Case Studies on Transport Policy 25 (2026) 101829

Contents lists available at ScienceDirect

Case Studies on Transport Policy

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

Multimodal two-echelon city logistics under space and time limitations: A case study of Amsterdam

Bas Bijvoet [a,*,1], Cigdem Karademir [b,1], Bilge Atasoy [b]

[a] National Road Traffic Data Portal, Utrecht, The Netherlands
[b] Department of Maritime and Transport Technology, Faculty of Mechanical Engineering, Delft University of Technology, Delft, 2628 CD, The Netherlands

ARTICLE INFO

Keywords:
Multimodal city logistics
Two-echelon vehicle routing
Service network design
Scenario analysis
Case study

ABSTRACT

This study explores how integrating inland waterways into multimodal distribution systems can enhance city logistics, alleviate street-level congestion, and ultimately improve the livability of the urban environment. To capture the operational complexities of such systems under severe space scarcity and regulatory access constraints, we formulate a multi-trip two-echelon vehicle routing problem. The model explicitly accounts for the physical limitations of dense city centers by incorporating storage-free satellites and spatial constraints for vehicle occupancy. While the first requires precise spatiotemporal synchronization between interacting vehicles during the transshipment operation from one mode to the other, the latter bounds the maximum number of transshipments occurring at the same time at a satellite. To evaluate system performance under these conditions, we develop an optimization framework driven by an iterative decomposition-based heuristic. The approach integrates a capacitated assignment model with routing heuristics through an adaptive workload-bounding feedback loop, ensuring that downstream routing constraints actively shape upstream customer-to-satellite assignments to find a feasible solution that attains the target service level while heuristically minimizing resource consumption. The methodology is demonstrated through a large-scale case study of a multimodal distribution system in Amsterdam, serving over 750 HoReCa businesses. To derive strategic insights from this operational model, we conduct a comprehensive scenario analysis of 10560 instances. The proposed framework identifies the minimum operational resources required to guarantee service coverage by systematically evaluating diverse strategic decisions in terms of network design, transshipment modalities, workforce levels, and city access time windows. The results illustrate practical trade-offs for urban logistics planning: relaxing full-coverage service targets (e.g., to 90%) provides substantial infrastructure savings, while denser satellite networks reduce street-level travel distances and increase zero-emission walking deliveries. Furthermore, enabling parallel transshipments with an adaptive workforce allows a significantly smaller network to maintain full service coverage.

1. Introduction

Global urbanization is an ongoing trend. The share of the world's population living in urban areas has increased from 30% in 1950 to 55% in 2018, and it is projected to reach 68% by 2050 (United Nations, 2019). This shift is accompanied by increased urban transportation demands. The resulting congestion, air and noise pollution, and traffic accidents have become concerns for policymakers (Browne et al., 2012). At the same time, growing consumer expectations and the rise of e-commerce have intensified the demand for fast and flexible delivery services, increasing freight activity and further straining urban logistics systems (Savelsbergh and Van Woensel, 2016). Enhancing city logistics while mitigating negative impacts on accessibility, safety, and sustainability has therefore become a shared challenge for European cities (European Commission, 2017).

This study is motivated by the urban logistics challenges faced by Amsterdam, the capital city of the Netherlands (Gemeente Amsterdam, 2020a, 2021b). In addition to increased transport movements driven by population and tourism growth, heavy logistics traffic also overburdens the city's historic infrastructure, which includes roughly 1800 bridges and 600 km of quay walls. Many of these structures are over a century old and were not designed to handle such heavy loads (Gemeente Amsterdam, 2020b,d). This has led to widespread structural deterioration, with around 850 bridges and 200 km of quay walls requiring extensive repair and renewal (Gemeente Amsterdam, 2022a). These

* Corresponding author.
E-mail addresses: bjbijvoet@gmail.com (B. Bijvoet), cigdem.karadem@gmail.com (C. Karademir), b.atasoy@tudelft.nl (B. Atasoy).
[1] Co-first authors, these authors contributed equally to this work.

https://doi.org/10.1016/j.cstp.2026.101829
Received 11 January 2026; Received in revised form 1 April 2026; Accepted 29 April 2026
Available online 13 May 2026

infrastructure works are expected to take several decades and will likely cause prolonged logistical disruptions. Meanwhile, freight operations are facing tightening restrictions on vehicle access hours, size, and emissions, further complicating urban logistics.

To address these multifaceted challenges, Amsterdam is exploring smart mobility solutions and new logistics strategies to enhance the city's livability and accessibility, while meeting the growing needs of residents and visitors (Gemeente Amsterdam, 2022b, 2019). A key initiative is the integration of the city's extensive inland waterway network of canals for urban freight transport. The municipality actively supports this modal shift to reduce road congestion and promote environmentally friendly transport. This effort is complemented by the development of logistical hubs on the city's periphery, enabling the bundling and transshipment of goods onto smaller, cleaner vehicles for last-mile distribution (Gemeente Amsterdam, 2021a). The municipality aims to develop clear regulatory frameworks, including designated loading and unloading locations and appropriate access time windows. Ongoing pilots and studies are examining the practical aspects and implications of this modal shift, aiming to build a reliable and efficient logistics system for both service providers and policymakers (Gemeente Amsterdam, 2020c).

To enable multimodal city logistics, distribution networks increasingly adopt multi-echelon structures, wherein goods pass through successive transshipment points before final delivery (Cuda et al., 2015). Typically, large vehicles transport goods from distribution centers (DCs) located outside the city center to intermediate transshipment facilities (also known as satellites) closer to customers within the city center. Then, smaller and low-emission vehicles complete the final leg of the delivery by transporting goods from satellites to customers (Cattaruzza et al., 2017). In addition to enabling the use of different vehicles across network levels, multi-echelon systems also enable shipment consolidation at DCs. This reduces delivery frequency, lowers logistics costs for suppliers and carriers, and minimizes environmental impacts (Savelsbergh and Van Woensel, 2016).

We model multimodal urban freight logistics as a space–time constrained two-echelon vehicle routing problem (2E-VRP). As illustrated in Fig. 1, a two-echelon distribution network involves DC-to-satellite transportation in the first echelon (FE) by first-echelon vehicles (FEVs), followed by satellite-to-customer transportation in the second echelon (SE) by second-echelon vehicles (SEVs). To capture the operational complexities of city logistics, the formulated 2E-VRP jointly incorporates storage-free satellites, vehicle occupancy limits, load- and SEV-type dependent transshipment times, and city access time restrictions. Furthermore, it also considers heterogeneous SEVs and multi-trip operations for all vehicles on both echelons. By integrating these features, the model is tailored to dense urban environments, where spatial and temporal limitations restrict the maximum transshipment throughput achievable at satellites.

To analyze the strategic trade-offs inherent in such systems, we develop an evaluation framework based on an iterative decomposition heuristic. We apply this framework to a large-scale case study of a multimodal distribution system serving hotels, restaurants, and cafés (HoReCa) in Amsterdam. By systematically evaluating diverse operational scenarios, we quantify the impact of varying service network designs, fleet compositions, and city access time restrictions. Ultimately, this study aims to provide insights for policymakers and logistics operators on how to balance infrastructure investments, labor, and urban space to achieve sustainable multimodal city freight distribution.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature and highlights the contributions of this study. Section 3 presents the problem statement, detailing the operational constraints and problem variants based on the available resources at satellites. Section 4 outlines the iterative decomposition-based heuristic developed to integrate allocation and routing decisions. Section 5 introduces the case study setup, while Section 6 discusses the computational experiments and derived operational and strategic insights. Finally, Section 7 provides concluding remarks and outlines directions for future research.

