Tank container operators' profit maximization through dynamic operations planning integrated with the quotation-booking process under multiple uncertainties

Xinjie XING[1], Paul R. DRAKE[1], Dongping SONG[1*] and Yang ZHOU[2]
[1]School of Management, University of Liverpool, Chatham Street, Liverpool, L69 7ZH, UK
[2]School of Economics and Management, Bejing Jiaotong University, Beijing, China

Abstract: Tank Container Operators (TCOs) are striving to maximize profit through the integration of their global Tank Container (TC) operations with the job quotation-booking process. However, TCOs face a set of unique challenges not faced by general shipping container operators, including the process uncertainties arising from TC cleaning and the use of Freight Forwarders (FFs). In this paper, a simulation-based two-stage optimization model is developed to address these challenges. The first stage focuses on tactical decisions of setting inventory levels and control policy for empty container repositioning. The second stage integrates the dynamic job acceptance/rejection decisions in the quotation-booking processes with container operations decisions in the planning and execution processes, such as job fulfilment, container leasing terms, choice of FFs considering cost and reliability, and empty tank container repositioning. The solution procedure is based on the simulation model combined with heuristic algorithms including an adjusted Genetic Algorithm, mathematical programming, and heuristic rules. Numerical examples based on a real case study are provided to illustrate the effectiveness of the model.

Keywords: OR in maritime industry; Profit maximization; Tank container management; Dynamic planning horizon; Uncertainty.

1. Introduction

The petrochemical industry has been growing steadily over recent decades, and up to 2014 the size of the global petrochemical market reached 490.5 million tons and is forecasted to grow at a Compound Annual Growth Rate (CAGR) of 5.1% from 2015 to 2022 (Grand View Research, 2016). As the biggest consumer, China accounted for 26.7% of global consumption in 2014 and is expected to witness growth of 6.2% from 2015 to 2022 (ibid). In terms of market value, 419.4 billion US dollars were traded in 2015, and the high demands are majorly coming from the automotive, textile, construction, industrial, medical pharmaceuticals, electronics and consumer goods industries. With the growth in the petrochemical industry, associated transport demands are also growing. As one of the key transport modes in this industry, Tank Containers (TCs) play an important role due to their convenient handling, safety, and environmental friendly features. Similar to Dry Containers (DCs), they are designed for intermodal transport, so they can be moved easily by truck, train and ship. According to the International Tank Container Organisation (ITCO, 2016), the global fleet size of TCs was estimated as 458,200 units in 2016, and it is maintaining a steady growth rate of 10% per year. Erera et al. (2005) concluded that the major advantages of TCs that have resulted in this growth are:

i. they are safer and produce less leakage during transportation and handling;
ii. they provide better space utilization compared to other modes, e.g. 43% more volume than drums stowed in DCs;
iii. no additional specialized port-side infrastructure is required when handling both DCs and TCs;
iv. they can be used to provide a reliable liquid storage device, particularly at the customer-end post-transport.

Although the physical features of TCs are similar to DCs, so that they are compatible with standardized cargo handling equipment and intermodal transport, their operations are quite different due to the special features of this industry. Dry Container Operators (DCOs) are normally shipping companies, who manage their own containers or long-term leased containers (can be regarded as self-ownership) with their own liner service. In contrast, Tank Container Operators (TCOs) offer a complete logistics service to customers in the petrochemical supply chain, but do not own ships, and their customer demands are satisfied by a so-called "*quotation-booking*" process (Erera et al. 2005). DCs are used in much higher volumes, creating large regular flows, through aggregation, that the large shipping companies can then match with regular routes and their own ships. It makes business sense for the shipping companies to own and manage DCs as it fits with their economies of scale business model. TCOs in contrast are used in much smaller volumes to provide far more specialist services, often with irregular flows. This is a low-volume high variety market less suited to the large shipping companies to own and manage TCs themselves. Instead, smaller specialist petrochemical logistics companies offer TC logistics and then piggyback on the container ships of the larger shipping companies. As a result, TCOs tend to emphasize profit (or revenue) maximization instead of cost minimization.

Customers book logistics services from TCOs with expected itinerary and execution time. TCOs need to respond quickly by developing a quotation through negotiation with external resource providers and analysis of their own resources. In this process, TCOs are challenged by how to deal with the uncertainties arising from the time gap between quotation development and service delivery. Furthermore, the high reliance on external resources magnifies these uncertainties. The following significant key features of the TC management problem have been left unaddressed.

i. The time gap between demand receipt and execution has not been modelled appropriately. As the customer request for a price quotation is often received well in advance of the demand execution time, so TCOs have to decide whether to issue a price quotation without accurate information on TC availability at the demand execution time. In addition, the demand receipt is revealed gradually over time. Erera et al. (2005) emphasized the "*quotation-booking*" process in TC management, but assumed all demands are known and deterministic in the planning horizon.

ii. There is a lack of decision support methods for developing quotations to meet individual customer demands. Support is required in determining precisely how to service individual demands, calculating expected costs and subsequently maximizing profits through the quotation process. This problem becomes even more complex with the option to lease containers, which can take the form of planned leasing or spot/emergent-leasing, in more real-time, with their different costs.

iii. Process uncertainties need to be included. For example, TCs are transported by third parties, so TCOs face significant uncertainties from Freight Forwarders (FFs) and shipping companies (as discussed later). Also, as practitioners from the TC industry pointed out, it is difficult to finish TC cleaning on time between different commodities, so deterministic, standardized cleaning times are not realistic.

iv. Empty Container Repositioning (ECR) is a critical task because the global flows of loaded containers are not balanced geographically. Regions with net outflows need ECR to replenish container stocks from net inflow regions. Empty Tank Container Repositioning (ETCR) is particularly expensive as TCOs have no ships and there are the third-party sources of uncertainty mentioned above. The planned and forecast execution of booked customer demands in the future may influence the volume of ETCR at the present time, but something unforeseen in the future may make the current ETCR ineffective.

Considering the above points, the research into TC management presented here, especially the quotation-booking process, is carried out on two levels. At the tactical level, this research will help TCOs adopt appropriate inventory control policies to maintain effective empty container repositioning and to cope with mid-term uncertainties. At the operational level, this research gives decision-making support to plans formulated by the quotation-booking process, particularly decisions about how to satisfy customer demands and how to manage the container fleet on a daily basis in the presence of uncertainties.

The rest of this paper is organized as follows. In Section 2, relevant literature is reviewed and the research gaps identified. In Section 3, the underlying problems are discussed in more detail and formulated mathematically. Following this, in Section 4 an optimization model is developed to integrate the two-level planning. In Section 5, a numerical test is conducted to demonstrate the application of the optimization model and the associated justifications of the model. The final section concludes with a summary of the main outcomes of this research and outlines future work.

2. Literature Review

The relevant literature is organized as follows. First, studies of TC operations are reviewed. However, due to limited research reported in this area and the unaddressed features of the problems in TC operations, the literature review has been expanded to general container operations. Note, at the tactical and operational planning levels, among the raised issues in Section 1, only ECR is explicitly and extensively discussed in the literature, while the other key issues raised are either not investigated or jointly discussed in the context of ECR studies (see SteadieSeifi et al., 2014; Song and Dong, 2015). Hence the second part of the literature review focuses on exploring ECR related studies in dry container operations.

In the TC operations field, only two relevant studies have been found. Erera et al (2005) have used the term 'quotation-booking' to describe the key process of TC operation and highlighted its significance. They employed a time-space deterministic model to optimize TC network flow with multiple commodities within an intermodal environment. They demonstrated the economic benefit of integrating TC booking and routing decisions with ECR decisions. Although their research has elaborated the special features associated with TC operations, the designed model is deterministic and assumed all demands in the planning horizon are known in advance and must be satisfied. The model is not applicable to support developing quotations to meet individual customer demands. More importantly, uncertainties such as container cleaning time and FF reliability are not considered, although these are important features in TC operations according to our communication with industry experts. Karimi et al. (2005) studied the TC operation using a different approach. They formulated the TC operation in a Just-in-Time (JIT) fashion and proposed a linear programming optimization model based on event-driven simulation. This helps in optimizing TC movement at the operational level, but again, it assumes a deterministic situation with the JIT setting that is barely the case in the real TC industry.

In the dry container operations field, many studies have been conducted in the last two decades (e.g. see the review papers: Braekers et al., 2011; Song and Dong, 2015). With the emphasis on empty container management models, the relevant literature may be classified into two groups: deterministic models and stochastic models. In the first group, for example, Choong et al. (2002) simulated ECR for an intermodal transportation network and examined the influence of the length of planning horizon. Meng and Wang (2011) developed a mixed-integer linear programming for shipping routes design and ECR optimization; Song and Dong (2013) considered route structure design, ship deployment and ECR jointly in a three-stage optimization process; Zheng et al. (2016) studied the container leasing price, with ECR consideration, from the point of view of the lessee. Although the above papers address several interesting issues such as length of planning horizon, laden and empty container routing and container leasing pricing, their models are not directly applicable to TC operations due to the higher levels and variety of uncertainties in the processes of quotation, booking and execution.

In the second group, Crainic et al. (1993) were among the first to extend deterministic formulations to stochastic programing models for ECR. Cheung and Chen (1998) applied a stochastic quasi-gradient method and a stochastic hybrid approximation procedure to deal with uncertainties in their two-stage optimization. They dealt with stochastic supply, demand and the residual capacity on vessels simultaneously in the optimization. Erera et al. (2009) used the adjusted robust optimization framework for dynamic empty repositioning when demands and future supply of empty containers are uncertain. Specifically, decisions and plans in this model are continuously adjusted when uncertainties are realized, i.e. they are dynamic.

Similarly, Di Francesco et al. (2009, 2013) developed a scenario-based formulation to solve stochastic problems on a rolling-basis. It only makes here-and-now decisions, and plans are updated as the planning horizon is rolling forward. Erera et al. (2009) and Di Francesco et al. (2009, 2013) demonstrated an ability to cope with uncertainties in a multi-stage decision making environment. However, as Epstein et al. (2012) noted, mathematical programming models are not suitable for large solution spaces mainly due to the computational complexity. Therefore, as TCOs normally operate on a global basis within large networks, it would be challenging to optimize them with stochastic mathematical models.

Alternatively, simulation-based optimization models have been developed widely to tackle container management problems for which computational complexity can be avoided at the cost of obtaining approximate solutions. For example, Lam et al. (2007) used a simulation-based approximate policy iteration algorithm to obtain an optimal average cost for ECR over an infinite planning horizon. Dong and Song (2009) used a simulation-based method to optimize a threshold control policy for ECR and container fleet size under stochastic demands. Song and Dong (2011) used simulation to evaluate the effectiveness of an empty container repositioning policy with flexible destination ports. Yun et al. (2011) built an (*s, S*) inventory policy for an inland transportation system in dealing with uncertain demand. Dang et al. (2013) took both ECR and leasing options into account with the optimization of a double threshold policy in an inland-depot system with dynamic order-arrival time. These papers present decision-making rules associated with system dynamic states, such as inventory levels of empty containers, and implement the decisions in the same time period as when they are made. The advantage of the inventory-based container management policies is ease of operating on a dynamic basis that can accommodate uncertainties. However, none of the studies in this stream have explicitly addressed the ECR decisions in relation to the quotation-booking process.

In the TC industry, customers normally request price quotations first and then make a booking with TCOs. Once the quotation is issued to the customer and the booking is confirmed, it is hard to change the demand fulfilment on the execution date. This is very different from the dry container shipping industry where neither shippers nor shipping lines would guarantee the cargo or the slot even after the booking confirmation. It should also be pointed out that there may be a significant gap between customer booking time and demand execution time, and uncertainties and the emergence of more information during the time gap can change what is expected (e.g. availability of empty containers at demand execution time). On the other hand, TCOs have the bargaining power to reject certain jobs during the quotation stage if they believe that the jobs are not sufficiently profitable within the business operations circumstances. Based on our communication with the industry, at present TCOs mainly rely on experience and manual calculation to decide whether to reject/accept jobs during the quotation stage. There is a need to develop decision support tools to assist in the quotation-booking process in relation to TC fleet management.

Moreover, although container leasing has been explicitly or implicitly studied by many authors (e.g. Moon et al., 2010; Dong and Song, 2012; Olivo et al., 2013), the difference between planned leasing and emergent leasing has not. Here planned leasing is defined as leasing that the TCO requests from lessors at least one day before the actual required time, whereas emergent leasing is requested on the same day as the actual use of the TC. In practice, planned leasing (pre-booked leasing) is cheaper than emergent leasing. This concept is analogous to the 'advanced purchase discount model', which is widely applied in the airline industry or other asset leasing activities (Gale and Holmes, 1992; Dana, 1998). From a supply chain coordination perspective, planned leasing contributes to information sharing under an uncertain environment (Tang and Girotra, 2017). TC lessors provide incentives to encourage their customers to do so. Taking the planned leasing and emergent leasing into consideration in the quotation-booking process enables TCOs to make strategic choices between these two options. In particular, when TCOs expect there will not be enough inventory to execute the demands received, they can arrange planned leasing to avoid higher emergent leasing costs. Furthermore, a cheaper leasing option could provide the opportunity for TCOs to plan to serve some demands with leasing containers to maintain more balanced container flows overall.

He (2013) modelled ECR with the participation of FFs and demonstrated that the total cost of ECR operations is affected by the choice of FFs. It differentiates FFs with different costs when the service depot of a FF is changed. However, the uncertain service levels and reliability provided by different FFs are not considered. In general, higher commission-charging FFs provide higher service levels and more reliable service. The use of FFs (or shipping lines directly) to book shipping slots may lead to uncertainty in realizing the empty container repositioning plan. In addition, after completing each job, empty TCs have to be cleaned at a dedicated depot in preparation for carrying different chemicals. This cleaning process may take several days and the precise time is a source of uncertainty. This uncertainty affects forecast TC availability that needs to be taken into consideration in planning ETCR. To the best of our knowledge, studies taking into consideration FF reliability and cleaning time uncertainty are not found in either the container operations or the maritime FF research literature.

This paper investigates how TCOs manage their container flows within the uncertain environment, and subsequently a decision-making support model is developed, based on an inventory-control based policy, to enable TCOs to achieve higher profits through the quotation-booking process.

3. Problem Description

TCOs have no ownership of maritime transportation services, instead serving customer demands through contracts with third-party transport providers. In the daily operation, customer demands are received including job start date, origin and destination. TCOs need to plan on these, developing corresponding quotations. TCOs exploit known information about costs and profits to decide how much they need to charge customers and how the demands should be served using three types of jobs; self-container jobs using TCOs' self-owned containers, planned leasing jobs and emergent leasing jobs.

Demands are not executed on their receiving date. Instead, they have a demand execution date set by the planning process, and the gap between the two dates varies from one demand to another. Once an execution date is set it is fixed, i.e. plans made each day have no influence over previously made plans as these have been returned already to the customer. The only change allowed to a planned job is if there is not enough inventory when the execution date arrives for a self-container job, which is then replaced by an emergent-leasing job. If at the planning stage it is forecast that there will not be enough self-containers on the execution date, then planned-leasing containers are scheduled for use, or the job can be rejected on profitability grounds. To simplify the narrative and formulation, it is assumed that quotation request, quotation return or rejection, and booking confirmation occur on the same day. For all the demands received on a given day, the latest of their execution dates forms the limit of the planning horizon on that day, so the planning horizon is dynamic and varying between days. Overall, this process is the 'booking-quotation process' illustrated in Figure 1.

Considering the quotation-booking process, it is challenging to make effective decisions for the following three reasons in particular:

i) Uncertain events occur along the supply chain, especially during the container return leg. However, once a quotation is returned to the customer it cannot be changed, so if there are not enough self-containers available for self-container jobs on a given execution day, TCOs will have to emergent-lease TCs. This increases the cost greatly.
ii) Leasing is in practice essential to provide flexibility without having excess capacity of self-containers and excessive just-in-case ETCR. However, pre-booked planned-leasing is much cheaper than emergent leasing. Therefore, TCOs need to consider not only how to avoid emergent leasing but also whether to use planned-leasing to achieve lower leasing and ETCR costs.
iii) In their niche market, TCOs have the bargaining power to reject some customer demands without losing future business. A job might be advantageously rejected if it would have knock-on effects or interactions with other jobs causing higher costs and lower profits. However, the time gap

between returning a quotation and actual execution makes it difficult for TCOs to evaluate whether or not to reject.

