Flex Serv Manuf J (2017) 29:433–503
DOI 10.1007/s10696-017-9278-6

Re-marshalling in automated container yards with terminal appointment systems

Filip Covic[1]

Published online: 3 February 2017

Abstract As a result of scarce land availability, growing competition and throughput, container terminals are increasing the stacking height of yard blocks to fulfil the demand for storage area. Due to inadequate retrieval information at initial stacking, shuffle moves can occur during retrieval operations as containers may be stacked in a sequence which does not correspond to the actual retrieval sequence. Automated stacking cranes can perform re-marshalling during periods of no crane workload to shift unproductive moves during retrieval operations to phases of idle time. Terminal appointment systems (TAS) enhance landside sequence information when external trucks (XT) announce their arrival beforehand. Under these circumstances, it is beneficial for terminal planers to understand the effects of using re-marshalling in combination with TAS. The purpose of this work is to introduce an online rule-based solution method for the re-marshalling problem with and without TAS. A simulation model of a fully operating yard block is used as environment to compare the proposed method with a benchmark heuristic from the literature. All tests are conducted for single and multiple Rail-Mounted-Gantry-Crane systems with different yard block sizes. It is also shown that solving the re-marshalling problem with the proposed algorithm generates results that reduce shuffle moves by 30% on average and by up to 50% in the best case, while always performing better in the worst case in comparison with not performing re-marshalling. Afterwards, influences on the method of selected TAS parameters are evaluated numerically. Results show that imprecise XT arrival information, not deviating above a certain threshold, significantly contribute to reducing congestion by mitigating XT waiting time and levelling arrival peaks. These benefits can be achieved without imposing restrictions on the arrival schedule preferred by XT companies.

✉ Filip Covic
filip.covic@uni-hamburg.de

[1] Institute for Operations Research, HBS Hamburg Business School, University of Hamburg, Moorweidenstr. 18, 20146 Hamburg, Germany

Keywords Container terminal · Automated stacking cranes · Re-marshalling · Terminal appointment system · Expert systems · Simulation

1 Introduction

In 2014, 80% of global trade volume was moved by deep sea vessels and seaport container terminals making the container terminal the central point for the handling of cargo worldwide (UNCTAD 2015). Nowadays, terminals face problems of scarce land availability at seaports, an intense competition between terminal operators and increased throughput. In a typical terminal, numerous containers arrive at the same time while numerous others should leave the terminal. With a limited amount of equipment (e.g. cranes, trucks), resources and personnel a terminal manager should devise an operational plan for the terminal by deciding about the following issues at the same time in order to have the best achievable performance: (1) yard crane scheduling, (2) container stacking, (3) yard crane routing and (4) relocation of containers. These issues are interrelated and by deciding about one of them, there is a limited degree of freedom to decide about the others. Unfortunately in most of the cases, a partial solution to each of these issues may not match each other or, if they match by a small chance, there would be a poor terminal plan.

One consequence of a poor terminal plan is the increase of the stacking height within a yard block to fulfil the demand for storage area by expanding the block capacity (Kemme 2012a). However, this approach leads to several difficulties when retrieving a container from the block. Depending on the (technical) design and layout of the yard block, typical stack heights of up to 1-over-6 can be observed while only the topmost container can be accessed directly (Brinkmann 2011). With increasing stack height, the risk of unproductive moves (shuffle moves) during retrieval operations becomes imminent as containers may be stacked in a sequence that does not correspond to the actual retrieval sequence. A *mis-overlay* is defined if a container A is stacked on top of another container B while container A leaves the yard block later than container B. In this paper, a method for re-marshalling is proposed to overcome the high number of shuffle moves and the resulting vehicle waiting times at handover areas.

Moreover, it is rarely possible to stack containers in an appropriate order during initial storage because of non-existent or incomplete information about their future retrieval times. The lack of adequate information is apparent especially for containers incoming or outgoing by external trucks (XTs) as their arrivals are perceived as almost completely random by the terminal operator (Dekker et al. 2007). Thus, the ensuing mis-overlays create shuffle moves for containers being picked up by all vehicle types and in turn induce waiting times for all jobs as a result of misused crane capacities. Additionally, the arising traffic congestion caused by XTs in and around the terminal has become a prominent issue in terms of environmental pollution in past years (Morais and Lord 2006). For this purpose, this work examines a combination of a container handling method with a system for vehicle arrival management to improve the processing of XTs in the terminal.

Seaports around the world are trying to mitigate these problems by implementing a diverse range of advanced technological systems to speed up XT gate processing, spreading XT arrivals over time and improving container storage and retrieval information. The *terminal appointment system (TAS)*[1] is suggested in practical applications and scientific literature as one method to cope with inadequate arrival information of XTs. By applying TAS, XT companies book a certain time slot by announcing their arrival at the terminal a predefined period beforehand. This enhancement of information can be utilised by terminals to adjust their operation and, as a consequence, streamline container handling. The major resulting problem within the yard block concerns the rearrangement of containers in a block to match the XT arrival sequence obtained by the TAS. This problem can be defined as the *re-marshalling problem (RMP)* which needs to be solved continuously while yard cranes, which handle containers in the block, become idle. The goal of this paper is to propose a solution method for the RMP that can improve the container arrangement without additional TAS information while also being able to incorporate it if available. Based on this approach, a detailed analysis of TAS parameters and their influence on the objectives is conducted.

The remainder of the paper is structured as follows: Sect. 2 gives an introduction to terminal design and operational considerations relevant for this study. It defines the basic working principle of a TAS and formalises the RMP and related objectives. A short literature review is provided in Sect. 3. Next, Sect. 4 elaborates on the methodology used to solve the RMP and the modelling of its testing environment. Section 5 states the experimental design and discusses results obtained. Finally, Sect. 6 concludes the paper and gives an outlook on future research.

2 Problem description

This section briefly describes the characteristics of the block design and crane systems addressed by the developed solution method. Subsequently, features and parameters of a TAS are discussed.

2.1 Problem environment

The RMP as subsequently defined arises typically in front-end yard blocks with automated Rail-Mounted-Gantry-Cranes (RMGCs). The low operating costs of RMGCs make them suitable for performing re-marshalling jobs during their idle times. Besides Single-RMG-systems (SRMG), multi-RMG-systems can be operated per block to increase overall productivity. In this context, Twin-RMG-systems (TRMG), Double-RMG-systems (DRMG) and Triple-RMG-systems (TriRMG) can be operated in these blocks. However, it should be noted that this leads to increased planning complexity for routing the cranes in order to avoid collisions and mitigate

[1] The term “Vehicle Booking System (VBS)” may also be encountered in literature (e.g. Davis (2009)).

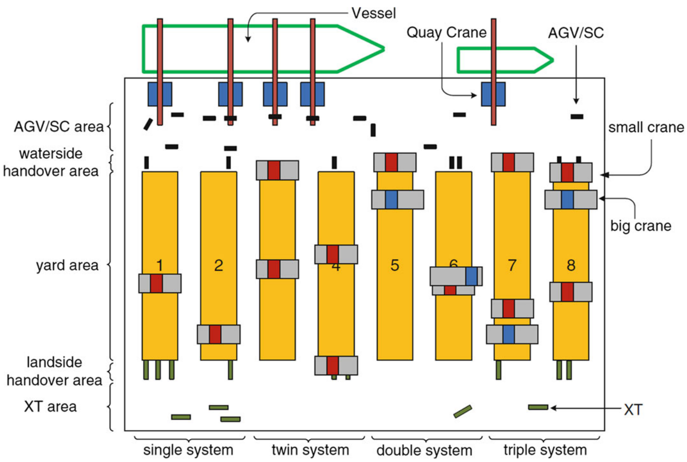

Fig. 1 Overview of a front-end terminal with S-, T-, D- and TriRMG (Kemme 2012b, p. 568)

blocking time. An illustrative overview of the container terminal system with all four RMG-types employed throughout this paper can be obtained from Fig. 1.

The Terminal Appointment System (TAS) can be integrated into this container terminal environment by the Terminal Operating System (TOS), which manages all terminal operations, strategies and container data, to make use of improved container arrival and retrieval times. XT companies can book a time slot for delivering or picking-up containers via web-interface. The terminal operator must set the framework for these bookings by stipulating specific parameters like the length of the time slot wherein XTs may arrive and the XT announcement time prior to its arrival. In general, two usage cases of the TAS can be considered:

– Reducing vehicle waiting time by increasing yard productivity through improved arrival time information (the case analysed here).
– Levelling the distribution of XT arrivals throughout the day by imposing a maximum number (capping) of XTs per time slot to avoid long phases of peak load and idle time.

2.2 Problem formulation

The aforementioned mis-overlays of containers within a stack occur due to incomplete information about the retrieval sequence at the time of storage. Without re-marshalling, i.e. no rearrangement by the yard cranes during storage in the block, shuffle moves must be performed at the time of container retrieval to access a container which leaves the terminal before the container on top of it. As a consequence, these moves decrease the crane productivity regarding storage and retrieval jobs resulting in longer vehicle waiting times. While a container dwells in a block, the terminal may receive updated information about it. By considering this

Table 1 Performance indicators for algorithm evaluation

Type	Symbol	Indicator description
Container handling	$\boldsymbol{n}_{shu}$, ($\boldsymbol{N}_{shu}$)	Number of shuffle moves per container, (total number of shuffle moves)
	$\boldsymbol{n}_{rm}$, ($\boldsymbol{N}_{rm}$)	Number of re-marshalling moves to eliminate a shuffle move, (total number of re-marshalling moves)
Vehicle waiting times at handovers	$\overline{\omega}_{tot}$	Mean total vehicle waiting time
	$\overline{\omega}_{lsout}$, $\overline{\omega}_{wsout}$	Mean XT and AGV waiting time for retrieval
	$\overline{\omega}_{lsin}$, $\overline{\omega}_{wsin}$	Mean XT and AGV waiting time for delivery
Standard deviation of vehicle waiting times	σ_{tot}	Standard deviation of total vehicle waiting time
	σ_{lsout}	Standard deviation of XT waiting time for retrieval

information, periods of no yard crane workload can be utilised to shift unproductive moves during retrieval operations to intervals when yard cranes are idle.

Due to the uncertainty of vehicle arrivals, it is not possible to determine the beginning or end of idling periods in which it is possible to perform re-marshalling. These periods occur multiple times a day with varying period lengths. Thus, this problem is an *online* problem that needs to be solved continuously and frequently throughout a planning period.

Formally, the RMP is a combinatorial (re-)optimisation problem that can be defined as follows:

Definition 1 Find a minimal sequence of container marshalling moves within a yard block in order to transform an initial block configuration into a target configuration so that the number of shuffle moves (and crane moving distance) during future retrieval operations is minimised.

In order to evaluate algorithm performance and TAS parameter sensitivity, the systematisation of performance indicators is used as shown in Table 1.

Firstly, a combination of RMP and TAS primarily addresses the elimination of mis-overlays by a minimal number of preemptive relocation moves during the time cranes are idle. Thus, container handling metrics are included based on single container and shuffle moves to enable a more general assessment independent of the planning horizon or block layout. In this context, the number of re-marshalling moves to eliminate a shuffle move n_{rm} is used as a measure for the re-marshalling quality to account for finding the minimal sequence of re-marshalling moves with respect to Definition 1. Secondly, the superordinate goal is to reduce vehicle waiting time at the handovers of the block. For this purpose, the total vehicle waiting time is split into vehicle and job types to examine pronounced algorithm and TAS influences. In this context, the proposed solution approach is designed to improve XT waiting time while maintaining priority on waterside jobs.[2] Hence, there is a

[2] Shipping carriers operating deep sea vessels are considered as the main customers by container terminals. The priority to minimise their berthing times is highest (Kemme 2012a; Davis 2009).

particular focus on (XT) retrieval jobs which are influenced by shuffle moves and TAS information more strongly. Thirdly, the standard deviation of total vehicle waiting times is used as secondary performance indicator to evaluate potential impacts on crane workload balance. Although the TAS usage case analysed in this work does not employ capping in order to balance crane workload, there might be effects induced by an increased yard productivity without imposing any schedule for XT arrivals.

In this context, the leading indicator to evaluate any algorithm performance should be based mainly on the superordinate goal to minimise vehicle waiting times at the handover areas. This goal is the primary input from yard block operation to increase overall terminal productivity as the handover areas constitute the interface between the yard block and the remaining terminal operation. However, the vehicle waiting times at the handover areas are directly dependent on the container configuration of the yard block which is essentially determining the yard crane workload and, thus, the ability of the cranes to serve the handover areas efficiently. As a result of the RMP being a yard block problem, the proposed solution method for the RMP in this work does not directly target the minimisation of vehicle waiting times but the operationalised objective of reducing mis-overlays which result in shuffle moves of the yard crane. A strong correlation between the number of shuffle moves and vehicle waiting times at handover areas is ascertained by Kemme (2012a). Moreover, the performance indicator for the re-marshalling quality helps to assess a good re-marshalling schedule which needs fewer re-marshalling moves to eliminate the same number of shuffle moves. Based on the discussion of performance indicators above and in Table 1, the evaluation *hierarchy of performance indicators* can be stated in order to determine the superiority of one solution method over another.

Hierarchy of performance indicators for terminal productivity: $\overline{\omega}_{tot} \succ \overline{\omega}_{wsout} \succ (\overline{\omega}_{lsout}) \succ n_{shu} \succ n_{rm}$
Directly targeted and operationalised objectives by the proposed algorithm: n_{shu}, n_{rm}

Another issue which must be acknowledged concerns the interruption of the re-marshalling schedule. For example, if a waterside storage job is announced during or immediately after the execution of a job close to the landside, re-marshalling may lead to artificial prolongation of the AGV waiting time at the handover as the crane must finish the re-marshalling job first and move the entire distance from landside to waterside afterwards. SRMG-systems with high workload are particularly affected by this due to overall low crane resources and the inability to split the workload. Note that the RMP is not restricted to one bay or a SRMGC like the related pre-marshalling problem (PMP) or the blocks relocation problem (BRP) (see literature review in Sect. 3). Therefore, further criteria like the minimisation of a stored container's distance to its respective handover area and considerations how to leave enough suitable slots open for future re-marshalling and stacking operations must be covered by a solution approach as well.

3 Literature review

The two-fold structure of the analysis calls for a structured discussion of current research which distinguishes between studies chiefly focusing on TAS or literature predominantly centring on the RMP.

3.1 Terminal appointment system

The optimisation of landside operation has experienced increased attention in recent years. It unveils unrealised potentials for improving terminal performance as means to cope with growing competition and regulatory requirements for mitigating traffic congestion (Morais and Lord 2006). Initially, TAS research was qualitatively focused observing several difficulties in application: A comprehensive study of TAS is conducted by Morais and Lord (2006) who employ surveys, field visits and quantitative analysis of greenhouse gas emissions caused by XT queuing to recommend a suitable TAS design for reducing traffic congestion at terminal gates. They conclude that a TAS is essential for balancing terminal capacity by spreading XT arrivals. With reference to the case study of the ports in Los Angeles and Long Beach, Giuliano et al. (2007) analyse the effectiveness of the TAS and related parameters on reducing XT turnaround time. They conduct surveys with terminal and XT company staff to assess the general perception and usefulness of the implemented measures. Maguire et al. (2010) examine different strategies to reduce congestion at terminal gates. By reviewing various implementations in North America, they identify the slot length and the flexibility of appointment handling as important factors for the practicability of a TAS to reduce congestion and minimise the stochastic elements of yard operation. All studies generally concur that XT companies experience obstacles adopting an appointment system, partly due to a reduction of their planning flexibility but also as a result of poor TAS implementation and insufficient usage of arrival data for planning by terminal operators (Davis 2009).

Quantitative TAS research can be divided according to the operational area of the terminal it primarily relates to: firstly, the hinterland operation or drayage is affected through XT routing and scheduling decisions. XT dispatching must be planned to satisfy customer demand at respective locations while incorporating TAS time slots or caps as additional constraints. For this purpose, Namboothiri and Erea (2008) develop a heuristic based on column generation that minimises transportation costs either by minimising the number of deployed XTs or the number of slots used at the terminal. They conclude that access capping at terminal gates increases XT companies' planning complexity while the length of the time slot and cap size has a significant productivity impact on XT operation. Phan and Kim (2015) examine the negotiation process between XT companies and terminal operators for booking a time slot through TAS. They develop a mathematical programming formulation for centralised and decentralised decision making in order to reduce congestion at terminal gates while minimising the impact on the planning flexibility of XT companies. The decentralised approach is numerically shown to provide best results and lowest computational time. The second field is concerned with optimising terminal gate operation to streamline entry and exit of containers by XT. Typically, queuing theory is applied to model gate transactions, utilisation and XT waiting time. For instance, Zhang et al. (2013) and Chen et al. (2013) use a non-stationary queuing model for gate operation and, afterwards, optimise TAS capping by a Genetic Algorithm. The latter evaluate two types of TAS: a static system based on an input of intended XT arrivals calculated a day before and a dynamic system which substitutes the input with a real-time web-interface for XT appointment requests and terminal acceptance or postponement. It is stated that appropriate capping can reduce gate queuing significantly while the dynamic TAS leaves more planning flexibility to XT companies. Similarly, Guan and Liu (2009) apply a multi-server queuing model to evaluate XT waiting time at terminal gates. First, gate operation is analysed without a TAS through sensitivity analysis. Afterwards, the application of a TAS to this model substantially reduces XT waiting times and costs in comparison to the case without TAS. Finally, enhanced XT arrival information may have considerable influence on yard block operation. Huynh (2005) analyses a yard block in order to evaluate TAS impact on yard crane utilisation and XT turnaround time. His study mainly focuses on applying simulation-optimisation to find an appropriate capping per time slot which is robust to XT no-shows. Results show that crane workload smoothing can contribute to lower turnaround times when the cap is chosen within specific intervals. Zhao and Goodchild (2010) use simulation to assess the impact of different degrees of information quality for container re-handling. They show that with partial XT arrival information the number of shuffle moves can be reduced significantly. Lastly, Zehendner and Feillet (2014) propose a deterministic multi-commodity min-cost flow model and a discrete-event simulation model to ascertain the optimal number of appointments per slot based on crane workload and available crane capacity. Simulation experiments demonstrate a reduction of vehicle waiting time for all modes as a result of smoothed XT arrivals through capping with mandatory bookings.

3.2 Re-marshalling problem

Research on the RMP is still scarce in comparison to its related problems, the PMP (Lee and Hsu 2007) and BRP (Caserta et al. 2012). Possibly, this results from the diverse scope of addressing the RMP by the literature. Depending on the technical design of the terminal, various modelling approaches regarding yard block layout, stacking restrictions, container types and objectives to be minimised can be observed. As a result, a formal and linking description addressing the RMP, for instance, in the form of a holistic mathematical programme has not been proposed in the literature. As the PMP and BRP are restricted to a bay by definition, a formalisation independent of the block design is feasible whereas it seems more challenging to cover all aspects within one (mathematical) model for the RMP. Moreover, assumptions about the availability of container retrieval times are inconsistent, where many publications assume complete information to demonstrate the theoretical effectiveness of their proposed solution approach. Lastly, there is variation in the literature concerning the considered interaction with other strategies in the yard block and their influence on the input for re-marshalling. In due regard to these irregular problem specifications, a numerical comparison of the few RMP solution methods found in the literature is exacerbated. Still, Yu et al. (2009) decompose the RMP into two separate tailored integer programming formulations (IP) for stacking and moving containers, respectively. However, only a target bay for the container to be re-marshalled is identified and not a concrete stack. Further approaches to solve an IP formulation of a variant of the RMP could not be found in the literature. One reason for neglecting this approach is the underlying complexity of the RMP which Caserta et al. (2011) show to be NP-hard by stating that the BRP is NP-hard and that the RMP with a single crane is a generalisation of it. A similar decomposition of the problem only for export containers with fixed quay crane assignment is performed by Park et al. (2009) who apply a co-evolutionary algorithm to integrate the solution of one problem into the calculation of the other. The heuristic by Kemme (2012a) identifies a container to be re-marshalled by the current distance from its target handover area and stacks the chosen container using the same category stacking he applies to initial container storage. Kang et al. (2006) apply partial order graphs and simulated annealing to create a re-handling free ordering of stacks. However, this approach requires an empty target stack in the block or terminal where the containers can be re-marshalled to. Asperen et al. (2013) evaluate the behaviour of stacking rules employed for re-handling under different degrees of available XT look-ahead times. They conclude that many basic stacking strategies can benefit from arrival announcement of only 4h in advance to reduce shuffle moves. In contrast, more sophisticated strategies approximating or estimating XT arrival times do not benefit from these announcements noticeably. Nonetheless, they do not provide a systematised identification of the container to be relocated. Finally, in the recent work by Choe et al. (2015) a container to be re-marshalled is inserted into a yard crane job schedule based on an estimation of the retrieval duration of the container. In this context, possible interruptions of the re-marshalling schedule by incoming and outgoing containers are considered in this study explicitly. By fixing a look-ahead horizon, the proposed algorithm reschedules the re-marshalling iteratively while determining the job sequence with a Genetic Algorithm. The subsequent stacking of the container is performed with general container stacking rules based on previous approaches. Results show an improvement of AGV waiting times, while XT waiting times stay the same in comparison to no re-marshalling.

3.3 Contribution and comparison with previous studies from literature review

In this work, two major contributions, that have been identified from the previous literature review, are proposed and experimentally tested. As the nature of both topics is highly related, both are integrated and analysed comprehensively in order to address open questions about the nature of re-marshalling in front-end yard blocks in combination with a usage case of the TAS scarcely covered in literature.

Firstly, a new solution method for the re-marshalling problem is proposed which primarily targets the reduction of shuffle moves and which can be fully integrated into yard block operation. The review paper on container handling by Caserta et al. (2011) calls for fast online solution approaches for the RMP and the integration of container handling problems in general. These topics are specifically addressed by the proposed algorithm in this work along with further distinctive features of the algorithm:

1. Handling of interruptions of the re-marshalling schedule by incoming and outgoing containers.
2. Explicit assignment of target stacks in a block without the need for designated (empty) stacks.
3. Import, export and transshipment containers as targets for re-marshalling.
4. Systematic strategy to evaluate most suitable stacks/containers to be re-marshalled.
5. Flexibility to incorporate updated container retrieval information (e.g. through TAS).
6. Assumptions about retrieval sequences are based on practically available data only.
7. Fast solution time for online real-time decision making.
8. Ability to implement the algorithm for different RMG-systems with no or low adjustment costs.
9. Testing the performance of the algorithm on different RMG-systems for practical front-end block layouts ranging from low to high capacity over a planing period going beyond the clearance of one bay or stack.
10. Testing the performance of the algorithm embedded in full yard block operation and interacting with a multitude of operational policies and solution methods to other problems.

Table 2 shows a comparison of this work with previous studies regarding the ten features (numbered 1–10) mentioned above. It can be observed that these features have not been considered holistically and that testing for practical terminal layouts has not been systematically conducted in literature, yet. By covering the largest share of features, Kemme (2012a) is the only study suitable for numerical comparison with the proposed algorithm of this work as both are founded on the

Table 2 Comparison of features of re-marshalling algorithm in this work with previous studies

Feature	1	2	3	4	5	6	7	8	9	10
Yu et al. (2009)						✓	✓			
Park et al. (2009)		✓	✓			✓	✓	✓		
Kang et al. (2006)							✓			
Kemme (2012a)	✓	✓	✓	✓		✓	✓	✓	✓	✓
Asperen et al. (2013)		✓			✓	✓		✓	✓	
Choe et al. (2015)	✓	✓	✓			✓	✓			
This work	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

same assumptions. However, it should be noted that a fair comparison between the algorithm by Kemme (2012a) and the one of this work is only possible without TAS as the former is not suitable for TAS application.

Secondly, influences of different TAS parameter settings on the re-marshalling behaviour for 36 RMG-system and block layout combinations are analysed by extensive simulation (four RMG-systems combined with nine layouts). On the one hand, this supports the understanding of the developed algorithm and deepens the evaluation of its performance. On the other hand, it serves as basis to derive insights for practical implementation of a TAS by terminal operators. The evaluation of a wide scope of performance indicators should aid terminal operators in assessing the relevant metrics and in revealing the interdependencies when targeting specific performance indicators and TAS usage cases.

As shown in the literature review about TAS, its application is often confined to utilising the TAS as a tool to balance XT arrivals by imposing caps for each time slot. Still, the combination of TAS and re-marshalling allows for a further analysis not conducted by any of the aforementioned studies: The influence of updated arrival information on unproductive moves and vehicle waiting times without enforcing any restrictions on XT arrivals. In addition to increasing information quantity, this enables an assessment of effects of information quality and reliability by varying TAS parameters concerning arrivals on time, length of the time slot and the obligation to book a slot.