2. Literature review

The 2E-VRP models freight distribution systems that employ two distinct vehicle fleets delivering goods from a central depot to customers via intermediate facilities known as satellites (Perboli et al., 2011). This hierarchical structure provides enhanced flexibility, environmental benefits, and operational cost reductions relative to traditional single-echelon systems (Agnimo et al., 2023). The 2E-VRP generalizes the standard VRP, an NP-hard problem. This intrinsic complexity is further exacerbated by the strong interdependence between first- and second-echelon vehicle routes, necessitating the use of heuristic and metaheuristic solution methods for practical applications (Cuda et al., 2015; Sluijk et al., 2023; Soares et al., 2024). Although exact optimization techniques offer valuable insights for small instances, their computational demands limit their applicability to large-scale and real-world problems typical of urban logistics (Marques et al., 2022; Karademir et al., 2025). Given our focus on large-scale and realistic applications, this review focuses on heuristic approaches developed to address more intricate 2E-VRP variants, highlighting key modeling features and the research gaps that this study aims to fill.

2.1. Problem variants and attributes

This section provides a structured overview of key problem attributes commonly addressed in 2E-VRPs. We discuss critical modeling dimensions, reflecting diverse assumptions on satellite and fleet capacity management, and relevant implications for feasibility, coordination, and flexibility.

2.1.1. Satellite capacity constraints

In 2E-VRPs, assumptions regarding satellite capacity for storing goods, accommodating vehicles, and performing transshipment operations play a central role in shaping synchronization requirements between FEVs and SEVs.

No Satellite Storage. In the strictest synchronization regime, satellites function purely as transshipment locations without any intermediate storage, requiring exact spatial and temporal synchronization between FEVs and SEVs to enable direct transshipment of goods. The lack of storage capacity can cause vehicles to experience waiting times at the satellite, leading to increased operational costs. This strict synchronization framework is formalized in Grangier et al. (2016) for urban logistics, where satellites do not accommodate storage. Similarly, Anderluh et al. (2017, 2021) model two-echelon systems with storage-free satellites, thereby introducing complex synchronization challenges and enforcing tighter temporal coupling between routing decisions across both echelons.

Limited Satellite Storage. These models allow for partial decoupling of vehicle arrivals by buffering inbound and outbound flows, subject to finite capacity constraints at the satellites. For example, Li et al. (2018) introduce a dynamic transshipment capacity model bounded by a maximum threshold, enabling more flexible coordination and relaxing strict synchronization requirements. In a related context, Li et al. (2020) consider a two-echelon van-drone delivery system where satellites are not fixed locations but are instead represented by FEVs paired with specific SEVs, effectively functioning as mobile transshipment points. Building on the concept of limited storage, Dumez et al. (2023) propose a variant featuring capacitated satellites supporting temporary storage, where competing forward and reverse flows are regulated by time-dependent capacity constraints.

Unlimited Satellite Storage. An alternative modeling paradigm assumes unlimited or effectively unconstrained storage capacity at satellites, thereby removing the need for strict synchronization between FEV and SEV operations (Li et al., 2021; Zhou et al., 2022; Lehmann and Winkenbach, 2024; Zamal et al., 2025). This assumption facilitates asynchronous drop-offs and pickups, with flow constraints ensuring

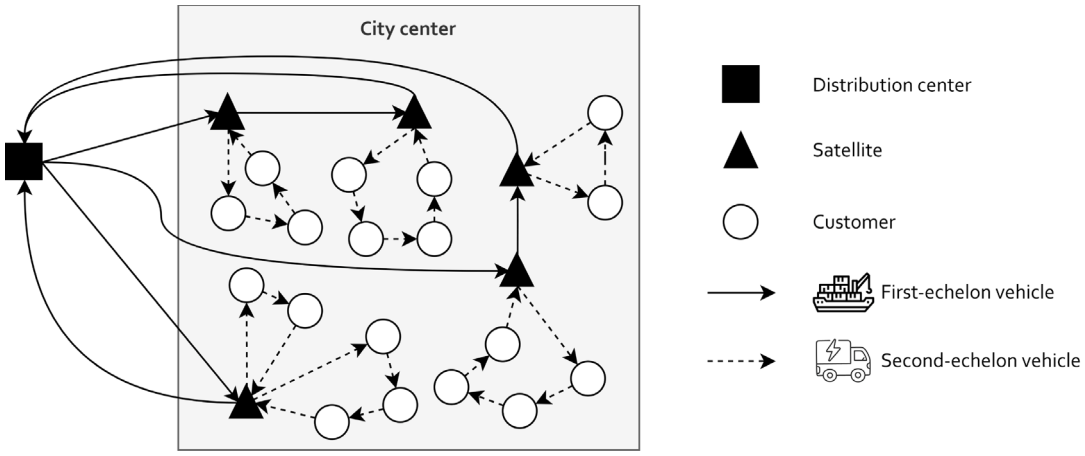


Fig. 1. Illustration of the multimodal two-echelon vehicle routing problem network. First-echelon vessels transport goods between the DC and satellites (solid lines), while second-echelon street vehicles perform routes connecting satellites to customers (dashed lines) within the city center area.

that FEV deliveries precede SEV pickups, thus decoupling vehicle arrivals and departures to a high degree. Although this approach simplifies vehicle coordination and enhances operational flexibility, it may underestimate the spatial limitations in dense urban logistics environments, making it less realistic for such constrained contexts.

Vehicle Occupancy Constraints. To enable two-echelon distribution in urban logistics without requiring major infrastructure changes, satellites are often assumed to operate without storage, leveraging existing spaces such as parking lots (Sluijk et al., 2023). These locations typically face spatial limitations or restricted access that constrain the number of FEVs and SEVs that can be present at the same time. Transshipment operations are often modeled to occur in parallel, allowing an FEV to transfer goods to multiple SEVs simultaneously, thereby simplifying the synchronization requirements. While Anderluh et al. (2021) enforce maximum waiting times for the vehicles at the satellites to reduce the congestion, the explicit representation of satellite occupancy and spatial constraints on vehicle presence remains unexplored within storage-free systems (Grangier et al., 2016; Anderluh et al., 2017). As a result, transshipment is typically represented as a flow through the satellite facility, with an emphasis on time coordination and flow balance rather than direct vehicle-to-vehicle interaction. This does not limit the system in terms of the number of transshipment operations at any time and raises concerns about the generalizability of the research findings to the urban settings where space is scarce. Incorporating such constraints introduces additional modeling and algorithmic complexity to manage vehicle presence at the satellite.

Transshipment Time Variability. The modeling of transshipment time in 2E-VRPs varies widely across the literature, ranging from instantaneous transfers to more complex, load-dependent, or synchronization-driven delays. Grangier et al. (2016) and Li et al. (2020) adopt a simplification where transshipment is considered instantaneous or negligible. Most studies assume a fixed non-zero transfer time, representing a constant duration for the exchange of goods, as seen in Anderluh et al. (2017, 2021), Li et al. (2021), Dumez et al. (2023), Lehmann and Winkenbach (2024), and Zamal et al. (2025). A more detailed approach involves load-dependent modeling, where transfer duration scales with the volume handled, as in Li et al. (2018) and Zhou et al. (2022).

Gap and Distinction. Existing studies that model exact synchronization without satellite storage typically assume parallel transshipment operations with no occupancy constraints (Grangier et al., 2016; Anderluh et al., 2017, 2021). This simplification, combined with instantaneous or fixed transshipment times, can overestimate the achievable transshipment throughput at satellites and is unrealistic for constrained urban environments. Our study addresses this critical gap by introducing a 2E-VRP variant with storage-free satellites that includes explicit satellite occupancy constraints, where limited space allows only one FEV to be present at any time. Furthermore, we explore both parallel and sequential transshipment modalities, and incorporate load-dependent (and SEV-type-dependent) transshipment times to accurately capture the scheduling complexity induced by resource-constrained operations.

2.1.2. Routing and fleet constraints

Fleet capacity is a critical determinant of operational feasibility and efficiency in 2E-VRP models. Choices regarding second- and first-echelon routing and fleet sizing reflect practical considerations in urban logistics while balancing computational tractability.

