Bayram, Vedat, Aydogan, Ciya and Kargar, Kamyar (2025) *Multi-Period Hub Network Design from a Dual Perspective: An Integrated Approach Considering Congestion, Demand Uncertainty, and Service Quality Optimization.* European Journal of Operational Research, 326 (1). pp. 78-95. ISSN 0377-2217.

Discrete Optimization

Multi-period hub network design from a dual perspective: An integrated approach considering congestion, demand uncertainty, and service quality optimization

Vedat Bayram [a,b,*], Çiya Aydoğan [c], Kamyar Kargar [b,d]

[a] *University of Kent, Kent Business School, Department of Analytics, Operations and Systems, Centre for Logistics and Sustainability Analytics, Canterbury, Kent, United Kingdom*
[b] *Department of Industrial Engineering, TED University, Ankara, Turkey*
[c] *Department of Industrial Engineering, Middle East Technical University, Ankara, Turkey*
[d] *Lancaster University, Management School, Department of Management Science, Lancaster, LA1 4YX, United Kingdom*

ARTICLE INFO

ABSTRACT

This study introduces a hub network design problem that considers three key factors: congestion, demand uncertainty, and multi-periodicity. Unlike classical models, which tend to address these factors separately, our model considers them simultaneously, providing a more realistic representation of hub network design challenges. Our model also incorporates service level considerations of network users, extending beyond the focus on transportation costs. Service quality is evaluated using two measures: travel time and the number of hubs visited during travel. Moreover, our model allows for adjustments in capacity levels and network structure throughout the planning horizon, adding a dynamic and realistic aspect to the problem setting. The inherent nonlinear nonconvex integer programming problem is reformulated into a mixed-integer second-order cone programming (SOCP) problem. To manage the model's complexity, we propose an exact solution algorithm based on Benders decomposition, where the sub-problems are solved using a column generation technique. The efficacy of the solution approach is demonstrated through extensive computational experiments. Additionally, we discuss the benefits of each considered feature in terms of transportation costs and their impact on network structure, providing insights for the field.

1. Introduction

Hub-and-spoke configuration is the backbone of telecommunication and transportation networks, particularly in situations where direct connections between points are either too expensive, inefficient, or impossible. Hubs undertake different functionalities in various real-world applications. In transportation networks, hubs consolidate, sort and redistribute the flow (e.g., passengers and freight) between origin–destination (OD) pairs (Alumur & Kara, 2008). In telecommunication systems, they act as routers, concentrators and switch centers (Klincewicz, 1998). Hubs also arise at the center of many innovative transportation and telecommunication services such as Bike-and-Ride (Tavassoli & Tamannaei, 2020), crowdsourced delivery (Macrina et al., 2020), mobility as a service (Jittrapirom et al., 2017), unmanned aerial vehicles (drones) delivery (Macias et al., 2020) and physical internet (Montreuil, 2011).

Hub networks provide two major benefits: economies of scale and efficient utilization of limited resources. However, there are potential drawbacks associated with these networks that cannot be overlooked. Firstly, consolidating the flow leads to congestion at hub centers, which can increase the overall cost and result in poor service quality. Hub airports are a prime example where congestion costs have been estimated to be more than $18.5 billion annually. In addition to this, congestion also leads to more than one million metric tons of carbon dioxide emissions (Forbes, 2019). A report by the Federal Aviation Administration (FAA, 2020) reveals that the total cost of delays in the US for 2019 was estimated to be around $33 billion. Secondly, hub networks lead to longer travel distances and times as compared to direct transportation, which can be a major issue for time-sensitive items or passenger transportation. Travel distance is particularly crucial for the distribution of perishable items like fresh fish (Ghaffarinasab et al.,

2022). Thirdly, demand fluctuations over the planning horizon can pose a major challenge while designing hub networks. Ignoring these variations can lead to idle capacity at some hubs and congestion at others, leading to higher total costs and lower service quality. Existing studies in the literature address these issues separately and propose models and solution methodologies to mitigate the negative effects. However, minimizing transportation costs, congestion, and travel distance are conflicting objectives that must be considered simultaneously throughout the entire planning horizon.

This study seeks to bridge the existing gap by presenting a comprehensive hub network design problem that concurrently addresses the aforementioned issues within a unified mathematical framework. Firstly, our model deals with congestion both at strategic and operational levels using a Kleinrock function that represents service/transit time (congestion cost) at the hubs. At the strategic level, hub capacities are determined by considering the demand uncertainty and congestion cost, and at the operational level, the available capacity is used in the most efficient way utilizing dynamic routing and hub assignments. Secondly, while common practice in the literature to avoid the long travel distances is to model the problem with service-oriented objectives such as hub center and maximal covering hub problems (Alumur & Kara, 2008), this approach shifts the focus on service level and looks at the problem from the perspective of the network user. To strike a balance between the concerns of the network owner and those of the network users, our proposed model minimizes the total cost from the network owner's perspective but at the same time guarantees that the service level will not exceed a given threshold. This threshold value is bounded by two criteria: a multiple of the shortest distance between the corresponding OD pair and the maximum number of hubs visited throughout the path. Thirdly, we incorporate a multi-period setting into our model to ease the undesired consequences of demand volatility during consecutive periods. In the proposed multi-period stochastic model, demand uncertainty at each period is captured by a finite set of scenarios with known probabilities. Our model has the flexibility to relocate, open new hubs, or close the existing ones. Given that we are considering capacitated hubs, the model also provides the possibility for adjusting the capacity of the operating hubs over time.

Moreover, diverging from the common practice observed in the literature, where OD path typically connects with only two intermediary hubs, our model introduces a problem parameter dictating the maximum allowable number of hubs visited along the path for each OD pair. To address this extension, we have devised a path-based mixed-integer second-order cone programming (MISOCP) formulation. These distinctive features endow our model with the capability to address a spectrum of practical scenarios, such as inter-modal public transportation featuring multiple stops in an itinerary (Marín et al., 2002) and telecommunication networks characterized by specific backbone structures (Klincewicz, 1998).

Acknowledging the NP-hard nature of many hub location problems in the literature (Alumur & Kara, 2008), the incorporation of the described extensions introduces additional computational challenges. Consequently, we propose a Benders Decomposition-based (BD; Benders, 1962) algorithm to yield high-quality solutions. The original nonlinear nonconvex integer programming problem is reformulated into a mixed-integer second-order cone programming (SOCP) problem. Subsequently, the Benders subproblem, which is a SOCP problem, is solved through the column generation technique to determine the optimal flow routes. A labeling algorithm with dominance rules is employed to generate pricing columns. The generation of new columns and Benders cuts requires employing the duality results of SOCP (Bayram & Yaman, 2018). Additionally, two sets of constraints (valid inequalities) are introduced to the master problem to ensure the generation of capacity-feasible solutions.

This study stands as the pioneering effort to simultaneously address congestion, service level, and stochasticity within the domain of hub location in a multi-period setting. Within this research framework, our objectives include examining the evolution of optimal network structure over the temporal horizon in the presence of demand uncertainty, understanding the advantages associated with the acquisition and adjustment of capacity in hub centers, exploring the impact of integrating service level considerations on optimal network topology and transportation costs and elucidating the collective benefits derived from the concurrent consideration of congestion, service level, and uncertainty within the analytical framework.

The remainder of the paper is organized as follows. Section 2 reviews the related literature. In Section 3, we propose a mixed-integer nonlinear programming formulation for the problem and discuss its transformation into a mixed-integer second-order conic programming one. Section 4 provides an overview of the proposed BD algorithm. In Section 5, we discuss the column generation technique used for solving sub-problems. We represent computational results in Section 6 and conclude the paper in Section 7.

2. Literature review

Ever since the pioneering works of O'Kelly (1986) and Campbell (1994), the hub location literature has been enriched in various directions. While early studies focused on innovative modeling and the development of efficient solution techniques for the classical hub location problem, this attention gradually shifted to the consideration of extensions that better reflect the relevant real-world characteristics. For a comprehensive overview of the hub location problem, including its historical evolution and application characteristics, we direct interested readers to the review articles by Alumur and Kara (2008), Farahani et al. (2013), Contreras (2015), and Alumur et al. (2021). Despite the existence of rich literature on hub location, it still lacks studies that jointly take into account the three main challenges with hub networks: congestion, service quality, and demand fluctuation over the planning horizon. The current study is a first step in incorporating these three key difficulties under a unified framework. The following sections present the literature review for each area.

2.1. Congestion in HLP

In the literature, one common approach to avoid congestion and/or reflect operational restrictions is to impose the capacity limit on the total flow passing through hubs and/or arcs. This method is adopted in many capacitated variants of the hub location problems. Some examples are Correia et al. (2010), Serper and Alumur (2016), Karimi (2018), Taherkhani et al. (2020) and Ghaffarinasab (2022). However, this approach fails to properly simulate the exponential behavior of the congestion effect. In practice, the congestion effect accelerates as the traffic flow through a hub approaches its capacity. For this reason, modeling the congestion using nonlinear functions will mimic this behavior more accurately. There exists a limited number of studies in the literature that use a nonlinear function to explicitly consider congestion cost in the hub location and based on the function type they fall into two classes: power law and Kleinrock type functions.

The power law function is presented in the form of au^b, where a and b are positive constants and u is the flow through a hub. It is employed in some studies such as De Camargo et al. (2011), Elhedhli and Hu (2005), Kian and Kargar (2016), de Camargo and Miranda (2012) and Alkaabneh et al. (2019). Different from the power law congestion function, some research model the congestion effect using the Kleinrock average delay functions (Kleinrock, 2007). The Kleinrock function has the form $\frac{u_k}{C_k - u_k}$, where u_k and C_k are flow and capacity level at hub k, respectively. Compared to the power law cost function, Kleinrock functions are more effective in capturing the congestion effect as they consider the relative difference between hub flow and hub capacity rather than the hub flow alone. Reflecting the congestion cost through a Kleinrock function, where each hub is modeled as an M/M/1 queue in the steady-state condition, has been previously employed in hub

location models (Bayram et al., 2023; Elhedhli & Wu, 2010). In this paper, we also adopt this approach to properly represent the congestion effect.

Guldmann and Shen (1997) is the first study in the literature that includes the congestion cost in the objective function using a Kleinrock-type congestion cost function. Later, Elhedhli and Wu (2010) adopted a similar approach and proposed a Lagrangian relaxation-based heuristic to solve the problem. Azizi et al. (2018) consider a problem where hubs are modeled as spatially distributed $M/G/1$ queues and use a piecewise linear approximation technique to linearize the resulting nonlinear model. They present an exact method based on the cutting plane technique and a heuristic approach using a genetic algorithm to solve the problem. For a similar problem, Dhyani Bhatt et al. (2021) propose two alternative second-order conic programming reformulations and evaluate their performance through computational experiments. Marianov and Serra (2003) and Mohammadi et al. (2011) employ a chance-constrained method to restrict the probability of congestion occurrence at hubs. de Camargo and Miranda (2012) examine congestion from two distinct perspectives: that of the network owner and the network user. The network owner is focused on achieving a cost-effective network design, while the network user's perspective minimizes the maximum congestion effect. There are also several studies addressing congestion through service time calculation (Alumur, Nickel, & Saldanha-da-Gama, 2018; Domínguez-Bravo et al., 2024; Ishfaq & Sox, 2012). Domínguez-Bravo et al. (2024) consider a discrete set of congestion levels in their model, similar to the approach taken by Alumur, Nickel, and Saldanha-da-Gama (2018), using a time-sensitive demand setting. Recently, Najy and Diabat (2020) incorporated both congestion and flow-dependent economies of scale into the uncapacitated hub location problem. They use a piecewise linear function to approximate the congestion cost function and develop a Benders decomposition algorithm to solve the problem. Bütün et al. (2021) consider the problem in the liner shipping sector and propose a tabu search heuristic to solve the problem. In their study, Jayaswal and Vidyarthi (2023) examine hub node and hub arc location problems, incorporating congestion and service level constraints for two shipment classes: express and regular. They propose a cutting plane-based solution algorithm to solve these problems.

None of the above research considers the demand fluctuation and its effects on congestion and network configuration over the planning horizon. While only Guldmann and Shen (1997), Elhedhli and Wu (2010), Alumur, Nickel, and Saldanha-da-Gama (2018) and Azizi et al. (2018) consider capacity acquisition decisions, no study exists that can balance the flow and reduce the effects of congestion by adjusting capacity in hub centers at different periods. Bayram et al. (2023) is the only study that examines the combined effects of demand uncertainty and congestion. However, unlike our approach, their model assumes the entire planning horizon as a single period and does not allow for any adjustments to the network structure (number, locations, and capacities of hubs), over time, and they do not consider service levels.

2.2. Service level in HLP

In the hub network design problem, service level is commonly measured by travel distances and times. Most of the studies in the literature that explicitly address service level, consider constraints on path lengths. Some example of such studies are Campbell (2009), Ishfaq and Sox (2010), Yaman and Elloumi (2012), Campbell (2013) and Lin and Lee (2018). Imposing bounds on travel distance may partially improve service level but the number of visited hubs is another important factor that affects the travel time and service quality. Delays in hubs due to various reasons (e.g., congestion, layovers, modal shift, etc.) are inevitable and their effects on service quality may be more significant than travel distance. In this study, we assess service quality by considering both travel distance and the number of hubs visited along the path. To the best of our knowledge, no prior literature

has evaluated service quality using these specific criteria. Our proposed model incorporates these considerations by imposing constraints. Specifically, for each origin–destination (OD) pair, the allocated path length must not exceed a predetermined multiple of the shortest path length, i.e., the length of its direct connection, and the number of visited hubs cannot surpass a predefined limit. Failure to meet these criteria prompts the establishment of a direct connection between the corresponding OD pair.

We also want to note that another approach to address service level issues in HLP is to develop hub center or hub covering models (see, e.g., Ghaffarinasab et al., 2022; Yaman et al., 2007). These models prioritize optimizing worst-case service scenarios, disregarding the overall transportation cost. Therefore, by integrating both cost and service quality, our proposed model would provide better insights into many to many transportation carriers (trucking companies, airlines, etc.) for whom the monetary issues cannot be disregarded easily. Also, some studies consider guarantees for alternative paths in case of failure (see, for example, Blanco et al., 2023). However, this topic is more aligned with hub location problems under disruption.

2.3. Demand uncertainty and multi-period HLP

Demand uncertainty constitutes a prevalent aspect in real-world applications of HLP. Existing studies addressing uncertainty in problem parameters predominantly adopt either of two approaches: Stochastic programming (Alumur et al., 2012; Ghaffarinasab et al., 2023; Hu et al., 2021; Kargar & Mahmutoğulları, 2022; Peiró et al., 2019; Sener & Feyzioglu, 2022; Yang & Chiu, 2016), or robust optimization (de Sá et al., 2018; Ghaffarinasab, 2018; Ghaffarinasab et al., 2020; Wang et al., 2020). These investigations typically assume that demand remains constant after realization, modeling the problem as a single-period HLP. However, in practice, problem parameters such as demand and cost may exhibit temporal variations as the network evolves over the planning horizon. The dynamic nature of HLPs implies that the optimal network structure for the initial period may not remain optimal or even feasible for subsequent periods. To adequately address these dynamic changes in environmental parameters, certain researchers have proposed multi-period hub location models.

The study by Campbell (1990) is the first to explore HLPs in a dynamic context, presenting a continuous approximation model for a general freight carrier servicing a fixed region with increasing demand density. Similarly, Khaleghi and Eydi (2023) study the bi-objective version of the continuous-time horizon hub location problem, aiming to minimize network costs and maximize responsiveness. Gelareh and Nickel (2008) contribute a multi-period hub location problem model tailored for public transport applications, using a greedy neighborhood search heuristic for solving the problem. In the context of uncapacitated multiple assignment hub location problems, Contreras et al. (2011) introduce a multi-period model, employing a branch-and-bound algorithm for its solution. Gelareh et al. (2015) consider a budget constraint that accounts for installations, maintenance, and facility closures at each period. Correia et al. (2018) introduce a stochastic multi-period model for the capacitated hub location problem, incorporating demand uncertainty through a finite set of scenarios with associated probabilities. Neamatian Monemi et al. (2021) model the humanitarian aid distribution problem as a multi-period hub location problem with serial demands reflecting user behavior. Aloullal et al. (2023) address the single-allocation, multi-period hub location–routing problem, assuming that, in each period, new hubs and inter-hub connections can be opened, but neither an open hub nor an active connection can be closed. Additionally, demand is assumed to be deterministic.