4 Simulation model and solution method for the re-marshalling problem

This section gives a short overview of the simulation model employed for the numerical test conducted in this work. Afterwards, it describes the details of the algorithmic procedure and the surrounding ideas.

4.1 Main features of the simulation model

As a result of the complex, stochastic and dynamic environment of a yard block, a simulation model of a single front-end block for SRMG-, TRMG-, DRMG- and TriRMG-systems is used to implement the TAS and solution method for the RMP. For a detailed description of the verified and validated model features and a systematic analysis of its parameters it is referred to Kemme (2012a). Therefore, only the most relevant default characteristics and methods are described here.

Three types of transport modes for container pick-up and delivery are considered: deep sea vessels that follow a user-generated weekly repeated *vessel-call pattern* (VCP), feeder vessels and XTs that are randomly generated.

One simulation run covers a period of 42 days in total consisting of 14 days warm-up time and 28 days for data generation. By varying the length (bays), width (lanes) and height (tiers) of the block, it is possible to examine different block layouts and capacities.

In addition to the *re-marshalling problem (RMP)* targeted in this work, three main problems must be solved to operate the yard block. For the purpose of a representative performance analysis of the proposed algorithm for the RMP, experimentally tested solution methods by Kemme (2012a) are employed for these other three problems in order to evaluate the proposed solution method for the RMP within full yard block operation. The interaction of these problems with the re-marshalling problem is stated subsequently. Firstly, based on the solution to the *container stacking problem*, a re-marshalling schedule is devised by the proposed algorithm of this study. The block configuration after re-marshalling serves as input data for the container stacking problem again to stack an incoming container. Secondly, the identified re-marshalling jobs are passed to the *crane scheduling problem* as input data to distribute and sequence jobs (i.e. storage, retrieval, shuffle and re-marshalling moves) for the yard cranes. Thirdly, based on this schedule, the *crane routing problem* is solved to mitigate yard crane interference in the case of multi-RMG-systems.

In general, it must be noted that the algorithmic procedure for the RMP, the experimental design for performance testing and sensitivity analysis and all results reported here are generated in the framework of the study conducted for this paper. The proposed algorithm and TAS were implemented in the existing model from Kemme (2012a) with existing solution methods for other problems. Best practices for choosing the appropriate algorithm for the other problems, which are relevant for the interaction with the proposed algorithm, were applied based on previous studies.

4.2 Solution method for the RMP

The RMP entailing the entire yard block represents a multi-criteria decision problem with a need to balance the potential for additional mis-overlays and long empty crane movement while leaving space for future shuffle-free container stacking. The online nature requires fast decision-making about the container to be re-marshalled and its new target stack. To identify a container, the suitability of each stack for re-marshalling must be evaluated. Depending on the layout, the number of stacks per block ranges from about 150–480 stacks with a total capacity of about 300–3000 containers subject to the number of tiers. Hence, a decision method is needed that offers a framework for scoring the stacks and matching their top containers with a target stack in real-time. To satisfy the above mentioned requirements, the proposed solution approach is a heuristic based on *Fuzzy Complex Queries (FCQ)*[3] defining rules for container handling implemented as expert system. FCQ can be associated with crisp relational databases. They are applied when a specific attribute of a database entry cannot be determined exactly. In this context, the entries are queried according to the degree of membership to a certain attribute they belong to. Therefore, results contain multiple entries to choose from with different degrees of membership to a query. The term ‘complex’ indicates the possibility to combine

[3] For a detailed description of the generic method see (Zimmermann 2001, pp. 268–277).

attributes within a query to restrict results to entries fulfilling the combination to a certain degree.

4.2.1 Algorithmic foundation

The principle of degrees of membership can be employed for assessing the suitability of a stack for re-marshalling within the multi-criteria environment. Let $s \in S$ denote a stack from the set of all stacks S in the block, $c \in C$ a container from the set of all containers C in the block and g a yard crane in the block. Moreover, let $SS \subseteq S$ and $TS \subseteq S$ be the subsets of stacks that belong to the query tables *Source Stacks* and *Target Stacks*, respectively. The underlying framework to execute the FCQ-algorithm can be described as follows:

1. Depiction of every stack as database entry with attributes classified in two *Tables*:

   a) Source Stacks (identification of containers to be re-marshalled)
   b) Target Stacks (finding new stacking position)

2. Each attribute is described by a linguistic variable A consisting of linguistic terms lt.
3. Properties of each stack are defined as base variables X_s^A and mapped onto linguistic terms by means of normalised membership functions $\mu_s^{lt} \in [0, 1]$. (For a related mapping approach see Ries et al. (2014)).

Based on this verbalisation of container and stack properties, rules are deduced to perform the container re-marshalling. Broadly speaking, these rules are causing a container movement from an *unordered* stack to an *ordered* stack. To determine the stack ordering, mis-overlays are anticipated based on vehicle departure times which can either be approximated by weekly repeated vessel call patterns for deep sea vessels and TAS information for XTs or estimated using the average container dwell time in case of totally random vehicle departures. Formalising the idea described above, a query for the identification of stacks is performed by the following command:

$$
\begin{aligned}
&SELECT\,[Stack\, s]\\
&FROM\,[Table]\\
&WHERE\,[A \text{ is } lt]
\end{aligned}
$$

The resulting list contains the stacks with its top containers that satisfy the query with their respective membership degree to each queried attribute. Multiple attributes can be inquired by using the connectives *and* and *or* in the *WHERE* predicate. However, as a crane can perform one re-marshalling job at a time, an aggregation operator must be used to merge the membership degree of each attribute to represent a single membership value indicating the suitability for re-marshalling of a stack. Normally, Fuzzy Set Theory expresses the *and*-connective by using the *min*-operator and the *or*-connective by using the *max*-operator. Regarding the application described in this work, a stack's suitability for re-marshalling would be determined by its lowest or highest membership degree to one queried attribute while neglecting the others using these aggregation operators. Thus, the compensatory *weighted averaging* operator is more apt to enable a contribution by each relevant attribute to the result (μ_s^{agg}). The difficulty of finding appropriate weights is overcome by goal prioritisation and consecutive extensive testing by simulation.

4.2.2 Linguistic variables

Firstly, the algorithmic foundation above involves the verbalisation of stack properties relevant for re-marshalling. Based on the problem description and objective definition in Sect. 2, Table 3 lists the attributes considered as linguistic variables for the solution method.

By specifying subsets of these attributes, a container to be re-marshalled and the target stack can be identified.

Linguistic terms

Secondly, each attribute can be described by linguistic terms based on the current state of a stack:

$$\text{Ordering} = \{\textit{ordered}, \textit{unordered}\}$$

$$\text{Urgency} = \{\textit{critical}, \textit{uncritical}\}$$

$$\text{Distance}(\text{Crane to Stack}) = \{\textit{near}, \textit{far}\}$$

$$\text{Distance}(\text{Stack to Handover}) = \{\textit{close}, \textit{distant}\}$$

$$\text{Priority} = \{\textit{high}, \textit{medium}, \textit{low}\}$$

$$\text{Occupancy} = \{\textit{full}, \textit{half}, \textit{empty}\}$$

The terms characterise the stacks individually by assigning a certain manifestation to each attribute.

Base variables and mapping

Thirdly, available numerical data about stacks in the block must be processed appropriately in order to associate them with the respective attributes. The processed data serve as base variables for the linguistic variables. The connection between both is performed by mapping the base variables onto linguistic terms

Table 3 Definition of linguistic variables

Linguistic variable *A*	Symbol	Attribute description
Ordering	Ord	To what extent the current ordering agrees with the expected
Urgency	Urg	How fast a blocked container must be freed up
Distance (Crane to Stack)	DCS	How far a crane must drive to reach a stack
Distance (Stack to Handover)	DSH	How far a container must be moved to reach its handover
Priority	Prio	How long until a container is retrieved from the block
Occupancy	Occu	How many containers are in a stack

through membership functions. Depending on the context, various types of functions can be employed to model the relationship. However, it is common to apply linear and/or triangular functions due to ease of implementation, often limited information about linguistic terms and overall good performance in control and business decision applications (Pedrycz 1994). A description of an attribute can be amplified by adding the modifier *very* before the respective linguistic term. This can be modelled by squaring the dedicated membership function of the not amplified term. As a result, stacks that show an exposed property are assigned a disproportionally high membership degree to its linguistic term.

As shuffle moves are evoked by container mis-overlays and the proposed algorithm should be suited to incorporate TAS information, departure times of deep sea vessels, feeders and XTs are fundamental for defining base variables. The three modes can be characterised by the availability of their departure times:

- *Deep sea vessel*: continuously repeated and known VCP; retrieval of container can be approximated within a time slot of vessel berthing.
- *Feeder*: random vehicle arrivals; retrieval of container can only be estimated using the average dwell time in the block.
- *XT*: two possible scenarios to determine retrieval time:

  a. Random vehicle arrivals; retrieval of container can only be estimated using the average dwell time in the block.
  b. TAS; retrieval of container can be approximated within a time slot booked by XTs.

Hence, following subsets $C_i \subseteq C$ with respect to the departure time category i of container c can be distinguished:

C_1: Subset of containers with *fixed* times (time slot based)
C_2: Subset of containers with *assumed* times (dwell time based)

The different accuracy of information for these two categories necessitates a separate handling of their containers by the algorithm. Table 4 shows the calculation parameters for the base variables and the fuzzy symbols that are the basis for running the algorithm:

In contrast to the other data, mis-overlays m cannot be obtained directly but need a pre-processing step beforehand. This is performed by evaluating the retrieval sequence within a stack. However, as two categories of retrieval times were defined before, a case analysis must be conducted. For this purpose, t_c^{out} is used as retrieval time for containers departing by deep sea vessel (approximated by the estimated departure time in the VCP) and for containers leaving by XT which booked a slot through the TAS (approximated by the start of the time slot). Due to the higher deviation of retrieval times for the second category of *assumed* times, following distinction is made when calculating mis-overlays showing the most consistent behaviour in simulation experiments:

Table 4 Notation for FCQ-algorithm

c	Index of container, a sourced container is indicated by c^*
s	Index of stack
SS	Set of source stacks
TS	Set of target stacks
X_s^A	Base variable for stack s with reference to linguistic variable $A \in \{Ord, Urg, DCS, DSH, Prio, Occu\}$
μ_s^{lt}	Membership function for stack s with reference to linguistic term lt
Q_k	Query result of query k
t_c^{out}	Retrieval time of container $c \in C_1$ from block, **if time slot based** (*VCP and TAS*)
t_c^{in}	Arrival time of container $c \in C$ in block
b_s	Bay of stack s
l_s	Lane of stack s
h_s	Height/tier of stack s
b_g	Current bay of crane g
l_g	Current lane of crane g
$\overline{T}^{rm}$	Mean time interval between two re-marshalling jobs
$\overline{d}^{rm}$	Mean duration of a re-marshalling job
$\overline{\delta}$	Mean container dwell time
m_c^{adj}	=1, if a container on top of container $c \in C$ is retrieved after container $c \in C$ (*adjacent mis-overlay*)
m_c^{dep}	Number of containers on top of container $c \in C_1$
m_s^{dep}	Number of containers on top of the deepest blocked container in stack s (*depth mis-overlay*)

Case 1: Two category 1 containers are stacked on each other ($c_2 \in C_1$ on top of $c_1 \in C_1$):

$$m_{c_1}^{adj} = 1, \quad if \quad t_{c_1}^{out} < t_{c_2}^{out} \tag{1}$$

Case 2: A category 2 container is stacked on a category 1 container ($c_2 \in C_2$ on top of $c_1 \in C_1$):

$$m_{c_1}^{adj} = 1, \quad if \quad t_{c_1}^{out} < t_{c_2}^{in} + \overline{\delta} \tag{2}$$

Case 3: Two category 2 containers are stacked on each other ($c_2 \in C_2$ on top of $c_1 \in C_2$):

$$m_{c_1}^{adj} = 1, \quad if \quad \lfloor t_{c_1}^{in} + \overline{\delta} \rceil < \lfloor t_{c_2}^{in} + \overline{\delta} \rceil \tag{3}$$

Case 4: A category 1 container is stacked on a category 2 container ($c_2 \in C_1$ on top of $c_1 \in C_2$):

$$m_{c_1}^{adj} = 1, \quad if \quad \lfloor t_{c_1}^{in} + 1 \rfloor < t_{c_2}^{in} \tag{4}$$

In Case 1 a mis-overlay is indicated if the top container has a later *fixed* retrieval time than the container below. A mis-overlay arises in Case 2 if the top container has a later dwell time-based retrieval time than the *fixed* time of the container below. As a result of the increased deviation for dwell time-based estimation, only time classes of containers with *assumed* times are considered in Case 3. This is done by rounding the container arrival times to the closest full day. Lastly, in Case 4 a mis-overlay is recorded if a top container of category 1 has a later arrival time than the arrival time rounded to the next highest full day of the category 2 container below. The rounding represents a buffer for reducing unnecessary re-marshalling moves resulting from the uncertainty of category 2 container retrieval times.

Using the preprocessed parameters the base variables are calculated as follows:

1. To characterise the ordering, the sum of depth and adjacent mis-overlays in stack s is used:

$$X_s^{Ord} = m_s^{dep} + \sum_{c \in s} m_c^{adj} \qquad \forall s \tag{5}$$

2. To represent urgency, the maximum difference in a stack between the retrieval time of a blocked container and estimated time to free up the container is used (slack). The smaller the slack the higher the urgency to free up the container. This does not apply to category 2 containers due to the high uncertainty of their retrieval. Hence, only category 1 containers are prioritised:

$$X_s^{Urg} = \max_{c|c \in s, c \in C_1} \{t_c^{out} - m_c^{dep} \cdot (\overline{T}^{rm} + \overline{d}^{rm})\} \qquad \forall s \tag{6}$$

3. For the distance variables, the total movement by portal (and trolley) is calculated:

$$X_s^{DCS} = |\ b_g - b_s\ | + |\ l_g - l_s\ | \qquad \forall s \in SS \tag{7}$$

$$X_s^{DSH} = |\ b_s - b_{handover}\ | \qquad \forall s \in TS \tag{8}$$

4. Priority is depicted by the *fixed* departure time of the top container's pick-up vehicle. Category 2 containers are assigned the highest priority in the block by default. This enables a separation of category 1 and 2 containers, afterwards:

$$\begin{aligned} X_s^{Prio} &= t_c^{out} && \forall s,\ c = c^{top},\ c \in C_1,\ c^{top} \in s \\ X_s^{Prio} &= \max_{c \in C_1} \{t_c^{out}\} && \forall s,\ c = c^{top},\ c \in C_2,\ c^{top} \in s \end{aligned} \tag{9}$$

5. Occupancy can be represented by the current height of the stack:

$$X_s^{Occu} = h_s \qquad \forall s \tag{10}$$

Figure 2 specifies the types of membership functions for mapping the base variables. In case of two linguistic terms per linguistic variable, linear or quadratic relationships are used. As it can be expected that a re-marshalling schedule will

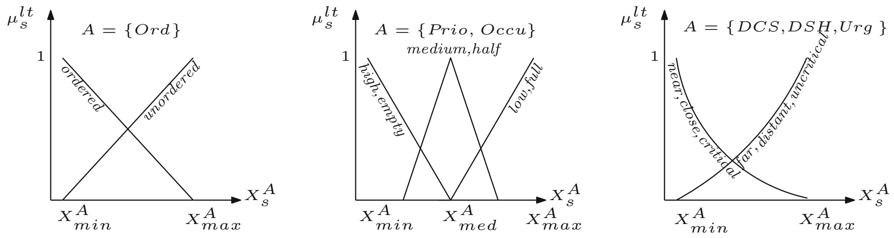

Fig. 2 Overview of membership functions applied for mapping of base variables

be interrupted by incoming or outgoing containers at one time, the Urgency and Distance variables are modelled as quadratic functions in order to guarantee that *very critical* and *very near* stacks are addressed first to quickly clean up these stacks and minimise the duration of re-marshalling jobs. Variables encompassing three terms are modelled by triangular membership functions. Some base variables are discrete, e.g. X_s^{Ord} and X_s^{Occu} as mis-overlays and stack heights are integer values only. Nevertheless, these are modelled by continuous functions to facilitate an automatic mapping for different block layouts without manual adjustment by the user. As a result, only certain values can be assumed within this continuous relationship. The supports are given by the minimum (X_{min}^A), medium (X_{med}^A) and maximum (X_{max}^A) value of the respective base variable.

Stacking restrictions

In spite of describing stacks by degrees of suitability for re-marshalling, there are crisp stacking restrictions that must be obeyed, in addition. On the one hand, these are activated for reasons of feasibility ensuring that only appropriate containers are stacked onto each other (restrictions 1a and 1b). On the other hand, they confine new container stackings to *ordered* target stacks (restriction 2a) and prohibit the creation of an adjacent mis-overlay with the top container of the target stack (restriction 2b). The latter is applied to avoid turning fully *ordered* stacks ($\mu_s^{ordered} = 1$) into partially *ordered* stacks ($\mu_s^{ordered} < 1$):

1. Stacking restrictions for feasibility: avoid container stacking
   a. if the maximum height of the candidate stack has been reached or
   b. if the size of the top container in the candidate stack does not match the container to be re-marshalled.

2. Stacking restrictions for ordering: avoid container stacking
   a. if Ordering of the candidate stack $\mu_s^{ordered} \leq 0.5$ (stack is not more *ordered* than *unordered*) or
   b. if a mis-overlay between the top container of the candidate stack and the container to be re-marshalled will be induced.

Rule-based algorithmic procedure

By employing the framework above, rule-based queries are performed to identify a container c^* to be re-marshalled (*FCQ-Rule 1*) and to find a suitable target stack (*FCQ-Rules 2 and 3*). After presenting the rules, the interaction of the performed queries is described in the subsequent pseudocode formulation in Algorithms 1 and 2.

$$
Rule1: \begin{array}{ll}
SELECT & [s^{Source}] \\
FROM & [SS] \\
WHERE & [\text{Ordering is } \textit{unordered} \text{ and} \\
 & \;\text{Urgency is } \textit{very critical} \text{ and} \\
 & \;\text{Distance (Crane to Stack) is } \textit{very near} \text{ and} \\
 & \;\text{Priority is } \textit{high}]
\end{array}
$$

Rule 1 identifies stacks[4] that need to be cleaned up as a result of blocked containers within them. The rationale behind this rule is to find a stack s^{Source} with many mis-overlays where a blocked container in this stack is estimated to leave the block soon. Thereby, it serves the purpose of eliminating as many mis-overlays as possible by targeting the stacks which will potentially have the most shuffle moves in the near future. This rationale is supplemented by an additional consideration of yard crane (empty) movement in order to account for the distance minimisation with reference to the RMP. Additionally, stacks with a high priority top container are preferred because it is easier to stack them onto other stacks without creating a new mis-overlay, thus, increasing the number of potentially suitable target stacks. Note that it is still possible to choose a top container with medium or low priority if the stack's other attributes are particularly pronounced.

$$
Rule2: \begin{array}{ll}
\multicolumn{2}{l}{\text{if Priority of } c^* \text{ is a) } \textit{high}, \text{ b) } \textit{medium}, \text{ c) } \textit{low} \text{ then}} \\
SELECT & [s^{Target}] \\
FROM & [TS] \\
WHERE & [\text{Occupancy is a) } \textit{full}, \text{ b) } \textit{half}, \text{ c) } \textit{empty}]
\end{array}
$$

Rule 2 maintains enough space for future re-marshalling jobs. The rationale behind this rule considers that high priority containers will leave the block sooner while low priority containers will dwell in the block longer. Thus, low priority containers should be stacked near to ground slots so that further containers of higher priority can be stacked on top of it. If a high priority container was stacked near to a ground slot, the stack would be restricted for all containers with lower priority in order to avoid an additional mis-overlay although there would be additional slots left in this stack.

[4] The top container c_i^* to be re-marshalled is subsequently chosen from these stacks.

$$Rule3: \begin{array}{ll} SELECT & [s^{Target}] \\ FROM & [TS] \\ WHERE & [\text{Ordering is } \textit{ordered} \text{ and} \\ & \ \text{Urgency is } \textit{very uncritical} \text{ and} \\ & \ \text{Distance (Stack to Handover) is } \textit{very close}] \end{array}$$

Finally, *Rule 3* identifies the opposite stacks of Rule 1 by contrastingly mirroring its query to find clean stacks for putting the re-marshalled container onto. The rationale behind this rule is to find a stack s^{Target} with few mis-overlays where a blocked container in this stack is estimated to leave the block late. Thereby, it serves the purpose of avoiding as many mis-overlays as possible by targeting the stacks which will potentially have the fewest shuffle moves in the near future. This rationale is supplemented by an additional consideration of yard crane (empty) movement in order to account for the distance minimisation with reference to the RMP.

It should be noted that the main objective of FCQ is to minimise the number of blocked container (future shuffle moves) which will be retrieved soon, i.e. to target *unordered* and *critical* stacks. For this reason, these query attributes are weighted more strongly while the distance minimisation is added to avoid inordinate empty crane movement for balancing purposes.

Altogether, the pseudocodes of Algorithms 1 and 2 provide the full procedure applying *Fuzzy Complex Queries*.

Algorithm 1 Pseudocode formulation of Fuzzy Complex Query (FCQ) - *Decision 1*

1: **if** crane g is idle and set of all plannable jobs for g: $J_g^p = \{\}$ **then**
2:     **for all** s that have changed since last re-marshalling job **do** ▷ Pre-processing
3:         update $X_s^A \mid A = \{Ord, Urg, DCS, Prio, DSH, Occu\}$
4:         $\mu_s^{lt} \leftarrow f(X_s^A)$
5:     **end for**
6:     $SS \leftarrow$ all $s \mid \mu_s^{unordered} > 0$ and $\mu_s^{crtical} > 0$
7:     $TS \leftarrow$ all $s \mid \mu_s^{ordered} > 0.5$ and $\mu_s^{full} < 1$
8:     $Q_1 \leftarrow$ query $s \in SS$ with *FCQ-Rule 1* ▷ *Decision 1*
9:     **for** $s \in Q_1$ **do**
10:         compute μ_s^{agg}
11:     **end for**
12:     $s^{Source} \leftarrow \arg\max_s \{\mu_s^{agg} \mid s \in Q_1\}$
13:     $c^* \leftarrow c^{top} \in s^{Source}$
14: **end if**
15: **goto** *Algorithm 2*

Algorithm 1 describes the procedure to determine a **source** stack for re-marshalling the respective top container. It is initiated when a crane is idle (line 1). As a consequence, the preprocessing is performed to evaluate every stack by mapping base variables onto linguistic terms for each linguistic variable (lines 2–5). Moreover, source and target candidates to query from are determined (lines 6–7). In this context, the set SS contains all stacks with at least one mis-overlay ($\mu_s^{unordered} > 0$) and a *critical* container (membership function $\mu_s^{critical} > 0$). The set TS contains all stacks whose maximum height has not been reached yet (membership function $\mu_s^{full} < 1$) and are more ordered than unordered (membership function $\mu_s^{ordered} > 0.5$). Lines 8–11 apply *Rule 1* to identify a suitable stack as source (query result Q_1) and aggregate the different membership degrees to one value (membership function μ_s^{agg}) for evaluation. In lines 12–13 the top container c^{top} of the most suitable stack is chosen to enter the second phase.

