The hierarchical multimodal hub location problem for cross-border logistics networks considering multiple capacity levels, congestion and economies of scale

Wang, Zhenjie; Zhang, Dezhi; Tavasszy, Lóránt; Fazi, Stefano

DOI
10.1016/j.tre.2025.103972

Publication date
2025

Document Version
Final published version

Published in
Transportation Research Part E: Logistics and Transportation Review

Citation (APA)
Wang, Z., Zhang, D., Tavasszy, L., & Fazi, S. (2025). The hierarchical multimodal hub location problem for cross-border logistics networks considering multiple capacity levels, congestion and economies of scale. *Transportation Research Part E: Logistics and Transportation Review*, *196*, Article 103972. https://doi.org/10.1016/j.tre.2025.103972

Transportation Research Part E 196 (2025) 103972

Contents lists available at ScienceDirect

Transportation Research Part E

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

The hierarchical multimodal hub location problem for cross-border logistics networks considering multiple capacity levels, congestion and economies of scale

Zhenjie Wang [a,b], Dezhi Zhang [a,*], Lóránt Tavasszy [b,c], Stefano Fazi [c]

[a] School of Traffic and Transportation Engineering, Central South University, 410075, Changsha, Hunan, China
[b] Department of Transport & Planning, Delft University of Technology, Stevinweg1, Delft 2628 CN, The Netherlands
[c] Faculty of Technology, Policy and Management, Delft University of Technology, P.O. Box 5015, 2600 GA Delft, The Netherlands

ARTICLE INFO

Keywords:
Hierarchical multimodal hub location problem
Flow-dependent economies of scale
Congestion
Hub capacity
Metaheuristic

ABSTRACT

The continuous growth of international container trade calls for logistics networks that seamlessly connect cross-border, domestic, and local transport services. In the design of these networks with various hubs and modes of transport, the consideration of both economies of scale for multimodal transport and congestion is essential since they can significantly impact the location of the hubs and their size. Thereby, in this paper, we study these features within a multimodal hub location problem for international trade that considers a hierarchy in the network structure. We develop a mixed-integer linear programming formulation, minimizing infrastructural, operational, and congestion costs. A hybrid adaptive variable neighborhood search algorithm with tailored operators and speed-up strategies is proposed to solve large-scale instances. Numerical experiments are conducted for China's New Western Land-Sea Corridor case and provide new managerial insights for designing hierarchical, multi-modal, cross-border logistics networks.

1. Introduction

The continuous growth of international trade poses challenges to logistics networks regarding improving connectivity and reducing cost (Wei and Dong, 2019). This is particularly relevant in export-focused countries, such as China, where the transportation volumes are more concentrated and logistics and manufacturing centers can be far away from the seaports and border crossings (Tsao and Thanh, 2019). Dry ports are a prominent solution to facilitate multimodal transport and to connect with seaports and cross-borders inland ports in order to cover such large distances and keep transportation costs low (Qiu et al., 2015). However, one of the main challenges within multimodal network designs lies in the trade-off between generating economies of scale by consolidating freight and limiting hubs' congestion. These factors can significantly impact the location of the hubs and their size.

In this paper, we tackle a hierarchical multimodal hub location problem for both export and import transport flows. In particular, we consider a set of origins and destinations of freight located in distinct regions and a set of potential locations for various types of intermediate hubs with corresponding capacity levels and three transportation modes available (road, railway, maritime). Due to the absence of direct links between origins and destinations as well as a fully connected international hub network, the freight follows a certain sequence, i.e., hierarchy, in the type of hubs. Moreover, the flow from one hub can be split and diverted to multiple hubs (henceforth, *multiple allocation*). The aim is to find the optimal locations and capacity levels of different hubs, design the connections,

* Corresponding author.
E-mail address: dzzhang@csu.edu.cn (D. Zhang).

https://doi.org/10.1016/j.tre.2025.103972
Received 22 April 2024; Received in revised form 19 November 2024; Accepted 13 January 2025
Available online 25 February 2025

and assign the freight flows to minimize infrastructural and operational costs. The consolidation of flow will simultaneously cause congestion at hubs and generate flow-dependent economies of scale on inter-hub arcs, which need to be balanced. We name this problem the hierarchical multimodal hub location problem with congestion and economies of scale (HMHLP-C-EOS).

Several papers have investigated the hub location problem and its variants. However, to our knowledge, no study simultaneously considers the effect of hub congestion, economies of scale, and multiple allocations in the context of hierarchical multimodal transportation. Distinct efforts can be found in the work of Alkaabneh et al. (2019), investigating a single-allocation hub-and-spoke network system with economies-of-scale and congestion, and in Najy and Diabat (2020) and Khosravi et al. (2024), focusing on a generic hub location problem with multiple allocations. Following the line of their work, we investigate the problem in the context of international multimodal transportation networks by capturing several innovative characteristics. These include (1) Hierarchy: the network structure has three layers and specific hub connecting rules, e.g., the exported cargo needs to be transported through seaports or inland cross-border ports to reach another region, indicating that the cargo should be transported through restricted paths with more than one hub; (2) Multiple transportation modes: the unit transportation cost varies with different modes and depends on the level of consolidation; and (3) Multiple hubs: the capacity setting of various types of hubs influences hub congestion costs, which typically have a nonlinear relation with the used hub capacities (Elhedhli and Wu, 2010; Bayram et al., 2023). Furthermore, the establishment costs of different hubs also impact the network design.

To tackle this extended hub location problem, we develop a mixed-integer linear program (MILP) using linearization techniques for the non-linearity of the cost function, making the model suitable for commercial solvers. To solve the larger-scale instances, we propose a tailored hybrid adaptive variable neighborhood search algorithm. The algorithm includes a network construction phase and a local search phase, both integrating a MILP solver to solve the flow assignment sub-problem. Speed-up strategies are integrated to reduce the computing effort. Numerical experiments first compare the performance of the MILP and the algorithm. Furthermore, real-world instances derived from the New Western Land-Sea Corridor in China are generated to provide a practical setting and managerial insights.

All in all, the contribution of this paper is threefold:

- We introduce and model a new hierarchical multimodal hub location problem with congestion and flow-dependent economies of scale (HMHLP-C-EOS) for cross-border logistics networks. Differently from the related studies, this work investigates the EOS and congestion in a hierarchical network topology containing multiple types of hubs and their capacity levels, as well as multiple flow allocation strategies. We linearize this problem into a tractable mixed integer linear programming formulation.
- We develop a tailored hybrid adaptive variable neighborhood search (HAVNS) algorithm with customized operators and speed-up strategies. This algorithm provides an efficient solution for solving large-scale HMHLP-C-EOS.
- A real-world case study on the New Western Land-Sea Corridor in China shows the applicability of the model and helps generate managerial insights for designing hierarchical multimodal hub networks. Primarily, the work confirms that the model enables decision-makers to optimize hub locations, capacities, and network configurations for balanced transport cost reduction and effective congestion management.

This paper is structured as follows. In Section 2, we review the related literature and position our study. The problem statement and the mathematical formulations are presented in Section 3. In Section 4, we describe the algorithm. Section 5 shows the numerical experiments, and in Section 6, we discuss the results and draw managerial insights. Finally, the conclusions and the directions for future research are presented in Section 7.

2. Literature review

The Hub location problem (HLP) tackles both the location of hubs and the network design decisions, including link selection and flow distribution. The objectives either focus on cost or service level, and the constraints vary with the requirements of real-world systems. From the seminal work of O'Kelly (1986, 1987) proposing continuous and discrete models, HLPs have been broadly explored. For a summary of the main developments of HLPs, we refer the readers to the review papers of Alumur and Kara (2008), Campbell and O'Kelly (2012), and Farahani et al. (2013); for a more recent comprehensive review, we refer to Alumur et al. (2021) and Saldanha-da-Gama (2022).

The need to go beyond the classical HLP and build models that better reflect real-world characteristics has been emphasized by several researchers (Alumur and Kara, 2008; Alumur et al., 2021; Zhou et al., 2023). As an important extension of HLP, the hierarchical hub location problem (HHLP) investigates the situations where hubs are arranged in a hierarchical structure, which determines the sequence of visits (Contreras and O'Kelly, 2019). Research on the application of the HHLP can be found in telecommunication networks (e.g., Yaman, 2009) and also in passenger transport networks (e.g., Real et al., 2018; Zhong et al., 2018; Torkestani et al., 2018).

In the context of freight transportation, Alumur et al. (2021) studied a hierarchical multimodal hub location problem (HMHLP) with a three-level hub network in a cargo delivery system. With time-defined deliveries and multiple modes, they proposed a MILP to minimize total costs. Dukkanci and Kara (2017) studied a multimodal hub-covering problem with service times in a hierarchical network and developed a heuristic solution algorithm to solve the problem. Shang et al. (2020) introduced a stochastic HMHLP modeling framework for cargo delivery systems to handle the uncertainty in construction cost and travel time. Eydi et al. (2022) tackled the HMHLP in a deterministic environment and developed an exact and heuristic solution method. Shang et al. (2021) tackled a bi-objective hierarchical multimodal hub location problem to minimize the total costs and the maximum delivery time in the system. Li et al. (2023) contributes to the hierarchical hub location problem in the context of urban and rural logistics. Irawan

et al. (2024) studied a dry port hub network design problem considering uncertain demand. A two-stage stochastic program model was developed to minimize the total costs and carbon emissions by deciding the location of feeder and hub dry ports.

All the above papers consider that each node of the network can be linked to at most one other node (*single allocation*). Few studies model multiple allocations in the HMHLP, which is a more realistic setting. Wei and Dong (2019) optimized the freight flow allocations in a cross-border logistics network with multiple allocations and transport modes and proposed an adaptive-weight Genetic Algorithm to solve the problem. However, they considered the hub locations as given. Ma et al. (2020) tackled a hierarchical multimodal hub location problem with time restrictions for the China Railway Express network, where the demand nodes can reach the rail hubs and also the gateway nodes with multiple allocations. In this study, the location of rail and gateway hubs and the allocation of flow must be determined. Li and Wang (2023) proposed a MILP to determine the location of the newly-built secondary hubs in a hierarchical multimodal hub network considering carbon emissions, while the locations of primary hubs are pre-determined. Ma et al. (2023) designed an entire network for the cross-border multimodal container transport system based on inland ports with the consideration of uncertainty, which was converted into an equivalent deterministic formulation and solved by commercial solvers. However, none of these studies capture the effects of flow consolidation, which can have a large impact on the cost function, especially in international cross-border logistics settings due to the large volume of freight.

With respect to the economies of scale, these can be achieved through flow consolidation on arcs between hubs. The economies of scale (EOS) in the hub location problem (HLP) are generally modeled in two ways (Rostami et al., 2022). The first way is to model EOS as a fixed discount factor ρ between hub–hub connections. The second, more nuanced approach incorporates a flow-dependent discount factor, offering a more realistic depiction of how cost savings increase with higher flow volumes. O'Kelly and Bryan (1998) first considered the flow-dependent economies of scale using a piecewise function in the hub location problem, which achieves an accurate approximation using a simple linear model. Hereafter, a large body of work followed this line and extended it under different variants, and we refer the reader to Alumur et al. (2021) and references therein for an overview of modeling economies of scale in the HLPs. We use the piecewise function in this work due to its established use in HLP literature, computational efficiency, and flexibility in modeling cost savings across different flow levels. More recent studies can be found in, for example, Ghaffarinasab et al. (2023) introduced a risk-averse single allocation hub location problem with flow-dependent economies of scale and solved it with an efficient solution algorithm based on Benders decomposition and scenario grouping. Freitas et al. (2023) addressed a single allocation hub location problem with heterogeneous economies of scale in each arc. A General Variable Neighborhood Search meta-heuristic was proposed to handle the large instances.

Concerning congestion at hub nodes caused by flow consolidation, Alkaabneh et al. (2019), Najy and Diabat (2020) and Khosravi et al. (2024) suggested that it should also be considered when capturing the positive effect of economies of scale in HLPs. In the literature, one way to avoid congestion is to impose explicit capacity constraints in the model (see, for example, Yaman and Carello, 2005; Contreras et al., 2009), which do not adequately reflect the exponential behavior of the congestion effect. Another way is to penalize the congestion cost in the objective function in different ways, in which two types of congestion cost functions are commonly used to capture the hub flow congestion HLPs (Alkaabneh et al., 2019; Bayram et al., 2023). In the first class, congestion is modeled as a convex flow function; see, for example, Elhedhli and Hu (2005), de Camargo and Miranda (2012) and Özgün-Kibiroğlu et al. (2019). In the second approach, the congestion function is derived based on the queuing models, which is more realistic since it captures the non-linear relationship between congestion costs and hub capacity Elhedhli and Wu (2010); see, for example, De Camargo et al. (2011), Azizi and Salhi (2022), Bütün et al. (2021), Dhyani Bhatt et al. (2021) and Bayram et al. (2023). We use queuing theory to model the congestion in this study as it is a realistic method for modeling congestion, especially when considering hub capacities.

Efforts to incorporate both congestion and flow-dependent economies of scale simultaneously in the HLP can be found in the work of Alkaabneh et al. (2019) for single allocation, and Najy and Diabat (2020) as well as Khosravi et al. (2024) for multiple allocation. Specifically, Alkaabneh et al. (2019) considered the M/M/1 queue and the power-law function to model the congestion function within a capacitated single allocation p-hub median location problem. A Lagrangian heuristic and a Greedy Randomized Adaptive Search Procedure (GRASP) were developed to solve the problem. Najy and Diabat (2020) investigated an uncapacitated HLP, using a piecewise linearization approach to model the two effects, and devised a Bender decomposition approach to solve the problem. Khosravi et al. (2024) presented a model for hub-and-spoke network design where the congestion is just limited through constraints. The differences between our study and these works are as follows. Firstly, they limited their study to a general hub-and-spoke network with a single transport mode, whereas we tackle a hierarchical multimodal HLP, which brings more complexity since the flows are allowed to go through more than two hubs, and the hierarchy causes more difficulties in finding a feasible solution. Secondly, we also investigate the decision in hub capacity level settings since the effect of congestion is not only related to the volume of flow but also determined by its proximity to the hub capacity, whereas Alkaabneh et al. (2019) considered the hub capacity as a parameter in their extensive analysis. The capacity level setting of different hubs will affect the routing of flow, creating more complexity in analyzing the interaction between the effect of congestion and the economies of scale. Furthermore, we additionally consider the fixed cost of arcs with different modes in order to analyze the trade-off between transport costs and fixed costs since multiple allocation strategies typically result in high fixed costs and complex networks (Hoff et al., 2017).

3. Problem statement and formulation

3.1. Problem description

This section introduces the multimodal hierarchical hub location problem considering hub congestion and economies of scale (EOS) in the context of an international cross-border multimodal logistics network. The network is defined by a set of origin–destination (OD) pairs with containerized cargo demands and a set of different types of potential intermediate hubs, including dry

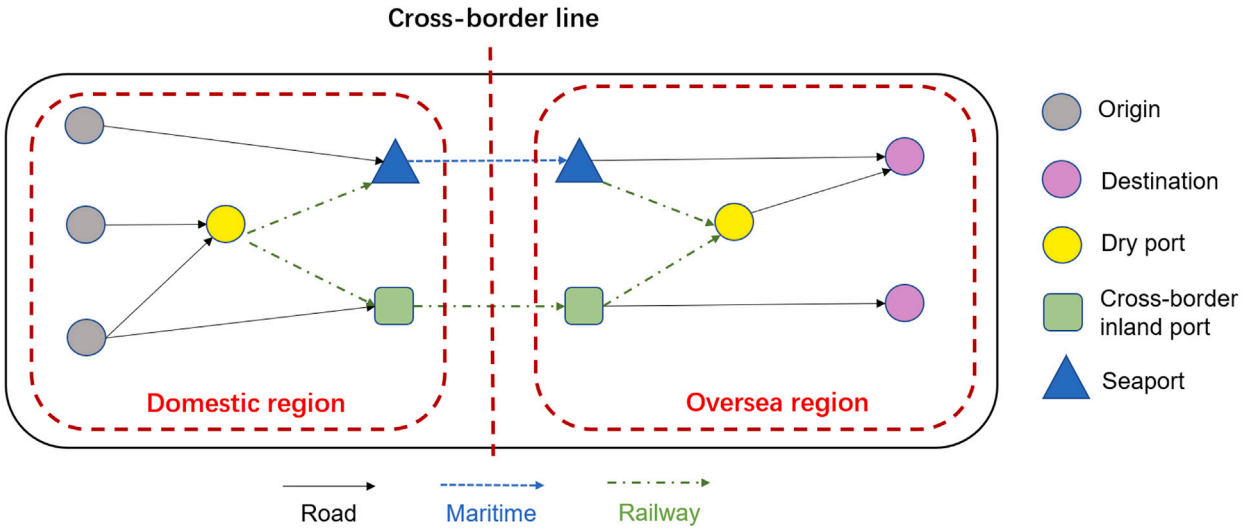


Fig. 1. Export transport process of the international multimodal container transport network.

ports, cross-border inland ports, and seaports. Nodes are distributed both in the domestic and foreign regions. Dry ports are inland hubs that provide consolidation and customs clearance for the import and export containers in the hinterland. A dry port can be used as an inland freight terminal for seaports and also as a starting point for cross-border land transportation for inland regions (Wei and Dong, 2019). Seaports and cross-border inland ports are considered gateways for cargo entering and exiting a region and also provide customs clearance and consolidation services.