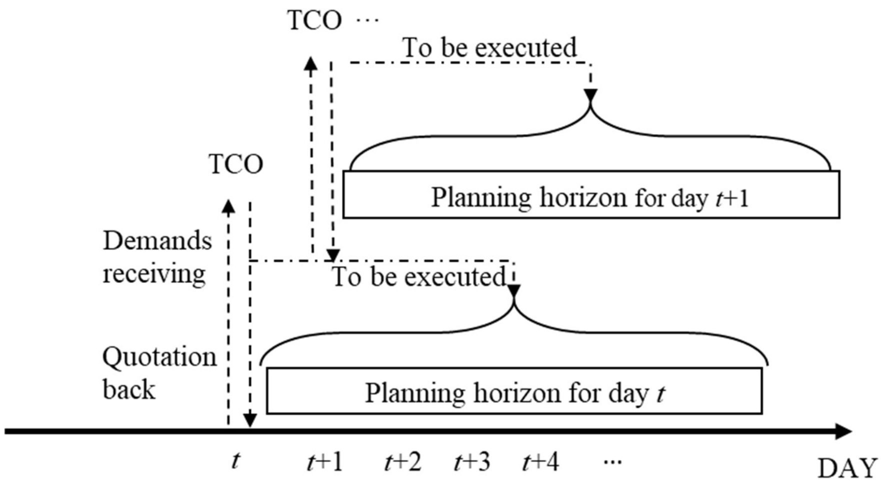


Figure 1. The Quotation-Booking Process

TCOs need effective strategies for the complex decision-making involved in dealing with these challenges. Unlike the dry container industry, TCOs need FFs to book external transport services for their TCs. Since maritime transport companies have limited capacity on specified routes, they will prioritize bookings for FFs with whom they have closer relationships. FFs with low priority will be less able to guarantee booking requests, so some of their TCOs' container transportation may not be completed as planned. The model developed here translates this into higher costs for FFs to maintain close relationships with transport companies, and these costs are passed on to TCOs, i.e. the higher the cost of an FF the higher its booking success rate. FFs providing 100% successful booking rates are defined as 'best FFs'. When TCOs choose FFs that are cheaper than best FFs, they will have the possibility of unsuccessful bookings. The successful booking rate is modelled as a discrete random variable that takes the value 50% or 100%, and the probability of a 100% successful rate is given by the cost of the chosen FF divided by that of the best FF. For example, if the cost of the best FF is £100, while the cost of the chosen FF is £60, the successful booking rate has a 60% chance of being 100% and 40% chance of being 50%. In the model, the choice of FF is made only for ETCR because only best FFs are used when meeting customer demands to avoid unsuccessful bookings for confirmed jobs. For ETCR, TCOs may choose appropriate inventory control policies with less than 100% successful booking rates to reduce costs. The safety stocks of TCs at each depot provide a buffer to guard against uncertainties caused by cheaper FFs (and other uncertainties explained later). According to the TC industry, inter-regional ETCR is far more expensive and seldom adopted, so the model categorizes depots into different regions geographically and only intra-regional ETCR is allowed.

In observed practise, TCOs can estimate accurately the container outflows from every depot to their destinations two weeks ahead. Hence, the data from the two-week customer-demand forecast is considered here in the decision-making. Figure 2 summarizes the key aspects of TC assets management.

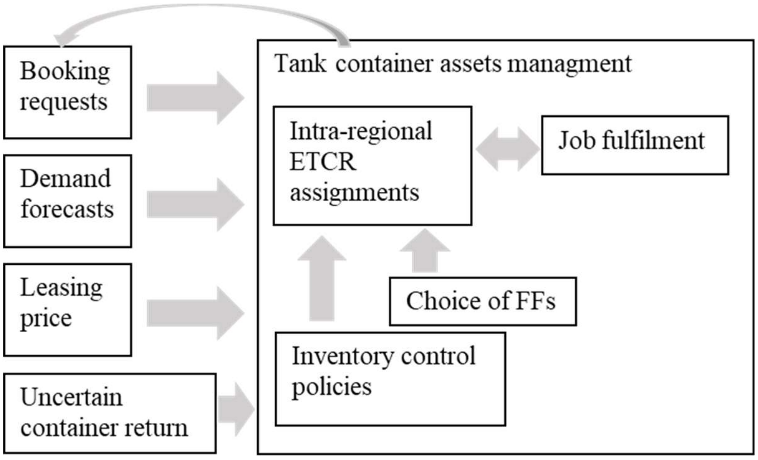


Figure 2. Tank container assets management overview

TCOs' operations can be different in many aspects. The following simplifying assumptions are made here to make the problem tractable:

1. Only the 20-foot equivalent unit (TEU) TC is used.
2. TC lessors have infinite fleets and leasing demands are met immediately.
3. Once a container-cleaning process has started, the cleaning time for that container becomes known.
4. FF cost is positively correlated with the shipping-slot booking success rate.
5. Selected FF will not vary from depot/region to depot/region, so only one FF will be used for global ETCR planning each day.
6. ETCR is only intra-regional, on routes available between any two depots.
7. Unloaded containers must be cleaned before reuse, with random duration in range 3 to 7 days.
8. Execution dates of customers' demands are later than their received dates.
9. Emergent leasing is more expensive than planned leasing, and leased containers are returned to lessees immediately after jobs.
10. The customer demand pattern remains similar annually. TCOs can forecast customer demands accurately two weeks ahead.
11. Self-owned TCs are always used first to meet customer demands during the demand-planning phase.

3.1 Notations

To formulate the system, the following notation sets are introduced:

Indices

i, j	Indices of TC depots.
s, t	Indices of date.
r	Index of regions.
d	Index of customer demands.
y	Index of predicted customer demands that is used in Stage 2 model.

Sets

T	Set of time periods for Stage 1 model; each element in T represents a day.
R	Set of regions.
P	Set of depots.
D_s	Set of customer demands received on day s. A customer demand is a tuple d, which contains the information of journey origin, destination, job received date, job start date, and number of containers. It is denoted as (O_d, D_d, S_d, T_d, M_d), where

$O_d, D_d \in P, O_d \neq D_d, S_d = s < T_d, M_d = 1$. Note that in TC operations, one demand is usually one unit. It is therefore assumed M_d = 1. However, the model can be modified easily to handle the case M_d >1.

D — Set of customer demands received on the days in T, which will be used in Stage 1 model.

Y_t — Set of customer demands for the next two weeks forecasted on day *t*. Each predicted demand y is denoted as $(O_y, D_y, S_y, T_y, M_y)$ and it represents a demand from depot O_y to D_y to be received on date S_y, where $O_y, D_y \in P, O_y \neq D_y; S_y \in [t+1,\ t+14]; M_y = 1$. Consistent with the above tuple, T_y represents the execution date of the forecasted demand, but its actual value cannot be forecasted and it is unknown at time *t*.

Input parameters

N — TC fleet size.

C_i^h — Inventory holding cost per TEU per day at depot i, where $i \in P$.

C_{ij}^p — Penalty cost for unmet demands per TEU from depot i to depot j, where $i, j \in P, i \neq j$.

C_i^o — Lifting-on cost per TEU at depot i, where $i \in P$.

C_i^f — Lifting-off cost per TEU at depot i, where $i \in P$.

C_b^t — Cost per TEU for choosing the best FF to move empty TCs on day *t*.

C_{ij} — Transportation cost per TEU (for both laden and empty) from depot i to depot j, where $i, j \in P, i \neq j$.

C_i^c — TC cleaning cost per TEU at depot i, where $i \in P$.

C_i^l — Planned (pre-booked) leasing cost per TEU per day at depot i, where $i \in P$.

C_i^{le} — Emergent leasing cost per TEU per day at depot i, where $i \in P$.

E_d — Revenue of demand *d*.

a_{ij} — Transportation time in days from depot i to j, where $i, j \in P, i \neq j$.

Inventory state and intermediate variables

$S_i(t)$ — Inventory level of depot *i* at the beginning of day *t*, where $i \in P$.

$S_i^m(\mathrm{t})$ — Adjusted inventory level of depot *i* on day *t* after confirmed container flow is completed, where $i \in P$.

Derived variables

b_i — Cleaning time in days at depot i, where $i \in P$. It is a random variable.

β^t — Shipping slot booking success rate on day *t*. A discrete random variable that takes two values: β^t = 100% with probability f_t/C_b^t; β^t = 50% with probability $(1 - f_t/C_b^t)$.

M_i^t — Length of each dynamic planning horizon, which equals the number of days from day *t* to the latest execution date in the demands received on day *t* at depot i, where $i \in P$.

Decision variables

W_d — Equals 1 if demand d is rejected, otherwise equals zero.

X_d^p — Equals 1 if demand d is planned to be delivered by self-container, otherwise equals zero.

X_d^a — Equals 1 if demand d is actually delivered by self-container on day T_d, otherwise equals zero.

Y_{ij}^t — Amount of ETCR containers from depot i to depot j on day *t*, where $i, j \in P, i \neq j$.

Z_d^p	Equals 1 if demand d is planned to be delivered by leased container, otherwise equals zero.
Z_d^e	Equals 1 if demand d is actually delivered by emergent-leasing container on day T_d, otherwise equals zero.
f_t	FF cost per TEU at day t subject to $f_t \in [2/5C_b^t, C_b^t]$. It determines the reliability of FF to complete the ETCR activity.
$[\mathrm{L}_i, \mathrm{U_i}]$	Upper and lower bounds of container inventory control policy at depot i, used for determining whether depot i is a surplus or deficit depot, where $i \in P$.

3.2 Outline of the Methodology

A two-stage simulation-based optimization approach is proposed to achieve two goals. The first goal is an optimized inventory control policy that leads to more effective ETCR at the tactical level by assuming all demands are accepted. The second is a decision-making support tool at the operational level for determining how new customer demands will be served every day to maximize profit by integrating with container operations planning. The different goals and their different preliminary settings require separate optimization processes. They are different in their planning levels. The inventory control policies are normally obtained through analysis of long-term statistics, which in turn, enables their adaptation to the associated environment. According to Braekers et al. (2011), inventory-control based optimization is tactical planning as it aims at ensuring the efficiency and rationale of existing resources over a medium horizon. Practically, once the inventory control policies are established, they will direct a series of operations, i.e. transportation, replenishment planning and production etc. Therefore, they are often maintained for a certain period of time to ensure the continuity of operations. In contrast, the second goal is at the operational level dealing with day-to-day operations. Customer demands are received on a daily basis, and associated decisions are made using current information. Therefore, only 'best-decisions-for-now' can be made when new demands are received, while demands received later can be planned using any subsequently available information, e.g. previously uncertain information may become certain.

Another critical reason why the two processes should be decoupled is that the two goals have different focuses. The inventory-control optimization seeks a long-term solution to TC management facing imbalanced trade flows and uncertain cleaning times by maximizing profit for the entire planning horizon. Whereas, the decision-making support tool is maximizing profit in serving customer orders (job quotation, planning and execution) on a daily basis within a dynamic planning horizon. The outputs of Stage 1 are used as inputs to Stage 2. On the other hand, the evaluation and optimization of the inventory-control policies rely on the simulation of simplified daily operations over the entire planning horizon.

The inventory-control based optimization is Stage 1 of the proposed simulation model. Specifically, a double-threshold inventory control policy is optimized through simulation of the entire planning horizon. To reflect industrial practice of daily operations, a special rolling-horizon approach is introduced that is different from the traditional rolling-horizon. As defined by Di Francesco et al. (2013), a 'rolling-horizon' refers to how a time-extended optimization model plans all the decisions for all periods of the planning horizon, but it will only implement the decisions for the first period and the model will be run again to plan and implement new decisions in the next period, when new information becomes available. Therefore, as the model runs forward, the total length of the planning horizon decreases by one period each period, so that the planning horizon at period t is (t, |T|). In contrast, the length of the rolling planning horizon in our model is determined dynamically. Planning happens every day that new demands are received, and the planning horizon is defined by the latest execution date (M_i^t) of the newly received demands. At every decision-making point, plans are made for the horizon (t, t+M_i^t). After this point TCOs can only adjust 'how' these scheduled demands will be served (self-owned containers or emergent-leased containers). They cannot alter execution times or reject jobs later on. This dynamic rolling-horizon is tailored to reflect the TC quotation-booking practice and, as it has not been seen in the literature, we believe it is novel.

Since historical data is used for the simulation at Stage 1, all the customer demands are accepted using either self-owned containers or planned-leasing containers. After all the known container flows are completed on this day, ETCR performed by following the inventory control policies. According to the initial inventory level of empty containers, every depot is classified as being either in surplus, in deficit or 'normal'. The deficit depots call for ETCR from surplus depots in their own region until either all the deficit depots are filled up to their lower bound threshold or all the surplus depots have repositioned out their TCs down to their upper bound. In Stage 1, it is assumed ETCR is 100% reliable as the 'best FF' is selected. By applying an Adapted Genetic Algorithm (AGA), a series of near optimal threshold-pairs can be generated through simulation using the historical data. With the completion of the Stage 1 simulation, the optimized inventory control policies are obtained.

In Stage 2, the optimized inventory control policies are implemented for ETCR. Whenever new demands arrive, all the demands received in the same period will be planned together. Similar to Stage 1, the demands are planned with the new dynamic rolling-horizon. However, Stage 2 seeks the most profitable way to serve the newly received customer demands, with demand rejection considered within the context of a two-week demand prediction. Experience has given industrial practitioners confidence in two-week time predictions, therefore they are used here when making decisions on customer demands. TCOs also need to decide which FFs are hired for ETCR on a given day, which incurs an additional process uncertainty in the reliability of ETCR. A more 'standard' GA is applied to select FFs on a daily basis within a dynamic planning horizon along the overall planning horizon.

3.3 Model at Stage 1: The Threshold Policy Optimization

Events in Stage 1

Stage 1 aims to find the optimal inventory control policies for all depots based on historical data. One year's daily operational data is used. It consists of the following four events and its mathematical model is formulated below.

1. *Inflows*. Inventory at each depot is updated with inflows of self-owned TCs from finished jobs and ETCR. Leased TCs are not counted because they are returned directly to the lessors. Since containers need to be cleaned after jobs only ETCR containers go directly into inventory. Cleaning times are modelled as a random variable that according to industrial experience varies from 3 to 7 days.
2. *Outflows*. Container outflow occurs for demands planned already for execution on that day. Although the 'to-be-executed' self-container and leased container jobs are planned, uncertainties may cause container unavailability. Once actual inventory cannot cover self-container jobs, emergent leasing is required.
3. *ETCR*. The remaining inventory in every depot is gauged with the specified inventory control policies, and ETCR determined accordingly. The real inventory levels in every depot must be modified by including the expected overall future container inflows and outflows within the planning horizon before comparison with associated threshold values. This avoids lead-time-caused repetitive ETCR and yields better inventory availability for upcoming demands.
4. *New Demands*. The dynamic rolling horizon for executing new demands is from the next day to the latest execution date of the new demands. Following a chronological sequence within the rolling horizon, the model simulates the expected container inflows and outflows on every demand execution date. Inventory on the demand execution dates is checked to see if there is enough to satisfy the 'to-be-executed' demands. If yes then demands are served by self-containers, otherwise planned leasing is required.

Mathematical model of Stage 1

Event 1: Inbound flow to receive self-owned containers on day t

At the beginning of day *t*, the inventory level for depot i is updated by adding the ETCR containers that have arrived and those that have returned from cleaning. Once the container cleaning process is started, the cleaning time becomes known. Let τ_d represent the cleaning time for job *d*, which is a realized sample of random variable b_i, then:

$$S_i(t)' = S_i(t) + \sum_{j\in P} Y_{ji}^{t-a_{ji}} + \sum_{d\in D} \sum_{D_d=i} \sum_{O_d=j, j\in P} \sum_{T_d=\mathrm{t}-\tau_d-a_{ji}} \mathrm{X}_d^{\mathrm{a}}; \quad (1)$$

Equation (1) indicates the expected inventory level for depot i after adding in TCs returning from ETCR or cleaning.

Event 2: Outbound flow to execute jobs on day *t*
The inventory level is updated with the planned container outflows for day *t*. Due to uncertain cleaning times, the actual inventory level may not satisfy all planned outflows. Therefore, emergent leasing may be required, so the most cost-effective way to assign the jobs among self-containers and emergent-leased containers must be determined. Let:

$$S_i(t)'' = \mathrm{Max}\ \{0, S_i(t)' - \sum_{d\in D} \sum_{O_d=i} \sum_{T_d=\mathrm{t}} \mathrm{X}_d^{\mathrm{p}}\}; \quad (2)$$

$$\text{If } S_i(t)'' > 0, \text{ then } \mathrm{X}_d^{\mathrm{a}} = \mathrm{X}_d^{\mathrm{p}}, \mathrm{Z}_d^{\mathrm{e}} = 0; \quad (3)$$

If $S_i(t)'' = 0$, then the assignment of jobs among self-containers and emergent-leasing containers is determined by solving the following mathematical programming problem:

$$\mathrm{Min} \sum_{d\in D} \sum_{O_d=i} \sum_{D_d=j, j\in P,} \sum_{T_d=\mathrm{t}} \mathrm{Z}_d^{\mathrm{e}} * \mathrm{C}_i^{\mathrm{le}} * a_{ij} \quad (4)$$

Subject to:

$$\sum_{d\in D} \sum_{O_d=i} \sum_{T_d=\mathrm{t}} \mathrm{X}_d^{\mathrm{a}} \le S_i(t)',$$

$$\mathrm{X}_d^{\mathrm{a}} + \mathrm{Z}_d^{\mathrm{e}} = \mathrm{X}_d^{\mathrm{p}}; \text{ for } d\in D \text{ with } T_d = t. \quad (5)$$

Equation (2) gives the potential inventory level for depot i after the job associated TC outbound flow. Equation (3) determines whether or not the current inventory is still able to cover the planned self-container jobs. Equation (4) determines how to assign self-container jobs and emergent-leasing jobs when the current inventory is unable to cover the planned self-container jobs.

