcal{B}_v(k)) \leq K - \mathcal{D}_v \tag{8}$$

$$\sum_{k=1}^{K} (1 - \mathcal{B}_v(k)) = \mathcal{A}_v - \mathcal{D}_v. \tag{9}$$

To express this case distinction as a decision variable, and to formulate constraints that relate $\mathcal{A}$, $\mathcal{D}$ and $\mathcal{B}$ we introduce for each vessel v a binary variable $\mathcal{E}_v$

$$\mathcal{E}_v = \begin{cases} 1 & \text{if } \mathcal{A}_v \geq \mathcal{D}_v, \\ 0 & \text{otherwise.} \end{cases}$$

This allows us to combine Eqs. 6 and 9 into

$$\sum_{k=1}^{K} (\mathcal{B}_v(k) - \mathcal{E}_v) = \mathcal{D}_v - \mathcal{A}_v \quad \forall v. \tag{10}$$

Since $\mathcal{B}_v(k)$ is a binary variable that will be 1 for at least 1 time slot in any solution, $\mathcal{E}_v$ is 1 if and only if $\mathcal{D}_v \leq \mathcal{A}_v$. Hence we can combine Eqs. 4 and 7, and Eqs. 5 and 8 to

$$1 - \mathcal{A}_v \leq k \cdot (\mathcal{B}_v(k) - \mathcal{E}_v) \leq \mathcal{D}_v - 1 \quad \forall v, k \tag{11}$$

and

$$\mathcal{D}_v - K \leq (K - k) \cdot (\mathcal{B}_v(k) - \mathcal{E}_v) \leq K - \mathcal{A}_v \quad \forall v, k. \tag{12}$$

Note that in this formulation, each vessel arrives exactly once during the considered cycle. In practice however, some vessel lines may call multiple times during the same cycle, in particular when the cycle lengths of the various loops are different. This can easily be incorporated in the model by introducing auxiliary variables for the arrival times of the additional (second, third, and so on) vessels of such a loop within the same cycle and defining the fixed inter-arrival time between them. Although the number of constraints will grow in

such a formulation, the number of decision variables will stay the same and therefore the complexity of the problem does not increase much.

2.2.2 Quay and crane capacity allocation

To formulate constraints for quay we introduce the auxiliary variable

$\mathcal{M}_{tv}(k) =$Amount of quay consumed in terminal t by vessel v during time slot $k, \forall t, v, k$

Vessel v requires an amount of quay meters M_v at a terminal t during time slot k, if and only if the vessel is actually berthed during time slot k at terminal t. Since vessel v does not consume space at a terminal where it does not berth and since it never consumes more than M_v at the terminal where it berths we have

$$\mathcal{M}_{tv}(k) \leq M_v \cdot \mathcal{X}_{tv} \qquad \forall t, v, k. \tag{13}$$

Also, over all terminals, vessel v requires M_v space if and only if it is berthed at time slot k.

$$\sum_{t=1}^{T} \mathcal{M}_{tv}(k) = M_v \cdot \mathcal{B}_v(k) \quad \forall v, k. \tag{14}$$

Finally, the sum of lengths of all vessels berthed at a terminal during any time slot should be less than or equal to the total quay length of that terminal

$$\sum_{v=1}^{V} \mathcal{M}_{tv}(k) \leq L_t \qquad \forall t, k. \tag{15}$$

In a similar vein we restrict the allocation of quay cranes capacity to vessels. We introduce an auxiliary continuous variable

$\mathcal{Q}_{tv}(k) =$ Amount of quay cranes processing vessel v in terminal t in time slot k, $\forall t, v, k$

Since a vessel v can only be operated on in a terminal if that vessel calls at that terminal, and since no more than S_v quay cranes can operate on that vessel simultaneously we have

$$\mathcal{Q}_{tv}(k) \leq S_v \cdot \mathcal{X}_{tv} \quad \forall t, v, k. \tag{16}$$

Moreover, a vessel can only be operated on, if it is actually berthed. As opposed to quay meters used, however, it is not necessarily the case that the maximum amount of quay cranes is employed on a vessel at any time.

$$\mathcal{Q}_{tv}(k) \leq S_v \cdot \mathcal{B}_v(k) \quad \forall t, v, k. \tag{17}$$

On the other hand, a vessel has to be fully processed during the cycle, which means that all its import containers (destined for the hinterland or other vessels) have to be discharged and all its outbound containers have to be loaded. For each import and export container a crane handling is required and hence:

$$\sum_{t=1}^{T} \sum_{k=1}^{K} \eta_v \lambda_t \cdot \mathcal{Q}_{tv}(k) = \sum_{z=0}^{V} I_{vz} + O_v \quad \forall v. \tag{18}$$

We want to minimize the maximum number of quay cranes in a terminal that are ever required during the cycle. Therefore, we introduce an auxiliary variable

$\mathcal{N}_t =$ The minimum number of quay cranes required in terminal t, $\quad \forall t$,

together with the constraint

$$\sum_{v=1}^{V} \mathcal{Q}_{tv}(k) \leq \mathcal{N}_t \quad \forall t, k. \tag{19}$$

The variable $\mathcal{N}_t$ is present in the objective function. When the intention is to determine the optimal number of quay cranes at each terminal, this variable $\mathcal{N}_t$ is unconstrained. However, quay cranes are typically not a resource that can be acquired or disposed of easily. Moreover, there may be physical limitations for the maximum number of quay cranes that can operate in one particular terminal, e.g. due to maintenance. Therefore, we use a hard upper bound N_t for the maximum number of quay cranes available at a terminal. The maximum number of quay cranes ever required in terminal t during the cycle cannot be larger than the number of quay cranes actually available in terminal t:

$$\mathcal{N}_t \leq N_t \quad \forall t. \tag{20}$$

Note on quay and quay crane capacities   Restriction (15) is a necessary but not a sufficient condition to guarantee that all vessels allocated to a certain terminal in a certain time slot can actually berth at that terminal. We have to simplify the problem in order to be able to solve it, and yet the simplified problem should still have practical relevance. This particular simplification achieves both: The resulting problem can be solved in a reasonable amount of time, and for typical quay utilizations (0.2–0.5) the actual allocation of vessels to the quay turns out to be still feasible in the many random experiments that we conducted (Hendriks 2009). This is in particular true for the port of Antwerp case study. In the case of very high quay crane utilizations, the solution may yield a solution that at a tactical level turns out not to be feasible.