The aim of the problem is to design a network that can satisfy the cargo demand at minimum costs. The cost consists of (i) fixed costs of establishing hubs depending on their type and capacity levels, (ii) fixed cost of establishing arcs, (iii) transportation costs between nodes, (iv) customs clearance costs, (v) transshipment cost, (vi) congestion costs in hubs. The decisions address the trade-off between congestion and economies of scale within a hierarchical network, the location of different hubs, their respective capacities, the allocation of demand nodes to hubs, the allocation between hubs, and the assignment of cargo flows.

In the hierarchical system, considering multiple allocations for cargo, the origin and destination nodes (level 3) are allowed to connect to dry ports (level 2) or directly to the seaports (level 1) or the cross-border inland ports (level 1) in the same region by roads either in domestic or overseas areas. Dry ports can be connected to seaports and cross-border inland ports by railway. Seaports can also be connected to seaports in other areas through maritime. On the other hand, cross-border inland ports are connected by railway. Fig. 1 illustrates an example of an export transport process of the international multimodal container transport network in the domestic area, which is symmetrical with the import process in the domestic area.

Considering the export process in the domestic region, there are two ways for a cargo to arrive at the overseas region from its origin with this configuration: (a) With the construction of dry ports, the cargo can be transported firstly from origins to dry ports through roads for consolidation and customs clearance, then to domestic seaports (or cross-border inland ports) via railways, and finally to the overseas area by reaching abroad seaports (or cross-border inland ports) through maritime transport (or via railway). (b) Without going through the dry ports, the cargo can also be transported directly from its origin to seaports (or cross-border inland ports) for consolidation and customs clearance by road and finally to the overseas area by reaching abroad seaports (or cross-border inland ports) through maritime (or through railway). The inland transport routes are symmetrical in the abroad region. (a) is attractive to shippers as the transportation cost can be reduced due to the possibility of consolidating cargo and generating economies of scale (Roso, 2013) by operating railway services between hubs. However, it will also lead to extra transfer costs and establishment costs of hubs with the increase in the number of such hubs. Furthermore, it also gives substantial congestion costs if a proper number of dry ports is not established.

3.2. Formulation

The formulation consists of the following features:

Sets. We consider a network $G = (N, A)$ with N the set of nodes and with A the set of arcs. Let $O \subseteq N$ and $D \subseteq N$ denote the set of demand and destination nodes, respectively, $H \subseteq N$ the set of potential hubs that consists of dry ports DP, cross-border inland ports CB, and seaports SP. Set A is composed of different arc sets: A_{ODP}, A_{OCB}, and A_{OSP} for road services between origins/destinations and dry ports, between origins/destinations and cross-border inland ports, and between origins/destinations and seaports, respectively; A_{DPCB} and A_{DPSP} for railway services between dry ports and cross-border inland ports, and between dry ports and seaports, respectively. The relationship between different arc sets in the model explains the modality change process, which refers to the transportation of containers changing from one mode to another in the hubs after a series of transshipment

operations (e.g., unloading and loading). Besides, we call arcs between hubs as hub arcs, and arcs between demand nodes and hubs are non-hub arcs. Let $Q = Q_{DP} \cup Q_{CB} \cup Q_{SP}$ be the set of capacity levels of different types of hubs. Let K denote the set of OD pairs containing the information about the origin, destination, and demand volume.

Parameters. We denote with W_k the cargo demands of OD pair with $t \in T$. The fixed cost of establishing hub $u \in H$ with capacity level $q \in Q$ is given by F^h_{uq} and the cost of connecting an arc between nodes for $u, v \in N$ is represented by F^a_{uv}. Let c_{uv} be the unit transportation cost between nodes with $u, v \in N$. We define the following transfer costs for mode change: θ_1 between road and railway, θ_2 between road and maritime, and θ_3 between railway and maritime. cu_{uv} is the unit customs clearance cost when cargo is transported from origin $u \in O$ to hub $u \in H$. The capacity of a hub $u \in H$ with a level $q \in Q$ is denoted by C^q_u.

Decision variables. y^q_u is the binary location variable, which equals one if a hub $u \in H$ with capacity level $q \in Q$ is installed, 0 otherwise. x_{uv} is the binary node assignment variable that is equal to 1 if node $u \in H$ is connected with $v \in H$, and 0, otherwise. Let f^k_{uv} be the proportion of demand W_k routed through two nodes $u \in N$ and $V \in N$ in sequence.

Economies of scale. Several scholars modeled economies of scale on the arc between hubs using the piece-wise linear function, see O'Kelly and Bryan (1998), Klincewicz (2002), Najy and Diabat (2020), Azizi and Salhi (2022), and Rostami et al. (2022) for example. In this paper, we adopt a similar linearization method. $|P|$ number of piece-wise linear functions are used to linearize the concave function of economies of scale with the index referring to each linear piece in the set. We introduce binary variables $z^p_{uv} \in \{0, 1\}$ corresponding to whether segment P is chosen and also replace f^k_{uv} with f^k_{uvp} to represent the proportion of demand W_k that is routed through two different hubs $u \in H$ and $v \in H$ in sequence on segment $p \in P$. Therefore, the total cost on the arc between hub $u, v \in H$ is given by a linear function $c_{uv}(W_k f^k_{uvp} l^p_{uv} + b^p_{uv} z^p_{uv})$, where l^p_{uv} and b^p_{uv} are the slope and the intercept of the pth linear function respectively.

Congestion. Congestion is related to the relative difference between hub flows and capacity since the closer the hub flow is to the capacity, the higher the congestion levels. Therefore, the congestion at the hub can be modeled as the ratio of the total flow to the remaining capacity by treating the hub as an M/M/1 queue (Elhedhli and Wu, 2010; Alkaabneh et al., 2019; Bayram et al., 2023). In this paper, we follow this line of reasoning and model the congestion cost ω_u at a hub u as:

$$\omega_u = \psi_u \frac{g_u}{\sum_{q \in Q} C^q_u y^q_u - g_u + \varepsilon} \tag{1}$$

where g_u and $\sum_{q \in Q} C^q_u$ are the newly introduced variables to represent the flow and the capacity at a hub $u \in H$ respectively. ε is an arbitrarily small positive number to avoid a denominator equal to zero, and it can be adjusted to change the ease with which flows approach hub capacity. This is because the rate at which congestion costs increase with the flow becomes slower as ε gets progressively larger. ψ_u is a given congestion cost factor, also providing flexibility to adjust the magnitude of congestion costs to the actual conditions.

Finally, we formulate the problem as a mixed-integer nonlinear programming as follows:

$$\begin{aligned} min \sum_{u \in H} \sum_{q \in Q} y^q_u F^h_{uq} + \sum_{u \in N} \sum_{v \in N} x_{uv} F^a_{uv} + \sum_{k \in K} \sum_{(u,v) \in A_{ODP} \cup A_{OCB} \cup A_{OSP}, u \neq v} W_k f^k_{uv} c_{uv} \\ + \sum_{k \in K} \sum_{u \in H} \sum_{v \in H, u \neq v} \sum_{p \in P} c_{uv}(W_k f^k_{uvp} l^p_{uv} + b^p_{uv} z^p_{uv}) + \sum_{k \in K} \sum_{(u,v) \in A_{ODP} \cup A_{OCB} \cup A_{OSP}, u \neq v} W_k f^k_{uv} cu_{uv} \\ + \theta_1 \sum_{k \in K} \sum_{(u,v) \in A_{ODP} \cup A_{OCB}, u \neq v} W_k f^k_{uv} + \theta_2 \sum_{k \in K} \sum_{(u,v) \in A_{OSP}, u \neq v} W_k f^k_{uv} + \\ \theta_3 \sum_{k \in K} \sum_{(u,v) \in A_{DPSP}, u \neq v} W_k f^k_{uvp} + \sum_{u \in H} \omega_u \end{aligned} \tag{2}$$

Subject to:

$$\sum_{q \in Q} y^q_u \leq 1 \qquad \forall u \in H \tag{3}$$

$$\sum_{v \in H} f^k_{uv} - \sum_{v \in H} f^k_{vu} = 1 \qquad \forall k \in K, u \in O \tag{4}$$

$$\sum_{p \in P} \sum_{v \in H} f^k_{uvp} + \sum_{v \in N} f^k_{uv} = \sum_{p \in P} \sum_{v \in H} f^k_{vup} + \sum_{v \in N} f^k_{vu} \qquad \forall k \in K, u \in H \tag{5}$$

$$\sum_{v \in H} f^k_{uv} - \sum_{v \in H} f^k_{vu} = -1 \qquad \forall k \in K, u \in D \tag{6}$$

$$f^k_{uv} \leq x_{uv} \qquad \forall k \in K, (u,v) \in A_{ODP} \cup A_{OCB} \cup A_{OSP}, u \neq v \tag{7}$$

$$x_{uv} \leq \sum_{q \in Q} y^q_u \qquad \forall u \in H, v \in N, u \neq v \tag{8}$$

$$\sum_{p \in P} z^p_{uv} \leq x_{uv} \qquad \forall u, v \in H, u \neq v \tag{9}$$

$$f^k_{uvp} \leq z^p_{uv} \qquad \forall k \in K, p \in P, (u,v) \in H, u \neq v \tag{10}$$

$$g_u = \sum_{p \in P} \sum_{k \in K} \sum_{v \in H} f^k_{uvp} W_k + \sum_{k \in K} \sum_{v \in O \cup D} f^k_{uv} W_k \qquad \forall u \in H \tag{11}$$

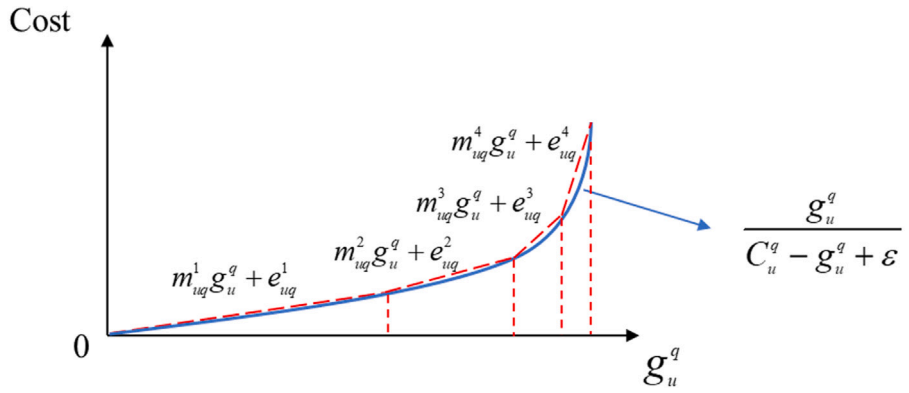


Fig. 2. An illustration of the piecewise linear approximation of congestion cost.

$$g_u \le \sum_{q\in Q} y_u^q C_u^q \qquad \forall u \in H \tag{12}$$

$$y_u^q \in \{0,1\} \qquad \forall u \in H, q \in Q \tag{13}$$

$$x_{uv} \in \{0,1\} \qquad \forall u,v \in N, u \neq v \tag{14}$$

$$z_{uv}^p \in \{0,1\} \qquad \forall u,v \in N, u \neq v, p \in P \tag{15}$$

$$f_{uvp}^k \ge 0 \qquad \forall k \in K, p \in P, u,v \in H, u \neq v \tag{16}$$

$$f_{uv}^k \ge 0 \qquad \forall k \in K, (u,v) \in A_{ODP} \cup A_{OCB} \cup A_{OSP}, u \neq v \tag{17}$$

$$g_u \ge 0 \qquad \forall u \in H \tag{18}$$

The objective function (2) minimizes the total cost consisting of fixed costs of establishing hubs and arcs, transportation cost of flow on the hub and non-hub arcs, cargo transfer costs, customs clearance costs, and hub congestion costs. Constraint (3) restricts the selection of only one capacity level for every hub. Constraints (4)–(6) are the flow balancing equations. Constraints (7) ensure that cargos can only go through the activated non-hub arcs. Constraint (8) imposes that the arc can only be connected to a hub when the hub is installed. Constraints (9) ensure that no segment on the arc between hubs u and v will be active if the arc is not built. Constraint (10) links the flow on segment p to its binary switch variable, which means that the proportion of flow due to segment p will be zero if the corresponding segment is not chosen. The flow at hub u is calculated by constraint (11). Constraint (12) limits the flow at each hub to its capacity. Variables domains are defined by constraints (13)–(18).

3.3. Linearization of formulation

To ease the complexity of the model, we linearize the nonlinear congestion cost term using a piecewise-linear approximation to get a MILP formulation. Following the approach proposed in Bayram et al. (2023), we first transform this cost into a convex function by introducing a variable g_u^q as the total flow through a hub $u \in H$ with capacity $q \in Q$ and replace g_u with it. Through this process, we can define the following convex function:

$$\sum_{u\in H} \omega_u = \sum_{u\in H} \sum_{q\in Q} \Psi_u \frac{g_u^q}{C_u^q - g_u^q + \varepsilon} \tag{19}$$

We also update constraints (11), (12), and (18) with the following ones:

$$\sum_{q\in Q} g_u^q = \sum_{p\in P} \sum_{k\in K} \sum_{v\in H} f_{uvp}^k W_k + \sum_{k\in K} \sum_{v\in O\cup D} f_{uv}^k W_k \qquad \forall u \in H \tag{20}$$

$$g_u^q \le y_u^q C_u^q \qquad \forall u \in H, q \in Q \tag{21}$$

$$g_u^q \ge 0 \qquad \forall u \in H \tag{22}$$

Next, we use a piecewise-linear approximation to linearize the nonlinear congestion cost term shown in Eq. (19); we do this without introducing any additional binary variable, similarly to Najy and Diabat (2020). As an example, we show the piecewise-linear approximation of the convex nonlinear congestion cost function with 4 segments in Fig. 2. A number $|R|$ of piece-wise linear functions is used to linearize the congestion cost with the index r referring to each linear piece in the set. Let m_{uq}^r and e_{uq}^r denote the slopes and intercepts of the rth linear function of hub u with capacity q, respectively. Since the congestion cost of hub u will

take the value of the largest output of any segment, the cost is given by:

$$\psi_u max(m^r_{uq} g^q_u + e^r_{uq}) \tag{23}$$

Thus, the linear congestion cost term ω_u at hub u can be expressed as:

$$min\{\omega_u : \omega_u \geq \psi_u(m^r_{uq} g^q_u + e^r_{uq}), \forall q \in Q, r \in R\} \tag{24}$$

Finally, the total congestion cost term $\sum_{u \in H} \omega_u$ can be calculated by incorporating the following constraints:

$$\omega_u \geq \psi_u(m^r_{uq} g^q_u + e^r_{uq}), \forall u \in H, q \in Q, r \in R \tag{25}$$

4. Solution framework

The hierarchical multimodal hub location problem with congestion and economies of scale (HMHLP-C-EOS) is an NP-hard problem as it is an extension of the hub location problem (HLP) (Alumur and Kara, 2008). Moreover, the multiple allocation and capacity setting contribute to the inherent complexity of the problem, leading to an even larger solution space. MILP solvers can only solve small instances of the studied problem. Hence, a metaheuristic approach is a recommended choice.

From the computational experience of previous studies on hub location problems with multiple allocations, Variable Neighborhood Search (VNS) metaheuristics appear to be a prominent alternative. We refer interested readers to Monemi et al. (2021), Fontes and Goncalves (2021) and Zhang et al. (2023a) for recent efforts. From a preliminary analysis, other available algorithms appear to have shortcomings. For example, studies related to the hub location problem generally apply the Tabu Search algorithm to single-allocation problems, e.g., Bütün et al. (2021). However, preliminary experiments showed that the different features of our problem such as the multiple allocations and the interaction of two effects (economies of scale and congestion) lead to ineffective tabu tables. With respect to Simulated Annealing, we noticed that randomly accepting worse solutions does not lead to fast local optima escapes. Finally, population evolutionary heuristics may lead to too many partial solutions, which can considerably increase the computational effort needed.