Event 3: ETCR
ETCR is driven by inventory control policies every day, but intrinsic problems may emerge. Before in-transit ETCR containers arrive at a deficit depot, the 'to-be-replenished' depots will still be in deficit and will keep asking for ETCR from surplus depots. If no intervention is made, repetitive ETCR assignments will occur. Also, since part of the future container flow information is already known, it makes no sense to reposition TCs out of a depot that is surplus today but will soon be a non-surplus depot because of planned jobs. Likewise, there is less need of ETCR for a deficit depot if TCs will be available soon from finished jobs or previously arranged ETCR. Consequently, the need for inventory adjustment arises. First, the horizon length of adjusted inventory needs to be decided, i.e. how far into the future does information on planned operations need to be taken into account? Since the main target of the adjusted inventory process is to enable effective ETCR, while the target of ETCR is to ensure better container availability to meet the received demands, the latest execution day of the received demands will be used to define the adjusted inventory horizon length. Then, the imminent inventory adjustments described above need to be calculated.

Within the determined horizon, the future container arrivals and confirmed container outflow are the main adjustments. The future container arrivals come from finished jobs and previous ETCR. For any depot *i*, the future container arrival of previous ETCR planning is the sum of all ETCR from other depots to depot *i* that departed before the decision-making day and will arrive at depot *i* within the horizon. Another adjusted component is the containers returned from finished jobs. Since self-owned containers need to be cleaned before their next job, they face two scenarios. One, cleaning has already started and the container will return to the depot within the planning horizon. Two, cleaning has not started, but it is expected to be finished and the container returned to the depot within the planning horizon period. For the first scenario, the return day is certain. For the second, since cleaning has not started, the cleaning duration is a random number that needs to be estimated (Figure 3). To simplify the computation, the mean value of the cleaning duration is used. Finally, since no customer demands will be rejected at Stage 1, the overall container outflows are estimated by the demands that arrived on or before the decision-making day, while their execution dates are within the planning horizon.

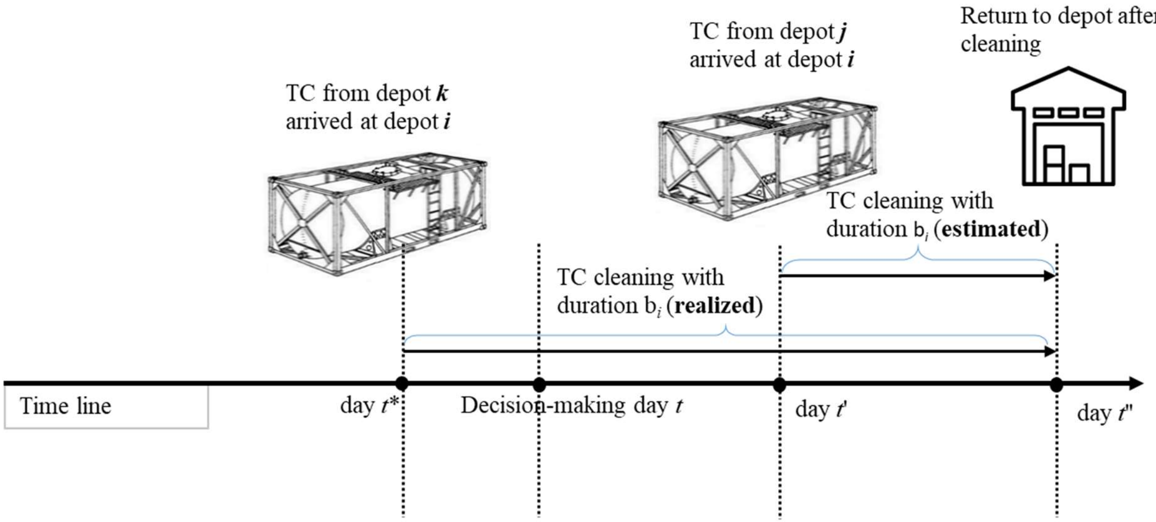


Figure 3. Two scenarios of container return after cleaning

Following the above discussion, let $\Omega^t_{ji,0}$ represent the time periods before time *t* and the deployed ETCR containers from depot *j* to depot *i* will be available in time period $t + 1$ *to* $t + M^t_i$, then $\Omega^t_{ji,0} = \{ s \in T | t + 1 - a_{ji} \le s \le min\,(\mathrm{t} + M^t_i - a_{ji}, t - 1)\}$. In addition, let $\Omega^t_{i,1}$ represent the containers that have finished jobs and started cleaning, and will be available in time period $t + 1$ to $t + M^t_i$, then $\Omega^t_{i,1} = \{d \in D | D_d = i, O_d = j, j \in P, T_d + a_{ji} \le t, t < T_d + a_{ji} + \tau_d \le t + M^t_i\}$; $\Omega^t_{i,2}$ represents the containers that are still fulfilling jobs but expected to be available in the time period t+1 to t+M^t_i, then $\Omega^t_{i,2} = \{d \in D | D_d = i, O_d = j, j \in P, S_d < t, T_d + a_{ji} > t, T_d + a_{ji} + \bar{\mathrm{b}}_i \le t + M^t_i\}$; $\Omega^t_{i,3}$ represents the containers that are planned to use self-containers but not shipped out yet, and are expected to be available in the time period $t + 1$ to $t + M^t_i$, then $\Omega^t_{i,3} = \{d \in D | D_d = i, O_d = j, j \in P, S_d < t, T_d > t, T_d + a_{ji} + \bar{\mathrm{b}}_i \le t + M^t_i\}$. The adjusted inventory level is:

$$S^m_i(t) = S_i(t)'' + \sum_{j \in P} \sum_{s \in \Omega^t_{ji,0}} Y^s_{ji} + \sum_{d \in \Omega^t_{i,1}} \mathrm{X}^{\mathrm{a}}_d + \sum_{d \in \Omega^t_{i,2}} \mathrm{X}^{\mathrm{a}}_d + \sum_{d \in \Omega^t_{i,3}} \mathrm{X}^{\mathrm{p}}_d - \sum_{d \in D} \sum_{O_d = i, S_d \le t} \sum_{T_d = \mathrm{t}+1}^{T_d = \mathrm{t} + M^t_i} M_d. \quad (6)$$

On the right-hand-side of Equation (6), the second to fifth terms are the container inflows specified above. The last term is the overall self-owned container outflows received on and before time *t*, and to-be-executed from $t + 1$ to $t + M^t_i$.

After the inventory levels are adjusted for all depots, ETCR assignments need to be determined. As inter-regional repositioning is not used, all ETCR is within the same region as follows. Let $P^{s}_{r,t}$ denote the set of surplus depots in the selected region r at time t, namely, $P^{s}_{r,t} := \{ i \in P_r \mid S_i^m(t)\text{-U}_i > 0\}$, where P_r is the set of depots in region r. Similarly, let $P^{d}_{r,t}$ denote the set of deficit depots in the same region r at time t, i.e. $P^{d}_{r,t} := \{ i \in P_r \mid \text{L}_i - S_i^m(t) > 0\}$. The ETCR assignments $\{Y^t_{ij}\}$ are determined by solving the following mathematical programming problem:

$$\text{Min} \sum_{i\in P^s_{r,t}} \sum_{j\in P^d_{r,t}} Y^t_{ij} * C_{ij}; \tag{7}$$
s.t.

$$\sum_{i\in P^s_{r,t}} \sum_{j\in P^d_{r,t}} Y^t_{ij} = Min\,[\sum_{i\in P^s_{r,t}} (S^m_i(\text{t}) - \text{U}_i)\ , \sum_{j\in P^d_{r,t}} (\text{L}_j - S^m_j(\text{t}))]. \tag{8}$$

After this event, the inventory levels at surplus depots are updated, which determines the inventory levels at the beginning of the next period:

$$S_i(t)''' = S_i(t+1) = S_i(t)'' - \sum_{j\in P^d_{r,t}} Y^t_{ij};\ \text{for } i \in P^s_{r,t}. \tag{9}$$

Event 4: Planning execution of new demands received on day t
This event plans the most profitable way to fulfil the new demands arriving on day t. Although execution of these jobs will be in the future, how this will be done must be decided on the receiving day. At the current stage, there are self-container jobs and planned-leasing jobs. When a job's execution date arrives, it will be executed as planned unless there are not enough self-containers, in which case an emergent-leasing job will arise. When time moves to the next day (t+1), the process is repeated and the new decisions are built on top of all the old plans without affecting them (Figure 4), i.e. plans once made are set firm and cannot be modified in the light of new demands or other data on subsequent days.

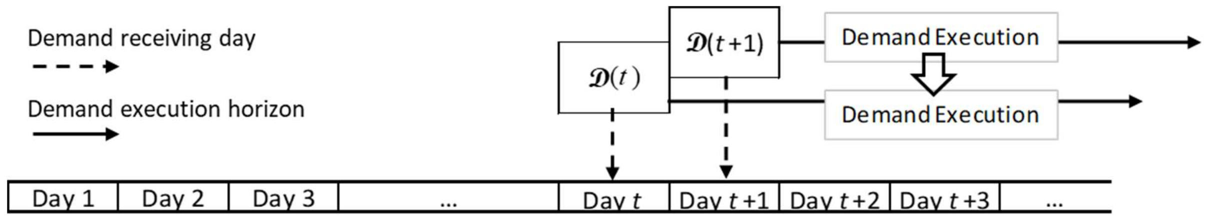


Figure 4. Overview of new demands receiving and planning

For a series of new demands, the rules for their planning are as follows. First, the latest execution date among these demands defines the current length of the planning horizon. Then, demands from the earliest execution date until the latest execution date will be planned. Second, within the planning horizon, if depot i has demands to be executed on day $t + q$, the inventory level of depot i is first updated with all the known information. This process is similar to the inventory adjustment in the previous event. $\Omega^t_{ji,0}$ is used to represent the time that ETCR activities have been arranged and those containers from depot j to depot i are expected to be available in time period $t + 1$ to $t + q$; $\Omega^t_{i,1}$, $\Omega^t_{i,2}$, $\Omega^t_{i,3}$ represent the sets of jobs with respect to different status. $\Omega^t_{i,1}$ comprises containers that have finished jobs and started cleaning, and will be available in time period $t + 1$ to $t + q$; $\Omega^t_{i,2}$ comprises containers that are still fulfilling their jobs and are expected to be available in the time periods from $t + 1$ to $t + q$; $\Omega^t_{i,3}$ represents the demands that have been planned to use self-containers but have not been shipped out yet, but are expected to be available in the time period from $t + 1$ to $t + q$. Their mathematical definitions are:

$$\Omega_{ji,0}^{t} = \{ s \in T | t + 1 - a_{ji} \leq s \leq \min(t + q - a_{ji}, t - 1)\}; \quad (10)$$
$$\Omega_{i,1}^{t} = \{d \in D | D_d = i, O_d = j, j \in P, T_d + a_{ji} \leq t, t < T_d + a_{ji} + \tau_d \leq t + q\}; \quad (11)$$
$$\Omega_{i,2}^{t} = \{d \in D | D_d = i, O_d = j, j \in P, S_d < t, T_d + a_{ji} > t, T_d + a_{ji} + \bar{b}_i \leq t + q\}; \quad (12)$$
$$\Omega_{i,3}^{t} = \{d \in D | D_d = i, O_d = j, j \in P, S_d < t, T_d > t, T_d + a_{ji} + \bar{b}_i \leq t + q\}. \quad (13)$$

If the updated self-containers are enough to cover all the 'to-be-executed' demands at that depot, those demands are planned as self-container jobs. If not, planned-leasing containers are needed. Mathematically, the assignments are described as follows:

$$S_i(t+q)' = S_i(t)''' + \sum_{j \in P} \sum_{s \in \Omega_{ji,0}^{t}} Y_{ji}^{s} + \sum_{d \in \Omega_{i,1}^{t}} X_d^{a} + \sum_{d \in \Omega_{i,2}^{t}} X_d^{a} + \sum_{d \in \Omega_{i,3}^{t}} X_d^{p} - \sum_{d \in D} \sum_{O_d = i} \sum_{S_d \leq t} \sum_{T_d = t+1}^{T_d = t+q} X_d^{p}. \quad (14)$$

On the right-hand side of Equation (14), the second term represents the accumulated ETCR jobs that have been scheduled and will arrive between time $t + 1$ to $t + q$. From the third to the fifth term are the accumulative container inflows related to self-container jobs between time $t + 1$ to $t + q$. The last term is all the scheduled container outflows between time $t + 1$ to $t + q$.

If $S_i(t+q)' \geq \sum_{d \in D} \sum_{O_d = i} \sum_{S_d = t} \sum_{T_d = t+q} M_d$, then $X_d^{p} = 1$ for any $d \in \{d \in D | O_d = i, S_d = t, T_d = t+q\}$; (15)

If $S_i(t+q)' < \sum_{d \in D} \sum_{O_d = i} \sum_{S_d = t} \sum_{T_d = t+q} M_d$, then the self-container jobs and planned-leasing jobs are determined by solving the following mathematical programming problem:

$$\text{Min} \sum_{d \in D} \sum_{O_d = i} \sum_{D_d = j, j \in P} \sum_{S_d = t} \sum_{T_d = t+q} Z_d^{p} * C_i^{l} * a_{ij}; \quad (16)$$

s.t.

$$\sum_{d \in D} \sum_{O_d = i} \sum_{D_d = j, j \in P} \sum_{S_d = t} \sum_{T_d = t+q} X_d^{p} \leq S_i(t+q)'; \quad (17)$$
$$X_d^{p} + Z_d^{p} = M_d \text{ and } \{d \in D | O_d = i, S_d = t, T_d = t+q\}; \quad (18)$$

Equations (15) and (16) define the two scenarios of demand assignments by comparing the inventory level and customer demands. Specifically, if there are not enough self-containers, planned-leasing containers are used. Equation (16) assigns the different types of jobs. Equation (17) and (18) define the constraints for the optimization equation.

Inventory control policy optimization

The objective of this model at Stage 1 is to find the optimal inventory control policy that leads to the most profitable TC operations, with profit defined as total revenue minus total cost. Here, the cost components include container-holding cost, laden and empty container moving cost, leasing cost, container-handling cost and container-cleaning cost. The optimal threshold values $\{[L_i, U_i] | i \in P\}$ are found by maximizing the following expected profit:

$$\text{Max} \quad EXP \quad \{\sum_{d \in D} M_d * E_d - \sum_{t \in T} \sum_{i \in P} S_i(t) * C_i^{h} - \sum_{d \in D} \sum_{O_d = i, i \in P} \sum_{D_d = j, j \in P} M_d * (C_{ij} + C_i^{o} + C_j^{f}) - \sum_{t \in T} \sum_i \sum_j Y_{ij}^{t} * (C_{ij} + C_i^{o} + C_j^{f}) - \sum_{d \in D} \sum_{D_d = j, j \in P} X_d^{a} * C_j^{c} - \sum_{d \in D} \sum_{O_d = i, i \in P} \sum_{D_d = j, j \in P} Z_d^{p} * C_i^{l} * a_{ij} - \sum_{d \in D} \sum_{O_d = i, i \in P} \sum_{D_d = j, j \in P} Z_d^{a} * C_i^{le} * a_{ij})\}. \quad (19)$$

3.4 Model at Stage 2: customer demands fulfilment

Stage 2 assists decision-making in terms of how the new customer demands will be served every day to make better profits, whilst facing the additional uncertainties caused by FFs' abilities to fulfil ETCR. The focus is on operational decisions, and the ETCR inventory-control policies from Stage 1 are inputs.

Events in Stage 2

There are four events with Events 1 and 2 being similar to those in Stage 1, whereas Events 3 and 4 are more complicated due to choosing FFs, job rejections and future demand forecasting.