Second-Echelon Routing. The modeling of SE routing exhibits substantial variation across the literature, particularly in terms of trip frequency, vehicle-to-satellite assignment, and the incorporation of backhaul requirements. A majority of studies support multi-trip operations for SEVs, enhancing fleet utilization and reducing the number of required vehicles. Within this paradigm, Grangier et al. (2016), Anderluh et al. (2017, 2021), and Dumez et al. (2023) enable flexible vehicle reassignment, allowing SEVs to initiate successive trips from different satellites. Li et al. (2020) address a 2E-VRP where SEVs are being dispatched from FEVs to serve customers, effectively allowing SEVs to operate in a free-floating manner across the service area within their paired routes (with each SEV paired to a specific FEV). In contrast, several studies impose dedicated satellite assignments, restricting each SEV to operate from a single satellite only, including the studies of Li et al. (2018) and Lehmann and Winkenbach (2024). This implies that SEV routes are planned independently per satellite, reflecting operational setups with limited coordination between satellite facilities or to simplify routing and synchronization. A more restrictive modeling approach is adopted by Li et al. (2021), Zhou et al. (2022), and Zamal et al. (2025), where SEVs are limited to single-trip operations, inherently implying a fixed assignment to a single satellite. These distinctions in SEV routing flexibility and operational constraints not only shape algorithmic design but also reflect assumptions about practical operations in dense urban logistics networks.

First-Echelon Routing. Multi-trip FEV routing, where vehicles return to the depot multiple times for replenishment between successive deliveries and pickups, remains scarce in 2E-VRP. Only Dumez et al. (2023), who model capacitated satellites in an urban context, allow for multi-trip FEV routing. Other studies restrict FEVs to a single trip per service type (e.g., one delivery and, if applicable, one pickup tour per vehicle) or further restrict operations by limiting each FEV to a single satellite visit. However, in urban contexts, where distances are typically short, imposing a single-trip restriction on vehicles may unnecessarily compromise vehicle utilization and necessitate a larger fleet of capacitated

vehicles to meet service requirements. Enabling multi-trip operations allows the feasibility of additional trips to be endogenously determined by operational constraints like time windows and satellite availability, rather than being imposed as a structural modeling assumption.

Fleet Sizing. A key distinction lies in whether the fleet size is fixed a priori or optimized as part of the objective function. Several studies adopt an unbounded fleet approach, where the number of vehicles is not constrained but instead minimized or penalized through fixed costs. For instance, Grangier et al. (2016) employ a lexicographic objective function that prioritizes minimizing the number of vehicles on each echelon before optimizing routing costs, while others include fixed vehicle costs in the objective function (Anderluh et al., 2017, 2021; Li et al., 2020, 2021; Zamal et al., 2025). This allows the model to endogenously determine not only the cost-efficient but also the required fleet size as a result of the optimization process. In contrast, Li et al. (2018) and Zhou et al. (2022) fix the number of vehicles on both echelons and focus on minimizing operational costs without directly penalizing vehicle usage, effectively decoupling fleet sizing from the optimization objective. Dumez et al. (2023) and Lehmann and Winkenbach (2024) adopt a hybrid approach, where fleet sizes are bounded but still subject to optimization within those limits via cost-based incentives. These modeling choices greatly affect operational feasibility. When fleet sizes are unconstrained, additional limits on transshipments operations and vehicle presence at satellites are needed to ensure solutions remain operationally realistic.

Gap and Distinction. Despite the growing interest in multi-trip routing, existing 2E-VRP literature does not integrate multi-trip operations for both SEVs and FEVs within storage-free systems. Rigid pairings between satellites, FEVs, and SEVs hinder vehicle repositioning and operational flexibility in space-limited urban settings. Furthermore, models of storage-free systems that omit both satellite occupancy constraints and fleet size limits risk compensating for complex synchronization requirements by deploying an unrealistically large number of vehicles. Our study closes these gaps by modeling multi-trip operations for both FEVs and SEVs. We also integrate a heterogeneous SEV fleet to enhance operational flexibility by providing trade-offs between fleet decisions and capacity requirements in the satellite network. Crucially, the FEV fleet deployment is endogenously constrained by satellite availability and occupancy limitations. On the other hand, the SEV fleet size is either fixed or optimized within upper bounds to reflect the congestion limits of the city center. This combined modeling of routing, fleet flexibility, and realistic resource constraints represents a novel and necessary extension for urban logistics.

2.2. Heuristic solution methods

Most heuristic solution methods for the 2E-VRP begin by building an initial solution, often by first creating SE routes from satellites to customers. Customers are typically assigned to their closest satellite, and a separate VRP is solved for each satellite. The total demand at each satellite and SE arrival and departure times at satellites then shape the FE routes. A key challenge is managing synchronization between these two echelons to ensure feasible transfers and coordinated timing, with methods differing primarily in whether synchronization is addressed sequentially through staged procedures or jointly within integrated frameworks. While various heuristics have been proposed, no single method consistently outperforms others across all problem variants (Sluijk et al., 2023), as diverse and complex problem extensions shape solution requirements differently.

Some heuristics use a staged approach, breaking the problem into separate phases for each echelon. For instance, Anderluh et al. (2017, 2021) develop methods that first construct and optimize SE routes before building FE routes, handling synchronization through flexible waiting times and post-optimization adjustments. Li et al. (2018, 2020, 2021) adopt a similar staged structure but introduce a cyclic refinement strategy, using Variable Neighborhood Search (VNS) or Adaptive Large Neighborhood Search (ALNS) to iteratively improve both echelons. Their approach, while more tightly coupled, still relies on enforcing synchronization through iterative feedback and repair mechanisms.

In contrast, some heuristics adopt an integrated optimization approach, jointly tackling both echelons within a single framework. Grangier et al. (2016) propose an ALNS for the 2E-VRP with exact synchronization, jointly optimizing FE and SE routes while embedding synchronization directly into the search process using a precedence graph with forward-time slacks adapted from Masson et al. (2013). Building on this concept of feasibility checks for neighborhood moves using a precedence graph with forward-time slacks, Dumez et al. (2023) and Lehmann and Winkenbach (2024) employ metaheuristics that combine heuristic search with exact optimization. Dumez et al. (2023) introduce an iterative fix-and-optimize procedure integrating two LNS components for each echelon with an MIP model, whereas Lehmann and Winkenbach (2024) couple an ALNS for the second echelon with an exact MILP for the first.

Gap and Distinction. Integrated optimization methods achieving constant time feasibility evaluation using forward time slacks on precomputed precedence graphs are not applicable to our setting (Grangier et al., 2016; Dumez et al., 2023; Lehmann and Winkenbach, 2024). To avoid the high complexity of variable synchronization, existing models typically assume transfer times are either zero (Grangier et al., 2016; Li et al., 2020) or a fixed duration (Anderluh et al., 2017, 2021; Li et al., 2021; Dumez et al., 2023; Lehmann and Winkenbach, 2024; Zamal et al., 2025). Conversely, the few studies that do incorporate load-dependent times rely on temporary satellite storage to decouple vehicle arrivals (Li et al., 2018; Zhou et al., 2022). Since our setting requires exact synchronization for load-dependent transfers at strictly storage-free satellites, the synchronization window between vehicles becomes variable per transfer. Thus, the proposed problem variants, which require jointly optimizing routes in this resource-constrained and space–time limited satellite network, result in non-linear feasibility functions. To address this computational complexity, this study introduces an iterative decomposition-based heuristic. While problem decomposition is necessary to handle non-linear operational constraints, it risks losing integration if the tactical assignment and operational routing phases are not tightly coupled. Rather than relying solely on route-level perturbations like local search or destroy-and-repair operators (Grangier et al., 2016; Zhou et al., 2022), or on post-hoc trip recombination (Dumez et al., 2023; Lehmann and Winkenbach, 2024; Zamal et al., 2025) to indirectly adjust assignments, our framework restores integration through a structural workload-bounding feedback loop. This framework decomposes the integrated problem into an exact capacitated customer-to-satellite assignment problem and hierarchical routing and scheduling subproblems. By employing an iterative allocation-routing-bounding loop, downstream routing infeasibilities explicitly update the upstream workload bounds for the next iteration. Consequently, the framework dynamically refines satellite workload bounds and reallocates customers in response to operationally overloaded satellites. Through this feedback mechanism, the iterative process progressively guides the solution toward a higher service level.

