Transportation Research Part B 197 (2025) 103236

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Liner fleet deployment and slot allocation problem: A distributionally robust optimization model with joint chance constraints

Tao Zhang [a], Shuaian Wang [b], Xu Xin [a,b,*]

[a] School of Economics and Management, Tongji University, Shanghai, 200092, PR China
[b] Department of Logistics and Maritime Studies, Faculty of Business, The Hong Kong Polytechnic University, Hung Hom, 999077, Hong Kong Special Administrative Region of China

ARTICLE INFO

Keywords:
Liner fleet planning
Slot allocation
Distributionally robust optimization
Uncertain demand
Joint chance constraints
Outer approximation algorithm
Two-stage optimization

ABSTRACT

In this paper, we address the classical liner fleet deployment and slot allocation joint optimization problem in the maritime field with uncertain container transportation demand. We relax the assumption in existing studies that the demand distribution function is known because container transportation demand is deeply affected by the world's economic and political landscape. With the help of advances in distributionally robust optimization theory, we develop a two-stage data-driven robust chance-constrained model. This distribution-free model requires only limited historical demand data as input and jointly optimizes the class (i.e., capacity) and number of liners assigned on each route and the scheme for allocating containers on each leg to maximize the profit (container transportation revenue minus fleet operating costs, voyage costs, and capital costs) of the liner company. The joint chance constraint in the model requires that the transportation demand of the contract shipper be satisfied with a pre-determined probability. We then reformulate the model as a second-order cone programming and design a customized algorithm to explore the global optimal solution based on the outer approximation algorithm framework. This paper can serve as a baseline distribution-free model for solving liner fleet deployment and slot allocation joint optimization problems.

1. Introduction

Maritime transportation is a fundamental aspect of international trade, facilitating the efficient and cost-effective movement of cargos across the globe (Chen et al., 2023; Xiang et al., 2025). Currently, over 80% of cargo is transported by ocean to all parts of the world in a convenient and efficient manner, contributing to the establishment of 'factories of the world' worldwide (Xin et al., 2023). Among the various modes of ocean transportation, container shipping occupies a significant position. Container transportation accounts for more than 70 % of the cargo value (Weerasinghe et al., 2023). Over the last 30 years, container transportation has grown at a rate of 8 % per year, and more than 173.8 million containers were transported by the ocean in 2023 (Weerasinghe et al., 2023). The existence of container transportation facilitates the purchasing of raw materials by companies and contributes to the formation of a

* Corresponding author at: Department of Logistics and Maritime Studies, Faculty of Business, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region of China.
E-mail addresses: taozhang@tongji.edu.cn (T. Zhang), hans.wang@polyu.edu.hk (S. Wang), xinxu@tongji.edu.cn, xu-david.xin@connect.polyu.hk (X. Xin).

https://doi.org/10.1016/j.trb.2025.103236
Received 19 June 2024; Received in revised form 3 March 2025; Accepted 22 April 2025
Available online 10 May 2025

globalized economy (Liu et al., 2021).

In practice, liner companies provide container transportation services by establishing a liner fleet comprising different classes of liners and deploying appropriate number and class of liners to operate on the liner shipping network (i.e., various routes) according to potential transportation demand (Wang et al., 2013). Subsequently, liners make port calls at regular intervals (e.g., once a week or twice a week) on these routes according to pre-designed service frequencies (Meng et al., 2012). This decision-making problem is also referred to as the **L**iner **F**leet **D**eployment **P**roblem (LFDP) in the field of maritime, which is generally carried out once every 3 to 6 months in order to ensure the stability of the service (Wang et al., 2013; Xin et al., 2023). Once the deployment is completed, shippers can access the information on the routes operated by liner companies, including service frequency, freight rates, and liner information, through the official channels provided by liner companies (e.g., the official website; see https://www.oocl.com for an example). Should shippers require the purchase of shipping services, they must contact the relevant liner company (typically via a freight forwarder). The liner company will then determine whether to accept the order based on the number of available slots (i.e., space on board a liner occupied by a container) (Meng et al., 2012), a process also known as *slot allocation*. As liner companies operate multiple routes simultaneously in their networks, the key decision for liner companies is to assign containers (orders) to the limited number of slots available on the liner on the optimal route. This decision problem is also called **S**lot **A**llocation **P**roblem (SAP). There are times when a liner company may choose to utilize different routes to fulfill an order, and this may be accomplished by utilizing slots on liners sailing on multiple routes. In the freight market, some shippers may sign long-term contracts with liner companies, also known as *contract shippers*, owing to their stable and large demand for container transportation. Other shippers, known as *spot shippers*, only have temporary transportation demands and can only passively accept quotes from liner companies (Wang et al., 2020). In light of this, the strategic management departments of liner companies typically prioritize meeting the transportation demands of contract shippers when making SAP decisions, which is considered a *loyal strategy* (Liu and Yang, 2015).

One can notice that in the process of addressing both LFDP and SAP for liner companies, container transportation demand is a critical parameter. However, in contrast to other types of transportation, container liner transportation activities are deeply coupled with the global political and economic environment and are frequently influenced by a number of uncontrollable factors, such as delayed container arrivals at ports due to bad weather (Wang et al., 2013). In practice, fluctuations in container transportation demand are sometimes so dramatic that it is almost impossible for liner companies to estimate them accurately (Meng et al., 2012; Liu et al., 2025). Consequently, since the beginning of the 2010s, scholars have gradually reached a consensus that the uncertain factor of container transportation demand needs to be taken into consideration when handling LFDP, SAP, or both, and a significant number of models considering demand uncertainty have been constructed via stochastic programming (SP) theory, including Wang et al. (2013) and Wang et al. (2020). However, the majority of established studies that consider uncertain demand assume that transportation demand is a random variable with a known distribution function (typically assumed to be a normal distribution function). On the basis of this assumption, several scenarios can be generated, and effective methods, such as sample sampling approximation (SAA), can be used to approximate the SP model by a deterministic problem with a limited number of scenarios. In practice, however, it is challenging to obtain the true distribution of container transportation demand in the context of complex political and economic circumstances (Zhao et al., 2022). This leads to the possibility that the traditional method of using prior knowledge to estimate the parameter distribution and solve a scenario-based SP may have large deviations (Xin et al., 2025a).

Motivated by the abovementioned background, we attempt to develop a two-stage distribution-free robust chance-constrained model to address the joint optimization of LFDP and SAP under demand uncertainty as an extension of Meng et al. (2012). In the first stage (also referred to *here-and-now* stage), the liner company is responsible for determining the liner fleet deployment scheme. Once the decision-making in the first stage is known and the random variables are realized, the liner company optimizes the slot allocation scheme in the second stage (*wait-and-see* stage). In contrast to previous studies, we do not assume that we have the distribution information of demand. Distributionally robust optimization (DRO) theory, which is a tool for effectively addressing uncertainty, allows us to identify the optimal liner deployment scheme and slot allocation scheme that perform the best in the worst-case distribution. This is achieved by requiring that the true distribution function be within an *ambiguity set* that contains possible distribution functions with similar characteristics, such as moment conditions. Although obtaining a precise description of container transportation demand distribution functions can be challenging, liner companies can collect partial statistical information or empirical data. Equipped with recent theories and techniques in the DRO field (e.g., theories in Delage and Ye (2010) and Chen et al. (2010)), this paper establishes a data-driven distributionally robust joint chance-constrained model, where the ambiguity sets rely only on mean and covariance information obtained from historical data. Although ambiguity sets are introduced to the model to characterize transportation demand uncertainty, solving such a model is still computationally intractable. Thus, we use the characteristics of the *conditional value at risk* (CVaR) to convert the original model into another model called the *second-order cone programming model* (SOCPM) and utilize an outer approximation algorithm (OAA) to solve this SOCPM exactly.

Overall, the contributions of this paper are as follows. (1) This paper contributes to the current literature by addressing a distributionally robust joint chance-constrained model for solving the joint optimization of LFDP and SAP. To our knowledge, this is the first paper to extend data-driven DRO theory to the joint optimization of two sub-problems in container liner shipping. (2) Based on the moment and covariance information of historical data, we transform the original distribution-free model into a standard SOCPM to make it computationally tractable by taking advantage of the properties of CVaR. (3) We design an OAA for the SOCPM that can be solved in a targeted manner to achieve an efficient and obtain the global optimal solution in a reasonable time.

The remainder of this paper is organized as follows. In Section 2, we review the relevant literature on LFDP and SAP. Section 3 provides a detailed description of the LFDP and SAP joint optimization problem involved in this paper and defines the concepts of liner shipping networks, transportation demand uncertainty and container transportation paths. This section helps readers better understand the daily operational decisions of liner companies. A data-driven distributionally robust joint chance-constrained model is

developed for solving the joint optimization problem in Section 4. Section 5 reformulates the model into a standard SOCPM, taking advantage of the properties of CVaR. In Section 6, an OAA is designed to achieve the exact solution of the SOCPM. Numerical experiments and sensitivity analyses are carried out in Section 7. Finally, in Section 8, we summarize the conclusions of our paper and propose several potentially interesting future research directions.

Notation. The following notation is used throughout this paper: bold letters are used to represent matrices and vectors. We let $\boldsymbol{\theta} = (\theta_w)$ and $\boldsymbol{\vartheta} = (\vartheta_w)$ represent two random vectors. The notation $[K] := \{1, 2, \cdots, K\}$ $(K \geq 1)$ represents the set containing integers from 1 to K. Finally, we use $\mathscr{F}$ and $\mathscr{G}$ to denote two ambiguity sets whose elements (i.e., distribution functions) are represented as F and G, respectively.

2. Literature review

Currently, the classical liner shipping network design problem is categorized by the academic community into three interrelated and interacting sub-problems (Wang and Meng, 2017). The first subproblem is called LFDP, which requires liner companies to deploy appropriate classes and numbers of liners for each route. This is a tactical-level decision-making problem. In this problem, liner companies may also consider optimizing the speed of their liners, chartering in and out, and optimizing the status of their liners (i.e., whether they are in service or in storage). The second sub-problem is called the service network design problem (SNDP), which is also a tactical-level decision-making problem in which liner companies optimize the port of call and calling sequence of their fleet. The third sub-problem is the SAP at the daily operational level. Liner companies need to make acceptance–rejection decisions on daily orders to accommodate limited liner slots. At first, scholars often studied one of the sub-problems in isolation. Subsequently, with an in-depth understanding of the problem, studies considering the joint optimization of two or even three sub-problems have been proposed. Readers interested in further information are encouraged to consult the following review papers: Christiansen et al. (2020) and Ksciuk et al. (2023). As mentioned before, the problem involved in this paper is the composite problem of LFDP and SAP. Therefore, this section primarily reviews the research on these two sub-problems and composite problems formed from the abovementioned three sub-problems.

2.1. Research on LFDP

Research on LFDP can be traced back to the 1990s. Perakis and Jaramillo (1991) and Jaramillo and Perakis (1991) are considered the earliest studies discussing LFDP. In these two papers, the authors formally defined LFDP. The former constructed a linear programming model, whereas the latter developed an integer linear programming model. In light of the research conducted by these two authors, scholars have gradually constructed models that consider practical factors. These include liner lay-up costs (Powell and Perkins, 1997), liner speed optimization (Gelareh and Meng, 2010), multi-period decisions (Meng and Wang, 2011), and cargo transshipments (Meng and Wang, 2012). In these preliminary studies, the authors typically assumed that container transportation demand is known, which is a relatively ideal assumption given the limitations of optimization theory.

After the 2010s, several scholars have begun to examine the LFDP in light of demand uncertainty. This type of research first originated from Meng and Wang (2010), who developed a chance-constrained programming model and simply assumed that the container transportation demand between two ports was a random variable with known mean-variance. The authors were then able to transform individual chance constraints (ICCs) in the model into several regular constraints by employing the cumulative density function (CDF) of the normal distribution. At this point, the model can be solved efficiently via commercial solvers. In contrast, Wang et al. (2013) considered the same problem but proposed a joint chance-constrained programming model. The authors believed that the premise for using ICCs is that the transportation demands between different port pairs are independent and may not be sufficiently rigorous. Moreover, the theory proposed by Meng and Wang (2010) is not applicable to the case where the random variable (i.e., transportation demand) follows a non-normal distribution. To handle the joint chance constraint, the authors applied a scenario-based stochastic programming solution method, namely, SAA, which was first proposed by Luedtke and Ahmed (2008). On the basis of the assumption that the demand distribution function is known (specifically, a log-normal distribution), 3000 and 10,000 independent Monte Carlo samples were generated in different experimental scenarios. This scenario-based method reformulates the joint chance constraint into a regular constraint (in each scenario).

In the aforementioned studies, the authors typically impose a strong assumption to address demand uncertainty; that is, the decision maker has information associated with the distribution function (e.g., the normal distribution function). This assumption is often justified by noting that (log)normal distributions are commonly used in problems involving demand uncertainty. However, in reality, the true distribution of random variables is usually difficult to obtain. In particular, the maritime industry is highly connected to the world economy, and demand fluctuations in this field are often highly related to the political and economic conditions of various countries around the world (Chen et al., 2023). This has prompted the academic community to begin attempting to design distribution-free models. For example, in 2014, Ng (2014) proposed a distribution-free approach with ICCs, requiring only the mean, variance and upper bound of demand as model inputs. These ICCs were subsequently transformed into tractable constraints via the well-known Markov inequality. In Ng (2015), the ICCs in Ng (2014) were replaced by joint chance constraints. The research objective of this type of research is to use probability theory to transform the model into a robust optimization model, which tries to optimize the worst-case scenario obtained by collecting historical data. However, the bound provided by the Markov inequality is usually too loose, and robust optimization is often criticized for obtaining overly conservative solutions (Xin et al., 2025b). Therefore, Ng and Lin (2018) introduced the concept of conditional demand into the model to prevent the previous model from being overly conservative. Very recently, with the development of DRO, scholars have begun to construct DRO models to solve LFDP. Zhao et al. (2022) represent a

pioneering study in this field. The authors constructed two robust joint chance constrained models on the basis of the model established by Ng (2015), assuming that the container transportation demand distribution function is contained in an ambiguity set. However, all these authors designed random variables on a route basis and ignored the voyage cost (which can be as high as 60 % of the voyage cost; see examples in Stopford (2008)) when constructing models. This simplification made the historical data after route changes no longer applicable, and the resulting calculation results could not be realistically applied. For more detailed information related to this type of research, see Wang and Meng (2017) and Ksciuk et al. (2023).

2.2. Research on SAP

SAP research commenced in the 1990s and was often combined with the freight rate pricing problem, which is regarded as a special type of revenue management problem. Maragos (1994) pioneered the joint optimization problem of slot allocation and pricing in the field of liner shipping. At the beginning of the 21st century, Ting and Tzeng (2004) attempted to optimize the container transportation volume between each port pair and proposed a mixed-integer programming (MIP) model. The model established by Feng and Chang (2010) considers allocating both laden and empty containers to specific routes. Zurheide and Fischer (2012) proposed a new SAP by dividing orders into urgent and non-urgent orders and taking into account the transshipment of containers. These two authors further considered adjusting the slot allocation strategy from the booking limit to the bid price (Zurheide and Fischer, 2015). The transportation demand in the maritime industry often exhibits seasonal fluctuations. In light of this, Wang et al. (2015) constructed an MIP model and developed a customized branch-and-bound (B&B) algorithm. In summary, early studies all assumed that container transportation demand is fixed and known. Based on the known demand, the authors considered various strategies and the operating behavior of liner companies to allocate slots.

As with the research on LFDP, research on SAPs has also gradually shifted from deterministic problems to uncertain problems in recent years. For example, Wang et al. (2020) distinguished shippers as contract shippers and spot shippers, and two types of slot optimization models were proposed based on different marketing psychology (i.e., loyal strategy and expansive strategy). The authors believed that a liner company only needs to meet the demands of the contract shipper among the two types of shippers, which serves as the foundation for the assumptions of our paper. Considering the uncertainty of the liner staying time at ports and container transportation demand, the authors constructed two SP models to optimize the slot allocation scheme and freight rate pricing scheme. Assuming that each random variable obeys a uniform distribution with known parameters, a customized SAA embedded in the reformulation linearization technique was designed to solve the model. Wang et al. (2021) subsequently established a two-stage stochastic MIP model to address the SAP with a freight rate pricing problem while considering demand uncertainty and the empty container repositioning factor. To address uncertainty, the authors assumed that the distribution function of random demand between each port pair is known (i.e., a normal distribution) and transformed the model into a deterministic one by applying the SAA method. To address this problem, the authors performed dual decomposition on the original model and designed a Lagrangian relaxation algorithm. Very recently, Liang et al. (2023) designed several slot allocation policies considering two types of shippers (i.e., contract shippers and spot shippers). Like the loyal strategy proposed by Wang et al. (2020), the authors also assumed that priority should be given to ensuring that the demands of the contract shipper are met. Under the assumption that the distribution function of the transportation demand of the two types of shippers between each port pair is known (i.e., a truncated normal distribution), the authors constructed a scenario-based SP model. Several policies have also been developed to allocate slots, including the first-come-first-serve policy, minimum quantity commitment policy, and threshold-based perturbation approach. Based on 5,000 scenarios generated from this known distribution, the authors conducted numerical experiments for different policies. In summary, most of the current studies that consider random factors in SAPs focus on transportation demand uncertainty. As the research progresses, additional factors (e.g., differentiating between different types of shippers, empty container repositioning, demurrage charges for stranded containers, liner staying time at the port) are incorporated into the model framework. To make the stochastic model tractable, the authors without exception assumed that the demand distribution function is known (e.g., a normal distribution or uniform distribution). However, as Ksciuk et al. (2023) noted, transportation demand uncertainty is caused by changes in economic policy, and combined with the lack of historical data, accurately estimating its true distribution is extremely difficult.