Event 3 plans ETCR. Since this happens before demand planning (Event 4), all received customer demands and future demand prediction are considered in adjusting the inventory levels. Event 3 plans the ETCR deployment but not the amount, which will be influenced by the choice of FF in Event 4 (see Equation (28)).

Event 4 makes decisions on satisfying demands in terms of choice of FFs, self-container jobs, planned-leasing jobs and demand rejections. FFs are chosen by an iterative procedure, with the other decisions being made following this selection, within each iteration. Figure 5 illustrates this iterative procedure and its mathematical formulation is given below.

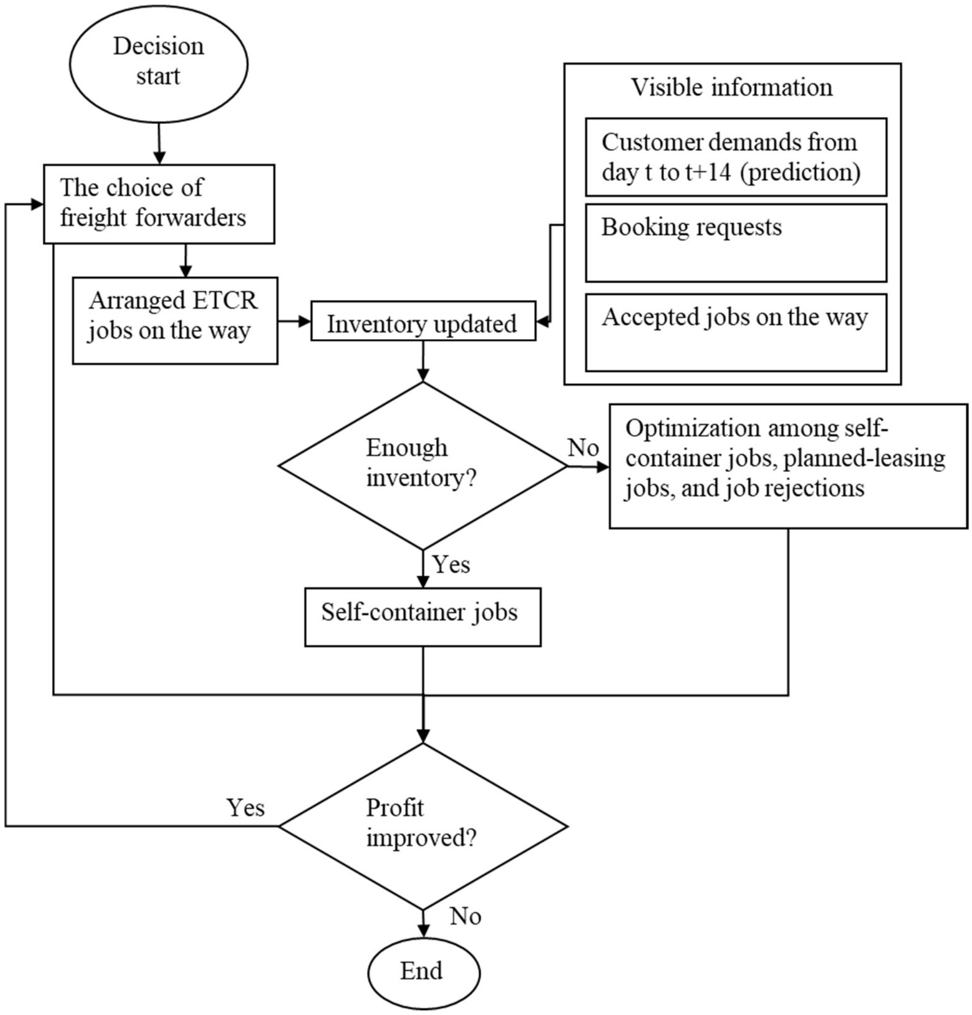


Figure 5. Decisions on new demands and choice of FFs

Mathematical model of Stage 2

Event 1: Inbound flow to receive self-owned containers on day *t*.
This event is the same as Event 1 in Stage 1 except the amount of ETCR is influenced by the choice of FFs. Also, since the FFs for the inflow ETCR are decided already, the associated booking success rate is known. Likewise, the cleaning duration for newly available containers is known. Let τ_d denote the known cleaning duration sampled from random variable b_i, and β_s $(s < t)$ be the known booking success rate on day *s*. The inventory level on day *t* is updated after Event 1 using Equation (20):

$$S_i(t)' = S_i(t) + \sum_{j\in P} Y_{ji}^{t-a_{ji}} * \beta_{t-a_{ji}} + \sum_{s=0}^{s=t-1} \sum_{d\in D_s} \sum_{D_d=i} \sum_{O_d=j,j\in P} \sum_{T_d=\mathrm{t}-\tau_d-a_{ji}} \mathrm{X}_d^{\mathrm{a}}, \tag{20}$$

Event 2: outbound flow to execute jobs on day *t*.

$$S_i(t)'' = \max\ \{0, S_i(t)' - \sum_{s=0}^{s=t} \sum_{d\in D_s} \sum_{O_d=i} \sum_{D_d=j,j\in P} \sum_{T_d=\mathrm{t}} \mathrm{X}_d^{\mathrm{p}}\}; \tag{21}$$

If $S_i(t)'' > 0$, then $\mathrm{X}_d^{\mathrm{a}} = \mathrm{X}_d^{\mathrm{p}}$, $\mathrm{Z}_d^{\mathrm{e}} = 0$, where $d \in D_s$; (22)

If $S_i(t)'' = 0$, then the assignments of the actual self-container jobs and emergent-leasing jobs are determined by solving the following sub-optimization problem:

$$\text{Min } \sum_{s=0}^{s=t-1} \sum_{d\in D_s} \sum_{O_d=i} \sum_{D_d=j,j\in P,} \sum_{T_d=\mathrm{t}} \mathrm{Z}_d^{\mathrm{e}} * \mathrm{C}_i^{\mathrm{le}} * a_{ij}; \tag{23}$$
s.t.

$$\sum_{s=0}^{s=t-1} \sum_{d\in D_s} \sum_{O_d=i} \sum_{T_d=\mathrm{t}} \mathrm{X}_d^{\mathrm{a}} \le S_i(t)',$$
$$\mathrm{X}_d^{\mathrm{a}} + \mathrm{Z}_d^{\mathrm{e}} = \mathrm{X}_d^{\mathrm{p}};\ \text{for } d\in D_s,\ s<t, \text{ and } T_d = t \tag{24}$$

Equations (21)-(24) jointly determine the laden TC outflows at depot *i* on day *t*.

Event 3: ETCR deployments.
ETCR is guided by the optimized inventory control policies obtained from Stage 1, while the actual process is the same as Event 3 in Stage 1. $\Omega_{ji,0}^t$ is used to represent the time periods before time *t* and the deployed ETCR containers from depot *j* to depot *i* will be available in time period $t + 1$ *to* $t + M_i^t$, then $\Omega_{ji,0}^t = \{ s \in T | t + 1 - a_{ji} \le s \le \min(t + M_i^t - a_{ji}, t - 1)\}$; $\Omega_{i,1}^t$ represents the containers that have finished jobs and started cleaning, and will be available in time period $t + 1$ to $t + M_i^t$, then $\Omega_{i,1}^t = \{d \in D_s | D_d = i, O_d = j, j \in P, T_d + a_{\mathrm{ji}} \le t, t < T_d + a_{ji} + \tau_{\mathrm{d}} \le t + M_i^t\}$; $\Omega_{i,2}^t$ represents the containers that are still fulfilling jobs but expected to be available in time period $t + 1$ to $t + M_i^t$, then $\Omega_{i,2}^t = \{d \in D_s | D_d = i, O_d = j, j \in P, s < t, T_d + a_{ji} > t, T_d + a_{ji} + \bar{\mathrm{b}}_i \le t + M_i^t\}$; $\Omega_{i,3}^t$ represents the containers that are planned for use in self-container jobs but have not shipped out yet, and are expected to be available in time period $t + 1$ to $t + M_i^t$, then $\Omega_{i,3}^t = \{d \in D_s | D_d = i, O_d = j, j \in P, s < t, T_d > t, T_d + a_{ji} + \bar{\mathrm{b}}_i \le t + M_i^t\}$. Equation (25) gives the adjusted inventory level.

$$S_i^m(t) = S_i(t)'' + \sum_{j\in P} \sum_{s\in\Omega_{ji,0}^t} Y_{ji}^s * \beta_s + \sum_{d\in\Omega_{i,1}^t} \mathrm{X}_d^{\mathrm{a}} + \sum_{d\in\Omega_{i,2}^t} \mathrm{X}_d^{\mathrm{a}} + \sum_{d\in\Omega_{i,3}^t} \mathrm{X}_d^{\mathrm{p}} -$$
$$\sum_{s=0}^{s=t-1} \sum_{d\in D_s, O_d=i} \sum_{T_d=\mathrm{t}+1}^{T_d=\mathrm{t}+M_i^t} M_d - \sum_{y\in Y_t} \sum_{O_y=i} \sum_{S_y=\mathrm{t}+1}^{S_y=\mathrm{t}+14} M_y \tag{25}$$

Equation (25) is the adjusted inventory process, and it is similar to Stage 1 except the booking successful rate, rejected jobs and predicted demands are included. β_s is the realized value for the successful booking rate of ETCR on day *s*. The last term is the total predicted container outflow for the following 14 days, i.e. the two-week forecast used in industry.

After every inventory level is adjusted, ETCR assignments are determined. If $P^{s}_{r,t}$ is taken as a set of surplus depots in region *r* at time *t*, any depot *i* of $P^{s}_{r,t}$ needs to be $\{i \in P^{s}_{r,t} | S^{m}_{i}(\mathrm{t})\text{-}\mathrm{U}_{i} > 0\}$. Likewise, if $P^{d}_{r,t}$ is the set of deficit depots in the same region, any depot *i* of $P^{d}_{r,t}$ should be $\{i \in P^{d}_{r,t} | \mathrm{L}_{i} - S^{m}_{i}(\mathrm{t}) > 0\}$. Equation (26) determines the ETCR assignments at time *t*.

$$\mathrm{Min} \sum_{i \in P^{s}_{r,t}} \sum_{j \in P^{d}_{r,t}} Y^{t}_{ij} * C_{ij}; \tag{26}$$

s.t.

$$\sum_{i \in P^{s}_{r,t}} \sum_{j \in P^{d}_{r,t}} Y^{t}_{ij} = Min\left[\sum_{i \in P^{s}_{r,t}} (S^{m}_{i}(\mathrm{t}) - \mathrm{U}_{i}), \sum_{j \in P^{d}_{r,t}} (\mathrm{L}_{j} - S^{m}_{j}(\mathrm{t}))\right]. \tag{27}$$

Event 4: Decisions towards new demands.

Since the consideration of FFs has been introduced at Stage 2, the determined ETCR amount from Event 3 may not be the actual repositioned amount due to the unreliability of the selected FF. To achieve greater profits, the choice of FFs will be optimized together with the decisions on meeting customer demands. However, without knowing the choice of FFs, this event cannot proceed. Hence, an FF is randomly selected with a cost of $f_t \in [\frac{2\mathrm{C}^{t}_{b}}{5}, \mathrm{C}^{t}_{b}]$. Based on the chosen FF, the associated booking success rate β_t can be realized, and the inventory level can be further updated to:

$$S_i(t)''' = S_i(t)'' - \sum_{j \in P^{d}_{r,t}} Y^{t}_{ij} * \beta_t \tag{28}$$

Then, taking the newly received demands D_t with execution date of $t + q$ as an example for the job assignments process, and using $\Omega^{t}_{ji,0}$ to represent the time periods before time *t* and the deployed ETCR containers from depot *j* to depot *i* will be available in the time periods from *t* + 1 *to t* + *q*, then $\Omega^{t}_{ji,0} = \{s \in (0,t) | t + 1 - a_{ji} \le s \le \min(\mathrm{t} + q - a_{ji}, t - 1)\}$; $\Omega^{t}_{i,1} = \{d \in D_s | D_d = i, O_d = j, j \in P, 0 \le s < t, T_d + a_{ji} \le t, t < T_d + a_{ji} + \tau_d \le t + q\}$ represents the containers that have finished jobs and started the cleaning process, and will be available in time period *t* + 1 to *t* + *q*; $\Omega^{t}_{i,2} = \{d \in D_s | D_d = i, O_d = j, j \in P, 0 \le s < t, T_d + a_{ji} > t, T_d + a_{ji} + \bar{\mathrm{b}}_i \le t + q\}$ represents the containers that are still fulfilling jobs but are expected to be available in time period *t* + 1 to *t* + *q*; $\Omega^{t}_{i,3} = \{d \in D_s | D_d = i, O_d = j, j \in P, 0 \le s < t, T_d > t, T_d + a_{ji} + \bar{\mathrm{b}}_i \le t + q\}$ represents the containers that are planned for use in self-container jobs but have not shipped out yet, and are expected to be available in time period *t* + 1 to *t* + *q*:

$$S_i(\mathrm{t} + \mathrm{q})' = S_i(t)''' + \sum_j \sum_{s \in \Omega^{t}_{ji,0}} Y^{s}_{ji} * \beta_s + \sum_{d \in \Omega^{t}_{i,1}} \mathrm{X}^{\mathrm{a}}_{d} + \sum_{d \in \Omega^{t}_{i,2}} \mathrm{X}^{\mathrm{a}}_{d} + \sum_{d \in \Omega^{t}_{i,3}} \mathrm{X}^{P}_{d} - \sum_{s=0}^{s=t} \sum_{d \in D_s} \sum_{O_d = i} \sum_{T_d = \mathrm{t}+1}^{T_d = \mathrm{t}+\mathrm{q}} \mathrm{X}^{P}_{d}; \tag{29}$$

Equation (29) is used to calculate the expected inventory level for depot *i* on day *t* + *q* after the planned container inflows and outflows are finished. The second term is used to obtain the amount of ETCR arrivals in time period *t* + 1 *to t* + *q*, and parameter β_s is the known value of the booking success rate for every ETCR arrangement on its associated day.

If $S_i(\mathrm{t} + \mathrm{q})' \ge \sum_{d \in D_s, s=t} \sum_{O_d = i} \sum_{T_d = \mathrm{t}+\mathrm{q}} M_d$, $\forall\ \mathrm{X}^{\mathrm{p}}_{d} = 1$, $\mathrm{Z}^{\mathrm{p}}_{d} = \mathrm{W}_d = 0$, for $\forall d \in \{d \in D_s | O_d = i, s = t, T_d = \mathrm{t} + \mathrm{q}\}$. (30)

If $S_i(\mathrm{t}+\mathrm{q})' < \sum_{d\in D_s, s=t}\sum_{O_d=i}\sum_{T_d=\mathrm{t+q}} M_d$, the assignment of self-container jobs, planned-leasing jobs and job rejections are determined by solving the following mathematical programming problem:

$$\text{Min} \quad \sum_{d\in D_s, s=t}\sum_{O_d=i}\sum_{D_d=j, j\in P}\sum_{T_d=\mathrm{t+q}}(\mathrm{X}_d^{\mathrm{p}}+\mathrm{Z}_d^{\mathrm{p}}) * \mathrm{C_{ij}} + \sum_{d\in D_s, s=t}\sum_{O_d=i}\sum_{D_d=j, j\in P}\sum_{T_d=\mathrm{t+q}} \mathrm{Z}_d^{\mathrm{p}} * \mathrm{C}_i^{\mathrm{l}} * a_{\mathrm{ij}} + \sum_{d\in D_s, s=t}\sum_{O_d=i}\sum_{D_d=j, j\in P}\sum_{T_d=\mathrm{t+q}} \mathrm{W}_d * \mathrm{C}_{ij}^{p}; \tag{31}$$

s.t.

$$\mathrm{X}_d^{\mathrm{p}}+\mathrm{Z}_d^{\mathrm{p}}+\mathrm{W}_d = 1 \text{ for } \forall d\in \{d\in D_s | O_d = i, s = t, T_d = t+q\}; \tag{32}$$

$$\sum_{d\in D_s, s=t}\sum_{O_d=i}\sum_{T_d=\mathrm{t+q}} \mathrm{X}_d^{\mathrm{p}} = S_i(t+q)'; \tag{33}$$

$$S_i(t+q)'' = S_i(t+q)' - \sum_{d\in D_s, s=t}\sum_{O_d=i}\sum_{T_d=t+q} \mathrm{X}_d^{\mathrm{p}}. \tag{34}$$

Equations (30) and (31) are the rules for assigning self-container jobs, planned-leasing jobs and rejections, and if there are not enough self-containers, the specific assignments are obtained by solving Equation (31). Equations (32)-(34) are the constraints for planning container outflows, if there are self-container jobs, planned-leasing jobs or rejections. The above steps, Equations (28)-(34), are repeated to finish the job assignments for all the demands received on day *t*.