2.3. Research contributions

A structured comparison of recent contributions to the 2E-VRP literature is provided in Table 1, highlighting their treatment of satellite constraints, transshipment operations, multi-trip routing, and fleet characteristics. This study extends the existing literature by addressing several critical, interconnected operational constraints relevant to real-world city logistics. The main contributions of this research are:

- We introduce multi-trip 2E-VRP variants tailored for city logistics to enhance operational flexibility and efficiency in dense

Table 1
Comparison of the proposed 2E-VRP variant with related literature.

Paper	Satellite constraints		Transshipment operations			Multi-trip		Fleet type[a]		Fleet limit		Application	
	Storage	Occupancy	Time	Sequential	Parallel	SE	FE	SE	FE	SE	FE	Dataset	Customers
Grangier et al. (2016)	✗	✗	Instant		✓	✓	✗	HM	HM	✗	✗	Synthetic	200
Anderluh et al. (2017)	✗	✗	Fixed		✓	✓	✗	HM	HM	✗	✗	Case study	125
Li et al. (2018)	✓	✗	Proportional			✓	✗	HM	HM	✓	✓	Case study	900
Li et al. (2020)	✓	✗	Instant			✓	✗	HM	HM	✗	✗	Synthetic	100
Anderluh et al. (2021)	✗	✗	Fixed		✓	✓	✗	HM	HM	✗	✗	Case study	100
Li et al. (2021)	✓	✗	Fixed			✗	✗	HT	HM	✗	✗	Case study	4080
Zhou et al. (2022)	✓	✗	Proportional			✗	✗	HM	HM	✓	✓	Synthetic	200
Dumez et al. (2023)	✓	✗	Fixed			✓	✓	HM	HM	✓	✓	Synthetic	100
Lehmann and Winkenbach (2024)	✓	✗	Fixed			✓	✗	HM	HM	✓	✓	Synthetic	100
Zamal et al. (2025)	✓	✗	Fixed			✗	✗	HM	HM	✗	✗	Case study	2150
This study	✗	✓	Proportional	✓	✓	✓	✓	HT	HM	✓	✗	Case study	750

[a] HM = homogeneous fleet, HT = heterogeneous fleet

urban areas under realistic transshipment operations. The problem incorporates resource-constrained, storage-free satellites that explicitly account for limited vehicle occupancy, labor capacity, load-dependent transshipment times, and distinct transshipment modalities, requiring precise spatiotemporal synchronization between vehicles.
- We develop a flexible iterative decomposition-based heuristic approach to manage the complexity of this rich variant. To maintain coordination between subproblems, the modular framework integrates a capacitated assignment model with routing heuristics via an adaptive workload-bounding feedback loop. It evaluates real-life operational constraints, specifically satellite workload and workforce allocation, by using downstream routing infeasibilities to dynamically tighten upstream workload bounds to progressively guide the solution toward a higher service level.
- We demonstrate our methodology through a large-scale real-world case study of a multimodal distribution system (integrating waterborne and road-based transport) serving 750 HoReCa businesses in Amsterdam, requiring the scheduling of approximately 300 synchronized transshipment operations. To derive strategic insights from the operational model, we conduct a comprehensive scenario analysis. By evaluating the system across varying satellite infrastructure, workforce levels, and time window policies, the framework quantifies resource trade-offs. Developed in collaboration with public authorities, this approach identifies the minimum operational resources required to guarantee feasibility, offering practical guidelines for strategic network design.

3. Problem statement

This section defines the problem by detailing key assumptions and operational activities at both echelons and satellites for synchronization. The main characteristics of the problem include multi-trip routes on both echelons, storage-free and resource-constrained satellites, and distinct transshipment modalities. The research scope and main assumptions are summarized as follows:

- A single DC is considered, with upstream supply to the DC excluded from the study.
- The distribution of goods between the DC, satellites, and customers is executed by carriers under a centralized decision-making authority that determines routes and schedules.
- The planning horizon spans one day, with all customer demand known at the time of route planning, reflecting orders placed prior to the requested delivery date.
- Empty containers and other reusable items collected by SEVs must be returned to the waiting FEV before it can depart the satellite.
- Spatial constraints at satellites limit vehicle occupancy to a single FEV at any given time.
- Operations are bounded by a uniform city access time window $\mathcal{W}$ that regulates logistics activities within the city center.

Although this problem is formulated around a waterborne and road-based transport network, it addresses core operational challenges relevant to a broader class of urban logistics systems. The requirements for storage-free satellites necessitating precise spatiotemporal synchronization for the transshipment of goods, and strict vehicle occupancy limits at satellites, both represent constraints common in dense urban environments. The proposed problem setting can be applied to other multimodal systems that face structurally analogous challenges.

3.1. FE logistics operations

The FE consists of a fleet of homogeneous FEVs that transport goods between the DC and satellites. Each FEV can perform multiple trips, always starting and ending at the DC. The multi-trip attribute of FEVs may reduce the required fleet size to serve larger demands. During a trip, an FEV may visit one or more satellites to deliver and collect goods from SEVs. At the DC, the time to load and unload one unit of goods onto and from an FEV is denoted by $\theta_{DC}^{FEV\rightarrow}$ and $\theta_{DC}^{FEV\leftarrow}$, respectively. The travel time between two locations a and b for FEVs is denoted by $\tau_{a,b}^{FEV}$.

Due to the limited physical space assumption at satellites in the FE, only one FEV can occupy a satellite at any given time. If a satellite is occupied by another FEV, subsequent FEVs must wait until it becomes available. Although this waiting incurs no cost and is not subject to time restrictions, the scheduling and routing of FEVs aim to minimize such waiting times. Direct deliveries from the DC to customers by FEVs are not permitted.

3.2. Satellite transshipment operations

Transshipment operations involve the logistics of goods between FEVs and SEVs. These operations are subject to critical resource constraints and, therefore, require synchronization at satellites, which impacts the system's operational efficiency and feasibility.

- **Storage constraints:** Satellites do not have storage capacity, requiring that FEVs and SEVs be present simultaneously at the same satellite for a transshipment to occur. This requirement imposes strict spatial and temporal synchronization between FEVs and SEVs.
- **Vehicle occupancy constraints:** Satellites can be visited multiple times by the same or different FEVs. However, only one FEV can occupy the satellite at any given time due to limited physical space. When the combined demand of customers assigned to a satellite exceeds the capacity of FEVs, multiple supply visits to that satellite are necessary, potentially causing waiting times for FEVs until space becomes available.

- **Transshipment time variability:** In addition to being load-dependent, the time required to transship goods may vary per satellite and SEV type, reflecting differences in handling operations. The time to load one unit of goods from an FEV onto an SEV of type k at satellite s for delivery is denoted by $\theta_s^{k\rightarrow}$, and the time to unload one unit of collected goods from an SEV of type k back onto an FEV at satellite s is denoted by $\theta_s^{k\leftarrow}$.
- **Transshipment modality:** Satellite resources, such as space for SEVs and lifting capabilities at satellites or FEVs, can limit the number of units transshipped concurrently. Transshipment operations are categorized into two distinct variants based on how FEVs transfer goods to SEVs: one-to-many (`1TM`) and one-to-one (`1T1`). In the `1TM` variant, an FEV is capable of servicing multiple SEVs simultaneously, facilitating parallel transshipment operations. In contrast, the `1T1` variant restricts transshipments to a sequential process where goods are transferred to a single SEV at a time. This sequential constraint can lead to waiting times for SEVs and necessitates addressing a specific scheduling problem at the satellite, as the sequence of operations directly impacts efficiency.

3.3. SE logistics operations

At the SE, a set of customers C must be served within the planning horizon. Each customer $c \in C$ has delivery demand d_c, and a pickup demand representing empty containers from previous deliveries. To ensure practical and efficient operations, pickups are restricted to customers with delivery demand. Furthermore, the pickup demand is assumed to be less than or equal to the customer's delivery demand to maintain balanced vehicle trip loads. The following practical considerations motivate this design choice:

- **Minimized route detours:** Limiting pickups to delivery locations eliminates detours to isolated pickup points, resulting in shorter and more efficient SE routes, particularly beneficial in congested urban areas.
- **Streamlined customer interaction:** Coordinating pickups with deliveries consolidates interactions and eliminates the need for separate appointments, as customers know that pickups occur during deliveries.
- **Coordinated satellite operations and SEV–FEV load synchronization:** Balanced pickup and delivery loads enable more predictable flows, supporting improved load consolidation and alignment with FEV capacity.