Algorithm 2 Pseudocode formulation of Fuzzy Complex Query (FCQ) - *Decision 2*

1: **get** c^* from *Algorithm 1*
2: **if** $c^* \neq 0$ **then**
3: $\quad Q_2 \leftarrow$ query $s \in TS$ with *FCQ-Rule 2* $\setminus \{s \mid Size(c^{top}) \neq Size(c^*),\ c^{top} \in s\}$ ▷ *Decision 2*
4: $\quad Q_3 \leftarrow$ query $s \in TS$ with *FCQ-Rule 3* $\setminus \{s \mid Size(c^{top}) \neq Size(c^*),\ c^{top} \in s\}$
5: $\quad Q_{23} \leftarrow Q_2 \cap Q_3$
6: $\quad$ **for** $s \in Q_{23}$ **do**
7: $\quad\quad$ compute μ_s^{agg}
8: $\quad$ **end for**
9: $\quad$ sort Q_{23} by decreasing μ_s^{agg} and set $j = 1$
10: $\quad$ **while** $(t_{c^*}^{out} > t_{c^{top}}^{out} \mid c^* \in C_1,\ c^{top} \in Q_{23}[j])$ and $\mu_{Q_{23}[j]}^{low} \neq 1$ and $\mu_{Q_{23}[j]}^{empty} \neq 1$ **do**
11: $\quad\quad Q_{23}[j] \leftarrow Q_{23}[j+1]$
12: $\quad$ **end while**
13: $\quad$ **if** $\mu_{Q_{23}[j]}^{agg} > 0.5$ **then**
14: $\quad\quad s^{Target} \leftarrow Q_{23}[j]$
15: $\quad\quad$ perform re-marshalling job for c^* from s^{Source} to s^{Target} ▷ Container movement
16: $\quad$ **end if**
17: **end if**
18: **goto** *Algorithm 1*

If a top container c^* is received as input (lines 1–2), the next phase begins by querying with *Rules 2 and 3* (lines 3–5) as shown by Algorithm 2. Stacks are removed from the query results Q_2 and Q_3 whose top container is not of the same size as c^* (Q_{23}). After aggregating the relevant membership degrees in lines 6–8, lines 9–12 choose a **target** stack based on the restriction that an additional adjacent mis-overlay must not be induced by the re-marshalling job. On the one hand, this is guaranteed by stacking category 1 containers only on other category 1 containers with later retrieval times. On the other hand, category 1 and 2 containers may be stacked on containers that have the lowest priority in the block or in ground slots (line 10). The query results in Q_{23} are sequentially checked in decreasing order until a suitable stack is found (lines 11–12). If the identified stack is more suitable than unsuitable (aggregated membership function of the $j-th$ query result in query list Q_{23} $\mu_{Q_{23}[j]}^{agg} > 0.5$, lines 13–14), the container is re-marshalled and the procedure is re-initiated to check for an idle crane again (lines 15-17).

5 Implementation

This section reports on simulation setups and experimental results with a subsequent discussion of the implications.

5.1 Experimental setup

Parameters and policies of the simulation model are set to default as described by Kemme (2012a). Noticeably, export, import and transshipment (30%)[5] containers are regarded while the ratio of 40 TEU to 20 TEU containers is 3:2. The planned average block occupancy rate is set to 75%. As five tiers have been identified as most suitable stack height by Duinkerken et al. (2001), the tested block layouts are varied with respect to this parameter. The capacities considered range from 864 to

[5] The transshipment factor regulates the number of feeder vessels picking-up and delivering containers.

2640 containers with selected combinations of $bays \in \{36, 44\} \times lanes \in \{6, 8, 10\} \times tiers \in \{4, 5, 6\}$ representing different block dimensions (subsequently referred to as *bays/lanes/tiers*). The basic framework for the layout designs is based on Kemme (2012a). However, different combinations are chosen in this work to account for the RMP properties. Firstly, small layouts with 36 bays and 6 lanes are tested with all numbers of tiers. It serves to show the behaviour of the algorithm in **high** stacks (in particular, 36/6/6). Secondly, the three levels are varied in layouts with 36 bays and 8 lanes to investigate whether the algorithm is able to work properly in layouts **balancing** the layout dimensions. Thirdly, **large** layouts are tested with an increased number of bays (44) and lanes (10) to test for long moving distances with respect to crane portal and trolley.

Simulations of the above mentioned RMG-system and block layout combinations are conducted in four versions to provide a benchmark for the FCQ-algorithm performance regarding the identified performance indicators in Sect. 2.2.

1. No re-marshalling (noRM) without any rearrangement of containers during periods when the cranes are idle.
2. With re-marshalling where the ensuing RMP is solved with the benchmark heuristic Heuristical Housekeeping Stacking (HHS) (Kemme 2012a).
3. With re-marshalling where the ensuing RMP is solved with the proposed FCQ-algorithm and no TAS is used.
4. With re-marshalling where the ensuing RMP is solved with the proposed FCQ-algorithm and TAS is used with full information within a slot of 1 h duration one day prior to XT arrival.

The first version is used to assess the influence of solving the RMP with FCQ on overall yard block productivity in comparison to no re-marshalling. The second version is used to assess how well the FCQ solves the RMP in comparison to a suitable benchmark heuristic from literature.

As stated in Sect. 3, the nature of other re-marshalling algorithms except for HHS do not allow a direct comparison with the FCQ due to their assumptions about yard block environment, available stacks and re-marshalling schedule, among others. Thus, noRM and HHS are used as benchmarks for the influences of re-marshalling in general and the heuristic proposed in this work. Following the performance comparison, TAS parameters are varied to analyse the FCQ behaviour under different conditions and to recommend a suitable TAS design for terminal operators and planners. In Table 5 the scheme of TAS parameter variation is shown. All settings described here are applied for nine different and practically relevant block layouts with four RMG-systems each. One simulation experiment with fixed parameter settings is tested in ten runs with altering seed values for the random number generator. This number of runs shows a good balance in the trade-off between stable mean values of indicators and computational time. Depending on the RMG-system and block layout under consideration, one run can last up to 20 min averaging about 8 min for all layout-system configurations. A larger number of runs per experiment was evaluated up to 15 runs which showed no considerable changes in mean values for the tested experiments. More importantly, the overall

Table 5 Parameter settings for simulation of the terminal appointment system

Parameter	Setup	Default	No. of settings	Represented property
Share of booked slots	0 to 100%	100%	7	Information quantity
Share of correct arrivals	0 to 100%	100%	7	Information reliability
Slot length	0.25 to 48 h	1 h	6	Information quality
Announcement time	1 to 120 h	24 h	6	Information quantity
Arrival deviation	1 to 48 h	–	5	Information uncertainty

observations of the performance of the algorithms were consistent with the experiments conducted with ten runs. For the purpose of providing a wide scope of experimental settings with practical and realistic RMG-system and layout configurations with different TAS settings, ten runs per experiment were deemed appropriate based on the pre-evaluation of simulation length and changes in mean value. This enables the study of performance indicators for statistically significant changes ($\alpha = 5\%$). All in all, 144 experiments (1440 runs) are conducted for performance comparison while additional 1116 experiments (11,600 runs) are carried *ceteris paribus* out to test for TAS parameter influence, totalling 1260 experiments (12,600 runs).

The default TAS setup simulates precise and accurate arrival information in order to mitigate interdependencies between the parameters when analysing their influence. For instance, if all XTs book a slot and arrive on time, the slot length can be studied independently of influences from arrival deviation and uncertainties about XT bookings. To examine the impact of arrival deviation the share of correct arrivals is set to 0% by default based on the same reasoning. The share of booked slots and the share of correct arrivals are addressed jointly by examining the participation share which is defined as the product of the two. In order to assess the isolated influence of each share on the defined performance indicators, either the share of booked slots or the share of correct arrivals is held at 100% while the other is being varied. In general, the entire range of information quality and quantity is investigated by employing this experimental design. The variation begins with a very accurate and detailed knowledge about retrieval times of outgoing containers by XT and moves on by deteriorating it until only vague arrival information is known to the terminal regarding the respective parameter. As QQ-plots showed non-normal distributions of all performance indicators and the different strategies are tested with the same random numbers per instance (common random numbers), the non-parametric pairwise Wilcoxon signed-rank test is used to test for statistical significance in the performance analysis. Moreover, 48 comparisons (six comparisons between strategies multiplied by eight comparisons of performance indicators) are made. This necessitates an adjustment of the chosen significance level $\alpha = 5\%$ for multiple comparisons which is done by the Benjamini-Hochberg correction. Regarding the TAS parameter sensitivity analysis, the Kendall-tau-b correlation coefficient ($\alpha = 5\%$), which is identified as suitable for non-normal distributions, is calculated for each parameter and the examined performance indicator (Kendall 1938).

5.2 Numerical results and discussion

Firstly, this subsection presents a performance comparison of FCQ with the HHS when solving the RMP. It also analyses the impacts of performing re-marshalling based on these algorithms with performing no re-marshalling. Secondly, the behaviour of the FCQ-algorithm under different TAS settings is numerically analysed and discussed. The terms in brackets throughout this subsection denote the relative improvement induced by a given algorithm and RMG-system: (*average of all layouts; best of all layouts; worst of all layouts*).

5.2.1 Performance of FCQ-algorithm

According to the performance indicator types identified in Table 1, the hypotheses in Table 6 are formulated. They support the understanding of the underlying intention of the algorithm design when addressing the interdependencies of the indicators.

Distance-based as well as mis-overlay-based re-marshalling was identified in Sect. 4 as means to target the superordinate goal of reducing vehicle waiting times at the handover as a measure for block productivity. While HHS focuses on the former, FCQ is designed to focus on the latter albeit integrating distance for balancing purposes. Hypothesis 1.1 states that HHS considers only distance-based improvements of containers. Thus, substantial changes in the number of shuffles moves per container n_{shu} in comparison to noRM cannot be expected because the container retrieval sequence within a stack is not considered with this heuristic. Adversely, FCQ in both versions should yield significant results as eliminating mis-

Table 6 Hypotheses for FCQ performance analysis

Type	No.	Hypothesis
Container handling	1.1	The tested re-marshalling versions can be ordered as follows according to n_{shu}: $n_{shu}(noRM) \approx n_{shu}(HHS) > n_{shu}(FCQ) > n_{shu}(FCQ/TAS)$
	1.2	FCQ/TAS improves the re-marshalling quality (n_{rm}) the most
	1.3	The tested RMG-systems can be ordered as follows according to N_{rm}: $N_{rm}(SRMG) < N_{rm}(TRMG)$, $N_{rm}(DRMG) < N_{rm}(TriRMG)$
Vehicle waiting times at handovers	1.4	FCQ without TAS improves $\overline{\varpi}_{wsout}$ compared to noRM and HHS
	1.5	FCQ/TAS improves $\overline{\varpi}_{lsout}$ compared to noRM and HHS
Standard deviation of vehicle waiting times	1.6	FCQ/TAS improves σ_{lsout} the most

n_{shu} number of shuffle moves per container, n_{rm} number of re-marshalling moves to eliminate a shuffle move, N_{rm} total number of re-marshalling moves, $\overline{\varpi}_{wsout}$ mean AGV waiting time for retrieval jobs, $\overline{\varpi}_{lsout}$ mean XT waiting time for retrieval jobs, σ_{lsout} standard deviation of XT waiting time for retrieval jobs overlays is the central rationale behind this heuristic. Next, as TAS enhances arrival information quantity and quality, re-marshalling moves are expected to be fitting without the need for future corrections due to initial lack of arrival information. Hence, Hypothesis 1.2 states that the minimal sequence of re-marshalling moves should be obtained by applying FCQ/TAS. Hypothesis 1.3 states that with increasing crane resources more idle crane time can be expected offering more opportunities to perform re-marshalling.

If no XT arrival information is available, the FCQ-algorithm generates only re-marshalling jobs that benefit export containers as they are the only containers forming category 1 of *fixed* departure times (see Sect. 4.2). Due to the random arrivals of other modes, containers leaving by deep sea vessel are solely regarded as *critical* and classified in *SS*. Therefore, Hypothesis 1.4 articulates this expected improvement in waterside retrieval jobs $\overline{\omega}_{wsout}$. Likewise including TAS, landside containers are included in *SS* making them available for re-marshalling as well. Based on this addition, landside retrieval jobs $\overline{\omega}_{lsout}$ should benefit which is stated in Hypothesis 1.5. Hypotheses about storage jobs ($\overline{\omega}_{wsin}$, $\overline{\omega}_{lsin}$) are not made as the interdependencies are not clear beforehand. Although FCQ/TAS continues admitting containers leaving by deep sea vessel in *SS*, the influence of additional re-marshalling of import containers on the relocation of export containers cannot be fully assessed, thus, exacerbating a holistic view on total vehicle waiting time $\overline{\omega}_{tot}$ as well. Generally, all tested re-marshalling versions should demonstrate an a priori unspecified performance improvement over noRM.

The standard deviation of waiting times serves as indicator if any workload smoothing can be observed as a result of applying a specific re-marshalling version. By reducing $\overline{\omega}_{lsout}$, the accumulation of waiting XTs at the handover is mitigated resulting in less peak workload. As a consequence, Hypothesis 1.6 states that there should be an only indirectly targeted improvement in σ_{lsout} when applying TAS.

A. SRMG-system

The numerical results for relevant performance indicators are given in Table 7 for the simulation experiments applying different re-marshalling versions in all layouts with SRMG-systems.

A1. Container handling

As expected, the simulation results show only minor changes in $\overline{n}_{shu}$ (2.0; 5.6; −1.7%) for HHS, while being at par or even slightly worse for some layouts. Contrary, by addressing mis-overlays and distance simultaneously, both versions of the FCQ are able to reduce $\overline{n}_{shu}$ significantly ranging from 2.2% in the worst case to 54.4% in the best case with an average improvement of 11.8 and 36.9% over noRM, respectively. In spite of the highly significant improvements for FCQ without TAS over HHS, there is a substantial performance gap between FCQ and FCQ/TAS. By applying FCQ/TAS, it is possible to avoid the need to shuffle every container in seven out of nine layouts with respect to $\overline{n}_{shu}$. These results underline the importance of XT arrival information to perform a more accurate mis-overlay calculation and find truly suitable stacks for containers. This can also be observed from the number of re-marshalling moves needed to eliminate one shuffle move $\overline{n}_{rm}$ which improves

Table 7 Strategy comparison of the means of performance indicators for 10 simulation runs in SRMG-systems

Layout	Strategy	$\overline{n}_{shu}$ ($\overline{N}_{shu}$)	$\overline{n}_{rm}$ ($\overline{N}_{rm}$)	$\overline{\overline{\omega}}_{tot}$	$\overline{\overline{\omega}}_{wsin}$	$\overline{\overline{\omega}}_{wsout}$	$\overline{\overline{\omega}}_{lsin}$	$\overline{\overline{\omega}}_{lsout}$	$\overline{\sigma}_{tot}$	$\overline{\sigma}_{lsout}$
36 × 6 × 4	noRM	1.14 (2402)	– (0)	3.38	2.21	5.03	1.32*	4.48	3.86	4.10*
	HHS	1.10 (2321)	10.75 (876)	3.27	2.44*	4.68	1.34*	4.43	3.59	3.74*
	FCQ	0.90 (1891)	3.03 (1550)	3.02	2.43*	3.48*	1.37*	5.00	3.32	3.96*
	FCQ/TAS	0.52 (1098)	1.95 (2547)	2.54	2.18	3.46*	1.32*	3.23	2.73	3.17
36 × 6 × 5	noRM	1.44 (3764*)	– (0)	7.09*	4.22*	11.17	3.47*	9.22*	10.39	10.92*
	HHS	1.44 (3787*)	– (1055)	6.70*	4.82	9.89	3.22*	8.50+	9.40	10.27*
	FCQ	1.25 (3305)	4.36 (1997)	5.63	4.45*	6.93*	2.70	8.64*+	7.65	8.86+
	FCQ/TAS	0.81 (2102)	1.69 (2804)	4.59	3.33	6.79*	2.35	5.66	6.59	7.81+
36 × 6 × 6	noRM	1.64 (4937*)	– (0)	15.64*	8.64	24.11	10.91*	18.95*	24.37*	26.10*
	HHS	1.66 (5148)	– (1136)	15.54*	10.62*	22.29	10.56*	18.50*	23.14*	25.55*
	FCQ	1.59 (4957*)	– (2072)	14.95*	10.08*	20.34	10.42*	19.37*	23.01*	25.40*
	FCQ/TAS	1.17 (3533)	1.55 (2179)	12.13	6.86	19.07	8.36	13.89	20.32	22.32
36 × 8 × 4	noRM	1.11 (3117)	– (0)	5.08	3.75	7.32	2.25	6.66	6.74	7.26
	HHS	1.07 (2999)	8.77 (1035)	4.62	4.19	6.05	2.04*	5.88	5.66	6.04*
	FCQ	0.89 (2507)	2.84 (1732)	4.20	4.08	4.45*	2.03*	6.41	5.09	6.12*
	FCQ/TAS	0.52 (1458)	1.88 (3114)	3.41	3.12	4.35*	1.83	4.15	4.09	4.85
36 × 8 × 5	noRM	1.40 (4876)	– (0)	16.89	10.14*	25.46	11.81	19.65	26.20*	27.89*
	HHS	1.41 (4949)	– (1198)	16.10	11.91	22.31	11.04	18.42	24.55*	26.66*
	FCQ	1.29 (4478)	5.41 (2149)	13.23	10.62*	16.20*	9.16	17.10	20.59	23.03
	FCQ/TAS	0.93 (3175)	1.62 (2755)	10.99	7.41	16.13*	7.53	12.36	18.41	20.06
36 × 8 × 6	noRM	1.54 (5989*)	– (0)	36.44*	23.65*	51.15	28.36*	43.82*	49.23*	53.21*
	HHS	1.57 (6151)	– (1066)	35.94*+	26.26+	46.89*	28.07*	43.60*	47.70*	53.27*
	FCQ	1.50 (5874*)	17.88 (1989)	34.03+’	24.79*+	43.12+	27.03*+	42.60*	47.16*	51.99*

Table 7 continued

Layout	Strategy	$\overline{n}_{shu}$ ($\overline{N}_{shu}$)	$\overline{n}_{rm}$ ($\overline{N}_{rm}$)	$\overline{\overline{\omega}}_{tot}$	$\overline{\overline{\omega}}_{wsin}$	$\overline{\overline{\omega}}_{wsout}$	$\overline{\overline{\omega}}_{lsin}$	$\overline{\overline{\omega}}_{lsout}$	$\overline{\sigma}_{tot}$	$\overline{\sigma}_{lsout}$
	FCQ/TAS	1.23 (4837)	2.02 (2317)	32.95’	21.84	46.06*+	25.89+	38.61	47.45*	52.12*
44 × 10 × 4	noRM	1.03 (4297)	− (0)	25.10	20.78	31.89	17.55	29.18	35.36	39.54
	HHS	0.97 (4076)	5.21 (1105)	20.44	21.65	21.81	13.98*	23.81*	27.83	32.46
	FCQ	0.91 (3817)	3.45 (1655)	18.56	20.07	17.89*	13.35*	23.38*	26.17*	31.11*
	FCQ/TAS	0.67 (2820)	2.15 (3177)	17.05	17.28	18.50*	12.17	19.92	25.11*	29.50*
44 × 10 × 5	noRM	1.19 (5381)	− (0)	55.46	41.09*	70.03	40.10	71.06	72.19	71.06
	HHS	1.13 (5157)	3.94 (884)	48.73	41.46*	55.08	36.37	63.90*	64.15*	75.82*
	FCQ	1.05 (4786)	2.68 (1594)	43.89*	40.74*	42.99	34.67*	61.73*	59.72	72.41+
	FCQ/TAS	0.83 (3883)	2.01 (3018)	44.34*	38.25	48.73	33.66*	58.88	62.57*	74.79*+
44 × 10 × 6	noRM	1.19 (5341)	− (0)	72.27	55.88*	86.47	52.06	98.91	91.51	102.90
	HHS	1.15 (5222)	6.05 (714)	64.73	55.53*	69.99	48.61	91.90	83.89	99.03*
	FCQ	1.02 (4665)	2.05 (1387)	56.18	54.38	50.05	44.99*	88.10	78.07	99.48*
	FCQ/TAS	0.81 (3810)	1.82 (2785)	57.71	51.55	59.23	44.05*	83.80	80.27	98.26*

*,+,’ Two or more results marked with the same symbol within a layout are not significantly different ($\alpha = 5\%$)

$\overline{n}_{shu}$ number of shuffle moves per container, $\overline{N}_{shu}$ total number of shuffle moves, $\overline{n}_{rm}$ number of re-marshalling moves to eliminate a shuffle move, $\overline{N}_{rm}$ total number of re-marshalling moves, $\overline{\overline{\omega}}_{tot}$ mean total vehicle waiting time, $\overline{\overline{\omega}}_{wsin}$ mean AGV waiting time for storage jobs, $\overline{\overline{\omega}}_{wsout}$ mean AGV waiting time for retrieval jobs, $\overline{\overline{\omega}}_{lsin}$ mean XT waiting time for storage jobs, $\overline{\overline{\omega}}_{lsout}$ mean XT waiting time for retrieval jobs, $\overline{\sigma}_{tot}$ standard deviation of total vehicle waiting time, $\overline{\sigma}_{lsout}$ standard deviation of XT waiting time for retrieval jobs accordingly. Note that $\overline{N}_{rm}$ increases with TAS due to a more granular mis-overlay calculation. It results in categorising more stacks in *SS* by the FCQ-algorithm which leads to causing more re-marshalling jobs. Ideally, exactly one re-marshalling move is needed to eliminate its corresponding shuffle move. Such a re-marshalling schedule can be constructed if there is complete information on container retrievals which is, however, not the present case. Thus, the growth in $\overline{n}_{rm}$ exceeds the decline of $\overline{n}_{shu}$ as uncertain and imprecise data about containers induce unnecessary and unfitting re-marshalling jobs. With TAS information available one day in advance, 1.5–2 re-marshalling moves are needed per shuffle move which is close to the ideal point in comparison to FCQ without TAS and HHS. Additional moves result from uncertain feeder and landside import containers in the block that are not booked.

A2. Vehicle waiting times
Overall, $\overline{\overline{\omega}}_{tot}$ can be significantly improved by HHS in six layouts compared to noRM (7.3; 18.6; 0.6%). FCQ without and with TAS reduces $\overline{\overline{\omega}}_{tot}$ even more with significant improvements compared to noRM by (16.7; 26.0; 4.4%) and (25.8; 35.2; 9.6%), respectively, performing better than HHS in every layout. However, the improvements in $\overline{n}_{shu}$ do not directly translate to vehicle waiting times with the same magnitude. This can be explained by the growing N_{rm} with FCQ(/TAS) because more containers can be described as *critical* and become candidates in *SS* for re-marshalling. As more re-marshalling jobs are performed there is a higher risk that new retrieval or storage jobs interrupt the schedule and must wait an additional amount of time until the crane finishes its current re-marshalling job. Depending on when a retrieval or storage job is generated, substantial additional waiting time may be induced for individual jobs. In addition, if there is a particularly suitable target slot during re-marshalling, the solution generated by the algorithm may choose a new stack further away from the handover of the candidate container, despite FCQ is accounting for distance. As a consequence, the crane has to move a longer distance at the time of retrieval.

In detail, FCQ is more able to improve waiting time for $\overline{\overline{\omega}}_{wsout}$ previously identified to be of higher importance. This is due to the fact that the algorithm makes extensive use of approximated time slots based on the weekly repeated vessel call pattern and online stowage planning whereas it does not stack according to raw container categories. Thus, export containers are continuously arranged according to their retrieval sequence. Regarding landside jobs, it can be observed that the application of TAS leads to significant improvements for $\overline{\overline{\omega}}_{lsout}$ in all layouts in comparison to no information (21.6; 35.5; 4.6%). Agreeing with the explanation for Hypothesis 1.4, FCQ without TAS does not address or prioritise import containers resulting in a performance on par or slightly worse than HHS concerning $\overline{\overline{\omega}}_{lsout}$. Improvements of $\overline{\overline{\omega}}_{wsin}$ and $\overline{\overline{\omega}}_{lsin}$ using FCQ can be explained by freed up crane resources previously occupied by the high $\overline{n}_{shu}$ with noRM and HHS. Hence, storage jobs benefit indirectly from FCQ with a pronounced effect using TAS. In accordance with observations for the container handling above, the distance-based evaluation by HHS particularly benefits retrieval jobs at both handover areas while there are generally minor improvements in $\overline{\overline{\omega}}_{lsin}$ and significant deterioration in

$\overline{\overline{\omega}}_{wsin}$. With regard to balancing waiting time between waterside and landside, total waiting time for waterside jobs are not adversely affected by the inclusion of TAS maintaining the benefits of an isolated application of the FCQ. This property is of importance as productivity improvements at the landside should not interfere with the terminal's priority customers at the waterside.

A3. Standard deviation of vehcile waiting times

In most cases, a significant improvement of the total standard deviation $\overline{\sigma}_{tot}$ can be noticed when applying FCQ instead of noRM (17.1; 26.4; 4.2%) or HHS (8.2; 18.6; 0.5%). Surprisingly, a significant improvement in $\overline{\sigma}_{lsout}$ by TAS can not be observed compared to FCQ without TAS in five layouts. The reason for this may be again the interruption of the re-marshalling schedule as described for the vehicle waiting times above. If a waterside and a landside job arrive simultaneously during the execution of a re-marshalling job, because it is more important the priority is given to the waterside job by default. Thus, there is a principal impact on the standard deviation of landside waiting time in adverse. The range of improvements in $\overline{\sigma}_{tot}$ nearly matches the range of $\overline{\overline{\omega}}_{tot}$. As the latter is decreasing, the accumulation of waiting vehicles at the handovers is mitigated resulting in less peak workload. Thus, vehicles can be steadily served due to increased crane productivity balancing the crane workload indirectly. Moreover, fewer shuffle moves are performed during retrieval operation leaving more time available for the actual pick up of the target container.