The multi-period HLP with a star–star network structure is explored by Tikani et al. (2021), who employ genetic algorithm and particle swarm optimization methods to derive high-quality solutions. Khaleghi and Eydi (2021) propose a robust model addressing the sustainable multi-period hub location problem with consideration for

Table 1
Review of Related Literature.

Authors/ Year	Congestion	Demand Uncertainty	Multi-Period	Capacity Adjustment	Network Adjustment	Service Quality	Solution Method	Instance (Size)
Gelareh et al. (2015)			✓		O/C		BD, MH	Gen.30
Correia et al. (2018)		✓	✓	I	O		Solver	CAB25
Azizi et al. (2018)	✓	✓					MH, CP	CAB25, TR81
Alumur, Nickel, and Saldanha-da-Gama (2018)	✓						Solver	AP40
Alkaabneh et al. (2019)	✓						MH, Lag.	CAB25, AP50
Peiró et al. (2019)		✓					H	CAB25, AP50
Najy and Diabat (2020)	✓						BD	CAB25, AP100
Neamatian Monemi et al. (2021)			✓				BD	Gen.100, CS27
Tikani et al. (2021)			✓	I/D		PL	MH	CS37
Dhyani Bhatt et al. (2021)	✓	✓					Solver	CAB25, AP50
Ghaffarinasab et al. (2023)		✓					BD	AP75
Rahmati et al. (2023)	✓	✓					BD	TR50
Aloullal et al. (2023)			✓		O		MH	AP50
Jayaswal and Vidyarthi (2023)	✓	✓				ST	CP	CAB25
Bayram et al. (2023)	✓	✓					BD	TR81, AP200
Khaleghi and Eydi (2023)	✓		Cont.	I/D	O/C		MH	CAB25, TR81, AP40
Domínguez-Bravo et al. (2024)	✓					ST	SF	CAB50
This work	✓	✓	✓	I/D	O/C	PL/NH	BD	CAB25, TR81, AP200

Notations:
Multi-Period: Cont.: Continuous-time planning horizon.
Capacity Adjustment: I: Option to increase capacity; **D**: Option to decrease capacity.
Network Adjustment: O: Option to open hubs; **C**: Option to close hubs.
Service Quality: ST: Service time; **NH**: Number of hubs; **PL**: Path length.
Solution Method: BD: Benders decomposition; **SF**: Strengthened formulation; **Lag.**: Lagrangian; **MH**: Metaheuristic; **CP**: Cutting-plane; **H**: Heuristic.
Instance: Gen.: Generated instances; **CS**: Case study.
Instance size: Refers to the highest number of network nodes for which a solution is reported. This does not necessarily imply the solution was found to optimality.

time-dependent demand. Rahmati et al. (2023) focus on sustainability in the hub location problem by controlling carbon emissions through a carbon cap policy, presenting both deterministic and stochastic formulations without incorporating a multi-period aspect.

Except for Correia et al. (2018) and Khaleghi and Eydi (2021), in all these studies, demand at each period is assumed to be deterministic and available to the decision-maker at the beginning of the planning horizon. This assumption is unrealistic in most real-world applications where uncertainty is unavoidable and the demand at each period is partially dependent on realized demands in previous periods. Correia et al. (2018) is the only study that includes stochastic programming in the multi-period HLP. However, this study only focuses on the phase-in side of the problem. This means the model neither considers the possibility of closing hubs nor decreasing their capacity and at each period only the establishment of new hubs and capacity expansion of the existing ones are allowed. According to the authors, the reason for this restriction is to avoid the extra difficulty that would emerge from the inclusion of the phase-out decisions. To provide a more realistic model we consider both phase-in and phase-out decisions in our problem framework. Therefore, at each period the proposed model has the flexibility to increase or decrease the capacity level of the operating hubs, as well as to determine whether to establish new hubs or remove existing ones. Table 1 compares most recent relevant literature which at least have one common feature with our work.

2.4. Contributions

Considering the current gap in the literature, this study makes several significant contributions to the field of hub network design, which we highlight as follows:

• We present a unified framework and mathematical model for hub network design that integrates congestion, service levels (travel time and number of hubs visited), and demand uncertainty within a multi-period setting. The model jointly optimizes strategic (hub locations and capacities) and operational (hub assignments and routing) decisions, employing a Kleinrock function to capture service and transit times, thereby managing congestion at both levels. It ensures cost minimization while maintaining service levels within defined thresholds based on OD pair distances and hub visits. The multi-period framework addresses demand volatility through scenario-based uncertainty, enabling dynamic hub operations, including relocations, openings, closures, and capacity adjustments. Unlike traditional models with fixed two-hub connections, our approach permits a variable number of intermediary hubs, accommodating real-world scenarios such as inter-modal transportation and complex telecommunication networks where the triangular inequality assumption does not hold.

• To address the NP-hard nature of the problem and the additional complexities of our model, we develop a Benders Decomposition-based algorithm. This method involves reformulating the problem into a mixed-integer second-order cone programming (MISOCP) problem and solving it using a column generation technique. The generation of new columns and Benders cuts requires the use of the duality results of SOCP. We incorporate a set of valid inequalities in the master problem to ensure that capacity-feasible solutions are generated for the subproblems. Additionally, the generation of new columns necessitates a sophisticated labeling algorithm that solves an elementary shortest path problem with two resource constraints. This algorithm employs dominance rules to retain Pareto-optimal labels and exclude non-optimal ones.

• We conduct an extensive computational study on realistic medium-to-large-scale instances to analyze trade-offs between network design cost and congestion. The results demonstrate the efficiency of our proposed algorithm and provide insights into network evolution, capacity acquisition benefits, service level impacts, and the advantages of an integrated problem-solving approach.

In the next section, we define the problem and provide its formulation, elucidating the process by which we convert the formulation into a mixed-integer SOCP formulation.

3. Problem definition and formulation

This section begins by presenting the problem setting, introducing the notation, and outlining the underlying assumptions. It then provides a mathematical formulation for the problem under study.

The problem integrates both strategic and operational decisions in hub network design — covering hub locations, capacities, assignments, and routing — within a multi-period framework to minimize total costs,
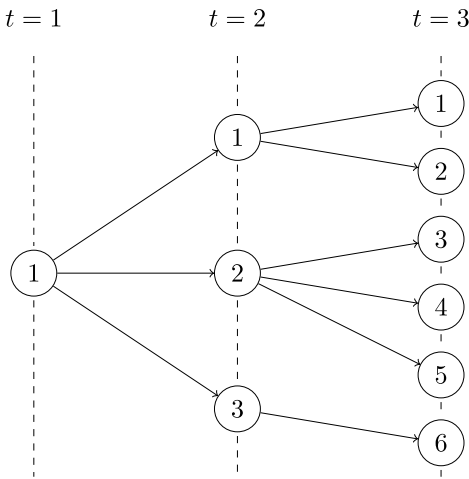
Fig. 1. Scenario tree example.

taking into account congestion, service levels, and demand uncertainty. It is formulated on a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ consisting of a set of nodes $\mathcal{N}$ and a set of arcs $\mathcal{A}$. The graph $\mathcal{G}$ may be either complete or incomplete, with no specific requirements regarding the inter-hub network topology. Let $\mathcal{H} \subseteq \mathcal{N}$ denote the set of potential hub locations, where $\mathcal{H}_0^+ \subseteq \mathcal{H}$ and $\mathcal{H}_0^- \subseteq \mathcal{H}$ represent the set of hubs that are open and closed at the start of the first period, respectively.

The set of periods is represented by $\mathcal{T}$, and for each period $t \in \mathcal{T}$, the set of scenarios is denoted by S_t. The probability of scenario s occurring in period t is denoted by π^{st}. To represent the precedence of scenarios between consecutive periods, we define $\Gamma_t(s)$ for scenario s in period t, where $\Gamma_t(s)$ specifies the scenario in period $t - 1$ that directly precedes scenario s in period t. For example, consider the scenario tree in Fig. 1. In this tree, the preceding scenario for scenario 5 in period 3 is scenario 2 in period 2, so $\Gamma_3(5) = 2$.

Let $\mathcal{K}$ denote the set of commodities associated with origin–destination pairs in $\mathcal{N}$. In this context, $\mathcal{P}_k^{ts}$ represents the set of all possible paths for commodity $k \in \mathcal{K}$ in period $t \in \mathcal{T}$ under scenario $s \in S_t$. User preferences or service quality are reflected by two parameters: (1) τ, the maximum number of hubs (connections) that users are willing to encounter along their path, and (2) $\kappa\rho_k^*$, the maximum allowable length (time or distance) of an alternative path, where ρ_k^* is the direct connection distance for commodity $k \in \mathcal{K}$, and κ is a multiplier indicating user tolerance. Given this tolerance level, the network owner is restricted from assigning a user to a path that exceeds κ times the direct connection length for commodity $k \in \mathcal{K}$ in period $t \in \mathcal{T}$ under scenario $s \in S_t$. We define $\mathcal{P}_k^{ts\kappa} = \{p \in \mathcal{P}_k^{ts} : \rho^p \leq \kappa\rho_k^*, \Lambda_p \leq \tau\}$ as the set of acceptable paths for commodity $k \in \mathcal{K}$ in period $t \in \mathcal{T}$ under scenario $s \in S_t$ at tolerance level κ, where ρ^p denotes the length of path p, and Λ_p denotes the number of connections in path p.

Our formulations/models are constructed using paths and path variables rather than focusing on flows through individual arcs. A path $p \in \mathcal{P}_k^{ts\kappa}$ is defined as an ordered set of nodes $\{o_k = n_p^0, n_p^1, \ldots, n_p^{m_p}, n_p^{m_p+1} = d_k\} \subseteq \mathcal{N}$, where the first node o_k and the last node d_k, represent the origin and destination nodes of commodity $k \in \mathcal{K}$, respectively and the intermediate nodes $\{n_p^1, \ldots, n_p^{m_p}\} \subseteq \mathcal{H}$ are hubs. This assumption is made without loss of generality by allowing the duplication of origin and/or destination nodes if they are hubs, with the transportation cost between the original and duplicated nodes set to zero.

The decision variable v_p^{ts} denotes the fraction of flow routed through path p in period t under scenario s. The transportation cost for a path $p \in \mathcal{P}_k^{ts\kappa}$ in period t under scenario s is given by $c_p^{ts} = w_k^{ts} \sum_{i=0}^{m_p} \hat{c}_{n_p^i n^{i+1} p}^t$, where w_k^{ts} represents the demand for commodity k in period t under scenario s, $\hat{c}_{mn}^t$ is the transportation cost between nodes $m \in \mathcal{N}$ and $n \in \mathcal{N}$ in period t, α is the discount factor for inter-hub transportation,

and $\hat{c}_{mn}^t$ is defined as follows:

$$\hat{c}_{mn}^t = \begin{cases} \hat{c}_{mn}^t & \text{if } m \notin \mathcal{H}, n \in \mathcal{H} \\ \alpha\hat{c}_{mn}^t & \text{if } m \in \mathcal{H}, n \in \mathcal{H} \\ \hat{c}_{mn}^t & \text{if } m \in \mathcal{H}, n \notin \mathcal{H}. \end{cases} \tag{1}$$

If no feasible path exists for a commodity, the model ensures that the commodity is directly transported from its origin to its destination at a cost equal to a multiple C_{direct} of its original connection cost.

We define $\mathcal{L}$ as the set of capacity levels, and we assume that hubs can acquire any discrete capacity level Δ_ℓ from this set at a cost of $g_\ell^t \Delta_\ell$, where g_ℓ^t is the unit capacity acquisition cost at level ℓ in period t. The binary decision variable y_h^{ts} indicates whether candidate hub h is open in period t under scenario s. The binary decision variable $z_{h\ell}^{ts}$ takes a value of 1 if a capacity level ℓ is acquired for hub h in period t under scenario s, and 0 otherwise. Hubs $h \in \mathcal{H}$ can be opened or closed at the beginning of period t at costs A_h^t and B_h^t, respectively, as represented by decision variables a_h^{ts} and b_h^{ts} under scenario s. We assume that the closing cost B_h^t can be negative if partial reimbursement is possible (e.g., relocation of equipment to a new hub facility or recovery gains from the sale of equipment), with the condition that $|B_h^t| < |A_h^t|$. For a complete list of the notation used, please refer to Table 2.

To model the congestion effect at hubs, we use the following congestion cost function as presented in Bayram et al. (2023).

$$\lambda^t \frac{u_h^{ts}}{\sum_{\ell \in \mathcal{L}} \Delta_\ell z_{h\ell}^{ts} - u_h^{ts} + \epsilon}, \quad h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t, \tag{2}$$

where u_h^{ts} is the flow on, and $\sum_{\ell \in \mathcal{L}} \Delta_\ell z_{h\ell}^{ts}$ is the designated capacity of hub $h \in \mathcal{H}$ in period $t \in \mathcal{T}$ in scenario $s \in S_t$, respectively, λ^t is a congestion cost scaling factor, and ϵ denote an arbitrarily small positive number introduced to circumvent scenarios involving division by zero. Below, we present a path-based multi-stage stochastic mixed-integer nonlinear programming (MINLP) formulation for the problem:

$$\min \sum_{t \in \mathcal{T}} \sum_{s \in S_t} \pi^{ts} \left(\sum_{h \in \mathcal{H}} \left(A_h^{ts} a_h^{ts} + B_h^{ts} b_h^{ts} \right) + \sum_{h \in \mathcal{H}} \sum_{\ell \in \mathcal{L}} g_\ell^t \Delta_\ell z_{h\ell}^{ts} \right.$$
$$\left. + \sum_{h \in \mathcal{H}} \lambda^t \frac{u_h^{ts}}{\sum_{\ell \in \mathcal{L}} \Delta_\ell z_{h\ell}^{ts} - u_h^{ts} + \epsilon} + \sum_{k \in \mathcal{K}} \sum_{p \in \mathcal{P}_k^{ts\kappa}} c_p^{ts} v_p^{ts} \right) \tag{3}$$

$$\sum_{p \in \mathcal{P}_k^{ts\kappa}} v_p^{ts} = 1 \qquad \forall k \in \mathcal{K}, t \in \mathcal{T}, s \in S_t \tag{4}$$

$$\sum_{k \in \mathcal{K}} \sum_{\substack{p \in \mathcal{P}_k^{ts\kappa}: \\ h \in p}} w_k^{ts} v_p^{ts} = u_h^{ts} \qquad t \in \mathcal{T}, s \in S_t \tag{5}$$

$$u_h^{ts} \leq \sum_{\ell \in \mathcal{L}} \Delta_\ell z_{h\ell}^{ts} \qquad \forall h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t \tag{6}$$

$$\sum_{\ell \in \mathcal{L}} z_{h\ell}^{ts} = y_h^{ts} \qquad \forall h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t \tag{7}$$

$$y_h^{0,1} = 1 \qquad \forall h \in \mathcal{H}_0^+ \tag{8}$$

$$y_h^{0,1} = 0 \qquad \forall h \in \mathcal{H}_0^- \tag{9}$$

$$y_h^{ts} - y_h^{t-1,\Gamma_t(s)} \leq a_h^{ts} \qquad \forall h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t \tag{10}$$

$$y_h^{ts} \geq a_h^{ts} \qquad \forall h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t \tag{11}$$

$$y_h^{t-1,\Gamma_t(s)} - y_h^{ts} \leq b_h^{ts} \qquad \forall h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t \tag{12}$$

$$y_h^{t-1,\Gamma_t(s)} \geq b_h^{ts} \qquad \forall h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t \tag{13}$$

$$y_h^{ts} + y_h^{t-1,\Gamma_t(s)} + b_h^{ts} \leq 2 \qquad \forall h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t \tag{14}$$

$$v_p^{ts} \geq 0 \qquad \forall k \in \mathcal{K}, p \in \mathcal{P}_k^{ts\kappa}, t \in \mathcal{T}, s \in S_t \tag{15}$$

$$u_h^{ts} \geq 0 \qquad \forall h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t \tag{16}$$

$$y_h^{ts} \in \{0, 1\} \qquad \forall h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t \tag{17}$$

$$z_{h\ell}^{ts} \in \{0, 1\} \qquad \forall h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t, \ell \in \mathcal{L} \tag{18}$$

$$a_h^{ts} \in \{0, 1\} \qquad \forall h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t \tag{19}$$

$$b_h^{ts} \in \{0, 1\} \qquad \forall h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t \tag{20}$$

Table 2

Description of the notation used.