2.3. Research on the joint optimization of LFDP, SNDP and SAP

As previously stated, a reasonable *liner fleet planning scheme* is necessary for liner companies to reduce costs. However, the *network design scheme* must be known before fleet planning can be carried out. Moreover, once the fleet planning scheme is determined, an excellent *slot allocation scheme* can further increase the income of the liner company. These three types of decisions are interconnected, and as a result, joint optimization studies of LFDP, SNDP and SAP have sprung up. Mourão et al. (2002) is regarded as the earliest study on collaborative optimization of LFDP and SAP, as the authors considered container transshipment operations in an LFDP. Container transshipment refers to the transportation of a container via one or more routes, with the associated cost varying according to the loading and unloading activities involved. Consequently, the liner company must allocate containers to the limited slots of the liner operating on a single route or liners on multiple routes (in the case of transshipment). However, the authors considered a liner shipping network comprising only one pair of ports and two routes, which is a highly simplified representation of reality. Some scholars have also considered combining LFDP, SNDP and SAP (with regard to container transshipment operations). For example, Agarwal and Ergun (2008) constructed an MIP model for this joint optimization problem and designed a meta-heuristic, a two-stage Benders decomposition algorithm, and a column generation-based algorithm. Papers that consider similar factors also include Gao et al. (2022) and Gelareh et al. (2010). Though the model frameworks incorporate more realistic factors, such as liner alliances, shippers' choice inertia,

and carbon taxes, the authors invariably assume that container transportation demand is known in order to solve these complex models.

3. Problem description

3.1. Liner shipping network

We consider that a liner company operates a liner shipping network that includes several pre-determined routes, and the set composed of them is denoted as $\mathscr{R}$. Each route calls several ports in sequence. We let $\mathscr{P}$ denote the set composed of ports of call and use r and p to represent the elements in $\mathscr{R}$ and $\mathscr{P}$, respectively. A route is usually a loop structure, which we can define on the basis of the calling sequence of the liner. For example, we can set:

$$p_r^1 \rightarrow p_r^2 \rightarrow p_r^3 \rightarrow \cdots \rightarrow p_r^{k_r} \rightarrow p_r^1$$

where $p_r^i \in \mathscr{P}$ represents the i^{th} port of call on route $r \in \mathscr{R}$ and k_r is the number of ports of call on route $r \in \mathscr{R}$. When a liner departs from a certain port (e.g., p_r^1), sails around the loop route and returns to this port, we say that the liner has completed a *voyage*. The part of the route from port p_r^i to port p_r^{i+1} is also called *leg* i of route $r \in \mathscr{R}$. The set composed of legs on route r is denoted as $\mathscr{L}_r$.

Fig. 1 shows the shape of a route in the real world. The liner company first selects 5 ports (i.e., Port 1 to Port 4 and Port 6) from 6 candidate ports to form a route. One thing worth noting is that some ports may be called more than once in a voyage (e.g., Port 2 in Fig. 1), which are usually called *butterfly nodes*. This is a significant difference between container liner shipping and land transportation, such as bus transportation. In view of this, further distinctions between different parts of a route are usually made. Typically, a portion of a route is called a forward sub-route and backward sub-route. For example, Chinese liner companies usually split their routes according to sailing from China to foreign countries and returning from foreign countries to China. The routes operated in the real world are very similar to those in Fig. 1.

To facilitate readers' understanding, we obtained Fig. 2 from the official website of the Orient Overseas Container Line (OOCL; see https://www.oocl.com). The sub-route formed by the blue arrow is a forward sub-route, and the sub-route formed by the yellow arrow is a backward sub-route. It takes 77 days for a liner to complete a voyage. For each port, the liner company guarantees a liner call every week.

3.2. Definition of container transportation demand uncertainty

We use $\mathscr{W} = \{(o,d) | o,d \in \mathscr{P}\}$ to represent the set of origin-destination (O-D) pairs of transportation demand whose elements are denoted by $w = (o, d)$. In practice, transportation demand is generated by two types of shippers (Liang et al., 2023; Wang et al., 2020). One type of shipper is called a *contract shipper*. He/she typically has regular and substantial transportation demands. Therefore, liner companies often aim to secure the transportation demands of these shippers (referred to as 'contract demand' therein), as they can generate stable income (Liang et al., 2023). We consider that the contract demand is often generated on a weekly basis, so we let random variable ϑ_w be the weekly contract demand expressed in twenty-foot equivalent units (TEUs) between O-D pair w, i.e., the number of containers generated by the contract shippers to transport between O-D pair w each week.

Another type of shipper is called a *spot shipper*. The transportation demands of such shippers (known as 'spot demand' therein) are usually urgent, and they have no right to bargain with liner companies. Therefore, liner companies usually make order acceptance–rejection decisions on the basis of slot utilization, without the need to guarantee that the transportation demands of each shipper are met. To account for this type of transportation demand, we use a random variable θ_w to represent the total transportation demand (including both contract demand and spot demand) for O-D pair w during the planning period. Note that in this paper, we need to collect only historical information on the total transportation demand θ_w during the entire planning period instead of the weekly spot demand. This is because we do not impose constraints on weekly spot demand.

3.3. Definition of container transportation paths

In the daily operation of a liner company, it needs to choose a container transportation path for container delivery on the O-D pair w. We use set $\mathscr{K}^w$ to represent the set of container transportation paths between the O-D pair w. In other words, for each O-D pair, there

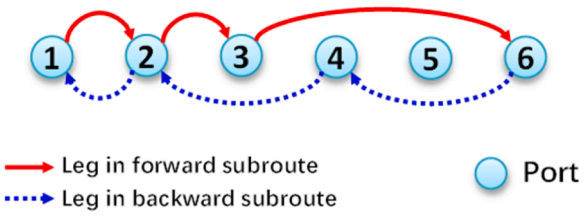


Fig. 1. A toy shipping route containing 6 candidate ports.

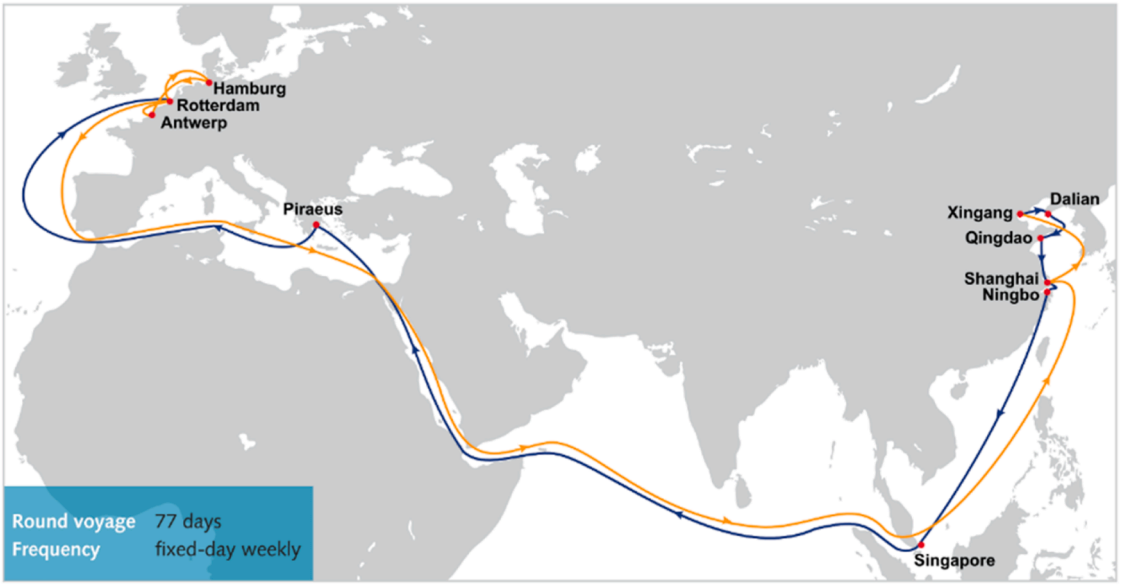


Fig. 2. A real-world route operating by OOCL.

are several corresponding transportation paths. Here, one container transportation path can be part of a route or a combination of several parts of multiple routes. Fig. 3 illustrates this definition in detail. To transport one container from Port 2 to Port 4 in Fig. 3, the liner company can choose from two container transportation paths:

$$h_1^{\text{Port 2–Port 4}} = p_3^1(\text{Port 2}) \xrightarrow{\text{Ship route 3}} p_3^2(\text{Port 4}),$$

and

$$h_2^{\text{Port 2–Port 4}} = p_1^1(\text{Port 2}) \xrightarrow{\text{Ship route 1}} p_1^2(\text{Port 3}) \mapsto p_2^2(\text{Port 3}) \xrightarrow{\text{Ship route 2}} p_2^3(\text{Port 4}).$$

If the liner company chooses the first transportation path $h_1^{\text{Port2–Port4}}$, this container will be loaded on a liner at port 2 and will be unloaded after the liner sails to port 4. In contrast, when the liner company selects $h_2^{\text{Port2–Port4}}$, this container is loaded at port 2 onto a liner deployed on route 1 and unloaded at port 3, and later, this container is reloaded (also referred to as *transshipped*) at port 3 onto a liner deployed on route 3 and finally unloaded at port 4. We can observe that the container handling costs incurred in transporting this container are different when the liner company chooses different container transportation paths due to the presence of container transshipment. In this paper, we use c_{h^w} to denote the handling cost incurred for transporting one TEU on path $h^w \in \mathscr{H}^w$.

3.4. LFDP and SAP

In our joint optimization problem, one of the decisions that the liner company needs to make is liner fleet deployment (to solve the LFDP). We let $\mathscr{S}$ be the set of liner classes and let CAP_s represent the capacity of an s-class liner (in terms of TEUs). The liner company must make decisions regarding the class and number of liners to deploy to each route, but not all liners it own need to be used. In practice, the liner company may use either its own liners or charters in liners on the charter market. Meanwhile, the liner company also

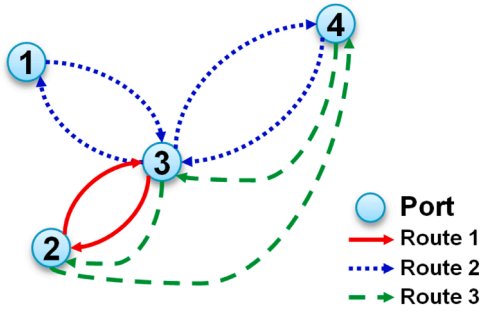


Fig. 3. A toy liner shipping network with 4 ports of call and 3 routes.

has the option of charting out liners to other liner companies. The charter market usually offers chartering options, including time charters, voyage charters and bareboat charters. The last type of charter is the simplest. The reason is that in this type of charter, the charterer is responsible for the full management of the liner, bearing all associated costs (with the exception of capital costs, taxes, and depreciation). Consequently, the shipowner is relieved of any financial obligation and is solely responsible for collecting the agreed-upon rent. For the sake of simplicity, we assume that the liner company is only able to charter liners through a bareboat charter. The number of class $s \in \mathscr{S}$ liners chartered in and out are denoted as n_s^{IN} and n_s^{OUT}, respectively. The charter rate (including maintenance costs and insurance) for a class $s \in \mathscr{S}$ liner during the entire planning horizon is c_s^{IN}. The income that can be earned from chartering out a class $s \in \mathscr{S}$ liner is c_s^{OUT}. Because the charter rate also includes maintenance costs and insurance, we have $c_s^{\mathrm{OUT}} < c_s^{\mathrm{IN}}$. We let NO_s^{MAX} and NC_s^{MAX} represent the number of class $s \in \mathscr{S}$ liners owned by the liner company and available in the market, respectively. Given these candidate liners, the company can arrange the appropriate class (number) of liners on the appropriate route. Unlike tramp shipping, liner shipping needs to provide regular service (e.g., weekly service) to add consistency for customers and usually has fixed timetables (i.e., arrival/departure times at each port of call). Therefore, months before the actual journey begins, shippers can browse timetables and freight rates published by liner companies on the internet. Considering this operational characteristic, we assume that the liner company assigns the same class liners to one route and that different classes of liners can be deployed on different routes. Each route must maintain a calling frequency of at least once a week. The optimization objective of LSDP is to establish a liner deployment scheme that minimizes costs and is subject to certain constraints (e.g., service frequency constraints).

After completing the assignment of the fleet on each route, the slots (capacity) available per week for each route are known. At this time, the liner company must carefully split the container transportation demands between each O-D pair to each container transportation path to solve the SAP. We let q_{h^w} denote the proportion of containers transported on container transportation path $h^w \in \mathscr{H}^w$ to total demand between O-D pair $w \in \mathscr{W}$. The optimization objective of the SAP is to create a container transportation scheme, say q_{h^w}, which assigns both ϑ_w and θ_w to each container transportation path. This scheme is able to satisfy the uncertain transportation demand ϑ_w with a pre-determined probability and maximize the expected profit of the liner company (associated with uncertain transportation demand θ_w) over the planning horizon.

Before concluding this section, we list the assumptions used in this paper. (1) The liner company endeavors to meet the weekly container transportation demands of contract shippers. (2) We consider only bareboat chartering as one type of charter. (3) The queuing of liners for berthing at each port is not considered. (4) The relocation of empty containers and inland transportation are not considered; it is assumed that there are always sufficient empty containers for cargo. (5) For all O-D pairs, all corresponding container transportation paths can meet the time limit requirements of the shipper. Moreover, the notations frequently used in this paper are listed below.

Sets	
$\mathscr{R}$	Set of routes, where $r \in \mathscr{R}$
$\mathscr{P}$	Set of ports of call, where $p \in \mathscr{P}$
$\mathscr{L}_r$	Set of legs on route $r \in \mathscr{R}$, where $l \in \mathscr{L}_r$
$\mathscr{S}$	Set of liner classes, where $s \in \mathscr{S}$
$\mathscr{W}$	Set of original-destination (O-D) pairs with container transportation demand, where $w = (o,d) \in \mathscr{W}$
$\mathscr{H}^w$	Set of container transportation paths associated with O-D pair $w \in \mathscr{W}$, whose elements are denoted as h^w
$\mathscr{T}$	The ambiguity set (also known as *distributional set*) used to describe the distribution function F
$\mathscr{G}$	The ambiguity set used to describe the distribution function G
$\mathbb{Z}_+$	Set of non-negative integers
Parameters	
T	The planning horizon (Unit: day) of the liner company
c_s^{OP}	The operating cost (Unit: 10^4 \$/day) of a class s liner
c_{sr}^{VOL}	The voyage cost (Unit: 10^4 \$/voyage) of a class s liner on route $r \in \mathscr{R}$
c_s^{IN}	The charter cost paid (Unit: 10^4 \$/day) for chartering in a class s liner
c_s^{OUT}	The revenue received (Unit: 10^4 \$/day) for chartering out a class s liner
NO_s^{MAX}	Number of class s liner company owned
NC_s^{MAX}	Number of class s liner in the market
t_{sr}	The voyage time (Unit: day) of a class s liner on route $r \in \mathscr{R}$
CAP_s	The capacity (Unit: 10^2 TEU) of a class s liner
$\delta_{h^w}^{lr}$	A binary parameter, which equals 1 if path $h^w \in \mathscr{H}^w$ contains leg $l \in \mathscr{L}_r$ in the shipping network, and 0 otherwise
r_w	The freight rate (Unit: 10^4 \$/TEU) for transporting one container between O-D pair $w \in \mathscr{W}$
c_{h^w}	The handling cost (Unit: 10^4 \$/TEU) incurred for transporting one container on path $h^w \in \mathscr{H}^w$
f_r^{MIN}	The minimum calling frequency required on route $r \in \mathscr{R}$ (Unit: times/week)
Random variables	
θ_w	Random container transportation demand (Unit: 10^2 TEU) from both contract shippers and spot shippers of the planning horizon between O-D pair$w \in \mathscr{W}$, where $\boldsymbol{\theta} = (\theta_w)$
ϑ_w	Random container transportation demand (Unit: 10^2 TEU) from contract shippers of each week between O-D pair$w \in \mathscr{W}$, where $\boldsymbol{\vartheta} = (\vartheta_w)$
Decision variables	
n_{sr}^{TOTAL}	Number of class $s \in \mathscr{S}$ liners deployed on route $r \in \mathscr{R}$
n_s^{IN}	Number of class $s \in \mathscr{S}$ liners chartered in
n_s^{OUT}	Number of class $s \in \mathscr{S}$ liners chartered out
n_{sr}^{VOL}	Number of voyages completed by class $s \in \mathscr{S}$ liner deployed on route $r \in \mathscr{R}$
f_r^{call}	The service frequency on route $r \in \mathscr{R}$ (Unit: times/week)

(*continued on next page*)

(*continued*)

$\boldsymbol{n}$	A vector representing the decision variables involved in the first stage model, where $\boldsymbol{n} = \left(x_{sr}, n_{sr}^{\mathrm{TOTAL}}, n_s^{\mathrm{IN}}, n_s^{\mathrm{OUT}}, n_{sr}^{\mathrm{VOL}}, f_r^{\mathrm{call}}\right)^{\mathrm{T}}$
x_{sr}	A binary variable, which equals 1 if class $s \in \mathscr{S}$ liner is deployed on route $r \in \mathscr{R}$, and 0 otherwise
q_{h^w}	The proportion of total transportation demand between O-D pair w transported on path $h^w \in \mathscr{H}^w$

4. Model establishment

Our original model **P** is a two-stage distributionally robust optimization model (2-DRO). In model **P**, each stage involves one type of decision. In other words, two types of decisions are made in a sequential fashion. In the first-stage (also referred to *here-and-now* stage), a decision related to liner fleet deployment, i.e., $\boldsymbol{n} = \left(x_{sr}, f_r^{\mathrm{call}}, n_{sr}^{\mathrm{TOTAL}}, n_s^{\mathrm{IN}}, n_s^{\mathrm{OUT}}, n_{sr}^{\mathrm{VOL}}\right)^{\mathrm{T}}$, is made before the uncertainty $\boldsymbol{\theta}$, an $|\mathscr{W}|$-dimensional random vector, is realized. After $\boldsymbol{\theta}$ realizes as a fixed value, the decision maker moves to the second-stage (also referred to as the *wait-and-see* stage) and determines q_{h^w} (related to slot allocation) by solving an optimization problem. In the following, we denote the second-stage cost by the function $g(\boldsymbol{n}, \boldsymbol{\theta})$.