The variable Q_{tv} is a continuous variable while in reality the number of quay cranes is an integer. However, Q_{tv} should be interpreted as an average and hence non integer assignments can always be realized by assigning integer numbers of cranes for partial time periods. Example an assignment of 3.6 quay cranes over a certain time period t may be realized as an assignment of 4 during $0.6t$ and of 3 during $0.4t$. However, care must be taken to make sure an available quay crane is not prevented from doing work because it cannot cross its neighboring quay cranes on the quay. In all random experiments that we conducted with typical quay crane utilization (0.5–0.7) and also in the actual port of Antwerp case, the resulting strategic solution always turned out to have a subsequent tactical solution.

2.2.3 Container flows

The problem to solve is how best to organize the flow of containers into and out of the terminals. A vessel berthing in some terminal brings containers into that terminal, destined for the hinterland and other vessels, requiring yard space in that terminal, in addition to quay and quay crane capacity. Also, it takes a number of containers from that terminal, which should be present in that terminal at the moment of loading. In order to capture all container movements during a cycle, we also have to express the non-vessel related moves. These split into two flows, the flow of export containers from the hinterland into the terminals, and the inter-terminal flow of containers. Notice that we decided to not explicitly model the import containers from the vessels to the hinterland. We assume that inbound containers with destination 0 (hinterland) are transported into the hinterland some time after their arrival. These containers will only affect the storage levels in a terminal and can be considered a base storage level in a terminal. This can be compensated for in the maximum storage level W_t for each terminal. We need two additional auxiliary variables to express the flow from the hinterland into each terminal

$\mathcal{H}_{tv}(k) = $ The number of containers from the hinterland arriving into terminal t with destination v during time slot k, $\quad \forall t, v, k$.

and the flow between the various terminals

$\mathcal{F}_{stv}(k) = $ The number of containers transported from terminal s to terminal t with destination v during time slot k, $s \neq t$, $\quad \forall s, t, v$.

The first constraint is that the total number of export containers should be absorbed by the combined terminals during the cycle, i.e.,

$$\sum_{k=1}^{K} \sum_{t=1}^{T} \mathcal{H}_{tv}(k) = H_v \quad \forall v. \tag{21}$$

Containers that are destined for a certain vessel do not necessarily enter the system at the terminal where that vessel berths. If a terminal is stacked up to its capacity, containers from the hinterland destined for vessel v temporarily may have to be stacked in a terminal different from the one that vessel v calls at. In order to express that the maximum yard capacity should not be exceeded we need to express the storage level of each terminal, the Work In Process or WIP, at every time slot

$\mathcal{W}_{tv}(k) = $ WIP in terminal t with destination v at the end of time slot k, $\quad \forall t, v, k$.

Difference equations are applied to update the storage levels over time. More precisely, the amount of containers in a terminal with a particular destination v at the end of time slot k is equal to the amount of containers in that terminal with that destination at the end of the previous time slot plus all incoming flows for that destination (from vessels, from other terminals, and from the hinterland) minus all

outgoing flows for that destination (to vessels and to other terminals) during time slot k. We assume that the amount of containers to be discharged from a vessel with a particular destination or to be loaded onto a vessel is spread evenly over the time slots during which vessel v is actually berthed. For vessel v, the total number of containers to be handled during a particular call is $\sum_{z=0}^{V} I_{vz} + O_v$. Of this total number, we define ι_{vz} to be the fraction to be discharged with destination z, and $ø_v$ to be the fraction to be loaded. That is, $\iota_{vz} = \frac{I_{vz}}{\sum_{z=0}^{V} I_{vz} + O_v}$ and $ø_v = \frac{O_v}{\sum_{z=0}^{V} I_{vz} + O_v}$. Hence the difference equation for $\mathcal{W}_{tv}$ becomes

$$
\begin{aligned}
\mathcal{W}_{tv}(k) = \mathcal{W}_{tv}(k-1) + \sum_{w=1}^{V} \iota_{wv}\eta_w\lambda_t \cdot \mathcal{Q}_{tw}(k) - ø_v\eta_v\lambda_t \cdot \mathcal{Q}_{tv}(k) \\
+ \mathcal{H}_{tv}(k) + \sum_{s=1}^{T} \mathcal{F}_{stv}(k-\Delta_{st}) - \sum_{s=1}^{T} \mathcal{F}_{tsv}(k) \quad \forall t, v, k
\end{aligned}
\tag{22}
$$

where the arguments of $\mathcal{W}_{tv}$ and $\mathcal{F}_{stv}$ are understood to be increased by K, if less than or equal to 0, due to the cyclic nature of the system. In order to prevent the storage levels from exceeding the capacity during a time slot, we have to put an upper bound on $\mathcal{W}_{tv}(k)$. Assuming the worst case scenario, viz. in a time slot all containers arrive at a terminal before any container leaves, the following constraint is required:

$$
\sum_{v=1}^{V} \left(\mathcal{W}_{tv}(k-1) + \sum_{w=1}^{V} \iota_{wv}\eta_w\lambda_t \cdot \mathcal{Q}_{tw}(k) + \mathcal{H}_{tv}(k) + \sum_{s=1}^{T} \mathcal{F}_{stv}(k-\Delta_{st}) \right) \leq W_t
\tag{23}
$$
$$
\forall t, k.
$$

On the other hand, to prevent the storage level from becoming negative during time slot k, again assuming the worst case scenario, the following constraint should also be satisfied

$$
\mathcal{W}_{tv}(k-1) - ø_v\eta_v\lambda_t \cdot \mathcal{Q}_{tv}(k) - \sum_{s=1}^{T} \mathcal{F}_{tsv}(k) \geq 0 \quad \forall t, v, k.
\tag{24}
$$

By summing over k one can check that the container flow constraints together with the quay crane allocation constraints generate inter-terminal flows $\mathcal{F}_{stv}(k)$ that ensure that all containers arriving at any terminal destined for vessel v will arrive in time at the terminal where v berths, and will be loaded onto the vessel during its call.