Sets:	
$\mathcal{N}$	Set of nodes
$\mathcal{H}$	Set of potential hub locations; $\mathcal{H} \subseteq \mathcal{N}$
$\mathcal{H}_0^+$	Set of open hubs at the beginning of first period; $H_0^+ \subseteq \mathcal{H}$
$\mathcal{H}_0^-$	Set of closed hubs at the beginning of first period; $\mathcal{H}_0^- \subseteq \mathcal{H}$
$\mathcal{K}$	Set of commodities corresponding to origin–destination pairs in $\mathcal{N}$
$\mathcal{T}$	Set of periods
S_t	Set of scenarios in period $t \in \mathcal{T}$
$\mathcal{P}_k^{ts}$	Set of all paths for commodity k in period t under scenario s
$\mathcal{P}_k^{ts\kappa}$	Set of acceptable paths
$\mathcal{L}$	Set of capacity levels for commodity $k \in \mathcal{K}$ in period $t \in \mathcal{T}$ under scenario $s \in S_t$ at tolerance level κ

Parameters:	
$\hat{c}_{mn}^t$	Transportation cost from node $m \in \mathcal{N}$ to node $n \in \mathcal{N}$ in period t
c_p^{ts}	Transportation cost for path $p \in \mathcal{P}_k^{ts}$ in period t under scenario s
α	Discount factor for inter-hub transportation
τ	The maximum number of hubs (connections) along the paths
ρ^p	The length of path p
Λ_p	The number of connections in path p
ρ_k^*	The direct connection distance for commodity $k \in \mathcal{K}$
$\kappa \rho_k^*$	The maximum allowable length (time or distance) of paths for commodity $k \in \mathcal{K}$
o_k	Origin node of commodity $k \in \mathcal{K}$
d_k	Destination node of commodity $k \in \mathcal{K}$
A_h^t	Cost of opening hub h at the beginning of period t
B_h^t	Cost of closing hub h at the beginning of period t
g_ℓ^t	Unit capacity acquisition cost at level ℓ in period t
λ^t	Congestion cost scaling factor in period t
π^{st}	Probability of occurrence of scenario s in period t
$\Gamma_t(s)$	Preceding scenario in period $t-1$ for scenario s in period t
Δ_ℓ	Capacity amount at level $\ell \in \mathcal{L}$
C_{direct}	Direct connection cost multiplier
w_k^{ts}	Total demand for commodity k in period t under scenario s

Decision Variables:	
a_h^{ts}	Determines whether candidate hub h is opened or not at the beginning of the period t under scenario s
b_h^{ts}	Determines whether candidate hub h is closed or not at the beginning of the period t under scenario s
$z_{h\ell}^{ts}$	Indicates level l capacity acquisition for hub h in period t under scenario s
$u_{h\ell}^{ts}$	Total flow passing through hub h with capacity level l in period t under scenario s
v_p^{ts}	The fraction of total flow transferred by path p in period t under scenario s
y_h^{ts}	Determines whether candidate hub h is open or not in period t under scenario s.

The objective function (3), minimizes the total expected cost related to opening and closing hubs, capacity acquisition, congestion, and transportation during the planning horizon. Constraints (4) guarantee demand satisfaction for every commodity $k \in \mathcal{K}$ in period $t \in \mathcal{T}$ in scenario $s \in S_t$. Constraints (5) compute the total flow passing through each hub in period $t \in \mathcal{T}$ in scenario $s \in S_t$. The flow through each hub is limited by its capacity, which is ensured by constraints (6). Constraints (7) assure that only open hubs can be allocated capacity. Constraints (8)–(9) specify the open and closed hubs at the beginning of the time horizon. Constraints (10)–(14) ensure consistency between the state of the hub in the previous and current periods and opening/closing decisions. In particular, constraints (10)–(11) ensure that the variable a_h^{ts} takes the value of 1 if hub h was closed in period $t-1$ and is opened in period t, and 0 otherwise. Similarly, Constraints (12)–(14) ensure that the variable b_h^{ts} takes the value of 1 if hub h was open in period $t-1$ and is closed in period t, and 0 otherwise. Finally, constraints (15)–(20) define the domains for the decision variables. The following lemma establishes that the variables **a** and **b** can be relaxed from their original binary domain.

Lemma 3.1. *The decision variables a_h^{ts} and b_h^{ts} always take the value of 0 or 1, when they are relaxed.*

Proof. The values of variables a_h^{ts} and b_h^{ts} are determined by the values of variables y_h^{ts} and $y_h^{t-1,\gamma_t(s)}$. Because y_h^{ts} and $y_h^{t-1,\gamma_t(s)}$ are binary variables, four possible cases exist.

If $y_h^{t-1,\gamma_t(s)} = 0$ and $y_h^{ts} = 1$ then constraints (10)–(14) are satisfied only with $a_h^{ts} = 1$ and $b_h^{ts} = 0$. In another case if $y_h^{t-1,\gamma_t(s)} = 1$ and $y_h^{ts} = 0$ then $a_h^{ts} = 0$ and $b_h^{ts} = 1$ is the only solution for the constraints

(10)–(14). In the other case where $y_h^{t-1,\gamma_t(s)} = 0$ and $y_h^{ts} = 0$ constraints (10)–(14) are satisfied only with $a_h^{ts} = 0$ and $b_h^{ts} = 0$. In the case where $y_h^{t-1,\gamma_t(s)} = 1$ and $y_h^{ts} = 1$ constraints (10)–(14) are satisfied only with $b_h^{ts} = 0$. In this case, a_h^{ts} will take of value 0 because its coefficient A_h^t in the objective is defined as positive. $\square$

This modeling approach yields a nonlinear nonconvex objective function, and the inclusion of ϵ may lead to numerical difficulties. To mitigate these issues, we adopt the modeling approach outlined by Bayram et al. (2023), which reformulates the congestion cost function into the following nonlinear but convex form:

$$\lambda^t \frac{u_{h\ell}^{ts}}{\Delta_\ell - u_{h\ell}^{ts}} \quad h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t. \tag{21}$$

Subsequently, we need to make further modifications to the model as follows:

- Define $u_{h\ell}^{ts}$ as the total flow passing through hub h with capacity level l in period t under scenario s.
- Replace the objective function (3) with a convex function as specified below.

$$\min \sum_{t \in \mathcal{T}} \sum_{s \in S_t} \pi^{ts} \left(\sum_{h \in \mathcal{H}} \left(A_h^{ts} a_h^{ts} + B_h^{ts} b_h^{ts} \right) + \sum_{h \in \mathcal{H}} \sum_{\ell \in \mathcal{L}} \left(g_\ell^t \Delta_\ell z_{h\ell}^{ts} + \lambda^t \frac{u_{h\ell}^{ts}}{\Delta_\ell - u_{h\ell}^{ts}} \right) \right.$$
$$\left. + \sum_{k \in \mathcal{K}} \sum_{p \in \mathcal{P}_k^{ts\kappa}} c_p^{ts} v_p^{ts} \right) \tag{22}$$

- Update the constraints (5), (6) and (16) with (23), (24) and (25), respectively, defined as follows:

$$\sum_{k\in\mathcal{K}}\sum_{\substack{p\in P_k^{ts\mathcal{K}}:\\h\in p}} w_k^{ts} v_p^{ts} = \sum_{\ell\in\mathcal{L}} u_{h\ell}^{ts} \qquad \forall k\in\mathcal{K}, t\in\mathcal{T}, s\in S_t \tag{23}$$

$$u_{h\ell}^{ts} \leq \Delta_\ell z_{h\ell}^{ts} \qquad \forall h\in\mathcal{H}, t\in\mathcal{T}, s\in S_t, \ell\in\mathcal{L} \tag{24}$$

$$u_{h\ell}^{ts} \geq 0 \qquad \forall h\in\mathcal{H}, t\in\mathcal{T}, s\in S_t, \ell\in\mathcal{L} \tag{25}$$

4. Benders decomposition

In this section, we introduce a Benders decomposition (BD) approach designed to solve the proposed model efficiently. BD is a decomposition technique tailored for mixed-integer linear and nonlinear programming problems, which partitions the original problem into a master problem (MP) and a series of subproblems (SPs). The MP, obtained by eliminating the second-stage decision variables through projection, initially contains only a subset of constraints referred to as Benders cuts (BCs). Solving the MP yields a lower bound on the optimal solution, as not all BCs are incorporated at the outset. The algorithm proceeds iteratively by solving the MP to determine hub locations and capacities, fixing these decisions, and subsequently solving the SPs for each scenario. Dual information from the SPs is used to generate BCs, which are then added to the MP. This process is repeated until convergence is achieved. For a comprehensive overview of the Benders decomposition algorithm, including its application in combinatorial optimization and various enhancement strategies targeting its key components, the reader is referred to the state-of-the-art survey by Rahmaniani et al. (2017).

In Sections 4.1 and 4.3, we provide the Benders reformulation of the model and outline our approach for solving the subproblems by reformulating them as second-order cone programming (SOCP) problems.

4.1. Primal subproblem

In the Benders reformulation, the master problem includes variables associated with hub opening, closing, and capacity decisions. By fixing the values of these variables as determined by the master problem, the primal subproblem ($PS^{\bar{t}\bar{s}}$) corresponding to scenario $\bar{s}$ in period $\bar{t}$ can be formulated as follows:

$$\min \quad \pi^{\bar{t}\bar{s}}\left(\sum_{h\in\mathcal{H}}\sum_{\ell\in\mathcal{L}}(\lambda^{\bar{t}}\frac{u_{h\ell}^{\bar{t}\bar{s}}}{\Delta_\ell - u_{h\ell}^{\bar{t}\bar{s}}}) + \sum_{k\in\mathcal{K}}\sum_{p\in\mathcal{P}_k^{\bar{t}\bar{s}\mathcal{K}}} c_p^{\bar{t}\bar{s}} v_p^{\bar{t}\bar{s}}\right) \tag{26}$$

$$\sum_{p\in\mathcal{P}_k^{\bar{t}\bar{s}\mathcal{K}}} v_p^{\bar{t}\bar{s}} = 1 \qquad \forall k\in\mathcal{K} \tag{27}$$

$$\sum_{k\in\mathcal{K}}\sum_{\substack{p\in\mathcal{P}_k^{\bar{t}\bar{s}\mathcal{K}}:\\h\in p}} w_k^{\bar{t}\bar{s}} v_p^{\bar{t}\bar{s}} = \sum_{\ell\in\mathcal{L}} u_{h\ell}^{\bar{t}\bar{s}} \qquad \forall h\in\mathcal{H} \tag{28}$$

$$u_{h\ell}^{\bar{t}\bar{s}} \leq \Delta_\ell \bar{z}_{h\ell}^{\bar{t}\bar{s}} \qquad \forall h\in\mathcal{H}, \ell\in\mathcal{L} \tag{29}$$

$$u_{h\ell}^{\bar{t}\bar{s}} \geq 0 \qquad \forall h\in\mathcal{H}, \ell\in\mathcal{L} \tag{30}$$

$$v_p^{\bar{t}\bar{s}} \geq 0 \qquad \forall k\in\mathcal{K}, p\in\mathcal{P}_k^{\bar{t}\bar{s}\mathcal{K}}, \tag{31}$$

where $\bar{z}_{h\ell}^{\bar{t}\bar{s}}$ are values for variables $z_{h\ell}^{ts}$ from the MP. The $PS^{\bar{t}\bar{s}}$ has a nonlinear (convex) objective function and to solve it efficiently we reformulate it as a second-order cone programming (SOCP) model.

4.2. Reformulation of PS as a SOCP

The Second-Order Cone Programming represents a cutting-edge technique in mathematical programming, widely utilized for solving various convex optimization problems (Alizadeh & Goldfarb, 2003; Ben-Tal & Nemirovski, 2001; Lobo et al., 1998; Nesterov et al., 1994). In our approach, the PS is reformulated as SOCP problem by converting

the nonlinear (convex) objective function (26) into a series of second-order cone constraints. This transformation is achieved by introducing auxiliary decision variables $r_{h\ell}^{\bar{t}\bar{s}}$ for each hub and each capacity level in the PS, enabling the formulation to take the following form:

$$r_{h\ell}^{\bar{t}\bar{s}} \geq \frac{u_{h\ell}^{\bar{t}\bar{s}}}{\Delta_\ell - u_{h\ell}^{\bar{t}\bar{s}}} \qquad \forall h\in\mathcal{H}, \ell\in\mathcal{L} \tag{32}$$

$$r_{h\ell}^{\bar{t}\bar{s}} \geq 0 \qquad \forall h\in\mathcal{H}, \ell\in\mathcal{L} \tag{33}$$

The constraints (32) can be represented as second-order cone constraints by multiplying both sides by Δ_ℓ and augmenting $(u_{h\ell}^{\bar{t}\bar{s}})^2$ to both sides. This manipulation yields:

$$(u_{h\ell}^{\bar{t}\bar{s}})^2 \leq (\Delta_\ell r_{h\ell}^{\bar{t}\bar{s}} - u_{h\ell}^{\bar{t}\bar{s}})(\Delta_\ell - u_{h\ell}^{\bar{t}\bar{s}}) \qquad \forall h\in\mathcal{H}, \ell\in\mathcal{L} \tag{34}$$

The constraints presented in Eq. (34) can be represented by hyperbolic inequalities, defined as $\zeta \leq \xi_1\xi_2$, with $\zeta, \xi_1, \xi_2 \geq 0$. This inequality can be converted into a quadratic form, specifically, $\|2\zeta, \xi_1 - \xi_2\| \leq \xi_1 + \xi_2$, where $\|.\|$ denotes the Euclidean norm (Lobo et al., 1998, Alizadeh & Goldfarb, 2003). As a result, Eq. (34) can be articulated as a second-order cone constraint (Günlük & Linderoth, 2008, Salimian, 2013) as follows:

$$\|2u_{h\ell}^{\bar{t}\bar{s}}, \Delta_\ell r_{h\ell}^{\bar{t}\bar{s}} - \Delta_\ell\| \leq \Delta_\ell r_{h\ell}^{\bar{t}\bar{s}} + \Delta_\ell - 2u_{h\ell}^{\bar{t}\bar{s}} \qquad \forall h\in\mathcal{H}, \ell\in\mathcal{L} \tag{35}$$

Now we can reformulate the PS as a SOCP (PS_SOCP) using the above transformations. For scenario $\bar{s}$ in period $\bar{t}$, $PS\_SOCP^{\bar{t}\bar{s}}$ is given below.