First-stage of P:

$$\max obj = \left\{ \begin{array}{c} -T\sum_{r\in\mathscr{R}}\sum_{s\in\mathscr{S}} c_s^{\mathrm{OP}} n_{sr}^{\mathrm{TOTAL}} - \sum_{r\in\mathscr{R}}\sum_{s\in\mathscr{S}} c_{sr}^{\mathrm{VOL}} n_{sr}^{\mathrm{VOL}} - T\sum_{s\in\mathscr{S}} c_s^{\mathrm{IN}} n_s^{\mathrm{IN}} \\ +T\sum_{s\in\mathscr{S}} c_s^{\mathrm{OUT}} n_s^{\mathrm{OUT}} + \inf_{F\in\mathscr{F}, G\in\mathscr{G}} \mathbb{E}_{F,G}[g(\boldsymbol{n}, \boldsymbol{\theta})] \end{array} \right\} \tag{1}$$

$$s.t. \sum_{s\in\mathscr{S}} x_{sr} = 1 \qquad \forall r \in \mathscr{R} \tag{2}$$

$$n_{sr}^{\mathrm{TOTAL}} \le \left(NO_s^{\mathrm{MAX}} + NC_s^{\mathrm{MAX}}\right) x_{sr} \qquad \forall s \in \mathscr{S}, r \in \mathscr{R} \tag{3}$$

$$n_s^{\mathrm{IN}} \le NC_s^{\mathrm{MAX}} \qquad \forall s \in \mathscr{S} \tag{4}$$

$$\sum_{r\in\mathscr{R}} n_{sr}^{\mathrm{TOTAL}} = NO_s^{\mathrm{MAX}} + n_s^{\mathrm{IN}} - n_s^{\mathrm{OUT}} \qquad \forall s \in \mathscr{S} \tag{5}$$

$$n_{sr}^{\mathrm{VOL}} \le n_{sr}^{\mathrm{TOTAL}} \times \left\lfloor \frac{T}{t_{sr}} \right\rfloor \qquad \forall s \in \mathscr{S}, r \in \mathscr{R} \tag{6}$$

$$f_r^{\mathrm{call}} = \sum_{s\in\mathscr{S}} n_{sr}^{\mathrm{VOL}} \Big/ \left\lceil \frac{T}{7} \right\rceil \qquad \forall r \in \mathscr{R} \tag{7}$$

$$f_r^{\mathrm{call}} \ge f_r^{\mathrm{MIN}} \qquad \forall r \in \mathscr{R} \tag{8}$$

$$f_r^{\mathrm{call}} \in \mathbb{Z}_+ \qquad \forall r \in \mathscr{R} \tag{9}$$

$$n_s^{\mathrm{IN}}, n_s^{\mathrm{OUT}} \in \mathbb{Z}_+ \qquad \forall s \in \mathscr{S} \tag{10}$$

$$n_{sr}^{\mathrm{TOTAL}}, n_{sr}^{\mathrm{VOL}} \in Z_+ \qquad \forall s \in \mathscr{S}, r \in \mathscr{R} \tag{11}$$

$$x_{sr} \in \{0, 1\} \qquad \forall s \in \mathscr{S}, r \in \mathscr{R} \tag{12}$$

where vector $\boldsymbol{n} = \left(x_{sr}, f_r^{\mathrm{call}}, n_{sr}^{\mathrm{TOTAL}}, n_s^{\mathrm{IN}}, n_s^{\mathrm{OUT}}, n_{sr}^{\mathrm{VOL}}\right)^{\mathrm{T}}$ contains all the decision variables associated with the liner deployment scheme; $\mathbb{Z}_+$ denotes the set of non-negative integers; $\lfloor a \rfloor$ represents the maximum integer not greater than a; and $\lceil a \rceil$ is the minimum integer greater than or equal to a.

Objective function (1) calculates the net profit of the liner company during the planning period, which is obtained by subtracting its cost from its revenue. Here, $\inf_{F\in\mathscr{F}, G\in\mathscr{G}} \mathbb{E}_{F,G}[g(\boldsymbol{n}, \boldsymbol{\theta})]$ is the worst-case expected wait-and-see revenue. Constraint (2) requires that only one class of liners can be deployed on each route. Constraint (3) gives the relationship between decision variables n_{sr}^{TOTAL} and x_{sr}. This constraint requires that the number of liners deployed cannot exceed the sum of the number of company-owned liners and the number of liners available in the market. Constraint (4) requires that the number of liners chartered by the company in the market cannot exceed the prescribed upper limit. Constraint (5) gives the relationship between n_{sr}^{TOTAL}, n_s^{IN} and n_s^{OUT}. Constraint (6) imposes that the number of voyages completed by s class liners on route r cannot exceed the maximum number of voyages they can complete. Constraint (7) gives the relationship between the number of completed voyages and the calling frequency for each route. Constraint (8) imposes that the actual calling frequency of each route meets the established frequency lower limit. Constraints (9) to (11) are integer constraints on the decision variables. Finally, constraint (12) requires x_{sr} is a binary decision variable.

Given a first-stage decision, i.e., $\boldsymbol{n} = \left(x_{sr}, f_r^{\text{call}}, n_{sr}^{\text{TOTAL}}, n_s^{\text{IN}}, n_s^{\text{OUT}}, n_{sr}^{\text{VOL}}\right)^{\text{T}}$, and realization of $\boldsymbol{\theta}$, the second-stage cost function $g(\boldsymbol{n}, \boldsymbol{\theta})$ can be established as:

Second-stage of P:

$$g(\mathbf{n}, \boldsymbol{\theta}) = \max \sum_{w \in \mathscr{W}} \theta_w \sum_{h^w \in \mathscr{H}^w} (r_w - c_{h^w}) q_{h^w} \tag{13}$$

$$s.t. \sum_{h^w \in \mathscr{H}^w} q_{h^w} = 1 \qquad \forall w \in \mathscr{W} \tag{14}$$

$$\mathbb{P}_G\left(\sum_{s \in \mathscr{S}} n_{sr}^{\text{VOL}} CAP_s \Big/ \left\lceil \frac{T}{7} \right\rceil \geq \sum_{w \in \mathscr{W}} \vartheta_w \sum_{h^w \in \mathscr{H}^w} q_{h^w} \delta_{h^w}^{lr}, \qquad \forall l \in \mathscr{L}_r, r \in \mathscr{R}\right) \geq \alpha \qquad \forall G \in \mathscr{G} \tag{15}$$

$$q_{h^w} \geq 0 \qquad \forall h^w \in \mathscr{H}^w, w \in \mathscr{W}. \tag{16}$$

Function (13) calculates the *profit* earned by the liner company from transporting containers during the planning period. Constraint (14) requires that for all O-D pairs, all container transportation demands generated (both contract demand and spot demand) need to be allocated to each transportation route every week. Constraint (15) is a joint chance constraint. This constraint means that the liner company's deployed capacity (i.e., slots) on all legs can meet the weekly contract demand with the probability not less than α, where α represents the service level set by the liner company. Constraint (16) is a non-negative constraint on variable q_{h^w}.

5. Model reformation

5.1. Objective function reformulation

Function (13) indicates that the objective value *obj* is linear in random vector $\boldsymbol{\theta}$ and that the optimal solution of the second-stage model is influenced solely by vector $\boldsymbol{\theta}$ and two parameters (i.e., freight rate r_w and container handling cost c_{h^w}) when the vector $\boldsymbol{n}$ is known. Therefore, $\mathbb{E}_{F,G}[g(\boldsymbol{n}, \boldsymbol{\theta})]$ is determined only by the random vector $\boldsymbol{\theta}$. To describe the distribution function F of $\boldsymbol{\theta}$, we use the first moment information to establish the ambiguity set $\mathscr{F}$. In this paper, we construct $\mathscr{F}$ based on an ellipsoid set with estimated mean-covariance information (i.e., $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma} \succeq 0$). Specifically, we set

$$\mathscr{F} := \left\{F : (\mathbb{E}_F[\boldsymbol{\theta}] - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbb{E}_F[\boldsymbol{\theta}] - \boldsymbol{\mu}) \leq \epsilon^2\right\} \tag{17}$$

where ϵ is a pre-determined positive parameter that can reflect the size of set $\mathscr{F}$.

The observation of set $\mathscr{F}$ reveals that the value of parameter ϵ is crucial. Its determination requires an accurate estimation of the first moment of $\boldsymbol{\theta}$. To this end, we apply the data-driven approach proposed in Section 3 of Delage and Ye (2010), which allows for a reasonable estimation of ϵ based on several independent samples. We set the mean of the random vector $\boldsymbol{\theta}$ to be $\widehat{\mu}$, and the covariance matrix to be $\widehat{\Sigma}$, respectively, and the known M samples are denoted as $\{\boldsymbol{\theta}^i\}_{i=1}^M$. If there are two parameters, say $R \geq 0$ and $\beta > 0$, such that $\mathbb{P}\left\{(\boldsymbol{\theta} - \widehat{\mu})^{\text{T}} \widehat{\Sigma}^{-1} (\boldsymbol{\theta} - \widehat{\mu}) \leq R^2\right\} = 1$, then the following inequality holds with a probability greater than $1 - \beta$:

$$(\boldsymbol{\mu}_0 - \widehat{\mu})^T \widehat{\Sigma}^{-1} (\boldsymbol{\mu}_0 - \widehat{\mu}) \leq \eta(\beta) \tag{18}$$

where $\boldsymbol{\mu}_0 = 1/M \sum_{i=1}^M \boldsymbol{\theta}^i$, $\eta(\beta) = \left(R^2/M\right)\left[2 + \sqrt{2\ln(1/\beta)}\right]^2$.

When the above sample information is known, we use the following **Proposition 1** to reformulate the two-stage model **P**.

Proposition 1. Model **P** can be reformulated as:

$$\max obj'_1 = \left\{ \begin{array}{c} -T \sum_{r \in \mathscr{R}} \sum_{s \in \mathscr{S}} c_s^{\text{OP}} n_{sr}^{\text{TOTAL}} - \sum_{r \in \mathscr{R}} \sum_{s \in \mathscr{S}} c_{sr}^{\text{VOL}} n_{sr}^{\text{VOL}} - T \sum_{s \in \mathscr{S}} c_s^{\text{IN}} n_s^{\text{IN}} \\ + T \sum_{s \in \mathscr{S}} c_s^{\text{OUT}} n_s^{\text{OUT}} + \boldsymbol{\mu}^{\text{T}} \boldsymbol{q} - \epsilon \sqrt{\boldsymbol{q}^{\text{T}} \boldsymbol{\Sigma} \boldsymbol{q}} \end{array} \right\} \tag{19}$$

s.t. (2), (3), (4), (5), (6), (7), (8), (9), (10), (11), (12), (14), (15), (16),

$$q_w = \sum_{h^w \in \mathscr{H}^w} (r_w - c_{h^w}) q_{h^w} \qquad \forall w \in \mathscr{W} \tag{20}$$

$$q_w \geq 0 \qquad \forall w \in \mathscr{W} \tag{21}$$

where $\| \cdot \|_n$ denotes the n-norm and $\boldsymbol{q} = (q_w)$ is a non-negative auxiliary vector.

Proof. To begin with, we reformulate the inf part of (1). As the random vector $\boldsymbol{\vartheta} = (\vartheta_w)$ does not affect the objective function (1), we omit the subscript G, which denotes the distribution function of $\boldsymbol{\vartheta}$ for clarity. Moreover, we introduce an auxiliary variable, i.e., $q_w =$

$\sum_{h^w\in\mathscr{H}^w}(r_w - c_{h^w})q_{h^w}$. With this simplification, we have

$g(\boldsymbol{n},\boldsymbol{\theta}) = \max\sum_{w\in\mathscr{W}}\theta_w\sum_{h^w\in\mathscr{H}^w}(r_w - c_{h^w})q_{h^w} = \max \boldsymbol{q}^{\mathrm{T}}\boldsymbol{\theta}$, and we can reformulate $\mathbb{E}_F[g(\boldsymbol{n},\boldsymbol{\theta})]$ in objective function (1) as

$$\mathbb{E}_F[g(\boldsymbol{n},\boldsymbol{\theta})] = (\boldsymbol{q}^*)^{\mathrm{T}}\mathbb{E}_F(\boldsymbol{\theta}) = \max_{\boldsymbol{q}\in\Omega(\boldsymbol{n})}\boldsymbol{q}^{\mathrm{T}}\mathbb{E}_F(\boldsymbol{\theta}) \tag{22}$$

where set $\Omega(\boldsymbol{n})$ denotes the feasible region of the second-stage model, and vector $\boldsymbol{q}^*$ represents an optimal solution of the second-stage model. From the minimax inequality, we have

$$\inf_{F\in\mathscr{F}}\mathbb{E}_F[g(\boldsymbol{n},\boldsymbol{q})] = \inf_{F\in\mathscr{F}}\max_{\boldsymbol{q}\in\Omega(\boldsymbol{n})}\boldsymbol{q}^{\mathrm{T}}\mathbb{E}_F(\boldsymbol{\theta}) = \min_{\mathbb{E}_F(\boldsymbol{\theta})\in\Pi}\max_{\boldsymbol{q}\in\Omega(\boldsymbol{n})}\boldsymbol{q}^{\mathrm{T}}\mathbb{E}_F(\boldsymbol{\theta}) \geq \max_{\boldsymbol{q}\in\Omega(\boldsymbol{n})}\min_{\mathbb{E}_F(\boldsymbol{\theta})\in\Pi}\boldsymbol{q}^{\mathrm{T}}\mathbb{E}_F(\boldsymbol{\theta})$$

where $\Pi = \{\chi\colon (\chi-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\chi-\boldsymbol{\mu}) \leq \epsilon^2\}$. At the same time, we can obtain

$$\min_{\mathbb{E}_F(\boldsymbol{\theta})\in\Pi}\max_{\boldsymbol{q}\in\Omega(\boldsymbol{n})}\boldsymbol{q}^{\mathrm{T}}\mathbb{E}_F(\boldsymbol{\theta}) = \min_{\mathbb{E}_F(\boldsymbol{\theta})\in\Pi}(\boldsymbol{q}^*)^{\mathrm{T}}\mathbb{E}_F(\boldsymbol{\theta}) \leq \max_{\boldsymbol{q}\in\Omega(\boldsymbol{n})}\min_{\mathbb{E}_F(\boldsymbol{\theta})\in\Pi}\boldsymbol{q}^{\mathrm{T}}\mathbb{E}_F(\boldsymbol{\theta}).$$

The first equality holds because $\boldsymbol{q}^*$ is an optimal solution. Meanwhile, since $\boldsymbol{q}^*\in\Omega(\boldsymbol{n})$, the second inequality is straightforward. Considering the mean-covariance information of the random vector $\boldsymbol{\theta}$ as well as the ambiguity set $\mathscr{F}$ defined as (17), we can conclude that

$$\min_{\mathbb{E}_F[\boldsymbol{\theta}]:(\mathbb{E}_F[\boldsymbol{\theta}]-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbb{E}_F[\boldsymbol{\theta}]-\boldsymbol{\mu})\leq\epsilon^2}\boldsymbol{q}^{\mathrm{T}}\mathbb{E}_F(\boldsymbol{\theta}) = \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{q} - \epsilon\sqrt{\boldsymbol{q}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{q}}.$$

At this time, objective function (1) can be reformulated as:

$$\max obj_1 = \left\{\begin{array}{c} -T\sum_{r\in\mathscr{R}}\sum_{s\in\mathscr{S}}c_s^{\mathrm{OP}}n_{sr}^{\mathrm{TOTAL}} - \sum_{r\in\mathscr{R}}\sum_{s\in\mathscr{S}}c_{sr}^{\mathrm{VOL}}n_{sr}^{\mathrm{VOL}} - T\sum_{s\in\mathscr{S}}c_s^{\mathrm{IN}}n_s^{\mathrm{IN}} \\ +T\sum_{s\in\mathscr{S}}c_s^{\mathrm{OUT}}n_s^{\mathrm{OUT}} + \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{q} - \epsilon\sqrt{\boldsymbol{q}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{q}} \end{array}\right\} \tag{23}$$

Therefore, we prove **Proposition 1** by introducing the vector $\boldsymbol{q}$. (Q.E.D.)

5.2. Constraints reformulation

The general idea of dealing with a joint chance constraint is to first decompose it into several ICCs. We follow the same idea as above in this section. To characterize the distribution function G of random vector $\boldsymbol{\vartheta}$, we first construct the ambiguity set $\mathscr{G}$, which contains the possible distributions of $\boldsymbol{\vartheta}$, on the basis of the estimated mean $\mathbf{u}$ and covariance matrix $\boldsymbol{\Gamma}$ of $\boldsymbol{\vartheta}$. To avoid excessive conservatism, an ellipsoidal set is used. The specific expression of $\mathscr{G}$ is provided below:

$$\mathscr{G} := \left\{\begin{array}{c} G : (\boldsymbol{\vartheta}-\mathbf{u})^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}(\boldsymbol{\vartheta}-\mathbf{u}) \leq \widetilde{\epsilon}^2 \\ \mathbb{E}_G[\boldsymbol{\vartheta}] = \mathbf{u} \\ \mathbb{E}_G\left[\boldsymbol{\vartheta}^2\right] = \mathbf{u}^{\mathrm{T}}\mathbf{u} + \boldsymbol{\Gamma} \end{array}\right\} \tag{24}$$

where the pre-determined parameter $\widetilde{\epsilon} > 0$ is used to control the size of the ambiguity set $\mathscr{G}$.