In Table 2 we summarize.

2.2.4 Objective function

For each terminal t linear costs c_t are assigned to the required quay crane capacity in that terminal. A linear unit penalty cost c_{st} is assigned when containers are transported from terminal s to terminal t. The variables can be represented in a vector and the objective function becomes

$$\min \quad \sum_{k=1}^{K}\sum_{s=1}^{T}\sum_{t=1}^{T}\sum_{v=1}^{V} c_{st}\mathcal{F}_{stv}(k) + \sum_{t=1}^{T} c_t\mathcal{N}_t. \tag{25}$$

Remark   In the solution of this MILP it could be that the storage level in one or more terminals has a minimum larger than zero, i.e. an arbitrary amount of containers is stored in a certain terminal during the entire cycle. This can be prevented by assigning a small cost for each stored container. This storage cost could also be used to model the fact that containers require a minimum amount of yard space throughout a cycle.

2.3 Modeling additional constraints

In the previous section we have laid the foundation for a MILP to minimize the quay crane and inter-terminal costs by allocating vessels to terminals for a duration that will allow proper discharging and loading for the expected number of containers to be handled. In practice, however, the freedom to allocate any vessel to any terminal, and choosing the best arrival and departure times for these vessels is usually somewhat restricted. Some vessels may need to berth in a particular terminal, for example for reasons of draft or quay crane limitations. Also, arrival and departure times may be flexible, but only within certain limits, because a vessel may have a previous or a next port of call with certain arrival and departure times. Hence, flexibility is not unlimited, and the potential reduction of the required number of quay cranes and the costs for inter-terminal transport should take into account the limited freedom to allocate a vessel to another terminal or give it different arrival or departure times. Therefore we slightly expand the MILP derived in the previous section, such that for a chosen/negotiated level of flexibility, the required quay crane capacity and inter-terminal transport is minimized. Although the model does not incorporate costs for certain modifications (since they are hard to quantify), we are at least able to quantify the savings in crane capacities and inter-terminal transport induced by these modifications, or put differently, the costs associated with existing schedules.

Hence in reality the goal is to minimize the number of quay cranes required for the current throughput in a set of terminals, and at the same time reduce the costs for inter-terminal transport, by adapting the current terminal and time allocations. Therefore, let us assume that we are given X_v, the preferred terminal of vessel v, and A_v and D_v, its preferred berthing and departure times, respectively. We would expect that changing the current terminal allocation and slightly shifting the preferred berthing interval in time for a couple of vessels will already lead to a significant reduction in both the objectives.

There are many ways to allow limited flexibility. In a practical setting, it might be the case that a terminal allocation is fixed for a vessel, but that its arrival and departure times are flexible, or the other way around. When arrival and departure time are flexible, it might be the case that the length of the berthing interval is still fixed, or also flexible. All of these flavors can be captured in extensions of the base model. In order to show the potential of this approach, we opt for the following

extension of the model and define two sets: the set S is the set of vessels whose current terminal and berthing allocation are allowed to be adapted, and the set F is the set of vessels whose berthing terminal and intervals are fixed.

We have to specify additional constraints (while all the others remain valid) for the vessels in the different sets. First, for the vessels in F, the time-position of the berthing interval should be fixed:

$$\mathcal{A}_i = A_i, \quad \forall i \in F \tag{26}$$

$$\mathcal{D}_i = D_i, \quad \forall i \in F \tag{27}$$

and the current terminal allocation for the vessels in F should be fixed

$$\mathcal{X}_{ij} = 1, \quad \forall i \in F \wedge j = X_i. \tag{28}$$

The vessels in set S are free to be allocated to any one of the terminals according to (2), so no additional constraint is necessary. Next, let P_v be defined as the number of time slots of the preferred berthing interval, i.e. $P_v = D_v - A_v$, for all v. (As usual, we add K when necessary in order to make sure that $1 \leq P_v \leq K$.) We want to allow some freedom in the time allocation of a berthing interval to the vessels in S, but not in the number of time slots of berthing:

$$\sum_{k=1}^{K} \mathcal{B}_v(k) = P_v, \quad \forall v \in S. \tag{29}$$

Instead, we want to allow some, but not arbitrary, flexibility in where to place this berthing interval within the cycle. Therefore, we introduce parameter G_v to express that vessel $v \in S$ can be placed at most G_v time slots earlier or G_v time slots later with respect to its preferred arrival as shown in Fig. 2.