$$(PS\_SOCP^{\bar{t}\bar{s}}): \quad \min \quad \pi^{\bar{t}\bar{s}}\left(\sum_{h\in\mathcal{H}}\sum_{\ell\in\mathcal{L}}\lambda^{\bar{t}} r_{h\ell}^{\bar{t}\bar{s}} + \sum_{k\in\mathcal{K}}\sum_{p\in\mathcal{P}_k^{\bar{t}\bar{s}}} c_p^{\bar{t}\bar{s}} v_p^{\bar{t}\bar{s}}\right) \tag{36}$$

$$\sum_{p\in\mathcal{P}_k^{\bar{t}\bar{s}}} v_p^{\bar{t}\bar{s}} = 1 \qquad \forall k\in\mathcal{K} \tag{37}$$

$$\sum_{k\in\mathcal{K}}\sum_{\substack{p\in\mathcal{P}_k^{\bar{t}\bar{s}\mathcal{K}}:\\h\in p}} w_k^{\bar{t}\bar{s}} v_p^{\bar{t}\bar{s}} = \sum_{\ell\in\mathcal{L}} u_{h\ell}^{\bar{t}\bar{s}} \qquad \forall h\in\mathcal{H} \tag{38}$$

$$u_{h\ell}^{\bar{t}\bar{s}} \leq \Delta_\ell \bar{z}_{h\ell}^{\bar{t}\bar{s}} \qquad \forall h\in\mathcal{H}, \ell\in\mathcal{L} \tag{39}$$

$$2u_{h\ell}^{\bar{t}\bar{s}} - t_{1h\ell}^{\bar{t}\bar{s}} = 0 \qquad \forall h\in\mathcal{H}, \ell\in\mathcal{L} \tag{40}$$

$$\Delta_\ell r_{h\ell}^{\bar{t}\bar{s}} - t_{2h\ell}^{\bar{t}\bar{s}} = \Delta_\ell \qquad \forall h\in\mathcal{H}, \ell\in\mathcal{L} \tag{41}$$

$$\Delta_\ell r_{h\ell}^{\bar{t}\bar{s}} - 2u_{h\ell}^{\bar{t}\bar{s}} - t_{3h\ell}^{\bar{t}\bar{s}} = -\Delta_\ell \qquad \forall h\in\mathcal{H}, \ell\in\mathcal{L} \tag{42}$$

$$(t_{2h\ell}^{\bar{t}\bar{s}})^2 + (t_{2h\ell}^{\bar{t}\bar{s}})^2 \leq (t_{3h\ell}^{\bar{t}\bar{s}})^2 \qquad \forall h\in\mathcal{H}, \ell\in\mathcal{L} \tag{43}$$

$$r_{h\ell}^{\bar{t}\bar{s}}, t_{3h\ell}^{\bar{t}\bar{s}} \geq 0 \qquad \forall h\in\mathcal{H}, \ell\in\mathcal{L} \tag{44}$$

$$u_{h\ell}^{\bar{t}\bar{s}} \geq 0 \qquad \forall h\in\mathcal{H}, \ell\in\mathcal{L} \tag{45}$$

$$v_p^{\bar{t}\bar{s}} \geq 0 \qquad \forall k\in\mathcal{K}, p\in\mathcal{P}_k^{\bar{t}\bar{s}\mathcal{K}}, \tag{46}$$

Constraints (40)–(44) are equivalent constraints of the cone constraint (35). To obtain the associated optimality cuts for the MP we define $\gamma_k^{\bar{t}\bar{s}}, \delta_h^{\bar{t}\bar{s}}, \eta_{h\ell}^{\bar{t}\bar{s}}, \mu_{h\ell}^{\bar{t}\bar{s}}, \sigma_{h\ell}^{\bar{t}\bar{s}}$ to be the dual variables corresponding to the constraints (37)–(39), (41), (42), respectively. Also let $\bar{\gamma}_k^{\bar{t}\bar{s}}, \bar{\delta}_h^{\bar{t}\bar{s}}, \bar{\eta}_{kh}^{\bar{t}\bar{s}}, \bar{\mu}_{h\ell}^{\bar{t}\bar{s}}, \bar{\sigma}_{h\ell}^{\bar{t}\bar{s}}$ be the optimal values of these variables at each iteration. Then the optimality cuts (47) for the MP are as follows:

$$\Omega^{tsj} \geq \sum_{k\in\mathcal{K}}\bar{\gamma}_k^{tsj} + \sum_{h\in\mathcal{H}}\sum_{\ell\in\mathcal{L}}\bar{\eta}_{h\ell}^{tsj}\Delta_\ell z_{h\ell}^{\bar{t}\bar{s}}$$

$$+ \sum_{h\in\mathcal{H}}\sum_{\ell\in\mathcal{L}}\bar{\mu}_{h\ell}^{tsj}\Delta_\ell - \sum_{h\in\mathcal{H}}\sum_{\ell\in\mathcal{L}}\bar{\sigma}_{h\ell}^{tsj}\Delta_\ell \quad \forall t\in\mathcal{T}, s\in S_t, j\in J, \tag{47}$$

where Ω^{ts} are the surrogate decision variables used in the MP to represent the congestion and routing costs of each subproblem under scenario s in period t, and J is the set of optimal multiplier vectors.

4.3. Master problem

Decisions on opening and closing of the hubs, as well as capacity acquisition, are made in MP. We also define surrogate decision variables Ω^{ts} to represent congestion and routing cost of subproblems in period t under scenario s. To ensure that the master problem (MP)

produces high-quality solutions in the absence of constraints from the subproblems and to guarantee the feasibility of the subproblems, we incorporate a set of valid inequalities into the MP. By permitting direct shipments between all origin–destination pairs, we ensure that any solution generated by the MP will always result in feasible subproblems. Specifically, the valid inequalities enforce that every open hub must possess sufficient capacity to accommodate the flow passing through it. This restriction is addressed by the constraint set (48).

$$\sum_{\ell \in \mathcal{L}} \Delta_\ell z_{h\ell}^{ts} \geq \sum_{\substack{k \in \mathcal{K}: \\ o_k = h \\ \text{or} \\ d_k = h}} w_k^{ts} y_h^{ts} \quad \forall h \in \mathcal{H}, t \in \mathcal{T}, s \in S_t \tag{48}$$

After adding valid inequalities (48) the final formulation of MP is given below.

$$\min \sum_{t \in \mathcal{T}} \sum_{s \in S_t} \sum_{h \in \mathcal{H}} \pi^{ts} \left(A_h^{ts} a_h^{ts} + B_h^{ts} b_h^{ts} + \sum_{\ell \in \mathcal{L}} g_\ell^t z_{h\ell}^{ts} \right) + \sum_{t \in \mathcal{T}} \sum_{s \in S_t} \Omega^{ts} \tag{49}$$

(7)–(14), (17)–(20), (47), (48)

5. A column generation algorithm for solving the subproblem

$PS\_SOCP^{\bar{t}\bar{s}}$ assumes that $\mathcal{P}_k^{\bar{t}\bar{s}\kappa}$ includes all possible acceptable paths for commodity k. However, generating the complete set of acceptable paths for each commodity will significantly increase the problem size, and consequently make it intractable. To overcome this challenge, we developed a column generation (CG) algorithm, where we can start to solve the problem with an initial path set instead of explicitly listing every possible path for commodity k. Then progressively, the paths that have the potential to improve the objective function are added to the problem. To employ the CG algorithm we need to split $PS\_SOCP^{\bar{t}\bar{s}}$ into two problems: the master problem and the subproblem. The master problem is the original column-wise formulation of the $PS\_SOCP^{\bar{t}\bar{s}}$ with only a subset $\tilde{\mathcal{P}}_k^{\bar{t}\bar{s}\kappa}$ of the set of all acceptable path variables $\mathcal{P}_k^{\bar{t}\bar{s}\kappa}$, i.e., $\tilde{\mathcal{P}}_k^{\bar{t}\bar{s}\kappa} \subset \mathcal{P}_k^{\bar{t}\bar{s}\kappa}$. In each iteration of the algorithm, we solve the master problem and obtain the values for its dual variables. Then, we solve the subproblem to identify a new promising variable that is eligible to enter the basis and improve the objective function value of the master problem. The subproblem is an optimization problem called the pricing problem (PP) which is solved for each commodity k and aims to find the variable with the most negative reduced cost to enter the basis. We continue this procedure until no new column with a negative reduced cost is detected.

5.1. Pricing problem

The reduced cost of a path p in the $PS\_SOCP^{\bar{t}\bar{s}}$ is represented by $\tilde{c}_p^{\bar{t}\bar{s}}$ and computed as follows:

$$\tilde{c}_p^{\bar{t}\bar{s}} = \pi^{\bar{t}\bar{s}} c_p^{\bar{t}\bar{s}} - \gamma_k^{\bar{t}\bar{s}} - \sum_{\substack{h \in \mathcal{H}: \\ h \in p}} w_k^{\bar{t}\bar{s}} \delta_h^{\bar{t}\bar{s}} \tag{50}$$

To detect the candidate paths with negative reduced costs, the PP has to be solved for each commodity k in $PS\_SOCP^{\bar{t}\bar{s}}$. The PP seeks an acceptable path with the most negative reduced cost $\tilde{c}_p^{\bar{t}\bar{s}}$ from o_k to d_k. In our problem setting, the PP is an elementary shortest path problem with two resource constraints (ESPPRC), which is known to be NP-Hard (Dror, 1994). The resource constraints include the limitation on the number of visited arcs (connections) along the path, and the length of the path bounded by parameter τ, and $\kappa \rho_k^*$, respectively.

ESPPRC is solved for the subgraph $\mathcal{G}_k$ of $\mathcal{G}$ with node set $\mathcal{N}_k = \mathcal{H} \cup \{o_k, d_k\}$ and arc set $\mathcal{A}_k = \{(m, n) \in \mathcal{A} : m, n \in \mathcal{N}_k \times \mathcal{N}_k\}$. The reduced cost of the arc $(m, n) \in \mathcal{A}_k$ on $\mathcal{G}_k$ is shown as $\tilde{c}_{kmn}^{\bar{t}\bar{s}}$ and calculated

as follows:

$$\tilde{c}_{kmn}^{\bar{t}\bar{s}} = \begin{cases} \pi^{\bar{t}\bar{s}} w_k^{\bar{t}\bar{s}} \hat{c}_{mn}^{\bar{t}} - \gamma_k^{\bar{t}\bar{s}} - w_k^{\bar{t}\bar{s}} \delta_n^{\bar{t}\bar{s}} & \text{if } m = o_k, n \in \mathcal{H}, \\ \pi^{\bar{t}\bar{s}} w_k^{\bar{t}\bar{s}} \hat{c}_{mn}^{\bar{t}} - w_k^{\bar{t}\bar{s}} \delta_n^{\bar{t}\bar{s}} & \text{if } m, n \in \mathcal{H}, \\ \pi^{\bar{t}\bar{s}} w_k^{\bar{t}\bar{s}} \hat{c}_{mn}^{\bar{t}} & \text{if } m \in \mathcal{H}, n = d_k \end{cases} \tag{51}$$

By obtaining the values of $\gamma_k^{\bar{t}\bar{s}}$ and $\delta_h^{\bar{t}\bar{s}}$ from the current $PS\_SOCP^{\bar{t}\bar{s}}$ we can formulate the $PP_k^{\bar{t}\bar{s}}$ for each commodity k as follows:

$$(PP_k^{\bar{t}\bar{s}}): \quad \min \sum_{(m,n) \in \mathcal{A}_k} \tilde{c}_{kmn}^{\bar{t}\bar{s}} x_{mn} \tag{52}$$

$$\sum_{(m,n) \in \mathcal{A}_k} x_{mn} \leq \tau \tag{53}$$

$$\sum_{(m,n) \in \mathcal{A}_k} t_{mn} x_{mn} \leq \kappa \rho_k^* \tag{54}$$

$$\sum_{\substack{n: \\ (m,n) \in \mathcal{A}_k}} x_{mn} - \sum_{\substack{n: \\ (m,n) \in \mathcal{A}_k}} x_{nm} = \begin{cases} 1 & \text{if } m = o_k \\ -1 & \text{if } m = d_k \quad \forall n \in \mathcal{N}_k \\ 0 & \text{otherwise} \end{cases} \tag{55}$$

$$\varphi_m - \varphi_n + |\mathcal{N}_k| x_{mn} \leq |\mathcal{N}_k| - 1 \qquad \begin{aligned} &\forall (m,n) \in \mathcal{A}_k: \\ &m \neq o_k \end{aligned} \tag{56}$$

$$x_{mn} \in \{0, 1\} \qquad \forall (m,n) \in \mathcal{A}_k \tag{57}$$

$$\varphi_n \geq 0 \qquad \forall n \in \mathcal{N}_k \setminus \{o_k\}, \tag{58}$$

where the binary variable x_{mn} indicates whether the arc (m, n) is in the path or not. The objective function (52) minimizes the total reduced cost of arcs in a candidate path. Constraint (53) ensures that the number of arcs (connections) in a candidate path does not exceed the upper bound τ. Constraint (54) guarantees that the travel time for the path does not exceed the maximum allowed time, where t_{mn} denotes the time spent on the arc (m, n). Constraints (55) enforce flow balance at the nodes. When calculating the reduced cost of an arc for the PP, some arcs may have negative reduced cost values. As a result, solving the shortest path problem could lead to cycles. To address this, we incorporate Miller–Tucker–Zemlin subtour elimination constraints (56) into the Pricing Problem to prevent such cycles, where the variable φ_n indicates the order of node n excluding the origin. Constraints (57) and (58) define the domain of the variables.

Upon solving the $PP_k^{\bar{t}\bar{s}}$, if its objective function value is negative, the corresponding path (column) is appended to the current $PS\_SOCP^{\bar{t}\bar{s}}$. Subsequently, in the next iteration, $PS\_SOCP^{\bar{t}\bar{s}}$ is resolved to find the values of $\gamma_k^{\bar{t}\bar{s}}$ and $\delta_h^{\bar{t}\bar{s}}$. This procedure continues until no new columns price out. To enhance computational efficiency, we employ a streamlined labeling algorithm to solve the PP. The details of this algorithm are presented in the following section.

5.2. Labeling algorithm

The ESPPRC can be solved using a dynamic programming approach, specifically a labeling algorithm (LA). The LA extends the Bellman–Ford shortest path algorithm considering resource constraints (Feillet et al., 2004; Irnich & Desaulniers, 2005). In LA, labels represent partial paths that start from the origin node. The LA starts with a starting label at the origin node and creates new labels by including accessible nodes on the graph using resource extension functions. A node can be added to a path if it does not make the path infeasible. The procedure of adding new labels terminates when all partial paths are processed and as a result, all feasible paths have reached the destination node. However, this procedure can generate an exponentially large number of labels and be inefficient even under tight resource constraints. To deal with this challenge, we keep Pareto-optimal labels and exclude the non-optimal labels. A dominance rule is applied to decide which labels are excluded. Labels, the extension function, and the dominance rule are defined further below.

A label is represented by a tuple $L = (n, \tilde{c}, \hat{\mathcal{N}}, \tilde{t})$ where $n \in \mathcal{N}$ shows the last node of the associated partial path, $\tilde{c}$ is the reduced cost of the path, $\hat{\mathcal{N}}$ is the set of nodes in the path, and $\tilde{t}$ is the length of the path. Initial label is defined as $L^0 = \{o_k, 0, \{o_k\}, 0\}$. The LA extends a label $L = \{n, \tilde{c}, \hat{\mathcal{N}}, \tilde{t}\}$ along arc $(n, n') \in \mathcal{A}_k$, $n' \notin \hat{\mathcal{N}}$ to construct a new label $L' = (n', \tilde{c}', \hat{\mathcal{N}}', \tilde{t}')$ in which $\tilde{c}' = \tilde{c} + \tilde{c}_{knn'}$, $\hat{\mathcal{N}}' = \hat{\mathcal{N}} \cup \{n'\}$, and $\tilde{t}' = \tilde{t} + t_{nn'}$. For each adjacent node of n, a new label is created. To avoid generating unacceptable paths, we do not extend L if it already contains τ number of arcs or if its length will exceed $\kappa \rho_k^*$. Also, note that adding an already visited node is not permitted.