According to the event description at the beginning of this stage, the optimized choice of FFs can be searched for within the range of f_t by running the loop from Equations (28)-(34) to maximize profit in Equation (35). β_t is the realized booking success rate for each loop:

$$\max_{f_t} \; EXP\,\{\sum_{d\in D_t}\sum_{T_d=\mathrm{t+1}}^{T_d=t+M_i^t}(M_d - W_d) * E_d - \sum_{d\in D_t}\sum_{i\in P}\sum_{T_d=\mathrm{t+1}}^{T_d=t+M_i^t} S_i^{T_d} * C_i^h - \sum_{d\in D_t}\sum_{O_d=i, i\in P}\sum_{D_d=j, j\in P}\sum_{T_d=\mathrm{t+1}}^{T_d=t+M_i^t}(M_d - W_d) * (C_{ij} + C_i^o + C_j^f + \mathrm{C}_b^t) - \sum_{i\in P}\sum_{j\in P} Y_{ij}^t * \beta_t * (\mathrm{C_{ij}} + C_i^o + C_j^f + f_t) - \sum_{d\in D_t}\sum_{D_d=j, j\in P}\sum_{T_d=\mathrm{t+1}}^{T_d=t+M_i^t} X_d^a * C_j^c - \sum_{d\in D_t}\sum_{O_d=i, i\in P}\sum_{T_d=\mathrm{t+1}}^{T_d=t+M_i^t} Z_d^p * C_i^l * a_{ij} - \sum_{d\in D_t}\sum_{O_d=i, i\in P}\sum_{D_d=j, j\in P}\sum_{T_d=\mathrm{t+1}}^{T_d=t+M_i^t} \mathrm{W}_d * C_{ij}^p\}. \tag{35}$$

4. Solution methods

4.1 The needs

The Stage 1 and Stage 2 models are difficult to solve analytically. First, they involve random variables and a large number of operational decisions (taking integer values). Second, to reflect the practices of real TCOs, these operational decisions need to be determined on an event-driven basis, which is difficult to formulate in a single mathematical programming model. Hence, the solution method proposed is a simulation-based heuristic, which allows the handling of workflow and constraints as well as the searching of the solution space and the execution of associated evaluations. However, some of its required input data comes from local mathematical programming optimizations, e.g. the everyday ETCR assignments are determined by linear programming. Therefore, hybrid elements are introduced to make the heuristic method a mixed optimization solution. For example, during Events 2, 3 and 4 in both simulation stages, linear programming is used jointly with certain rules to optimize ETCR amounts and job assignments. This allows:

i) an increase in computational tractability by using a heuristic method;
ii) an increase in effectiveness and efficiency by using a mathematical optimization method to find the local optimum.

Similarly, a math-heuristic is another hybridized optimization algorithm that uses interoperation of heuristics and mathematical programming. For example, Rath and Gutjahr (2014) propose a math-heuristic to optimize warehouse location routing. They use a mixed-integer linear programming formulation as the backbone and a constraint pool heuristic to reduce the expensive computational part for dealing with large problem spaces. Chen and Lau (2011) use a math-heuristic for resource scheduling in maritime logistics. They decompose their problem into two sub-problems, using heuristics for their machine-scheduling sub-problem, while using linear integer programming for their equipment allocation sub-problem. Comparing math-heuristics to the solution method applied in this paper, no matter how the mathematical and heuristic parts are structured, they are not built upon event-driven simulation. They are just an extension to either heuristic or mathematical programming methods to combine both of their advantages. In this paper, a math-heuristic can hardly be applied. This is because without the simulation process, it is hard to formulate the dynamic traits of the changing planning horizon and variable container cleaning times causing different container returns, and it is hard to handle some subtle issues such as the 2-week demand forecast and different groups of job-finished containers etc. Instead, a novel mixed optimization method is designed here to address the problem formulated in Section 3.

The simulation-based optimization method developed here consists of a GA search module and a simulation-based decision-making module. The latter uses a discrete event model of the operational level process of TC management and flows. This allows tracing of the TC holding cost, laden and empty container transportation costs, planned and emergent leasing costs, FF cost (Stage 2), container lifting-on/off cost, container cleaning cost and job rejection penalty of each order at each region and depot. It outputs the profit of a given solution, whereas the GA searches for better solutions.

4.2 Simulation module

The structure of the simulator is described in Appendix 1 and Appendix 2. In Stage 1 and Stage 2, it simulates the same daily process of each depot simultaneously (receive containers returned from cleaning and repositioning; arrange outflow containers to execute customer demands, determine ETCR and leasing; cope with new customer demands and planning etc.). It takes a set of input data including customer demands, inventory thresholds and initial net stock, and the shipping network with distances between regions and depots. It interacts with the decision-making module to receive its outputs for use in executing the four events. It records and allows tracing of the storage, loading, transhipment and unloading processes of each job and the inventory level of each depot. It outputs an operational level performance measure; the total profits.

In Stage 1, the decision-making module is limited to the assignment of self-container jobs and planned-leasing jobs and linear optimization of the order in which to take the jobs. In Stage 2, decisions are made with consideration of the 2-week demand forecast, and job rejections and choice of FFs are considered jointly. The job assignments are again made by linear optimization, and the output performance measure is used to determine the best choice of FFs. ETCR is the same for both stages. It first determines a depot's status as deficit or surplus, by comparing the current inventory, 'on the way' containers and deterministic future demands against the thresholds. Then, the ETCR activities (the quantities, origins and destinations of reposition containers) are deployed by solving a classic assignment problem. This decision-making process is performed dynamically in an event-driven module based on the input threshold values, customer demands and dynamic information obtained from the simulator.

To evaluate the performance of the system, in Stage 1 the relevant costs including handling costs, transport costs, leasing costs and inventory costs are calculated. The Stage 1 simulation terminates when the defined total simulation days are reached, in this case 180 days. In Stage 2, the FF costs and job rejection penalties are also calculated. Whenever ETCR is required on a given day, the FF optimization is run, and the overall simulation terminates when the defined total simulation days are reached, as for Stage 1.

4.3 The Heuristic Search Method (HSM)

To emulate observed industrial practice a heuristic is introduced to determine the threshold values in the inventory control policy. This utilizes the statistics of customer demands and inventory dynamics across the whole planning horizon. First, all the depots are grouped into surplus and deficit depots according to their overall TC net flow (e.g. a net import depot is a surplus depot). Next, the following key statistical indicators are estimated:

i. Average jobs per day in depot i (μ_i);
ii. Standard deviation of demands in depot i (σ_i);
iii. Least Inventory Level (LIL_i) for every depot i based on given container flow information;
iv. Largest Backlog Order (LBO_i) for depot i;
v. Largest Consecutive Container Net Outflow ($LCCNO_i$) for depot i.

Specifically, μ_i, σ_i, LBO_i and LIL_i can be obtained simply from demand information, while $LCCNO_i$ is determined as follows. Each depot's container net flow is monitored daily and, when its first net outflow occurs, the amount is recorded as the first Continuous Container Net Outflow ($CCNO_i$); this is a negative number. $CCNO_i$ is updated according to the net flow in the following days. Once $CCNO_i$ is updated as a positive number, it is returned to zero and this round of $CCNO_i$ updating is finished. The next round of $CCNO_i$ updating starts with the next net outflow. This is repeated until the end of the planning horizon, and for each depot. During the first iteration $LCCNO_i$ is set as the largest negative $CCNO_i$. At each further iteration, if there is a larger negative $CCNO_i$ then $LCCNO_i$ is updated, so it is the largest across all iterations. Appendix 3 summarises this process.

$LCCNO_i$ is the inventory a depot requires to meet all its customer demands. If it has less, leasing is required. If it has more, these can be fed to other depots. Therefore, the $LCCNO_i$ values are used as the upper threshold values for surplus depots, encouraging them to transfer TCs. The lower threshold values for surplus depots are decided from LIL_i, LBO_i and $\mu_i + \sigma_i$. If $LIL_i > 0$, the lower threshold value for the depot is set as 1, which means, with the safeguard of the upper threshold, it needs no external help to meet its demand. If $LIL_i = 0$ and $LBO_i > 0$, then even though this is a surplus depot, it still has a stock-out risk on a given day. Thus, minimum (LBO_i, $\mu_i + \sigma_i$) determines whether this depot should call for help based on the inventory level falling below the level required to meet its average demand.

For deficit depots, apart from the statistical indicators used above, Most Inventory Level (MIL_i) and 'Largest that can be Repositioned Amount' (LRA) in this region are also needed. MIL_i is the highest inventory level that this depot has ever reached. LRA is the total number of containers available for repositioning in the region. Maximum (MIL_i, $\mu_i + \sigma_i$) determines the lower bound for the deficit depot, helping it to call for more ETCR to increase the number of self-container jobs. The upper bound for deficit depots is set as minimum ($LCCNO_i$, LRA). This is because ETCR can only be intra-regional for the TC industry, therefore, $LCCNO_i$ is the amount that allows the deficit depots to meet all demands with self-owned containers, but it cannot exceed LRA. Appendix 4 summarises the heuristic for threshold values.

4.4 The Adapted Genetic Algorithm (AGA)

Alternatively, the threshold value in Stage 1 can be obtained and optimized using an AGA; one of the most commonly used meta-heuristic optimization approaches in container operations research (e.g. Dong and Song, 2009; Dang et al., 2013). The AGA used here is built upon a modification of the 'standard' or default Genetic Algorithm (GA) implemented in Matlab® using scattered crossover and Gaussian mutation (MathWorks, 2018). It is illustrated in Appendix 5.

As the first operation, the standard GA is performed with respect to the underlying problem. For the genetic representation (chromosome), the candidate solution consists of the double threshold values for each depot, coded as a vector of non-negative integers denoted as $\{[L_i, U_i] \mid i \in P\}$, where L_i and U_i are the lower and upper bounds of the inventory thresholds for depot *i*. A valid chromosome should satisfy the constraint $0 \le L_i \le U_i < N$. The initial population of solutions is generated randomly. Since the optimization is to maximize profit, the higher the objective function value (profit), the higher the solution fitness value should be. To achieve this, $E(q)$ is used to represent the total profits under the solution represented by chromosome *q*, then the fitness value of chromosome *q* is defined as $F_q = E(q) - \min\{E(q): 1 \le q \le N_p\}$, where N_p is the population size. For the parent selection process, roulette wheel sampling is used; each of two parents is selected from a binary tournament, which randomly picks two individuals from the entire population and retains the fittest. The two selected parents generate a child using scattered crossover. Fourth, probabilities are selected for crossover and mutation, and also, since pairs of elements in Stage 1 are formed by the lower and upper inventory bounds of a specific depot, these must be copied together to the offspring as a pair during crossover. Finally, all the parent and offspring chromosomes are sorted into descending fitness order and only the chromosomes with sequence numbers less than or equal to N_p are carried into the next generation.

After the setting of the standard GA, the next operation will run the simulation module iteratively to find an improved variable range for the target variables. Because, as a pilot study indicated, the variable range bordered by the current constraints (i.e. $0 \le L_i \le U_i < N$) is too broad to find a good result within an acceptable computation time, especially when the problem scale is large. Therefore, the range needs to be more precise (narrower) to help the GA to evolve fitter solutions within a shorter time. Specifically, this operation involves three major steps to reduce the variable constraints range and to fit the standard GA. First, the initial variable range is used to run the GA for a fixed number of generations to generate the first series of 'optimized' results. A value of 70 is used as the initial upper bound for the variable range because beyond this value the rate of convergence to optimality slowed down greatly in pilot experiments. Second, the upper threshold values (i.e. U_i) are gradually reduced concurrently, e.g. 65 to 61, and the simulation module is re-run to see if performance is affected. If it is not, it means the current value is too large and the range should be reduced further. This process is repeated until the evaluation results change, then values from the last run that made no changes to overall evaluation are used as the new variable range, and the GA optimization solver is run again to obtain the new series of 'optimized' results. In the final step, the GA parameters such as crossover and mutation probabilities, population size, stall generation limit (stop limit for no improvement) and selection methods are re-evaluated to determine the final results.

The above AGA is needed only for the Stage 1 threshold-value optimization problem, as the standard GA in Matlab® is effective and efficient enough for the Stage 2 FF optimization. In Stage 1 the population size is 50, and the GA terminates after 100 generations or when the improvement in best fitness < 0.001 for 10 consecutive generations. Stage 2 uses a population size of 20, and terminates after 20 generations or when the improvement in best fitness < 0.001 for 5 consecutive generations. Crossover rate is 0.8 and mutation rate is 0.2 for both stages.

5. Numerical examples

Computational tests of the model have been conducted with 'real' operational data from a major, global TCO. These tests have three purposes:

i. to investigate the feasibility of the model in solving realistically sized problems with basic PCs;
ii. to benchmark against general practices to understand the economic significance of the proposed tool in achieving better decision-making. (Stage 1 compares the performance of the optimization system with the practices in managing TC inventory, Stage 2 compares it with the practices choosing FFs and using different customer demand predictions);
iii. to quantify the influences of different factors on operational profits to generate managerial insights.

5.1 Model initialization

Due to the different objectives of the two stages, two different simulation environments were created. For both stages, the simulation horizon (i.e. overall planning horizon) is 180 days discretized in days. In delineating the global operation, nine depots are picked across three regions as shown in Figure 6.

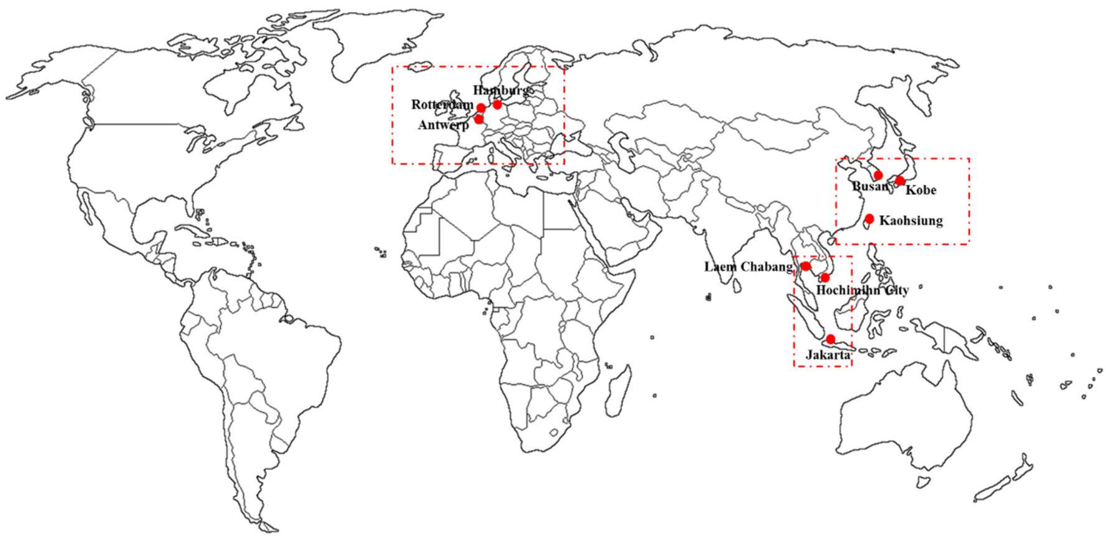


Figure 6. The depots and their related regions

Among the nine depots, the travel distances between any two are known and measured as shipping days. Transportation between any two depots is available but due to cost considerations, ETCR is only intra-regional.

During the 180 days in Stage 1, 1,003 customer demands occur. Every single demand represents one booking request and only one container is needed per booking. Demands are specified with origins, destinations, receiving date, execution date and expected revenues. The unit costs of inventory, lift-on/lift-off, container cleaning and job rejection penalty are listed in Table 1. The transportation cost per self-owned TC between two ports is assumed to be the transit time in days multiplied by a constant of £10. If the job is fulfilled with pre-planned leasing containers, the pre-planned leasing cost is £100 per day. For emergent leasing containers, the cost is £130 per day. The revenue per container ranges from £287 to £6,769. These values are generalized from the case TCO's data.

The initial inventory levels at the depots are uniformly distributed. The initial fleet size is designed to match the overall demands. Taking the average demand per day, the demand standard deviation and the average duration time for one job into consideration, the fleet size is rounded to 135 units in total.

Container cleaning duration is modelled as a random variable with a uniform distribution in the range 3 to 7 days. Again, this is a generalization of industrial data. Later, the truncated Normal distribution is used to evaluate the influence of different variances.

The model is implemented using Matlab® 2015 and a PC with a four-core 3.30 GHz processor.