B. Multi-RMG-systems

As no systematic differences can be observed regarding the algorithm behaviour, the TRMG-, DRMG- and TriRMG-system are analysed collectively. Subsequently, the focus is put on additional insights gained for the three systems in comparison to SRMG.

B1. Container handling

The higher crane resources of multi-RMG-systems are noticeable in the increased $\overline{N}_{rm}$ for bigger block layouts compared to the SRMG-system. In the TriRMG-system case, there is continued growth in $\overline{N}_{rm}$ particularly for the largest and six-tiered layouts. It appears that the SRMG-system has enough resources to manage the necessary re-marshalling schedule in four-tiered layouts while the advantage of multi-RMG-systems is fully exploited in six-tiered layouts. The increase of $\overline{N}_{rm}$ in the latter seems to correspond to a decrease in $\overline{N}_{shu}$ as $\overline{n}_{rm}$ stays approximately constant in comparison to the SRMG-system. As a result of more fittingly executable re-marshalling jobs, the improvement over noRM in $\overline{n}_{shu}$ is higher for the FCQ-algorithm without TAS and particularly with TAS compared to SRMG. Repeatedly, TAS exhibits more sensitivity for improvements unfolding more potentials to utilise idle crane resources adequately. The FCQ without TAS shows rather similar results to the SRMG-system indicating that SRMG-systems may have sufficient resources to handle waterside re-marshalling exclusively.

B2. Vehicle waiting times

In spite of a stronger reduction of mis-overlays with multiple crane systems, the relative performance improvement of $\overline{\overline{\omega}}_{tot}$ over noRM for FCQ and FCQ/TAS is lower on average than for the SRMG-system. Nonetheless, the improvements for the TRMG- and DRMG-systems are more robust throughout the layouts performing better in the worst case layouts. The TriRMG-system shows already high productivity without re-marshalling due its high crane resources. Despite the very low waiting time for all layouts in this system, both versions of the FCQ are still able to unlock final productivity potentials by translating the elimination of shuffle moves into a reduction of vehicle waiting time. Additionally, this system shows the most consistent behaviour over all layouts as it is the only one in which FCQ is able to gain additional improvements in the largest layout by employing TAS. This can be explained by the fact that additionally generated re-marshalling moves resulting from enhanced information by TAS do not severely interfere with retrieval or storage jobs as triple crane resources are available, the crane workload is already balanced due to low waiting time in general and the crossing capability of one crane ensures a more balanced distribution of re-marshalling jobs between the cranes. Moreover, an isolated inclusion of distance-based re-marshalling (HHS) does not show any effects with TriRMG. HHS shows no improvement in $\overline{\overline{\omega}}_{tot}$ for five layouts and even a significant deterioration in four layouts. Similar observations can be made regarding the systems with two cranes with the TRMG-system showing no improvements in six layouts and the DRMG-system in four layouts. This is due to the fact that reductions gained for retrieval jobs are depleted by a waiting time increase for storage jobs emphasising the principal influence of mis-overlays to significantly reduce total vehicle waiting time.

B3. Standard deviation of vehicle waiting times

Multi-RMG-systems show a more consistent behaviour regarding the standard deviation of vehicle waiting times. Notably pronounced impacts by TAS can be observed in all systems in comparison to SRMG. Only the large layouts for the TRMG- and DRMG-system show no beneficial effects while six layouts show a clear reduction in this indicator. The TriRMG is the only system where FCQ/TAS is able to improve $\overline{\sigma}_{lsout}$ throughout all layouts showing the most consistent behaviour in avoiding peak workload. This can be explained by the workload split between multiple cranes offering the opportunity to balance waterside and landside jobs more smoothly. The priority for waterside jobs in the simulation model is not as distinct as with a SRMG-system as one crane in the TRMG- and TriRMG-system is responsible for managing the landside, thus, reducing the previously mentioned disadvantages for these jobs.

C. Block layout

While the previous investigation focused on the algorithms' properties regarding different RMG-systems, the following examination promotes a deeper insight into their behaviour for different block layouts by presenting *all instances* of the simulation experiment graphically with respect to the layout. For this purpose,

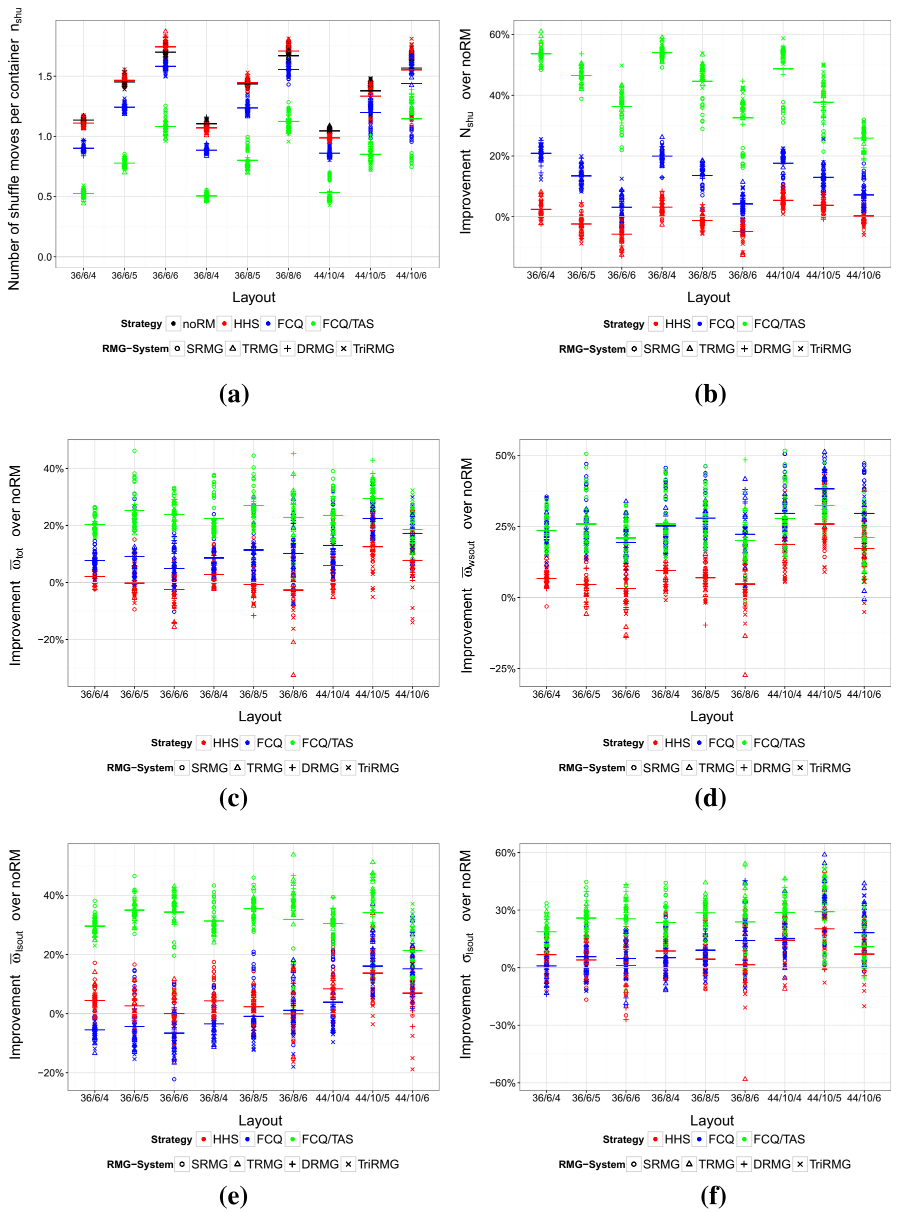

Fig. 3 Comparison of tested algorithms for selected performance indicators for all RMG-system and layouts. **a** Number of shuffle moves per container, **b** number of shuffle moves, **c** total vehicle waiting time, **d** AGV waiting time for retrieval jobs, **e** XT waiting time for retrieval jobs, **f** standard deviation of XT vehicle waiting time for retrieval jobs

Fig. 3 shows all runs conducted for all RMG-layout combinations. The horizontal/transverse lines show the corresponding mean values of the performance indicators for all RMG-systems per layout.

C1. Container handling

Figure 3a, b display all instances for n_{shu} and the relative improvement of N_{shu} for the tested algorithms over no re-marshalling, respectively. It can be seen that the main layout dimension affecting performance of all algorithms is the stack height. Both versions of FCQ follow the same pattern regarding the layout performing significantly better than noRM in nearly every layout. While the layouts with four tiers achieve an improvement of $\sim 50\%$ independently of their capacity, a drop of the results in the layouts with five tiers ($\sim 45\%$) and an even stronger one in the layouts with six tiers (~ 30–35%) can be observed. Moreover, the improvements in the case of the SRMG are at the lower limit of all RMG-systems in the five- and six-tiered layouts indicating that one crane may not have enough resources to fully translate additional TAS information into a complete re-marshalling schedule. The more stable performance for the largest layout with all RMG-systems (SRMG in particular) compared to the other two layouts with six tiers can be explained by the stacking density. In the largest layout, a sorting into *TS* for clean candidate stacks is facilitated as there are potentially more suitable stacks available. In contrast, the other two layouts have the same height but contain fewer stacks. In combination, high stacking with narrow and short layouts lead to a higher share of highly mis-overlayed stacks obstructing the finding of feasible re-marshalling moves. Besides, HHS shows issues in N_{shu} in high and medium layouts with five and six tiers dropping below noRM. Although HHS uses category stacking to mitigate future shuffle moves, it appears that an exclusively distance-based approach to re-marshalling can even induce additional mis-overlays.

C2. Vehicle waiting times

Improvements over noRM for the vehicle waiting time in total, for waterside and landside retrieval jobs are depicted in Fig. 3c–e. The FCQ without TAS slightly improves with increasing capacity while the FCQ/TAS has issues maintaining its performance for the layouts with six tiers, their means meeting in the largest layout for $\overline{\varpi}_{tot}$. While a high number of stacks in the largest layout is beneficial for eliminating shuffle moves by re-marshalling it is disadvantageous for algorithms not primarily focusing on distance reduction to handover areas. The elevated amount of re-marshalling jobs induced by additional TAS information poses the risk of schedule interruption described before. This effect is particularly pronounced for layouts with 44 bays where long crane portal movements arise which diminish advantages gained by the reduction of mis-overlays through enhanced TAS information. The same argument may be applied to $\overline{\varpi}_{wsout}$ for FCQ/TAS compared to FCQ in Fig. 3d. In RMG-systems with low crane resources like the SRMG, it is possible that some necessary re-marshalling jobs of export containers are not performed due to more re-marshalling jobs of import containers. Therefore, the re-marshalling capacity of such systems is insufficient leading to a slight deterioration of AGV waiting time with FCQ/TAS in bigger layouts. Note that $\overline{\varpi}_{wsin}$ with TAS is improving over FCQ (see Table 7), thus, outbalancing unfavourable waiting time effects at the waterside. For the same reason, the HHS behaves similarly with regard to the shuffle moves for compact to medium layouts while the pattern for large layouts aligns with FCQ as the distance-based re-marshalling becomes more beneficial in this case. Regarding $\overline{\omega}_{lsout}$ in Fig. 3e, the highly beneficial effect of TAS can be explained with the previously observed drop to an improvement of 20% on average in the largest layout. Notably, the effect of neglecting import containers by FCQ without TAS abates with increasing block capacity even leading to an improvement for large layouts.

C3. Standard deviation of vehicle waiting times

In Fig. 3f, TAS demonstrates a highly significant improvement of σ_{lsout} over noRM for all layouts. As with the other indicators, FCQ/TAS shows a robust performance except for the drop in the largest layout. In particular, FCQ without TAS is able to balance the crane workload with respect to the landside even for layouts where $\overline{\omega}_{lsout}$ is deteriorating. In cases where FCQ has a beneficial effect on $\overline{\omega}_{lsout}$, the reductions in σ_{lsout} are comparatively higher. Although this may suggest a more balanced crane workload for landside retrieval jobs, it also implies a higher overall level of workload as the waiting time increases and the standard deviation of this waiting time is mitigated. However, this does not contradict the general assumption that increasing the yard productivity as a whole may smooth crane workload at both handover areas.

Summary of performance of the FCQ-algorithm

Following the preceding analysis of re-marshalling influence in general and algorithm performance in particular, an evaluation of the formulated hypotheses can be conducted.

It can be concluded that by shifting unproductive moves from times of container retrieval to times of no crane workload a significant and substantial reduction of vehicle waiting time at the handover areas can be achieved. Enhanced information of XT arrivals lead to more opportunities for re-marshalling which also increase the number of fitting re-marshalling moves. The SRMG-system seems to lack resources to perform a complete re-marshalling schedule in some configurations, thus, not benefiting entirely from TAS information especially in large layouts. The analysis shows that the FCQ independent of TAS information is able to improve the relevant metrics for container terminal operation. While the FCQ is tailored to target waterside retrieval jobs which are prioritised by terminal operators, the heuristic is also able to incorporate updated XT arrival information to improve landside retrieval jobs while mainly maintaining the performance on the waterside. Finally, the results for σ_{lsout} are ambiguous depending on the RMG-system used. Nevertheless, significant improvements in the standard deviation of total vehicle waiting time can be observed with TAS in comparison to without TAS due to expanded yard productivity. However, contrasting FCQ/TAS with noRM and HHS, highly significant reduction in the standard deviation of all vehicle waiting times can be observed. Hence, peaks are mitigated and crane workload balance can be achieved without interfering with XT arrival plans. In summary, Table 8 shows the hypotheses results according to the conducted evaluation. In cases, where a hypothesis (formulated as alternative hypothesis in the test) is confirmed for all system-layout combinations it is stated so, otherwise only the system-layout combination for which the hypothesis is not confirmed is listed.

Table 8 Results of hypotheses testing for FCQ performance analysis

Type	No.	Hypothesis test result
Container handling	1.1	Confirmed
	1.2	Confirmed
	1.3	Not confirmed for all RMG-systems in compact layouts
		Not confirmed for multi-RMG-systems in medium layouts
Vehicle waiting time at handovers	1.4	Confirmed
	1.5	Confirmed
Standard deviation of vehicle waiting time	1.6	Not confirmed for SRMG
		Not confirmed for TRMG and DRMG in large layouts

5.2.2 TAS parameter sensitivity analysis

After demonstrating the effectiveness of the proposed solution approach, it is subsequently possible to conduct a sensitivity analysis concerning the quantity, quality and reliability of XT arrival information through TAS. On the one hand, it is intended to identify critical TAS parameters for the performance of FCQ. On the other hand, it is aimed at supporting and facilitating the design and operation of TAS by terminal operators. Each of the parameters discussed in the experimental setup (Sect. 5.1) is investigated by incrementing stepwise from FCQ without TAS to FCQ/TAS as used for the performance evaluation above. For reasons of clarity and space constraints, the results of the medium 36/8/4 layout will be presented graphically.

In more detail, the TAS parameter sensitivity analysis is organised as follows: as a **pre-evaluation** step, the approximate magnitude and direction of the Kendall-tau-b correlation coefficient for selected performance indicators and the respective TAS parameter are presented. This can be used to assess if first patterns regarding the layout types can be observed while also giving an overview of the subsequent more focused analysis. Afterwards, **general observations** with respect to layouts and RMG-systems are discussed in more detail. In this context, the principle focus is put on the main performance indicators of container handling (N_{shu}, N_{rm}, n_{rm}) and XT waiting time for retrieval jobs ($\overline{\omega}_{lsout}$), which is the directly targeted indicator by TAS. Subsequently, particularities of **compact layouts** (36/6/*) and **large layouts** (44/10/*) in comparison to medium layouts (36/8/*) are briefly investigated within a separate paragraph of each parameter discussion. For this purpose, the examined performance indicators resulting form a specific parameter setting are normalised by dividing them by the value of this indicator in case of no TAS. Thus, the figures show the relative magnitude in comparison to no XT arrival information. Finally, the slot length and announcement time must be analysed with reference to the average container dwell time in the block which is chosen as five days by default. In correspondence to the strategy comparison, the hypotheses in Table 9 are

Table 9 Hypotheses for TAS parameter sensitivity analysis

Type	No.	Hypothesis
Participation share	2.1	Increasing the participation share decreases $\overline{\omega}_{lsout}$
	2.2	Increasing the share of booked slots increases N_{rm}
	2.3	Increasing the share of correct arrivals decreases n_{rm}
	2.4	SRMG benefits the least with increasing share of booked slots in six-tiered layouts
Slot length	2.5	Increasing the slot length increases $\overline{\omega}_{lsout}$
	2.6	Increasing the slot length decreases N_{rm}
	2.7	Increasing the slot length increases N_{shu}
Announcement time	2.8	Increasing the announcement time decreases $\overline{\omega}_{lsout}$
	2.9	SRMG benefits the least with increasing announcement time in six-tiered layouts
Deviation from the announced time	2.10	Increasing the deviation from the announced time increases n_{rm}
	2.11	Increasing the deviation from the announced time increases σ_{tot}

N_{shu} total number of shuffle moves, n_{rm} number of re-marshalling moves to eliminate a shuffle move, N_{rm} total number of re-marshalling moves, $\overline{\omega}_{lsout}$ mean XT waiting time for retrieval jobs, σ_{tot} standard deviation of total vehicle waiting time

formulated about the behaviour of FCQ with TAS for different TAS settings and designs:

Firstly, a high participation share should support a reduction in landside waiting time. The more (accurate) information is available about incoming XTs the more precise is the calculation of mis-overlays for import containers classifying their stacks as *critical* in *SS* and making them potential re-marshalling candidates (Hypothesis 2.1). As a consequence, the amount of XT arrival information affects the amount of re-marshalling jobs as import containers are gradually added to *SS* with increasing share of booked slots (Hypothesis 2.2). In contrast, the share of correct arrivals within the booked time slot should reduce the needed number of re-marshalling moves to mitigate shuffling because re-marshalling moves do not have to be corrected in the future due to inaccurate or changing XT arrival information (Hypothesis 2.3). The previously conducted algorithm evaluation shows performance issues regarding vehicle waiting times of SRMG-systems in six-tiered layouts when full TAS information is accessible. As a result, Hypothesis 2.4 states that the SRMG should not benefit as much as high resource RMG-systems with expanding candidate stacks in *SS* due to TAS.

Secondly, the slot length impacts the granularity of mis-overlay calculation for import containers. For instance, import containers can only be grouped in stacks with same booked slots if XT arrival data are available on a daily basis whereas they can be arranged according to the specific sequence within the stack in case of hourly slots. Hence, it is possible to realise the retrieval sequence in a stack more accurately with shorter slots reducing $\overline{\omega}_{lsout}$ (Hypothesis 2.5). The influence on the re-marshalling quality is not clear a priori. It can only be assumed that having less detailed XT arrival data generates fewer calculated mis-overlays which leads to a decrease in potentially *unordered* stacks in *SS*. As a consequence, the amount of re-marshalling moves will be smaller (Hypothesis 2.6). In contrast, less precise mis-overlay calculation due to longer slot lengths will lead to more undetected future shuffle moves by FCQ, thus, increasing shuffling (Hypothesis 2.7).

Ideally, if XTs announce their arrival time prior or close to the initial storage time of the container, stacking in the block can be carried out without inducing any mis-overlays eliminating the need for re-marshalling. All experiments in this work are conducted with an average container dwell time of five days making it impossible for XT companies to adopt a planning horizon of that length to enable an announcement of container pick-ups immediately. However, the earlier XT companies set their vehicle schedule the sooner mis-overlay calculations for an extended number of import containers can be performed. As a consequence, it becomes easier to establish the appropriate retrieval sequence in stacks leading to a reduction in $\overline{\omega}_{lsout}$ (Hypothesis 2.8). In accordance with the argument for the share of booked slots, the enlarged pool of candidate stacks in *SS* with longer announcement times aggravates a full execution of the re-marshalling schedule for low resource systems like the SRMG leading to Hypothesis 2.9. Statements about the behaviour concerning the split in container handling are difficult to make. The announcement time does not only increase the amount of available XT arrival information like the share of booked slots but also specifies the look-ahead time for re-marshalling making the mis-overlay calculation more precise. While the former property tends to increase the number of re-marshalling moves, the latter property makes the moves more fitting, thus, reducing the needed amount.

Finally, it is important to consider the deviation from the announced time of XT arrivals (if not on time). In particular, high deviations of more than one day, which may be regarded as no-shows, conflict with the rearranged container sequence in accordance with the actually booked data. This should have a disadvantageous impact on the re-marshalling quality, as containers with late retrievals or no-shows must be subsequently relocated again (Hypothesis 2.10). According to Hypothesis 2.11 this may also lead to an imbalance in crane workload as seemingly *ordered* stacks are not regarded in re-marshalling schedules anymore, although being actually *unordered*. Thus, containers with late XT arrivals or no-shows can substantially block containers due to the mis-classification of the entire stack inducing high waiting time for any container type beneath.

A. Participation share

By definition, the participation share has a two-fold structure. Therefore, the analysis serves two functions for evaluating the TAS design and working principle: Firstly, the share of booked slots indicates for how many import containers in the block mis-overlays can be calculated within a given look-ahead time (see announcement time below) and, consequently, how extensive the re-marshalling schedule will be. Thus, it is a measure for **information quantity**. Secondly, the share of correct arrivals does not influence the amount of information but it indicates how dependable the given arrival information is. Accordingly, it is a parameter for **information reliability**.

Table 10 Overview of Kendall-tau-b correlation ($\alpha = 5\%$) between selected performance indicators and participation share

Participation share	RMG-system	Layout type	N_{shu}	N_{rm}	$\overline{\omega}_{tot}$	$\overline{\omega}_{wsout}$	$\overline{\omega}_{lsout}$	σ_{tot}
Share of booked slots	SRMG	Tiers: 4/5/6	↓/↓/↓	↑/o/-	↘/→/-	o/o/→	↓/↘/→	→/o/-
	TRMG		↓/↓/↓	↑/→/o	↓/-/-	o/o/o	↓/↓/-	↘/→/o
	DRMG		↓/↓/↓	↑/→/o	↓/-/-	o/o/o	↓/↓/-	↘/→/-
	TriRMG		↓/↓/↓	↑/↗/→	↓/↓/↓	o/o/o	↓/↓/↓	↓/↓/-
Share of correct arrivals	SRMG	Tiers: 4/5/6	↓/↘/→	↘/→/→	↘/o/o	o/o/o	-/o/o	o/o/o
	TRMG		↓/↓/↓	↘/↘/→	↓/↘/o	o/o/o	↓/↓/-	-/o/o
	DRMG		↓/↓/↓	↘/→/→	↓/↘/-	o/o/o	↓/↓/-	↘/→/-
	TriRMG		↓/↓/↓	↘/↘/→	↓/↓/-	o/o/o	↓/↓/↓	↘/↘/-

↓: very strong negative correlation $\tau_b \leq -0.7$ | ↘: strong negative correlation $-0.7 < \tau_b \leq -0.5$ | →: moderate negative correlation $-0.5 < \tau_b \leq -0.3$ | o: no/weak negative correlation $-0.3 < \tau_b \leq 0$ | o: no/weak positive correlation $0 \leq \tau_b < 0.3$ | →: moderate positive correlation $0.3 \leq \tau_b < 0.5$ | ↗: strong positive correlation $0.5 \leq \tau_b < 0.7$ | ↑: very strong positive correlation $0.7 \leq \tau_b$, | -: no identified pattern | symbols for N_{rm} are not coloured as either a negative or positive correlation may be beneficial for re-marshalling

N_{shu} - total number of shuffle moves, N_{rm} - total number of re-marshalling moves, $\overline{\omega}_{tot}$ - mean total vehicle waiting time, $\overline{\omega}_{wsout}$ - mean AGV waiting time for retrieval jobs, $\overline{\omega}_{lsout}$ - mean XT waiting time for retrieval jobs, σ_{tot} - standard deviation of total vehicle waiting time

Pre-evaluation of correlation

Prior to evaluating the layouts and RMG-systems for the participation share in detail, the approximate magnitude and direction of the Kendall-tau-b correlation coefficients ($\alpha = 5\%$) are listed in Table 10 for selected performance indicators with the participation share.