Therefore, we propose a customized hybrid algorithm integrating AVNS and a MILP solver for the multiple allocations part of the problem. In particular, the metaheuristic will decide upon the configuration of the network in terms of terminal locations, size, and connections, and the MILP solver will complement the solution by assigning the cargo flows. In contrast to standard VNS algorithm implementations, we develop tailored operators for our problem. First, we implement a path-based expanding strategy with a customized coefficient to add new hubs in network expanding and shaking operators. Second, capacity feasibility checking and repair procedures are activated when border hubs are removed, or their capacity is decreased in a hierarchical hub network. In addition, we activate arcs when implementing operators on hubs, contributing to generating feasible solutions.

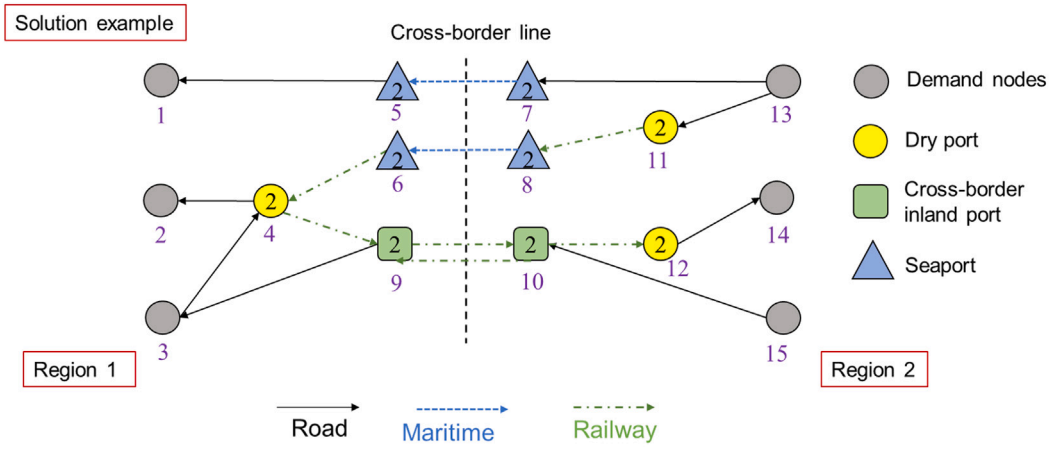
In this section, we describe the proposed algorithm. First, we show the representation of the solution and summarize the framework of the hybrid adaptive variable neighborhood search (HAVNS) algorithm in Section 4.1. Next, the initialization process is presented in Section 4.2, followed by the main algorithm in Section 4.3, which includes a network expanding operator in 4.3.1, a shaking procedure in 4.3.2, a local search procedure in Section 4.3.3, an adaptive operator selection mechanism in 4.3.4, and speed-up strategies in 4.3.5.

4.1. Overall framework of the HAVNS

The basic idea to solve HMHLP-C-EOS is to decompose it into a master problem and a sub-problem. The master problem is the location problem, where we jointly determine which hubs and capacity levels to establish and which arcs to activate. The sub-problem is the flow assignment problem, nested within the master problem and solved exactly with a branch-and-cut algorithm in CPLEX. The flow assignment problem can be easily solved by CPLEX, for small instances, but it can be unbearable for larger ones due to the presence of binary variables related to the activation of economies of scale on hub arcs. Hence, a stopping criteria for the optimality gap GAP_Limit within CPLEX is set to prevent the optimizer from spending too much time proving the optimality of the integer solution. The underlying MILP model is the one shown in Section 3.2, with variables y and x being fixed.

The algorithm works as follows. After the parameters are initialized, an initial feasible solution, S_{init}, is generated by an initialization algorithm described in 4.2 and set as the current solution $S_{current}$ and best-so-far solution S_{HAVNS}. The initialization method is decided by the parameter $Init\_method$. Next, the ALNS algorithm is run for a fixed time T_{max} within a main loop. Let G and $main\_iter_{max}$ be the number of consecutive iterations without improvement and the maximum number allowed of such iterations in the main loop, respectively. If $G \leq main\_iter_{max}$, a so-called hub-added operator N_{hub_add} is applied and generates a new hub network solution S_{hub_add} after launching the subproblem. Otherwise, if $G > main\_iter_{max}$, a shaking procedure is applied to the solution $S_{current}$.

After either procedure, local search operators N_{LS} (see Section 4.3.3) are applied to the new hub network solution in the local search (LS) process to improve the solution further, using either the best-improvement or the first-improvement strategy within. The former looks for the best solution in the neighborhood, whereas the latter stops when a better solution is found. Generally, a first-improvement strategy is implemented for medium and larger instances due to the large solution space. The LS operators are



S=({arr_y[4][2], arr_y[5][2], arr_y[6][2], arr_y[7][2], arr_y[8][2], arr_y[9][2], arr_y[10][2], arr_y[11][2], arr_y[12][2], arr_y[13][2], arr_y[14][2]}, {arr_x[5][1], arr_x[7][5], arr_x[13][7], arr_x[4][2], arr_x[6][2], arr_x[8][6], arr_x[11][8], arr_x[13][11], arr_x[3][4], arr_x[4][9], arr_x[9][10], arr_x[10][12], arr_x[12][14], arr_x[9][2], arr_x[10][9], arr_x[15][10]})

Fig. 3. Representation of a hub network setting (Numbers within the hubs define the capacity level).

launched until a maximum number of iterations without improvement ($inner\_iter_{max}$) is reached, and they are selected based on an adaptive operator selection mechanism (see Section 4.3.4).

After any network change in the LS, we launch the subproblem again to get S_{LS}. From the constructed solution, we remove hubs and arcs without flows, obtaining S'_{LS}. This way, $f(S'_{LS})$ is used to evaluate the solution and possibly update the current best. If the new solution S'_{LS} is better than the current solution $S_{current}$, that is $f(S'_{LS}) < f(S_{current})$, then $S_{current} = S'_{LS}$; otherwise it is rejected. If the current solution is better than the best-so-far solution S_{HAVNS}, then $S_{HAVNS} = S_{current}$. The pseudocode of the main framework is provided in Algorithm 1 in Appendix A.

With respect to solution representation within the algorithm, we use the following notation. For each solution S, we define the tuple $S = (Y, X)$. Y collects all hubs information with the array $arr\_y[u][q]$ which equals 1 if potential hub u with capacity level q is installed, and 0 otherwise. X stores the information about arcs with the array $arr\_x[u][v]$, which equals 1 if node u connects with node v, and 0 otherwise. Fig. 3 shows a hub network example and the representation of the solution associated with this network.

4.2. Initial solution

We propose two initialization greedy algorithms. Initialization 1 generates a large initial network, whereas initialization 2 generates one as small as possible to meet all OD demands.

4.2.1. Initialization 1

The steps of the algorithms are defined as follows:

Step 1: Sort all ODs in descending order based on the demand volume in a list L. Create an empty list D.

Step 2: Load OD pairs to the network sequentially.

(2.1) If list L is empty, go to **Step 3**. Otherwise, select the first available OD pair $k \in L$, and let i be the origin node. Set $i' = i$ as the latest searched node, and $W_k^{i'} = W_k$ as the demand of OD pair k waiting for assignment at node i'. Go to Step (2.2).

(2.2) Find the closest node h to i'. If h is the destination node of k, go to (2.7); otherwise, if it is a hub, go to (2.3).

(2.3) If hub h is not opened, go to (2.4); otherwise, go to (2.6).

(2.4) Establish hub h with capacity level Cap_h equal to the maximum allowed, and set $Cap\_free\_h = Cap\_h$, go to (2.5).

(2.5) If Cap_free_h is larger than the demand $W_k^{i'}$, allocate the demand of i' to hub h and go to (2.2). Let hub h be i', repeat (2.2). Otherwise, allocate as much demand as possible, and set hub h as full ($Cap\_h\_mark = 1$). Add i' to list D along with the remaining demand ($remain\_flow\_i' = W_t^{i'} - Cap\_free\_h$). Let hub h be i', repeat (2.2).

(2.6) Check the free capacity Cap_free_h of h. If $Cap\_free\_h > 0$, go to (2.5); otherwise, find the nearest hub h' to i' other than h, let h' be h, and then go to (2.3).

(2.7) Check list D. If the list is empty, remove k from list L and go back to (2.1) for another OD assignment. Otherwise, select the first i', and go to (2.2)

Step 3: We calculate the objective function based on the network and its flow allocation obtained in **Step 1** and **Step 2**, and output the initial solution S_{init}.

The pseudocode for initialization 1 is presented in Algorithm 2 in Appendix A.

4.2.2. Initialization 2

The hub candidates u other than dry ports are ranked in non-increasing order of a priority index τ_u in list H'. The index is inspired by Bütün et al. (2021), and is calculated as follows:

$$\tau_u = \left(\frac{md_u}{\sum_{u\in H} md_u}\right)\left(\frac{\sum_{u\in H} mc_u}{mc_u}\right) \tag{26}$$

$$md_u = \sum_{i\in O}\sum_{j\in D}\left(\frac{w_{ij}}{d_{iu}+d_{uj}}\right) \tag{27}$$

$$mc_u = \sum_{n\in N} c_{un}d_{un} \tag{28}$$

where w_{ij} is the freight demand volume generated from node i and node j; md_u (modified distance) is the ratio between the demand of all OD pairs and the sum of the distances between the hub u and all OD nodes; mc_u (modified cost) is the product of the distance from the hub u to all nodes n by the corresponding unit transportation cost. Hence, the hub candidates with relatively large modified distances and smaller modified costs are favored.

We check the hub type of the first hub in the list H' (the cross-border inland port or the seaport), and based on τ_u, iteratively add hubs of the same type with the largest capacity to both regions until the sum of the hub capacities on both sides is greater than or equal to the total freight demand. If the capacity does not meet the demand, we follow the same procedure for another hub type. Afterward, demand nodes are connected with hubs in the same regions, and hubs in two different regions are connected bi-directionally. Finally, the subproblem is launched to obtain an initial solution S_{init}.

4.3. Main algorithm

Before entering the local search procedure, the algorithm implements a network-expanding operator or a shaking one after a number of consecutive unsuccessful iterations. In this section, we first explain these two preliminary operators and then describe the local search ones.

4.3.1. Network expanding operator

The network expanding operator N_{hub_add} adds a pre-defined number num_{hubadd} of hubs to the network and is devised to generate more feasible hierarchical connections. The hub candidates u are ranked in non-increasing order of a priority index τ'_u. The index is calculated as follows:

$$\tau'_u = \tau_u * hub\_typ\_reg_u \tag{29}$$

where τ_u is the index derived in (26) and $hub\_typ\_reg_u$ is a coefficient that counts the number of the same type of hubs used in the same region found in the previous solution, which contributes to reflecting which kinds of hubs were more popular on which regions in the last iteration.

Afterward, we iteratively add the hub candidates based on τ'_u and set their capacity to the highest level. To generate more feasible paths, we adopt a path-based strategy to expand the network, which generates paths bi-directionally from the newly added nodes and through the connected hubs towards origins of the same region, direction 1, and towards destinations in other regions, direction 2. Specifically, as shown in Fig. 4, taking a newly added hub h in region 1 as an example, we first connect it to a predefined number $num_{hubexpa}$ of closest hubs h' (e.g., $num_{hubexpa} = 2$ in Fig. 4) in both directions. The arc buildings between h and h' are bi-directional. Next, we iteratively connect *hierarchically* the $num_{hubexpa}$ hubs h' to their closest $num_{hubexpa}$ neighbors, if available, and so on, until reaching either the destination or the origin nodes.

4.3.2. Shaking procedure

To avoid local optima, we devise a shaking procedure to generate a new solution S_{shake}. This operator works similarly to N_{hub_add} with adding new hubs, but we use the roulette-wheel mechanism to iteratively select a number num_{hubadd} of new hubs h based on τ'_u. One other difference is that to further increase the diversity of the solutions, we also target unestablished hubs h'' when implementing the path-based strategy. As shown in Fig. 5, if a hub h' is one of the nearest hubs, we link it with h in both directions. If a hub h'' is one of the nearest hubs, we activate it with the largest capacity level and continue to form the path until it reaches the demand nodes in both directions. This entails that the number of newly added hubs may be larger than num_{hubadd}.

4.3.3. Local search procedure

The local search algorithm takes either S_{hub_add} or S_{shake} as an initial solution, which is processed by CPLEX for flow assignment to get S_{new}. We develop three operators that either remove hubs or change their capacity and that are selected by a roulette-wheel mechanism based on the performance in the last iteration.

Operator 1

Operator N_{LS_1} increases the capacity level of an established hub h that has not reached the highest capacity level. The aim is to both reduce congestion costs and attract flows to generate economies of scale. All the established hubs are first ranked in a non-increasing order based on the congestion cost. After increasing the capacity of hub h, N_{LS_1} can activate arcs to attract more demand. All the unactivated arcs between hub h and other connectable nodes h' are found and ranked in non-decreasing order based on an index that equals the length of the arc multiplied by the unit transportation cost. The first predefined number num_{arcAct} of arcs is activated in two directions from this list. The illustration of the operator is shown in Fig. 6.

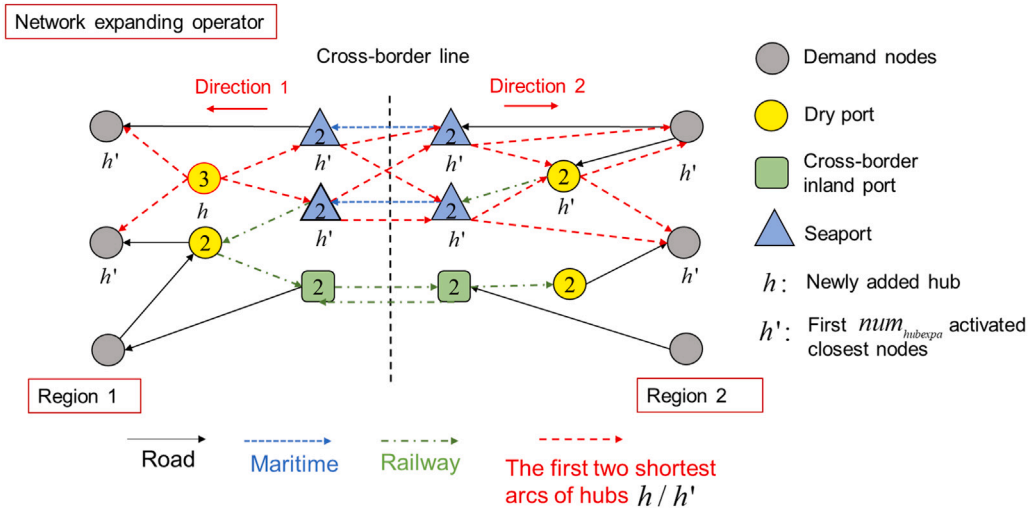


Fig. 4. An illustration of the network expanding operator.

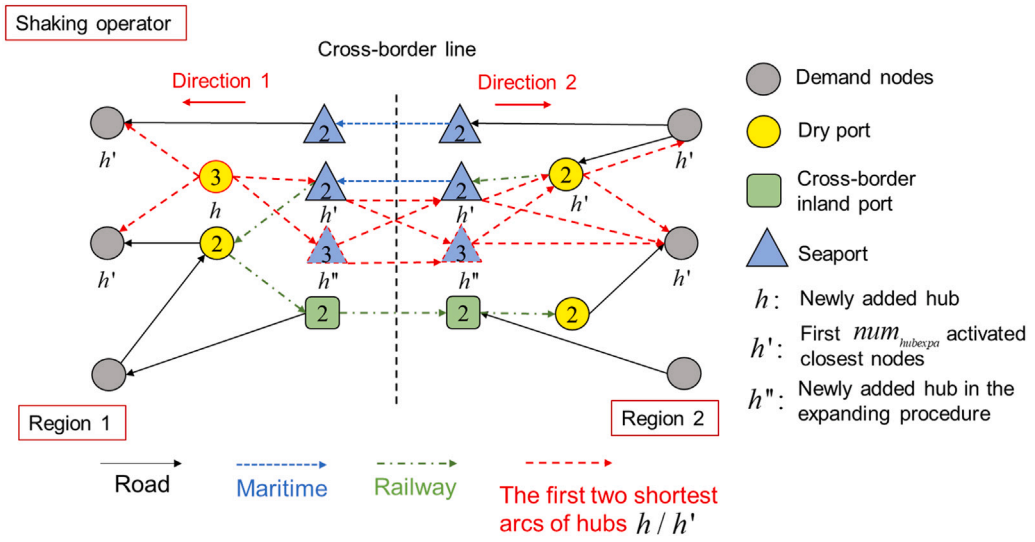


Fig. 5. An illustration of the shaking operator.