Table 1. Cost parameters per TC

Inventory holding cost	Lift on & off	Job-Rejection Penalty	Cleaning	Self-container transportation	Planned leasing	Emergent leasing	FF cost
£3/day	£20	£200	£20	£10/day	£100/day	£130/day	£40

5.2 Preliminary computational experiments

This section presents preliminary computational experiments to test the models and solution procedures. To allow result reproducibility and future research, data used in this section (e.g. demands, revenues of individual jobs, job quote dates, job start dates, job origin and destination locations, travel times) is publicly available in Mendeley Data (https://data.mendeley.com/datasets/3jst3k2fyr/1) with DOI: 10.17632/3jst3k2fyr.1. It should be noted that due to the Non-Disclosure Agreement with the industrial partners, the data used in this section are generated randomly using the statistical attributes of the 'real' data used in later sections.

Firstly, the model is tested at Stage 1 focusing on seeking the 'optimal' inventory threshold values for each depot using the different ETCR methods, i.e. ETCR with AGA, ETCR with HSM, ETCR with RAIL and No ETCR. The results for operational performance in Table 2 show that in terms of total cost and profits, ETCR with AGA clearly outperforms the other methods.

Table 2. Comparison of overall results for the model at Stage 1 with different ETCR methods

Indicator	ETCR with AGA	No ETCR	% change from ETCR with AGA	ETCR with HSM	% change from ETCR with AGA	ETCR with RAIL	% change from ETCR wit AGA
Revenue	£1,434,639	£1,434,639	0%	£1,434,639	0%	£1,434,639	0%
Total cost	£365,390	£469,001	+28.4%	£496,046	+35.8%	£434,607	+18.9%
Profits	£1,069,249	£965,638	-9.7%	£938,593	-12.2%	£1,000,032	-6.5%

Secondly, the model is tested at Stage 2 for the three types of FF selection criteria (optimal, best and lowest cost). Based on the Stage 1 results, the inventory control policy obtained with ETCR with AGA is chosen as an input to Stage 2. A new series of demands is randomly generated, sharing the same statistical attributes as the data used in the first preliminary test. The TCO makes decisions on servicing new demands every day. Table 3 gives the operational performance under each FF selection criterion by averaging over ten experimental samples. Throughout the ten experiments, all the stochastic inputs (e.g. container cleaning time and successful repositioned containers) are randomly generated. By doing so, the operational performance tends to be robust statistically.

Table 3 Comparison of results (average of 10 samples) with different FF selection criteria

Indicators	Optimal FF	Best FF	% Diff. to optimal FF	Lowest cost FF	% Diff. to optimal FF
Self-container jobs	768	766	-0.41%	769	-0.41%
Planned-leasing jobs	186	186	+0.57%	184	0.00%
Emergent leasing jobs	84	87	-1.33%	85	+1.33%
Rejected jobs	32	30	+4.55%	32	+3.03%
Revenue	£1,387,483	£1,390,646	+0.23%	£1,387,761	+0.02%
Cost	£271,100	£284,741	+5.03%	£282,510	+4.21%
Profit	£1,116,382	£1,103,906	-1.12%	£1,098,251	-1.62%
Average FF level (measured by successful booking rate)	0.62	1	N/A	0.4	N/A

Tables 2 and 3 illustrate the effectiveness of conducting the optimisation processes in both Stage 1 and Stage 2, especially the cost reduction from applying ETCR with AGA at Stage 1 and using the optimal FF at Stage 2. Thus, albeit the different series of demands and uncertainties, higher profits are achieved with the proposed optimisation techniques for experiments from both stages.

In the following sections, 'real' industrial data is used in performing a full and detailed analysis and exploration of the proposed models and solutions. However, this real data will be not be made publicly available due to data protection.

5.3 Computation results in Stage 1

Experimental results in Table 4 compare ETCR with AGA and No ETCR. Also, it is compared with two other ETCR approaches seen in TCOs. First, ETCR is guided by a 'Regional Average Inventory Level' (RAIL) for each region. RAIL is obtained by averaging the up-to-date inventory level of all depots in the region. Depots with inventories lower than RAIL are fed by depots with inventories above RAIL. Figure 7 illustrates the process. Second, the threshold values of ETCR are determined heuristically by HSM.

Table 4. Comparison of results for No ETCR, ETCR with AGA, RAIL, and HSM

Indicators	ETCR with AGA	No ETCR	% change from ETCR with AGA	ETCR with RAIL	% change from ETCR with AGA	ETCR with HSM	% change from ETCR with AGA
Self-container jobs	730	685	-6.2%	784	+7.4%	708	-3.0%
Planned-leasing jobs	230	242	+5.2%	157	-31.7%	228	-0.9%
Emergent-leasing jobs	43	76	+76.7%	62	+44.2%	67	+55.8%
Number of ETCR	76	n/a	n/a	458	+502.6%	21	-72.4%
Total costs	£320,469	£401,130	+25.2%	£387,239	+20.8%	£373,009	+16.4%
Total profits	£1,105,154	£1,024,493	-7.3%	£1,038,384	-6.0%	£1,052,614	-4.8%
Total inventory costs	£37,869	£40,860	+7.9%	£35,469	-6.3%	£39,609	+4.6%
Cost of self-container jobs	£84,890	£77,070	-9.2%	£89,900	+5.9%	£79,810	-6.0%
Cost of planned-leasing jobs	£121,900	£165,200	+35.5%	£87,400	-28.3%	£148,800	+22.1%
Cost of emergent leasing jobs	£38,870	£84,240	+116.7%	£109,590	+181.9%	£69,940	+79.9%
Utilization	67.1%	47.3%	-29.5%	75.6%	+12.7%	61.7%	-8.0%
Utilization for jobs	62.9%	47.3%	-24.8%	58.4%	-7.2%	59.2%	-5.9%

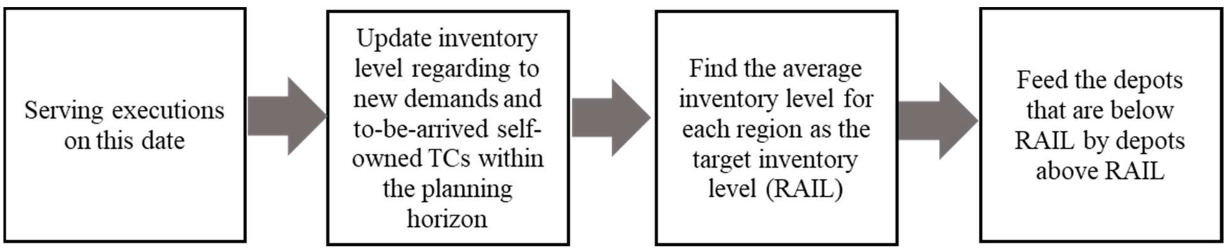


Figure 7. The process for RAIL-based ETCR

Compared to ETCR with AGA, the profit with No ETCR is 7.3% lower, ETCR with RAIL is 6% lower and ETCR with HSM is nearly 5% lower. The improvement with ETCR with AGA is mainly due to reductions in planned-leasing and emergent leasing costs. Specifically, No ETCR yields 35.5% higher planned-leasing cost and 116.7% higher emergent leasing cost, ETCR with RAIL yields 181.9% higher emergent leasing cost (reduction of planned-leasing cost for ETCR with RAIL is not enough to compensate for higher emergent leasing cost), and ETCR with HSM yields 22.1% higher planned-leasing cost and 79.9% higher emergent leasing cost. Strikingly, ETCR with RAIL results in approximately 6 times more ETRC movements than ETCR with AGA.

To understand what is happening, first note that No ETCR yields 5.2% more planned-leasing jobs than ETCR with AGA, but the increase in the planned-leasing cost is nearly 7 times more at 35.5%. To understand this, consider an example with two depots starting with inventory levels of 2 and 8 self-owned containers respectively. Suppose each depot has 10 demands, 4 of which are long duration and 6 short duration. Without ETCR, the first depot would have to service 2 long duration demands with planned-leasing containers. If instead the inventory had been rebalanced by ETCR, say to inventory levels of 4 and 6, then there would be no need to service long duration jobs with planned-lease containers. Therefore, although the overall number of planned-leasing jobs would remain the same at 10, the cost would fall significantly. This means ETCR is not so much reducing the number of planned-leasing jobs as focusing them on to shorter duration demands by having more balanced inventories across the depots.

The more balanced inventories are also spreading out the ability of inventory to provide a buffer to protect against stochasticity and subsequent emergent leasing, with ETCR with AGA reducing the number of emergent leasing jobs by 43% (No ETCR is 76.7% higher), and their cost by 54% (No ETCR is 116.7% higher). If inventory is not balanced then low-inventory depots will arise and these are more likely to need emergent leasing.

The very high volume of ETCR movements for ETCR with RAIL (approximately 6 times ETCR with AGA) means more inventory balancing is occurring, resulting in more self-container jobs, and therefore less planned-leasing, because the self-containers are more often in the right place for outflows. However, compared to ETCR with AGA this does not translate into higher profits. This is predominantly because ETCR with RAIL yields a big increase in the number (44.2%) and cost (181.9%) of emergent jobs in addition to the greatly increased amount of ETCR, which is not offset by a sufficient reduction in planned leasing. The increase in emergent-leasing costs is more than four times the increase in number of emergent-leasing jobs. This means that ETCR with RAIL not only yields more emergent leasing, but this tends to be for more expensive longer duration jobs, i.e. a double or amplified shortcoming. Two more metrics were introduced to further analyse this phenomenon. 'Utilization' is the average total time TCs spend on job related activity or ETCR during the whole planning horizon (180 days), with job activity including laden delivery, holding by receiver, cleaning and return to inventory. 'Utilization-for-jobs' is just the average time TCs spend on job related activity excluding ETCR. In Table 2, these show that while ETCR with RAIL is keeping TCs very busy, much of this activity is taken up with ETCR with the result that the Utilization-for-jobs is less than for ETCR with AGA, which in turn is better than ETCR with HSM.

What we are seeing is that ETCR with RAIL results in hugely excessive TC repositioning. It does yield higher profits than No ETCR, but these are still 6% less than ETCR with AGA yields with a sixth of the amount of repositioning. This is a very important result as ETCR with RAIL is a natural way for industry to work, demonstrating the practical value of the new ETCR with AGA. ETCR with RAIL is too focused on immediate rebalancing of inventories rather than planning using a longer-term perspective of net flows and inventory levels. For example, excessive ETCR can be caused when a long-term deficit depot temporarily has sufficient inventory, or a long-term surplus depot is temporarily deficient. ETCR with AGA looks further ahead, making more considered decisions, rather than rushing to reposition based on current inventory levels.

ETCR with HSM yields far fewer ETCR movements than ETCR with AGA (-72.2%) but far more emergent leasing (+55.8%) with an even bigger increase in emergent leasing costs (+79.9%). Quite simply, ETCR with HSM is simply not doing enough repositioning and this is resulting in a big increase in emergent leasing to cover for local shortages. Clearly, ETCR with AGA is yielding better results by achieving a better balance between over and under repositioning, compared with ETCR with RAIL and ETCR with HSM.

The cleaning duration is stochastic. In order to evaluate the ETCR policy's robustness and sensitivity to the spread of the cleaning times within the range [3,7], the time is modelled using a Normal distribution with mean 5 days and truncated beyond the [3,7] range. Then, three experiments were run with the variance set to 0.5, 1 and 2 respectively, with the results in Table 5.

Table 5. Comparison of results under normal distribution with different standard deviations

Indicators	AGA with normal cleaning $b_i \sim N(5, 0.5)$	AGA with normal cleaning $b_i \sim N(5, 1)$	Difference from $b_i \sim N(5, 0.5)$	AGA with normal cleaning $b_i \sim N(5, 2)$	Difference from $b_i \sim N(5, 0.5)$
Self-container jobs	732	734	+0.3%	727	-0.7%
Planned-leasing jobs	227	227	+0.0%	230	+1.3%
Emergent-leasing jobs	44	42	-4.6%	46	+4.6%
No. of ETCR movements	78	75	-3.9%	73	-6.4%
Total costs	£309,439	£311,549	+0.7%	£316,434	+2.3%
Total profits	£1,116,184	£1,114,704	-0.1%	£1,109,189	-0.6%
Inventory costs	£37,389	£37,569	+0.5%	£37,704	+0.8%
Cost of self-container jobs	£85,770	£85,660	-0.1%	£85,200	-0.7%
Cost of planned-leasing jobs	£122,200	£120,000	-1.8%	£121,900	-0.3%
Cost of emergent leasing jobs	£27,040	£31,330	+15.9%	£34,840	+28.9%

Table 3 shows that the key effect of increased variability in cleaning times is a shift in costs to emergent leasing. This is understandable as emergent leasing is used to cope with unavailability of self-owned TCs. The practical implication is that TCOs should increase the reliability and certainty of the container cleaning process, not just the mean duration, to reduce emergent leasing costs and increase profits.

5.4 Computation results in Stage 2

In this stage, the model is advanced to apply the joint decision-making process associated with ETCR, job fulfilments and choice of FFs on a day-to-day basis. To reduce the computation complexity, the cost of the best FF is fixed at £40 per job (see Table 1) across all regions. In reality, the cost of an FF may be different from region to region or even from route to route. However, the simplified value used here is sufficient to demonstrate the effectiveness of the model. During real-time decision-making, TCOs do not need to simulate such a long period as in the tests here, so their computation time will be less, allowing them to increase data complexity.

This stage introduces a penalty cost for rejecting demands to achieve greater profits. It is first set as £200 per job and varied later to test the model's sensitivity to it. The two-week look-ahead approach observed in industry is used, so in ETCR calculations the adjusted inventory level will allow for predicted customer demands. The demands' mean and standard deviations from Stage 1 are used to generate new demands. Decisions are then made on job fulfilment for demands received daily and the FF for ETCR based on the threshold values obtained from Stage 1. The simulation length is again 180 days, and performance is evaluated with indicators at the end of the planning horizon. Since the influence of different FFs over ETCR is subject to the stochasticity in the model, the simulation is run 10 times and the results averaged for each scenario.

To articulate the significance of optimizing the FF choice with the proposed model, the best, lowest-cost, and random FFs (uniform probability) are also applied for comparison. Using best FFs represents TCOs who wish to guarantee smooth execution of their plans, i.e. to compete on service quality, although this is expensive. Using lowest-cost FFs represents TCOs competing on price by offering low-cost services, but to the detriment of service quality/reliability. Random FF represents TCOs with limited access to the FF market and limited market power in making choices; they have to take whatever they can get due to capacity constraints in the industry.

Table 4 shows best FF yields better profits and job fulfilment than random FF and lowest-cost FF, but optimal FF yields the best profit. This is achieved by big reductions in FF cost (best FF is 114.9% higher) and the cost of emergent leasing jobs (best FF is 19.9% higher). Underlying the improvement is a substantial reduction in ETCR movements (best FF is 11.4% higher). From a strategic management perspective, this has advantages beyond just an increase in profit. It also means that the TCO is not beholden to just the best FF as another better FF can be identified due to its lower costs. Even if the profit differences are small, having a feasible alternative opens up competition that could drive costs lower, and having options in service providers is always strategically important.

In order to evaluate the continued robustness of the model at this stage, a simulation was run with no ETCR and no job rejection, yielding a total profit of £1,042,336. This is clearly less than that achieved across Table 6, demonstrating the continued effectiveness of ETCR.

Table 6. Results for Optimized and Non-Optimized FF choices

	Indicators	Optimal FF	Best FF	% Diff. to optimal FF	Random FF	% Diff. to optimal FF	Lowest cost FF	% Diff. to optimal FF
Self-container jobs		746	750	+0.5%	717	-3.9%	708	-5.1%
Planned-leasing jobs		230	231	+0.4%	241	+4.8%	258	+12.2%
Emergent leasing jobs		46	43	-6.5%	59	+28.3%	49	+6.5%
Rejected jobs		53	51	-3.8%	58	+9.4%	60	+13.2%
No. of ETCR		79	88	+11.4%	51	-34.5%	34	-57.0%
Total revenue		£1,387,056	£1,391,380	+0.3%	£1,381,441	-0.4%	£1,376,596	-0.8%
COSTS	Inventory	£38,184	£38,151	-0.1%	£39,129	+2.5%	£39,717	+4.0%
	FF	£1,638	£3,520	+114.9%	£1,427	-12.9%	£544	-66.8%
	Penalty	£10,600	£10,200	-3.8%	£11,600	+9.4%	£12,000	+13.2%
	Self-container jobs	£94,300	£94,220	-0.1%	£92,760	-1.6%	£91,350	-3.1%
	Planned-leasing jobs	£51,600	£52,200	+1.2%	£55,400	+7.4%	£52,300	+1.4%
	Emergent leasing jobs	£46,800	£56,100	+19.9%	£57,880	+23.7%	£59,400	+26.9%
	Total (not sum of above)	£280,852	£292,511	+4.2%	£293,871	+4.6%	£290,655	+3.5%
Total profits		£1,106,203	£1,098,869	-0.7%	£1,087,570	-1.7%	£1,085,941	-1.8%

The optimized and lowest cost FFs are a source of stochasticity due to their random reliability value. To see their effect, Figure 8 presents the profits from 10 repeated experiments with the same randomly generated stream of container cleaning times, but different random values for FF reliability. As the reliability of the best FF is constant at 100%, its profits are constant. The optimized FF gives higher profitability than the best FF in 7 experiments. The lowest cost FF yields the highest profit in 2 experiments when by random chance it produces high reliability, but it clearly gives the lowest profits on 5 occasions. The random FF yields a dynamic profit-making ability, but with lower average profits close to that of the lowest cost FF.