It can be seen that both shares have a high influence on shuffling while only four-tiered layouts are highly sensitive concerning re-marshalling. Overall, the TriRMG-system seems to be most sensitive to parameter changes with respect to container handling. $\overline{\omega}_{tot}$ is mainly influenced by landside operation as the waterside seems to be nearly unaffected by changes in the participation share. However, the pattern for the six-tiered layouts are not consistent showing mostly weak or moderate negative correlations for $\overline{\omega}_{tot}$ and $\overline{\omega}_{lsout}$. Generally, a beneficial tendency can be observed for σ_{tot} with increasing participations share, although showing only a weak/moderate negative correlation in comparison to the other indicators. The low crane resource SRMG-system seems to experience the lowest advantages with respect to this indicator while the TriRMG has the highest sensitivity again. Altogether, the pre-evaluation indicates that the participation share is highly relevant for assessing the FCQ behaviour and TAS design by being (highly) correlated with the targeted performance indicators. A further analysis with respect to the number of tiers seems appropriate.

A.1 General obervations (container handling and landside retrieval effects)

In Fig. 4 the influence of the participation share on container handling and vehicle waiting times for the medium layouts with four tiers is depicted. The results for each RMG-system are supplemented with a linear regression. Consistent trends can be observed for all indicators. With reference to container handling in Fig. 4a, b two effects can be distinguished. On the one hand, if the share of booked arrivals increases (green), i.e more retrieval data for import containers become available, the number of re-marshalling moves increases accordingly as more import containers are potentially suitable to be queried from *SS*. The high number of moves seems to be beneficial for the elimination of shuffle moves as they are reduced with more booked slots. On the other hand, if all XTs book a slot (red) a decline in re-marshalling moves is observable with increasing accuracy while shuffle moves decrease simultaneously. In total, fewer re-marshalling moves are needed to reduce

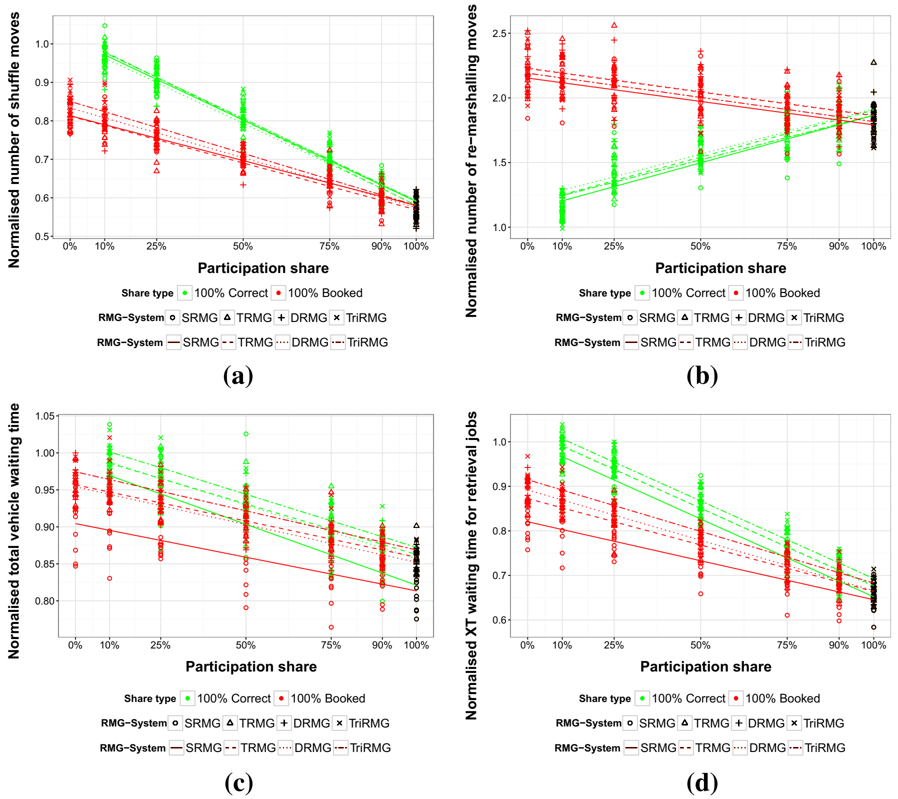

Fig. 4 FCQ behaviour for participation shares in 36/8/4. **a** Shuffle moves, **b** re-marshalling moves, **c** total vehicle waiting time, **d** XT waiting time for retrieval jobs

shuffling even further, resulting in a substantial benefit in n_{rm}. Regarding vehicle waiting times, a behaviour corresponding to shuffling can be observed. Primarily, TAS supports the elimination of mis-overlays of import containers as evident in Fig. 4d for both types of shares. These benefits translate to $\overline{\omega}_{tot}$ in Fig. 4c, although with a smaller magnitude since waterside operation is not significantly affected by changes in the participation share. Another point of consideration alludes to apparent advantages of having 100% of arrivals booked in contrast to accessing 100% correct data. As the red instances demonstrate the behaviour of the former, it can be seen that these results outperform the latter with respect to shuffling and waiting times. For these experiments, the default maximum deviation from the announced time in case of incorrect arrivals is set to 12 h. Taking this into account, it can be argued that the availability of XT arrival information seems preferable to highly precise information in cases of low to mid maximum deviation from the announced time. Naturally, the performance gap diminishes with increasing participation by XT as the two shares approach each other steadily.

The benefits of TAS in five-tiered layouts are still significant for all RMG-systems, albeit with different behaviour between SRMG- and multi-RMG-systems. The SRMG struggles to obtain significant improvements with increasing share of correctness of XT arrivals in this case. Notably, SRMG-systems seem to be insensitive to variations regarding the share of correct arrivals as a result of difficulties in translating TAS information into vehicle waiting time improvements with 100% booked arrivals.

Layouts with six tiers mostly continue the development observed for the performance indicators with increasing stacking height. However, the amount of re-marshalling moves decreases with increasing participation independent of the share analysed. The reduction in re-marshalling moves can be explained by the risk of additional mis-overlays in layouts with six tiers that impede re-marshalling as the number of candidate target stacks *TS* in relation to the number of containers is smaller than in blocks with a lower number of tiers. As a consequence, this leads to an increased intricacy of finding an *ordered* stack to re-marshal to. Due to the lack of sufficiently available *ordered* stacks, re-marshalling moves are skipped until a new stack reaches the threshold of being more *ordered* than *unordered* ($\mu_s^{ordered} > 0.5$). Additionally, high resource multi-RMG-systems execute more re-marshalling jobs within a given period than the SRMG. Thus, they are able to perform the feasible re-marshalling schedules fully for stacks where stacking restrictions are not active. In combination, the few re-marshalling jobs, that are possible, are performed completely by these systems while additional jobs are not initiated due to the low number of *ordered* stacks and their respective stacking restrictions. However, the executed re-marshalling moves are more fitting with increasing share of booked slots illustrated by a decrease in N_{shu}. With reference to shuffling in multi-RMG-systems, its magnitude of elimination is still similar to the five-tiered layouts indicating the ability of these systems to increase the re-marshalling quality with additional and more precise TAS information independent of block density or stacking height. In contrast, there are pronounced difficulties in the SRMG case to achieve benefits even for landside retrieval jobs, thus, diminishing the beneficial effects on $\overline{\omega}_{tot}$ substantially.

A.2 Observations for compact layouts

In general, compact layouts show the most consistent patterns regarding the participation share. In these layouts the SRMG is most able to match the performance of the multi-RMG-systems. This may be due to a less disadvantageous impact of re-marshalling interruptions by incoming storage and retrieval jobs because crane portal and trolley distances are smaller in comparison to other layouts.

A.3 Observations for large layouts

Large layouts, especially with five and six tiers, exhibit a different preference of having 100% booked slots and 100% correct arrivals. While the former was identified as the more preferable for compact and medium layouts, the latter shows better result for large layouts regarding relevant performance indicators. This can be explained by the phenomenon, that more accurate information is more relevant for layouts where a higher potential of long crane movements is prevalent. If there is less XT arrival information, less re-marshalling moves are induced meaning that frequent crane movements over long distances during re-marshalling are mitigated.

Moreover, accurate arrival information helps to reduce correcting re-marshalling moves as a result of unfitting prior ones. Hence, the reduction of (unsuitable) re-marshalling has a notably beneficial effect for large layouts where interruptions by incoming retrieval or storage jobs are more critical due to long crane movements. Another aspect is the insensitivity of the largest layout (44/10/6) to the participation share. This insensitivity is most prevalent for the SRMG-system where almost all performance indicators show no change with increasing participation. The TRMG exhibits significant results only for N_{rm} and N_{shu}, while the DRMG- and even more the TriRMG-system benefit decidedly from an increased participation share. The better performance of the DRMG over the TRMG may be explained by the crossing ability of the former. When the landside crane of the TRMG-system is performing re-marshalling, a landside retrieval or storage job must wait until the re-marshalling job is finished as a result of the inability of the waterside crane to take the job instead. In the DRMG-system, a crane can serve both handover areas and, thus, can be assigned a job if it is disadvantageous to wait for the other crane to finish its re-marshalling job. Generally, this indicates that only high crane resource systems with workload sharing (and crossing ability in large layouts) are able to cope with an extensive re-marshalling schedule while fulfilling storage and retrieval jobs at the handover areas.

B. Slot length

While the participation share impacts both, quantity (share of booked slots) as well as reliability (share of correct arrivals) of TAS information, the slot length is a parameter determining the precision of given information. Thus, it characterises the TAS design regarding the **information quality**.

Pre-evaluation of correlation

Prior to evaluating the layouts and RMG-systems for the slot length in detail, the approximate magnitude and direction of the Kendall-tau-b correlation coefficients ($\alpha = 5\%$) are listed in Table 11 for selected performance indicators with the slot length. A systematic distinction between the different layouts according to the number of tiers cannot be made as was the case for the participation share. Thus, the analysis is conducted with respect to the three layout types (compact, medium, large).

Broadly, all performance indicators seem to be distinctly less sensitive to changes of the slot length. Re-marshalling is only moderately correlated in most cases while

Table 11 Overview of Kendall-tau-b correlation ($\alpha = 5\%$) between selected performance indicators and slot length

RMG-system	Layout type	N_{shu}	N_{rm}	$\overline{\omega}_{tot}$	$\overline{\omega}_{wsout}$	$\overline{\omega}_{lsout}$	σ_{tot}
SRMG	compact/medium/large	o/o/→	-/→/o	o/o/o	→/→/o	o/o/o	o/o/o
TRMG		→/→/o	→/↘/→	→/o/o	→/-/o	-/-/-	→/o/o
DRMG		↗/↗/↗	→/→/→	↗/→/-	→/↗/→	↗/→/→	→/→/o
TriRMG		↗/↗/↑	→/→/→	→/↗/↗	o/o/o	→/↗/↗	→/→/→

↓: very strong negative correlation $\tau_b \leq -0.7$ | ↘: strong negative correlation $-0.7 < \tau_b \leq -0.5$ | →: moderate negative correlation $-0.5 < \tau_b \leq -0.3$ | o: no/weak negative correlation $-0.3 < \tau_b \leq 0$ | o: no/weak positive correlation $0 \leq \tau_b < 0.3$ | →: moderate positive correlation $0.3 \leq \tau_b < 0.5$ | ↗: strong positive correlation $0.5 \leq \tau_b < 0.7$ | ↑: very strong positive correlation $0.7 \leq \tau_b$, | -: no identified pattern | symbols for N_{rm} are not coloured as either a negative or positive correlation may be beneficial for re-marshalling

N_{shu} - total number of shuffle moves, N_{rm} - total number of re-marshalling moves, $\overline{\omega}_{tot}$ - mean total vehicle waiting time, $\overline{\omega}_{wsout}$ - mean AGV waiting time for retrieval jobs, $\overline{\omega}_{lsout}$ - mean XT waiting time for retrieval jobs, σ_{tot} - standard deviation of total vehicle waiting time

shuffling is disadvantageously affected although not to the same degree as with the participation share. The SRMG- and TRMG-systems seem to be particularly insensitive to the slot length, which also translates to the weak correlation for $\overline{\omega}_{tot}$ and $\overline{\omega}_{lsout}$. Vehicle waiting times seem to be more correlated with the slot length in DRMG- and TriRMG-systems, albeit not reaching the level of the participation share as well. Notably, the SRMG benefits from higher slot length concerning N_{shu} in large layouts and $\overline{\omega}_{lsout}$ for all layouts which appears counter-intuitive as the information becomes less precise. Otherwise, the pre-evaluation of performance indicators predominantly suggests a moderate disadvantageous effect of increasing the slot length.

B1. General observations (container handling and landside retrieval effects)
In Fig. 5 the influence of the slot length on container handling and waiting time in the medium 36/8/4 layout is shown roughly representing the behaviour of all layouts with 36 bays. Concerning re-marshalling depicted in Fig. 5a, b reduction in re-marshalling moves with increasing slot length can be observed. Nonetheless, there is no significant difference for slot lengths of up to 6 h. Only the longest two slot lengths show a decline in re-marshalling moves compared to shorter ones. Similarly,

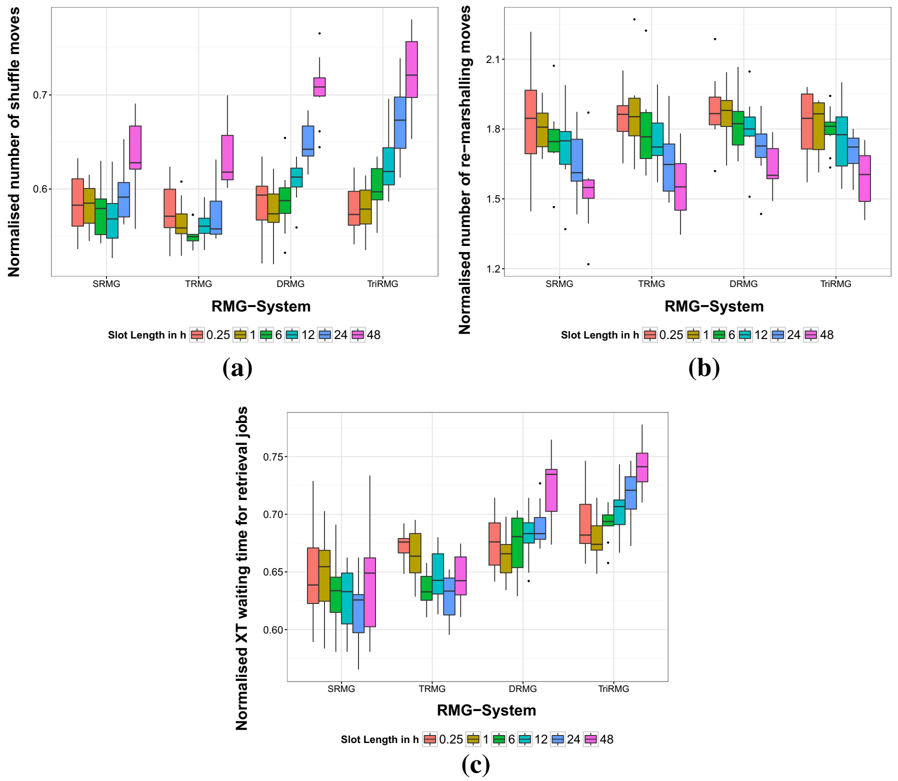

Fig. 5 FCQ behaviour concerning container handling and waiting time for slot lengths in 36/8/4. **a** Shuffle moves, **b** re-marshalling moves, **c** XT waiting time for retrieval jobs

shuffling in layouts with 36 bays is significantly influenced by high sloth lengths while the algorithm exhibits stable behaviour for slots of up to 12 h (Fig. 5a). The most pronounced effects appear for the DRMG- and TriRMG-system which are more sensitive to lengths starting from 12 h in all layouts. These systems with crossing capabilities have more resources to perform a full re-marshalling schedule and consequently benefit from a more precise mis-overlay calculation induced by a higher resolution of XT arrival data. The insensitivity to slot lengths between 0.25 and 12 h can be explained as follows: small slot lengths characterise more stacks as *unordered* because the mis-overlay calculation is more precise putting more stacks in *SS*. One resulting effect is that more stacks could be transformed into *ordered* stacks. Another contrasting effect reduces the number of *ordered* stacks in *TS* due to the more accurate mis-overlay calculation. Particularly small slot length shift this balance towards *SS* while long slot lengths shift it towards *TS*. Medium lengths keep a balance between this trade-off. Thus, only exceptionally long slots induce a disadvantageous behaviour of the FCQ regarding shuffling. With reference to the XT waiting time for retrieval jobs, the SRMG- and TRMG-system do not show (highly) significant impacts by a change of the slot length while the DRMG is only affected by the 48 h slot (Fig. 5c). The TriRMG-system exhibits the most consistent behaviour for all medium layouts independent of the number of tiers, being significantly sensitive to disadvantageous increases of almost any slot length.

B2. Observations for compact layouts
Compact layouts show the most pronounced balancing conflict between *SS* and *TS* with changing slot length. The tendency of a U-shaped behaviour similar to Fig. 5a in the medium layouts is particularly noticeable for N_{shu} and $\overline{\omega}_{lsout}$. This effect is due to the low number of stacks in comparison to the stack height of these layouts. Thus, intermediate slot lengths show the best results, balancing the need for a precise mis-overlay calculation and a fitting activation of stacking restrictions. In addition, there is a distinct reduction in transforming shuffle moves into re-marshalling moves in compact layouts. N_{rm} strongly decreases while N_{shu} strongly increases in comparison to the medium layouts. The lower number of stacks in compact layouts leads to an overall lower number of stacks in *SS*, reducing the pool of potential containers to be re-marshalled. Imprecise mis-overlay calculation resulting from long slot lengths leads to a lower number of detected mis-overlays within the limited number of stacks in compact layouts, thus, fitting re-marshalling moves are induced less frequently. As a consequence, an increase of shuffling can be observed.

B3. Observations for large layouts
FCQ is more insensitive to changes in the slot length for large layouts, in particular for SRMG-systems. Significant changes can be observed only for 48 h most of the time. Shuffling shows a similar behaviour for multi-RMG-systems, the high resource TriRMG being the only system showing an increase in shuffle moves starting from 12 h. Interestingly, the SRMG shows an unexpected behaviour in large layouts. Although the precision of arrival information is decreasing, shuffling drops as well. This observation and the reduced sensitivity of multi-RMG-systems may be

due to long crane movements in large layouts. The increased amount of re-marshalling jobs with short slot lengths increases the risk of interrupting the re-marshalling schedule when the crane is far away from the incoming storage or retrieval job. Resulting high crane empty movements deplete the available time for re-marshalling which would be particularly useful for short slot lengths with precise mis-overlay calculation. As the SRMG has the lowest crane resources, re-marshalling schedules are performed incompletely. Interruptions occur frequently with short slot lengths. These impede the full clean up of a specific stack while stacking restrictions are activated due to partially *ordered* stacks invalidating them as potential stacks which offer target slots for re-marshalling. The argument of balancing *SS* and *TS* described above for the container handling can be used to explain the strong insensitivity of FCQ for the SRMG- and TRMG-system up to 24h regarding $\overline{\omega}_{lsout}$ in large layouts. The TriRMG-system shows significant results starting from 12 h in most cases maintaining the findings that the high resource system is able to make use of precise XT arrival information. While the DRMG-system performs on par with the TriRMG-system for compact and medium layouts, it aligns with the other systems for the large layout. Overall, in comparison to the participation share whose variation can influence shuffle moves up to 30%-points, the slot length has a rather moderate range of up to 15%-points in the worst case of a practically unrealistic slot length of 48 h. It results from the aforementioned analysis of the trade-off in classification between the *SS* and *TS*.

C. Announcement time

The announcement time of XT arrivals specifies the look-ahead horizon for the mis-overlay calculation and the resulting re-marshalling frequency. With long announcement times, the number of containers assigned to time slot based fixed departure times (*category 1*) increases accordingly. Thus, the announcement time is mainly a parameter influencing the **information quantity** of available XT arrivals.

Pre-evaluation of correlation

Prior to evaluating the layouts and RMG-systems for the announcement time in detail, the approximate magnitude and direction of the Kendall-tau-b correlation coefficients ($\alpha = 5\%$) are listed in Table 12 for selected performance indicators with the announcement time. For this TAS parameter, neither the distinction between tiers nor between layout types is as pronounced as for the two previous parameters.

Table 12 Overview of Kendall-tau-b correlation ($\alpha = 5\%$) between selected performance indicators and announcement time

RMG-system	Layout type	N_{shu}	N_{rm}	$\overline{\omega}_{tot}$	$\overline{\omega}_{wsout}$	$\overline{\omega}_{lsout}$	σ_{tot}
SRMG	Tiers: 4/5/6	↘/→/→	-/-/-	→/o/o	o/→/→	↘/o/o	-/o/o
TRMG		→/-/-	o/o/↘	→/o/o	o/-/-	↘/↘/o	o/o/o
DRMG		↘/-/-	o/o/↘	→/o/→	o/↗/↗	↘/-/↘	→/o/→
TriRMG		→/-/-	o/o/↑	→/o/o	o/→/→	↘/↓/-	→/o/o

↓: very strong negative correlation $\tau_b \leq -0.7$ | ↘: strong negative correlation $-0.7 < \tau_b \leq -0.5$ | →: moderate negative correlation $-0.5 < \tau_b \leq -0.3$ | o: no/weak negative correlation $-0.3 < \tau_b \leq 0$ | o: no/weak positive correlation $0 \leq \tau_b < 0.3$ | →: moderate positive correlation $0.3 \leq \tau_b < 0.5$ | ↗: strong positive correlation $0.5 \leq \tau_b < 0.7$ | ↑: very strong positive correlation $0.7 \leq \tau_b$, | -: no identified pattern | symbols for N_{rm} are not coloured as either a negative or positive correlation may be beneficial for re-marshalling

N_{shu} - total number of shuffle moves, N_{rm} - total number of re-marshalling moves, $\overline{\omega}_{tot}$ - mean total vehicle waiting time, $\overline{\omega}_{wsout}$ - mean AGV waiting time for retrieval jobs, $\overline{\omega}_{lsout}$ - mean XT waiting time for retrieval jobs, σ_{tot} - standard deviation of total vehicle waiting time

However, as the patterns regarding the tiers are more prevalent, the pre-evaluation is conducted according to this layout parameter.

The correlation of the announcement time is mainly weak or moderate for all performance indicators. Even N_{shu} which shows the highest sensitivity to the previous two TAS parameters demonstrates only moderate beneficial effects with increasing announcement time while the behaviour for five- and six-tiered blocks depends more on the layout type than on the number of tiers which is not depicted in Table 12. N_{rm} is even more insensitive to the announcement time than to the slot length indicating no correlation in the majority of cases. Four-tiered layouts seem to experience the most consistent moderate negative correlation while blocks with increased hight see no or weak correlation only. $\overline{\varpi}_{lsout}$ appears to be influenced by the announcement time the most, showing only a strong correlation, in particular for four-tiered layouts. Nonetheless, these effects are balanced by moderate to strong positive correlations at the waterside for five- and six-tiered layouts. Thus, $\overline{\varpi}_{tot}$ and σ_{tot} are weakly or not at all correlated with the announcement time in theses cases. Unlike for the other two parameters, the TriRMG-system does not have a distinguished pattern that could imply beneficial effects of high crane resources. In spite of influencing the information quantity like the share of booked slots, this pre-evaluation suggests that similar benefits cannot be obtained by increasing the announcement time using FCQ. Possible reasons for this ambivalent behaviour are investigated in the following.

C1. General observations (container handling and landside retrieval effects)
In Fig. 6 the results for container handling in medium layouts with four tiers are shown. A similar behaviour for all RMG-systems can be observed. There seems to be a parabolic pattern with increasing announcement time. This is particularly pronounced for re-marshalling depicted in Fig. 6b with approximately the same N_{rm} for very short and very long announcement times at a very low level while intermediate times ranging from 24 to 72 h exhibit no significant difference. This effect can be explained by the amount of available XT arrival information determined by the announcement time. With times of 1h there are fewer *category 1* containers, thus, the number of container in *SS* is reduced. As a consequence, fewer re-marshalling jobs will be initiated. In contrast, with 120 h of look-ahead time the re-marshalling quality improves leading to more fitting moves as a result of a better mis-overlay calculation and eliminating the need for future re-marshalling moves to correct the previous ones. Correspondingly, the intermediate announcement times show a behaviour in-between of these two extremes. Looking at the shuffle moves in Fig. 6a, no significant or only weakly significant differences between announcement times of up to 36 h can be observed. This corresponds to the behaviour of nearly all four-tiered layouts and all systems. Thus, it implies that FCQ is rather robust for practically feasible look-ahead times whereas substantial benefits only occur for announcement times close to the average container dwell time of five days. Most probably, the latter cannot be expected to be enforced in practice because XT companies are reluctant or even unable to schedule their vehicles on planning horizons so long [cf. Giuliano et al. (2007)]. It can also be seen that shuffling drops by 20%-points (30%-points in the large layout) when comparing short and intermediate times to the very long announcement time.