Operator 2

Operator N_{LS_2} decreases the capacity level of a hub h that has not reached the lowest capacity level from the current solution. In this case, all the established hubs are ranked in a non-decreasing order based on the congestion cost. To prevent the emergence of infeasible solutions, the following procedures are proposed to check and repair solutions.

Situation 1: hub h is a dry port

(1) Check if the new capacity new_cap_h of the dry port h with reduced capacity levels is greater or equal to the original hub flow g_h. If it is, stop and generate a new solution; otherwise, compute the remaining demand $demand\_rem = g_h\text{-}new\_cap\_h$ and go to (2).

(2) Based on distance, sum one by one the free capacity of other dry ports h' in the same region. Stop summing as soon as the partial sum $total\_freeCap\_h' \geq demand\_rem$. Connect dry ports h' with the inflow nodes h_in and outflow nodes h_out of h to generate a new solution. If $total\_freeCap\_h' \geq demand\_rem$, stop. Otherwise, connect the inflow nodes h_in and outflow nodes h_out of h to provide the possibility of direct transport to cover the remaining demand. See Fig. 7.

Situation 2: hub h is a cross-border inland port

(1) Check if the new capacity new_cap_h of the cross-border inland port h with reduced capacity levels is greater or equal to the original hub flow g_h. If it is, stop and generate a new solution; otherwise, compute the remaining demand $demand\_rem = g_h\text{-}new\_cap\_h$ and go to (2).

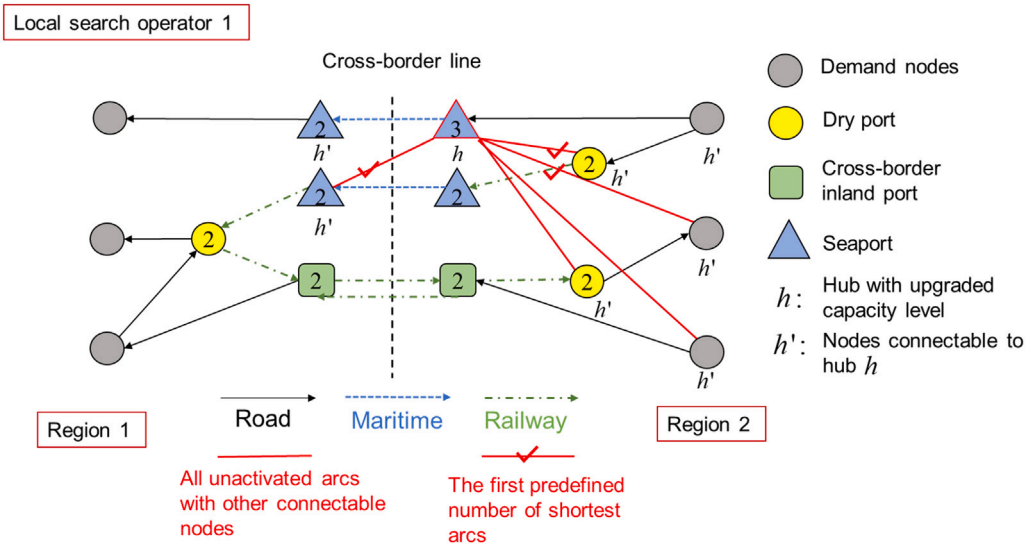


Fig. 6. An illustration of the operator N_{LS_1}.

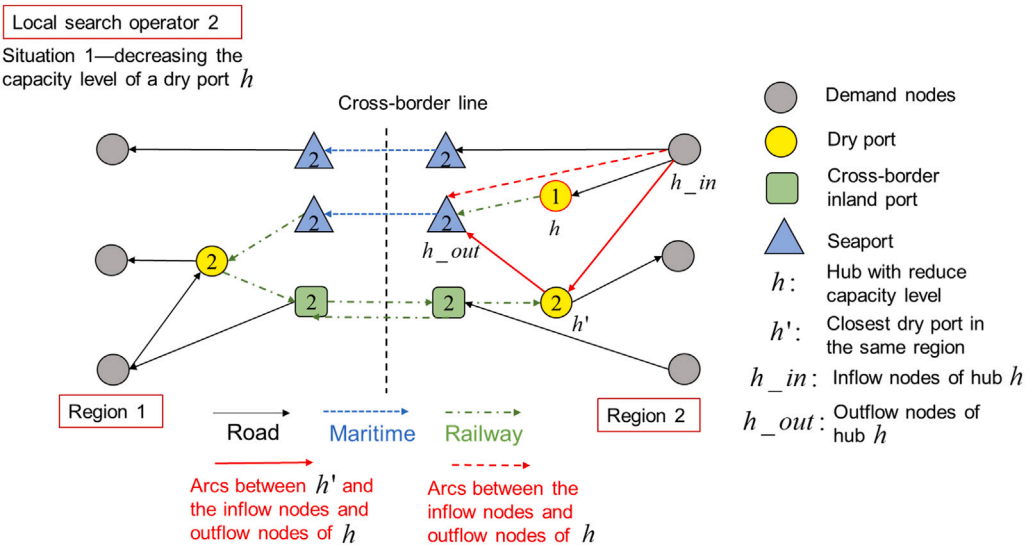


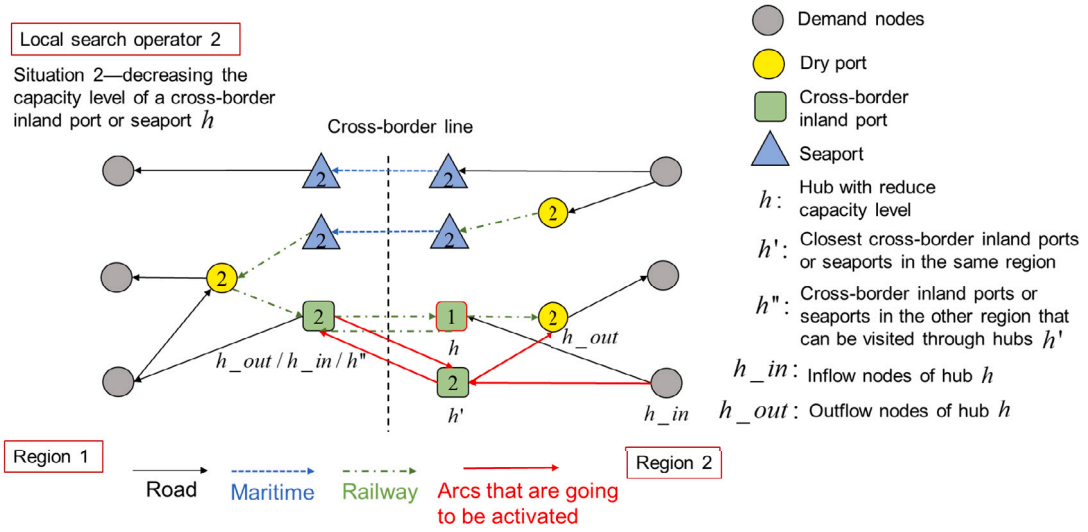
Fig. 7. An illustration of the operator N_{LS_2} in situation 1. (Dry port h capacity is decreased to level 1. The dashed red line is the direct connection between h_in and h_out and is established only if $total\_freeCap\_h' < demand\_rem$.).

(2) Make a list of other cross-border inland ports and seaports h' in the same region. Sort them by non-decreasing distance from h and list cross-border inland ports first. Sum the capacity of all hubs h' in the list or stop summing if the partial sum $total\_freeCap\_h' \geq demand\_rem$. If $total\_freeCap\_h' \geq demand\_rem$, go to (3); otherwise, stop because the solution is infeasible.

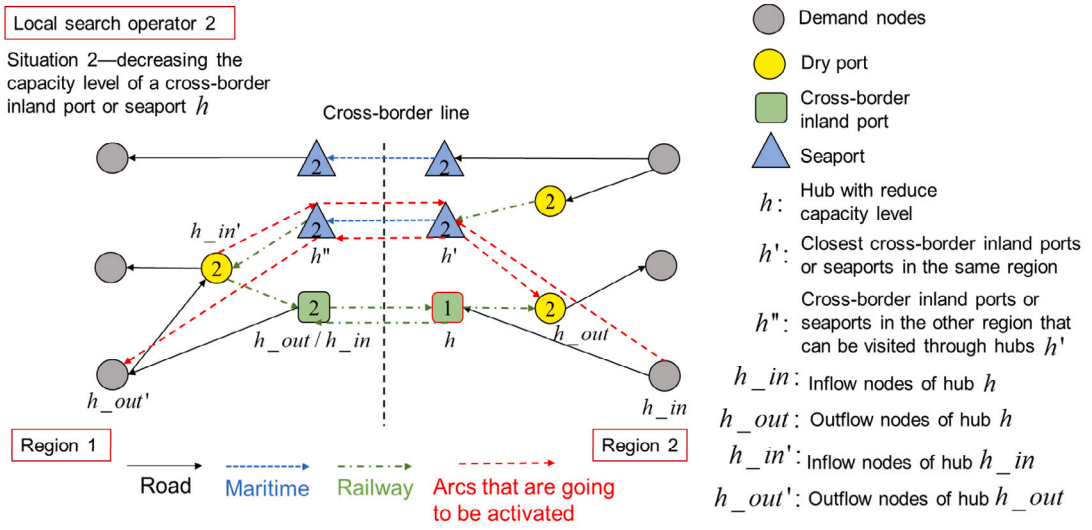
(3) Sum one by one the free capacity of cross-border ports and seaports h'' that can be visited through hubs h' in the other region. Sum the capacity of all hubs h'' or stop summing if the partial sum $total\_freeCap\_h'' \geq demand\_rem$. If $total\_freeCap\_h'' < demand\_rem$, stop because the solution is infeasible. Otherwise, connect hubs h' and h'' and connect all the inflow nodes and outflow nodes (h_in/h_out) of hubs h with h' following the flow direction. See Fig. 8(a). Furthermore, if some hubs h_in/h_out (cross-border or seaports) in the other region are not connectable with h' due to accessibility restrictions, connect h'' with the inflow and outflow nodes h_in'/h_out' of h_in/h_out to provide a feasible solution, see Fig. 8(b).

Situation 3: hub h is a seaport

Take a similar procedure to **Situation 2**, but prioritize checking the total remaining capacity of seaports.



(a) When $min\{total\_freeCap\_h', total\_freeCap\_h''\} \geq demand\_rem$.



(b) When $min\{total\_freeCap\_h', total\_freeCap\_h''\} \geq demand\_rem$ (h_in/h_out in the other region cannot be connected with h'. So, their respective inflow and outflow nodes h_in' and h_out' are connected to h'').

Fig. 8. An illustration of the operator N_{LS_2} in situation 2.

Operator 3

Finally, operator N_{LS_3} removes a hub from the current solution. The order of removing hubs is based on the congestion cost of the established hubs in a non-decreasing order. To prevent the emergence of infeasible solutions after removing the hub, we propose the same procedure of N_{LS_2} to check and repair solutions without step (1) because both the hub h and the connecting arcs are removed.

4.3.4. Adaptive operator selection mechanism

In the LS, we adopt the roulette-wheel principle as the adaptive mechanism to select neighborhoods. Each LS operator i is assigned a dynamic weight γ_i to calculate the selection probability. In particular, we define $\gamma_i = \gamma_i(1-r) + \mu\pi_i$, where π_i is either increased by σ_1 or decreased σ_2, if the solution is better or worse than the previous one, respectively. μ is a fixed reaction factor that controls how quickly the weight adjusts according to the operator's performance in the last iteration.

Table 1
Parameters used in the HAVNS.

Parameter	Description	Values
$main\_iter_{max}$	Maximum successive unimproved iterations in the main loop	3
$inner\_iter_{max}$	Maximum successive unimproved iterations in the local search loop	4
$Strategy$	Best-improvement or First-improvement	1 for small-size instances, 2 for larger instances.
T_max	The maximum running time	1800 s for small-size instances, 3600 s for larger instances, allowing to exceed the time limit to complete the current iteration.
GAP_Limit	A given stopping criteria within CPLEX	0 for small-size, 4% for medium-size and 8% for large-size instances
$Init\_method$	Methods to initialize a solution	1 for small-size instances, 2 for larger instances
num_{arcact}	A predefined number of arcs that need to be activated bi-directionally	4
num_{hubadd}	The number of newly added hubs each main loop iteration	2 for small-size, 3 for medium-size, and 4 for large-size instances
$num_{hubexpa}$	The number of nearest hubs searched in the network extension	3
σ_1	Score for a new best-so-far solution	10
σ_2	Score for obtaining a non-improving solution	1
μ	Reaction factor	0.6

4.3.5. Speed-up strategies

Solving the flow allocation subproblem for every new candidate solution requires a large computational effort. Hence, we propose two speed-up strategies to improve the algorithm's efficiency.

Hash table

We avoid repetitive calculations using a Hash Table that keeps track of the main elements of different solutions, with two columns named "key" and "value". "Key" is the location and capacity level of hubs y_u_q and the arcs activation x_u_v between node u and v. "Value" is the cost provided by y and x. Calling CPLEX will be skipped if a solution with a duplicate value in the Hash table is generated.

Preprocessing heuristic

We design a preprocessing heuristic to fix, for a solution obtained by an operator, some of the variables z_{uv}^{p} to reduce the size of the subproblem. We calculate the possible maximum flow $max\_flow\_arc\_uv$ on the arc from hub u to v since the segment p will not be chosen if its corresponding flow interval is larger than $max\_flow\_arc\_uv$. The procedure is as follows.

(1) Find all hubs h connected with O/D nodes directly and update the possible maximum flow max_flow_node of h as the sum of demand for all O/D nodes connected to this hub. Then update the possible maximum flow of dry ports h_1, and the possible maximum flow $max\_flow\_arc\_uv$ on the arc from dry ports u to cross-border inland ports or seaports v.

(2) Find cross-border inland ports h_2, and seaports h_3 that are not connected with O/D nodes directly but can be reached from dry ports h_1 in the same region. Subsequently, based on the max_flow_node obtained in (1), update the possible maximum flow of h_2 and h_3 by summing the max_flow_node of all h_1 connected to them. Afterward, update the possible maximum flow $max\_flow\_arc\_uv$ on the arc between h_2 and between h_3 based on max_flow_node.

(3) Update the possible maximum flow $max\_flow\_arc\_uv$ on the arc from cross-border inland ports h_2 or seaports h_3 to dry ports h_1 according to $max\_flow\_arc\_uv$ on the arc between h_2 and between h_3 obtained in (2).

(4) Finally, compare the capacity of hubs u and v with $max\_flow\_arc\_uv$, and the smaller value serves as the final possible maximum flow $max\_flow\_arc\_uv$. Then, set variables z_{uv}^{p}, whose corresponding flow interval exceeds the possible maximum flow, to 0.

5. Numerical experiments

We propose two experiments. In the first, presented in Section 5.1, we show the performance of the hybrid metaheuristic against CPLEX using a set of 18 randomly generated instances. In Section 5.2, the second experiment provides managerial insights into the hierarchical multimodal hub location optimization problem using instances inspired by a real-world case study.

The initialization algorithm and the hybrid meta-heuristic are coded in Java, and all the experiments are run on a machine with Intel(R) Core (TM) i5-12400 CPU and 32 GB RAM. ILOG CPLEX 12.6 is employed to solve the MILP formulation and the subproblem within the algorithm.

The hybrid adaptive variable neighbor search (HAVNS) parameters have been determined in preliminary experiments considering the trade-off between solution quality and computational time. The values are summarized in Table 1. HAVNS is applied to each instance 10 times to show the algorithm's stability. For CPLEX, we set the maximum computational time to 1800 s for small-sized instances and 7200 s for medium and large-sized instances.

Table 2
Results on randomly generated instances. For CPLEX, UB, Gap_{lb}, and T_{CP} describe the upper bound to the solution, the gap from the lower bound, and the computational time respectively. For HAVNS, F_{best} is the best solution, and F_{aver} is the average value after all runs. Finally, $Gap_{ub} = (F_{best} - UB)/UB$, $Gap_{aver} = (F_{aver} - UB)/UB$, $Gap_{dev} = (F_{aver} - F_{best})/F_{best}$ and T_{best} is the CPU times for obtaining the best solution. Gap_{300} is the gap between the best solution and the solution after 300 s. The gaps are reported in percentages.