The corresponding lower bound for arrival and upper bound for departure of vessel $v \in S$ can then be defined as $A_v^l = A_v - G_v$ and $D_v^u = D_v - G_v$, respectively. Since P_v is fixed, we can express the desired limited flexibility as a constraint on A_v only. The earliest allowed arrival is A_v^l and the latest $A_v^l + 2G_v$, where care has to be taken that the latter value is between 1 and K again. Without loss of generality we assume that $2G_v \leq K$, since that already allows us to express any degree of flexibility in the arrival time of vessel v. By this assumption we know that it suffices to subtract K simply once in order to make sure that $A_v^l + 2G_v$ becomes a proper time slot again. This suggests the following case distinction between the cases
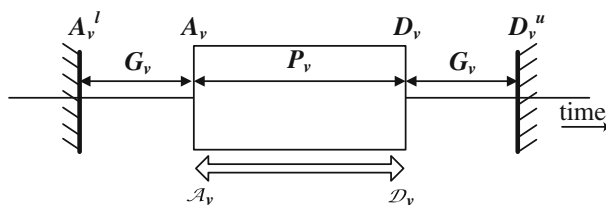

Fig. 2 Flexibility in the time-position of the berthing interval P_v of vessel $v \in S$

$A_v^l + 2G_v \leq K$ and $A_v^l + 2G_v > K$, which is known beforehand, since these are given parameters.

- $A_v^l + 2G_v \leq K$: In this case any allowed value of the decision variable $\mathcal{A}_v$ is larger than A_v^l. Hence, the constraints are

$$A_v^l \leq \mathcal{A}_v \leq A_v^l + 2G_v. \tag{30}$$

- $A_v^l + 2G_v > K$: In this case, it is possible that $\mathcal{A}_v \geq A_v^l$, in which case there is no additional restriction on its value, other than that it should be at most K, which is taken care of by the base model. The other possibility is that $\mathcal{A}_v < A_v^l$, in which case it should also be at most $A_v^l + 2G_v - K$, which is between 1 and K because of our assumption that $2G_v \leq K$. Therefore, we introduce the additional auxiliary variable $\mathcal{E}_v^a$ to distinguish these two cases. The technique to make this distinction is very similar to Eqs. 10 through 12. For each case exactly one of the Eqs. 31 and 32 will force $\mathcal{E}_v^a$ to assume the right value:

$$K \cdot \mathcal{E}_v^a \geq A_v^l - \mathcal{A}_v \tag{31}$$

$$-K \cdot (1 - \mathcal{E}_v^a) < A_v^l - \mathcal{A}_v. \tag{32}$$

Now the additional restriction on the arrival time, as phrased by the case distinction above can be stated as

$$\mathcal{A}_v \leq A_v^l + 2G_v - K \cdot \mathcal{E}_v^a. \tag{33}$$

2.3.1 A two-step approach

With this extended model we can solve problems of a practical size. Currently the allocation schedules in ports are planned on a 1 or 2 h time grid. However, if we try to solve the MILP formulation for such time grids ($K = 84$ and $K = 168$, respectively) the problem for a realistic number of vessels, say $V = 40$, takes too long to allow us to experiment with many different choices for the flexibility in arrival times. For a real scheduling situation in a port this may not be a problem, since such an optimization is done at most a few times a year and hence may run for a number of days. However, since the purpose of this research is to show that a MILP can be formulated and is very useful to highlight the (considerable) costs that are associated with the schedules that are currently used in practice, we need to experiment with many different settings. Therefore we perform the optimization in two sequential steps: First, the MILP is solved for time slots of 8 h. For a weekly cycle this means $K = 21$. The model can then be solved within minutes while generating allocations that are quite accurate, since port employees work in shifts and vessels commonly berth during multiples of a shift. In a second step, a similar MILP is built to refine the constructed allocation per terminal on a 1-h time grid. Solving this MILP takes less than a second.

The resulting two-step optimization approach enables us to efficiently investigate the dependency of the cost savings (in crane capacity and inter-terminal transport)

on the level of flexibility. We select a number of vessels and the maximum allowed modifications in their arrivals. We round the desired berthing time up to a multiple of 8 h to express the parameter P_v of the model introduced above, and solve it. Then, in the second step, we refine the constructed allocation per terminal from 8-h time slots into 1-h time slots, where the terminal allocation remains fixed. We set the berthing time interval for each vessel back to the original value (in hours instead of multiple of 8 h) to express the parameter P_v and require it to be positioned between the allocated arrival and departure times of the first step. So $\mathcal{A}_v^\star$, the optimal values of the first step optimization on the coarse time grid, play the same role as A_v^l in Eqs. 30 through 33. For $\mathcal{D}_v$ we introduce similar restrictions to bound the departure time from above, using $\mathcal{D}_v^\star$. Since the vessels are already distributed among the terminals, the remaining objective is to minimize the required crane capacity.

In Sect. 3 we conduct two representative experiments on realistic data to show that considerable costs may be involved in the current commercially driven allocation schedules.

3 Case study

We consider three interacting terminals ($T = 3$) in the port of Antwerp, where thirty-seven vessel lines have one of their vessels processed exactly once a week ($V = 37$). Furthermore, we assume that each vessel line has a preferred terminal and a preferred arrival and departure time, which fit best to their schedules (the current berth allocations). By these experiments the induced costs for the required number of quay cranes and inter-terminal transport are made explicit.

We are interested in potential reductions in required crane capacity and inter-terminal transport if we allow small modifications to the current allocations constructed by PSA Antwerp, which are primarily driven by commercial considerations. These reductions show how much additional container handling capacity is in principle still available without having to invest in new cranes. A simple visual analysis of these terminals (without any computation at all) reveals that each terminal shows heavy fluctuations in the workload distribution over the shifts of the week. At busy shifts, the entire quay crane capacity is fully occupied, whereas at quiet shifts no crane is working. A second observation is the relatively large number of transshipment containers that have to be transported from one terminal to another due to the allocation of connecting vessels to different terminals. (For the purposes of this paper it is not important who actually pays for this inter-terminal traffic. For example, a terminal's carbon footprint may be considered as a cost that needs reduction.)