The containers collected by the SEVs are transported back to the satellites. These containers are subsequently returned to the DC by FEVs, requiring strict synchronization to ensure all SEVs complete their deliveries and pickups and return to the satellite before the FEV departs. This waiting constraint is a modeling choice driven by the storage-free nature of the satellites. If vessels do not wait, backhaul items would require intermediate storage or frequent fragmented FEV trips to collect small batches of returns, which would increase vessel traffic and congestion. Therefore, requiring FEVs to wait until all SEVs have returned allows backhaul flows to be consolidated efficiently while preserving the storage-free operation of the satellites and avoiding additional vessel movements.

A heterogeneous fleet of SEV types $k \in \mathcal{K}$ with capacity Q^k is operated by personnel assigned to the satellites. Each satellite $s \in \mathcal{S}$ is staffed with a number of personnel $\mathcal{P}_s$ who perform SE trips. This number is also referred to as the workforce of a satellite. Each person $p \in \{1, \dots, \mathcal{P}_s\}$ is dedicated to a single satellite s, can operate any SEV type k, and can perform multiple trips within the planning horizon. Each SEV type k at satellite s has a service range ρ_s^k, limiting its use to customers within this area. If a customer's demand exceeds the SEV capacity, multiple visits are required to fully serve them. The service range supports decision-making about which SEV type is best suited to serve a given customer, with customers located closer to the satellite potentially served by different vehicle types than those farther away based on accessibility and vehicle capabilities.

An SE trip consists of serving one or more customers, starting and ending at the assigned satellite. The total trip time includes transshipment time at the satellite for loading goods to be delivered onto the SEV before departure, travel time from the satellite to the customer(s), service time at the customer(s) for delivering and collecting goods, travel time back to the satellite, and transshipment time to unload the collected goods onto the FEV upon return.

The service time at each customer is proportional to their demand and may vary by both customer and SEV type, reflecting differences in handling operations for deliveries and pickups. The per-unit service time at customer c for delivery by SEV type k is denoted by $\theta_c^{k\rightarrow}$, while the per-unit service time for backhaul pickup is denoted by $\theta_c^{k\leftarrow}$. The travel time between locations a and b for SEV type k is denoted by $\tau_{a,b}^k$.

3.4. City access and system-wide time constraints

Operational activities in the multimodal distribution system are bounded by a uniform city access time window span $\mathcal{W}$. This parameter reflects urban policies that regulate vehicle presence and logistics activities within a specific geographic area. Although the DC is located outside this restricted zone, the forward distribution of goods is restricted to fall within $\mathcal{W}$, starting from the initial loading of FEVs at the depot. This conservative modeling assumption prevents the use of unbounded pre-loading time and ensures that the entire forward process operates under a single time constraint. The return trip of the FEVs to the DC is treated differently and is not subject to this time restriction. The window $\mathcal{W}$ is defined to encompass all critical urban operations of an FEV, from initial loading at the DC to the completion of its last transshipment (including both delivery and backhaul collection) at the final satellite visited within the urban area, thus completing all associated SE activities. Consequently, any subsequent outbound travel or backhaul unloading at the DC falls outside the regulatory scope of urban access policies, as these post-operational activities neither impact city congestion nor occupy satellite resources.

3.5. Workforce allocation and transshipment modalities

Two problem variants are proposed to represent workforce capacity constraints at satellites within the multimodal distribution system: the Fixed Workforce (`FW`) variant and the Adaptive Workforce (`AW`) variant. In both variants, the actual number of personnel assigned to satellite s is denoted by $\mathcal{P}_s$. In the `FW` variant, this workforce is predefined and fixed as an input parameter $\mathtt{P}_s$, such that $\mathcal{P}_s = \mathtt{P}_s$, allowing for varying workforce levels across satellites but remaining constant throughout the optimization process. Conversely, the `AW` variant models workforce allocation dynamically, where $\mathcal{P}_s$ is a decision variable determined by the iterative framework, subject to an upper bound $\mathtt{P}_s^{\max}$ provided as an input parameter ($\mathcal{P}_s \leq \mathtt{P}_s^{\max}$). This flexible workforce allocation mechanism enables the model to adjust personnel distribution based on the workload of each satellite, balancing resource use according to operational demands. By employing a distinct formulation, the model supports different representations of workforce allocation management as a tactical decision in multimodal logistics operations.

The transshipment modality, denoted by $\mu_s \in \{\mathtt{1TM}, \mathtt{1T1}\}$ for each satellite $s \in \mathcal{S}$, determines whether transshipments at the satellite can be performed sequentially (`1T1`) or in parallel (`1TM`). Fig. 2 illustrates both transshipment modalities, highlighting how transshipments can overlap in the `1TM` setting but must be performed sequentially under the `1T1` setting, resulting in waiting time and a longer service time for the FEV to deliver and collect goods at the satellite. Since satellites are assumed to be public spaces, no infrastructural investments are proposed to facilitate transshipments at these locations. Hence, although the modality is defined per satellite, the actual lifting equipment is

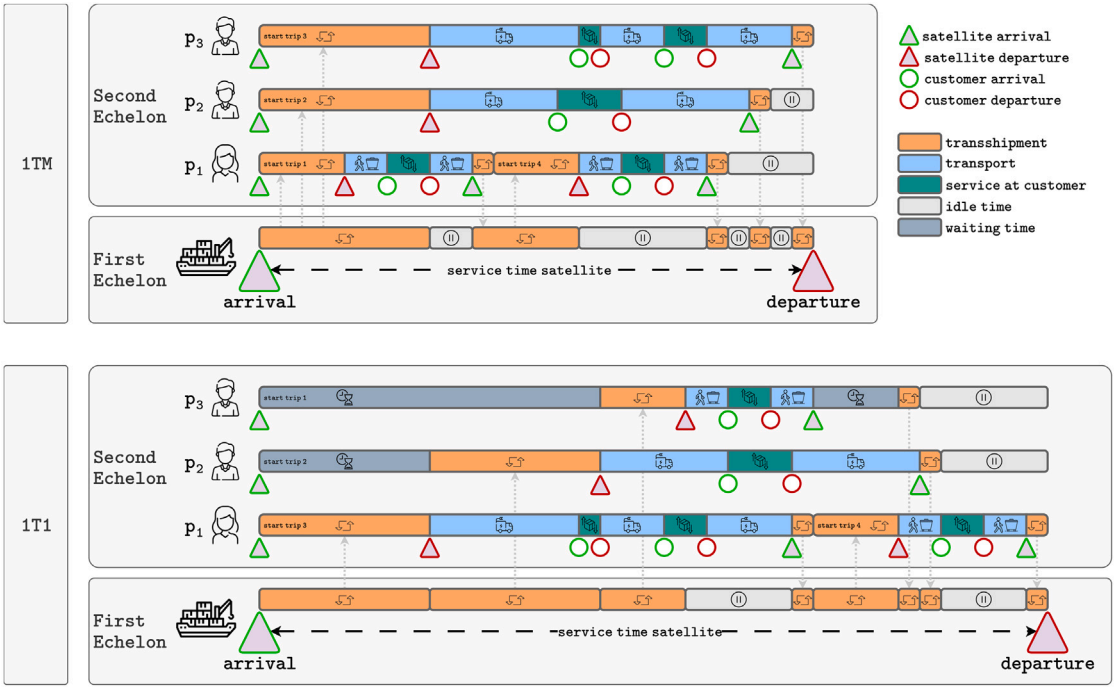


Fig. 2. Illustration of parallel one-to-many (1TM) and sequential one-to-one (1T1) transshipment modalities.

modeled as an attribute of the FEVs. The transshipment modality therefore serves a dual purpose: it represents lifting equipment limitations at FEVs and implicitly models space constraints for SEVs at satellites. In essence, the modality and workforce allocation jointly define the effective SEV space at each location, providing a realistic representation of a logistics system with constrained transshipment throughput.