First, based on $\mathscr{G}$, we decompose the joint chance constraints to obtain several ICCs and perform conic transformation. For convenience of subsequent description, we set $v(\widetilde{\mathbf{q}}_{lr}, \widetilde{x}_r) = \sum_{w\in\mathscr{W}}\vartheta_w\sum_{h^w\in\mathscr{H}^w}q_{h^w}\delta_{h^w}^{lr} - \sum_{s\in\mathscr{S}}n_{sr}^{\mathrm{VOL}}CAP_s/\left\lceil\frac{T}{7}\right\rceil = \boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr} - \widetilde{n}_r$, where $\widetilde{n}_r = \sum_{s\in\mathscr{S}}n_{sr}^{\mathrm{VOL}}CAP_s/\left\lceil\frac{T}{7}\right\rceil$ and $\widetilde{\mathbf{q}}_{lr} = \left(q_{h^w}\delta_{h^w}^{lr}\right)$. The conic reformulation for the ICC is inextricably linked to the upper bound of $\mathbb{E}[v(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r)^+]$, where $(\cdot)^+ := \max\{\cdot,0\}$. Therefore, we present the expression of the above upper bound in **Lemma 1** and then prove the sub-additive property associated with this bound in **Lemma 2**.

Lemma 1. Let $\mathbf{u}$ and $\boldsymbol{\Gamma}$ denote the mean and covariance matrix of vector $\boldsymbol{\vartheta}$, respectively. An upper bound of $\mathbb{E}[v(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r)^+]$ can be written as the following function $\pi(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r) = \frac{1}{2}\left(\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u} - \widetilde{n}_r\right) + \frac{1}{2}\sqrt{\left(\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u} - \widetilde{n}_r\right)^2 + \widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\boldsymbol{\Gamma}\widetilde{\mathbf{q}}_{lr}}$.

Proof. Given that $a^+ = 1/2(a + |a|)$, we obtain

$$\mathbb{E}[v(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r)^+] = \frac{1}{2}\mathbb{E}[v(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r) + |v(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r)|] \tag{25}$$

Applying Jensen's inequality to the convex function $\Theta(x) = x^2$ yields

$$
\begin{aligned}
&\left(\mathbb{E}[|\nu(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r)|]\right)^2 \le \mathbb{E}\left[|\nu(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r)|^2\right] = \mathbb{E}\left[\left|\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}-\widetilde{n}_r\right|^2\right]\\
&= \mathbb{E}\left[\left(\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}\right)^2+\widetilde{n}_r^2-2\widetilde{n}_r\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}\right]\\
&= \mathbb{E}\left[\left(\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}\right)^2\right]+\widetilde{n}_r^2-2\widetilde{n}_r\widetilde{\mathbf{q}}_{lr}^{\mathrm{T}}lr\mathbb{E}[\boldsymbol{\vartheta}]\\
&= \left(\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u}\right)^2+\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\boldsymbol{\Gamma}\widetilde{\mathbf{q}}_{lrlr}+\widetilde{n}_r^2-2\widetilde{n}_r^2\widetilde{\mathbf{q}}lr^{\mathrm{T}}{}_{lr}\mathbf{u}\\
&= \left(\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u}-\widetilde{n}_r\right)^2+\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\boldsymbol{\Gamma}\widetilde{\mathbf{q}}_{lr}.
\end{aligned}
\tag{26}
$$

Now, we have $\mathbb{E}[\nu(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r)^+] \le \frac{1}{2}\left(\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u}-\widetilde{n}_r\right)+\frac{1}{2}\sqrt{\left(\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u}-\widetilde{n}_r\right)^2+\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\boldsymbol{\Gamma}\widetilde{\mathbf{q}}_{lr}}$. (Q.E.D.)

Lemma 2. $\pi(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r)$ is a sub-additive function. This means that for $(\widetilde{\mathbf{q}}_{11},\widetilde{n}_1)$ and $(\widetilde{\mathbf{q}}_{22},\widetilde{n}_2)$, we have $\pi(\widetilde{\mathbf{q}}_{11},\widetilde{n}_1)+\pi(\widetilde{\mathbf{q}}_{22},\widetilde{n}_2) \ge \pi(\widetilde{\mathbf{q}}_{11}+\widetilde{\mathbf{q}}_{22},\widetilde{n}_1+\widetilde{n}_2)$.

Proof. We set $A = \widetilde{\mathbf{q}}^{\mathrm{T}}\mathbf{u}-\widetilde{n}$, the matrix $\mathbf{B} = (A,\widetilde{\mathbf{q}})$, and the matrix $\boldsymbol{\Gamma}' = \begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\Gamma}\end{pmatrix}$. Then, we have $\sqrt{(\widetilde{\mathbf{q}}^{\mathrm{T}}\mathbf{u}-\widetilde{n})^2+\widetilde{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\Gamma}\widetilde{q}} = \sqrt{A^2+\widetilde{\mathbf{q}}^{\mathrm{T}}\boldsymbol{\Gamma}\widetilde{q}} = \sqrt{\mathbf{B}^{\mathrm{T}}\boldsymbol{\Gamma}'B} = \|(\boldsymbol{\Gamma}')^{\frac{1}{2}}\mathbf{B}\|_1$. Therefore, function $\pi(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r)$ is sub-additive because the norm is a sub-additive function. (Q.E.D.)

Following **Lemma 1**, we can apply **Proposition 2** to perform a conic reformulation of the ICC decomposed by (15).

Proposition 2. The ICC

$$
\mathbb{P}\left\{\sum_{s\in\mathscr{S}} n_{sr}^{\mathrm{VOL}}CAP_s \Big/ \left\lceil\frac{T}{7}\right\rceil - \sum_{w\in\mathscr{W}}\vartheta_w\sum_{h^w\in\mathscr{H}^w} q_{h^w}\delta_{h^w}^{lr} \ge 0\right\} \ge 1-\varphi \qquad \forall l\in\mathscr{L}_r, r\in\mathscr{R}
\tag{27}
$$

can be conic transformed as

$$
\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u}-\widetilde{n}_r+\sqrt{\frac{(1-\varphi)\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\boldsymbol{\Gamma}\widetilde{\mathbf{q}}_{lr}}{\varphi}} \le 0 \qquad \forall l\in\mathscr{L}_r, r\in\mathscr{R}
\tag{28}
$$

where $\varphi = 1-\alpha$.

Proof. We utilize the convex approximation method based on the CVaR measure to address the ICC to make it computationally tractable (Ben-Tal and Nemirovski, 1998; Ben-Tal and Teboulle, 1986). This approach was extended by Rockafellar and Uryasev (2000). The CVaR of $\boldsymbol{\vartheta}$ at a confidence level $1-\varphi$ is formulated as $\tau_{1-\varphi}(\boldsymbol{\vartheta}) := \min_{\partial\in\mathbb{R}}\left\{\partial+\frac{1}{\varphi}\mathbb{E}[(\boldsymbol{\vartheta}-\partial)^+]\right\}$. Moreover, it is well known that $\mathbb{P}\{y(\boldsymbol{\vartheta})\} \ge 1-\varphi$ can be convex tightest approximated by $\tau_{1-\varphi}[y(\boldsymbol{\vartheta})] \le 0$, where $y(\boldsymbol{\vartheta})$ is affinely dependent on $\boldsymbol{\vartheta}$. Extending the above theory to our problem, the constraints specified in **Proposition 2** can be reformulated as:

$$
\begin{aligned}
&\min_{\partial\in\mathbb{R}}\left\{\partial+\frac{1}{\varphi}\mathbb{E}\left[\left(\sum_{w\in\mathscr{W}}\vartheta_w\sum_{h^w\in\mathscr{H}^w} q_{h^w}\delta_{h^w}^{lr}-\sum_{s\in\mathscr{S}} n_{sr}^{\mathrm{VOL}}CAP_s\Big/\left\lceil\frac{T}{7}\right\rceil-\partial\right)^+\right]\right\}\\
&\qquad = \min_{\partial\in\mathbb{R}}\left\{\partial+\frac{1}{\varphi}\mathbb{E}\left[\left(\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}-\widetilde{n}_r-\partial\right)^+\right]\right\} \le 0.
\end{aligned}
$$

The closed-form solution $\mathbb{E}[(\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}-\widetilde{n}_r-\partial)^+]$ obtained remains computationally intractable. Therefore, we transform it further via **Lemma 1**. Applying the bound suggested in **Lemma 1**, we have

$$
\begin{aligned}
&\tau_{1-\varepsilon}[\nu(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r)] \le \min_{\partial}\left[\partial+\frac{1}{\varphi}\pi\left(\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr},\widetilde{n}_r+\partial\right)\right]\\
&= \min_{\partial}\left[\partial+\frac{\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u}-\widetilde{n}_r-\partial}{2\varphi}+\frac{\sqrt{\left(\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u}-\widetilde{n}_r-\partial\right)^2+\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\boldsymbol{\Gamma}\widetilde{\mathbf{q}}_{lr}}}{2\varphi}\right]\\
&\qquad = \widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u}-\widetilde{n}_r+\sqrt{\frac{(1-\varphi)\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\boldsymbol{\Gamma}\widetilde{\mathbf{q}}_{lr}}{\varphi}}.
\end{aligned}
\tag{29}
$$

Constraint (29) represents a valid conic transformation of constraint (27) because the last equation is satisfied when $\partial = \partial^* = \frac{(1-2\varphi)}{2\sqrt{\varphi(1-\varphi)}}\sqrt{\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\boldsymbol{\Gamma}\widetilde{\mathbf{q}}_{lr}}$. For clarity, hereafter, we use

$$\varrho_{1-\varphi}[v(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r)]=\min_{\partial}\left[\partial+\frac{1}{\varphi}\pi\left(\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u},\widetilde{n}_r+\partial\right)\right]=\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u}-\widetilde{n}_r+\sqrt{\frac{(1-\varphi)\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{\Gamma}\widetilde{\mathbf{q}}_{lr}}{\varphi}} \tag{30}$$

to denote an upper bound of $\tau_{1-\varepsilon}[v(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r)]$. (Q.E.D.)

Now, we extend the above method to obtain the conic transformation of (15). Note that (15) is equivalent to

$$\mathbb{P}\left(\cup_{l\in\mathscr{L}_r,r\in\mathscr{R}}\left(\sum_{s\in\mathscr{S}}n_{sr}^{\mathrm{VOL}}CAP_s\Big/\left\lceil\frac{T}{7}\right\rceil-\sum_{w\in\mathscr{W}}\vartheta_w\sum_{h^w\in\mathscr{H}^w}q_{h^w}\delta_{h^w}^{lr}>0\right)\right)\le\varphi \tag{31}$$

where $\varphi=1-\alpha$. Based on Galambos (1977), which proposed Bonferroni inequality, (31) can be further decomposed as

$$\begin{aligned}&\mathbb{P}\left\{\bigcup_{l\in\mathscr{L}_r,r\in\mathscr{R}}\left(\sum_{s\in\mathscr{S}}n_{sr}^{\mathrm{VOL}}CAP_s\Big/\left\lceil\frac{T}{7}\right\rceil-\sum_{w\in\mathscr{W}}\vartheta_w\sum_{h^w\in\mathscr{H}^w}q_{h^w}\delta_{h^w}^{lr}>0\right)\right\}\\&\le\sum_{l\in\mathscr{L}_r}\sum_{r\in\mathscr{R}}\mathbb{P}\left(\sum_{s\in\mathscr{S}}n_{sr}^{\mathrm{VOL}}CAP_s\Big/\left\lceil\frac{T}{7}\right\rceil-\sum_{w\in\mathscr{W}}\vartheta_w\sum_{h^w\in\mathscr{H}^w}q_{h^w}\delta_{h^w}^{lr}>0\right)\le\varphi.\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}&\mathbb{P}\left(\sum_{s\in\mathscr{S}}n_{sr}^{\mathrm{VOL}}CAP_s\Big/\left\lceil\frac{T}{7}\right\rceil-\sum_{w\in\mathscr{W}}\vartheta_w\sum_{h^w\in\mathscr{H}^w}q_{h^w}\delta_{h^w}^{lr}>0\right)\le\varphi_{lr},\qquad\forall l\in\mathscr{L}_r,r\in\mathscr{R}\\&\Rightarrow\mathbb{P}\left(\sum_{s\in\mathscr{S}}n_{sr}^{\mathrm{VOL}}CAP_s\Big/\left\lceil\frac{T}{7}\right\rceil-\sum_{w\in\mathscr{W}}\vartheta_w\sum_{h^w\in\mathscr{H}^w}q_{h^w}\delta_{h^w}^{lr}\le0\right)\ge1-\varphi_{lr},\qquad\forall l\in\mathscr{L}_r,r\in\mathscr{R}\end{aligned} \tag{32}$$

where $\sum_{l\in\mathscr{L}_r}\sum_{r\in\mathscr{R}}\varphi_{lr}=\varphi$.

It is evident that the only differences between the chance constraints presented in (31) and (32) are found in φ_{lr} and φ. Therefore, if we use (32) directly to approximate (31), the challenge arises in identifying an adequately sized risk budget φ_{lr}. The techniques proposed in some papers (e.g., Nemirovski and Shapiro (2007)) suggests that we can set φ_{lr} to $\varphi/|\mathscr{L}|$ (where $|\mathscr{L}|$ denotes number of elements in set $\mathscr{L}:=\bigcup_{r\in\mathscr{R}}\mathscr{L}_r$). Nevertheless, according to Zymler et al. (2013), this approach tends to yield overly conservative results.

To address this issue, we utilize the approach in Chen et al. (2010). This approach involves creating an optimization model with a set, say $\mathscr{D}$, and a vector (including a number of positive scaling parameters), say $\mathbf{\Upsilon}$, to reformulate (15). Here, we define set $\mathscr{D}:=\{(l,r)|l\in\mathscr{L}_r,r\in\mathscr{R}\}$, index set $\mathscr{C}\subseteq\mathscr{D}$ and $\mathbf{\Upsilon}=(\Upsilon_{lr})>\mathbf{0}$ $(\forall l\in\mathscr{L}_r,r\in\mathscr{R})$. Following Zymler et al. (2013), the inclusion of Υ_{lr} has no effect on the feasible region of (15), and furthermore, serves to enhance its approximation. In the following **Proposition 3**, we establish this optimization model.

Proposition 3. Assume that $\mathscr{J}$ is the smallest closed convex set containing the support of $\boldsymbol{\vartheta}$. The function $\psi(\boldsymbol{n},\widetilde{\mathbf{q}},\Upsilon,\mathscr{C})$ is shown as

$$\psi(\boldsymbol{n},\widetilde{\mathbf{q}},\Upsilon,\mathscr{C}):=\min_{\mathbf{a}_{lr},b_r}\left\{\min_{\partial}\left[\partial+\frac{\pi(\mathbf{a}_{lr},b_r+\partial)}{\varphi}\right]+\frac{1}{\varphi}\left[\sum_{(l,r)\in\mathscr{C}}\pi(\Upsilon_{lr}\widetilde{\mathbf{q}}_{lr}-\mathbf{a}_{lr},\Upsilon_{lr}\widetilde{n}_r-b_r)\right]\right\}.$$

If $\psi(\mathbf{n},\widetilde{\mathbf{q}},\Upsilon,\mathscr{C})\le0$ and $\max_{\vartheta\in\mathscr{J}}(\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}-\widetilde{n}_r)\le0$ $(\forall(l,r)\in\mathscr{D}\setminus\mathscr{C})$, then the joint chance constraint (15) holds.

Proof. Since the robust counterpart $\max_{\vartheta\in\mathscr{J}}\left(\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}-\widetilde{n}_r\right)\le0$ $(\forall(l,r)\in\mathscr{D}\setminus\mathscr{C})$ implies $\mathbb{P}\left(\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}-\widetilde{n}_r>0\right)=0$ $(\forall(l,r)\in\mathscr{D}\setminus\mathscr{C})$ and with parameter $\mathbf{\Upsilon}=(\Upsilon_{lr})>0$, we have

$$\begin{aligned}\mathbb{P}\left(\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}-\widetilde{n}_r\le0,\qquad\forall(l,r)\in\mathscr{D}\right)&=\mathbb{P}\left(\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}-\widetilde{n}_r\le0,\qquad\forall(l,r)\in\mathscr{C}\right)\\&=\mathbb{P}\left(\max_{(l,r)\in\mathscr{C}}\left\{\Upsilon_{lr}\left[\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}-\widetilde{n}_r\right]\right\}\le0\right).\end{aligned}$$

In addition, in light of **Proposition 2**, it is evident that (28) represents the most precise convex approximation of the ICC. When $(\widetilde{\mathbf{q}},\widetilde{\mathbf{n}})^{\mathrm{T}}$ is feasible, we can obtain $\tau_{1-\varphi}\left[\max_{(l,r)\in\mathscr{C}}\{\Upsilon_{lr}[\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}-\widetilde{n}_r]\}\right]\le0$. In Meilijson and Nádas (1979), the authors proposed (33). If we further define $K=\mathbf{a}_{lr}^{\mathrm{T}}\boldsymbol{\vartheta}-b_r$, (34) is obtained.

$$\mathbb{E}\left[\max_{i=1,\cdots,n}\{A_i\}-\iota\right]^+\le\mathbb{E}[K-\iota]^+\sum_{i=1}^{n}\mathbb{E}[A_i-K]^+,\ \text{for any}\ r.v.\ K \tag{33}$$

$$
\begin{aligned}
&\tau_{1-\varphi}\left[\max_{(l,r)\in\mathscr{C}}\left[\mathrm{Y}_{lr}\left(\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}-\widetilde{n}_r\right)\right]\right]\\
&=\min_{\partial}\left\{\partial+\frac{1}{\varphi}\mathbb{E}\left[\left(\max_{(l,r)\in\mathscr{C}}\left[\mathrm{Y}_{lr}\left(\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}-\widetilde{n}_r\right)\right]-\partial\right)^{+}\right]\right\}\\
&\le\min_{\partial,\mathbf{a}_{lr},b_r}\left\{\partial+\frac{1}{\varphi}\left[\mathbb{E}\left[\left(\mathbf{a}_{lr}^{\mathrm{T}}\boldsymbol{\vartheta}-b_r-\partial\right)^{+}\right]+\sum_{(l,r)\in\mathscr{C}}\mathbb{E}\left[\left(\left[\mathrm{Y}_{lr}\widetilde{\mathbf{q}}_{lr}-\mathbf{a}_{lr}\right]^{\mathrm{T}}\mathbf{q}-\mathrm{Y}_{lr}\widetilde{n}_r+b_r\right)^{+}\right]\right]\right\}\\
&\le\min_{\partial,\mathbf{a}_{lr},b_r}\left\{\partial+\frac{1}{\varphi}\left[\pi(\mathbf{a}_{lr},b_r+\partial)+\sum_{(l,r)\in\mathscr{C}}\pi(\mathrm{Y}_{lr}\widetilde{\mathbf{q}}_{lr}-\mathbf{a}_{lr},\mathrm{Y}_{lr}\widetilde{n}_r-b_r)\right]\right\}\\
&=\min_{\mathbf{a}_{lr},b_r}\left\{\min_{\partial}\left[\partial+\frac{\pi(\mathbf{a}_{lr},b_r+\partial)}{\varphi}\right]+\frac{1}{\varphi}\left[\sum_{(l,r)\in\mathscr{C}}\pi(\mathrm{Y}_{lr}\widetilde{\mathbf{q}}_{lr}-\mathbf{a}_{lr},\mathrm{Y}_{lr}\widetilde{n}_r-b_r)\right]\right\}\\
&=\psi(\mathbf{n},\widetilde{\mathbf{q}},\mathrm{Y},\mathscr{C})\le 0
\end{aligned}
\tag{34}
$$

Therefore, if $\psi(\mathbf{n},\widetilde{\mathbf{q}},\mathrm{Y},\mathscr{C})\le 0$, then $\tau_{1-\varphi}\left[\max_{(l,r)\in\mathscr{C}}\{\mathrm{Y}_{lr}[\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}-\widetilde{n}_r]\}\right]\le 0$. (Q.E.D.)