The reliability of the optimized FFs during each experiment ranged from 50% to 80%, showing that this optimization is truly using the FF range and not just going for high reliability FFs.

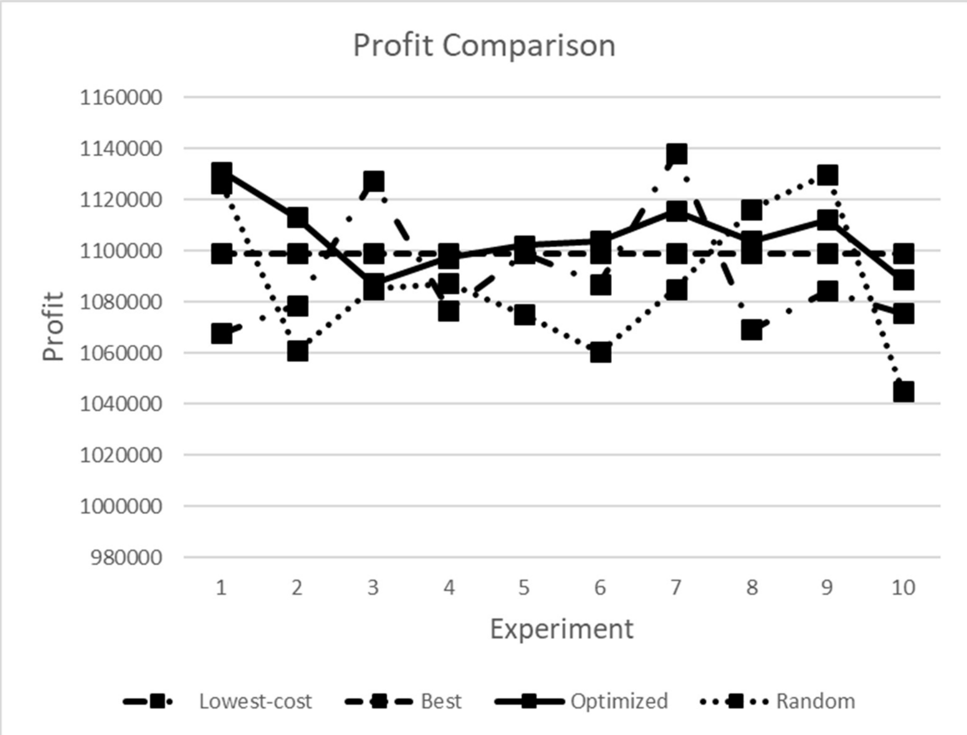


Figure 8. Profit by experiment for different FF criteria

Table 4 shows that optimizing the FF reduces leasing costs. If the penalty cost of rejecting a job were increased one would expect to see more leasing to accommodate a reduction in rejections. Table 7 presents the results when the penalty cost for rejecting jobs is varied. This shows that as the penalty cost, planned leasing cost or emergent leasing cost increase so FF optimization yields greater improvements in profit. This is due to large decreases in costs rather than increases in revenue that remain slightly lower in the experiments conducted.

Table 7. Effect of Key Cost Coefficients on Change in Profit for Optimized FF compared to Best FF

Penalty Cost	Planned leasing	Emergent leasing	Profit	Revenue	Cost
£100	£100	£130	-2.1%	-1.2%	+2.1%
£100	£120	£130	-0.7%	-0.3%	+0.9%
£100	£120	£150	+0.4%	-0.6%	-1.2%
£500	£100	£130	+1.1%	-0.5%	-5.3%
£500	£120	£130	+1.8%	-0.5%	-7.2%
£500	£120	£150	+2.9%	-0.3%	-8.4%
£1000	£100	£130	+2.4%	-0.2%	-8.1%
£1000	£120	£130	+3.1%	-0.3%	-8.9%
£1000	£120	£150	+4.4%	-0.4%	-9.6%

In line with industrial practice, a 2-week forecast was used in optimizing ETCR and inventory planning in Stage 2. To understand the effectiveness of incorporating the 2-week forecast into the optimization (2-week forecast + ETCR with AGA), it is compared with not using the forecast in Table 8. This shows that including the forecast yields a substantial decrease in leasing, penalty and ETCR costs with a corresponding increase in profits. The increased visibility given by the forecast is allowing the plans to achieve more with self-containers instead of resorting to leasing and excessive container repositioning.

To see if including the forecast changes the superiority of ETCR with AGA, Table 9 presents the change in performance seen when using ETCR with AGA compared to ETCR with HSM and ETCR with RAIL, with all three now using the 2-week forecast. This shows that ETCR with AGA is still the most profitable, increasing profit by 6.7% and 7.3% respectively, as it makes better use of self-containers resulting in lower leasing and ETCR costs.

Table 8. Change in ETCR with AGA performance after including 2-week forecast

Indicator	Change in performance with 2-week forecast
Revenue	+3.4%
Profit	+4.7%
Total Cost	-9.6%
Cost of self-container jobs	+2.1%
Cost of planned-leasing jobs	-6.4%
Cost of emergent-leasing jobs	-8.2%
Penalty cost	-9.2%
ETCR cost	-12.3%

Table 9. Change in performance when using ETCR with AGA compared to other optimization methods after including 2-week forecast

Indicator	ETCR with HSM & 2-week forecast	ETCR with RAIL & 2-week forecast
Revenue	+3.7%	+2.8%
Profit	+6.7%	+7.3%
Total Cost	-18.7%	-22.4%
Cost of self-container jobs	+9.4%	+3.5%
Cost of planned-leasing jobs	-15.7%	-34.7%
Cost of emergent-leasing jobs	-17.1%	-87.1%
Penalty cost	-14.1%	-19.1%
ETCR cost	-4.4%	-476.7%

Having seen the benefits of incorporating a 2-week forecast into the optimization, sensitivity to the length of the forecast period was investigated. Table 10 compares the results for 1, 2, 3 and 4-week forecasts with optimized FF. The 3-week forecast yields the highest profit through increasing the number of self-container jobs, resulting in decreased leasing jobs, and particularly the cost of these as it is longer more expensive jobs that are being switched to self-containers. The 1-week and 4-week forecasting periods result in increased ETCR. In the case of 1-week this is because it is too short to take into account future demands for the surplus depots, i.e. it is approaching no ETCR, so TCs are shipped to deficit depots too readily. In the case of 4-week forecasting, although more of the future demand forecast is considered, the forecast demand only tells where the origin is, but not the destination. Therefore, when the inventory is planned for the future, some of the future arrivals are not clear. However, if containers are reserved or repositioned for the whole 4-week forecast, too many TCs may be kept or moved, as the cleaned arrival TCs replenish the inventory as well.

Considering the current average job and cleaning durations, most containers will be ready for their next job within three weeks. Combining this with the above result the inference is that it is not beneficial to forecast beyond the typical job plus cleaning time. As the average job plus cleaning duration may be subject to change, due to changes in demand patterns, transport facilities or cleaning processes etc., it follows that TCOs should monitor this and adjust the forecast period accordingly.

Table 10. Sensitivity analysis to forecast period with optimized FF

Indicators		**Forecast Period**			
		1-Week	**2-Week**	**3-Week**	**4-Week**
Self-container jobs		724	746	763	733
Planned-leasing jobs		244	230	221	241
Emergent leasing jobs		49	46	43	46
Rejected jobs		58	53	48	55
No. of ETCR		92	79	71	88
Total revenue		£1,384,251	£1,387,056	£1,391,093	£1,385,596
Costs	Inventory costs	£35,772	£38,184	£36,594	£43,683
	FF cost	£1,940	£1,638	£1,149	£1,868
	Penalty costs	£11,600	£10,600	£9,600	£11,000
	Cost for self-container jobs	£91,630	£94,300	£98,820	£92,350
	Cost for planned-leasing jobs	£54,200	£51,600	£49,200	£53,300
	Cost for emergent leasing jobs	£48,100	£46,800	£38,740	£45,240
	Total (not sum of above)	£297,537	£280,852	£271,093	£294,651
Total profits		£1,086,714	£1,106,204	£1,120,000	£1,090,945

6. Conclusions

To improve Tank Container (TC) operations management, this study has proposed a two-stage model that enables optimization of a double-threshold inventory control policy for tank Container Operators (TCOs) to gain better operational profits, as well as demand fulfilment, during the quotation-booking process. On top of the optimized inventory policy, the model simulates and optimizes the choice of Freight Forwarders (FFs) under a more realistic operational environment including job rejections under a two-week demand forecast, on a rolling planning basis. The effectiveness of the two-stage model in optimization has been demonstrated through a series of numerical tests. Sensitivity analysis has articulated the managerial insights associated with the model with respect to different uncertainty levels, different FF costs and different demand forecast lengths.

A novel dynamic ‘rolling horizon’ for planning has been introduced to emulate the TC quotation-booking process seen in industry. The length of a given day’s planning horizon varies with the latest execution date of the future jobs being planned on that day. This rolling horizon is different to the static rolling horizons seen in the literature, e.g. Di Francesco et al. (2013).

In filling the academic vacuum in TC operations studies, this paper has built a more comprehensive picture of TC operations on top of the existing literature. Especially, it has modelled the overall quotation-booking process with the incorporation of more practical issues such as uncertain container cleaning times, choice of FFs, emergent leasing and customer demand forecast.

Due to the incorporation of various key real-world features of TC operations, this research has practical value for industry. Not only can practitioners benefit from more efficient and effective real-time booking quotations, they can also benefit from better decision-making support with the evaluation of different scenarios for both short-term and long-term TC operations. Further, within a dynamic external environment, the model can help TCOs to find a better position for their own interests when their bargaining power varies and the cost of FFs fluctuates.

Through numerical experiments understanding of the economics behind the decisions in this important industrial operation has been gained. Common practices including inventory control policy, choice of FFs and customer demand prediction have been emulated and their performance compared with that of the new optimization model presented here. Key findings from the experiments include:

i. The optimized inventory control policy has shown great advantage in controlling the inventory cost, increasing TC utilization for self-container jobs, reducing emergent leasing (which is particularly expensive) and improving overall profits under dynamic market conditions. From the strategic perspective, the optimized inventory control policy has produced more precise resource allocation, allowing TCOs to exploit market opportunities better in the long-term regardless of various uncertainties

ii. The benefits of optimizing with respect to profits the choice of FF for empty TC repositioning (ETCR), rather than always using the most reliable and therefore expensive FF, have been demonstrated. From a strategic management perspective, this has important advantages beyond just profits as it means that the TCO is not dependent on just the most reliable FF. Even if the profit differences are small, having a feasible alternative opens up competition that could drive costs down and service quality up.

iii. The importance of including stochastic TC cleaning times has been demonstrated, as this is a source of uncertainty that leads to emergent leasing when TCs are held up. Experiments have shown that increased reliability (reduced variation) in cleaning times results in higher profits due to reduced emergent leasing due to increased certainty in planning. This means that TCOs should aim for more reliable (less variable) cleaning times and not just shorter cleaning times.

iv. Taking into consideration a demand forecast in the optimization can reduce excessive ETCR that would cause higher costs and less profits. Experiments with ETCR guided by regional average inventory levels (ETCR with RAIL), which emulates a natural industrial practise, have revealed that not taking a longer-term perspective in planning and simply repositioning TCs based on current inventory levels results in hugely excessive repositioning, as well as more emergent-leasing and this tends to be for expensive, longer distance jobs. The greatly reduced repositioning, and thereby greater profits, achieved with ETCR with adapted GA (AGA) using demand forecasts, demonstrates the validity of the novel approach presented here, and in particular the value of taking a longer-term perspective of net flows and inventory levels in planning ETCR. To this end TCOs should aim to develop their forecasting capabilities to achieve more accurate forecasts. Results have shown that the forecast horizons should correspond to the typical TC job plus cleaning times for best results, and it is recommended that TCO's monitor their average job plus cleaning times with a view to revising forecast horizons accordingly.

Whilst this paper has contributed new fundamental knowledge to be used in improving profits from TC operations, it should also inspire others to research into this important industrial topic that has been largely neglected in the literature to-date.

This study could be expanded in two directions. First, at the operational level and in respect of demand fulfilment, the current simulation model assigns different job types in a rule-based fashion. Instead, decisions on self-container jobs, planned-leasing jobs and rejected jobs could be optimized simultaneously and solved by mathematical programming techniques. Second, at the strategic level, the TC fleet size in the experiments is developed from the given TCO data, but it is not optimized. The model could be developed further to optimize the fleet size, particularly as TCs are an expensive asset. In addition, different types of TCs could be considered.

References

BRAEKERS, K., JANSSENS, G.K. and CARIS, A., 2011. Challenges in managing empty container movements at multiple planning levels. *Transport Reviews,* **31**(6), pp. 681–708.

CHEN, H. AND LAU, H. 2011. A math-heuristic approach for integrated resource scheduling in a maritime logistics facility, *2011 IEEE International Conference on Industrial Engineering & Engineering Management,* pp. 195.

CHEUNG, R. K. and CHEN, C.Y. (1998). A two-stage stochastic network model and solution methods for the dynamic empty container allocation problem. *Transportation Science*, 32(2), 142–162.

CHOONG, S.T., COLE, M.H. and KUTANOGLU, E., 2002. Empty container management for intermodal transportation networks. *Transportation Research Part E,* **38**, pp. 423-438.

CRAINIC, T.G., GENDREAU, M. and DEJAX, P., 1993. Dynamic and stochastic models for the allocation of empty containers. *Operations Research,* **41**, pp. 102-126.

DANA, J.D., 1998. Advance-Purchase discounts and price discrimination in competitive markets. *The Journal of Political Economy,* **106**(2), pp. 395-422.

DANG, Q.V., NIELSEN, I.E. and YUN, W.-., 2013. Replenishment policies for empty containers in an inland multi-depot system. *Maritime Economics & Logistics,* **15**, pp. 120–149.

Di FRANCESCO, M., LAI, M. and ZUDDAS, P., 2013. Maritime repositioning of empty container under uncertain port disruptions. *Computers & Industrial Engineering,* **64**, pp. 827-837.

Di FRANCESCO, M., LAI, M. and ZUDDAS, P., 2009. The effect of multi-scenario policies on empty container repositioning. *Transportation Research Part E,* **45**, pp. 758-770.

DONG, J.X. and SONG, D.P., 2012. Lease term optimization in container shipping systems. *International Journal of Logistics Research and Applications,* **15**(2), pp. 87-107.

DONG, J.X. and SONG, D.P., 2009. Container fleet sizing and empty repositioning in liner shipping systems. *Transportation Research Part E,* **45**, pp. 860-877.

EPSTEIN, R., NEELY, A., WEINTRAUB, A.V.F., HURTADO, S., GONZALEZ, G., BEIZA, A., NAVEAS, M., INFANTE, F., ALARCON, F., ANGULO, G., BERNER, C., CATALAN, J., GONZALEZ, C. and YUNG, D., 2012. A strategic empty container logistics optimization in a major shipping company. *Informs,* **42**, pp. 5-16.

ERERA, A.L., MORALES, J.C. and SAVELSBERGH, M., 2009. Robust optimization for empty repositioning problems. *Operations Research,* **57**(2), pp. 468-483.

ERERA, A.L., MORALES, J.C. and SAVELSBERGH, M., 2005. Global intermodal tank container management for the chemical industry. *Transportation Research Part E,* **41**, pp. 551-566.

GALE, I.L. and HOLMES, T.J., 1992. The efficiency of advance-purchase discounts in the presence of aggregate demand uncertainty. *International Journal of Industrial Organization,* **10**(3), pp. 413.

GRAND VIEW RESEARCH, 2016. *Petrochemicals market analysis by product (Ethylene, Propylene, Butadiene, Benzene, Xylene, Toluene, Vinyls, Styrene, Methanol) by region (North America, Europe, China, Middle East, Africa, Latin America) and segment forecasts, 2014 - 2025.* 978-1-68038-560-1. U.S.: Grand View Research.

HE, P., 2013. Optimization of empty container allocation in depots with the participation of freight forwarders, *Proceedings of the 4Th International Conference on Transportation Engineering* 2013, pp. 1797-1802.