In order to identify and remove non-Pareto-optimal labels, we apply the following dominance rule. For any two labels $L^1 = \left(n^1, \tilde{c}^1, \hat{\mathcal{N}}^1, \tilde{t}^1 \right)$ and $L^2 = \left(n^2, \tilde{c}^2, \hat{\mathcal{N}}^2, \tilde{t}^2 \right)$, we say that L^1 dominates L^2 if all of the following conditions hold: (i) $n^1 = n^2$, (ii) $\tilde{c}^1 \leq \tilde{c}^2$, (iii) $\hat{\mathcal{N}}^1 \subseteq \hat{\mathcal{N}}^2$, and (iv) $\tilde{t}^1 \leq \tilde{t}^2$. In LA, we apply the dominance rule on a node whenever a new label is extended to that node. A label is discarded if it is dominated by at least one other label. Finally, the algorithm stops whenever no label can be extended and returns the non-dominated labels that have reached d_k.

6. Computational experiments

In this section, we first describe the problem instances and the design of experiments. Then, we evaluate the performance of the solution algorithm in Section 6.2. Finally, in Section 6.3, we present a managerial insights discussion, including: (i) an impact analysis that investigates the sensitivity of the model (network topology and corresponding costs) to variations in problem parameters; and (ii) an evaluation of the significance of uncertainty, congestion, and multi-period considerations in the context of the HLP.

We conducted the computational experiments on a workstation running 64-bit Windows, equipped with Intel Xeon E-2246G processors at 3.60 GHz and 16.0 GB of RAM. The coding was done in Java v18, using ILOG CPLEX v22.10. To integrate the Benders cuts, we utilized the lazy constraint callback feature of CPLEX. Additionally, a time limit (TL) of 48 h was imposed for solving each problem instance.

6.1. Test data and design of experiments

We utilized three widely recognized datasets from the hub location problem (HLP) literature to generate problem instances: (i) the Civil Aeronautics Board (CAB) dataset (Beasley, 2018), representing air passenger traffic between 25 US cities; (ii) the Turkish network (TR) dataset (Kara, 2011; Yaman et al., 2007), based on cargo flows between 81 Turkish cities; and (iii) the Australian Postal (AP) dataset (Beasley, 2018; Ernst & Krishnamoorthy, 1996), consisting of 200 nodes. In total, we solved 243 instances for a comprehensive computational analysis. From each dataset, 32 instances were generated for the proposed model, varying parameter settings to evaluate algorithm performance, derive managerial insights, and assess model behavior. An additional 147 instances were solved to examine the effects of congestion, demand uncertainty, and multi-periodicity, as detailed in Sections 6.3.2 to 6.3.4. The planning horizon consists of three periods, with each period having an equal probability of experiencing one of two demand scenarios: low or high. This results in four scenarios—high–high (HH), high–low (HL), low–high (LH), and low–low (LL). Low-demand scenarios reduce flows by 15%, while high-demand scenarios increase flows by 30%. A detailed scenario tree is provided in Fig. 2, where upward branches $((2,1), (3,1), (3,3))$ represent increased demand, and downward branches $((2,2), (3,2), (3,4))$ signify decreased demand.

We assume a uniform 5% cost increase between consecutive periods. At each candidate hub location, three capacity levels — large, medium, and small — are available, with maximum capacities of 30,000 for the AP and TR datasets and 6000 for the CAB dataset. Unit capacity acquisition costs are 0.1 for small, 0.09 for medium, and 0.08 for large
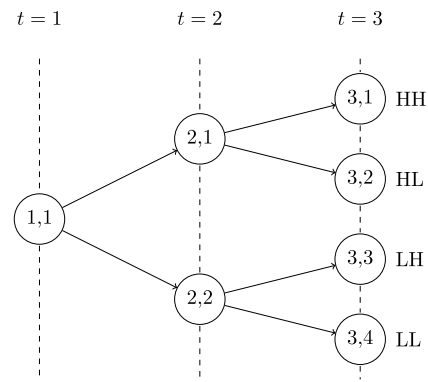

Fig. 2. Scenario tree illustration.

capacities. Seven candidate hub locations are selected from the network nodes based on the highest total inflow and outflow demand, W_i, calculated as: $W_i = \sum_{t \in \mathcal{T}, s \in S_t, k \in \mathcal{K}: o(k)=i \vee d(k)=i} w_k^{ts}$, where w_k^{ts} represents the demand for commodity k in time period t and scenario s. For the TR dataset, each of Türkiye's seven geographical regions can host at most one hub, limiting the candidate hubs to a maximum of seven. Each dataset begins with a Base Instance (BI) and 17 additional instances are generated by varying eight parameters and features. Parameter values follow Bayram et al. (2023) for common elements, with additional values assigned for new parameters in this study. Table 3 details the BI parameter settings for each dataset. The first three rows represent the discount factor, congestion cost, and maximum path length. Rows labeled A_h^t, κ, and B_h^t denote the multipliers defining the maximum travel time relative to the shortest OD path and the costs of opening and closing hubs, respectively. The C_{direct} row represents the direct connection cost multiplier, used to calculate transportation costs for flows routed directly between OD pairs by multiplying the coefficient with the direct distance. The final row indicates the network topology, either complete or incomplete. To construct the incomplete networks, we use a similar approach to that of Bayram et al. (2023). We first calculate a ranking index $\frac{pop_i \times pop_j}{d_{ij}}$ for each origin–destination (OD) pair (i, j), where pop_i and pop_j represent the populations of cities i and j, respectively, and d_{ij} is the shortest path distance between them for TR dataset. For AP and CAB dataset, the ranking index is calculated as $\frac{W_i \times W_j}{d_{ij}}$ where W_i total inflow and outflow demand values of node i. The OD pairs are then ranked in descending order according to their ranking index, and the top 33% of pairs are selected for connection in the IC problem instances. In an incomplete network, if a direct arc for an OD pair is missing, the cost multiplier for that direct path is set to $5 \times C_{direct}$. Table 4 presents the variations in parameters for each instance. For instance, I_1 maintains the same parameter values as the base instance except for α, which is set to 0.5 for all datasets. All subproblems start with a set of initial paths. We generate initial paths as in Bayram et al. (2023). For each commodity $k \in \mathcal{K}$, a direct path, along with single-hub and two-hub paths, is generated. Single-hub paths are limited to a maximum of eight connections from the origin to hubs, selected based on the minimum cost. Similarly, two-hub paths are constrained to a maximum of eight connections from the origin to hubs and from hubs to destinations, with each segment chosen based on the minimum connection cost.

6.2. Performance analysis of the solution algorithm

We present detailed computational performance metrics for the CAB, TR, and AP datasets in Table 6. This table reports the percent relative gap at termination (%Gap), solution time in seconds (CPU(s)), the number of optimality cuts added (#OptCuts), the number of columns generated by the subproblems (#Col.), and the number

Table 3

Parameter setting for the base instances.

Parameter	Dataset		
	AP	CAB	TR
α	0.75	0.75	0.75
λ^t	1000	1000	1000
τ	5	5	5
A_h^t	1	5	1
κ	2.5	2.5	2.5
B_h^t	−0.6	−0.6	−0.6
C_{direct}	4	4	4
NT	Incomplete	Incomplete	Incomplete

Table 4

Parameter setting for additional instances.

Instance	Parameter	Dataset		
		AP	CAB	TR
I_1	α	0.5	0.5	0.5
I_2		1.0	1.0	1.0
I_3	λ^t	0	0	0
I_4		2000	2000	2000
I_5	τ	2	2	2
I_6		3	3	3
I_7		4	4	4
I_8	A_h^t	2	2	10
I_9		3	3	15
I_{10}	κ	1.5	1.5	1.5
I_{11}		3.0	3.0	3.0
I_{12}		4.0	4.0	4.0
I_{13}	B_h^t	−0.8	−0.8	−0.8
I_{14}		−1.0	−1.0	−1.0
I_{15}	C_{direct}	2	2	2
I_{16}		6	6	6
I_{17}	NT	Complete	Complete	Complete

of nodes searched in the branch-and-bound tree, excluding the root node (#BBnode) across all 54 instances. The results obtained from the instances across all datasets indicate that the algorithm is efficient in achieving high-quality solutions, although its performance varies among datasets. The best performance is observed in the AP dataset, with an average gap of 0.74% (standard deviation: 0.23%) across all instances, followed by the CAB dataset with an average gap of 0.90% (standard deviation: 0.66%), and lastly, the TR dataset with an average gap of 2.76% (standard deviation: 1.35%). The maximum gap for each dataset is 1.1%, 2.76%, and 5.92%, respectively. The spread of the relative gap for all datasets is shown in Fig. 3. An analysis of the optimality cuts added during the solution procedure reveals that the TR dataset requires the fewest, averaging 1,642 cuts, followed by the CAB dataset with 2,412, and the AP dataset with the most at 3,281 on average across all instances. Similarly, the AP dataset generates the highest number of columns, averaging 61,711, followed by the TR dataset with 22,187 and the CAB dataset with 1,753. These results align with the network structures, where the CAB dataset features the fewest potential paths, and the AP dataset the most. We observe that instances associated with complete network settings demonstrate longer solution times compared to their incomplete network counterparts, with the CAB and TR datasets exhibiting the largest performance gaps. Additionally, I_3 instances, which assume zero congestion costs, are solved more efficiently and achieve optimality across all datasets. This contrast underscores the heightened complexity introduced by comprehensive models with extensive assumptions and parameters, such as ours, compared to classical hub location problems. Simplifications like ignoring congestion significantly reduce problem complexity, as reflected in the faster solution times for I_3 instances. To assess the impact of extended planning horizons and additional scenarios

on algorithm performance, we solved seven additional instances for each dataset. The planning horizon was extended to a maximum of six periods, with two scenario settings considered: two scenarios per period and three scenarios per period. The results highlight computational challenges in solving instances beyond four periods, particularly for larger datasets. For the AP dataset, the largest among the three, all cases encountered memory limitations. In the TR dataset, the algorithm successfully solved the instance with five periods and two scenarios, achieving a 10.76% optimality gap, but faced memory constraints in all other cases. In the CAB dataset, instances with a two-scenario setting were solved, yielding optimality gaps of 3.33% and 6.96% for five and six periods, respectively, while memory limitations arose in the three-scenario setting. Detailed results are provided in Table 7, presenting the percentage relative gap at termination for various numbers of planning horizons with two scenario options (Sce.) across all datasets. The term "Memory" indicates instances that could not be solved due to memory limitations. Similarly, we investigated the effect of the number of potential hub locations on algorithm performance by solving instances with up to 20 potential hub locations. The results indicate that for fewer than 7 hub locations, the algorithm solves the model to optimality across all datasets. For 7 potential hub locations, the algorithm achieves small optimality gaps of 0.69%, 1.54%, and 0.87% for the CAB, TR, and AP datasets, respectively. As the number of potential hub locations increases, the optimality gap also rises, though at varying rates across datasets. Specifically, for instances with 20 potential hub locations, the gap increases to 9.28% for CAB, 50.62% for TR, and 99.73% for AP. The detailed results are provided in Table 8, reporting the percentage relative gap at termination (%Gap) and the solution time in seconds (CPU(s)) for various numbers of potential hub locations (# Potential Hub) across all datasets.

6.3. Managerial insights

In this section, we begin by discussing the impact of parameter variations on the objective function value and the associated costs. We also examine the effects on the network structure, such as the number of operational hubs and capacity levels. Furthermore, we explore how these variations influence the direct shipment ratio, which reflects the interaction between operational constraints and service quality requirements. Then, in the subsequent three sections, we explore the impact of the three main assumptions of our model: congestion, uncertainty, and multi-periodicity.

6.3.1. Impact analysis

Economies of scale (α): As expected, the total cost decreases as α decreases. A smaller α makes routing through hubs more attractive, prompting the model to favor hub-based routes. This leads to a network with a bigger number of operational hubs and/or higher capacity levels, resulting in increased capacity and congestion costs. However, these increases are outweighed by the savings in routing costs achieved through hub-based routing. Consequently, the total cost decreases with lower economies of scale values. Figs. 4 and 5 illustrate the changes in the number of hubs and the total network capacity for all scenarios throughout the planning horizon for the CAB dataset.

Congestion Cost Coefficient (λ^t): The total cost exhibits a predictable inverse relationship with the congestion cost coefficient across all instances. As congestion costs rise, hub-based consolidation and distribution become increasingly expensive. In response, the model employs two key strategies to minimize hub center loads and manage congestion costs: first, it expands the total network capacity by upgrading existing hubs or establishing new ones; second, it redirects some flow from hub-based routes to direct shipments. Fig. 6 demonstrates how congestion costs influence both network capacity and direct shipment ratios across the datasets. The TR dataset shows increased utilization of both strategies—expanding capacity and increasing direct shipments. However, the AP and CAB datasets primarily rely on
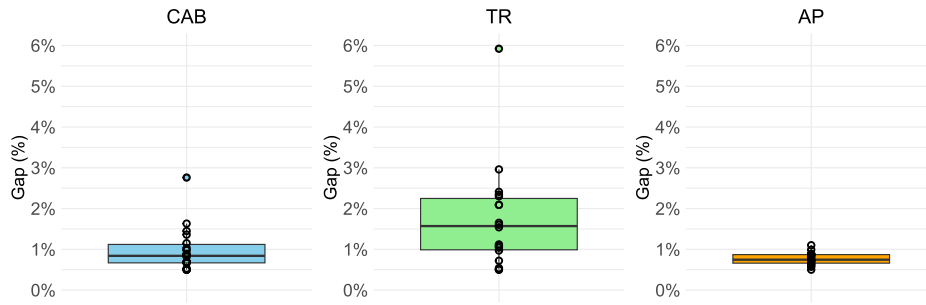
Fig. 3. Percent relative gap at termination for all instances across all datasets.
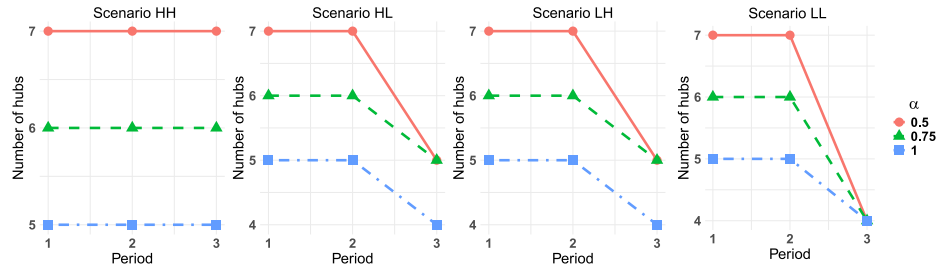
Fig. 4. Effect of α on the number of operating hubs in the CAB dataset.
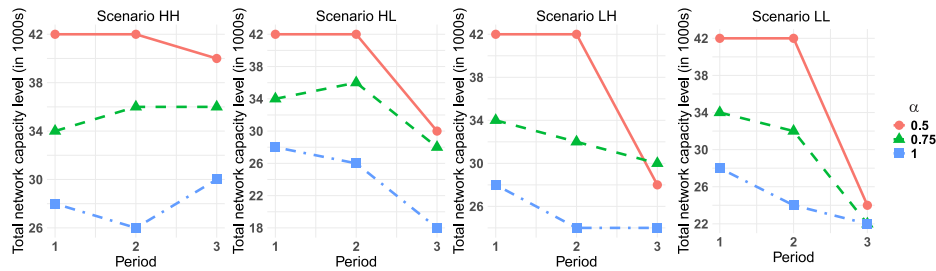
Fig. 5. Effect of α on the total network capacity level in the CAB dataset.
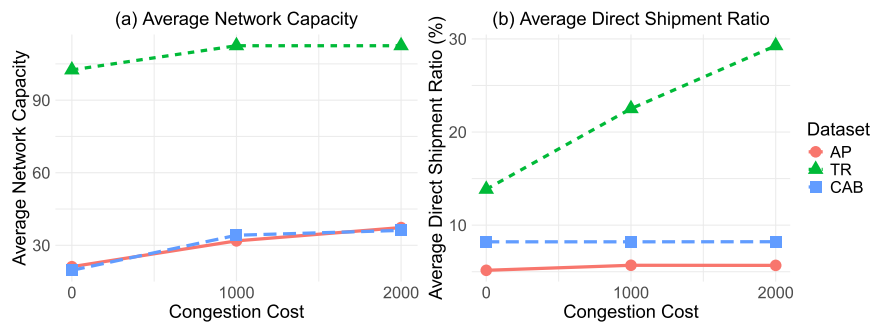
Fig. 6. Effect of λ^t on the average network capacity and direct shipment ratios across all datasets.

network capacity expansion rather than increasing direct shipments. This strategic difference stems from the greater distances in the AP and CAB datasets, which make direct shipping economically less viable compared to the TR dataset.