Lemma 3. $\psi(\mathbf{n},\widetilde{\mathbf{q}},\mathrm{Y},\mathscr{C})\le 0$ dominates constraint (28).

Proof. We have

$$
\begin{aligned}
0&\ge\min_{\partial,\mathbf{a}_{lr},b_r}\left\{\partial+\frac{1}{\varphi}\left[\pi(\mathbf{a}_{lr},b_r+\partial)+\sum_{(l,r)\in\mathscr{C}}\pi(\mathrm{Y}_{lr}\widetilde{\mathbf{q}}_{lr}-\mathbf{a}_{lr},\mathrm{Y}_{lr}\widetilde{n}_r-b_r)\right]\right\}\\
&\ge\min_{\partial,\mathbf{a}_{lr},b_r}\left\{\partial+\frac{1}{\varphi}[\pi(\mathbf{a}_{lr},b_r+\partial)+\pi(\mathrm{Y}_{lr}\widetilde{\mathbf{q}}_{lr}-\mathbf{a}_{lr},\mathrm{Y}_{lr}\widetilde{n}_r-b_r)]\right\}\\
&\ge\min_{\partial}\left\{\partial+\frac{1}{\varphi}[\pi(\mathrm{Y}_{lr}\widetilde{\mathbf{q}}_{lr},\mathrm{Y}_{lr}\widetilde{n}_r+\partial)]\right\}\\
&\overset{\mathrm{Y}_{lr}=1}{=}\min_{\partial}\left\{\partial+\frac{1}{\varphi}[\pi(\widetilde{\mathbf{q}}_{lr},\widetilde{n}_r+\partial)]\right\}\\
&=\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u}-\widetilde{n}_r+\sqrt{\frac{(1-\varphi)\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\boldsymbol{\Gamma}\widetilde{\mathbf{q}}_{lr}}{\varphi}}.
\end{aligned}
$$

According to **Proposition 3**, the first inequality holds, and the second inequality is a result of the fact that we select only one element from $\mathscr{C}$. The third one is derived based on **Lemma 2** proposed in this section. Finally, the two equations at the end are a consequence of $\Upsilon_{lr}=1$ and (29). (Q.E.D.)

Now, we introduce two auxiliary variables (i.e., t_0 and t_{lr}), and $\psi(\mathbf{n},\widetilde{\mathbf{q}},\mathrm{Y},\mathscr{C})\le 0$ is equivalent to $\exists b_r,t_0,t_{lr}\in\mathbb{R}$ such that constraints (35)–(37) hold.

$$
t_0+\frac{1}{\varphi}\sum_{(l,r)\in\mathscr{C}}t_{lr}\le 0 \tag{35}
$$

$$
\varrho_{1-\varphi}[\nu(\mathbf{a}_{lr},b_r+\partial)]\le t_0\qquad\forall(l,r)\in\mathscr{C} \tag{36}
$$

$$
\pi(\mathrm{Y}_{lr}\widetilde{\mathbf{q}}_{lr}-\mathbf{a}_{lr},\mathrm{Y}_{lr}\widetilde{n}_r-b_r)\le t_{lr}\qquad\forall(l,r)\in\mathscr{C} \tag{37}
$$

The robust counterpart, i.e., $\max_{\boldsymbol{\vartheta}\in\mathscr{J}}(\boldsymbol{\vartheta}^{\mathrm{T}}\widetilde{\mathbf{q}}_{lr}-\widetilde{n}_r)\le 0$ $(\forall(l,r)\in\mathscr{D}\setminus\mathscr{C})$, can be converted to (38) according to **Theorem 3(c)** in Chen and Sim (2009).

$$
\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u}-\widetilde{n}_r+\widetilde{\epsilon}\sqrt{\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\boldsymbol{\Gamma}\widetilde{\mathbf{q}}_{lr}}\le 0\qquad\forall(l,r)\in\mathscr{D}\setminus\mathscr{C} \tag{38}
$$

At this point, constraints (35)–(37) can dominate function $\psi(\mathbf{n},\widetilde{\mathbf{q}},\mathrm{Y},\mathscr{C})\le 0$, and with constraint (38), we can reformulate the original problem **P** in its deterministic equivalent form, i.e., **R-P**.

R-P:

$$
\max obj'_1=\left\{\begin{array}{c}
-T\sum_{r\in\mathscr{R}}\sum_{s\in\mathscr{S}}c_s^{\mathrm{OP}}n_{sr}^{\mathrm{TOTAL}}-\sum_{r\in\mathscr{R}}\sum_{s\in\mathscr{S}}c_{sr}^{\mathrm{VOL}}n_{sr}^{\mathrm{VOL}}-T\sum_{s\in\mathscr{S}}c_s^{\mathrm{IN}}n_s^{\mathrm{IN}}\\
+T\sum_{s\in\mathscr{S}}c_s^{\mathrm{OUT}}n_s^{\mathrm{OUT}}+\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{q}-\epsilon\sqrt{\boldsymbol{q}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{q}}
\end{array}\right\} \tag{39}
$$

$$
s.t.\ (2)-(12),(14),(16),(20),(21),(35)-(38).
$$

where $\widetilde{n}_r = \sum_{s\in\mathscr{J}} n_{sr}^{\mathrm{VOL}} CAP_s / \left\lceil \frac{T}{7} \right\rceil$.

Proposition 4. When $\widetilde{\epsilon} \le \sqrt{(1-\varphi)/\varphi}$, constraints (35)–(37) in **R-P** are redundant.

Proof. We realize the proof of this proposition by means of a counterfactual. To prove **Proposition 4**, we can equivalently prove that we can obtain an optimal solution of **R-P** with $\mathscr{C} = \varnothing$ when $\widetilde{\epsilon} \le \sqrt{(1-\varphi)/\varphi}$. Let us posit the existence of a counterexample. We assume that $((l^*,r^*) \in \mathscr{C})$ enables the **R-P** to achieve an optimal solution with $(l^*,r^*) \in \mathscr{C}$. Given that all elements within the set $\mathscr{C}$ are required to satisfy $\psi(\mathbf{n}, \widetilde{\mathbf{q}}, \mathrm{Y}, \mathscr{C}) \le 0$, it follows that there must be $\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u} - \widetilde{n}_r + \widetilde{\epsilon}\sqrt{\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{\Gamma}\widetilde{\mathbf{q}}_{lr}} \le \widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u} - \widetilde{n}_r + \sqrt{(1-\varphi)\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{\Gamma}\widetilde{\mathbf{q}}_{lr}/\varphi} \le \psi(\mathbf{n}, \widetilde{\mathbf{q}}, \mathrm{Y}, (l^*,r^*)) \le 0$. We then note that, according to $\widetilde{\epsilon} \le \sqrt{(1-\varphi)/\varphi}$, the first inequality holds. The second inequality is a consequence of the results presented in **Lemma 3**. The above series of inequalities shows that inequality $\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u} - \widetilde{n}_r + \widetilde{\epsilon}\sqrt{\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{\Gamma}\widetilde{\mathbf{q}}_{lr}} \le 0$ is less conservative than inequality $\psi(\mathbf{n}, \widetilde{\mathbf{q}}, \mathrm{Y}, (l^*,r^*)) \le 0$. It can thus be concluded that when $(l^*,r^*) \in \mathscr{S} \setminus \mathscr{C}$, a superior solution can be obtained in comparison to when $(l^*,r^*) \in \mathscr{C}$. This is inconsistent with the assumption. (Q.E.D.)

Remark. It can be observed that the aforementioned proposition renders the **R-P** equivalent to a standard parametric SOCPM (Zhang and Li, 2015). Following **Proposition 4**, it is evident that the **R-P** is independent of $\mathscr{C}$ and $\Upsilon_{lr} > 0$ $(\forall (l,r) \in \mathscr{C})$ when $\widetilde{\epsilon} \le \sqrt{(1-\varphi)/\varphi}$. Meanwhile, we introduce an auxiliary variable $\varpi \ge 0$ to eliminate the nonlinear term (i.e., $\epsilon\sqrt{\boldsymbol{q}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{q}}$) in the objective function. Consequently, **R-P** can be subjected to a further conversion process to yield **R-P-1**. The OAA framework can then be employed.

R-P-1:

$$\max obj'_1 = \left\{ \begin{array}{c} -T\sum_{r\in\mathscr{R}}\sum_{s\in\mathscr{J}} c_s^{\mathrm{OP}} n_{sr}^{\mathrm{TOTAL}} - \sum_{r\in\mathscr{R}}\sum_{s\in\mathscr{J}} c_{sr}^{\mathrm{VOL}} n_{sr}^{\mathrm{VOL}} - T\sum_{s\in\mathscr{J}} c_s^{\mathrm{IN}} n_s^{\mathrm{IN}} \\ +T\sum_{s\in\mathscr{J}} c_s^{\mathrm{OUT}} n_s^{\mathrm{OUT}} + \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{q} - \epsilon\varpi \end{array} \right\} \tag{40}$$

$$s.t. \ \widetilde{\epsilon} \le \sqrt{(1-\varphi)/\varphi} \tag{41}$$

$$\| \boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{q} \| \le \varpi \tag{42}$$

$$\varpi \ge 0 \tag{43}$$

(2), (3), (4), (5), (6), (7), (8), (9), (10), (11), (12), (14), (16), (20), (21), (38).

6. Solution methodology

In this subsection, we present a customized algorithm to solve **P** via the OAA framework. The **R-P-1** model we reformulate at the end of Section 5 is a special type of MINLP. Bonami et al. (2008) demonstrated that OAA can be used to explore the global optimal solution when the problem can be formulated as a *convex MINLP model*. An MINLP is considered *convex* if the continuous relaxation of this model is convex (Shahabi et al., 2014). This reinforces the need to verify that our **R-P-1** model satisfies the stated requirements prior to implementing the OAA. This is necessary to ensure the global optimality of the solution. In the following **Proposition 5**, we provides the proof.

Proposition 5. Functions

$$F_1(\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}, \widetilde{n}_r) = \widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u} - \widetilde{n}_r + \widetilde{\epsilon}\sqrt{\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{\Gamma}\mathbf{q}_{lr}}$$

$$F_2(\boldsymbol{q}, \varpi) = \sqrt{\boldsymbol{q}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{q}} - \varpi$$

are convex.

Proof. We know that the 1-norm is a convex function. In function $F_1(\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}, \widetilde{n}_r)$, we have $F_1(\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}, \widetilde{n}_r) = \widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u} - \widetilde{n}_r + \widetilde{\epsilon}\sqrt{\widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{\Gamma}\widetilde{\mathbf{q}}_{lr}} = \widetilde{\mathbf{q}}^{\mathrm{T}}{}_{lr}\mathbf{u} - \widetilde{n}_r + \widetilde{\epsilon}\| \mathbf{\Gamma}^{\frac{1}{2}}\widetilde{\mathbf{q}}_{lr} \|_1$. Moreover, in function $F_2(\boldsymbol{q},\varpi)$, we know that $\sqrt{\boldsymbol{q}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{q}} = \| \boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{q} \|_1$. Thus, we can conclude that they are both convex functions. (Q.E.D.)

6.1. Algorithm initialization

We should first define a feasible $\left(\widetilde{x}_{sr,0}, \widetilde{n}_{sr,0}^{\mathrm{Total}}, \widetilde{n}_{sr,0}^{\mathrm{VOL}}, \widetilde{n}_{s,0}^{\mathrm{IN}}, \widetilde{n}_{s,0}^{\mathrm{OUT}}\right)^{\mathrm{T}}$ to initialize our algorithm. To this end, it seems reasonable to posit that the liner company does not charter out liners, i.e., we set $\widetilde{n}_{s,0}^{\mathrm{OUT}} = 0$. The liner company deploys the liner with the highest capacity for each route, i.e., $\widetilde{x}_{s^*r,0} = 1$, $s^* = \{s | \max_{s\in\mathscr{J}} CAP_s\}$. On each route, the number of voyages completed by the liner is $\widetilde{n}_{s^*r,0}^{\mathrm{VOL}} = \left\lceil \frac{T}{7} \right\rceil$ and

$\widetilde{n}_{s^*r,0}^{\text{Total}} = \widetilde{n}_{s^*r,0}^{\text{VOL}} / \left\lfloor \frac{T}{t_{sr}} \right\rfloor$. The liner company prioritizes the use of owned liners, and after all owned liners have been deployed, it begins to seek to charter liners in the market. Note that if the number of s^* class liners deployed in the shipping network exceeds $NO_{s^*}^{\text{MAX}} + NC_{s^*}^{\text{MAX}}$ according to the above rule, the introduction of liners with lower capacity begins. According to the size of the route index, we prioritize the liner with a higher capacity for routes with a smaller route index.

6.2. Subproblem

The subproblem is concerned with determining the optimal solution for continuous variables in **R-P**. To solve this nonlinear subproblem, integer variables are fixed. The subproblem of **R-P** is written as **R-P-SP**. In this model, we introduce wavy lines and superscripts (i.e., i) to all integer decision variables, i.e., $\left(\widetilde{x}_{sr,0}, \widetilde{n}_{sr,0}^{\text{Total}}, \widetilde{n}_{sr,0}^{\text{VOL}}, \widetilde{n}_{s,0}^{\text{IN}}, \widetilde{n}_{s,0}^{\text{OUT}}\right)$, thereby indicating that these variables are incorporated into the subproblem at the i^{th} iteration.

R-P-SP:

$$\max sp = \left\{ \begin{array}{c} -T\sum_{r\in\mathscr{R}}\sum_{s\in\mathscr{S}} c_s^{\text{OP}} \widetilde{n}_{sr,i}^{\text{TOTAL}} - \sum_{r\in\mathscr{R}}\sum_{s\in\mathscr{S}} c_{sr}^{\text{VOL}} \widetilde{n}_{sr,i}^{\text{VOL}} - T\sum_{s\in\mathscr{S}} c_s^{\text{IN}} \widetilde{n}_{s,i}^{\text{IN}} \\ +T\sum_{s\in\mathscr{S}} c_s^{\text{OUT}} \widetilde{n}_{s,i}^{\text{OUT}} + \boldsymbol{\mu}^{\text{T}}\boldsymbol{q} - \epsilon\varpi \end{array} \right\} \tag{44}$$

$$s.t.\ \widetilde{n}_r = \sum_{s\in\mathscr{S}} \widetilde{n}_{sr,i}^{\text{VOL}} CAP_s \Big/ \left\lceil \frac{T}{7} \right\rceil \qquad \forall r \in \mathscr{R} \tag{45}$$

(14), (16), (20), (21), (38), (41), (43).

6.3. Master problem

The master problem is an MIP model that optimizes integer decision variables on the basis of the continuous variables obtained. Given the presence of nonlinear constraints in **P**, we use **Proposition 6** to perform linear approximations (also known as OA cuts) for constraints (38) and (42).