3.6. Objectives and optimization decisions

The primary objective of the framework is to find a feasible operational routing schedule that attains the configured target service level (η) while minimizing resource consumption. The framework's internal heuristics are designed to support this objective. Resources such as vessels and personnel are deployed incrementally as needed, while routing algorithms prioritize the construction of low-distance trips, maximizing load consolidation, and minimizing transshipment waiting times. This promotes a heuristically minimal travel distance and fleet sizes while pursuing the target service level.

Explicit upper bounds on the FEV fleet size are not imposed. Instead, the required number of FEVs is determined endogenously, naturally constrained by satellite availability, city access time windows, and satellite occupancy constraints. Because satellites are storage-free and can only accommodate one FEV at a time, the maximum throughput of the transshipment network inherently limits FEV deployment. Imposing an artificial cap on the fleet size could create unnecessary bottlenecks and compromise the attainable service level. Crucially, multi-trip operations on both echelons enable efficient demand fulfillment while avoiding the need for an excessively large fleet.

4. Methodology: An iterative decomposition-based heuristic

This study introduces an iterative decomposition-based heuristic designed to address the complexities of a 2E-VRP characterized by the proposed variants in realistic capacity settings. To manage the computational complexity of centralized transport planning that leverages collaboration benefits in city logistics (Barratt, 2004; Cruijssen et al., 2007; Gansterer and Hartl, 2018; Cleophas et al., 2019), the integrated problem is decomposed into satellite scheduling and routing subproblems for both echelons. This facilitates flexible adaptation of optimization methods and problem features for each subproblem and supports further extensions when required. The proposed modeling is highly flexible, incorporating satellite-, customer- and SEV type-specific lead times for transshipment and service operations, as well as allowing the allocation of the workforce to be specified separately for each satellite. The proposed framework integrates a capacitated customer-to-satellite assignment optimization model with routing and scheduling heuristics. The methodology is tailored to find feasible operational schedules by addressing the inherent interdependencies between customer allocation and subsequent routing, a common challenge in multi-echelon distribution networks.

As illustrated in Fig. 3, the procedure follows an iterative process structured into three main phases: the *Allocation phase*, the *Routing phase*, and the *Bounding phase*. At the core of the framework is an adaptive feedback loop that iteratively refines algorithmic parameters known as satellite workload bounds. These workload bounds constrain the total customer demand that can be assigned to each satellite by the optimization model for the capacitated assignment problem.

The primary objective of the framework is to attain the configured target service level (η) while minimizing resource consumption. Rather than optimizing a single cost function, the framework is designed to evaluate resource trade-offs ex-post using operational metrics such as fleet size, workforce, and total travel distance. During solution construction, internal heuristics guide decision-making by incrementally deploying resources and prioritizing low-distance routes, load consolidation, and reduced transshipment waiting times, while aiming to meet the target service level.

Determining the maximum transshipment throughput of each satellite, or in other words, its maximum feasible daily workload, is challenging, as it depends strongly on the interdependent routing and scheduling of vehicles under operational constraints. A single change in customer-to-satellite assignment can affect the entire system, triggering a cascade of adjustments to trip schedules. Solving the routing problems for a given customer-to-satellite allocation reveals whether the assigned workload is feasible under operational constraints, but not whether the satellite can potentially serve a higher workload, nor how much. In

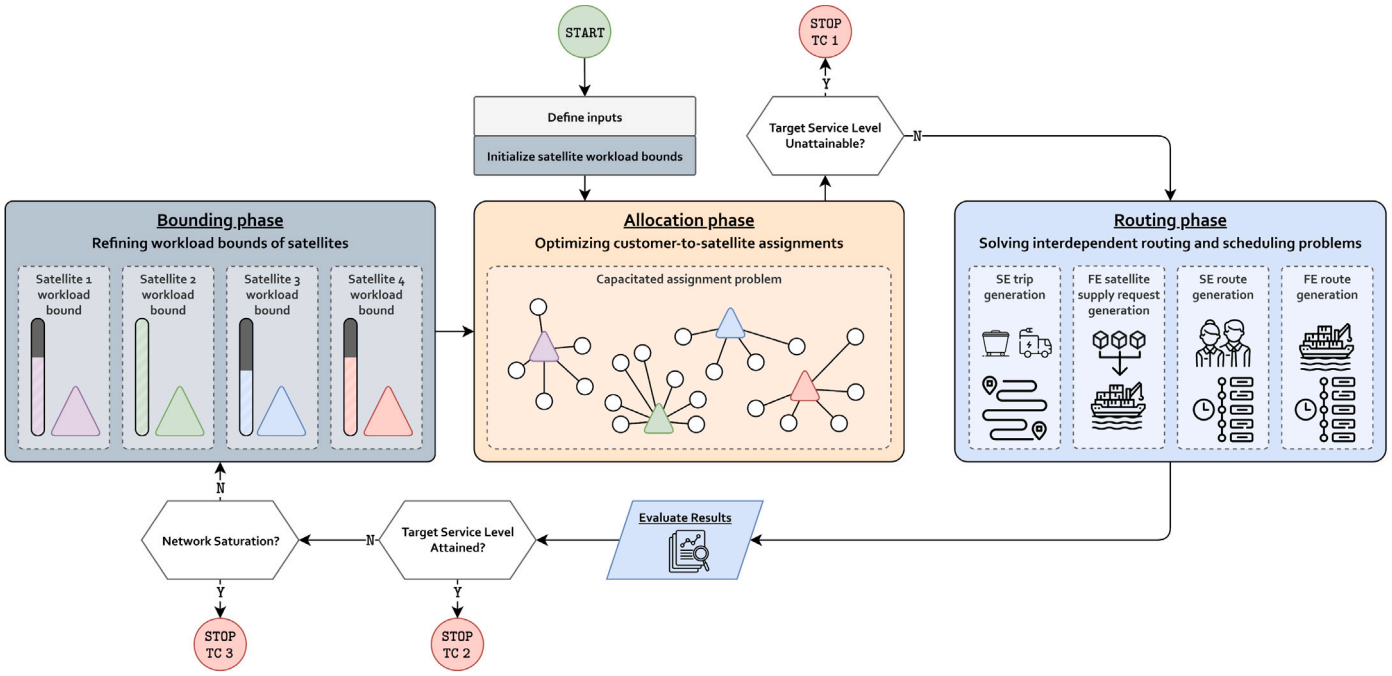


Fig. 3. Overview of the proposed framework.

other words, it remains unknown whether a satellite could handle a higher workload than assigned by the capacitated assignment model, since the existence of a feasible solution for a different allocation depends on the complex and highly constrained 2E-VRP. This can be assessed by solving the interdependent routing and scheduling problems for FEVs and SEVs and further refined to enhance the overall service level of the solution. Table 2 summarizes the notation used for the input sets, parameters, and decision variables.

4.1. Allocation phase: Optimizing customer-to-satellite assignments

The allocation phase of the proposed framework determines the assignments of customers to satellites, which serve as transshipment points within the service network. These assignments shape the spatial configuration of the network and directly influence synchronization requirements by specifying where transshipments must occur. The primary objective of this tactical-level assignment is to distribute customer demand in a manner that supports the construction of a feasible solution for both echelons of the 2E-VRP.

The capacitated assignment problem for allocating customers to satellites is formulated as an integer linear programming (ILP) model. The model assigns each customer to a single satellite while minimizing the total service distance and ensuring that the total demand assigned to any satellite does not exceed its workload bound for the current iteration. These bounds, denoted Q_s^i for satellite s in iteration i, are initialized based on best-case transshipment throughput assumptions (see Section 4.4.1) and iteratively refined based on routing feasibility in the bounding phase (see Section 4.3). This adaptive adjustment of workload bounds plays a key role in progressively guiding the iterative process toward an operationally feasible solution.

To formally define the capacitated assignment problem, let S represent the set of available satellites and C denote the set of customers. The demand for customer $c \in C$ is given by d_c. The allocation cost between customer c and satellite s is defined by $\tau_{s,c}$, corresponding to the SE distance. The decision variable $x_{c,s}^i$ is a binary variable, equal to 1 if customer c is assigned to satellite s in iteration i, and 0 otherwise.