Hub opening cost (A_h^t): We analyzed the effect of hub opening fixed cost using three different values. The model responds to cost increases in two ways: by opening fewer hubs initially or by closing some hubs during the planning horizon to regain part of the initial opening cost. As a result of operating fewer hubs, higher congestion at hub centers may occur, which must be addressed by acquiring additional capacities and/or incurring higher congestion costs. However, the changes in these costs do not follow the same trend across all

datasets, as shown in Fig. 7. For instance, in the AP dataset, congestion costs increase while capacity costs decrease as the opening cost rises. Conversely, in the TR dataset, a similar pattern is observed as the opening cost increases from 5 to 10, but this trend reverses when the opening cost further increases from 10 to 15.

Hub closing cost (B_h^t): We consider three hub closing cost coefficients: -0.6, -0.8, and -1. As the recovered portion of the initial fixed cost increases, the model gains flexibility to open more hubs at the beginning of the planning horizon and close them later as demand is realized. This increase in the number of operating hubs, in turn, leads to lower congestion and capacity acquisition costs. The closure patterns vary across the tested datasets but generally occur

Fig. 7. Effect of hub opening cost on various transportation cost components across all datasets.

Table 5

Effect of hub closing cost in the number of operating hubs throughout the planning horizon for each scenario across all datasets.

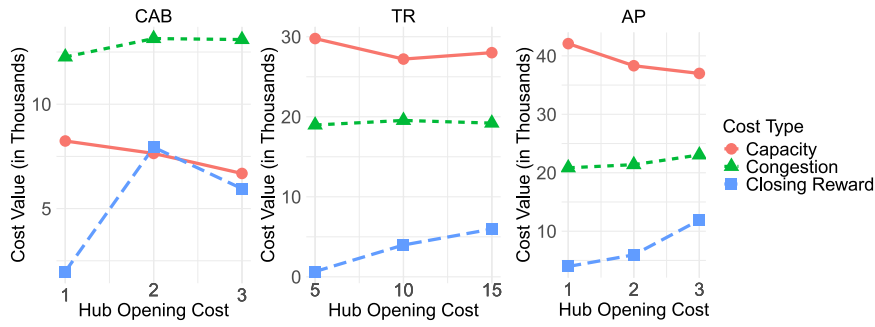| B_h^t	CAB				TR				AP			
	Scenario	Period			Scenario	Period			Scenario	Period		
		1	2	3		1	2	3		1	2	3
−0.6	HH	6	6	6	HH	6	6	6	HH	7	7	7
	HL	6	6	5	HL	6	6	6	HL	7	7	6
	LH	6	6	5	LH	6	6	6	LH	7	7	6
	LL	6	6	4	LL	6	6	4	LL	7	7	5
−0.8	HH	6	7	5	HH	6	6	6	HH	7	7	7
	HL	6	7	4	HL	6	6	6	HL	7	7	5
	LH	6	6	5	LH	6	6	4	LH	7	7	5
	LL	6	6	4	LL	6	6	4	LL	7	7	5
−1	HH	7	7	5	HH	6	6	6	HH	7	7	7
	HL	7	7	5	HL	6	6	4	HL	7	7	5
	LH	7	7	4	LH	6	6	4	LH	7	7	5
	LL	7	7	4	LL	6	6	4	LL	7	7	5

later in the planning horizon, particularly in scenarios with decreasing demand realizations. Table 5 illustrates the number of operating hubs throughout the planning horizon for each scenario across all datasets under different recovery coefficients. The table clearly demonstrates that as the recovery coefficient increases, the frequency of hub closures during the planning horizon also rises.

Maximum travel distance multiplier (κ): To assess the impact of the maximum travel distance commitment on network configuration and associated costs, we set four distinct values for the multiplier κ. As expected, providing a higher level of service, represented by a shorter maximum travel distance, leads to higher total transportation costs. Fig. 8 illustrates the percentage increase in total transportation costs as κ decreases. Although the percentage of increase varies across datasets, with AP being the least affected and TR the most affected on average, all datasets show a drastic cost increase when κ decreases from 2.5 to 1.5. Specifically, the cost increases are 37.56%, 107.79%, and 24.89% for the CAB, TR, and AP datasets, respectively. The direct effect of a shorter maximum travel distance on routing strategies is that some flows are shifted from hub-based routing to direct shipment. As a result, while congestion costs decrease, routing costs increase significantly. Fig. 9 illustrates these trends across all datasets.

Number of visited hubs: The results indicate negligible changes in solutions as the bound on the number of visited hubs decreases from 4 to 3 and then to 2. This is primarily because the restriction on path length dominates over the hub-visit constraint. However, when the threshold reduces to 1, the hub-visit restriction becomes active, leading to less inter-hub flow transfer. Consequently, fewer hubs or lower capacities are required, reducing congestion and capacity costs. However, reduced inter-hub transfers limit the benefit of discounted transportation costs between hubs, increasing routing costs. This routing cost increase outweighs the savings, resulting in higher total transportation costs. Fig. 10 shows the percentage increase in transportation costs as the hub-visit limit decreases. The TR dataset exhibits a substantial cost increase, nearly tripling, while the CAB dataset sees a moderate

7.6% rise, and the AP dataset shows minimal change with a 1% increase. These variations highlight how service quality requirements affect networks differently. Network structure and parameter values significantly influence the cost implications of such restrictions, making service quality standards economically feasible in some networks but prohibitively expensive in others, depending on their characteristics and operational parameters.

Direct connection cost (C_{direct}): We analyzed the impact of direct connection costs on total costs through their effect on the direct shipment ratio. For the AP dataset, the direct shipment ratio in the base instance averages 5.69%. Reducing C_{direct} from 4 to 2 increases this ratio to 8.38%. Similarly, in the CAB and TR datasets, the ratios rise from 8.22% to 9.05% and from 22.51% to 40.45%, respectively. Conversely, increasing C_{direct} from 4 to 6 does not reduce the ratio in the AP and CAB datasets but causes a 6.5% decrease in the TR dataset. Fig. 11 illustrates these changes across all scenarios for the TR dataset. The persistence of direct shipments, even with high C_{direct}, stems from service quality requirements mandating transportation via direct links despite higher costs. As direct shipment costs decrease, the model naturally adjusts by using fewer hubs and/or reducing total network capacity.

Network topology (NT): We further tested the model on a network with a complete graph topology. As expected, complete networks yield lower total transportation costs compared to incomplete ones due to a larger feasible region. However, the cost reduction varies with the network's origin–destination structure, including distances and flows. For example, switching from an incomplete to a complete topology reduces total costs by approximately 1.81% in the AP dataset but results in 55% and 50% reductions in the CAB and TR datasets, respectively. Fig. 12 illustrates that routing costs are most affected by topology changes. The tested datasets reveal distinct trends. In the AP and CAB datasets, both topologies maintain the same number of operating hubs across scenarios, except for one LH scenario where the complete network gains one additional hub in period 3. Conversely, in the TR

Table 6
Performance of the Algorithm Across All Datasets.

Ins.	CAB					TR					AP				
	%Gap	CPU(s)	#OptCuts	#Col.	#BBNode	%Gap	CPU(s)	#OptCuts	#Col.	#BBNode	%Gap	CPU(s)	#OptCuts	#Col.	#BBNode
BI	0.69	172,800	2766	1964	302,585,432	1.54	172,800	1761	24,273	383,519,600	0.87	172,800	3709	63,764	162,862,500
I_1	0.89	172,800	2432	3448	296,011,900	2.30	172,800	1776	27,614	410,541,800	0.76	172,800	2324	160,664	260,010,800
I_2	0.00	18,566	1859	776	44,158,082	1.08	172,800	1650	22,143	438,658,775	0.57	172,800	3530	7	176,006,700
I_3	0.00	119	496	2406	171,704	0.00	994	456	25,563	1,029,340	0.00	10,419	505	62,778	431,114
I_4	1.03	172,800	2439	1695	301,205,515	2.96	172,800	2250	23,540	334,810,900	0.89	172,800	3960	57,119	154,245,659
I_5	0.66	172,800	3018	0	281,013,000	0.54	172,800	1494	0	493,138,400	0.69	172,800	3487	0	198,336,800
I_6	0.69	172,800	2766	1964	305,160,300	2.41	172,800	1754	21,305	426,576,039	0.87	172,800	3700	63,764	167,365,200
I_7	0.69	172,800	2770	1964	303,884,500	1.65	172,800	1766	23,974	389,551,300	0.87	172,800	3700	63,764	166,956,188
I_8	1.63	172,800	2512	1976	283,698,420	1.12	172,800	1642	22,730	420,949,000	0.73	172,800	4285	66,268	127,304,554
I_9	1.36	172,800	2358	2052	324,395,400	2.09	172,800	1579	24,676	423,951,002	1.10	172,800	3424	66,124	155,529,804
I_{10}	0.00	58,258	2211	1022	130,138,668	0.72	172,800	1468	15,186	493,156,215	0.60	172,800	2948	61,727	246,381,420
I_{11}	0.86	172,800	2412	2064	294,269,961	0.97	172,800	1754	26,030	351,258,930	0.81	172,800	3568	64,651	153,474,600
I_{12}	0.98	172,800	2680	2276	259,569,069	0.00	103,415	1395	26,632	258,756,996	0.71	172,800	3807	65,492	157,866,286
I_{13}	1.15	172,800	2232	1773	308,870,733	2.34	172,800	1726	25,789	363,939,715	0.65	172,800	3069	63,461	185,050,100
I_{14}	0.82	172,800	1955	1838	391,631,600	1.60	172,800	1665	28,098	443,816,756	0.99	172,800	3307	62,836	191,817,885
I_{15}	1.45	172,800	2698	1835	305,214,600	1.04	172,800	1300	26,830	512,717,520	0.83	172,800	4014	64,602	145,437,380
I_{16}	0.53	172,800	2583	1906	302,351,357	2.09	172,800	1813	23,820	408,711,079	0.63	172,800	2858	61,494	236,399,549
I_{17}	2.76	172,800	3234	598	184,071,693	5.92	172,800	2298	11,164	272,383,172	0.73	172,800	2872	62,286	204,896,602

Table 7
Impact of Extended Planning Horizons and Additional Scenarios on Algorithm Performance.

Dataset	3 Periods		4 Periods		5 Periods		6 Periods	
	2 Sce.	3 Sce.	2 Sce.	3 Sce.	2 Sce.	3 Sce.	2 Sce.	3 Sce.
CAB	0.69	1.75	1.88	5.72	3.33	Memory	6.96	Memory
TR	1.54	5.78	6.41	11.17	10.76	Memory	Memory	Memory
AP	0.87	1.56	1.47	Memory	Memory	Memory	Memory	Memory

Table 8
Impact of the Number of Potential Hub Locations on Algorithm Performance Across All Datasets.

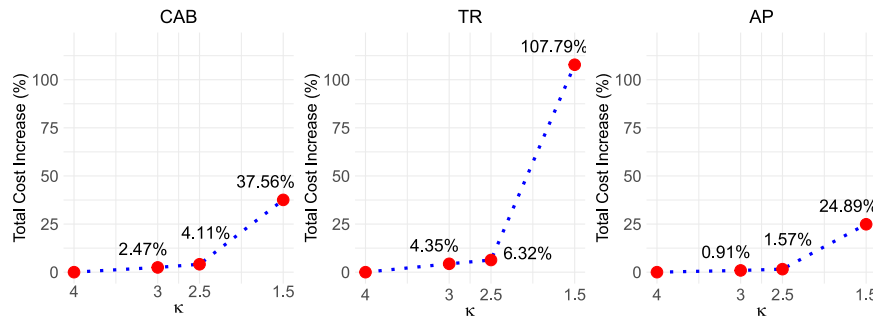| # Potential Hub	CAB		TR		AP	
	%Gap	CPU(s)	%Gap	CPU (s)	%Gap	CPU (s)
4	0.00	42.11	0.00	274.89	0.00	3406.97
5	0.00	1473.48	0.00	1473.83	0.00	9559.02
6	0.00	134448.39	0.00	24078.78	0.00	35800.55
7	0.69	172800.00	1.54	172800.00	0.87	172800.00
8	1.69	172800.00	8.05	172800.00	1.16	172800.00
10	2.94	172800.00	17.18	172800.00	2.25	172800.00
12	4.34	172800.00	24.48	172800.00	3.57	172800.00
14	6.19	172800.00	29.57	172800.00	6.29	172800.00
20	9.28	172800.00	50.62	172800.00	99.73	172800.00

Fig. 8. Effect of κ on percentage increase in total transportation costs across all datasets.

dataset, the complete network consistently operates 1–3 fewer hubs across all scenarios and periods. Regarding direct shipment ratios, complete networks in the AP and CAB datasets exhibit reductions of 0.02% and 3.91%, respectively, compared to incomplete networks. In the TR dataset, however, the complete network shows a 3.04% higher direct shipment ratio. These variations are important and highlight that we cannot expect uniform adjustments in routing strategies when transitioning from one network topology to another. Depending on model parameter values — such as the distance and flow between origin–destination pairs, congestion costs, capacity acquisition costs, and direct shipment costs — and their relative scales and interactions,

entirely different strategies may emerge. These could involve either prioritizing hub-based routing or reducing the number of operating hubs and favoring increased reliance on direct shipments.

6.3.2. Effect of congestion

This section examines the impact of congestion in our model setting. To achieve this, we first solve the model without considering congestion, determining hub locations and capacities in its absence. These decisions are then used as inputs in a subsequent iteration of the original model, where congestion is incorporated. This approach, termed the "Non-Congestion Model", enables a comparison of costs between
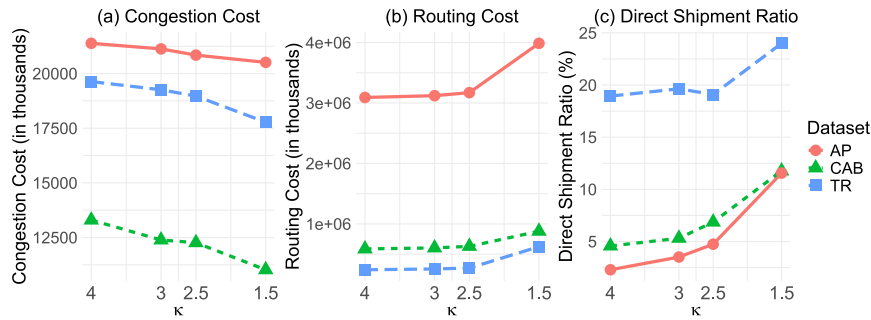
Fig. 9. Effect of κ on congestion and routing costs and average direct shipment ratios across all datasets.
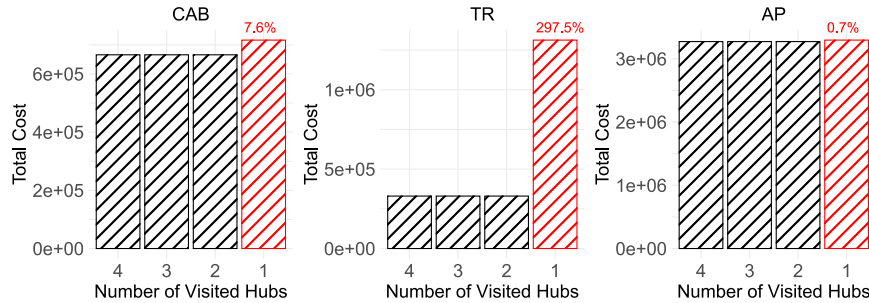

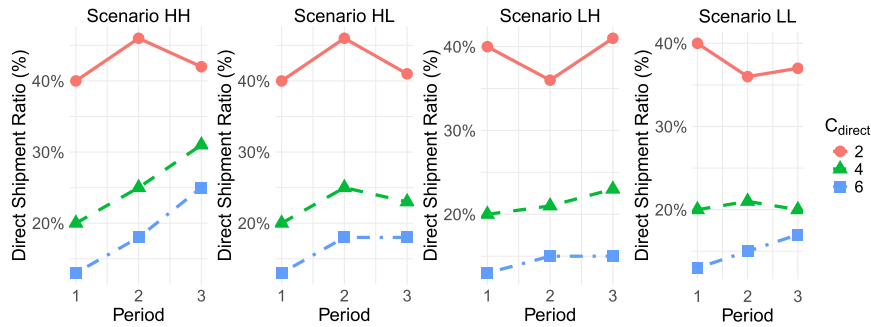**Fig. 10.** Effect of number of visited hubs on the total transportation cost across all datasets.