Instance $(O/D, DP, CB, SP, Q)$	CPLEX			HAVNS						
	$UB(kUSD)$	Gap_{lb}	$T_{CP}(sec.)$	$F_{best}(kUSD)$	$F_{aver}(kUSD)$	Gap_{ub}	Gap_{aver}	Gap_{dev}	$T_{best}(sec.)$	Gap_{300}
S1: 10,6,6,6,Tight	90,533	0	762	90,533	90,533	0	0	0	29	0
S2: 10,6,6,6,Loose	83,605	0	728	83,605	83,605	0	0	0	14	0
S3: 20,6,6,6,Tight	275,300	1.61	1800	274,879	274,957	−0.15	−0.09	0.06	62	0
S4: 20,6,6,6,Loose	257,694	0.77	1800	256,719	256,963	−0.38	−0.28	0.09	58	0
S5: 10,10,10,10,Tight	102,623	1.73	1800	102,576	102,921	−0.05	0.18	0.23	153	0
S6: 10,10,10,10,Loose	87,815	2.52	1800	87,377	87 625	−0.50	−0.22	0.28	344	0.51
S7: 20,10,10,10,Tight	315,010	1.54	7200	313,180	313,918	−0.58	−0.35	0.23	1182	1.5
S8: 20,10,10,10,Loose	296,544	0.69	7200	296,121	296,602	−0.14	0.02	0.16	1702	4.76
S9: 40,10,10,10,Tight	1,197,969	11.37	7200	1,173,721	1,176,382	−2.02	−1.80	0.23	2169	7.28
S10: 40,10,10,10,Loose	1,168,615	9.40	7200	1,163,419	1,166,314	−0.44	−0.20	0.25	2563	8.14
S11: 20,20,20,20,Tight	354,191	10.24	7200	342,503	343,521	−3.30	−2.99	0.30	2096	6.31
S12: 20,20,20,20,Loose	323,973	14.66	7200	299,963	302,037	−7.41	−6.76	0.69	2208	6.87
S13: 40,20,20,20,Tight	–	–	Memory	1,460,846	1,473,199	–	–	1.96	3165	16.65
S14: 40,20,20,20,Loose	–	–	Memory	1,444,809	1,467,885	–	–	1.60	3385	22.52
S15: 80,20,20,20,Tight	–	–	Memory	8,700,568	8,797,839	–	–	1.12	3076	23.99
S16: 80,20,20,20,Loose	–	–	Memory	7,756,026	7,925,101	–	–	2.18	3553	15.17
S17: 80,30,30,30,Tight	–	–	Memory	34,273,152	35,384,245	–	–	3.24	3382	13.83
S18: 80,30,30,30,Loose	–	–	Memory	32,914,288	34,126,785	–	–	3.68	3449	14.28

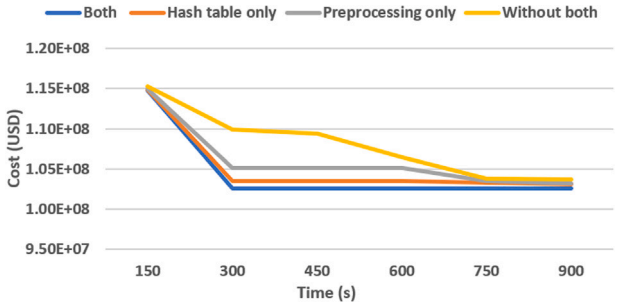


Fig. 9. The performance comparison of speed-up strategies in S5.

5.1. Randomly generated instances experiment

We randomly generate international hierarchical multimodal hub location problem (HMHLP) instances with three transport modes (road, railway, and maritime), four types of nodes (O/D nodes, dry ports, cross-border inland ports, and seaports), where the layout of the nodes between the two regions is symmetrical. The distances between nodes are calculated using the Euclidean distance. To replicate a possible cross-border international logistics network, we set specific rules for the location of nodes and transport processes: nodes are spread on a rectangular region that is further divided into two regions by the so-called "cross-border line"; cross-border inland ports and seaports are distributed along the cross-border line of a region; there is freight demand only between demand nodes located in two different regions to simulate the import and export demand; cross-border transportation is only available between cross-border inland ports or between seaports. A similar network structure can be seen in Fig. 1.

The instances are denoted with the string "$O/D, DP, CB, SP, Q$" to reflect the number of O/D nodes, dry ports, cross-border inland ports, seaports, and the type of hub capacity (tight or loose), respectively. In particular, we generate 18 problem instances denoted as $S1, S2, \ldots, S18$ with 2 levels of Q corresponding to $\{tight, loose\}$. Instances with less than 50 nodes are referred to as "small-size" instances, those between 50 and 100 nodes as "medium-size", and those with more than 100 nodes as "large-size".

Demand for each OD pair is generated randomly between 50 and 150 TEU. We define a based capacity C_{base} for capacity level setting, which is calculated by dividing total demand by the sum of the number of cross-border inland ports and seaports in a single region to guarantee feasible solutions. Based on C_{base}, we set the tight capacity level of dry ports as C_{base}, $2C_{base}$, $3C_{base}$, cross-border inland ports as $2C_{base}$, $3C_{base}$, $4C_{base}$, and seaports as $3C_{base}$, $4C_{base}$, $5C_{base}$. The corresponding loose capacity levels are set as twice the tight capacity levels. The fixed costs of different types and capacity levels of hubs are generated in proportion to the above capacity levels setting, with the base hub fixed cost for C_{base} being 10^6 USD. The variable transportation costs of road, railway, and maritime are set to 100, 10, and 1 USD/TEU.km, and the arc activation costs are set to 500, 5000, and 50 000 USD. The transfer costs of mode changing between road and railway is set to 100 USD/TEU, between road and maritime is 300 USD/TEU, and between railway and maritime is 500 USD/TEU. The customs clearance cost is set to 250 USD/TEU.

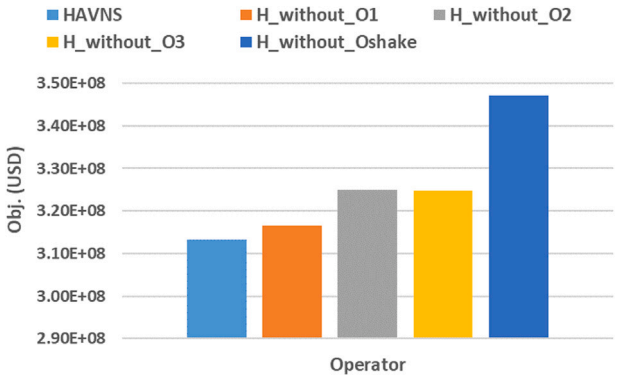


Fig. 10. The performance comparison of operators in S7.

Table 3
The costs and hub capacities parameters.

Parameter	Description	Value(units)
c_{uv}	Unit transport cost for domestic road/abroad road/domestic railway/abroad railway/maritime	1.23/0.97/0.31/0.92/0.21 (USD/TEU km)
θ_1	Unit transfer cost between road and railway	17 (USD/TEU)
θ_2	Unit transfer cost between road and maritime	68 (USD/TEU)
θ_3	Unit transfer cost between railway and maritime	70 (USD/TEU)
cu_{uv}	Unit customs clearance cost at dry ports or cross-border inland ports/seaports	430/500 (USD/TEU)
F^a_{uv}	Fixed cost of operating an arc for road/railway/maritime transport service	700/3000/8000 (USD)
F^h_{base}	Base fixed cost of establishing a hub	50 000 000 (USD)
C_{base}	Base capacity level of a hub	50 000 (TEU/year)
ψ_u	Congestion cost factor	2000

For the congestion cost function approximation, we generate piece-wise linear functions with a maximum error allowed of 20 between the original and linear functions, in which case the number of linear pieces varies with different maximum capacity levels. An average weight error $weighted\_error_{aver}$ is used to calculate the error more objectively and realistically, which can be calculated by:

$$weighted\_error_{aver} = \frac{1}{N}\sum_{i=1}^{N}\frac{F(x_i) - G(x_i)}{G(x_i) + \eta} \tag{30}$$

where $F(x)$ is the piece-wise linear approximation, $G(x)$ is the original congestion function, N is the number we discretize the defining domains to evaluate the error, η is a small constant to prevent division by zero. A preliminary experiment confirmed that the average weighted error is below 5% in all instances, which indicates a high-precision approximation. The congestion cost factor is set to 1000 USD. Regarding economies of scale (EOS), following O'Kelly and Bryan (1998), De Camargo et al. (2009), and Najy and Diabat (2020), we discretize the EOS nonlinear functions using a piece-wise linear function with four segments. A parameter α is introduced to indicate the decrement in the inclination of the function's segments, i.e., $\alpha = 0.2$ means slope $l_1 = 1$, $l_2 = 0.8$, $l_3 = 0.6$ and $l_4 = 0.4$. We do not fix the abscissa points where the piece-wise function slope changes since the effect of EOS should be varied with every instance, aiming to comprehensively analyze the relationship between EOS savings and congestion penalties. Similar to the approach proposed in Najy and Diabat (2020), we compute the value of the y-intercepts for EOS functions as $y_p = p\alpha g_0 + y_{p-1}$, for $p \geq 2$, with $g_0 = \sum_k W_k/h_0$ and h_0 being the number of hubs and set $y_1 = 0$.

The computational results of HAVNS and CPLEX are summarized in Table 2. CPLEX can find the optimal solution for one-third of the small instances (S1–S6) within 1800 s. Gap_{ub} shows that HAVNS can find solutions as good as or better than CPLEX for small instances. The Gap_{aver} with 0.28% the worst case and the Gap_{dev} below 0.28% exhibit the robustness of the HAVNS to get the high-quality solutions. The HAVNS can find the same or even better solution than CPLEX with less computational time in 900 s.

For the medium-sized instances (S7–S12), CPLEX fails to find the optimal solution within 7200 s. For these and the smaller instances, the algorithm can find better solutions ($Gap_{ub} < 0$), with a maximum gap of –7.41% within 3600 s. In addition, in terms of consistency, the algorithm shows a small Gap_{dev} all below 0.69%. CPLEX is out of memory when solving the large instances from S13 to S18 due to the formulation's large memory requirements. The Gap_{dev} values are below 3.68%, showing the robustness of the algorithm.

The last column shows the gap between the best solution found and the best solution after 300 s. In general, we can notice that after this time, the improvement is relatively small, showing that the algorithm is quite fast in reaching good solutions. In addition, we can appreciate that the combined deployment of speed-up strategies has a big impact on performance. As an example, Fig. 9 shows the reduced performance of the algorithm when dropping one or all strategies when solving a randomly picked instance.

To analyze the impact of the operators, a set of small computational experiments is carried out. These experiments consist of removing each operator from the executions of the hybrid adaptive variable neighbor search (HAVNS) while keeping the other

Fig. 11. Distributions of logistics nodes in the New Western Land-Sea Corridor. Note that the yellow nodes are counted as two because they represent the city as both a demand node and a dry port. (Map source: https://www.ldmap.net).

components. The version of the heuristic considering all the components is denoted by HAVNS, and the version without the operators N_i is denoted by $H\_without\_Oi$, i = 1, 2, 3, shake. We do not remove the network expanding operator as it is a basic component. The experiments are conducted in a randomly picked instance, and each version is carried out for 10 runs. Fig. 10 presents the best-found objective value returned by each version of the algorithm. Note that for all versions, the best result is achieved when all operators are used in the HAVNS. Furthermore, removing the shaking operator has the greatest impact on the ability to find best-found results. The presented results demonstrate that all operators are necessary for the computational experiments.

5.2. Real-world instances experiments

To show the applicability of the model and the algorithm, we propose a real-world case study concerning a cross-border multimodal transport logistics network in the New Western Land-Sea Corridor in China. We first describe the setting and the instance generation. Next, the results and the sensitivity analysis are discussed.

5.2.1. Setting

China recently released a plan to build the Western Land-Sea Corridor to provide faster transportation services between China's western inland areas and the surrounding ASEAN region. The cross-border multimodal transport logistics network involves nodes in China and nodes in the ASEAN region. We select candidate nodes, including O/D nodes, dry ports, cross-border inland

Table 4
Optimization results for three scenarios. In the superscript of the hub location result, D stands for dry port, C stands for cross-border inland port, and S stands for seaport. For HA, taking "r(0): 30" as an example, r represents railway transportation (m represents maritime), (0) indicates that the 0th scale economy interval has been reached (0 means that the flow has not yet generated scale economies), and 30 indicates that there are 30 arcs.

Scenarios	Hub selection (capacity)	HA	NHA
EOS only	$42(1)^D$, $46(1)^D$, $50(2)^D$, $58(2)^D$, $64(1)^D$, $68(3)^C$, $69(1)^C$, $71(1)^C$, $72(3)^C$, $73(1)^C$, $75(1)^C$, $76(1)^S$, $79(1)^S$	r(0):30, r(1):2, r(2):1, r(3):3 m(0):0, m(1):0, m(2):1, m(3):1	204
EOS + congestion	$42(3)^D$, $46(2)^D$, $50(2)^D$, $58(2)^D$, $62(3)^D$, $64(1)^D$, $68(2)^C$, $69(1)^C$, $70(1)^C$, $71(2)^C$, $72(2)^C$, $73(1)^C$, $74(1)^C$, $75(2)^C$, $76(1)^S$, $79(1)^S$	r(0):34, r(1):1, r(2):3, r(3):6 m(0):0, m(1):0, m(2):1, m(3):1	184
Congestion only	$42(2)^D$, $45(1)^D$, $46(1)^D$, $47(1)^D$, $49(2)^D$, $50(2)^D$, $53(1)^D$, $55(1)^D$, $57(1)^D$, $58(1)^D$, $62(3)^D$, $64(1)^D$, $68(2)^C$, $69(1)^C$, $70(2)^C$, $71(2)^C$, $72(2)^C$, $73(1)^C$, $74(2)^C$, $75(2)^C$, $76(1)^S$, $79(1)^S$	r(0):90 m(0):2	135

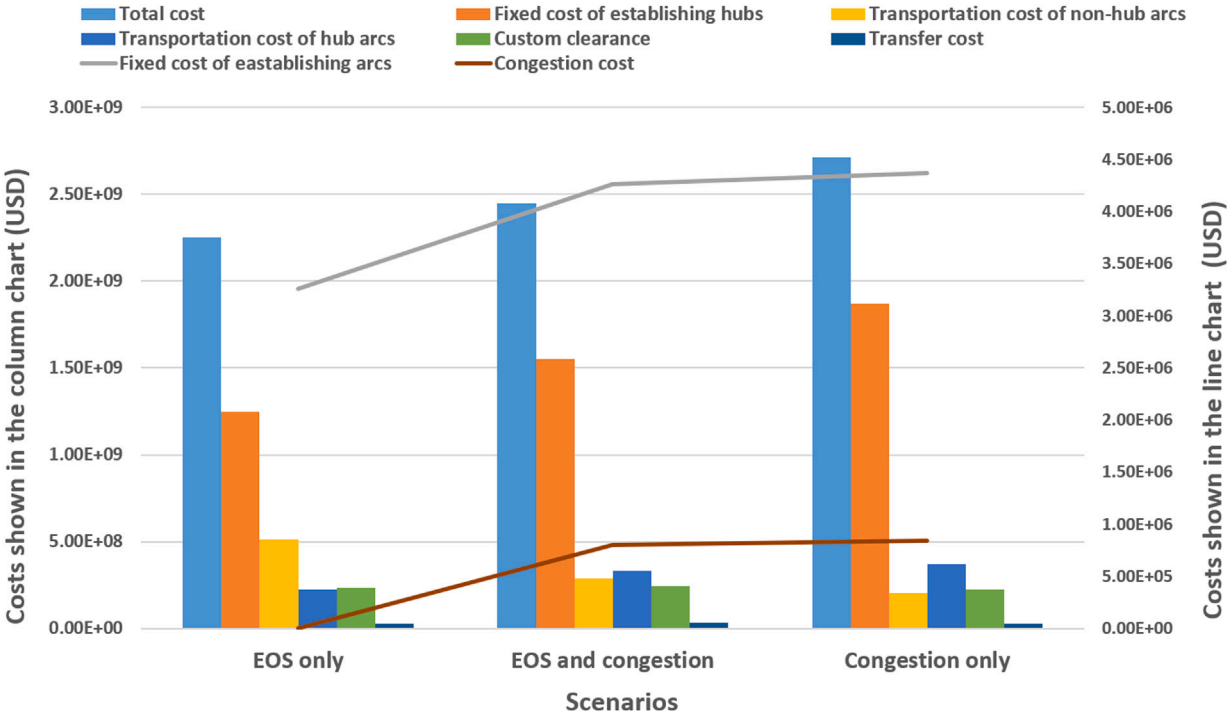


Fig. 12. The changes in the cost components under three different scenarios.