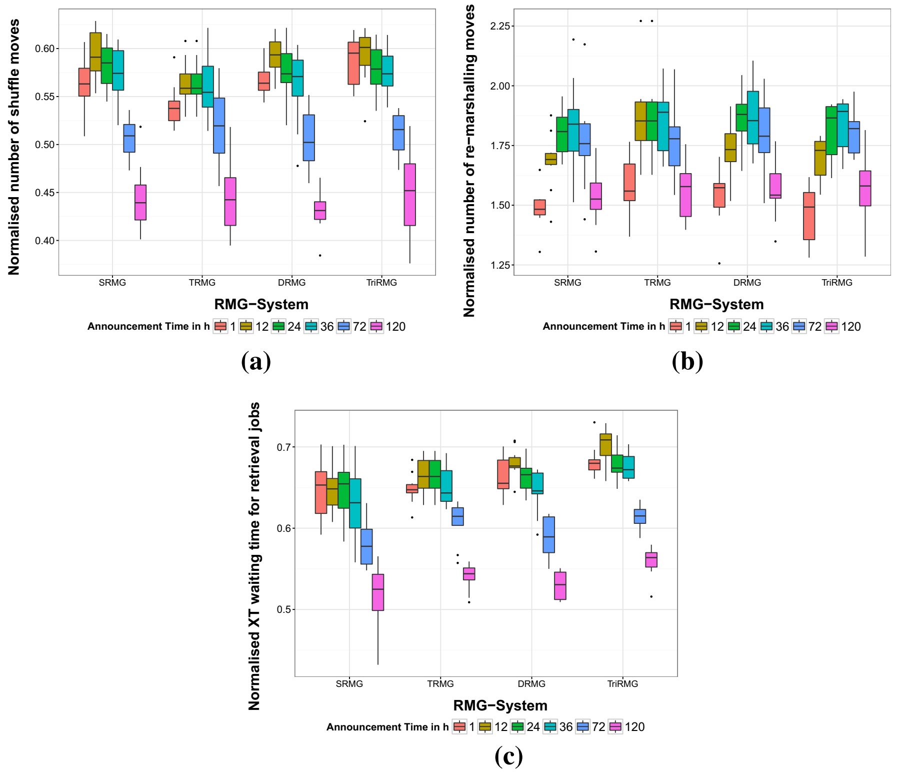

Fig. 6 FCQ behaviour concerning container handling and landside retrieval jobs for annnouncement times in 36/8/4. **a** Shuffle moves, **b** re-marshalling moves and **c** 36/8/4

Continuing with the five-tiered layouts, the differences between the announcement times are less distinct in comparison to the layouts with four tiers, although a generally similar behaviour of FCQ can be observed. In particular, the benefits of very long announcement times are not as pronounced as in the four-tiered case and very short times align more strongly with intermediate look-ahead times.

The pattern for six tiered-layouts aligns well for all layouts with respect to re-marshalling while there is no consistent behaviour regarding shuffling. Re-marshalling moves show a decidedly significant decrease with increasing announcement time for all systems. To assess the reason behind this, a look at the shuffle moves shows that the SRMG is still able to improve significantly this indicator with better look-ahead times while the multi-RMG-Systems are not affected. As the shuffling level of the latter is substantially lower, the SRMG still has more potential to re-marshal. In contrast, the multi-RMG-systems are generally more able to produce *ordered* stacks meaning that more stacking restrictions apply to prevent them turning *unordered* again. As a consequence, the following sequence of explanations can be employed summing up the main container handling behaviour of FCQ regarding the announcement time: with high announcement times there is a high amount of XT arrival information available leading to a crisp separation of *unordered* stacks in *SS* and *ordered* stacks in TS. While the former induce re-marshalling moves resulting from more precise mis-overlay calculation, the latter block re-marshalling by activating stacking restrictions due to a more precise mis-overlay calculation resulting from the utilisation of more XT arrival information.

Finally, the behaviour of FCQ regarding landside retrieval jobs is very similar in all layouts with the same number of tiers. In four-tiered layouts as depicted in Fig. 6c a direct translation of the pattern for shuffling to the $\overline{\omega}_{lsout}$ can be made. While there is a robust behaviour for short to intermediate look-ahead times in all layouts, highly significant improvements can be achieved only with impractically long announcement times. The magnitude of change in waiting time corresponds mainly to the magnitude of shuffle moves. With increasing stack height, the SRMG starts to become insensitive to announcement time changes while the multi-RMG-systems are still benefiting to approximately the same degree as the four-tiered layout. In the six-tiered layout the TriRMG-system is the only system showing this magnitude of improvement with increasing announcement time while the TRMG and DRMG become insensitive. In the 44/10/6 layout, even the TriRMG is not able to benefit from higher look-ahead times. The evaluation concerning $\overline{\omega}_{lsout}$ shows that long announcement times detect a high number of potential containers to be re-marshalled. Firstly, this leads to a more frequent activation of stacking restrictions impeding re-marshalling. Secondly, high announcement times may increase the crane workload, particularly in large and compact layouts, alleviating the benefits of additional XT arrival information even for the high resource TriRMG-system in one case.

C2. Observations for compact layouts

No specific patterns can be observed for compact layouts in comparison to medium and large layouts. However, in the most compact layout (36/6/6) the aforementioned division of *SS* and *TS* becomes most apparent. In this layout, re-marshalling decreases with increasing announcement time as no suitable re-marshalling moves can be found while shuffling deteriorates as a consequence of rigorous stacking restrictions interfering with stack clean up through re-marshalling.

C3. Observations for large layouts

In contrast to compact layouts, the SRMG-system shows the highest sensitivity to announcement time changes in large layouts compared to multi-RMG-systems. While the latter show significant improvements in $\overline{\omega}_{lsout}$ and shuffling starting from 24 h, the former shows highly significant reductions in both indicators for almost every announcement time change. The different behaviour of FCQ in large layouts may be explained by the additionally available time for re-marshalling due to longer look-ahead times. It was previously shown that the SMRG is struggling to match the re-marshalling performance of multi-RMG-systems in six-tiered and large layouts due to lower crane resources. By having longer look-ahead times, more fitting re-marshalling moves can be performed. (It can be observed that the re-marshalling quality n_{rm} is substantially improving with increasing announcement time. N_{rm} decreases heavily while N_{shu} stays constant or decreases as well). Thus, the reduced need for re-marshalling resources is particularly beneficial for large layouts where a risk of long crane distances is highest. As there is no workload sharing between cranes in the SRMG-system, this system is highly affected by this circumstance because one crane must perform all storage, retrieval, shuffle and re-marshalling jobs. As a consequence, increasing the re-marshalling quality in large layouts benefits the low resource SRMG the most while in multi-RMG-systems (especially for RMGs with crossing capabilities like DRMG- and TriRMG-systems) one crane can finish a re-marshalling job and the other takes over the incoming storage or retrieval job.

D. Deviation from the announced time

The deviation from the announced time by XT companies is closely tied to the share of correct arrivals discussed before. Naturally, the impact of the deviation is more pronounced with XT not arriving on time. It is assumed that no XT arrives on time for the experiments conducted here in order to assess the pure behaviour of FCQ with varying deviation. While the share of correct arrivals indicates how reliable the data for container retrieval estimation are, the deviation from the announced time specifies the scope of this unreliability by describing the underlying uncertainty of available information.

Pre-evaluation of correlation

Prior to evaluating the layouts and RMG-systems for the deviation from the announced time in detail, the approximate magnitude and direction of the Kendall-tau-b correlation coefficients ($\alpha = 5\%$) are listed in Table 13 for selected performance indicators with the deviation from the announced time. For this TAS parameter the patterns are more generic and cannot be specifically isolated with respect to the number of tiers or the layout type. Thus, if more than one correlation indicator is stated per cell, it only illustrates differences in patterns without assigning them to a specific tier or layout type.

The deviation from the announced time shows stronger correlations than the announcement time and slot length regarding the two chiefly targeted indicators N_{shu} and $\overline{\omega}_{lsout}$. In detail, the deviation from the announced time seems to have clear disadvantages for these two indicators. However, the SRMG-system shows inconsistent behaviour with respect to layout types and tiers. Again, re-marshalling appears to be only moderately or not affected. Correlations between $\overline{\omega}_{tot}$ and the

Table 13 Overview of Kendall-tau-b correlation ($\alpha = 5\%$) between selected performance indicators and deviation from the announced time

RMG-system	Layout type	N_{shu}	N_{rm}	$\overline{\omega}_{tot}$	$\overline{\omega}_{wsout}$	$\overline{\omega}_{lsout}$	σ_{tot}
SRMG	Unspecified	↗/o	→/o	-	o/o	-	o/o
TRMG		↑	→/o	↗/→	o	↗	o
DRMG		↑	o	↗/→	o	↗/↑	→/o
TriRMG		↑	→/o	↗	o	↑	→/o/↗

↓: very strong negative correlation $\tau_b \leq -0.7$ | ↘: strong negative correlation $-0.7 < \tau_b \leq -0.5$ | →: moderate negative correlation $-0.5 < \tau_b \leq -0.3$ | o: no/weak negative correlation $-0.3 < \tau_b \leq 0$ | o: no/weak positive correlation $0 \leq \tau_b < 0.3$ | →: moderate positive correlation $0.3 \leq \tau_b < 0.5$ | ↗: strong positive correlation $0.5 \leq \tau_b < 0.7$ | ↑: very strong positive correlation $0.7 \leq \tau_b$, | -: no identified pattern | symbols for N_{rm} are not coloured as either a negative or positive correlation may be beneficial for re-marshalling

N_{shu} - total number of shuffle moves, N_{rm} - total number of re-marshalling moves, $\overline{\omega}_{tot}$ - mean total vehicle waiting time, $\overline{\omega}_{wsout}$ - mean AGV waiting time for retrieval jobs, $\overline{\omega}_{lsout}$ - mean XT waiting time for retrieval jobs, σ_{tot} - standard deviation of total vehicle waiting time

deviation from the announced time arise almost excursively from the landside as there is only weak or no correlation at the waterside. The pre-evaluation suggests that the deviation from the announced time may have a substantial influence on FCQ behaviour. The subsequent discussion tries to have a detailed view on potential differences between low (late/early arrivals) and high (no-shows/ignoring the booked slot) deviations from the announced time.

In the following, values of respective performance indicators are depicted based on different maximum deviations from the announced time. In more detail, the deviations state the maximum difference between the announced time and the realised XT arrival time during the simulation run. Based on the generated XT arrival time in the simulation model, the announced time is generated with reference to the parameter slot length discussed previously. For example, if the realised arrival time is 01:23 pm and the slot length is 30 min, the announced time will be set between 01:00 pm to 01:30 pm. Thus, if the XT arrives during any time in this interval, it is considered to be on time. However, if the same XT announces to arrive between 12:00 pm and 12:30 pm and is still arriving at 01:23 pm, there is a deviation from the announced time of 1 h. The maximum deviation from the announced time is considered as a parameter in the simulation model in order to generate XTs arriving late. In this case, the realised deviation from the announced time is generated from a discrete uniform distribution with supports 0.5 and 1 h as its parameters, where the maximum deviation is parametrically set to 1 h ($U(0.5, 1)$). As the slot length stipulates a 30 min interval for arrival, the potential realisations of deviations from the announced time are 0.5 or 1 h only.

D1. General observations (container handling and landside retrieval effects)
Figure 7 shows the results for container handling in the medium 36/8/4 layout. Shuffling in Fig. 7b is significantly influenced by changes of deviation from the announced time. A significant upwards trend can be observed in all tested

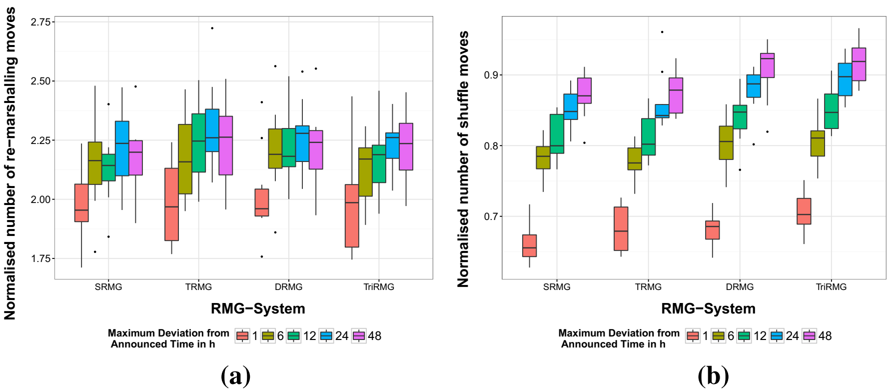

Fig. 7 FCQ behaviour concerning container handling for deviations from the announced time in 36/8/4. **a** Re-marshalling moves, **b** shuffle moves

layouts with almost all systems. In most cases a deterioration of performance by approximately 20%-points is evident comparing 1 and 48 h deviations. Higher uncertainty of XT arrivals makes the mis-overlay calculation less precise, thus, impeding an identification of fitting re-marshalling moves. Regarding the latter, Fig. 7a illustrates the typical pattern for seven out of nine tested layouts. It can be seen that the number of re-marshalling moves is mainly not or only slightly affected by changes in the deviation from the announced time. As FCQ is only taking into account the reported time by XT companies for the mis-overlay calculation, there is no particular influence on the generation of the re-marshalling schedule. If a container is picked up late, it remains highly prioritised by FCQ as it may be picked up any moment, thus, not creating calculated mis-overlays for containers beneath. Otherwise, if a container is retrieved early and is blocked, a shuffle move is the resulting move that occurs because the apparent mis-overlay is invisible to the mis-overlay calculation underlying FCQ. The significant difference between the deviation of 1h and the other values can be explained by the particularly low number of shuffle moves needed with the former which results from better accuracy of mis-overlay calculation. Contrary to re-marshalling moves that can be skipped if no suitable stack is available for a container to be re-marshalled, shuffling must be performed in order to free up containers to be retrieved. Thus with higher deviation from announced time, container stackings may be forced onto stacks that are not suitable and induce even more or new mis-overlays in *ordered* stacks. As a consequence, this potentially leads to more stacks being classified in *SS* inducing more re-marshalling moves with a higher number of shuffle moves. In combination, Fig. 7 indicates a deterioration in the re-marshalling quality n_{rm} with increasing deviation from the the announced time as N_{rm} stays nearly unaffected starting from 6 h while N_{shu} increases continuously for all settings of the deviation.

With reference to the XT waiting time in Fig. 8a, a significant influence of increasing deviation can be seen. The pattern corresponds to the shuffling discussed above being valid for almost all layouts and systems. Naturally, import containers are strongly affected by deviations from the announced time due to being stacked onto each other during the initial stacking by category stacking. As retrieval information of these containers becomes updated, new mis-overlays are calculated involving primarily import containers. As a result of the imprecise calculation due to high deviations from the announced time, import containers primarily suffer leading to increased waiting times at the landside handover area. Looking at the deviation of total vehicle waiting time, two patterns can be observed where Fig. 8b is showing the intermediate behaviour of them. In spite of indicating a significant increase from 1 to 48 h deviation, the change is not pronounced, especially, in comparison to shuffling and XT waiting time. Nonetheless, layouts with lower capacities show a more substantial separation between 1h deviation from the announced time and the other values regarding σ_{tot}. Adversely, layouts starting from 36/8/5 show no significant differences for all values of the deviation from the announced time for the SRMG-, TRMG- and DRMG-system, while the TriRMG shows no significant difference in the largest layout. In most cases, all significant

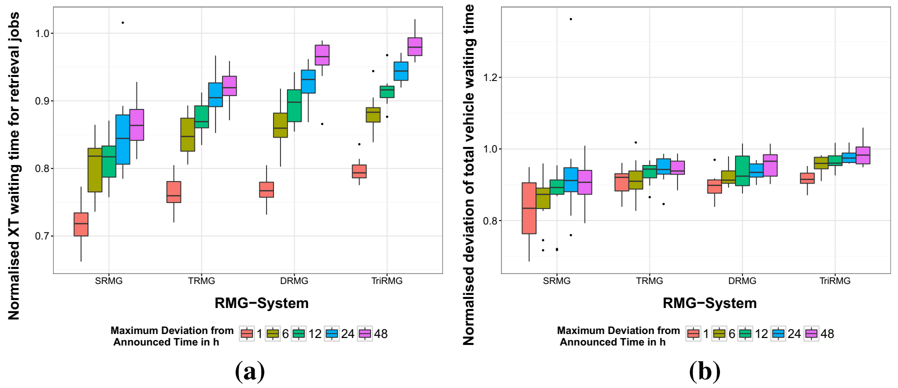

Fig. 8 FCQ behaviour concerning waiting times for deviations from the announced time in 36/8/4. **a** XT waiting time for retrieval jobs, **b** standard deviation of total vehicle waiting time

differences reported in this discussion regarding σ_{tot} are observed between 1h deviations and the other values. This indicates that FCQ seems to be insensitive to changes of the deviation starting at 6 h.

D2. Observations for compact layouts

The compact layouts with all three tiers show an analogous behaviour to Figs. 7 and 8 with respect to all evaluated performance indicators.

D3. Observations for large layouts

Large layouts appear to be insensitive to changes of the deviation from the announced time. In detail, the five- and six-tiered layouts with the SRMG-system show no (N_{shu}, $\overline{\omega}_{lsout}$, σ_{tot}) or only weakly (N_{rm}) significant results. Most notably, there is a high contrast between the effects of the deviation from the announced time for large and the smaller layouts with reference to shuffling and XT waiting time. While there is a clear and strong upwards pattern for nearly every increase of the deviation in compact and medium layouts, the SRMG stays unaffected in large layouts. This indicates that the low resource SRMG-system is not primarily impacted by the underlying uncertainty of available XT arrival information in large layouts where high crane resources are of particular need. The performance issues described in the previous analysis of the other TAS parameters appear to be more dominant regarding the re-marshalling quality of this system. As it is already unable to utilise high quality and quantity data to its full benefits, it is not possible to isolate a specific pattern resulting from uncertainty effects in the SRMG-system. This argument can be further verified by the behaviour of the TriRMG-system. It is the only system being slightly sensitive for σ_{tot} in the four- and five-tiered layouts showing also the highest sensitivity for the other performance indicators. Due to the availability of meaningful high crane resources with TriRMG, disadvantageous changes in all TAS parameters for all layouts are sensible because it is most able to obtain a high n_{rm} while utilising XT arrival information completely. The TRMG- and DRMG-systems show an ambivalent behaviour being in the range between SRMG and TriRMG, the TRMG leaning more towards behaviour of the SRMG and the DRMG tending to the patterns of the TriRMG.

Summary of TAS parameter sensitivity analysis
Following the preceding sensitivity analysis of the four relevant TAS parameter, an evaluation of the formulated hypotheses can be conducted. In total, FCQ seems to be most sensitive to the participation share regarding the direction and the magnitude of the performance influence, being particularly relevant when isolating the impacts by the share of booked slots. Thus, increasing this parameter has the ability to improve $\overline{\omega}_{lsout}$ the most. Although the share of correct arrivals does not exhibit the same magnitude of improvement, a similar behaviour of FCQ can be observed. However, it can be seen that the low resource SRMG-system is not able to fully benefit from this highly influential parameter in every layout while the TRMG already starts to indicate a potential disadvantage of the lack of its crossing capability in six-tiered layouts. It is mainly observed that FCQ is effective in translating additional amounts of booked slots into an expanded re-marshalling schedule to decrease N_{shu} while benefiting from more accurate information to reduce N_{rm}. Altogether, this results in a reduction of the number of re-marshalling moves needed to eliminate a shuffle move (re-marshalling quality n_{rm}) in all layout-system combinations. The slot length has a rather moderate effect on FCQ behaviour and relevant performance indicators. The SRMG- and TRMG-system are very insensitive to changes in this parameter. Only the crossing DRMG and very high resource TriRMG are capable of making use of highly precise arrival information. The announcement time shows the lowest influence of all parameters, in particular for the case of practically enforceable planning horizons for XT companies. Due to this reason, the pattern of resource capacity distribution between the RMG-systems is most apparent with this parameter. The SRMG struggles to utilise longer look-ahead times for all five- and six-tiered layouts, the TRMG for all six-tiered layouts, the DRMG for large five- and six-tiered layouts and the TriRMG only for large six-tiered layouts. This illustrates the general ability of the different systems to cope with extensive re-marshalling induced by a TAS. Finally, a high deviation from announced time has a highly significant disadvantageous effect on re-marshalling quality and $\overline{\omega}_{lsout}$ comparable in magnitude to the participation share. Deviations from announced times generate unfitting re-marshalling moves, thus, exhausting crane resources while not preventing shuffling. Nonetheless, only the TriRMG-system, where σ_{tot} is already very low, experiences unfavourable effects concerning this performance indicator. The other systems appear to continue having an inherently high σ_{tot} where impacts by the deviation from the announced time cannot be isolated. In summary, the results of the hypotheses testing conducted throughout the TAS parameter sensitivity analysis are listed in Table 14. In cases, where a hypothesis (formulated as alternative hypothesis in the test) is confirmed for all system-layout combinations it is stated so, otherwise only the system-layout combination for which the hypothesis is not confirmed is declared.

Table 14 Results of hypotheses testing for TAS parameter sensitivity analysis

Type	No.	Hypothesis test result
Participation share	2.1	Not confirmed for SRMG for share of correct arrivals: 4-tiered large layout
		Not confirmed for SRMG for participation share: 5-tiered medium/large, 6-tiered layouts
		Not confirmed for TRMG for participation share: 6-tiered large layout
	2.2	Not confirmed for multi-RMG in 6-tiered layouts
	2.3	Confirmed
	2.4	Confirmed
Slot length	2.5	Not confirmed for SRMG and TRMG
	2.6	Not confirmed for SRMG in large layouts
	2.7	Not confirmed for SRMG and TRMG
Announcement time	2.8	Not confirmed for SRMG in 5-tiered/6-tiered layouts
		Not confirmed for TRMG in 6-tiered layouts
		Not confirmed for DRMG in large 5-tiered/6-tiered layouts
		Not confirmed for TriRMG in large 6-tiered layouts
	2.9	Confirmed
Deviation from the announced time	2.10	Confirmed
	2.11	Not confirmed for SRMG, TRMG and DRMG

6 Conclusion

Firstly, this work proposes a solution method based on *Fuzzy Complex Queries* for the RMP which is embedded into a front-end block layout designed for single and multi-RMG systems. Rules are devised that enable automated yard cranes to work using expert decision systems. Relevant performance indicators are identified to assess the effectiveness of this algorithm in terms of the fundamental definition of the RMP (container handling), yard block productivity (vehicle waiting times at handover areas) and crane workload (standard deviation of vehicle waiting times). In this context, it is shown that FCQ has the ability to improve relevant metrics incorporating different degrees of Terminal Appointment System information for XT arrivals. Category stacking, being the stacking strategy mainly applied in theory and practice, is the main rationale behind the benchmark case of Heuristical Housekeeping Stacking (HHS) and the initial stacking in the case when no re-marshalling is performed (noRM). However, the results for FCQ without TAS show that further potentials to reduce mis-overlays can be unlocked by evaluating and rearranging export containers based on the estimated departure time of their corresponding deep sea vessel from the vessel call pattern. By adding complete TAS information a day in advance, there is a highly significant improvement of XT waiting time which previously suffered from the prioritisation of waterside operations. In spite of the landside productivity improvement, FCQ is able to maintain the enhanced performance of the waterside in nearly all cases. In terms of

Definition 1 of the RMP, FCQ/TAS is able to reduce shuffling significantly (n_{shu}) while also achieving to a reasonable degree the objective of minimising the number of re-marshalling moves (n_{rm}). Only the SRMG-system struggles to fully benefit from complete TAS information in high-tiered large layouts, where a stronger focus on distance-based re-marshalling could be more beneficial. Notably, strong reductions in the standard deviation of vehicle waiting times can be achieved as well. This indicates that crane workload can be balanced without directly interfering with XT companies' schedules. This may contribute to achieving better yard block productivity while guaranteeing higher XT flexibility which is not the case when slot capping is employed. From this point of view, it should facilitate a comprehensive adoption of TAS by XT companies in order for both sides' benefits and substantially mitigate the need for capping. All in all, following performance-wise conclusion can be drawn:

FCQ is highly suited to improve waterside operation, in general, while being able to incorporate TAS information to further improve overall productivity by targeting landside information.