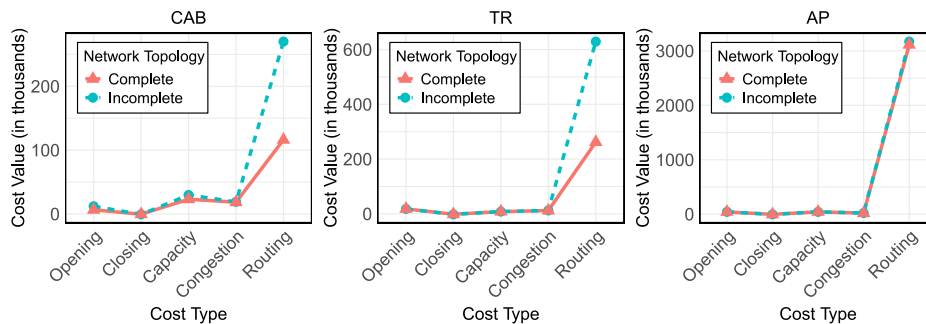**Fig. 11.** Effect of C_{direct} on the direct shipment ratio for the TR dataset.

Fig. 12. Various cost components comparison between complete and incomplete network topologies across all datasets.

solutions with and without congestion, highlighting the effects of neglecting congestion. Following Alumur, Nickel, and Saldanha-da-Gama (2018), we quantify the additional cost of adapting non-congestion-based solutions to scenarios with congestion, referred to as the Value of Congestion (VoC). The detailed results, including the objective function values for non-congestion models (Non-congestion OF) and the value of congestion (in percentages), are presented in Table 9 for all datasets. The results demonstrate that accounting for congestion affects total transportation costs to varying extents across datasets. The CAB dataset is most impacted, with an average cost increase of 2.72% (maximum 4.78%), followed by the TR dataset at 1.42% (maximum 3.35%), and the AP dataset at 0.78% (maximum 1.11%). Interestingly, these impacts do not align with the proportion of congestion costs in the original model, which are 1.92%, 5.54%, and 0.62% for CAB, TR, and AP, respectively. Despite TR having the highest congestion cost proportion, CAB experiences the greatest impact from neglecting congestion.
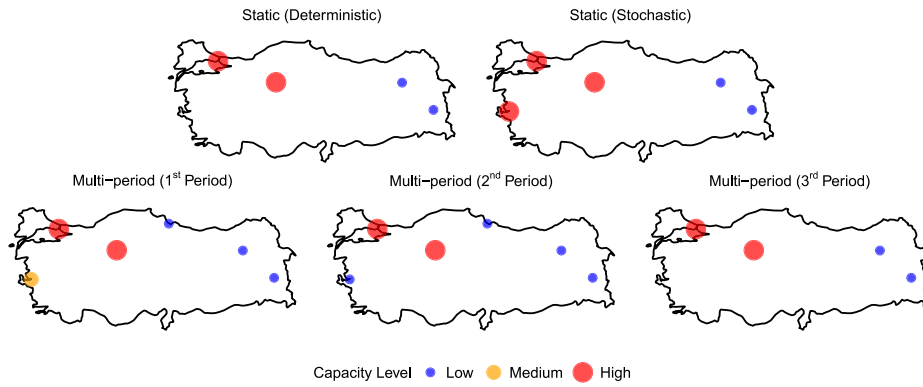
Fig. 13. Comparison of Static vs. Multi-period Networks for the LL Scenario (TR Base Instance).
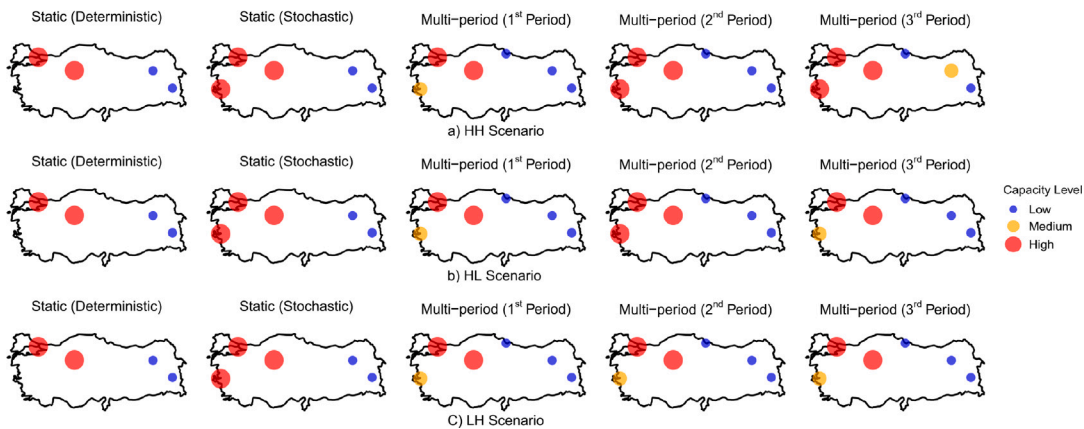

Fig. 14. Comparison of Static vs. Multi-period Networks for HH, HL, and LH scenarios (TR Base Instance).

Changes in network structure are also observed, with 15 out of 16 CAB instances, 9 TR instances, and 12 AP instances showing adjustments in the number of hubs and capacities. Hub capacities are modified in all instances, and routing decisions, including the volume of direct shipments, are affected in all but two AP instances.

6.3.3. Effect of demand uncertainty

This section outlines an experiment to evaluate the benefits of incorporating stochasticity into the model. First, we assume a specific scenario (e.g., LL) and solve it as a deterministic problem, disregarding uncertainty. The resulting hub locations and capacities are then applied to other scenarios, effectively examining the outcomes of designing the network for one scenario while facing another. This process is repeated for all four scenarios (LL, LH, HL, HH), yielding 16 objective function values. From these, we compute the average outcome under a deterministic approach, termed the "Expected Deterministic Cost". Comparing this value with the stochastic model's result quantifies the additional cost incurred, referred to as the "Value of Stochastic Modeling" (VoSM), representing the expected cost increase due to demand uncertainty. Table 10 provides a comparative assessment of the expected deterministic cost and the value of stochastic modeling (in percentages) across all instances. Unlike the congestion effect, the TR dataset shows the highest sensitivity to ignoring demand uncertainty, with an average cost increase of 0.73%, compared to 0.21% for CAB and 0.17% for AP. Notably, seven instances — one from CAB, four from TR, and two from AP — are rendered infeasible because the deterministic model's hub operations and network capacity cannot handle realized demand fluctuations. This highlights that a stochastic approach not only reduces costs but also ensures system robustness and resilience, mitigating potential inefficiencies.

6.3.4. Effect of multi-periodicity

This section evaluates the impact of a multi-period planning by comparing the proposed multi-period model with its static counterpart. Various approaches may be employed to define the static counterpart to a multi-period model (Bakker & Nickel, 2024). Following a two-step approach similar to Alumur, Nickel, Saldanha-da-Gama, and Verter (2012), we first solve the static model by setting demand to its average values over the planning horizon, determining hub locations and capacities. These decisions are then fixed and applied to the original multi-period model to compute the static counterpart's objective function value. The relative arithmetic difference between this value and that of the multi-period model, termed the "Value of Multi-Period Solution" (VoMPS), quantifies the expected cost decrease when transitioning from a static to a multi-period framework. Table 10 presents the detailed results for all datasets, reporting the Expected Deterministic (ED) Cost, the Value of Stochastic Modeling (VoSM) as percentages, the objective function values for static models (Static OF), and the Value of Multi-Period Solution (VoMPS) as percentages. The "" sign indicates that the problem is infeasible. The results indicate average cost increases of 2.54%, 2.45%, and 0.52% for the CAB, TR, and AP datasets, respectively, with maximum increases of 8.60% (CAB), 4.13% (TR), and 0.81% (AP). These maximum values for CAB and TR are associated with instances having complete network structures, where fewer operating hubs and lower capacities (compared to incomplete networks) make it harder to absorb demand fluctuations during the planning horizon. Notably, the static model consistently establishes fewer hubs and smaller capacities, particularly for the CAB and TR datasets, compared to the multi-period model, with these differences being more significant under high demand fluctuations. Fig. 14 illustrates these variations for the TR dataset base instance across all

Table 9

Objective Function Values for Non-congestion Models and Value of Congestion (VoC) as Percentages Across All Datasets.

Instance	CAB		TR		AP	
	Non-congestion OF	VoC (%)	Non-congestion OF	VoC (%)	Non-congestion OF	VoC (%)
BI	682483.10	2.52	336631.07	1.94	3293007.51	0.63
I_1	644416.00	3.37	334366.28	1.29	3099415.61	0.83
I_2	710268.97	2.36	335157.73	1.33	3405015.72	0.95
I_5	733568.59	2.42	1316702.49	0.31	3332274.08	1.11
I_6	682483.10	2.52	335476.80	1.58	3293007.51	0.63
I_7	682483.10	2.52	334974.15	1.41	3293007.51	0.63
I_8	697950.29	2.41	345807.68	1.31	3332221.07	0.72
I_9	714220.76	2.77	355785.94	1.25	3366713.92	0.74
I_{10}	936077.38	2.22	690075.18	0.57	4119072.63	0.79
I_{11}	659411.36	3.12	316379.59	1.86	3251320.75	0.92
I_{12}	640996.34	2.72	300037.61	0.81	3224573.53	0.99
I_{13}	680022.21	2.25	335261.58	1.49	3294180.84	0.73
I_{14}	678725.18	2.27	334302.29	1.36	3292811.44	0.73
I_{15}	484727.62	3.43	230976.10	1.07	3150447.37	0.64
I_{16}	879126.80	1.90	431944.07	1.76	3425366.29	0.65
I_{17}	312981.77	4.78	169119.14	3.35	3238626.56	0.79

Table 10

Values of Expected Deterministic (ED) Cost, Value of Stochastic Modeling (VoSM) as Percentages, Objective Function Values for Static and Value of Multi-Period Solution (VoMPS) as Percentages Across All Datasets.

Ins.	CAB				TR				AP			
	ED Cost	VoSM (%)	Static OF	VoMPS (%)	ED Cost	VoSM (%)	Static OF	VoMPS (%)	ED Cost	VoSM (%)	Static OF	VoMPS (%)
BI	667480.7319	0.27	678015.25	1.85	332692.5875	0.75	341350.99	3.37	3277210.263	0.15	3289323.06	0.52
I_1	624275.6925	0.14	640124.68	2.68	331855.9606	0.53	340906.84	3.27	3080771.243	0.23	3098606.77	0.81
I_2	694953.6913	0.15	696430.68	0.36	333324.6325	0.78	341665.04	3.30	–	–	3375341.36	0.07
I_3	–	–	–	–	–	–	–	–	–	–	–	–
I_4	678891.7075	0.29	688298.25	1.68	349214.3269	0.69	358811.04	3.45	3294801.249	0.14	3307500.27	0.52
I_5	717303.5056	0.15	726657.88	1.46	–	–	1318138.03	0.42	3301521.009	0.18	3313319.4	0.54
I_6	667480.7319	0.27	678015.25	1.85	332626.2569	0.72	341350.99	3.36	3277210.263	0.15	3289323.06	0.52
I_7	667480.7319	0.27	678015.25	1.85	332647.5731	0.70	341350.99	3.34	3277210.263	0.15	3289323.06	0.52
I_8	683103.9425	0.23	704597.68	3.39	343600.8475	0.67	349446.02	2.38	3313698.352	0.16	3318442.55	0.31
I_9	698006.0494	0.43	713597.68	2.68	354179.1863	0.79	357541.05	1.75	3344675.582	0.08	3365191.75	0.69
I_{10}	916059.0881	0.04	–	–	–	–	–	–	4092837.258	0.14	4115143.34	0.69
I_{11}	641021.8463	0.24	649408.39	1.56	312630.1638	0.66	316424.85	1.88	3229522.492	0.24	3243401.14	0.67
I_{12}	625201.5606	0.18	646429.4	3.59	300113.6669	0.83	303427.05	1.95	3198862.292	0.19	3207669.72	0.46
I_{13}	665924.7406	0.13	678015.25	1.95	332418.915	0.63	341350.99	3.33	3275450.984	0.16	3289323.06	0.58
I_{14}	664442.6481	0.12	678015.25	2.16	332133.7269	0.70	341350.99	3.49	3276445.849	0.23	3289323.06	0.62
I_{15}	469491.93	0.18	486783.74	3.87	229951.285	0.62	228983.12	0.20	3136338.711	0.19	3144724.14	0.46
I_{16}	864510.5775	0.21	873153.28	1.21	–	–	424278.96	0.00	3409291.428	0.18	3424085.21	0.61
I_{17}	299547.8213	0.29	324388.56	8.60	165408.8406	1.08	170410.14	4.14	3218122.731	0.15	3230765.83	0.55

scenarios. We observed operational failures in five instances — two each from the CAB and TR datasets and one from AP — where the static model's hub locations and capacities were insufficient to meet increased demand in later periods. This highlights the necessity of a multi-period approach, which adapts to fluctuating conditions over time, ensuring robustness and capacity to handle all scenarios. An alternative analysis considered a two-stage stochastic model, similar to Bayram et al. (2023), although their model excludes service quality concerns. Hub locations and capacities determined by their model were integrated into our multi-period model to evaluate the impact of short-term planning. The average Value of Multi-Period Solution (VoMPS) percentages were 0.92% (CAB), 4.10% (TR), and 0.31% (AP), with infeasibility occurring in one CAB instance. In terms of network structure, our model demonstrates flexibility in adjusting hub locations and capacities throughout the planning horizon to minimize the impact of fluctuations and uncertainty. For instance, Fig. 13 compares the network structure for the base instance of the TR dataset under LL scenarios. Similar comparisons for other scenarios are provided in Fig. 14. In the computational result illustrations, we refer to the static counterpart of the first approach as the "Deterministic Static" model and the static counterpart of the second approach as the "Stochastic Static" model.

In conclusion, it is worth noting that while one of the main objectives of this paper is to provide a unified and comprehensive framework. The numerical results demonstrate that incorporating each of the aforementioned features can lead to cost savings to varying extents, as well as enhanced reliability and robustness for the entire system. However, these effects should not be interpreted as universally applicable across all cases. Changes in any parameter or problem setting can influence the magnitude and direction of these effects. For instance, as discussed earlier, instances from the AP dataset exhibit lower sensitivity, whereas TR instances show more pronounced responses to parameter changes or model adjustments. Therefore, in real-world applications, selecting appropriate cost parameter ranges and conducting detailed sensitivity analyses are crucial for obtaining realistic insights and providing practical guidance closely aligned with the specific problem under investigation.