As an illustration Fig. 3 depicts the number of active quay cranes (scaled to 1) in one of the terminals for each hour of a weekly cycle ($K = 168$) according to an allocation representative for the situation in Antwerp. Note that the scaled total number of quay cranes available in the terminal can be larger than 1 (implying that never all available cranes are active at the same time). The black line represents the

Fig. 3 Current number of quay cranes active in terminal 1 during the hours of a 1-week-cycle
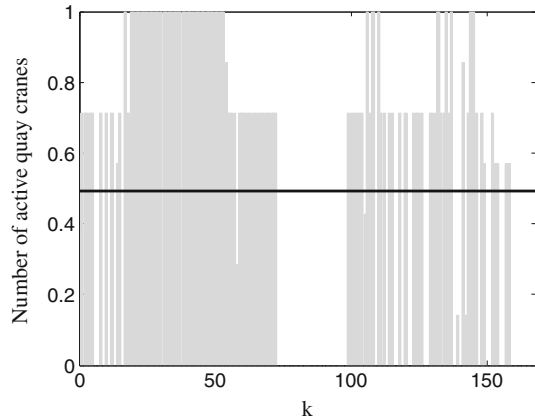
mean number of active quay cranes per hour. From the high fluctuations in the number of active quay cranes we conclude that i) for a few time slots a high amount of the crane capacity is needed and ii) during a lot of time slots a large percentage of crane capacity is not used.

The experiments have been conducted on an Intel-based quad core Linux server with a CPU-speed of 2.4GHz, and 4GB of RAM. In the first step, with 21 time slots of 8 h each, the number of binary variables is 888 and the number of constraints is a little over 2,300. In the second step, we solve the problem per terminal with 168 time slots of 1 h each per cycle. Both the number of binary variables and the number of constraints are approximately 2,200. Solving the first step takes a few minutes, solving the second step takes less than 1 s. The solver used is CPLEX 11.

In the first experiment, our hypothesis is that a slight modification of the allocation of only a couple of vessels already leads to large cost reductions. Hence, for a small selection of the vessels, considered by the terminal operator as practically feasible, we allow (1) a terminal allocation different from the preferred one, and (2) the berthing interval different from the preferred one. We allow the berthing interval to shift maximally G_v time slots to the left or to the right on the time axis as depicted in Fig. 2. For different values of these bounds, and for different ratios between the parameter value c_{st} and c_t in Eq. 25, we run the two-step optimization and from the results construct Pareto frontiers. Each Pareto frontier shows a front of optima for which the bound G_v is the same, but the ratio between c_{st} and c_t is different. It is a convenient way to visualize the trade-off between conflicting objectives.

In these Pareto frontiers the total number of quay cranes (over all terminals) is scaled to the number of quay cranes maximally active in a representative allocation in the three terminals in Antwerp. One of the points from one of the Pareto frontiers is highlighted to further illuminate the hypothesis. This allocation is the result of a two-step optimization after selecting less than one third of the vessels and allowing a change in their terminal allocation and a shift in the time position of the berthing interval of these vessels of no more than 24 h ($G_v = 24$). The results of the two-step optimization suggest that about 25% of the number of quay cranes can be saved

while at the same time the inter-terminal transportation costs are reduced by 3%. Although there might be proper (commercial) reasons for using the current schedules, at least with our approach, we can quantify the additional costs induced.

In the second experiment, we assume that the preferred arrival and departure times for all vessels are satisfied ($G_v = 0, \forall v$), but that they may be berthed at a terminal different from their preferred one. Consequently, the terminal allocation of *all* vessels is chosen to be flexible while the allocations of their berthing intervals are fixed to the existing one. Results suggest (see Fig. 6) that with the same number of quay cranes as in the current allocation, 40% of the costs for inter-terminal transport can be saved by adapting the terminal allocation.

The details and a discussion of the two experiments are presented in the next two sections.

3.1 Berthing time and terminal re-allocation

In this experiment we arbitrarily select a total of 11 vessels (out of the total set of vessels) from the busy peaks in terminals 1, 2 and 3 (busy peaks in terminal 1 for instance can be found simply by observing Fig. 3), and define this set to be S. For each value of $G_v \in \{0, 8, 16, 24, 48\}$ (in hours) for the vessels in S, a Pareto frontier of the total (sum over all terminals) number of active quay cranes versus the total costs for inter-terminal transport is constructed. This means that for a certain black or grey frontier G_v is the same, and each point on that frontier results from a single two-step optimization with a specific ratio between costs for quay cranes and costs for inter-terminal transport.

The results are depicted in Fig. 4a. As G_v increases, the frontiers move from the upper right to the lower left in this figure. That means that the black frontier starting at the upper right corresponds to the case where $G_v = 0$, the grey frontier below that one corresponds to the case where $G_v = 8$, the black frontier in the middle (overlapping with the grey ones) to the case where $G_v = 16$, the lower grey one to the case where $G_v = 24$, and the lower black one (almost fully overlapping with the grey one) to the case where $G_v = 48$. The cross represents the state of an allocation representative for Antwerp, i.e. it represents the number of quay cranes maximally active (possibly smaller than the number of quay cranes available) and the amount of inter-terminal transport in an allocation representative for Antwerp. From Fig. 4a we conclude:

- The frontier for the case where $G_v = 0$ already shows potential improvements in the amount of inter-terminal transport (on the frontier at the left side of the cross).
- In the same situation, i.e. $G_v = 0$ for $v \in S$, a reduction in the number of quay cranes is possible at the expense of higher inter-terminal costs transport (on the same frontier at the lower-right side of the cross). This means that even without changing the berthing intervals ($G_v = 0$ for $v \in S$), the costs for inter-terminal transport or the number of required quay cranes can be reduced. Apparently, an adaptation in the terminal allocation of the vessels in S is sufficient to achieve this.
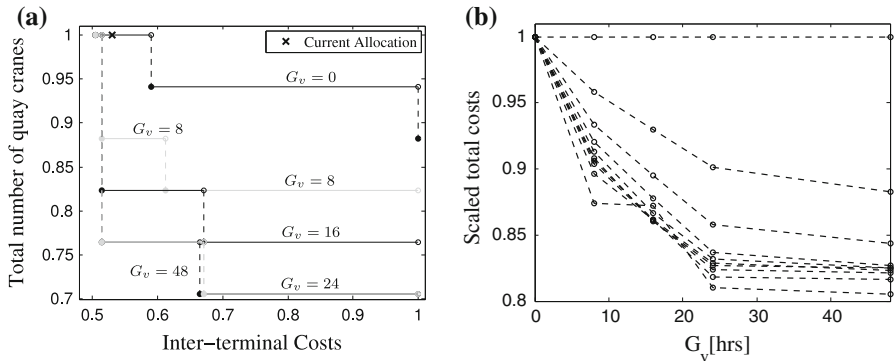
Fig. 4 Improvements by allowing certain flexibilities. **a** Pareto frontiers, for $G_v \in \{0, 8, 16, 24, 48\}$. As G_v increases, the front moves from the *upper right* to the *lower left* corner. **b** Costs versus flexibility level G_v for different cost ratios

- The improvements going from $G_v = 0$ to $G_v = 8$ are relatively large, whereas the improvements going from $G_v = 24$ to $G_v = 48$ are approximately zero (the lower grey ($G_v = 24$) and the lower black ($G_v = 48$) mostly overlap).
- All fronts intersect (at the upper left point) where the parameter for crane costs c_t are zero. Apparently, the inter-terminal costs cannot be further reduced even if G_v grows and the maximum number of cranes is used.
- The grey bullet [coordinates (0.51, 0.76)] on the lower grey frontier suggests that if $G_v = 24$ for $v \in S$ the number of maximally active quay cranes can be reduced by almost 25% and the costs for inter-terminal transport by about 3%. This means that besides a possible change in terminal allocation, the time allocation of only 11 vessels has to be shifted at most 1 day to gain significant improvements.

It is interesting to visualize the difference between the representative allocation in Antwerp (black cross in Fig. 4a) and the allocation represented by the grey bullet in the lower gray frontier having coordinates (0.51, 0.76) in a different way: Fig. 5a, c and e represent the scaled number of active quay cranes for each of the three terminals for each hour in the 1-week-cycle. Figure 5b, d and f represent the number of active quay cranes in each of the three terminals for each hour in the 1-week-cycle for the allocation found by using the two step optimization [i.e. the optimum represented by the grey bullet having coordinates (0.51, 0.76)]. The black lines represent the mean quay crane usage per hour in the different terminals. If we compare Fig. 5a, c, e with Fig. 5b, d, f, we notice the following:

- The workload in the generated allocation is better balanced (over the terminals and over time) than in the current allocation. This results in the previously mentioned reduction of almost 25% of the required quay cranes.
- At some points in time still some quay cranes are not active even in the allocation generated by the two step optimization. Introducing either a higher level of flexibility (by increasing G_v, $v \in S$) or including more vessels into the
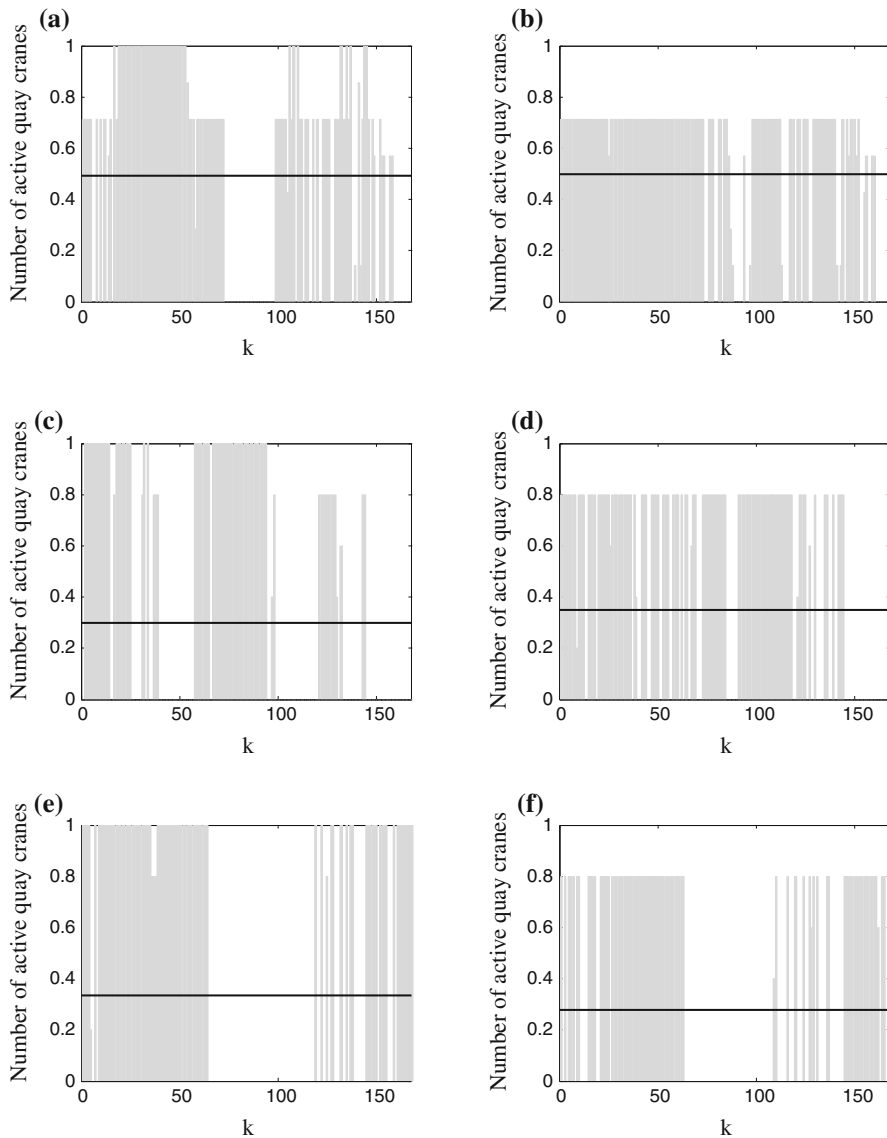
Fig. 5 Current and generated quay crane usage during the hours of a 1-week cycle. **a** Current allocation for terminal 1. **b** Generated allocation for terminal 1. **c** Current allocation terminal 2. **d** Generated allocation terminal 2. **e** Current allocation for terminal 3. **f** Generated allocation for terminal 3

set S would probably fill up these gaps and lead to an even better workload balance and a smaller number of required cranes.