Table 5
The slope of piece-wise linear concave functions.

Function 1	Function 2	Function 3
f_1	f_2	f_3
1.0	1.0	0.8
0.9	0.8	0.6
0.8	0.6	0.4
0.7	0.4	0.2

ports, and seaports, based on the following criteria: (1) cities with operating transport services (road, railway, and maritime) in the New Western Land-Sea Corridor; (2) hub cities presented in the national logistics hub layout planning by the Chinese government (National Development and Reform Commission, 2018); (3) official cooperative operating provinces and cities in the Overall Plan for the New Western Land-sea Corridor (National Development and Reform Commission, 2019). Finally, we select 83 cities as candidate nodes, as shown in Fig. 11. We also copied each cross-border inland port $u \in CB$ into u and u' to be consistent with the model structure; this also makes the results applicable to the example of two regions that do not border and more general settings. The information about nodes in the network is presented in Appendix B.

The international freight demand OD matrix in this study is generated using a gravity model (GM) commonly used in cargo demand estimation (Van Nguyen et al., 2020; Wei and Lee, 2021). The GM is expressed as follows:

$$w_{ij} = GM_i M_j d_{ij}^{-\beta} \tag{31}$$

where w_{ij} is the freight demand volume generated from node i and node j; M_i and M_j are the total trade import and export volume in 2022 in node i and node j, respectively, which can be obtained from the official website of General Administration of Customs of the People's Republic of China (http://www.customs.gov.cn/); d_{ij} represents the distance between node i and node j, which is calculated by the $Vincenty$ formulation based on the coordinates of nodes; $[\beta]$ is the gravity attention coefficient generally set to 2; G is the gravity coefficient, which is set to 1/5000 to maintain the total freight demand at the same level as the current approximate freight volume of the new Western land-sea Corridor (https://www.xibulhxtd.cn/html/2023/07/384.html) in 2022. The OD matrix is shown in Appendix B.

The data on the network, including the distance of different modes between two nodes, are obtained from Google Maps (http://map.google.com/), CHINA RAILWAY (http://www.china-railway.com.cn/), the China Railway Corporation website ($www.12306.cn$), and the Port Odometer Web (http://www.360doc/com/). Other costs and hub capacities parameters are obtained and adapted from the related literature (Zhao et al., 2018; Wei and Dong, 2019; Wang et al., 2023; Ma et al., 2023; Zhang et al., 2023b; Bayram et al., 2023), and development plans from official websites (http://www.ndrc.gov.cn/), which are shown in Table 3. Note that to reflect the cost and capacity differences between different types of hubs, we set the base capacity level $[C]_{base}$ and the base fixed cost of establishing a hub F^h_{base}. The capacity levels of dry ports are set as C_{base}, $2C_{base}$, $3C_{base}$, cross-border inland ports as $2C_{base}$, $4C_{base}$, $6C_{base}$, and seaports as $3C_{base}$, $6C_{base}$, $9C_{base}$. The fixed costs of establishing a dry port with capacity level q are set as F^h_{base}, $1.2F^h_{base}$, $1.4F^h_{base}$, cross-border inland port with capacity level q as $2F^h_{base}$, $1.2 * 2F^h_{base}$, $1.4 * 2F^h_{base}$, seaport with capacity level q as $3F^h_{base}$, $1.2 * 3F^h_{base}$, $1.4 * 3F^h_{base}$.

5.2.2. Network solution analysis under different scenarios

In this section, we analyze how economies of scale (EOS) and hub congestion affect the hierarchical multimodal hub network design. We compare the network solutions under three scenarios: only with the effect of EOS, only with hub congestion, and with the EOS and congestion. The experiment will show the impact on spatial patterns and expected costs.

The results are shown in Table 4, where column 2 lists the hub selection and the corresponding capacity level, and the third and fourth columns list the number of hub arcs (HA) and non-hub arcs (NHA), respectively. Fig. 12 shows the changes in various cost components under different scenarios.

From Table 4, we notice that from "EOS only" to "EOS+congestion", and then to "Congestion only", the fixed cost of establishing hubs and arcs increases since more hubs and hub arcs (HA) are added to spread the volumes. Constructing more hubs, on the one hand, makes the demand nodes closer to the hub to reduce the transportation costs of non-hub arcs. On the other hand, more flows are attracted to use the hub arc services and the EOS effect in the network tapers off from the scenario "EOS only" to the scenario "Congestion only", which increases the hub arc transportation cost (See Fig. 12). In essence, in the "EOS+congestion", the investments for hub construction to limit congestion are compensated by more opportunities to achieve economies of scale. In the "EOS only", these advantages can be achieved largely already with fewer hubs. Also, the "Congestion only" is even more likely to generate economies of scale due to the large number of hubs, even though these, eventually, cannot be exploited.

Since the "EOS and congestion" is the most realistic and complete scenario, it is interesting to evaluate the effects of applying the "EOS only" and "Congestion only" solutions to this setting. The results are shown in Fig. 13. The "Congestion only" solution is very close to the "EOS and congestion". As mentioned before, in this case, the construction of hubs creates a setting where economies of scale could be potentially achieved, and they finally are when inserting EOS. Hence, the penalty of applying this solution to a more realistic setting is somewhat limited. On the other hand, the congestion cost is the component that has the largest impact on the decisions of the models. In fact, when plugging in the solution of the "EOS only" scenario into the more realistic "EOS+congestion" scenario, the result is dramatic. Since the congestion cost is not considered, not having enough hubs may result in a substantial cost in the end. All in all, achieving high utilization of high-capacitated means of transport is a necessary consequence of all three scenarios to exploit the network, which somewhat diminishes the value of EOS compared to the congestion factor.

In terms of the network solution structure obtained in "EOS and congestion", the dry port "Chongqing" and the seaport "Beibu Gulf" are chosen as the hub locations in this optimized scheme, which matches the corridor's real-world setting. From the perspective of the structure and distribution scheme of the whole network, we can notice that the service range of different hubs overlaps since they serve the cargo demand of the same demand nodes. This indicates that the multi-allocation strategy provides good flexibility for cargo consolidation to better balance the hub congestion and the EOS in the network. In addition, dry ports are favored for being located closer to high and dense demand areas.

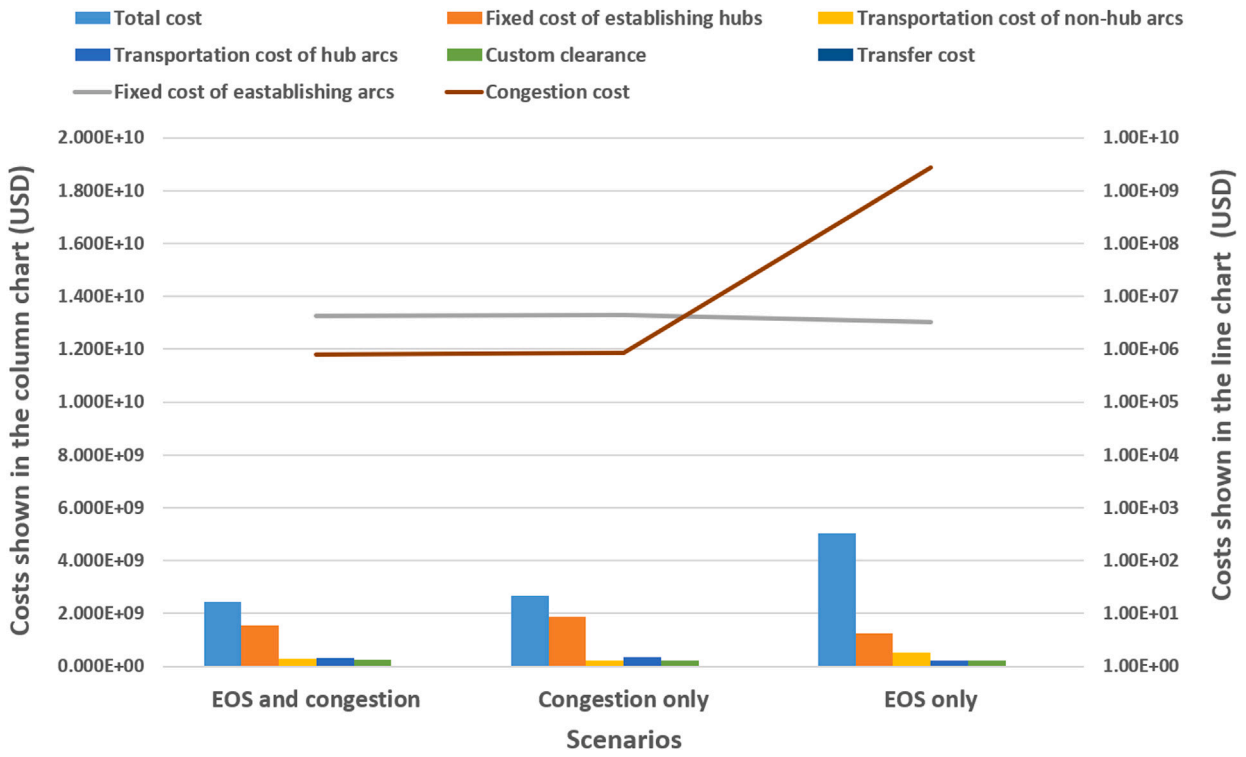


Fig. 13. The changes in the cost components under three different scenarios when further considering EOS and Congestion in the congestion-only and EOS-only solutions, respectively. (The vertical axis on the right is converted to a logarithmic scale.)

5.2.3. *Sensitivity analysis*

In this section, we discuss the network characteristics of solutions for real-world cases under different parameter settings to provide additional managerial insights about the application of HMHLP-C-EOS.

Sensitivity analysis on the EOS function

We test different degrees of EOS with three functions shown in Table 5 in the "EOS+congestion" scenario, referring to Klincewicz (2002), Alkaabneh et al. (2019), and Rostami et al. (2022). The different slopes of the three functions correspond to different levels of benefits from the EOS: modest (f1), median (f2), and aggressive (f3). The results are presented in the Appendix C in Table C.1 and Fig. C.1.

With more aggressive EOS, the total cost and fixed costs of establishing hubs and arcs decrease (see Fig. C.1 in the Appendix) since the number of hubs and hub arcs is gradually reduced to centralize the flows (shown in Table C.1). Furthermore, the capacity level of some hubs, especially in the scenario in $f3$, increases to help limit congestion and consolidate flows to achieve a more aggressive EOS on hub arcs. In addition, the decrease in the number of hubs and hub arcs lowers the transportation cost of the hub arcs and increases the transportation cost of the non-hub arcs. This is because fewer flows go through the hub arcs with the cheaper unit transportation cost due to the aggressive EOS, and more flows go to the non-hub arcs.

Sensitivity analysis on the congestion cost factor

We test different congestion cost factors, ψ_u, for the "EOS+congestion" scenario. The congestion cost factor starts from 0 and takes values of 0, 2000, 4000, and 8000 respectively. The results are presented in the Appendix C in Table C.2 and Fig. C.2.

Up to $\psi_u = 4000$, we register a similar behavior as shown in 5.2.2; namely, an increase in the number of dry ports and cross-border inland ports to decentralize the flow and avoid congestion. Notable is the case when the cost factor is the highest $\psi_u = 8000$, in which the number of seaports increases. The reason is that such a high congestion cost makes the inclusion of more high-capacity seaports necessary, helping to ease the aggregation of flows in border ports. It is worthy mentioning that although more flows go to the hub arcs due the increasing number of hubs and hub arcs, a slight decrease in the transportation cost of hub arcs can still be observed in Fig. C.2. This is because more hub arcs with lower unit transportation cost (the maritime service) are added to the

network. In addition, the increase in the fixed cost of establishing arcs and the decrease in the transportation cost of non-hub arcs are expected as more hubs are activated to spread flow consolidation.

6. Discussion

In light of the results, we draw the following managerial implications. These can be used as a starting point for guidelines for the practical design and optimization of networks based on the hierarchical multimodal hub location problem with hub congestion and economies of scale (HMHLP-C-EOS).

(1) Incorporating joint economies of scale and hub congestion is necessary and beneficial for decision-makers in practical hierarchical multimodal logistics network design. On the one hand, the network designers are supposed to design a more appropriate network with fewer hubs that utilize flow aggregation on large-capacity services (e.g., railway and maritime) to reduce transport costs. On the other hand, identifying congestion at key hubs in the network is necessary as it helps avoid operating under pressure, contributing to keeping the balance between achieving economies of scale and incurring congestion costs as a result of consolidating flows.

(2) This study provides guidance and suggestions for the location of potential hubs in China and the ASEAN region in the New Western Land-Sea Corridor. With the economic and trade development of China and ASEAN, the flow of cargo in the corridor will gradually increase, and the hybrid hub-and-spoke transport network based on the dry ports can effectively balance the hub congestion and the EOS effect in the network. The selection of the location of dry ports can consider the geographical location of demand nodes and the size of their freight demand. Nodes that are located closer to high and dense demand areas are given priority to establishing hubs. In addition, the multi-allocation strategy can provide great flexibility for flow consolidation, contributing to expanding the hub service coverage, alleviating the hub congestion, and utilizing the EOS effect.

(3) Higher EOS in the hierarchical multimodal network generates a design with fewer hubs but with hubs equipped with larger capacity levels. When the EOS effect is strong, the operation of hubs (e.g., cross-border inland ports and seaports) in the cross-border position can be prioritized. Decreasing the number of cross-border hubs or increasing their capacity helps to achieve more flow aggregation and better utilize the EOS in the network, thereby reducing the transportation cost of services between hubs.

(4) The congestion cost factor can significantly affect the construction of hubs. Increasing the number of dry ports can reduce the congestion in the network by dispersing flow across new hubs. Furthermore, more railway services generated by new dry ports simultaneously provide the possibility of achieving the EOS effect in the network, contributing to reducing transport costs. When the congestion cost factor is high, the model prioritizes network expansion to reduce bottlenecks by increasing the number or capacity of seaports. This model can help establish at which point networks should be extended by a growing involvement of high-capacity border ports. All in all, constructing well-equipped hubs in practice is conducive to improving the efficiency of handling cargo to weaken the congestion effects, consequently reducing the congestion and enhancing the EOS in the network.

(5) Network designers should flexibly design the combination of different hubs and capacity levels to cope with the change of the two effects rather than intuitively increasing or decreasing the number of hubs. The appropriate level of network investment in congestion reduction is given by the expected benefit of enhanced EOS and can be determined using this model.

7. Conclusions

In this paper, we study a novel and more realistic variant of the capacitated hub location problem where the capacity setting and the joint consideration of economies of scale and hub congestion are incorporated into the multiple-allocation formulation of the hierarchical multimodal cross-border network design. The problem is first modeled as a MILP through a set of linearization techniques. Since standard solvers cannot handle large instances, a tailored meta-heuristics approach was developed. The hybrid adaptive variable neighborhood search algorithm with tailored operators and speed-up strategies decomposes the problem into a master problem and a sub-problem. A set of computational experiments show the effectiveness of the proposed algorithm. Numerical experiments based on the case-study study of the New Western Land-Sea Corridor in China highlight the applicability of the proposed model, and the sensitivity analysis is also conducted to provide further managerial insights for HMHLP-C-EOS.

There are a few extensions that could be worthwhile investigating. We have shown that network design responds to changes in demand and cost levels with spatially significant shifts. One research direction involves the impact of uncertainty about future OD demands and costs on multimodal hub network design. This can be explored using stochastic programming and robust optimization approaches. Another promising extension is towards a multi-period HMHLP-C-EOS. Single-period planning may prove to be near-sighted as investments are not easily reversed, and changes over the years may create lock-ins for the longer term. Gains could be achieved with designs that explicitly consider developments over several periods. Last but not least, investigating the HMHLP-C-EOS in a bi-level programming framework could offer additional insights into the lower-level behaviors and the interaction process between multiple stakeholders with different objectives.

CRediT authorship contribution statement

Zhenjie Wang: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Data curation, Conceptualization. **Dezhi Zhang:** Writing – review & editing, Supervision, Project administration, Investigation, Funding acquisition, Data curation. **Lóránt Tavasszy:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Stefano Fazi:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Conceptualization.