7. Conclusion

This study presents a hub network design problem that integrates three pivotal factors: congestion, demand uncertainty, and multi-periodicity. Different from conventional approaches, which often address these factors independently, our methodology concurrently considers them, providing a more comprehensive understanding of the complexities inherent in hub network design. Moreover, our model incorporates service level considerations for network users, extending beyond the conventional focus solely on transportation costs. Service quality is evaluated using two metrics: travel time and the number of hubs visited during a journey. Furthermore, our model exhibits dynamic characteristics, allowing for adjustments in capacity levels and network configuration throughout the planning period, thereby introducing practical adaptability to the problem setting.

CRediT authorship contribution statement

Vedat Bayram: Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Çiya Aydoğan:** Validation, Software, Investigation, Formal analysis, Data curation, Conceptualization. **Kamyar Kargar:** Writing – review & editing, Writing – original draft, Visualization, Validation, Investigation, Formal analysis, Data curation.

Acknowledgment

References

Alizadeh, F., & Goldfarb, D. (2003). Second-order cone programming. *Mathematical Programming, 95*(1), 3–51.

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. *Transportation Research Part C (Emerging Technologies), 102*, 249–273.

Aloullal, A., Saldanha-da-Gama, F., & Todosijević, R. (2023). Multi-period single-allocation hub location-routing: Models and heuristic solutions. *European Journal of Operational Research, 310*(1), 53–70.

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 Journal of Operational Research, 291*(1), 1–17.

Alumur, S. A., & Kara, B. Y. (2008). Network hub location problems: The state of the art. *European Journal of Operational Research, 190*(1), 1–21.

Alumur, S. A., Nickel, S., Rohrbeck, B., & Saldanha-da-Gama, F. (2018). Modeling congestion and service time in hub location problems. *Applied Mathematical Modelling, 55*, 13–32.

Alumur, S. A., Nickel, S., & Saldanha-da-Gama, F. (2012). Hub location under uncertainty. *Transportation Research, Part B (Methodological), 46*(4), 529–543.

Alumur, S. A., Nickel, S., Saldanha-da-Gama, F., & Verter, V. (2012). Multi-period reverse logistics network design. *European Journal of Operational Research, 220*(1), 67–78.

Azizi, N., Vidyarthi, N., & Chauhan, S. S. (2018). Modelling and analysis of hub-and-spoke networks under stochastic demand and congestion. *Annals of Operations Research, 264*(1), 1–40.

Bakker, H., & Nickel, S. (2024). The value of the multi-period solution revisited: When to model time in capacitated location problems. *Computers & Operations Research, 161*, Article 106428.

Bayram, V., & Yaman, H. (2018). Shelter location and evacuation route assignment under uncertainty: A benders decomposition approach. *Transportation Science, 52*(2), 416–436.

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

Beasley, J. E. (2018). *OR-library: Hub location.* OR-Library, URL: http://people.brunel.ac.uk/~mastjjb/jeb/info.html. (Accessed 30 July 2024).

Ben-Tal, A., & Nemirovski, A. (2001). Lectures on modern convex optimization: Analysis, algorithms, and engineering applications. *MPS-SIAM Series on Optimization.*

Benders, J. F. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik, 4*(1), 238–252.

Blanco, V., Fernández, E., & Hinojosa, Y. (2023). Hub location with protection under interhub link failures. *INFORMS Journal on Computing, 35*(5), 966–985.

Bütün, C., Petrovic, S., & Muyldermans, L. (2021). The capacitated directed cycle hub location and routing problem under congestion. *European Journal of Operational Research, 292*(2), 714–734.

Campbell, J. F. (1990). Locating transportation terminals to serve an expanding demand. *Transportation Research, Part B (Methodological), 24*(3), 173–192.

Campbell, J. F. (1994). Integer programming formulations of discrete hub location problems. *European Journal of Operational Research, 72*(2), 387–405.

Campbell, J. F. (2009). Hub location for time definite transportation. *Computers & Operations Research, 36*(12), 3107–3116.

Campbell, J. F. (2013). A continuous approximation model for time definite many-to-many transportation. *Transportation Research, Part B (Methodological), 54*, 100–112.

Contreras, I. (2015). Hub location problems. In G. Laporte, S. Nickel, & F. Saldanha da Gama (Eds.), *Location science* (pp. 311–344). Cham: Springer International Publishing.

Contreras, I., Cordeau, J. F., & Laporte, G. (2011). The dynamic uncapacitated hub location problem. *Transportation Science, 45*(1), 18–32.

Correia, I., Nickel, S., & Saldanha-da-Gama, F. (2010). Single-assignment hub location problems with multiple capacity levels. *Transportation Research, Part B (Methodological), 44*(8), 1047–1066.

Correia, I., Nickel, S., & Saldanha-da-Gama, F. (2018). A stochastic multi-period capacitated multiple allocation hub location problem: Formulation and inequalities. *Omega, 74*, 122–134.

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

de Camargo, R. S., & Miranda, G. (2012). Single allocation hub location problem under congestion: Network owner and user perspectives. *Expert Systems with Applications, 39*, 3385–3391.

de Sá, E. M., Morabito, R., & de Camargo, R. S. (2018). Benders decomposition applied to a robust multiple allocation incomplete hub location problem. *Computers & Operations Research, 89*, 31–50.

Dhyani Bhatt, S., Jayaswal, S., Sinha, A., & Vidyarthi, N. (2021). Alternate second order conic program reformulations for hub location under stochastic demand and congestion. *Annals of Operations Research, 304*(1–2), 481–527.

Domínguez-Bravo, C. A., Fernández, E., & Lüer-Villagra, A. (2024). Hub location with congestion and time-sensitive demand. *European Journal of Operational Research, 316*(3), 828–844.

Dror, M. (1994). Note on the complexity of the shortest path models for column generation in VRPTW. *Operations Research, 42*(5), 977–978.

Elhedhli, S., & Hu, F. X. (2005). Hub-and-spoke network design with congestion. *Computers & Operations Research, 32*(6), 1615–1632.

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

Ernst, A. T., & Krishnamoorthy, M. (1996). Efficient algorithms for the uncapacitated single allocation p-hub median problem. *Location Science, 4*(3), 139–154.

FAA (2020). Cost of delay estimates. URL: https://www.faa.gov/data_research/aviation_data_statistics/media/cost_delay_estimates.pdf.

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

Feillet, D., Dejax, P., Gendreau, M., & Gueguen, C. (2004). An exact algorithm for the elementary shortest path problem with resource constraints: Application to some vehicle routing problems. *Networks, 44*(3), 216–229.

Forbes (2019). Airlines struggle to cope with rush-hour-style congestion. URL: https://www.forbes.com/sites/oliverwyman/2019/09/19/airlines-struggle-to-cope-with-rush-hour-style-congestion/?sh=7bd91b8a2594.

Gelareh, S., Monemi, R. N., & Nickel, S. (2015). Multi-period hub location problems in transportation. *Transportation Research, E: Logistics and Transportation Review, 75*, 67–94.

Gelareh, S., & Nickel, S. (2008). Multi-period public transport design: A novel model and solution approaches.

Ghaffarinasab, N. (2018). An efficient matheuristic for the robust multiple allocation p-hub median problem under polyhedral demand uncertainty. *Computers & Operations Research, 97*, 31–47.

Ghaffarinasab, N. (2022). Stochastic hub location problems with Bernoulli demands. *Computers & Operations Research, 145*, Article 105851.

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. *Transportation Research, Part B (Methodological), 167*, 32–53.

Ghaffarinasab, N., Kara, B. Y., & Campbell, J. F. (2022). The stratified p-hub center and p-hub maximal covering problems. *Transportation Research, Part B (Methodological), 157*, 120–148.

Ghaffarinasab, N., Zare Andaryan, A., & Ebadi Torkayesh, A. (2020). Robust single allocation p-hub median problem under hose and hybrid demand uncertainties: models and algorithms. *International Journal of Management Science and Engineering Management, 15*(3), 184–195.

Guldmann, J. M., & Shen, G. (1997). A general mixed integer nonlinear optimization model for hub network design. In *44th North American meeting of the regional science association international.*

Günlük, O., & Linderoth, J. (2008). Perspective relaxation of mixed integer nonlinear programs with indicator variables. In A. Lodi, A. Panconesi, & G. Rinaldi (Eds.), *Integer programming and combinatorial optimization* (pp. 1–16). Berlin: Springer.

Hu, Q. M., Hu, S., Wang, J., & Li, X. (2021). Stochastic single allocation hub location problems with balanced utilization of hub capacities. *Transportation Research, Part B (Methodological), 153*, 204–227.

Irnich, S., & Desaulniers, G. (2005). Shortest path problems with resource constraints. In G. Desaulniers, J. Desrosiers, & M. M. Solomon (Eds.), *Column generation* (pp. 33–65). Springer.

Ishfaq, R., & Sox, C. R. (2010). Intermodal logistics: The interplay of financial, operational and service issues. *Transportation Research, E: Logistics and Transportation Review, 46*(6), 926–949.

Ishfaq, R., & Sox, C. R. (2012). Design of intermodal logistics networks with hub delays. *European Journal of Operational Research, 220*(3), 629–641.

Jayaswal, S., & Vidyarthi, N. (2023). Multiple allocation hub location with service level constraints for two shipment classes. *European Journal of Operational Research*, *309*(2), 634–655.

Jittrapirom, P., Caiati, V., Feneri, A. M., Ebrahimigharehbaghi, S., González, M., & Narayan, J. (2017). Mobility as a service: A critical review of definitions, assessments of schemes, and key challenges. *Urban Planning*, *2*(2), 13–25.

Kara, B. Y. (2011). *Hub location*. Bilkent University, URL: https://ie.bilkent.edu.tr/~bkara/hub_location.php. (Accessed 04 September 2019).

Kargar, K., & Mahmutoğulları, A. I. (2022). Risk-averse hub location: Formulation and solution approach. *Computers & Operations Research*, *143*, Article 105760.

Karimi, H. (2018). The capacitated hub covering location-routing problem for simultaneous pickup and delivery systems. *Computers & Industrial Engineering*, *116*, 47–58.

Khaleghi, A., & Eydi, A. (2021). Robust sustainable multi-period hub location considering uncertain time-dependent demand. *RAIRO-Operations Research*, *55*(6), 3541–3574.

Khaleghi, A., & Eydi, A. (2023). Multi-period hub location problem considering polynomial time-dependent demand. *Computers & Operations Research*, *159*, Article 106357.

Kian, R., & Kargar, K. (2016). Comparison of the formulations for a hub-and-spoke network design problem under congestion. *Computers & Industrial Engineering*, *101*, 504–512.

Kleinrock, L. (2007). *Communication nets: Stochastic message flow and delay*. Courier Corporation.

Klincewicz, J. G. (1998). Hub location in backbone/tributary network design: a review. *Location Science*, *6*(1), 307–335.

Lin, C. C., & Lee, S. C. (2018). Hub network design problem with profit optimization for time-definite LTL freight transportation. *Transportation Research, E: Logistics and Transportation Review*, *114*, 104–120.

Lobo, M. S., Vandenberghe, L., Boyd, S., & Lebret, H. (1998). Applications of second-order cone programming. *Linear Algebra and its Applications*, *284*(1), 193–228.

Macias, J. E., Angeloudis, P., & Ochieng, W. (2020). Optimal hub selection for rapid medical deliveries using unmanned aerial vehicles. *Transportation Research Part C (Emerging Technologies)*, *110*, 56–80.

Macrina, G., Pugliese, L. D. P., Guerriero, F., & Laporte, G. (2020). Crowd-shipping with time windows and transshipment nodes. *Computers & Operations Research*, *113*, Article 104806.

Marianov, V., & Serra, D. (2003). Location models for airline hubs behaving as $M/D/c$ queues. *Computers & Operations Research*, *30*(7), 983–1003.

Marín, A., Nickel, S., Schöbel, A., & Sonneborn, T. (2002). Extensions of the uncapacitated hub location problem for applications in intermodal public transportation. In *13th mini-EURO conference and IX meeting of the EURO working group on transportation*.

Mohammadi, M., Jolai, F., & Rostami, H. (2011). An M/M/c queue model for hub covering location problem. *Mathematical and Computer Modelling*, *54*(11–12), 2623–2638.

Montreuil, B. (2011). Toward a physical internet: meeting the global logistics sustainability grand challenge. *Logistics Research*, *3*(2–3), 71–87.

Najy, W., & Diabat, A. (2020). Benders decomposition for multiple-allocation hub-and-spoke network design with economies of scale and node congestion. *Transportation Research, Part B (Methodological)*, *133*, 62–84.

Neamatian Monemi, R., Gelareh, S., Nagih, A., Maculan, N., & Danach, K. (2021). Multi-period hub location problem with serial demands: A case study of humanitarian aids distribution in Lebanon. *Transportation Research, E: Logistics and Transportation Review*, *149*, Article 102201.

Nesterov, Y., Nemirovskii, A. S., & Ye, Y. (1994). *Interior-point polynomial algorithms in convex programming*: *vol. 13*, (pp. 379–402). SIAM.

O'Kelly, M. E. (1986). The location of interacting hub facilities. *Transportation Science*, *20*(2), 92–106.

Peiró, J., Corberán, Á., Martí, R., & Saldanha-da-Gama, F. (2019). Heuristic solutions for a class of stochastic uncapacitated p -hub median problems. *Transportation Science*, *53*(4), 1126–1149.

Rahmaniani, R., Crainic, T. G., Gendreau, M., & Rei, W. (2017). The Benders decomposition algorithm: A literature review. *European Journal of Operational Research*, *259*(3).

Rahmati, R., Neghabi, H., Bashiri, M., & Salari, M. (2023). Stochastic regional-based profit-maximizing hub location problem: A sustainable overview. *Omega, 121*, Article 102921.

Salimian, M. (2013). *A mixed integer second order cone programming reformulation for a congested location and capacity allocation problem in supply chain network design* (Master's thesis), Turkey: METU.

Sener, N., & Feyzioglu, O. (2022). Multiple allocation hub covering flow problem under uncertainty. *Annals of Operations Research*, *320*(2), 975–997.

Serper, E. Z., & Alumur, S. A. (2016). The design of capacitated intermodal hub networks with different vehicle types. *Transportation Research, Part B (Methodological)*, *86*, 51–65.

Taherkhani, G., Alumur, S. A., & Hosseini, M. (2020). Benders decomposition for the profit maximizing capacitated hub location problem with multiple demand classes. *Transportation Science*, *54*(6), 1446–1470.

Tavassoli, K., & Tamannaei, M. (2020). Hub network design for integrated Bike-and-Ride services: A competitive approach to reducing automobile dependence. *Journal of Cleaner Production*, *248*, Article 119247.

Tikani, H., Ramezanian, R., Setak, M., & Van Woensel, T. (2021). Hybrid evolutionary algorithms and Lagrangian relaxation for multi-period star hub median problem considering financial and service quality issues. *Engineering Applications of Artificial Intelligence*, *97*, Article 104056.

Wang, S., Chen, Z., & Liu, T. (2020). Distributionally robust hub location. *Transportation Science*, *54*(5), 1189–1210.

Yaman, H., & Elloumi, S. (2012). Star p-hub center problem and star p-hub median problem with bounded path lengths. *Computers & Operations Research*, *39*(11), 2725–2732.

Yaman, H., Kara, B. Y., & Tansel, B. Ç. (2007). The latest arrival hub location problem for cargo delivery systems with stopovers. *Transportation Research, Part B (Methodological)*, *41*(8), 906–919.

Yang, T. H., & Chiu, T. Y. (2016). Airline hub-and-spoke system design under stochastic demand and hub congestion. *Journal of Industrial and Production Engineering*, *33*(2), 69–76.