ITCO, 2016. *ITCO tank container fleet survey.* ITCO.

KARIMI, I.A., SHARAFALI, M. and MAHALINGAM, H., 2005. Scheduling tank container movements for chemical logistics. *American Institute of Chemical Engineers,* **51**(1), pp. 178-197.

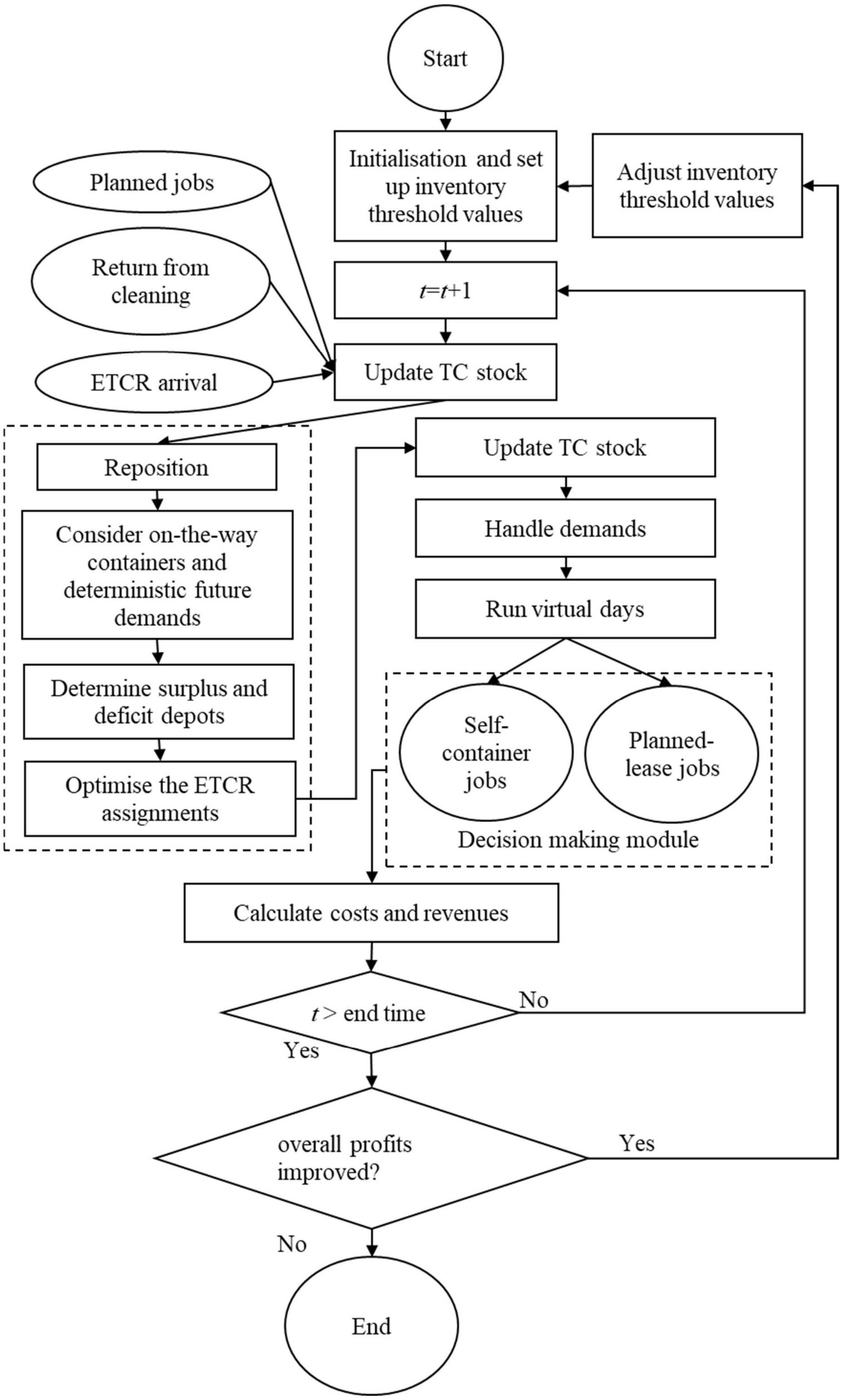
LAM, S.W., LEE, L.H. and TANG, L.C., 2007. An approximate dynamic programming approach for the empty container allocation problem. *Transportation Research Part C,* **15**, pp. 265–277.

MATHWORKS, 2018. *Matlab R2018a: Documentation-functions-gaoptimset*, Retrieved from URL: https://uk.mathworks.com/help/gads/gaoptimset.html

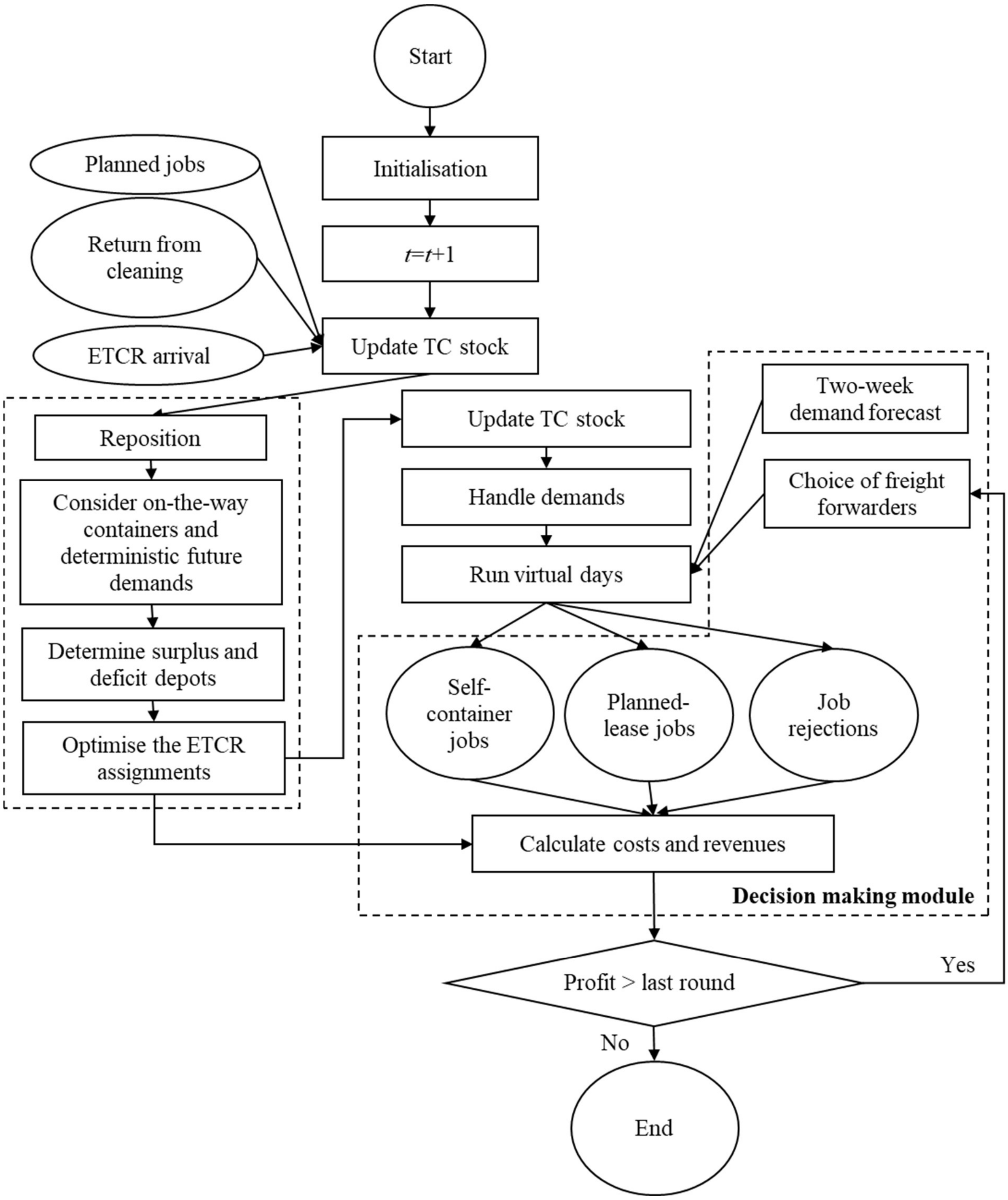
MENG, Q. and WANG, S., 2011. Liner shipping service network design with empty container repositioning. *Transportation Research Part E,* **47**, pp. 695-708.

MOON, I., NGOC, A.D.D. and HUR, Y.S., 2010. Positioning empty containers among multiple ports with leasing and purchasing considerations. *OR Spectrum,* **32**, pp. 765–786.
OLIVO, A., Di FRANCESCO, M. and ZUDDAS, P., 2013. An optimization model for the inland repositioning of empty containers. *Maritime Economics & Logistics,* **15**, pp. 309–331.
RATH, S. AND GUTJAHR, W., 2014, A math-heuristic for the warehouse location–routing problem in disaster relief. *Computers and Operations Research*, **42**, pp. 25-39.
SONG, D.P. and DONG, J.X. 2011. Effectiveness of an empty container repositioning policy with flexible destination ports. *Transport Policy*, **18**(1), 92–101.
SONG, D.P. and DONG, J.X., 2013. Long-haul liner service route design with ship deployment and empty container repositioning. *Transportation Research Part B,* **55**, pp. 188–211.
SONG, D.P. and DONG, J.X., 2015. Empty container repositioning. In: C.Y. LEE and Q. MENG, eds, *Handbook of Ocean Container Transport Logistics- Making Global Supply Chain Effective.* London: Springer, pp. 163-208.
STEADIESEIFI, M., DELLAERT, N.P., NUIJTEN, W., VAN WOENSEL, T. and RAOUFI, R., 2014. Multimodal freight transportation planning: A literature review. *European Journal of Operational Research,* **233**, pp. 1-15.
TANG, W. and GIROTRA, K., 2017. Using advance purchase discount contracts under uncertain information acquisition cost. *Production and Operations Management.*
YUN, W.Y., LEE, Y.M. and CHOI, Y.S., 2011. Optimal inventory control of empty containers in inland transportation system. *International Journal of Production Economics,* **133**(1), pp. 451-457.
ZHENG, J.F., SUN, Z. and ZHANG, F.J., 2016. Measuring the perceived container leasing prices in liner shipping network design with empty container repositioning. *Transportation Research Part E,* **94**, pp. 123-140.

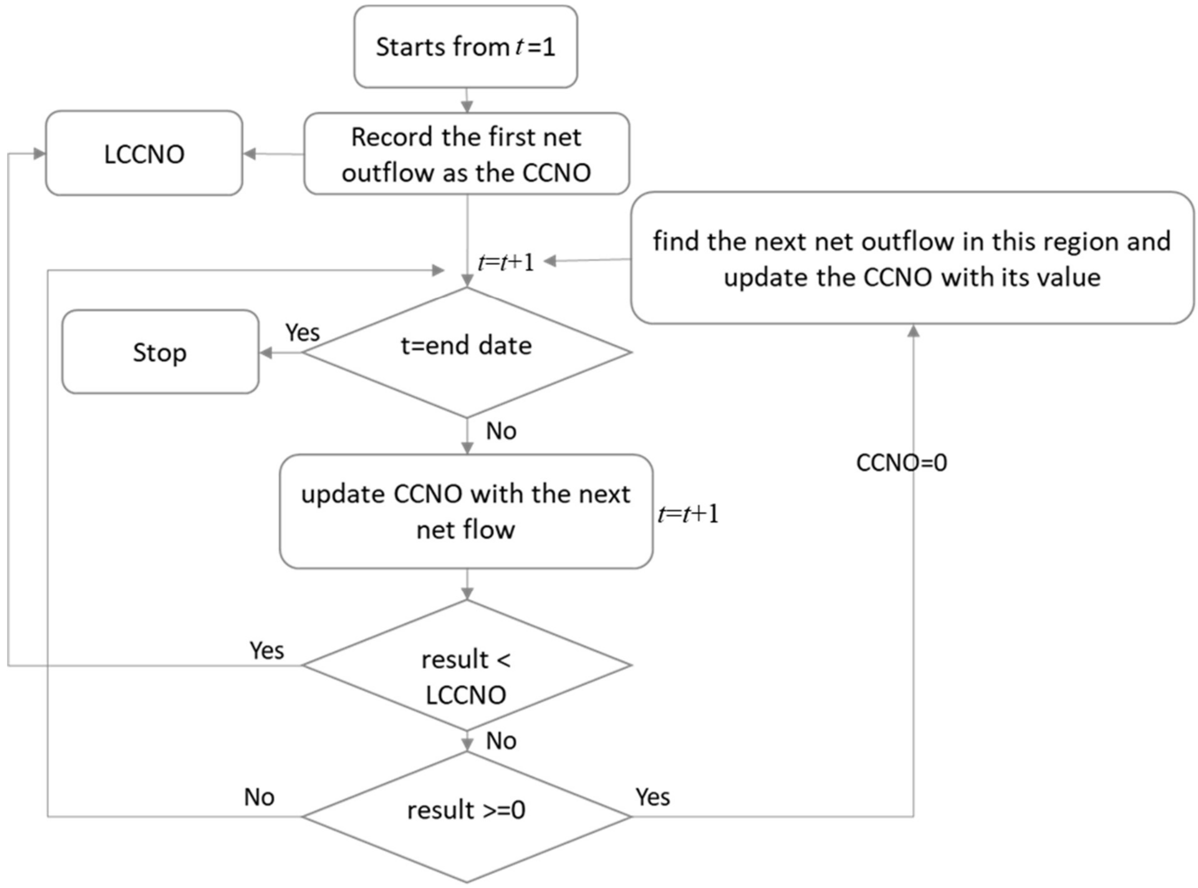
Appendix 1. The Simulation Module in Stage 1



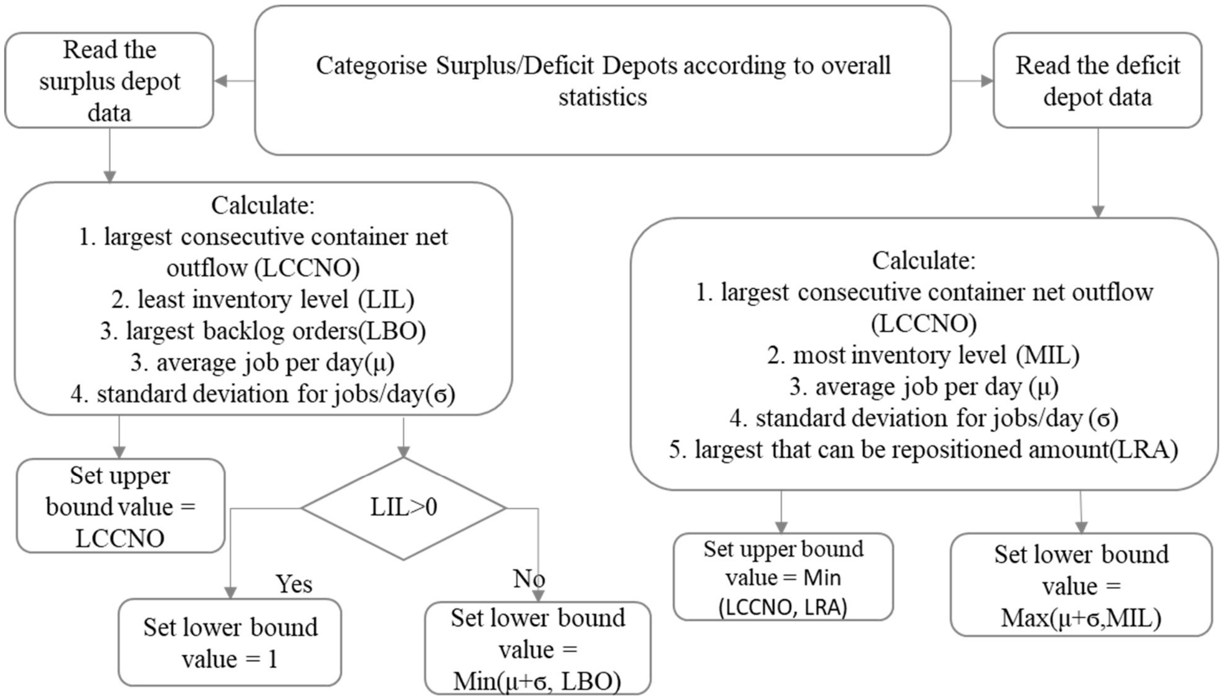
Appendix 2. The Simulation Module in Stage 2



Appendix 3. Process of obtaining LCCNO value



Appendix 4. The Heuristic Search Method



Appendix 5. The flow chart of AGA