Proposition 6. Let $\widehat{\mathbf{q}}^i$, $\widehat{\mathbf{q}}_{lr}^i$ and $\widehat{\varpi}^i$ denote the optimal solutions of **R-P-SP** at the i^{th} iteration and $\mathscr{I}$ represents the maximum number of iterations. The valid linear OA cut associated with constraints (38) and (42) can be specified as follows:

$$\left(\widetilde{\mathbf{q}}_{lr}^{\text{T}}\mathbf{u} - \widetilde{n}_r\right)\sqrt{\left(\widehat{\mathbf{q}}_{lr}^i\right)^{\text{T}}\boldsymbol{\Gamma}\widehat{\mathbf{q}}_{lr}^i} + \widetilde{\epsilon}\widetilde{\mathbf{q}}_{lr}^{\text{T}}\boldsymbol{\Gamma}\widehat{\mathbf{q}}_{lr}^i \le 0 \qquad \forall l \in \mathscr{L}_r, r \in \mathscr{R}, i \in \{1,2,\cdots,\mathscr{I}\} \tag{46}$$

$$\sqrt{\boldsymbol{q}^{\text{T}}\boldsymbol{\Sigma}\widehat{\boldsymbol{q}}^i} - \varpi\widehat{\varpi}^i \le 0 \qquad \forall i \in \{1,2,\cdots,\mathscr{I}\}. \tag{47}$$

Proof. From the convexity of these two functions shown in **Proposition 6**, we further perform a Taylor expansion of constraints (45) and (46), which yields

$$F_1\left(\widehat{\mathbf{q}}_{lr}^i, \widehat{n}_r^i\right) + \nabla F_1\left(\widehat{\mathbf{q}}_{lr}^i, \widehat{n}_r^i\right)^{\text{T}} \begin{bmatrix} \mathbf{q}_{lr} - \widehat{\mathbf{q}}_{lr}^i \\ \widetilde{n}_r - \widehat{n}_r^i \end{bmatrix} \le F_1(\mathbf{q}_{lr}, \widetilde{n}_r) \le 0,$$

and

$$F_2\left(\widehat{\boldsymbol{q}}^i, \widehat{\varpi}^i\right) + \nabla F_2\left(\widehat{\boldsymbol{q}}^i, \widehat{\varpi}^i\right)^{\text{T}} \begin{bmatrix} \boldsymbol{q} - \widehat{\boldsymbol{q}}^i \\ \varpi - \widehat{\varpi}^i \end{bmatrix} \le F_2\left(\widehat{\boldsymbol{q}}^i, \widehat{\varpi}^i\right) \le 0,$$

where $\nabla F_1\left(\widehat{\mathbf{q}}_{lr}^i, \widehat{n}_r^i\right) = \left[\mathbf{u} + \frac{\widetilde{\epsilon}\left(\widehat{\mathbf{q}}_{lr}^i\right)^{\text{T}}\boldsymbol{\Gamma}}{\sqrt{\left(\widehat{\mathbf{q}}_{lr}^i\right)^{\text{T}}\boldsymbol{\Gamma}\widehat{\mathbf{q}}_{lr}^i}}, -1\right]$ and $\nabla F_2\left(\widehat{\boldsymbol{q}}^i, \widehat{\varpi}^i\right) = \left[\frac{\left(\widehat{\boldsymbol{q}}^i\right)^{\text{T}}\boldsymbol{\Sigma}}{\sqrt{\left(\widehat{\boldsymbol{q}}^i\right)^{\text{T}}\boldsymbol{\Sigma}\widehat{\boldsymbol{q}}^i}}, -1\right]$. Then, we can obtain the closed-form solutions of (46) and (47) through simple algebra. (Q.E.D.)

Now, we introduce an auxiliary variable ξ and use BU^i (BL^i) to represent the upper (lower) bound of the model at the i^{th} iteration of the algorithm, respectively. On the basis of **Proposition 6**, we can propose **R-P-MP**, i.e., the master problem of **R-P**, as follows.

R-P-MP:

$$\min mp = \xi \tag{48}$$

$$s.t.\ \xi \ge T\sum_{r\in\mathscr{R}}\sum_{s\in\mathscr{S}} c_s^{\text{OP}} n_{sr}^{\text{TOTAL}} + \sum_{r\in\mathscr{R}}\sum_{s\in\mathscr{S}} c_{sr}^{\text{VOL}} n_{sr}^{\text{VOL}} + T\sum_{s\in\mathscr{S}} c_s^{\text{IN}} n_s^{\text{IN}} - T\sum_{s\in\mathscr{S}} c_s^{\text{OUT}} n_s^{\text{OUT}} - \boldsymbol{\mu}^{\text{T}}\boldsymbol{q} + \epsilon\varpi \tag{49}$$

$$\xi \le BU^i - \nu \qquad \forall i \in \{1,2,\cdots,\mathscr{I}\} \tag{50}$$

$$\xi \ge BL^i \qquad \forall i \in \{1,2,\cdots,\mathscr{I}\} \tag{51}$$

(2), (3), (4), (5), (6), (7), (8), (9), (10), (11), (12), (14), (16), (20), (21), (41), (43), (46), (47). where ν is a small positive number and $\widetilde{n}_r = \sum_{s\in\mathscr{S}} n_{sr}^{\mathrm{VOL}} CAP_s / \left\lceil \frac{T}{7} \right\rceil$.

In the aforementioned equations, function (46) represents the objective function of **R-P-MP**. In accordance with the methodology proposed by Fletcher and Leyffer (1994), we introduce (50) to ensure the solution of **R-P-MP** is less than the upper bound. The termination condition is that the **R-P-MP** is infeasible. Note that it is not necessary for the **R-P-MP** to be solved optimally during the iteration of the algorithm, provided that a new feasible integer solution is produced. The solution is designated as ν-optimal. The pseudocode is shown in Algorithm 1.

7. Numerical experiment

In this section, numerical experiments are carried out to verify the validity of the models and algorithms. By carrying out several sensitivity analyses in Section 7.2, we also attempt to gain managerial insights. Finally, in Section 7.3, out-of-sample analysis of the robust optimization model, the SP model, and 2-DRO empirically validates the out-of-sample performance of our model.

7.1. Case study description

We consider an Asia−Europe liner shipping network in this paper. The example consists of 10 routes, 46 ports, 63 legs, and 331 O-D pairs. Freight rate data are sourced from actual businesses. These routes were also chosen for the numerical experiments carried out by Meng et al. (2012) and Wang et al. (2013). For each route, the ports of call included and the voyage distances are reported in Table 1.

We consider that the liner company has five classes of liners, and their performance data (e.g., capacity, speed, and daily operation cost) are shown in Table 2. We apply the technique proposed in Delage and Ye (2010) to estimate ϵ in the ambiguity set $\mathscr{F}$. To this end, on the basis of a multivariate normal distribution, we collect $\overline{M}$ independent scenarios (i.e., samples) for random vector $\boldsymbol{\theta}$, i.e., $\{\boldsymbol{\theta}^i\}_{i=1}^{\overline{M}}$. After these samples are obtained, we calculate the mean ($\widehat{\boldsymbol{\mu}}$) and covariance matrix ($\widehat{\Sigma}$) of these samples. Then, on the basis of (17), β is set to 0.05 and $R^2 = \max_{i=1,2,\cdots,\overline{M}} \left(\boldsymbol{\theta}^i - \widehat{\boldsymbol{\mu}}\right)^{\mathrm{T}} \widehat{\Sigma}^{-1} \left(\boldsymbol{\theta}^i - \widehat{\boldsymbol{\mu}}\right)$, and we estimate parameter ε by calculating $\epsilon = R/\sqrt{\overline{M}}\left(2 + \sqrt{2\ln(1/\beta)}\right)$. In reality, the operational data (e.g., transportation demand data) of liner companies are often difficult to obtain because they involve their commercial secrets. Numerical experiments based on randomly generated transportation demand data are usually conducted. Therefore, we randomly generate 1000 sample data, i.e., $\overline{M} = 1000$ and $\{\boldsymbol{\theta}^i\}_{i=1}^{1000}$. The mean of each component in $\boldsymbol{\theta}$ is uniformly drawn from [10, 60], and the standard deviation is uniformly drawn from [0.1, 1]. In addition, we set the radius $\widetilde{\epsilon}$ of set $\mathscr{G}$ as $\widetilde{\epsilon} \in [0.001, 4]$, the mean of each component in $\boldsymbol{\vartheta}$ is uniformly drawn from [0.1, 12], and the standard deviation is uniformly drawn from [0.1, 2].

Before the formal case study begins, we first conduct experiments on computing time to ensure that the case size involved in the case is appropriate and that our method is suitable for larger shipping networks. To this end, the number of routes is set to 5, 10, and 20, and the number of ports is set to 20, 25, 30, 40, 50, 60, 80, 100, and 120. The number of legs is set to vary from 200 to 200. The number of O-D pairs is set to 20 % of the maximum possibility, considering that some O-D pairs may always have no transportation demand (e.g., two ports that are very close to each other). After combination, we obtain the 27 computing scenarios shown in Table 3.

For each computing scenario, we randomly generate 15 groups of calculation cases and obtain the average calculation time. In all the computing scenarios, the optimal solution of the model is explored within a limited computing time (which we set as 3,600 s), so the optimality gap is 0. The experimental results are in line with our expectations. As the problem size increases, the computation time

Algorithm 1
Outer approximation algorithm.

Outer approximation algorithm
Input: $\widetilde{x}_{sr,0}$, $\widetilde{n}_{sr,0}^{\mathrm{Total}}$, $\widetilde{n}_{sr,0}^{\mathrm{VOL}}$, $\widetilde{n}_{s,0}^{\mathrm{IN}}$, $\widetilde{n}_{s,0}^{\mathrm{OUT}}$: initial value;
BU^0: upper bound, equals to $+\infty$;
BL^0: lower bound, equals to $-\infty$;
$\mathscr{I}$: maximum number of iterations.
Procedure:
For i= 1: $\mathscr{I}$ **do**
  Solve **R-P-SP**. Obtain optimal solution of continuous decision variable. Denote the value of objective function as BU^i.
  Add OA cuts (46) and (47) with fixed $\widehat{q}^i$, $\widehat{\mathbf{q}}_{lr}^i$ and $\widehat{\varpi}^i$. Solve **R-P-MP**, get new feasible integer decision variables. Denote the value of objective function as BL^i.
  If R-P-MP is infeasible **then**
    Break and return the incumbent value, report: $\left(\widetilde{x}_{sr,i}, \widetilde{n}_{sr,i}^{\mathrm{Total}}, \widetilde{n}_{sr,i}^{\mathrm{VOL}}, \widetilde{n}_{s,i}^{\mathrm{IN}}, \widetilde{n}_{s,i}^{\mathrm{OUT}}\right)$, $\widehat{q}^i$, $\widehat{\mathbf{q}}_{lr}^i$ and $\widehat{\varpi}^i$;
  End if
End for

Table 1
Voyage distance and ports of call for different routes.

Route No.	Voyage distance (n mile)	Ports of call
R1	3,440	Yokohama → Tokyo → Nagoya → Kobe → Shanghai → Hong Kong → Yokohama
R2	2,361	Ho Chi Minh → Laem Chabang → Singapore → Port Klang → Ho Chi Minh
R3	8,591	Brisbane → Sydney → Melbourne → Adelaide → Fremantle → Jakarta → Singapore → Brisbane
R4	1,622	Manila → Kaohsiung → Xiamen → Hong Kong → Yantian → Chiwan → Manila
R5	1,997	Dalian → Xingang → Qingdao → Shanghai → Ningbo → Kwangyang → Busan → Dalian
R6	4,948	Chittagong → Chennai → Colombo → Cochin → Nhava Sheva → Chittagong
R7	6,400	Sokhna → Aqabah → Jeddah → Salalah → Karachi → Jebel Ali → Sokhna
R8	1,193	Southampton → Thamesport → Hamburg → Bremerhaven → Rotterdam → Antwerp → Zeebrugge → Le Havre → Southampton
R9	6,009	Port Klang → Singapore → Jakarta → Kaohsiung → Busan → Hong Kong → Chiwan → Port Klang
R10	20,478	Yantian → Hamburg → Sokhna → Jeddah → Port Klang → Singapore → Manila → Yantian

Table 2
Liner performance parameters of different classes.

Item	Liner class				
	1	2	3	4	5
Liner capacity (TEUs)	2808	3218	4500	5714	8063
Sailing speed of the liner (knots)	21.00	22.00	24.20	24.60	25.20
Daily operation cost (10^3 $)	19.80	22.50	30.90	38.80	54.20
Chartering in rate (million $)	2.02	2.63	3.51	4.76	6.09
Chartering out rate (million $)	1.82	2.34	3.21	4.32	5.12
NO_s^{MAX}	2	2	9	2	12
NC_s^{MAX}	5	5	6	5	10

Table 3
Computational time for different scale examples.

No.	Number of routes	Number of ports	Number of legs	Average calculation time (s)	Optimality gap (%)
1	5	20	20	24	0.00
2	5	20	25	33	0.00
3	5	20	30	37	0.00
4	5	25	25	32	0.00
5	5	25	30	35	0.00
6	5	25	35	42	0.00
7	5	30	30	37	0.00
8	5	30	40	41	0.00
9	5	30	50	48	0.00
10	10	40	40	62	0.00
11	10	40	50	94	0.00
12	10	40	60	115	0.00
13	10	50	50	108	0.00
14	10	50	70	156	0.00
15	10	50	90	191	0.00
16	10	60	60	188	0.00
17	10	60	80	202	0.00
18	10	60	100	224	0.00
19	20	80	80	631	0.00
20	20	80	110	678	0.00
21	20	80	140	712	0.00
22	20	100	100	721	0.00
23	20	100	140	835	0.00
24	20	100	180	957	0.00
25	20	120	120	924	0.00
26	20	120	160	1104	0.00
27	20	120	200	1346	0.00

of each computational scenario gradually increases. For larger-scale problems, our algorithm can still obtain the optimal solution to the problem in a reasonable amount of time. Notably, in reality, liner companies do not combine too many routes for unified deployment. This is because if routes involving different continents (e.g., the Asia–Europe route and the Europe–America route) are unified for liner deployment, the liner company may need to pay more costs and time to liners located in Europe to Asia.

7.2. Sensitivity analysis of parameters in the ambiguity set

7.2.1. Sensitivity analysis of the radius $\widetilde{\epsilon}$

The radius $\widetilde{\epsilon}$ of the ambiguity set $\mathscr{G}$ indicates the size of the uncertainty of the weekly contract demand, and the uncertain model reduces to the deterministic model when the radius $\widetilde{\epsilon}$ equals 0. Under different political and economic environments, the degree of uncertainty in the shipping industry's container transportation demand also varies. To explore the impact of changes in $\widetilde{\epsilon}$ (degree of uncertainty) on the optimal solution and objective function value, the radius $\widetilde{\epsilon}$ is set to be in the range of $\widetilde{\epsilon} \in [0.001, 4]$. We set the mean and covariance of contract demand to be the same as those of the benchmark. The computation time for each scenario is approximately 100 s. From Fig. 4, with an increase in the ambiguity set radius, the profit of the liner company decreases significantly from approximately $ 9400 million to over $ 2000 million. Here, a large change in the profit of the liner company indicates that the liner company must cover a much greater additional cost when facing a high level of uncertainty. In contrast to the deterministic scenario where the radius $\widetilde{\epsilon}$ is approximately 0, the profit of the liner company falls to approximately one-fifth of its original value when uncertainty is considered. As illustrated in Fig. 4, we also find that the trend is different when $\widetilde{\epsilon} \leq 1$ and $\widetilde{\epsilon} > 1$. The profit of the liner company decreases more quickly when the radius is lower than 1, with a total decrease of nearly 6000 million HK dollars when $\widetilde{\epsilon} \in [0, 1]$. This is mainly because of the destruction resulting from the emergence of uncertainty, which makes the current scheme unable to adapt to random contract demand. The total number of voyages completed by the fleet has also undergone similar changes; that is, when $\widetilde{\epsilon} \in [0, 1]$, the number of voyages fluctuates, and after $\widetilde{\epsilon} > 1$, it tends to stabilize. Table 4 shows the liner fleet planning scheme, the average capacity of the liner in the fleet and the profit that can be earned by the liner company in three typical scenarios. As $\widetilde{\epsilon}$ increases, the liner company gradually begins to use more large liners to cope with uncertainty, causing the average liner capacity in the fleet to increase.

7.2.2. Sensitivity analysis of mean contract demand $\mathbf{u}$

The mean of contract demand $\mathbf{u}$ also varies according to the different shipping seasons. To gain further insight into the impact of contract demand on the profitability of the liner company and model solution, we conduct several groups of numerical experiments on parameters in the ambiguity set $\mathscr{G}$. In this subsection, we adjust $\mathbf{u}$ to 60 % to 150 % of the benchmark value, and the interval is set to 10 %. The coefficient of variation of $\mathbf{u}$ is between 66.67 % and 166.67 % because the covariance of contract demand is unchanged. The radius $\widetilde{\epsilon}$ of the ambiguity set $\mathscr{G}$ is set to 2, and the risk tolerance φ is set to 0.05. Fig. 5 shows the results of the **R-P-1** model with different means of contract demand. As $\mathbf{u}$ increases, the profit of the liner company decreases, and this decrease is more obvious when the mean of the contract demand is lower than the benchmark value. Table 5 presents the specific liner fleet deployment plans of three typical scenarios and the corresponding average liner capacity of the fleet and the profit of the liner company. Similar to the situation in the previous subsection, as uncertainty increases, the liner company prefers to use larger liners, so the average fleet liner capacity increases. The number of voyages completed by the liner fleet remains relatively stable across scenarios, which shows that adjusting the liner class is the priority method for addressing transportation demand uncertainty.

7.2.3. Sensitivity analysis of covariance of contract demand $\boldsymbol{\Gamma}$

As the world economic situation changes, sometimes the demand for container transportation is relatively stable, and sometimes the fluctuations in demand are more obvious. In this subsection, we focus on the impact of changes in contract demand covariance on the optimal solution of our model. To this end, we adjust the value of $\boldsymbol{\Gamma}$ to 60 % to 150 % of the benchmark, and the radius $\widetilde{\epsilon}$ of the

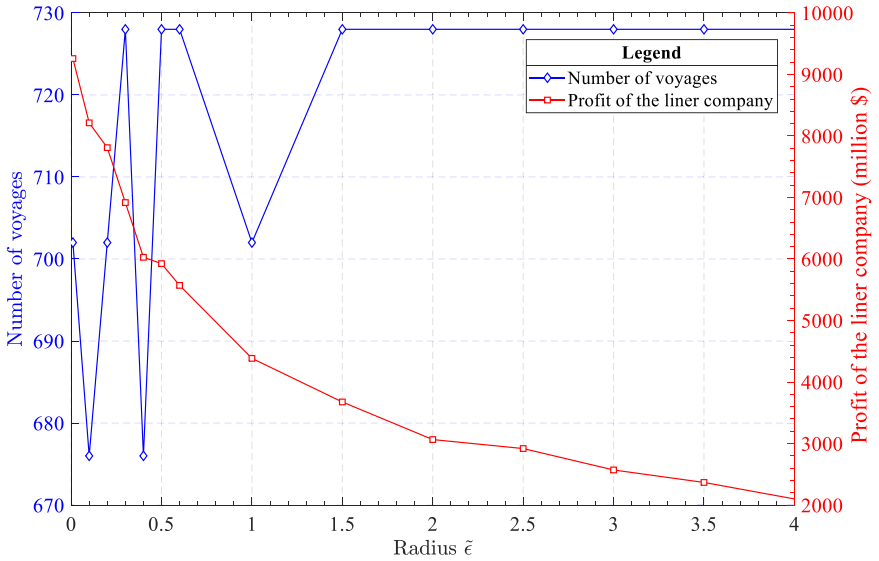


Fig. 4. Sensitivity analysis of the radius $\widetilde{\epsilon}$ in the ambiguity set $\mathscr{G}$.