To ensure feasibility under restrictive workload bounds, the model is augmented with a dummy node representing the DC that can absorb excess demand, yielding the augmented satellite set $S' = S \cup \{\mathrm{DC}\}$. Assigning customers to the DC incurs a large penalty cost $\tau_{\mathrm{DC},c} = \mathcal{M}$, ensuring that such assignments occur only when necessary. Customers assigned to the DC are treated as unserved demand and may be outsourced for direct deliveries by road without using the satellite network, but they do not participate in the multimodal 2E-VRP. The capacitated assignment problem is formulated as the following ILP:

$$\min \sum_{s \in S'} \sum_{c \in C} \tau_{s,c} x_{c,s}^i \tag{1}$$

subject to

$$\sum_{s \in S'} x_{c,s}^i = 1 \qquad \forall c \in C \tag{2}$$

$$\sum_{c \in C} d_c x_{c,s}^i \le Q_s^i \qquad \forall s \in S \tag{3}$$

$$x_{c,s}^i \in \{0,1\} \qquad \forall s \in S', c \in C \tag{4}$$

The objective function (1) minimizes the allocation costs between customers and satellites in the augmented satellite set S'. Constraints (2) guarantee that each customer is assigned to exactly one satellite within this augmented set. Constraints (3) ensure that the aggregate demand of all customers assigned to each satellite $s \in S$ does not exceed its current workload bound. Finally, constraints (4) impose binary values on the decision variables. This capacitated assignment optimization model determines the sets of customers, C_s, assigned to each satellite s, and its output is used in the routing phase that solves the routing and scheduling problems to evaluate the service level and the operational cost of these assignments within the multimodal system.

4.2. Routing phase: Solving interdependent routing and scheduling problems

The routing phase constructs a synchronized schedule for both echelons using a sequence of heuristics. The detailed pseudo-code for each of these steps is provided in Algorithms 1–4 in Appendix B. It is important to distinguish between the physical flow of goods, which operates top-down from the DC to the customer, and the algorithmic construction of the solution, which proceeds bottom-up to accurately capture resource constraints. The primary objective of the framework is to attain the configured target service level (η). The sub-objectives

Table 2
Notation overview: input sets, parameters, and decision variables.

Input sets	
S	Set of available satellites
$\mathcal{K}$	Set of SEV types
C	Set of customers with demand in the planning horizon
Input parameters	
$\mathcal{W}$	Permitted working hours; city access time window span
μ_s	Transshipment mode at $s \in S$: 1TM or 1T1
η	Target service level (fraction of total demand, $0 \leq \eta \leq 1$)
Q^k, Q^{FEV}	Capacity of SEV type k and of the FEV
$\mathrm{p}_s^{\max}$	Maximum workforce (number of personnel) at satellite s
d_c	Demand of customer $c \in C$
$\theta_s^{k\rightarrow}$, $\theta_s^{k\leftarrow}$	Per-unit transshipment time at satellite s by SEV type k (delivery flow $\rightarrow$, backhaul flow $\leftarrow$)
$\theta_c^{k\rightarrow}$, $\theta_c^{k\leftarrow}$	Per-unit service time at customer c by SEV type k (delivery flow $\rightarrow$, backhaul flow $\leftarrow$)
$\theta_{\mathrm{DC}}^{\mathrm{FEV}\rightarrow}$, $\theta_{\mathrm{DC}}^{\mathrm{FEV}\leftarrow}$	Per-unit loading ($\rightarrow$) and unloading ($\leftarrow$) time at the DC for the FEV
$\tau_{a,b}^k$, $\tau_{a,b}^{\mathrm{FEV}}$	Travel time from location a to b for SEV type k and the FEV, respectively
ρ_s^k	Service range of SEV type k at satellite s
Decision variables for routing	
C_s	Set of customers assigned to satellite s (output of allocation phase)
$\mathcal{P}_s$	Workforce size (number of personnel) at satellite s
q_t^{SE}, q_t^{FE}	Load carried during SE or FE trip t
δ_t^{SE}, δ_t^{FE}	Duration of SE or FE trip t (travel + service + waiting + transshipment)
$\gamma_t^{\rightarrow}$, $\gamma_t^{\leftarrow}$	Transshipment time for SE trip t (loading onto $\rightarrow$, unloading from $\leftarrow$ the SEV)
$\omega_t^{\rightarrow}$, $\omega_t^{\leftarrow}$	SEV waiting time at the satellite before transshipment for SE trip t (delivery flow $\rightarrow$, backhaul flow $\leftarrow$)
$\Omega_t^{\leftarrow}$	FEV waiting time at the satellite before the backhaul transshipment for SE trip t
$\Omega_{s,n}$	FEV waiting time before departing to serve supply request (s,n)
$\mathcal{T}_s$	Set of SE trips from satellite s
$\mathcal{T}_{s,n}$	Set of SE trips from satellite s for supply request n
σ_t^{SE}	Sequence of stops in SE trip t
$\sigma_{s,n}^p$	Sequence of SE trips from satellite s for request n assigned to person p
$\mathcal{R}$	Set of FE routes, each route as a sequence of FE trips
σ_r^{FE}	Sequence of FE trips in route $r \in \mathcal{R}$
σ_t^{FE}	Sequence of stops in FE trip t
$\mathcal{O}_s$	Scheduled FEV arrival and departure times per request served at satellite s
Decision variables for satellite supply requests	
$\mathcal{N}_s$	Set of supply requests for satellite s
$d_{s,n}$	Total demand for supply request n at satellite s
$\delta_{s,n}$	Total service time for supply request n at satellite s

within this phase, such as minimizing travel distance or maximizing load consolidation, function as complementary heuristic proxies that aim to enhance overall system throughput. Each sub-objective aims to free up critical resources like vehicle time and satellite space, for example by reducing waiting times or minimizing the duration of transshipment operations. This enables the construction of a denser operational schedule, allowing more delivery and supply trips to be completed within the limited time window, thereby increasing the overall service level.

Step 1: SE trip generation (distance minimization). The first step constructs the specific set of second-echelon trips required to serve all assigned customers. To minimize travel distance, a nearest-neighbor heuristic (NNH) sequentially builds feasible trips for the available street-level vehicles. This procedure explicitly accounts for the heterogeneous fleet by assigning customers to appropriate vehicle types based on their spatial service ranges. It iteratively groups nearby deliveries until specific vehicle capacities are reached, splitting a customer's demand into multiple visits if it exceeds the vehicle's capacity. By minimizing travel distance, this step produces efficient trips that consume less of the available time window, increasing the probability that more trips can be completed before the city access deadline.

Step 2: FE satellite supply request generation (load consolidation). The second step consolidates the generated second-echelon trips into supply requests for the vessels. Modeled as a bin-packing problem, a first-fit strategy groups the street-level trips while strictly respecting the physical volumetric capacity of the vessels. To promote efficient operations, the heuristic prioritizes packing quicker street-level trips first. By prioritizing load consolidation, this step aims to reduce the required frequency of vessel resupply visits. Since satellites can only accommodate one vessel at a time, fewer visits reduce the likelihood of vessels waiting for docking space. Resolving this spatial bottleneck ensures satellites can be resupplied faster, directly increasing the system's total throughput capacity. As an added benefit, managing fewer, fuller loads maximizes vessel fleet efficiency and reduces waterway congestion.

Step 3: SE route generation (service time minimization). The third step sequences the second-echelon trips in the consolidated supply requests at the satellite level, functioning as a machine scheduling problem. It schedules the required transshipment operations by dynamically assigning street-level trips to the available satellite personnel, aiming to balance the workforce workload. Crucially, the heuristic coordinates these assignments according to the specific transshipment modality (parallel or sequential), as illustrated in Fig. 2, to actively mitigate vehicle waiting times. By dynamically assigning street-level trips to personnel, balancing the workload among personnel, and coordinating transshipment modalities to mitigate vehicle waiting times, this step aims to minimize the total service time required for the supply request. Keeping this duration short reduces the time a vessel must occupy the satellite, freeing up the limited docking space faster and maximizing the satellite's availability for subsequent deliveries.