- The mean quay crane usage in a specific terminal can differ for the current allocation and the generated allocation. This can be explained by a difference in terminal allocation of the vessels in S. The total quay crane usage however remains the same.

Additionally, we depict the benefit of modification differently. For a constant ratio of quay crane costs and inter-terminal costs, the scaled total costs are plotted versus the level of flexibility G_v. Figure 4b presents the results for 10 ratios ($\frac{c_t}{c_{st}} \in \{0, 20, 40, 60, 80, 100, 120, 140, 160, \infty\}$). The "frontiers" move downwards with increasing ratio. So the upper (horizontal graph) represents the results for the costs ratio 0, the graph below that one represents the results for the ratio 20, and so on. From this figure we notice:

- For the ratio equal to 0 (no costs for quay cranes), the total costs (for this ratio only inter-terminal costs) are not affected by the level of flexibility. Apparently, the terminal length and crane capacity are not the limiting factor so that the best (lowest inter-terminal traffic) solution is already obtained at $G_v = 0$.
- For all ratios larger than zero the total costs decrease as the level of flexibility increases.
- As the ratio grows between 0 and 40, the total cost decrease becomes relatively large. Apparently, these are the sensitive ratios, where a slight increase of the crane costs already leads to a significant reduction in quay cranes. This suggests that the current number of quay cranes is far from minimal.
- Increasing the cost ratio beyond 40 does no longer lead to substantial cost savings. A further increase of quay crane costs does no longer affect possible cost reductions.

3.2 Terminal re-allocation for fixed timetable

In this experiment we assume that none of the vessel lines is prepared to change their berthing interval in Antwerp, but we also assume that each of them allows a different terminal allocation. We are interested in decreasing the current costs for inter-terminal transport, while the current arrival and departure times are kept the same. Additionally, we require that the maximum number of quay cranes active in a terminal is at most equal to the number of quay cranes currently present in that terminal. Hence, we allow a terminal adaptation for each vessel and fix its berthing interval in time to the current allocation ($G_v = 0$). Figure 6 shows the cumulative costs for inter-terminal transport for the current allocation and the generated allocation for each hour in the weekly cycle. The costs are scaled to the total costs in the current allocation. These results suggest that, with the same number of quay cranes, about 40% of the costs for inter-terminal transport can be saved.

4 Conclusions and recommendations

We considered a port consisting of a number of inter-acting terminals operated by one terminal operator. The strategic problem faced by the operator is to allocate vessels to a terminal for a certain time interval. We abstracted from position and quay crane allocations (problems faced at tactical and operational levels, respectively) and constructed a MILP to strategically allocate a terminal and a time window to each of the vessels in the cycle.
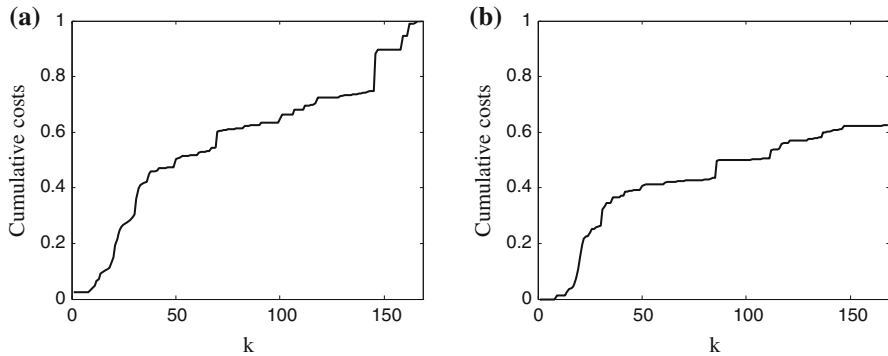
Fig. 6 Current and generated cumulative inter-terminal costs for a 1-week cycle. **a** Inter-terminal costs for the current allocation. **b** Inter-terminal costs for the generated allocation

Since the ability to solve real-life problems within reasonable time strongly depends on the resolution of the time grid, a two-step optimization approach was introduced. This approach was applied to perform a case study for the multi-terminal container operator PSA Antwerp, who operates multiple terminals in the port of Antwerp. The two-step approach enabled us to efficiently investigate the benefit of modifying an existing allocation, i.e. the potential crane and inter-terminal transport savings if the existing terminal and time allocations would be modified. Phrased differently, it shows the capability to accommodate additional traffic without having to increase the existing crane capacity.

Results suggest that a small modification of an existing allocation suffices to gain significant improvements: a reduction of almost 25% of the number of cranes and at the same time a reduction of more than 3% of the inter-terminal costs. Furthermore, if the current terminal allocation of all vessels can be modified while the current time allocation is fixed, costs for inter-terminal transport can be reduced by 40%. Although the model does not incorporate costs for the modifications, at least this tool enables us to quantify the cost savings or spare capacity that can be achieved by these modifications.