Secondly, when applying a rule based strategy like FCQ, several conclusions can be drawn with respect to a suitable TAS design and incentive creation for XT adoption:

A. A high participation by XT companies is essential for a proper functioning of FCQ as it delivers the baseline on which mis-overlays are calculated and, thus, import containers can be prioritised for re-marshalling. In most cases with intermediate deviation from the announced time and a low number of no-shows, a high amount of XT bookings is preferred to very reliable information. This is due to the fact, that having arrival information at all enables the initiation of a re-marshalling job, even if it may not be the best fitting one. However, the share of correct arrivals does have a substantial influence as well, particularly in terms of saving crane resources when re-marshalling. Overall, a mandatory TAS seems to have the best impact on re-marshalling. However, observing these reductions in landside waiting time, XT companies should have an intrinsic motivation to book slots voluntarily as long as queuing and streamlined handover processes benefit their schedule as well. For employing FCQ with TAS following guideline can be stated:

Incentivise a complete booking of XT arrivals to reduce shuffling of import containers while ensuring a minimal number of re-marshalling moves through reliable information.

B. The slot length has a rather moderate impact on FCQ behaviour. In most cases, FCQ can cope with an announcement within slots of half a day (based on five days average container dwell time) which may be appropriate to perform clean re-marshalling while ensuring planning flexibility for XT companies. With increasing crane resources and crossing capability, the value of higher resolutions of XT booking becomes the more relevant, the TriRMG being sensitive to 1h slot lengths in many cases. For employing FCQ with TAS following guideline can be stated:

The needed XT arrival precison determined by the slot length depends on the RMG-system layout. Low resource RMG-systems appear to be rather indifferent to low to intermediate slot lengths.

C. FCQ is least sensitive to changes in announcement time. Only practically unrealisable long times show a consistently high benefit for relevant performance indicators. With respect to potentially implementable designs, there seems to be a consistent pattern throughout all system-layout combinations referring to five days average dwell time:

Announcement times of 12–24 h balance the informational need by the terminal for re-marshalling and the flexibility desired by XT companies.

D. The deviation from the announced time shows a continuous deteriorating effect on XT waiting time, re-marshalling quality and shuffling irrespective of layout and RMG-system. The magnitude corresponds to the influence of the participations share emphasising that the deviation may affect the precision of mis-overlay calculation more strongly than the slot length. High deviations lead to false prioritisation of containers inducing more crane workload trough re-marshalling moves while these moves do not benefit or may even disbenefit shuffling:

Discourage high deviations from announced time that can be regarded as no-shows or neglection of the booked slot.

In conclusion, the sensitivity analysis recommends the most careful and detailed TAS design for DRMG- and TriRMG-systems while systems less suited to complete re-marshalling schedules are more stable to parameter changes.

This work has opened up many ideas for further research. Long distance crane movements are not only harmful when re-marshalling is interrupted by incoming storage and retrieval jobs, but may also increase the risk of crane blocking time in multi-RMG-systems. This is particularly relevant when more than one crane is performing re-marshalling. FCQ does not encompass any particular evaluation to account for potential blocking time. A component within the algorithm to anticipate and potentially prevent blocking time may show improved results. Moreover, it seems of interest to assess the influence of specific yard block parameters on re-marshalling and the FCQ behaviour. With reference to the general FCQ behaviour, increasing the planned average occupancy rate could be analysed to understand what happens when the number of mis-overlays is increasing and fewer stacks become available for re-marshalling with more containers in the block. With reference to TAS, it might be of interest to evaluate the importance of having TAS with different transshipment shares. As it can be expected that there is no need for TAS with a transshipment share of 100% because there are no XTs in this system, the question arises at which point a TAS warrants to be implemented. Penultimately, although being embedded into a complex, dynamic and stochastic simulation model, an analytical approach to solve an isolated and formalised model to optimality could help to benchmark FCQ. Finally, FCQ is designed to only use arrival information for XTs picking up containers. However, a TAS can also gather data for XTs delivering containers to the block. A further analysis of problems and cases how these data may be utilised seems appropriate.

Acknowledgements I am grateful to Nils Kemme who made his yard block simulation model available to me and has been a valuable adviser for this work. Additionally, I would like to thank the handling Editor and the two anonymous reviewers for their valuable comments which improved this manuscript.

Appendix 1: Strategy comparison of the means of performance indicators in multi-RMG-systems

See Tables 15, 16 and 17.

Table 15 Strategy comparison of the means of performance indicators for 10 simulation runs in TRMG-systems

Layout	Strategy	$\overline{n}_{shu}$ ($\overline{N}_{shu}$)	$\overline{n}_{rm}$ ($\overline{N}_{rm}$)	$\overline{\overline{\omega}}_{tot}$	$\overline{\overline{\omega}}_{wsin}$	$\overline{\overline{\omega}}_{wsout}$	$\overline{\overline{\omega}}_{lsin}$	$\overline{\overline{\omega}}_{lsout}$	$\overline{\sigma}_{tot}$	$\overline{\sigma}_{lsout}$
36 × 6 × 4	noRM	1.13 (2375)	– (0)	2.26	1.25*	3.60	0.86*	3.27	2.23	2.02*
	HHS	1.10 (2315)	16.62 (988)	2.19	1.51	3.29	0.81*	3.02	2.02	1.77
	FCQ	0.89 (1872)	3.07 (1541)	2.09	1.38	2.69*	0.91*	3.48	1.93	2.14*
	FCQ/TAS	0.51 (1070)	1.99 (2596)	1.82	1.28*	2.72*	0.84*	2.31	1.69	1.63
36 × 6 × 5	noRM	1.45 (3787)	– (0)	3.21*	1.75*	5.19	1.38*	4.43	3.41	3.00*
	HHS	1.47 (3893)	– (1364)	3.26*	2.20	5.00	1.39*	4.20	3.25	2.74
	FCQ	1.23 (3247)	4.43 (2390)	2.98	2.02	3.85	1.46*	4.78	2.91	2.97*
	FCQ/TAS	0.76 (1996)	1.90 (3408)	2.52	1.72*	4.00	1.21	2.90	2.54	2.25
36 × 6 × 6	noRM	1.72 (5223)	– (0)	4.92*	2.62	8.15*	2.53*	6.15*	6.01*	5.04*
	HHS	1.78 (5588)	– (1535)	4.99*	3.50	7.68*	2.44*	5.99*	5.59*+	4.77*
	FCQ	1.58 (5006)	15.44 (3358)	4.63	3.12	6.27+	2.54*	6.72	4.95+	4.73*
	FCQ/TAS	1.04 (3171)	1.59 (3268)	3.67	2.33	6.27+	1.83	3.78	4.41	3.49
36 × 8 × 4	noRM	1.10 (3086)	– (0)	2.61*	1.70*	3.95	1.14*	3.48	2.56	2.19
	HHS	1.06 (2990)	13.07 (1245)	2.58*	2.09	3.62	1.10*+	3.32	2.38*	2.00
	FCQ	0.88 (2470)	3.01 (1854)	2.48	1.95	3.04*	1.20	3.77	2.32*	2.29
	FCQ/TAS	0.49 (1395)	2.04 (3454)	2.12	1.71*	3.02*	1.07+	2.51	1.98	1.81
36 × 8 × 5	noRM	1.45 (5101)	– (0)	4.29*	2.73	6.79	2.13*	5.17	4.85*	4.00
	HHS	1.46 (5147)	– (1638)	4.35*	3.75	6.17	2.12*	4.90	4.55*	3.63*
	FCQ	1.21 (4264)	3.51 (2937)	3.91	3.23	4.93*	2.07*	5.41	3.97	3.65*
	FCQ/TAS	0.75 (2630)	1.82 (4486)	3.23	2.55	4.97*	1.70	3.26	3.42	2.65
36 × 8 × 6	noRM	1.70 (6852)	– (0)	8.41*	5.45	13.21*	5.32*	8.87*	11.30*	9.44*
	HHS	1.75 (7314)	– (1817)	8.61*	7.62	12.25*	5.10*	8.39*	11.05*	8.53*+
	FCQ	1.57 (6560)	12.81 (3734)	7.38	6.14	9.98+	4.46	8.57*	8.72+	7.19+’

Table 15 continued

Layout	Strategy	$\overline{n}_{shu}$ ($\overline{N}_{shu}$)	$\overline{n}_{rm}$ ($\overline{N}_{rm}$)	$\overline{\overline{\omega}}_{tot}$	$\overline{\overline{\omega}}_{wsin}$	$\overline{\overline{\omega}}_{wsout}$	$\overline{\overline{\omega}}_{lsin}$	$\overline{\overline{\omega}}_{lsout}$	$\overline{\sigma}_{tot}$	$\overline{\sigma}_{lsout}$
	FCQ/TAS	1.09 (4375)	1.54 (3809)	6.26	4.50	10.45+	3.56	5.34	9.32+	6.61’
44 × 10 × 4	noRM	1.06 (4511)	− (0)	4.55*	4.34	6.34	2.17	4.69*	5.12	3.60
	HHS	0.99 (4203)	5.40 (1661)	4.54*	6.19	5.06	1.94	4.24	4.62	2.95*
	FCQ	0.85 (3602)	2.75 (2499)	4.12	4.85	4.54	2.06	4.71*	3.86	3.06*
	FCQ/TAS	0.48 (2061)	1.95 (4783)	3.61	3.97	4.79	1.82	3.24	3.59	2.40
44 × 10 × 5	noRM	1.44 (7627)	− (0)	15.67	14.52*	23.54	8.78	12.84	23.31	17.48
	HHS	1.40 (7358)	7.64 (2050)	12.58	17.29	14.44*	6.50	9.79*	15.18*	11.52*
	FCQ	1.26 (6639)	3.41 (3363)	11.59	14.97*	13.79*	5.93	9.66*	16.07*+	9.88+
	FCQ/TAS	0.85 (4502)	1.66 (5179)	10.62	11.35	15.99	5.26	7.18	17.96+	10.86*+
44 × 10 × 5	noRM	1.69 (9788)	− (0)	38.93	33.35	58.57	22.87	34.63	55.17*	45.85
	HHS	1.64 (9540)	7.46 (1845)	33.03*	35.96*	42.81	19.26*	29.22	45.26	38.64*
	FCQ	1.53 (9002)	4.62 (3629)	34.17*	37.06*	48.12*	18.09*	26.38*	52.24*+	32.18
	FCQ/TAS	1.23 (7263)	1.81 (4565)	33.22*	31.34	50.13*	18.23*	26.48*	50.51+	39.81*

*,+,’ Two or more results marked with the same symbol within a layout are not significantly different ($\alpha = 5\%$)

Table 16 Strategy comparison of the means of performance indicators for 10 simulation runs in DRMG-systems

Layout	Strategy	$\overline{n}_{shu}$ ($\overline{N}_{shu}$)	$\overline{n}_{rm}$ ($\overline{N}_{rm}$)	$\overline{\overline{\omega}}_{tot}$	$\overline{\overline{\omega}}_{wsin}$	$\overline{\overline{\omega}}_{wsout}$	$\overline{\overline{\omega}}_{lsin}$	$\overline{\overline{\omega}}_{lsout}$	$\overline{\sigma}_{tot}$	$\overline{\sigma}_{lsout}$
36 × 6 × 4	noRM	1.13 (2373*)	– (0)	2.06*	1.01*	3.21	0.86*	3.15*	1.95	1.88*
	HHS	1.12 (2341*)	31.3 (999)	2.03*	1.15	3.00	0.85*+	3.11*	1.83	1.78
	FCQ	0.90 (1877)	2.98 (1479)	1.91	1.07	2.49*	0.88	3.34	1.73	1.88*
	FCQ/TAS	0.53 (1119)	2.04 (2551)	1.65	0.98*	2.49*	0.83+	2.25	1.46	1.56
35 × 6 × 5	noRM	1.46 (3808*)	– (0)	2.82*	1.29	4.46*	1.33*	4.30*	3.04*	2.92*
	HHS	1.47 (3899*)	– (1383)	2.86*	1.55	4.32*	1.31*	4.27*	2.92*	2.78*
	FCQ	1.24 (3280)	4.62 (2441)	2.63	1.45	3.44	1.34*	4.56	2.61	2.79*
	FCQ/TAS	0.78 (2019)	1.91 (3416)	2.15	1.20	3.39	1.13	2.82	2.12	2.14
36 × 6 × 6	noRM	1.72 (5225)	– (0)	4.11	1.74	6.59*	2.42*	5.87*	5.29*	4.97*
	HHS	1.78 (5553)	– (1527)	4.29	2.21*	6.63*	2.45*	5.94*	5.28*	5.08*
	FCQ	1.58 (5018)	15.70 (3252)	3.83	2.09*	5.19+	2.23	6.14	4.33	4.45
	FCQ/TAS	1.07 (3239)	1.68 (3341)	3.02	1.50	5.04+	1.75	3.68	3.53	3.48
36 × 8 × 4	noRM	1.10 (3105)	– (0)	2.42	1.29	3.61	1.17*	3.60	2.39	2.37
	HHS	1.07 (3005)	12.14 (1218)	2.36	1.49*	3.30	1.13*	3.53	2.19	2.12*
	FCQ	0.89 (2497)	2.96 (1801)	2.23	1.42*	2.78+	1.14*	3.75	2.03	2.16*
	FCQ/TAS	0.51 (1433)	2.00 (3339)	1.91	1.24	2.77+	1.07	2.49	1.72	1.74
36 × 8 × 5	noRM	1.45 (5079*)	– (0)	3.80*	1.87	5.88*	2.35*	5.24*	4.55*	4.28*
	HHS	1.45 (5097*)	– (1638)	3.87*	2.36	5.66*	2.24*+	5.14*	4.40*	4.01*+
	FCQ	1.22 (4315)	3.83 (2929)	3.39	2.15	4.22+	2.08+	5.40	3.61	3.86+
	FCQ/TAS	0.77 (2685)	1.81 (4327)	2.76	1.67	4.17+	1.74	3.35	3.03	2.99
36 × 8 × 6	noRM	1.72 (7006)	– (0)	7.60*	3.31	12.01	5.82*	9.31*	11.35*	11.11*
	HHS	1.74 (7261)	– (1706)	7.64*	4.31	11.27	5.69*	9.24*	10.73*	10.71*
	FCQ	1.60 (6669)	10.62 (3579)	6.37	3.71	8.55*	4.84	8.75	8.50	8.75
	FCQ/TAS	1.13 (4498)	1.45 (3629)	5.10	2.53	8.23*	3.76	5.60	7.34	7.11

Table 16 continued

Layout	Strategy	$\overline{n}_{shu}$ ($\overline{N}_{shu}$)	$\overline{n}_{rm}$ ($\overline{N}_{rm}$)	$\overline{\overline{\omega}}_{tot}$	$\overline{\overline{\omega}}_{wsin}$	$\overline{\overline{\omega}}_{wsout}$	$\overline{\overline{\omega}}_{lsin}$	$\overline{\overline{\omega}}_{lsout}$	$\overline{\sigma}_{tot}$	$\overline{\sigma}_{lsout}$
44 × 10 × 4	noRM	1.05 (4463)	– (0)	3.92	2.65	5.34	2.54	5.05*	4.39	4.35
	HHS	1.00 (4237)	6.86 (1552)	3.69	3.22	4.53	2.22*	4.70	3.65	3.48*
	FCQ	0.85 (3603)	2.73 (2352)	3.43	2.91	3.83*	2.16*	4.99*	3.31	3.53*
	FCQ/TAS	0.50 (2136)	1.96 (4566)	2.91	2.36	3.81*	1.97	3.39	2.77	2.82
44 × 10 × 5	noRM	1.43 (7564)	– (0)	12.56	7.25*	18.47	10.07	14.03	18.68	19.65
	HHS	1.40 (7329)	7.86 (1850)	10.67	8.61	13.55	8.19	11.90	14.39*	15.46*
	FCQ	1.28 (6777)	4.08 (3209)	9.41	7.39*	11.35*	7.56	11.42	12.66+	13.75+
	FCQ/TAS	0.93 (4909)	1.73 (4590)	8.37	5.36	11.99*	6.83	8.96	12.98*+	13.90*+
44 × 10 × 6	noRM	1.66 (9898)	– (0)	32.40	18.48*	48.23	24.67	37.90	45.24	48.52*
	HHS	1.67 (10033)	– (1651)	30.83	19.97	42.82	23.69	36.38	42.37*	47.37*
	FCQ	1.59 (9585)	10.76 (3366)	27.73*	18.85*	36.29	22.10*	33.96	38.05	42.66+
	FCQ/TAS	1.32 (7899)	1.95 (3889)	27.84*	16.43	41.27	21.50*	31.60	41.14*	45.41+

Table 17 Strategy comparison of the means of performance indicators for 10 simulation runs in TriRMG-systems

Layout	Strategy	$\overline{n}_{shu}$ ($\overline{N}_{shu}$)	$\overline{n}_{rm}$ ($\overline{N}_{rm}$)	$\overline{\overline{\omega}}_{tot}$	$\overline{\overline{\omega}}_{wsin}$	$\overline{\overline{\omega}}_{wsout}$	$\overline{\overline{\omega}}_{lsin}$	$\overline{\overline{\omega}}_{lsout}$	$\overline{\sigma}_{tot}$	$\overline{\sigma}_{lsout}$
36 × 6 × 4	noRM	1.14 (2396)	– (0)	1.87*	0.84*	3.03	0.62*	2.96*	1.72	1.67*
	HHS	1.12 (2338)	13.47 (780)	1.86*	0.97	2.86	0.63*+	2.93*	1.65	1.63*
	FCQ	0.91 (1911)	3.15 (1528)	1.78	0.88	2.55*	0.66+	3.13	1.58	1.65*
	FCQ/TAS	0.54 (1130)	2.07 (2625)	1.55	0.83*	2.52*	0.62*	2.15	1.34	1.39
36 × 6 × 5	noRM	1.46 (3813)	– (0)	2.41	1.03*	3.95*	0.90	3.83	2.45*	2.36*
	HHS	1.49 (3947)	– (1113)	2.49	1.22+	3.89*	0.96*	3.93	2.45*	2.41*
	FCQ	1.25 (3290)	4.76 (2487)	2.34	1.18+	3.25+	0.95*	4.18	2.22	2.36*
	FCQ/TAS	0.77 (1986)	1.86 (3406)	1.92	0.98*	3.22+	0.83	2.57	1.81	1.79
36 × 6 × 6	noRM	1.72 (5247)	– (0)	3.16	1.24	5.24*	1.44*	4.91	3.62*	3.46*
	HHS	1.76 (5512)	– (1313)	3.30	1.60*	5.19*	1.47*	5.04	3.54*	3.38*
	FCQ	1.57 (4980)	13.57 (3623)	3.08	1.56*	4.25+	1.47*	5.37	3.16	3.37*
	FCQ/TAS	1.05 (3169)	1.74 (3606)	2.48	1.19	4.34+	1.08	3.18	2.61	2.48
36 × 8 × 4	noRM	1.11 (3108)	– (0)	2.08*	1.01*	3.26	0.80*	3.24*	1.92*	1.86*
	HHS	1.08 (3030)	12.79 (995)	2.10*	1.16	3.11	0.81*	3.28*	1.89*	1.87*
	FCQ	0.88 (2460)	2.89 (1873)	1.99	1.12	2.72*	0.83*	3.40	1.74	1.83*
	FCQ/TAS	0.50 (1420)	2.00 (3381)	1.73	1.01*	2.71*	0.78*	2.31	1.49	1.52
36 × 8 × 5	noRM	1.43 (5015)	– (0)	2.88	1.34*	4.60*	1.33*	4.26	3.01*	2.77*
	HHS	1.47 (5137)	– (1444)	3.00	1.69	4.49*	1.37*	4.43	2.98*	2.82*
	FCQ	1.22 (4284)	4.24 (3101)	2.75	1.57	3.66+	1.34*	4.64	2.62	2.75*
	FCQ/TAS	0.75 (2606)	1.92 (4636)	2.28	1.32*	3.67+	1.12	2.89	2.19	2.14
36 × 8 × 6	noRM	1.71 (6973)	– (0)	4.36	1.95	6.94*	2.56*	6.07	5.35*	4.97*
	HHS	1.77 (7403)	– (1711)	4.70	2.70	7.00*	2.73*	6.42*	5.45*	5.13*
	FCQ	1.56 (6543)	9.99 (4300)	4.05	2.41	5.36+	2.32	6.45*	4.24	4.44
	FCQ/TAS	1.06 (4226)	1.59 (4359)	3.29	1.71	5.56+	1.84	3.82	3.81	3.45

Table 17 continued

Layout	Strategy	$\overline{n}_{shu}$ ($\overline{N}_{shu}$)	$\overline{n}_{rm}$ ($\overline{N}_{rm}$)	$\overline{\overline{\omega}}_{tot}$	$\overline{\overline{\omega}}_{wsin}$	$\overline{\overline{\omega}}_{wsout}$	$\overline{\overline{\omega}}_{lsin}$	$\overline{\overline{\omega}}_{lsout}$	$\overline{\sigma}_{tot}$	$\overline{\sigma}_{lsout}$
44 × 10 × 4	noRM	1.05 (4443)	– (0)	2.78*	1.78*	4.06	1.32*+	3.86*	2.57	2.42*
	HHS	1.00 (4242)	7.29 (1464)	2.80*	2.17	3.71	1.36*	3.91*	2.44	2.35*
	FCQ	0.84 (3556)	2.78 (2463)	2.66	2.06	3.30	1.29+	4.04	2.22	2.23
	FCQ/TAS	0.48 (2040)	2.01 (4825)	2.33	1.75*	3.38	1.18	2.82	1.95	1.87
44 × 10 × 5	noRM	1.45 (7759)	– (0)	5.41*	3.27	7.95	3.53	6.70	6.63	6.37
	HHS	1.41 (7440)	6.08 (1941)	5.23*	4.61	6.60	3.16	6.30*	5.54	5.14
	FCQ	1.21 (6414)	2.93 (3946)	4.44	3.88	5.18	2.64	6.13*	4.26	4.22
	FCQ/TAS	0.78 (4139)	1.75 (6334)	3.69	2.75	5.38	2.27	4.08	3.83	3.68
44 × 10 × 6	noRM	1.72 (10737)	– (0)	14.68*	8.38	22.30	11.30*	16.24*	22.20	22.57*
	HHS	1.74 (10983)	– (2098)	14.46*	11.07	19.25	10.98*	15.99*	19.62	20.82*
	FCQ	1.61 (10215)	9.10 (4752)	11.66	9.25	14.66	8.98	13.70	15.69	16.27
	FCQ/TAS	1.23 (7615)	1.64 (5107)	10.92	6.50	16.71	8.47	11.32	17.26	17.88

*,+,' Two or more results marked with the same symbol within a layout are not significantly different ($\alpha = 5\%$)

Appendix 2: Comparison of FCQ/TAS for different participation shares regarding main performance indicators

See Figs. 9 and 10.