Declaration of competing interest

None.

Acknowledgments

This research is supported by the Science and Technology Research and Development Plan Project of China National Railway Group Co. (Grant No. N2023X022), the China Scholarship Council (CSC) under Grant 202206370138, and the High-end Think Tank Project of Central South University (Grant No. 2021znzk08). We also gratefully thank the support from the Freight and Logistics Lab in the Department of Transport and Planning at Delft University of Technology.

Appendix A

See Algorithms 1 and 2.

Algorithm 1 HAVNS.

Input: a problem instance data, initialization parameter ($Init\_method = 1$: Initialization 1 and $Init\_method = 2$: Initialization 2), the improvement strategy in the LS iteration Strategy ($Strategy = 1$: best improvement and $Strategy = 2$: first improvement), the maximum successive unimproved iterations in the LS $inner\_iter_{max}$, the maximum running time T_{max}.
Output: The final solution S_{HAVNS}.

$G \leftarrow 0$
Initialization: Generate a feasible initial solution S_{init} according to $Init\_method$
$S_{HAVNS} \leftarrow S_{current} \leftarrow S_{init}$
while $t \leq T_{max}$ **do**
  if $G \leq main\_iter_{max}$ **then**
    $S_{hub\_add} \leftarrow MILPsubproblem \leftarrow N_{hub\_add}(S_{current})$
    $S_{LS} \leftarrow MILPsubproblem \leftarrow N_{LS}(strategy, S_{hub\_add})$
  else
    $S_{shake} \leftarrow MILPsubproblem \leftarrow Shake(S_{current})$
    $S_{LS} \leftarrow MILPsubproblem \leftarrow N_{LS}(strategy, S_{shake})$
    $G \leftarrow 0$
  end if
  $S'_{LS} \leftarrow$ removing hubs and arcs with zero flow from S_{LS}
  if $f(S'_{LS}) < f(S_{current})$ **then**
    $S_{current} \leftarrow S'_{LS}$
    if $f(S_{current}) < f(S_{HAVNS})$ **then**
      $S_{HAVNS} \leftarrow S_{current}$
    end if
  else
    G+1
  end if
end while

Algorithm 2 Initialization 1.

Input: a problem instance data, nodes set N, hub candidates set H, and ODs set K.
Output: initial solution S_{init}.

Step 1: Sort all ODs in set K in descending order based on the demand volume in List L. Create an empty list D
Step 2: Load OD pairs in L to the network sequentially.
for each OD pair $k \in L$ **do**
    Let i be the origin node of k. Set $i' \leftarrow i$ (latest searched node) and $W_k^{i'} \leftarrow W_k$ (demand for k waiting for assignment at node i').
    while true **do**
        Find the closest node h to i'.
        if h is the destination node of k **then**
            break
        else if h is a hub **then**
            if h is not opened **then**
                Establish hub h with the largest capacity level and set the free capacity Cap_free_h of hub h as Cap_h.
            end if
            if $Cap\_free\_h \geq W_k^{i'}$ **then**
                Load the $W_k^{i'}$ to h and allocate i' to hub h.
                Update $i' = h$.
            else
                Load a demand of the same volume as the free capacity Cap_free_h onto hub h, allocate the i' to hub h, and mark hub h is full ($Cap\_h\_mark \leftarrow 1$).
                Add i' to list D along with the remaining demand waiting for loading from node i' and keep it as $remain\_flow\_i'$ ($remain\_flow\_i' \leftarrow W_k^{i'} - Cap\_free\_h$).
                Update $i' = h$.
            end if
        else
            if hub h is full ($Cap\_h\_mark = 1$) **then**
                Find the nearest hub h' to i' other than h.
                Update $h = h'$.
            else
                Continue with hub h.
            end if
        end if
    end while
    if list D is empty **then**
        Remove k from list L.
        break
    else
        Select the first i' in D, and update k for further assignment.
    end if
end for
Step 3: Calculate the objective function based on the network and its flow allocation obtained in **Step 1** and **Step 2**, return S_{init} and $f(S_{init})$.

Appendix B

The label and name of the nodes in the network are presented in Table B.1. The export and import OD matrices in the network are presented in Tables B.2 and B.3, respectively.

Table B.1
Label and name of nodes in the network.

Name	Label	Node type	Region	Name	Label	Node type	Region
Chongqing	1	Demand node	1	Hengyang	45	Dry port	1
Hohhot	2	Demand node	1	Changsha	46	Dry port	1
Huaihua	3	Demand node	1	Zhanjiang	47	Dry port	1
Hengyang	4	Demand node	1	Liuzhou	48	Dry port	1
Yueyang	5	Demand node	1	Nanning	49	Dry port	1
Changsha	6	Demand node	1	Chengdu	50	Dry port	1
Zhanjiang	7	Demand node	1	Guiyang	51	Dry port	1
Liuzhou	8	Demand node	1	Zunyi	52	Dry port	1
Hezhou	9	Demand node	1	Kunming	53	Dry port	1
Yulin	10	Demand node	1	Lasa	54	Dry port	1
Baise	11	Demand node	1	Xian	55	Dry port	1
Guilin	12	Demand node	1	Urumqi	56	Dry port	1
Nanning	13	Demand node	1	Wuhan	57	Dry port	1
Chongzuo	14	Demand node	1	Zhengzhou	58	Dry port	1
Chengdu	15	Demand node	1	Nanchang	59	Dry port	1
Yibin	16	Demand node	1	Hanoi	60	Dry port	2
Luzhou	17	Demand node	1	Phnom Penh	61	Dry port	2
Meishan	18	Demand node	1	Bangkok	62	Dry port	2
Panzhihua	19	Demand node	1	Singapore	63	Dry port	2
Zigong	20	Demand node	1	Nay Pyi Daw	64	Dry port	2
Neijiang	21	Demand node	1	Yangon	65	Dry port	2
Mianyang	22	Demand node	1	Vientiane	66	Dry port	2
Guiyang	23	Demand node	1	Kuala Lumpur	67	Dry port	2
Zunyi	24	Demand node	1	Pingxiang	68	Cross-border inland port	1
Anshun	25	Demand node	1	Hekou	69	Cross-border inland port	1
Tongren	26	Demand node	1	Mohan	70	Cross-border inland port	1
Kunming	27	Demand node	1	Ruili	71	Cross-border inland port	1
Lincang	28	Demand node	1	Pingxiang’	72	Cross-border inland port	2
Qujing	29	Demand node	1	Hekou’	73	Cross-border inland port	2
Lasa	30	Demand node	1	Mohan’	74	Cross-border inland port	2
Xian	31	Demand node	1	Ruili’	75	Cross-border inland port	2
Yulin	32	Demand node	1	Beibu Gulf Port	76	Seaport	1
Lanzhou	33	Demand node	1	Zhanjiang Port	77	Seaport	1
Zhongwei	34	Demand node	1	Yantian Port	78	Seaport	1
Shizuishan	35	Demand node	1	Singapore Port	79	Seaport	2
Urumqi	36	Demand node	1	Hu Chi Minh	80	Seaport	2
Wuhan	37	Demand node	1	Hanoi	81	Demand node	2
Luohe	38	Demand node	1	Phnom Penh	82	Demand node	2
Zhengzhou	39	Demand node	1	Bangkok	83	Demand node	2
Nanchang	40	Demand node	1	Singapore	84	Demand node	2
Shijiazhuang	41	Demand node	1	Nay Pyi Daw	85	Demand node	2
Chongqing	42	Dry port	1	Vientiane	86	Demand node	2
Hohhot	43	Dry port	1	Kuala Lumpur	87	Demand node	2
Huaihua	44	Dry port	1				

Table B.2
Export OD matrix between demand nodes (unit: TEU).

	Hanoi	Phnom Penh	Bangkok	Singapore	Nay Pyi Daw	Vientiane	Kuala Lumpur
Chongqing	20 025	318	11 355	2386	2602	1401	8656
Hohhot	120	2	68	21	15	6	73
Huaihua	211	3	96	21	19	12	76
Hengyang	1583	21	629	156	117	70	545
Yueyang	1190	17	521	138	102	55	479
Changsha	9780	137	4105	1057	786	443	3687
Zhanjiang	3441	34	898	177	133	117	630
Liuzhou	1900	24	739	145	132	103	521
Hezhou	185	2	64	15	11	8	51
Yulin	386	4	122	24	20	16	86
Baise	3304	46	1605	254	309	290	940
Guilin	634	8	251	53	46	33	189
Nanning	8795	107	3290	565	555	529	2055
Chongzuo	24 190	301	9785	1504	1652	1814	5543
Chengdu	16 136	273	10 343	2136	2600	1238	7828
Yibin	832	14	517	99	126	69	365
Luzhou	468	8	279	55	66	36	201
Meishan	269	5	175	35	44	21	129
Panzhihua	172	3	133	21	37	21	78
Zigong	141	2	87	17	21	11	63
Neijiang	131	2	80	16	19	10	59
Mianyang	510	9	319	69	79	37	252
Guiyang	1790	27	957	176	206	139	646
Zunyi	386	6	206	41	45	28	148
Anshun	35	1	19	3	4	3	13
Tongren	126	2	60	13	12	7	46
Kunming	6077	103	4470	623	1145	861	2371
Lincang	132	3	132	16	44	25	61
Qujing	548	9	365	55	88	64	208
Lasa	80	2	77	16	28	7	62
Xian	6227	102	3468	903	787	351	3201
Yulin	42	1	24	7	6	2	24
Lanzhou	122	2	78	20	20	8	71
Zhongwei	32	1	20	5	5	2	19
Shizuishan	72	1	44	12	11	4	44
Urumqi	305	6	230	71	62	18	255
Wuhan	6604	97	2950	825	584	296	2854
Luohe	83	1	41	12	8	4	41
Zhengzhou	7449	118	3766	1095	798	358	3803
Nanchang	3386	47	1368	397	256	135	1359
Shijiazhuang	1244	21	657	203	142	59	704

Table B.3
Import OD matrix between demand nodes (unit: TEU).

	Hanoi	Phnom Penh	Bangkok	Singapore	Nay Pyi Daw	Vientiane	Kuala Lumpur
Chongqing	18 539	1360	8775	3205	1692	546	4104
Hohhot	193	15	91	49	17	4	60
Huaihua	37	2	14	5	2	1	7
Hengyang	1507	94	499	215	78	28	266
Yueyang	2512	167	918	421	151	49	518
Changsha	5907	381	2070	926	334	113	1141
Zhanjiang	11 154	516	2430	833	304	160	1046
Liuzhou	111	7	36	12	5	3	16
Hezhou	34	2	10	4	1	1	5
Yulin	213	11	56	19	8	4	24
Baise	416	27	169	46	27	15	61
Guilin	88	5	29	11	4	2	13
Nanning	15 150	848	4731	1411	672	384	1813
Chongzuo	7536	434	2545	680	361	238	884
Chengdu	17 990	1407	9625	3455	2036	581	4470
Yibin	767	58	398	133	82	27	172
Luzhou	531	40	264	90	52	17	117
Meishan	226	18	122	43	26	8	55
Panzhihua	134	11	86	23	20	7	31
Zigong	97	7	50	17	10	3	22
Neijiang	29	2	15	5	3	1	7
Mianyang	478	37	249	94	52	14	121
Guiyang	1700	118	758	243	137	56	314
Zunyi	222	16	99	34	18	7	44
Anshun	35	2	16	5	3	1	6
Tongren	58	4	23	9	4	1	11
Kunming	11 249	883	6907	1674	1489	671	2248
Lincang	179	16	149	31	42	15	43
Qujing	79	6	44	12	9	4	15
Lasa	8	1	7	2	2	0	3
Xian	6166	467	2867	1297	547	146	1623
Yulin	18	1	9	4	2	0	5
Lanzhou	322	26	173	76	37	9	96
Zhongwei	46	4	24	11	5	1	14
Shizuishan	19	2	10	5	2	0	6
Urumqi	163	15	103	55	23	4	70
Wuhan	7053	480	2630	1278	438	133	1561
Luohe	58	4	24	12	4	1	14
Zhengzhou	8540	627	3604	1822	643	173	2233
Nanchang	2309	150	779	393	123	39	475
Shijiazhuang	1113	85	491	264	89	22	323

Appendix C

In this appendix, we present the results of the sensitivity analysis.

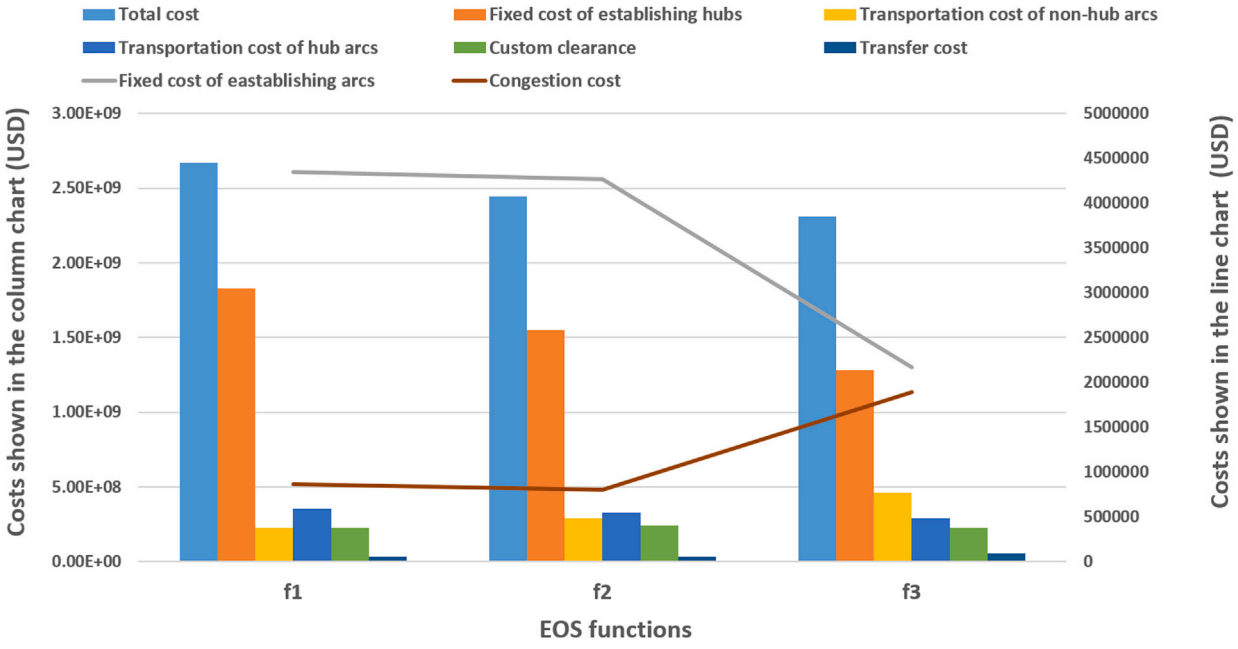


Fig. C.1. The changes in the cost components under different EOS functions.