Table 4
Liner fleet deployment scheme under different radii $\tilde{\epsilon}$ in the ambiguity set $\mathscr{G}$.

$\tilde{\epsilon}$	Liner class	Number of liners										Average capacity (TEU)	Total profit (million $)
		R1	R2	R3	R4	R5	R6	R7	R8	R9	R10		
0.01	1											6218.51	8819
	2										7		
	3					2		5					
	4	3											
	5		2	7	2		4		1	5			
0.10	1		1								7	6351.13	8213
	2												
	3						4						
	4	3											
	5			7	2	2		5	1	5			
2.00	1										7	6599.94	2015
	2												
	3								1				
	4		2			2							
	5	3		3	2		4	5		5			

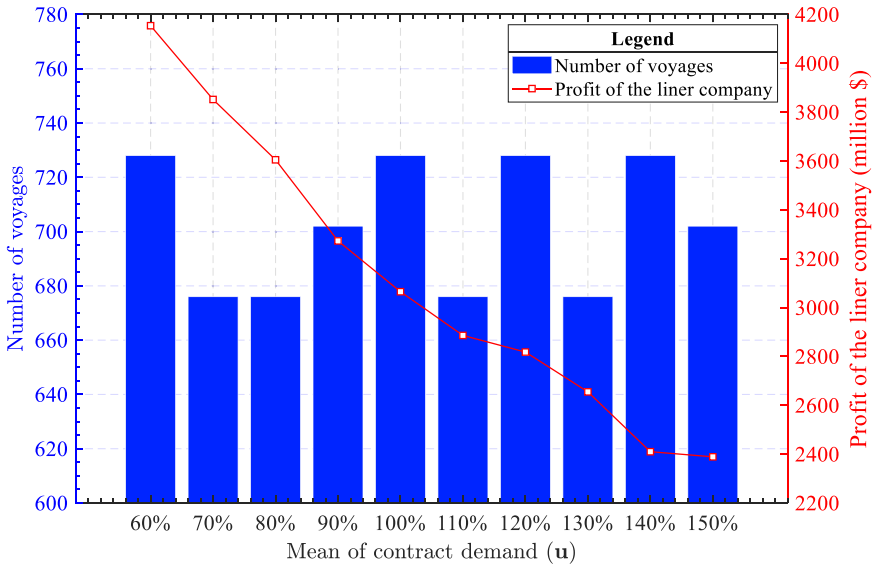


Fig. 5. Sensitivity analysis of mean contract demand **u**.

Table 5
Liner fleet deployment scheme under different mean contract demands **u**.

u	Liner class	Number of liners										Average capacity (TEU)	Total profit (million $)
		R1	R2	R3	R4	R5	R6	R7	R8	R9	R10		
60 %	1										7	6315.00	4430
	2												
	3			7									
	4		2										
	5	3			2	2	4	5	1	5			
70 %	1										7	6517.00	3852
	2												
	3		2										
	4							5					
	5	3		6	2	2	3		1	5			
110 %	1						3					6666.20	2886
	2												
	3										6		
	4	3	2										
	5			6	2	2		5	1	5			

ambiguity set $\mathscr{G}$ is set to 0.5. Similarly, the coefficient of variation of $\mathbf{\Gamma}$ also changes from 60 % to 150 % because the mean of contract demand is unchanged. Fig. 6 shows the relationship between the profit of the liner company and the covariance of contract demand. Similarly, the profit decreases slightly when the covariance of contract demand is greater than that in the benchmark scenario. Overall, the trends shown in Fig. 6 are basically the same as those in the previous subsection, which reflects the changes in the decision-making of the liner company under different levels of uncertainty. However, unlike in the previous subsection, the number of voyages that the liner fleet needs to complete in different scenarios is essentially the same. This shows that when parameter $\mathbf{\Gamma}$ fluctuates, the main way for liner companies to deal with uncertainty is to adjust the size of the liner. We also use Table 6 to show three typical scenarios. As seen in the first two scenarios in Table 6, as uncertainty increases, the liner company responds by adjusting the class of the liners. However, the average capacity of each liner has not changed much. When $\mathbf{\Gamma}$ becomes the benchmark value, the liner company instead chooses to introduce more liners to cope. It can be seen that when liner companies face uncertainty, a reasonable choice is to use a variety of tools to strategically adjust their fleets to adapt to unpredictable market conditions.

7.2.4. Sensitivity analysis of liner capacity CAP_s

In recent years, there has been a discernible trend toward larger liners. To increase the economies of scale in freight transportation, the capacity of liners will be further expanded in the future. Therefore, in this subsection, experiments are conducted to analyze the influence of the liner capacity on the calculation results. The means and covariances of both total demand and contract demand are set to be the same as those in the benchmark scenario, the radius ε is calculated through the method in Section 7.1, and the radius $\widetilde{\epsilon}$ is set to 0.6. The liner capacity is set to be in the range of 30 % to 120 % of the benchmark value with 10 % as the interval. Fig. 7 shows the results of the **R-P-1** model with different liner capacity settings. The profit of the liner company increases when the liner capacity increases, while the number of voyages remains basically unchanged, but there are certain fluctuations. The aforementioned changes are consistent with our original predictions. Increasing parameter CAP_s is equivalent to reducing the operating cost of each liner. Therefore, Table 7 shows that liner companies are increasingly willing to use larger liners. Doing so can help the company cope with uncertainty without incurring excessive costs. In Table 7, when parameter CAP_s doubles, the average size of the liner in the fleet also doubles. This also reflects the impact of technological innovation on liner companies. As liner operating costs decline, the cost of liner companies responding to uncertainty gradually decreases.

7.3. Out-of-sample analysis

As mentioned in Sections 1 and 2, the joint optimization of LFDP and SAP is usually solved via robust optimization or SP methods in existing studies. However, the results of robust optimization are often overly conservative because they optimize the worst case determined by the range of uncertain parameter values. When the worst case is poorly chosen (e.g., an extreme case is chosen as a sample), it often leads to a significant waste of resources. In addition, solving the SP model usually involves approximating this model via a deterministic problem with a limited number of scenarios (i.e., the SAA approach) on the basis of the assumption that the distribution function of uncertain parameters is known. However, the uncertain parameters we need to address (i.e., transportation demand) are highly coupled with the political and economic forms of the world. This prevents us from accurately identifying distribution functions with uncertain parameters. If the liner deployment decision is made under the wrong probability distribution function, the performance of this decision under the true probability distribution may be disappointing. This leads to the *optimization*

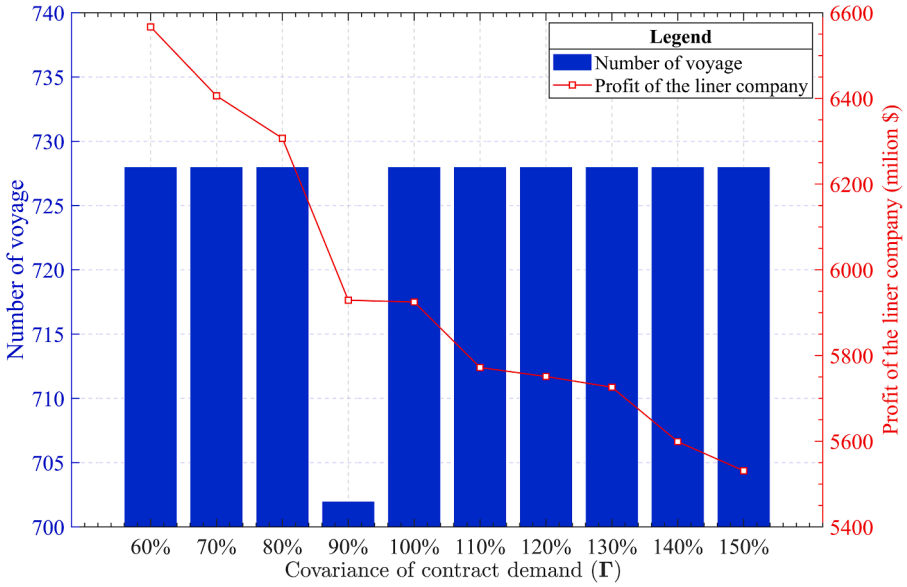


Fig. 6. Sensitivity analysis of the covariance of contract demand $\mathbf{\Gamma}$.

Table 6
Liner fleet deployment scheme under different covariances of contract demand **Γ**.

Γ	Liner class	Number of liners										Average capacity (TEU)	Total profit (million $)
		R1	R2	R3	R4	R5	R6	R7	R8	R9	R10		
60 %	1										7	6410.84	6567
	2												
	3						4						
	4							5					
	5	3	2	7	2	2			1	5			
70 %	1											6486.37	6406
	2										7		
	3						4						
	4							5					
	5	3	2	7	2	2			1	5			
100 %	1										7	6014.77	5925
	2						5						
	3		2										
	4							5					
	5	3		7	2	2			1	5			

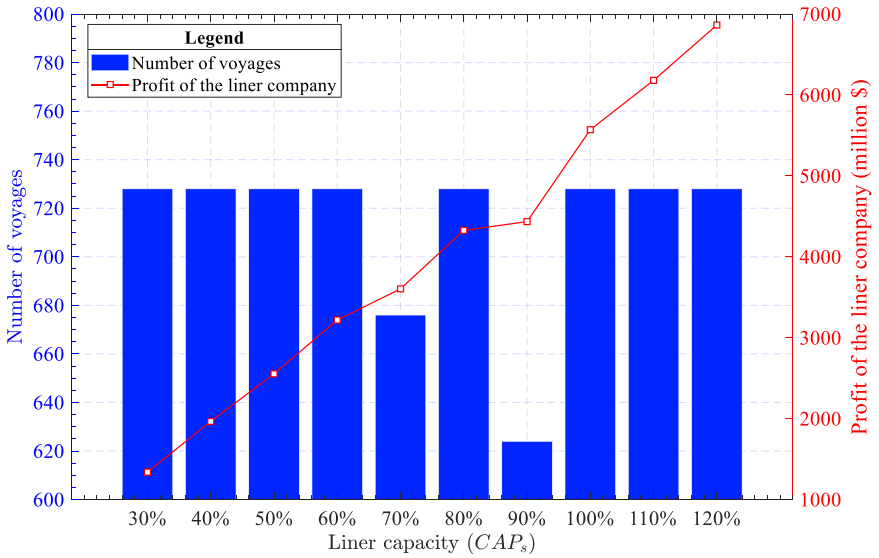


Fig. 7. Sensitivity analysis of liner capacity CAP_s.

Table 7
Liner fleet deployment scheme under different liner capacities CAP_s.

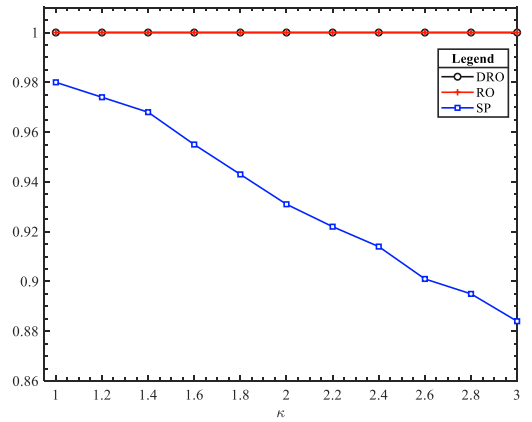
CAP_s	Liner class	Number of liners										Average capacity (TEU)	Total profit (million $)
		R1	R2	R3	R4	R5	R6	R7	R8	R9	R10		
60 %	1										7	3846.51	3218
	2												
	3						4						
	4							5					
	5	3	2	7	2	2			1	5			
100 %	1										7	6014.77	5571
	2						5						
	3		2										
	4							5					
	5	3		7	2	2			1	5			
120 %	1										7	7467.98	6864
	2												
	3		2				4						
	4							5					
	5	3		7	2	2			1	5			

bias phenomenon, where optimization on the basis of imperfectly distributed estimates leads to biased decisions. The abovementioned defects exposed by SP and robust optimization models (i.e., over-conservatism and optimization bias) can be empirically explored through out-of-sample analysis. Therefore, in this section, we empirically validate the benefits of the DRO approach by conducting out-of-sample analysis.

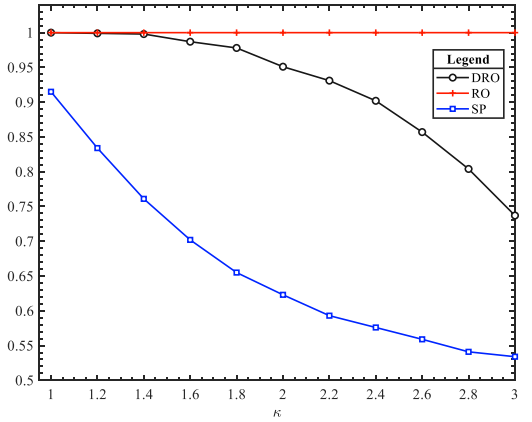
The steps of out-of-sample analysis can be summarized as follows. First, we solve the robust optimization model and the SP model to realize the joint optimization of the LFDP and SAP. Specifically, the method for solving the SP model is based on Meng et al. (2012). We generated 20 scenarios via the normal distribution function and the parameter setting scheme in Section 7.1. To handle the joint chance constraint (15) in the model, we apply the technique proposed in Section 4.1 of Wang et al. (2013). The parameters of the robust optimization model are derived from the abovementioned 20 scenarios. Secondly, two distribution functions are considered, and 10,000 test samples are randomly generated according to each distribution function. The reason for choosing these two distribution functions is that they are distribution functions that the uncertain parameters are usually assumed to obey in existing studies. Finally, we substitute each test sample into the joint chance constraint (15), where the values of the decision variables in this constraint are obtained by solving different models (i.e., robust optimization model, SP model and DRO model) to test the out-of-sample performance of different models. We use κ (a non-negative parameter) to control the degree of out-of-sample uncertainty and the distribution can be written as follows.

(1) *Truncated normal distribution* (TNO): For each O-D pair $w \in \mathscr{W}$, we set the random container transportation demand from the contract shippers of each week $\vartheta_w \sim N(\mu_w,(\kappa\sigma_w)^2)$ with non-negative marginal.
(2) *Truncated uniform distribution* (TUN): For each O-D pair $w \in \mathscr{W}$, we set the random container transportation demand from contract shippers of each week $\vartheta_w \sim U\left(\mu_w - \kappa\sqrt{3}\,\sigma_w, \mu_w + \kappa\sqrt{3}\,\sigma_w\right)$ with non-negative marginal.

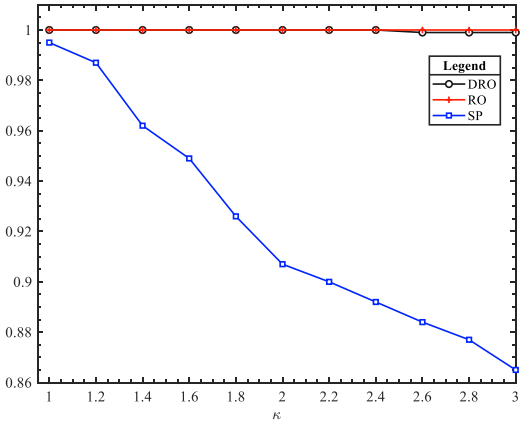
In the out-of-sample analysis, two indicators are employed for the evaluation of the out-of-sample performance of different models: average individual probability (AIP) and average joint probability (AJP). For a test sample, the former counts each constraint in the



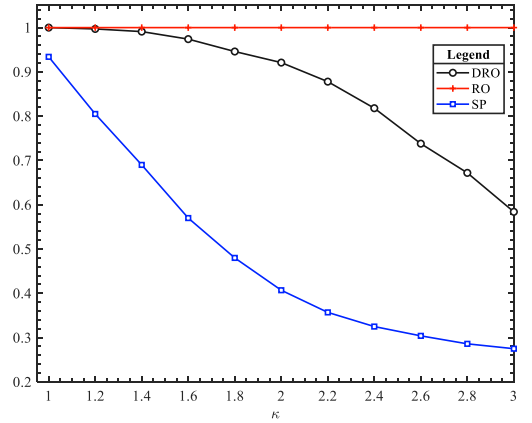
(*a*) AIP under truncated normal distribution



(*b*) AJP under truncated normal distribution



(*c*) AIP under truncated uniform distribution



(*d*) AJP under truncated uniform distribution

Fig. 8. Out-of-sample performance of different models with two distribution functions.

joint chance constraint (15) and takes a value of 1 if one constraint is not violated. In contrast, for each test sample, the latter takes a value of 1 if all the constraints within the joint chance constraint (15) are not violated. The calculation results of these two indicators are within the interval [0, 1]. The larger the value of the indicator is, the better the out-of-sample performance of the model. The calculation methods of AIP and AJP are as follows:

$$AIP = \frac{1}{N\sum_{r\in\mathscr{R}}(|\mathscr{L}_r||\mathscr{R}|)}\sum_{n\in[N]}\sum_{l\in\mathscr{L}_r,r\in\mathscr{R}}\mathbb{I}\left(\sum_{s\in\mathscr{S}}n_{sr}^{\mathrm{VOL},*}CAP_s \Big/ \left\lceil\frac{T}{7}\right\rceil \geq \sum_{w\in\mathscr{W}}\widehat{\vartheta}_w^n\sum_{h^w\in\mathscr{H}^w}q_{h^w}^*\delta_{h^w}^{lr}\right) \tag{52}$$

$$AJP = \frac{1}{N}\sum_{n\in[N]}\mathbb{I}\left(\sum_{s\in\mathscr{S}}n_{sr}^{\mathrm{VOL},*}CAP_s \Big/ \left\lceil\frac{T}{7}\right\rceil \geq \sum_{w\in\mathscr{W}}\widehat{\vartheta}_w^n\sum_{h^w\in\mathscr{H}^w}q_{h^w}^*\delta_{h^w}^{lr}, \qquad \forall l\in\mathscr{L}_r, r\in\mathscr{R}\right) \tag{53}$$

where $\widehat{\boldsymbol{\vartheta}} = (\widehat{\vartheta}_w^n)$ ($n \in [N]$) represents N test samples; $\mathbb{I}(\cdot)$ denotes the indicator function; and $n_{sr}^{\mathrm{VOL},*}$ and $q_{h^w}^*$ are optimal solutions obtained by solving different models (e.g., robust optimization model, SP model and 2-DRO).