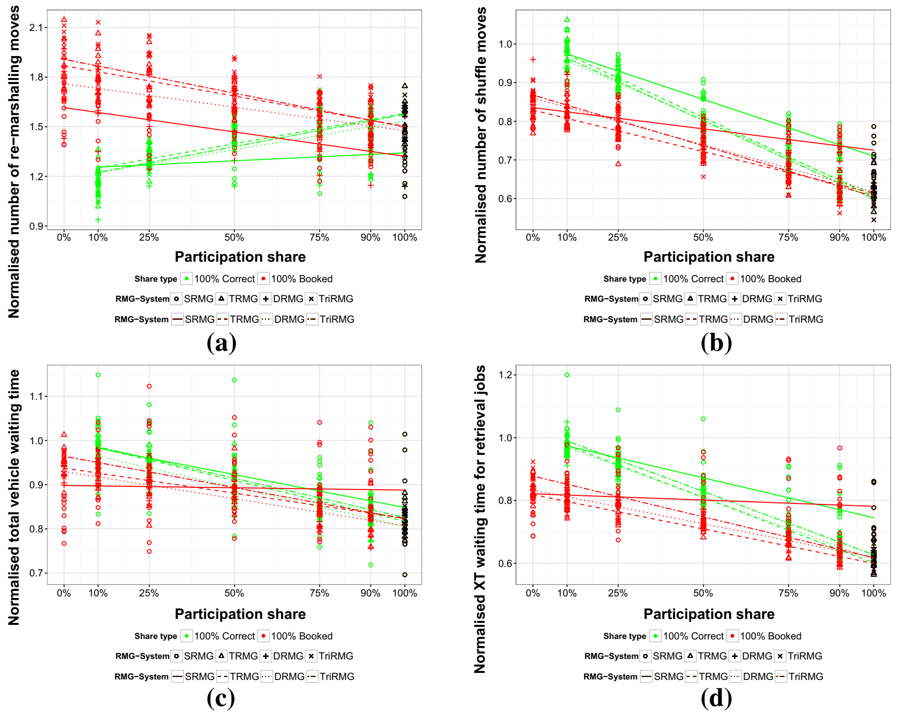

Fig. 9 FCQ behaviour for participation shares in 36/8/5. **a** Re-marshalling moves, **b** shuffle moves, **c** total vehicle waiting time, **d** XT waiting time for retrieval jobs

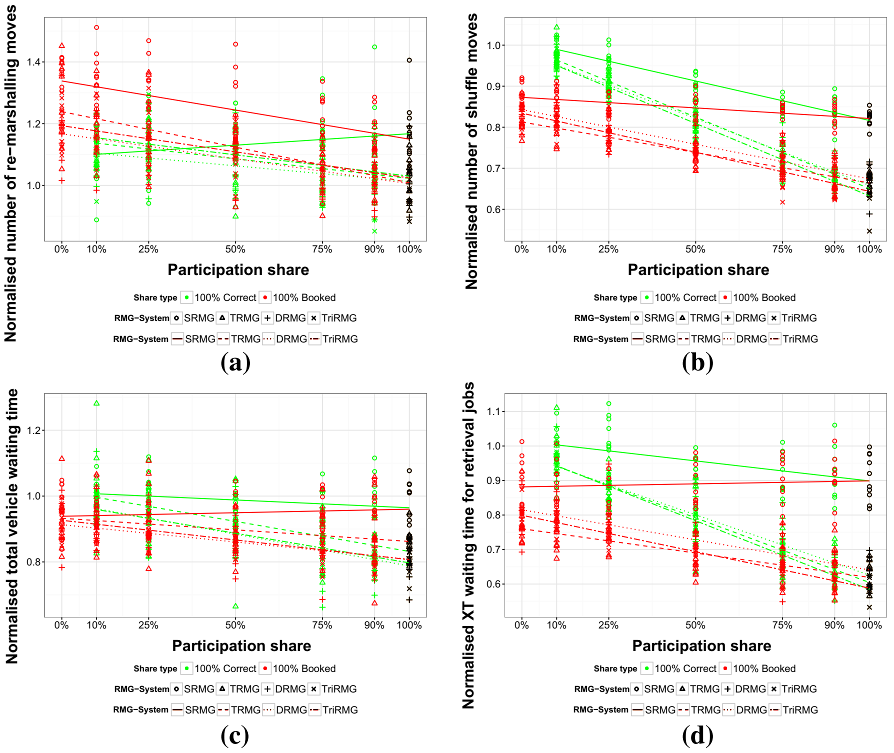

Fig. 10 FCQ behaviour for participation shares in 36/8/6. **a** Re-marshalling moves, **b** shuffle moves, **c** total vehicle waiting time, **d** XT waiting time for retrieval jobs

Appendix 3: Comparison of FCQ/TAS for different slot lengths regarding main performance indicators

See Fig. 11.

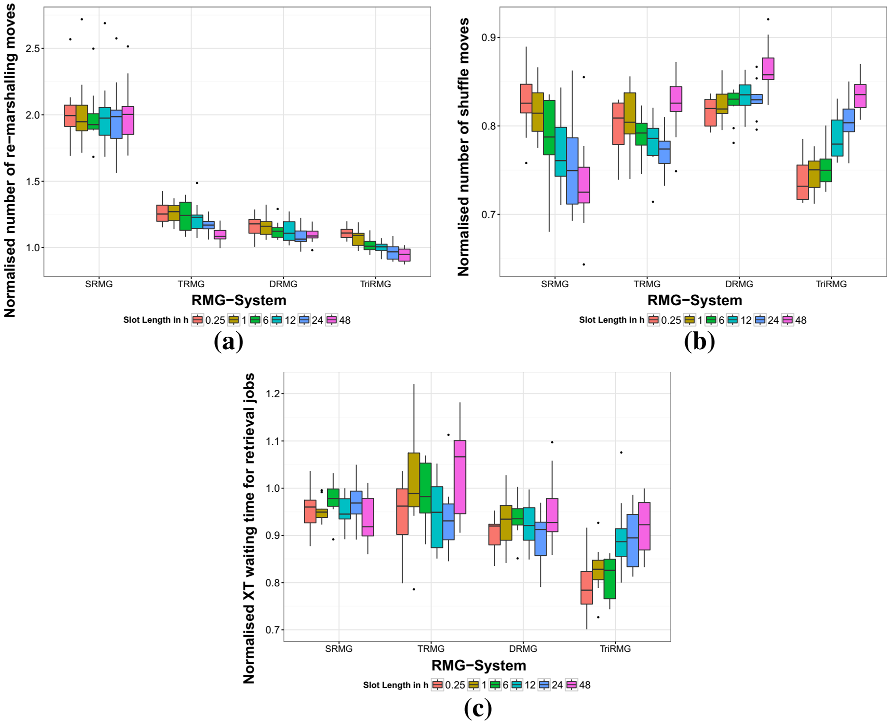

Fig. 11 FCQ behaviour concerning container handling and waiting time for slot lengths in 44/10/6. **a** Re-marshalling moves, **b** shuffle moves, **c** XT waiting time for retrieval jobs

Appendix 4: Comparison of FCQ/TAS for different announcement times regarding main performance indicators

See Figs. 12, 13 and 14.

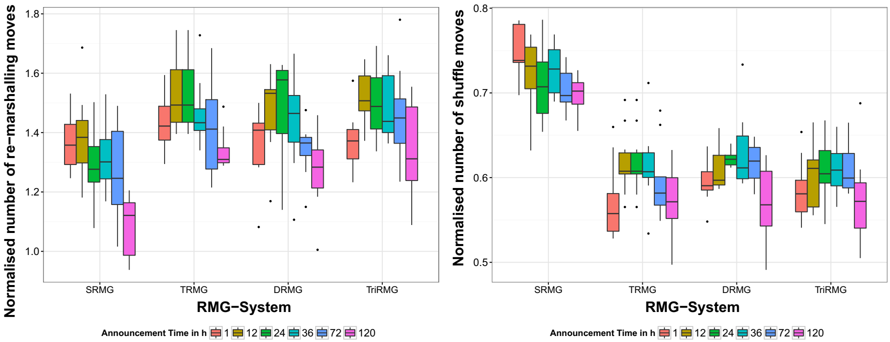

Fig. 12 FCQ behaviour concerning container handling for announcement times in 36/8/5. **a** Re-marshalling moves, **b** shuffle moves

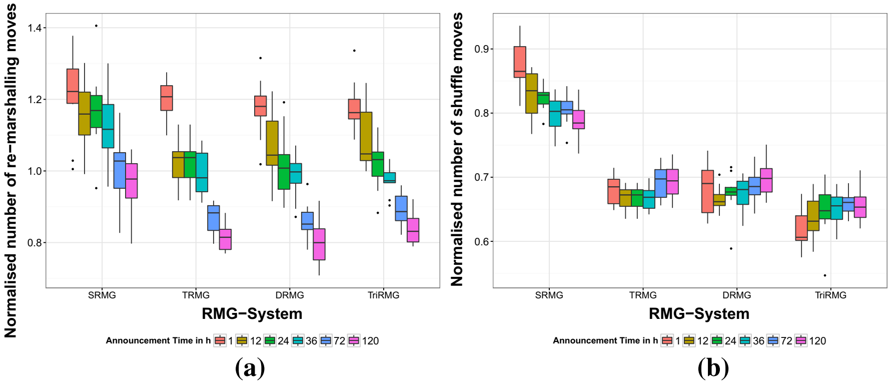

Fig. 13 FCQ behaviour concerning container handling for announcement times in 36/8/6. **a** Re-marshalling moves, **b** shuffle moves

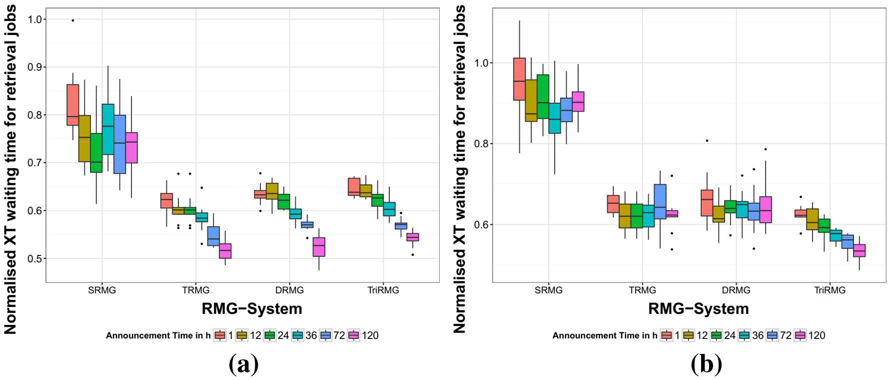

Fig. 14 FCQ behaviour concerning landside retrieval jobs for announcement times. **a** 36/8/5, **b** 36/8/6

Appendix 5: Detailed working example of Re-marshalling and TAS

After formalising the RMP and TAS separately, Fig. 15 gives an illustrative example of available container data and the generic mechanism that underlies the combined problem structure and serves as main working rationale for the subsequent development of the solution method. In more detail, it should provide an understanding of the fuzziness of assigning a container to a suitable slot in case of container retrieval data being available as intervals only. The decision rationale behind every optional container movement is explained for the movement of three containers. The three containers in this illustrative example are used as representatives for the movement types that can occur in the yard block. Hence, it shows the decisions that must be performed by an algorithm for solving the RMP with non-deterministic data. For reasons of clarity, only three bays for exchanging containers are considered here.[6] The three bays are situated close to the waterside (blue), block

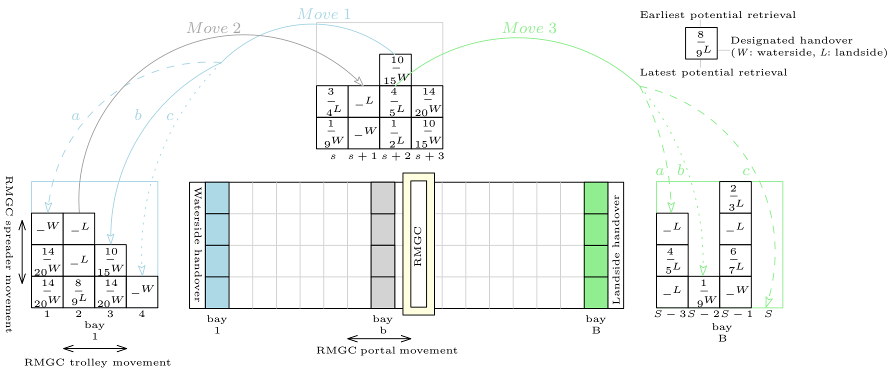

Fig. 15 Visualisation of the RMP with TAS information

[6] The actual RMP covers all bays in the block that must be dealt with simultaneously.

centre (grey) and landside (green), respectively to illustrate the necessity of including crane movements into the solution process for the RMP. In this instance, a bay *b* contains four stacks *s* that are numbered consecutively from 1 to *S* throughout all three considered bays with four tiers per stack available for storage. An approximated time interval of retrieval is stated for every container, if available. This is the case for containers departing by deep sea vessel (marked by '*W*') indicated by its berthing time and for XT (marked by '*L*') indicated by the booked time slot through TAS. If no time interval is specified, departure data of the container's retrieval vehicle are not available, i.e. for feeder vessels (marked by '*W*') and XTs that have not booked a time slot (until the observed time point). Note that the interval depends on the mode of transport and TAS parameter for the slot length.

At the time of no retrieval or storage jobs at the handover areas, the RMGC initiates its re-marshalling schedule. The aim is to find a blocked container in the yard that optimises crane movement and overall clearance of mis-overlays before a shuffle move of any container is induced. In the following, the RMP is decomposed into two main decision problems and a resulting procedure is described how a re-marshalling schedule may be constructed in Fig. 15:

Decision 1: *Identification of a source stack for the relocation of its top container*

Bay *b* might be used for sourcing a container being close to the current crane position and holding several stacks with potentially mis-overlayed containers in them. Analysing the approximated retrieval times, it is evident that no exact calculation of mis-overlays can be conducted as the intervals are overlapping. In this scenario, any container in bay *b* may be blocked by the container above depending on the exact retrieval time. An estimation of mis-overlays is particularly aggravated by containers departing by randomly arriving vehicles at the handover as in stacks $s + 1$. Hence, a crisp decision about a specific container is not always possible. A solution method must define criteria to evaluate the suitability based on these imprecise intervals. Conceivably, the bottom container in stack $s + 2$ seems to be a reasonable candidate for re-marshalling. Depending on the exact retrieval time of the stack's topmost container, the identified container is blocked by two containers and will leave the yard block as one of the first. Thus, the urgency to free up this container is high and the precision of intervals is more reliable in comparison to the other stacks.

Decision 2: *Stacking the sourced top container on a suitable stack*

In finding a suitable new stacking position, *Decision 2* faces the same issues with reference to overlapping time intervals as *Decision 1*. As a consequence, multiple options arise at this point. Firstly, the range of crane movement must be considered. On the one hand, there is a motivation to stay close to the source stack in order to avoid long moving distances. Notably, staying in the same bay *b* would avoid crane portal movement and the RMP would relate directly to the PMP. On the other hand, bay *b* is situated at the block centre implying that the crane has to move a long distance to the handover area of the respective container at the time of retrieval. As a result, the RMP does not only address the transformation of shuffle moves into re-marshalling moves but also shifts crane movement at the time of retrieval to movement during idle times. Following options arise to perform the needed re-marshalling move:

Move 1: Stacking a container leaving by deep sea vessel

Suppose the topmost container of stack $s + 2$ (subsequently referred to as c^{*1}) was chosen in *Decision 1*. c^{*1} is leaving the terminal by deep sea vessel as indicated by the given time interval and the index '*W*'. The blue arrow labelled 'Move 1' shows three possible stacking positions for c^{*1} (a, b, c):

a) Stack 1 could be used as there is only a small overlapping time interval of the two bottom containers with c^{*1} and the top container is indeterminate, thus, posing a potential not to induce any shuffle move in future (dashed line).
b) Stack 3 appears suitable as its top container has the same interval as c^{*1} indicating that they are departing by the same vessel and may be interchangeable during online stowage planning (see Sect. 4.1). The bottom container of stack 3 will most likely leave the yard block at a later time than c^{*1} (solid line).
c) Stack 4 has one container departing by feeder as there is no time interval available. Placing c^{*1} on this stack could either eliminate or induce a new shuffle move (dotted line).

Depending on the retrieval times of the two containers without time interval in stacks 1 and 4, an assessment of the stacking position can be made. The average container dwell time can be used as auxiliary estimator of the retrieval time for containers randomly leaving the block. However, the uncertainty in comparison to the fixed slots of berthing and TAS makes the mis-overlay calculation imprecise. Assuming similar arrival times of both containers without time interval in stacks 1 and 4, 'Move 1a' could be beneficial as the interval overlapping is small and this stack has potential to be a full stack without any mis-overlays. 'Move 1c' has also potential not to induce a mis-overlay. However, it seems to be less preferable since a full stack which is free of any mis-overlays leaves more slots open for future re-marshalling and stacking moves than a slightly occupied stack which is also free of any mis-overlays. If 'Move 1a' is chosen, there are more options to place containers on stack 4 that leave the block later than c^{*1}. In contrast, 'Move 1b' might be the best option in terms of balancing the uncertainty of inducing mis-overlays and leaving space for future handling moves. There is only slight overlapping of intervals so that most certainly this will not induce any mis-overlays.

Move 2: Stacking a container with no retrieval information and long distance to its handover area

While the bottom container of stack $s + 2$ is still blocked it does not need to be always advantageous to do an empty crane move from bay 1 to bay *b*. Depending on the urgency to free up this container, another container near the current crane position of bay 1 could be selected to be moved towards the direction of bay *b*. Therefore, *Decision 1* has to be performed again which leads to the conclusion that the only appropriate stack is the second one as it is the only one containing containers leaving by XT. If a mandatory TAS is applied, there is even certainty of a mis-overlay in this stack because the two containers without time intervals would have been booked if they left prior to the bottom container. As a consequence, the top container (subsequently referred to as c^{*2}) of stack 2 is chosen. With respect to the new stacking position of c^{*2}, there is only one possible stack in bay b that seems suitable eliminating the mis-overlay. The safest stack to place c^{*2} is $s+1$. Although there still might be a mis-overlay induced, 'Move 2' is separating containers with known time intervals from containers with unknown time intervals, thus, facilitating the calculation of mis-overlays within stacks with known time intervals. Additionally by applying this move, c^{*2} has passed half of the distance to its handover area outside of the time of retrieval. Distance-wise, it could be even more beneficial to move container c^{*2} to bay B being closest to the handover area. However, as the urgency to free up the bottom container in stack $s+2$ is high, it seems advisable to put c^{*2} near this stack. Thus, the crane is immediately close to its next re-marshalling job.

Move 3: Stacking a container leaving by XT with TAS information
The last type of container movement refers to the re-marshalling of containers leaving the block by XTs that booked a time slot through TAS. By dropping container c^{*2} on stack $s+1$ the current crane position is at bay b again. The bottom container in stack $s+2$ is still (urgently) blocked by one container on top of it. Thus, *Decision 1* should naturally lead to selecting the latter (subsequently referred to as c^{*3}) as suitable for re-marshalling. There are three potential slots for executing 'Move 3' towards the bay close to the relevant handover area (bay B) as stack $S-1$ is already fully occupied in this example:

a) Stack $S-3$ stores only containers leaving by XT. Two of them have not been booked yet implying that these containers will leave later than c^{*3} assuming a mandatory TAS. The middle container has the same interval as c^{*3}. Note that containers leaving by XT in the same interval are not interchangeable like in the case of deep sea vessels as they are picked up by different vehicles. As a consequence, this stacking move poses a risk for a shuffle move.
b) The time interval of c^{*3} is situated at the middle of the interval of the only container in stack $S-2$. The formation of a new mis-overlay cannot be stated definitely at this point.
c) A ground slot is available in stack S. Ideally, these should be occupied by containers with very late retrieval time to facilitate a flexible stacking of many containers on top of them. c^{*3} is a container leaving soon in comparison to the other containers in the block. However, stacking in this slot guarantees the elimination of future shuffle moves and opens the opportunity to yield the stack empty again soon.

In the case of 'Move 3', no clear identification of a target stack regarding the elimination of mis-overlays and leaving space for future handling moves can be made. Based on the generic analysis before, it can be stated that 'Move 3a' and 'Move 3c' seem to be preferable to 'Mover 3b' and that the actual decision depends on the design of the solution algorithm.

The representative instance discussed here illustrates the fuzziness in solving the RMP in terms of the three operationalised objectives for the overall goals introduced before:

– Reducing the potential for additional mis-overlays.
– Reducing empty crane movement.
– Leaving space for future shuffle-free container stacking.

References

Asperen E, Borgman B, Dekker R (2013) Evaluating impact of truck announcements on container stacking efficiency. Flex Serv Manuf J 25:543–556

Brinkmann B (2011) Operations systems of container terminals: a compendious overview. Handbook Of Terminal Planning. Springer, New York

Caserta M, Schwarze S, Voß S (2011) Container rehandling at maritime container terminals. Handbook of terminal planning. Science & Business Media, LLC, New York

Caserta M, Schwarze S, Voß S (2012) A mathematical formulation and complexity considerations for the blocks relocation problem. Eur J Oper Res 219:96–104

Chen G, Govindan K, Yang ZZ, Choi TM, Jiang L (2013) Terminal appointment system design by non-stationary $M(t)/Ek/c(t)$ queueing model and genetic algorithm. Int J Prod Econ 146:694–703

Choe R, Kim T, Kim T, Ryu K (2015) Crane scheduling for opportunistic remarshaling of containers in an automated stacking yard. Flex Serv Manuf J 27:331–349

Davis P (2009) Container terminal reservation systems. In: 3rd Annual METRANS National Urban Freight Conference

Dekker R, Voogd P, Asperen E (2007) Advanced methods for container stacking. Container terminals and cargo systems. Springer, Berlin

Duinkerken M, Evers J, Ottjes J (2001) A simulation model for integrating quay transport and stacking policies on automated container terminals. In: Proceedings of the 15th European Simulation Multiconference, p 909–916

Giuliano G, O'Brien T, Hayden H, Dell'aquila P (2007) The terminal gate appointment system at the ports of Los Angeles ans Long Beach: an assessment. Transp Res D 12(7):453–528

Guan C, Liu R (2009) Container terminal gate appointment system optimization. Marit Econ Logist 11:378–398

Huynh N (2005) Methodologies for reducing truck turn time at marine container terminals. Ph.D. thesis, The University of Texas at Austin

Kang J, Oh MS, Ahn E, Ryu K, Kim K (2006) Planning for intra-block remarshalling in a container terminal. Advances in Applied Artificial Intelligence. In: 19th International Conference on Industrial. Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE, p 1211–1220

Kemme N (2012a) Design and operation of automated container storage systems. Physica-Verlag HD, Berlin

Kemme N (2012b) Effects of storage block layout and automated yard crane systems on the performance of seaport container terminals. OR Spectr 34:563–591

Kendall GM (1938) A new measure of rank correlation. Biometrika 30:81–93

Lee Y, Hsu NY (2007) An optimization model for the container pre-marshalling problem. Comp Oper Res 34:3295–3313

Maguire A, Ivey S, Lipinski M, Golias M (2010) Relieving congestion at intermodal marine container terminals: review of tactical/operational strategies. In: 51st Annual Transportation Research Forum, Arlington, Viginia

Morais P, Lord E (2006) Terminal appointment system study. Transport Canada Publication TP 14570E, Transportation Development Centre of Transport Canada

Namboothiri R, Erea A (2008) Planning local container drayage operations given a port access appointment system. Transp Res Part E 44:185–202

Park K, Park T, Ryu K (2009) Planning for remarshaling in an automated container terminal using cooperative coevolutionary algorithms. SAC 2009. Honolulu, Hawaii, p 1098–1105

Pedrycz W (1994) Why triangular membership functions? Fuzzy Sets Syst 64:21–30

Phan M, Kim K (2015) Negotiating truck arrival times among trucking companies and a container terminal. Transp Res Part E 75:132–144

Ries J, Gonzalez-Ramirez R, Miranda P (2014) A fuzzy logic model for the container stacking problem at container terminals. In: Computational Logisitcs, 5th International Conference, ICCL 2014, vol 8760. Springer International Publishing, p 93–111

UNCTAD (2015) Review of Maritime Transport (2015) UN Symbol: UNCTAD/RMT/2015. United Nations Conference on Trade and Development. http://unctad.org/en/pages/PublicationWebflyer.aspx?publicationid=1374

Yu V, Cheng HY, Ting H (2009) Optimizing re-marshalling operation in export container terminals. In: Asia Pacific Industrial Engineering & Management Systems Conference, p 2934–2938

Zehendner E, Feillet D (2014) Benefits of a truck appointment system on the service quality of inland transport modes at a multimodal container terminal. Eur J Oper Res 235:461–469

Zhang X, Zeng Q, Chen W (2013) Optimization model for truck appointment in container terminals. Proced Soc Behav Sci 96:1938–1947

Zhao W, Goodchild A (2010) The impact of truck arrival information on container terminal rehandling. Transp Res Part E 46:327–343

Zimmermann HJ (2001) Fuzzy set theory and its applications, 4th edn. Kluwer Academic Publishing, Boston-Dordrecht-London

Filip Covic is currently in the final stage of his PhD program (2014 till now) at the Institute for Operations Research of the University of Hamburg. He has been working there as research assistant on a number of projects in collaboration with experts from different areas in Operations Research and Operations Management. He finished his study in Industrial Engineering at Master level (2011 to 2014) and at Bachelor level (2008 to 2011) jointly at the University of Hamburg, Helmut-Schmidt University and Hamburg University of Applied Sciences. He is interested in applying simulation and integer programming, especially decomposition techniques, in analysing container terminal planning problems, particularly container handling, yard block optimisation and managing vehicle arrival times.