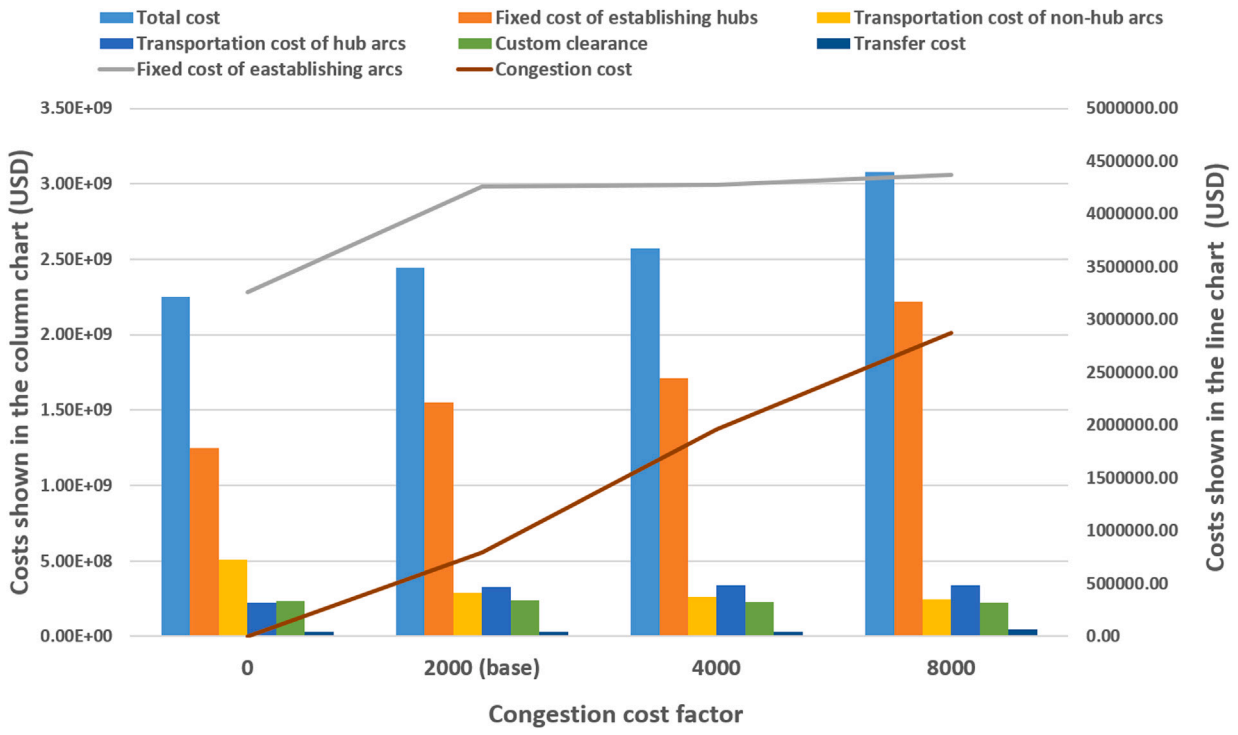


Fig. C.2. The changes in the cost components under different congestion factors.

Table C.1
Effect of EOS on hub network design. In the superscript of the hub location result, D stands for dry port, C stands for cross-border inland port, and S stands for seaport. For HA, taking "r(0): 73" as an example, r represents railway transportation (m represents maritime), (0) indicates that the 0th scale economy interval has been reached (0 means that the flow has not yet generated scale economies), and 73 indicates that there are 73 arcs.

Scenarios	Hub selection (capacity)	HA	NHA
f_1	$42(3)^D$, $45(1)^D$, $46(1)^D$, $47(1)^D$, $49(2)^D$, $50(2)^D$, $53(1)^D$, $57(1)^D$, $58(1)^D$, $62(3)^D$, $64(1)^D$, $68(2)^C$, $69(1)^C$, $70(2)^C$, $71(2)^C$, $72(2)^C$, $73(1)^C$, $74(2)^C$, $75(2)^C$, $76(1)^S$, $79(1)^S$	r(0):73, r(1):2, r(2):1, r(3):7 m(0):0, m(1):0, m(2):1, m(3):1	141
f_2	$42(3)^D$, $46(2)^D$, $50(2)^D$, $58(2)^D$, $62(3)^D$, $64(1)^D$, $68(2)^C$, $69(1)^C$, $70(1)^C$, $71(2)^C$, $72(2)^C$, $73(1)^C$, $74(1)^C$, $75(2)^C$, $76(1)^S$, $79(1)^S$	r(0):34, r(1):1, r(2):3, r(3):6 m(0):0, m(1):0, m(2):1, m(3):1	184
f_3	$42(2)^D$, $49(2)^D$, $57(3)^D$, $61(3)^D$, $62(3)^D$, $63(3)^D$, $68(2)^C$, $70(1)^C$, $72(3)^C$, $74(1)^C$, $76(3)^S$, $79(3)^S$	r(0):5, r(1):5, r(2):5, r(3):4 m(0):0, m(1):0, m(2):0, m(3):2	157

Table C.2
Effect of congestion cost factor on hub network design. In the superscript of the hub location result, D stands for dry port, C stands for cross-border inland port, and S stands for seaport. For HA, taking "r(0): 30" as an example, r represents railway transportation (m represents maritime), (0) indicates that the 0th scale economy interval has been reached (0 means that the flow has not yet generated scale economies), and 30 indicates that there are 30 arcs.

Scenarios	Hub selection (capacity)	HA	NHA
0	$42(1)^D$, $46(1)^D$, $50(2)^D$, $58(2)^D$, $64(1)^D$, $68(3)^C$, $69(1)^C$, $72(3)^C$, $73(1)^C$, $75(1)^C$, $76(1)^S$, $79(1)^S$	r(0):30, r(1):2, r(2):1, r(3):3 m(0):0, m(1):0, m(2):1, m(3):1	204
2000	$42(3)^D$, $46(2)^D$, $50(2)^D$, $58(2)^D$, $62(3)^D$, $64(1)^D$, $68(2)^C$, $69(1)^C$, $70(1)^C$, $71(2)^C$, $72(2)^C$, $73(1)^C$, $74(1)^C$, $75(2)^C$, $76(1)^S$, $79(1)^S$	r(0):34, r(1):1, r(2):3, r(3):6 m(0):0, m(1):0, m(2):1, m(3):1	184
4000	$42(2)^D$, $46(2)^D$, $49(2)^D$, $50(2)^D$, $51(1)^D$, $55(2)^D$, $62(3)^D$, $64(1)^D$, $68(3)^D$, $69(1)^C$, $70(2)^C$, $71(2)^C$, $72(2)^C$, $73(1)^S$, $74(2)^S$, $75(2)^S$, $76(1)^S$, $79(1)^S$	r(0):51, r(1):1, r(2):0, r(3):8 m(0):0, m(1):0, m(2):1, m(3):1	137
8000	$42(3)^D$, $44(1)^D$, $45(1)^D$, $46(2)^D$, $48(1)^D$, $49(2)^D$, $50(2)^D$, $51(1)^D$, $52(1)^D$, $57(1)^D$, $61(1)^D$, $62(3)^D$, $63(3)^D$, $64(1)^D$, $68(3)^C$, $69(1)^C$, $70(2)^C$, $71(1)^C$, $72(3)^C$, $73(1)^C$, $74(2)^C$, $75(1)^C$, $76(3)^S$, $79(1)^S$, $80(1)^S$	r(0):60, r(1):7, r(2):2, r(3):5 m(0):0, m(1):1, m(2):1, m(3):2	184

References

Alkaabneh, F., Diabat, A., Elhedhli, S., 2019. A Lagrangian heuristic and GRASP for the hub-and-spoke network system with economies-of-scale and congestion. Transp. Res. C 102, 249–273.

Alumur, S.A., Campbell, J.F., Contreras, I., Kara, B.Y., Marianov, V., O'Kelly, M.E., 2021. Perspectives on modeling hub location problems. European J. Oper. Res. 291 (1), 1–17.

Alumur, S., Kara, B.Y., 2008. Network hub location problems: The state of the art. European J. Oper. Res. 190 (1), 1–21.

Azizi, N., Salhi, S., 2022. Reliable hub-and-spoke systems with multiple capacity levels and flow dependent discount factor. European J. Oper. Res. 298 (3), 834–854.

Bayram, V., Yıldız, B., Farham, M.S., 2023. Hub network design problem with capacity, congestion, and stochastic demand considerations. Transp. Sci. 57 (5), 1276–1295.

Bütün, C., Petrovic, S., Muyldermans, L., 2021. The capacitated directed cycle hub location and routing problem under congestion. European J. Oper. Res. 292 (2), 714–734.

Campbell, J.F., O'Kelly, M.E., 2012. Twenty-five years of hub location research. Transp. Sci. 46 (2), 153–169.

Contreras, I., Díaz, J.A., Fernández, E., 2009. Lagrangean relaxation for the capacitated hub location problem with single assignment. OR Spectrum 31, 483–505.

Contreras, I., O'Kelly, M., 2019. Hub location problems. Locat. Sci. 327–363.

De Camargo, R.S., de Miranda, Jr., G., Ferreira, R.P., 2011. A hybrid outer-approximation/benders decomposition algorithm for the single allocation hub location problem under congestion. Oper. Res. Lett. 39 (5), 329–337.

De Camargo, R.S., de Miranda, Jr., G., Luna, H.P.L., 2009. Benders decomposition for hub location problems with economies of scale. Transp. Sci. 43 (1), 86–97.

de Camargo, R.S., Miranda, G., 2012. Single allocation hub location problem under congestion: Network owner and user perspectives. Expert Syst. Appl. 39 (3), 3385–3391.

Dhyani Bhatt, S., Jayaswal, S., Sinha, A., Vidyarthi, N., 2021. Alternate second order conic program reformulations for hub location under stochastic demand and congestion. Ann. Oper. Res. 304, 481–527.

Dukkanci, O., Kara, B.Y., 2017. Routing and scheduling decisions in the hierarchical hub location problem. Comput. Oper. Res. 85, 45–57.

Elhedhli, S., Hu, F.X., 2005. Hub-and-spoke network design with congestion. Comput. Oper. Res. 32 (6), 1615–1632.

Elhedhli, S., Wu, H., 2010. A Lagrangean heuristic for hub-and-spoke system design with capacity selection and congestion. INFORMS J. Comput. 22 (2), 282–296.

Eydi, A., Khaleghi, A., Barzegar, K., 2022. Ring hierarchical hub network design problem: Exact and heuristic solution methods. EURO J. Transp. Logist. 11, 100096.

Farahani, R.Z., Hekmatfar, M., Arabani, A.B., Nikbakhsh, E., 2013. Hub location problems: A review of models, classification, solution techniques, and applications. Comput. Ind. Eng. 64 (4), 1096–1109.

Fontes, F.F.d.C., Goncalves, G., 2021. A variable neighbourhood decomposition search approach applied to a global liner shipping network using a hub-and-spoke with sub-hub structure. Int. J. Prod. Res. 59 (1), 30–46.

Freitas, N.C., de Sá, E.M., de Souza, S.R., 2023. A GVNS algorithm applied to the single allocation hub location problem with heterogeneous economies of scale. Comput. Oper. Res. 159, 106350.

Ghaffarinasab, N., Çavuş, Ö., Kara, B.Y., 2023. A mean-CVaR approach to the risk-averse single allocation hub location problem with flow-dependent economies of scale. Transp. Res. B 167, 32–53.

Hoff, A., Peiro, J., Corberan, A., Marti, R., 2017. Heuristics for the capacitated modular hub location problem. Comput. Oper. Res. 86, 94–109.

Irawan, C.A., Salhi, S., Jones, D., Dai, J., Liu, M.J., 2024. A dry port hub-and-spoke network design: An optimization model, solution method, and application. Comput. Oper. Res. 167, 106646.
Khosravi, S., Bozorgi, A., Zahedi-Seresht, M., 2024. Hub-and-spoke network design considering congestion and flow-based cost function. Appl. Sci. 14 (15), 6416.
Klincewicz, J.G., 2002. Enumeration and search procedures for a hub location problem with economies of scale. Ann. Oper. Res. 110 (1), 107–122.
Li, Z.-C., Bing, X., Fu, X., 2023. A hierarchical hub location model for the integrated design of urban and rural logistics networks under demand uncertainty. Ann. Oper. Res. 1–22.
Li, H., Wang, Y., 2023. Hierarchical multimodal hub location problem with carbon emissions. Sustainability 15 (3), 1945.
Ma, Y., Shi, X., Qiu, Y., 2020. Hierarchical multimodal hub location with time restriction for China railway (CR) express network. IEEE Access 8, 61395–61404.
Ma, J., Wang, X., Yang, K., Jiang, L., 2023. Uncertain programming model for the cross-border multimodal container transport system based on inland ports. Axioms 12 (2), 132.
Monemi, R.N., Gelareh, S., Nagih, A., Jones, D., 2021. Bi-objective load balancing multiple allocation hub location: a compromise programming approach. Ann. Oper. Res. 296 (1), 363–406.
Najy, W., Diabat, A., 2020. Benders decomposition for multiple-allocation hub-and-spoke network design with economies of scale and node congestion. Transp. Res. B 133, 62–84.
National Development and Reform Commission, 2018. National logistics hub layout and construction plan.
National Development and Reform Commission, 2019. Overall planning of the new land-sea corridor in the west 2019–2035. https://www.gov.cn/xinwen/2019-08/15/5421375/files/345c17c4bbaf4606ac36f49b149cbaec.pdf.
O'Kelly, M.E., 1986. The location of interacting hub facilities. Transp. Sci. 20 (2), 92–106.
O'Kelly, M.E., 1987. A quadratic integer program for the location of interacting hub facilities. European J. Oper. Res. 32 (3), 393–404.
O'Kelly, M.E., Bryan, D., 1998. Hub location with flow economies of scale. Transp. Res. B 32 (8), 605–616.
Özgün-Kibiroğlu, Ç., Serarslan, M.N., Topcu, Y.İ., 2019. Particle swarm optimization for uncapacitated multiple allocation hub location problem under congestion. Expert Syst. Appl. 119, 1–19.
Qiu, X., Lam, J.S.L., Huang, G.Q., 2015. A bilevel storage pricing model for outbound containers in a dry port system. Transp. Res. E 73, 65–83.
Real, L.B., O'Kelly, M., de Miranda, G., de Camargo, R.S., 2018. The gateway hub location problem. J. Air Transp. Manag. 73, 95–112.
Roso, V., 2013. Sustainable intermodal transport via dry ports–importance of directional development. World Rev. Intermodal Transp. Res. 4 (2–3), 140–156.
Rostami, B., Chitsaz, M., Arslan, O., Laporte, G., Lodi, A., 2022. Single allocation hub location with heterogeneous economies of scale. Oper. Res. 70 (2), 766–785.
Saldanha-da-Gama, F., 2022. Facility location in logistics and transportation: An enduring relationship. Transp. Res. E 166, 102903.
Shang, X., Yang, K., Jia, B., Gao, Z., Ji, H., 2021. Heuristic algorithms for the bi-objective hierarchical multimodal hub location problem in cargo delivery systems. Appl. Math. Model. 91, 412–437.
Shang, X., Yang, K., Wang, W., Wang, W., Zhang, H., Celic, S., 2020. Stochastic hierarchical multimodal hub location problem for cargo delivery systems: Formulation and algorithm. IEEE Access 8, 55076–55090.
Torkestani, S.S., Seyedhosseini, S.M., Makui, A., Shahanaghi, K., 2018. The reliable design of a hierarchical multi-modes transportation hub location problems (HMMTHLP) under dynamic network disruption (DND). Comput. Ind. Eng. 122, 39–86.
Tsao, Y.-C., Thanh, V.-V., 2019. A multi-objective mixed robust possibilistic flexible programming approach for sustainable seaport-dry port network design under an uncertain environment. Transp. Res. E 124, 13–39.
Van Nguyen, T., Zhang, J., Zhou, L., Meng, M., He, Y., 2020. A data-driven optimization of large-scale dry port location using the hybrid approach of data mining and complex network theory. Transp. Res. E 134, 101816.
Wang, Z., Zhang, D., Tavasszy, L., Fazi, S., 2023. Integrated multimodal freight service network design and pricing with a competing service integrator and heterogeneous shipper classes. Transp. Res. E 179, 103290.
Wei, H., Dong, M., 2019. Import-export freight organization and optimization in the dry-port-based cross-border logistics network under the belt and road initiative. Comput. Ind. Eng. 130, 472–484.
Wei, H., Lee, P.T.-W., 2021. Designing a coordinated horizontal alliance system for China's inland ports with China railway express platforms along the silk road economic belt. Transp. Res. E 147, 102238.
Yaman, H., 2009. The hierarchical hub median problem with single assignment. Transp. Res. B 43 (6), 643–658.
Yaman, H., Carello, G., 2005. Solving the hub location problem with modular link capacities. Comput. Oper. Res. 32 (12), 3227–3245.
Zhang, X., Liu, C., Peng, Y., Lu, J., 2023b. Connectivity-based spatial patterns and factors influencing international container multimodal hubs in China under the belt and road initiative. Transp. Policy 143, 10–24.
Zhang, C., Sun, X., Dai, W., Wandelt, S., 2023a. Solving hub location problems with profits using variable neighborhood search. Transp. Res. Rec. 2677 (1), 1675–1695.
Zhao, L., Zhao, Y., Hu, Q., Li, H., Stoeter, J., 2018. Evaluation of consolidation center cargo capacity and loctions for China railway express. Transp. Res. E 117, 58–81.
Zhong, W., Juan, Z., Zong, F., Su, H., 2018. Hierarchical hub location model and hybrid algorithm for integration of urban and rural public transport. Int. J. Distrib. Sens. Netw. 14 (4), 1550147718773263.
Zhou, S., Ji, B., Song, Y., Samson, S.Y., Zhang, D., Van Woensel, T., 2023. Hub-and-spoke network design for container shipping in inland waterways. Expert Syst. Appl. 223, 119850.