We set the parameter κ to range from 1 to 3 with an interval of 0.2. The results of the out-of-sample performance are shown in Fig. 8 and Table 8. In Fig. 8, the combination of two distribution functions and two indicators resulted in four scenarios (i.e., Figs. 8(*a*) to 8 (*d*)). The three lines DRO, RO and SP represent the out-of-sample performance of the 2-DRO, the robust optimization model and the SP model under different κ values, respectively. We find that in all the scenarios, the values of the two indicators under the robust optimization model are 1. This shows that regardless of how the parameter κ changes, the robust optimization solution does not violate the constraints under out-of-sample data. Although this shows that the solution obtained by solving the robust optimization model can adapt to changes in the external environment (i.e., uncertain parameters), it also shows that the solution obtained is relatively conservative. The SP model performs the worst, and many constraints begin to be violated as κ increases. When κ increases to 3, the AJP of the SP model is <55 % under the truncated normal distribution and 30 % under the truncated uniform distribution, respectively. This shows that when there is a gap between the true distribution and the estimated distribution, the SP model exhibits obvious *Optimization Bias* phenomenon. Finally, we can observe that the out-of-sample performance of 2-DRO is between that of the robust optimization model and the SP model. When κ equals 3, the AJP indicators under the truncated normal distribution and the truncated uniform distribution still exceed 70 % and 50 %, respectively. The above experimental results empirically verify that the 2-DRO we proposed is robust. It can avoid the production of overly conservative solutions and the *optimization bias* caused by the SP model. Table 8 presents the values of each point in Fig. 8.

7.4. Managerial insights

According to the results of the above numerical experiments, several managerial insights into the deployment of liners and slot allocation can be provided to the operations managers of liner companies.

(1) *Multiple liner classes can help increase profit and realize adequate reliability*. As discussed in Section 7.2, when various uncertain parameters change, the liner company fleet liner class also changes simultaneously. Multiple liner classes mean that the liner company has more choices to adapt to the different scenarios in practical operations, which means that more strategies are available when facing uncertainty. Furthermore, more investment in the research and development of new maritime technologies may be an essential tool for liner companies to cope with the shock of uncertainty in the future.

(2) *The adjustment of the existing scheme is the key to addressing various demand scenarios*. We find that a change in the number of liners is not necessary under different parameter settings, while the adjustment of existing deployment and slot allocation schemes is the core measure for coping with the uncertainty of container demand. Consequently, liner companies should not expect to solve the problems resulting from uncertainty by purchasing more liners. Today, as liners become larger, liner companies should

Table 8
AIP and AJP values under two distribution functions.

κ	DRO				RO				SP			
	TNO		TUN		TNO		TUN		TNO		TUN	
	DIP	DJP	DIP	DJP	DIP	DJP	DIP	DJP	DIP	DJP	DIP	DJP
1.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.980	0.915	0.995	0.934
1.2	1.000	0.999	1.000	0.997	1.000	1.000	1.000	1.000	0.974	0.834	0.987	0.805
1.4	1.000	0.998	1.000	0.991	1.000	1.000	1.000	1.000	0.968	0.761	0.962	0.690
1.6	1.000	0.987	1.000	0.974	1.000	1.000	1.000	1.000	0.955	0.702	0.949	0.570
1.8	1.000	0.978	1.000	0.946	1.000	1.000	1.000	1.000	0.943	0.655	0.926	0.480
2.0	1.000	0.951	1.000	0.921	1.000	1.000	1.000	1.000	0.931	0.623	0.907	0.407
2.2	1.000	0.931	1.000	0.878	1.000	1.000	1.000	1.000	0.922	0.593	0.900	0.357
2.4	1.000	0.902	1.000	0.818	1.000	1.000	1.000	1.000	0.914	0.576	0.892	0.325
2.6	1.000	0.857	0.999	0.738	1.000	1.000	1.000	1.000	0.901	0.559	0.884	0.304
2.8	1.000	0.804	0.999	0.672	1.000	1.000	1.000	1.000	0.895	0.541	0.877	0.286
3.0	1.000	0.737	0.999	0.584	1.000	1.000	1.000	1.000	0.884	0.534	0.865	0.275

pay attention to the management of current resources and be confident in solving uncertain problems through limited liners and voyages. The liner alliance currently popular in the shipping industry is an example. A liner company forms partnerships with other companies to share resources and liner capacity and improve operational efficiency. This allows for better adaptation to market fluctuations and changes in demand.

(3) *Advance preparatory actions against uncertainty are meaningful for liner companies.* From the trend of the variation in the profit with the ambiguity set radius shown in Section 7.2, decisions that do not take uncertainty into account will lead to higher unforeseen costs when uncertainty occurs. It is obvious that there is a larger shock when unaware of uncertainty, whereas the additional cost caused by increasing uncertainty is lower when pre-considering uncertainty is considered. This reveals the value of decision-making with respect to preparedness for various uncertain scenarios in liner operations. Therefore, to reduce unexpected costs, a meaningful measure for liner companies is to address uncertainty in advance. For example, liner companies can establish a sound risk management mechanism to identify and assess risks in a timely manner and take corresponding measures to cope with them.

8. Conclusions and future work

Scientific and reasonable liner fleet design and slot allocation schemes are the keys to improving the profitability of liner companies. However, the severe demand fluctuations in the maritime industry make existing optimization methods (e.g., robust optimization and stochastic optimization) based on known demand distribution functions flawed. In this paper, we abandon the current mainstream modeling ideas and develop a two-stage distributionally robust optimization model (2-DRO) for liner fleet deployment and slot allocation joint optimization problems while considering the uncertainty of container transportation demand. To ensure a certain reliability of the liner fleet, a joint chance constraint ensuring demand satisfaction is introduced into this distribution-free model. Two ellipsoid ambiguity sets are constructed with first- and second-moment data to describe the distribution functions of contract demand and total demand. This 2-DRO is then reformulated as a second-order cone programming model. We also extend this formulation into a special case considering the connection between the ambiguity set radius and risk tolerance, and a customized outer approximation algorithm (OAA) is further proposed to explore the global solution to our 2-DRO. Numerical experiments are comprehensively conducted to investigate the performance of the 2-DRO. To further analyze the impact of various parameters in the model, sensitivity analysis experiments are carried out. The effect of uncertainty is shown to be significant for the liner deployment and allocation scheme. An out-of-sample analysis of three types of models (i.e., the robust optimization model, stochastic programming model, and 2-DRO) empirically verifies that our model avoids generating overly conservative solutions while circumventing the optimization bias phenomenon caused by the stochastic programming approach.

Finally, we summarize several possible extensions for future research on the basis of previous research and this paper. First, we consider a short-term planning horizon in our modeling framework. The dynamic properties of the long-term horizon present new challenges to our model. The combination of dynamic modeling and uncertainty may require new approaches to ensure computability. The construction of an ambiguity set in combination with dynamic features is also a challenge in this direction. Second, more factors related to liner shipping can be introduced into our model. For example, considering liner alliances may be an interesting direction. Third, other ambiguity sets, such as scenario-wise ambiguity sets, may lead to more insights into the operations and management of liners. There are distinct peak and low seasons in maritime transport, as well as temporal and spatial differences in the transportation demands of different commodities. Clustering uncertain transportation demand first via machine learning and other methods and then constructing ambiguity sets based on the clustering results is certainly a valuable direction for future research.

Data availability

Data will be made available on request.

CRediT authorship contribution statement

Tao Zhang: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Shuaian Wang:** Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Conceptualization. **Xu Xin:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Acknowledgements

This work was supported by the National Natural Science Foundation of China [Grant No. 72371221], and the Research Grants Council of the Hong Kong Special Administrative Region, China [Project number HKSAR RGC TRS T32-707/22-N].

References

Agarwal, R., Ergun, Ö., 2008. Ship scheduling and network design for cargo routing in liner shipping. Transp. Sci. 42 (2), 175–196.
Ben-Tal, A., Nemirovski, A., 1998. Robust convex optimization. Mathe. Oper. Res. 23 (4), 769–805.
Ben-Tal, A., Teboulle, M., 1986. Expected utility, penalty functions, and duality in stochastic nonlinear programming. Manage. Sci. 32 (11), 1445–1466.

Bonami, P., Biegler, L.T., Conn, A.R., Cornuéjols, G., Grossmann, I.E., Laird, C.D., Lee, J., Lodi, A., Margot, F., Sawaya, N., 2008. An algorithmic framework for convex mixed integer nonlinear programs. Discr. Optimiz. 5 (2), 186–204.
Chen, K., Yi, X., Xin, X., Zhang, T., 2023. Liner shipping network design model with carbon tax, seasonal freight rate fluctuations and empty container relocation. Sustain. Horiz. 8, 100073.
Chen, W., Sim, M., 2009. Goal-driven optimization. Oper. Res. 57 (2), 342–357.
Chen, W., Sim, M., Sun, J., Teo, C.-P., 2010. From CVaR to uncertainty set: implications in joint chance-constrained optimization. Oper. Res. 58 (2), 470–485.
Christiansen, M., Hellsten, E., Pisinger, D., Sacramento, D., Vilhelmsen, C., 2020. Liner shipping network design. Eur. J. Oper. Res. 286 (1), 1–20.
Delage, E., Ye, Y., 2010. Distributionally robust optimization under moment uncertainty with application to data-driven problems. Oper. Res. 58 (3), 595–612.
Feng, C.-M., Chang, C.-H., 2010. Optimal slot allocation with empty container reposition problem for Asia ocean carriers. Int. J. Shipp. Transp. Logist. 2 (1), 22–43.
Fletcher, R., Leyffer, S., 1994. Solving mixed integer nonlinear programs by outer approximation. Math. Program. 66, 327–349.
Galambos, J., 1977. Bonferroni inequalities. Ann. Probab. 577–581.
Gao, S., Xin, X., Li, C., Liu, Y., Chen, K., 2022. Container ocean shipping network design considering carbon tax and choice inertia of cargo owners. Ocean Coast. Manag. 216, 105986.
Gelareh, S., Meng, Q., 2010. A novel modeling approach for the fleet deployment problem within a short-term planning horizon. Transp. Res. Part E: Logist. Transp. Rev. 46 (1), 76–89.
Gelareh, S., Nickel, S., Pisinger, D., 2010. Liner shipping hub network design in a competitive environment. Transp. Res. Part E: Logist. Transp. Rev. 46 (6), 991–1004.
Jaramillo, D., Perakis, A.N., 1991. Fleet deployment optimization for liner shipping part 2. Implementation and results. Mar. Policy Manage. 18 (4), 235–262.
Ksciuk, J., Kuhlemann, S., Tierney, K., Koberstein, A., 2023. Uncertainty in maritime ship routing and scheduling: a literature review. Eur. J. Oper. Res. 308 (2), 499–524.
Liang, J., Li, L., Zheng, J., Tan, Z., 2023. Service-oriented container slot allocation policy under stochastic demand. Transp. Res. Part B: Methodol. 176, 102799.
Liu, D., Yang, H., 2015. Joint slot allocation and dynamic pricing of container sea-rail multimodal transportation. J. Traff. Transp. Eng. 2 (3), 198–208.
Liu, M., Xin, X., Wang, X., Zhang, T., Chen, K., 2025. Dual-channel slot sales strategy for container liner shipping companies with blockchain technology adoption. Transp. Policy 162, 200–220.
Liu, Y., Xin, X., Yang, Z., Chen, K., Li, C., 2021. Liner shipping network-transaction mechanism joint design model considering carbon tax and liner alliance. Ocean Coast. Manag. 212, 105817.
Luedtke, J., Ahmed, S., 2008. A sample approximation approach for optimization with probabilistic constraints. SIAM J. Optimiz. 19 (2), 674–699.
Maragos, S.A., 1994. *Yield Management for the Maritime Industry* (Doctoral dissertation. Massachusetts Institute of Technology).
Meilijson, I., Nádas, A., 1979. Convex majorization with an application to the length of critical paths. J. Appl. Probab. 16 (3), 671–677.
Meng, Q., Wang, S., 2012. Liner ship fleet deployment with week-dependent container shipment demand. Eur. J. Oper. Res. 222 (2), 241–252.
Meng, Q., Wang, T., 2010. A chance constrained programming model for short-term liner ship fleet planning problems. Mar. Policy Manage. 37 (4), 329–346.
Meng, Q., Wang, T., 2011. A scenario-based dynamic programming model for multi-period liner ship fleet planning. Transp. Res. Part E: Logist. Transp. Rev. 47 (4), 401–413.
Meng, Q., Wang, T., Wang, S., 2012. Short-term liner ship fleet planning with container transshipment and uncertain container shipment demand. Eur. J. Oper. Res. 223 (1), 96–105.
Mourão, M., Pato, M.V., Paixão, A., 2002. Ship assignment with hub and spoke constraints. Mar. Policy Manage. 29 (2), 135–150.
Nemirovski, A., Shapiro, A., 2007. Convex approximations of chance constrained programs. SIAM J. Optimiz. 17 (4), 969–996.
Ng, M., 2014. Distribution-free vessel deployment for liner shipping. Eur. J. Oper. Res. 238 (3), 858–862.
Ng, M., 2015. Container vessel fleet deployment for liner shipping with stochastic dependencies in shipping demand. Transp. Res. Part B: Methodol. 74, 79–87.
Ng, M., Lin, D.-Y., 2018. Fleet deployment in liner shipping with incomplete demand information. Transp. Res. Part E: Logist. Transp. Rev. 116, 184–189.
Perakis, A.N., Jaramillo, D., 1991. Fleet deployment optimization for liner shipping part 1. Background, problem formulation and solution approaches. Mar. Policy Manage. 18 (3), 183–200.
Powell, B., Perkins, A., 1997. Fleet deployment optimization for liner shipping: an integer programming model. Mar. Policy Manage. 24 (2), 183–192.
Rockafellar, R.T., Uryasev, S., 2000. Optimization of conditional value-at-risk. J. Risk 2, 21–42.
Shahabi, M., Unnikrishnan, A., Jafari-Shirazi, E., Boyles, S.D., 2014. A three level location-inventory problem with correlated demand. Transp. Res. Part B: Methodol. 69, 1–18.
Stopford, M., 2008. Maritime Economics 3e. Routledge.
Ting, S.-C., Tzeng, G.-H., 2004. An optimal containership slot allocation for liner shipping revenue management. Mar. Policy Manage. 31 (3), 199–211.
Wang, S., Meng, Q., 2017. Container liner fleet deployment: a systematic overview. Transp. Res. Part C: Emerg. Technol. 77, 389–404.
Wang, T., Meng, Q., Wang, S., Qu, X., 2021. A two-stage stochastic nonlinear integer-programming model for slot allocation of a liner container shipping service. Transp. Res. Part B: Methodol. 150, 143–160.
Wang, T., Meng, Q., Wang, S., Tan, Z., 2013. Risk management in liner ship fleet deployment: a joint chance constrained programming model. Transp. Res. Part E: Logist. Transp. Rev. 60, 1–12.
Wang, T., Tian, X., Wang, Y., 2020. Container slot allocation and dynamic pricing of time-sensitive cargoes considering port congestion and uncertain demand. Transp. Res. Part E: Logist. Transp. Rev. 144, 102149.
Wang, Y., Meng, Q., Du, Y., 2015. Liner container seasonal shipping revenue management. Transp. Res. Part B: Methodol. 82, 141–161.
Weerasinghe, B.A., Perera, H.N., Bai, X., 2023. Optimizing container terminal operations: a systematic review of operations research applications. Mar. Econ. Logist. 1–35.
Xiang, Z., Xin, X., Zhang, T., Chen, K., Liu, M., 2025. Asia–Europe liner shipping network design model considering Arctic route and black carbon tax. Ocean Coast. Manag. 261, 107492.
Xin, X., Wang, X., Zhang, T., Chen, H., Guo, Q., Zhou, S., 2023. Liner alliance shipping network design model with shippers' choice inertia and empty container relocation. Electr. Res. Arch. 31 (9), 5509–5540.
Xin, X., Zhang, T., Wang, X., He, F., Wu, L., 2025a. Risk-averse distributionally robust optimization for construction waste reverse logistics with a joint chance constraint. Comput. Oper. Res. 173, 106829.
Xin, X., Zhang, T., Xiang, Z., Liu, M., 2025b. Battery electric vehicle transportation network robust pricing-infrastructure location model with boundedly rational travelers. Appl. Energy 386, 125606.
Zhang, Z.-H., Li, K., 2015. A novel probabilistic formulation for locating and sizing emergency medical service stations. Ann. Oper. Res. 229, 813–835.
Zhao, Y., Chen, Z., Lim, A., Zhang, Z., 2022. Vessel deployment with limited information: distributionally robust chance constrained models. Transp. Res. Part B: Methodol. 161, 197–217.
Zurheide, S., Fischer, K., 2012. A revenue management slot allocation model for liner shipping networks. Mar. Econ. Logist. 14, 334–361.
Zurheide, S., Fischer, K., 2015. Revenue management methods for the liner shipping industry. Flex. Serv. Manufac. J. 27, 200–223.
Zymler, S., Kuhn, D., Rustem, B., 2013. Distributionally robust joint chance constraints with second-order moment information. Math. Program. 137, 167–198.