References

Bierwirth C, Meisel F (2010) A survey of berth allocation and quay crane scheduling problems in container terminals. Eur J Oper Res 202:615–627

Cordeau J-F, Laporte G, Legato P, Moccia L (2005) Models and tabu search heuristics for the berth-allocation problem. Transp Sci 39:526–538

Giallombardo G, Moccia L, Salani M, Vacca I (2010) Modeling and solving the tactical berth allocation problem. Transp Res Part B 44:232–245

Guan Y, Cheung RK (2004) The berth allocation problem: models and solution methods. OR Spectr 26:75–92

Hansen P, Oğuz C, Mladenović N (2008) Variable neighborhood search for minimum cost berth allocation. Eur J Oper Res 191:636–649

Hendriks MPM (2009) Multi-step optimization of logistics networks. Strategic, tactical and operational decisions. PhD thesis, Eindhoven University of Technology

Hendriks MPM, Laumanns M, Lefeber E, Udding JT (2010) Robust cyclic berth planning of container vessels. OR Spectr 32:501–517

Hendriks MPM, Lefeber E, Udding JT (2011) Simultaneous berth allocation and yard planning at tactical level (in preparation)

Imai A, Nagaiwa K, Chan WT (1997) Efficient planning of berth allocation for container terminals in Asia. J Adv Transp 31:75–94

Imai A, Sun X, Nishimura E, Papadimitriou S (2005) Berth allocation in a container port: using a continuous location space approach. Transp Res Part B 39:199–221

Imai A, Nishimura E, Papadimitriou S (2008a) Berthing ships at a multi-user container terminal with a limited quay capacity. Transp Res Part E 44:136–151

Imai A, Chen HC, Nishimura E, Papadimitriou S (2008b) The simultaneous berth and quay crane allocation problem. Transp Res Part E 44:900–920

Kim KH, Moon KC (2003) Berth scheduling by simulated annealing. Transp Res Part B 37:541–560

Lim A (1998) The berth planning problem. Oper Res Lett 22:105–110

Lim A, Rodrigues B, Xiao F, Zhu Y (2004) Crane scheduling with spatial constraints. Nav Res Logist 51:386–406

Meisel F, Bierwirth C (2009) Heuristics for the integration of crane productivity in the berth allocation problem. Transp Res Part E 45:196–209

Monaco MF, Sammarra M (2007) The berth allocation problem: a strong formulation solved by a Lagrangean approach. Transp Sci 41:265–280

Moorthy R, Teo C-P (2006) Berth management in container terminal: the template design problem. OR Spectr 28:495–518

Ottjes JA, Veeke HPM, Duinkerken MB, Rijsenbrij JC, Lodewijks G (2006) Simulation of a multiterminal system for container handling. OR Spectr 28:447–468

Papadimitriou CH, Steiglitz K (1982) Combinatorial optimization: algorithms and complexity. Prentice-Hall, Englewood Cliffs

Park Y-M, Kim KH (2003) A scheduling method for berth and quay cranes. OR Spectr 25:1–23

Pinedo M (1995) Scheduling: theory, algorithms and systems. Prentice-Hall, Englewood Cliffs

Rashidi H, Tsang EPK (2006) Container terminals: scheduling decisions, their formulation and solutions. J Sched (submitted)

Stahlbock R, Voß S (2008) Operations research at container terminals: a literature update. OR Spectr 30:1–52

Steenken D, Voß S, Stahlbock R (2004) Container terminal operation and operations research—a classification and literature review. OR Spectr 26:3–49

Vis IFA, de Koster R (2003) Transshipment of containers at a container terminal: an overview. Eur J Oper Res 147:1–16

Wang F, Lim A (2007) A stochastic beam search for the berth allocation problem. Decis Support Syst 42:2186–2196

Author Biographies

M. P. Maarten Hendriks received his MSc degree in Mechanical Engineering from Eindhoven University of Technology, The Netherlands, in 2004. In 2009 he received his PhD degree on the subject of multi-step optimization of logistics networks within the Systems Engineering group of Eindhoven University of Technology, The Netherlands. He is currently working as a senior consultant for the company Quintiq, a leading provider of Advanced Planning and Scheduling (APS) software solutions.

Dieter Armbruster received the PhD, degree in physics from the Universität Tübingen, Germany, in 1980. He was a postdoc in mathematics and in theoretical and applied mechanics at Cornell University, Ithaca, NY. He is a Professor in the School of Mathematical and Statistical Sciences, Arizona State University, Tempe, and a part time Professor in Mechanical Engineering at Eindhoven University of Technology, Eindhoven, The Netherlands. His research interests are broad based and range from dynamical systems theory and chaos to the dynamics of complex networks and production systems and the simulation and control of semiconductor fabs.

Marco Laumanns is a Research Staff Member in the Business Optimization Group of the IBM Zurich Research Lab. He holds a diploma degree in Computer Science from the Technical University of Dortmund (1999) and a PhD in Electrical Engineering from ETH Zurich (2003). Since 2010 he is a private lecturer for Operations Research in the Department of Management, Economics, and Technology at ETH Zurich. His main research interests are in optimization and control, in particular robust and stochastic optimization for decision making under uncertainty.

Erjen Lefeber received his MSc in Applied Mathematics from the University of Twente, Enschede, The Netherlands, in 1996. In 2000 he received his PhD from the University of Twente on the subject of tracking control of nonlinear mechanical systems. Since 2000, he has been an Assistant Professor at the Manufacturing Networks group, at the Department of Mechanical Engineering at Eindhoven University of Technology. His current research interests include modeling and control of manufacturing systems.

Jan Tijmen Udding received his MSc in Applied Mathematics from Eindhoven University of Technology, The Netherlands, in 1980. In 1984 he received his PhD degree in computing science from the same university on the design of delay-insensitive circuitry. He has spent his professional career both in academia, most recently as extraordinary professor at the Systems Engineering group, Eindhoven University of Technology, and the logistics industry, most recently as CIO of the container terminal operator PSA Antwerp. His research interests include optimization of logistic processes, concurrency, and programming.