Step 4: FE route generation (service level maximization). The final step constructs multi-trip routes for the vessels to fulfill the generated supply requests. This includes the selection of the constructed supply requests to be served and the FEV routes to serve them within the city access time window. A greedy insertion heuristic sequentially assigns these requests to the vessel fleet while respecting vessel capacities and the city access time window. To determine which request to serve next, a hierarchical sorting rule is applied that prioritizes candidates (s,n) based on three criteria, ensuring efficient scheduling and minimal FEV waiting time. Specifically, unserved requests are hierarchically sorted according to:

1. earliest possible arrival time ($\texttt{EPA}_{s,n}$): requests are first sorted in ascending order of the earliest possible arrival time at satellite s, accounting for travel time and potential waiting due to satellite occupancy.
2. total demand ($d_{s,n}$): ties are broken by favoring requests with higher demand, sorted in descending order.
3. service duration ($\delta_{s,n}$): remaining ties are resolved by favoring shorter service duration, in ascending order.

By densely scheduling the available working hours, this hierarchical assignment of supply requests aims to maximize service level and FEV utilization within the available time to access the satellites. As a result, this process finalizes the operational schedule for the corresponding second-echelon routes. This process finalizes the synchronized operational schedule and ultimately determines the key performance indicators for the current allocation: the overall service level, the required fleet sizes for both echelons, and the total distance traveled. The FE routing Algorithm 4 can also be configured to terminate early if the cumulative served demand meets the target service level η, providing lower estimates of the minimum resources required for scenarios not targeting full coverage.

4.3. Bounding phase: Refining workload bounds of satellites

The bounding phase serves as a corrective phase within the iterative framework, aiming to ensure feasible workload assignments across satellites. Upon completing the routing phase, the resulting solution reveals whether any satellites have been allocated workloads that exceed what they can feasibly serve. Formally, a satellite s is considered overloaded in iteration i if the total demand of customers assigned to it, calculated as $\sum_{c\in\mathcal{C}_s} d_c$, exceeds the actual demand served, calculated as $\sum_{n\in\mathcal{O}_s} d_{s,n}$ for all served supply requests n of satellite $s \in \mathcal{S}$. The set of overloaded satellites in iteration i is denoted as $\mathcal{S}^{\texttt{overload},i}$. The bounding phase uses this information to reduce the workload bounds of overloaded satellites, thereby guiding the allocation phase to reassign customers from overloaded satellites to those that can potentially serve a higher workload. This adaptive feedback loop is designed to converge on a feasible 2E-VRP solution for the given customer-satellite allocation that maximizes overall service level without overly balancing satellite workloads, as that increases SE routing distance.

The reduction of workload bounds for overloaded satellites depends on the extent to which the current workload bound exceeds the actual demand served. Satellites whose workload bounds substantially exceed their served demand undergo a more aggressive reduction, whereas those with smaller differences receive a more conservative adjustment. This tiered reduction strategy prevents unnecessary customer reassignment and reduces redundant iterations where updated bounds do not yet affect the allocation phase, a situation often caused by initially high workload bounds. The reduction factor α_s^i, which controls the workload bound adjustment, is formally defined as follows, where $\alpha_2, \alpha_1, \alpha_0, \Delta_1$, and Δ_0 are design parameters further detailed in the case study:

$$\alpha_s^i = \begin{cases} \alpha_2, & \text{if } Q_s^{i-1} - \sum_{n\in\mathcal{O}_s} d_{s,n} \geq \Delta_1 \\ \alpha_1, & \text{if } \Delta_0 < Q_s^{i-1} - \sum_{n\in\mathcal{O}_s} d_{s,n} < \Delta_1 \\ \alpha_0, & \text{otherwise} \end{cases} \quad \forall s \in \mathcal{S}^{\texttt{overload},i-1} \tag{5}$$

Accordingly, the workload bounds are updated for each overloaded satellite $s \in \mathcal{S}^{\texttt{overload},i-1}$ by subtracting the corresponding reduction factor, such that $Q_s^i \leftarrow Q_s^{i-1} - \alpha_s^i$.

4.4. Architecture of the iterative decomposition-based heuristic

As visualized in Fig. 4, the framework employs an iterative loop to progressively refine the solution. The following subsections detail the initialization of workload bounds, the improvement procedure, and the termination conditions.

4.4.1. Initializing satellite workload bounds

To use the proposed iterative process, an initial upper bound on the workload of each satellite must be specified, as the capacitated assignment optimization model requires these bounds to determine customer-to-satellite allocations. To avoid overly restrictive initial workload bounds that might reduce flexibility and lead to inefficient allocations, the bounds are initialized at a high level, reflecting a best-case transshipment throughput scenario. This assumes continuous transshipment operations at the satellites throughout the city access time window ($\mathcal{W}$), a simplification that ignores routing and synchronization constraints.

Depending on the transshipment modality, the initial workload bound reflects the best-case upper bound on the volume of goods a satellite can handle, as formalized in Eq. (6). For satellites operating under a 1T1 modality, the bound is obtained by dividing the available transshipment time (defined as the city access time window minus the FEV travel time from the DC to the satellite) by the minimum per-unit transshipment time required for both delivery and backhaul flows across SEV types, denoted as $\theta_s^{\min} = \min_{k\in\mathcal{K}} \left(\theta_s^{k\rightarrow} + \theta_s^{k\leftarrow}\right)$. For satellites operating under a 1TM modality, this bound is further scaled by the maximum available workforce ($\texttt{P}_s^{\max} = \texttt{P}_s$ in case of model variant FW) to account for parallel transshipments.

$$Q_s^1 = \begin{cases} \dfrac{\mathcal{W} - \tau_{\texttt{DC},s}^{\texttt{FE}}}{\theta_s^{\min}}, & \text{if } \mu_s = \texttt{1T1} \\ \dfrac{\mathcal{W} - \tau_{\texttt{DC},s}^{\texttt{FE}}}{\theta_s^{\min}} \cdot \texttt{P}_s^{\max}, & \text{if } \mu_s = \texttt{1TM} \end{cases} \quad \forall s \in \mathcal{S} \tag{6}$$

4.4.2. The improvement procedure

The improvement procedure consists of an iterative loop in which three main phases interact in each iteration i to progressively refine the solution.

1. **Allocation phase:** The capacitated assignment model allocates customers to satellites and yields a provisional spatial configuration of the transshipment network.
2. **Routing phase:** The routing phase decides the service level for the given satellite allocations and consists of routing and scheduling heuristics that ensure a synchronized solution to the proposed 2E-VRP variants under temporal, spatial, labor, and transshipment constraints.
3. **Bounding phase:** If any satellite is overloaded, meaning that allocated demand exceeds actual served demand, the corresponding workload bound is reduced using a tiered adjustment strategy to change the solution.

This iterative loop continues until the procedure reaches a termination condition. The iterative structure enables the system to reallocate demand across satellites in response to observed unserved demand, progressively converging toward a feasible routing schedule that improves the overall service level under the given operational constraints.

In each iteration, the assignment vector, $\mathbf{x}^i := \{x_{c,s}^i\}_{c\in\mathcal{C},s\in\mathcal{S}}$, is modified in the allocation phase. This vector is compared to the previous assignment, $\mathbf{x}^{i-1}$. When $\mathbf{x}^i = \mathbf{x}^{i-1}$, the unchanged allocation shows that the current workload bounds are not yet restrictive enough to enforce a change in the solution. To avoid computational redundancy, the routing phase is skipped and the workload bounds are tightened immediately, after which the assignment problem is re-solved until $\mathbf{x}^i \neq \mathbf{x}^{i-1}$.

The framework incorporates specific logic for model variant AW, in which the workforce per satellite is adaptive and treated as a decision variable within the iterative process. In this variant, the workforce $\mathcal{P}_s^i$ at each satellite s is adjusted dynamically based on its workload and is bounded by the predefined parameter $\texttt{P}_s^{\max}$. Rather than immediately reducing workload bounds in the presence of infeasibility, the framework prioritizes incrementally increasing the satellite workforce as a means of resolving overloads. Following the construction of SE trips, the minimum number of personnel required to serve all SE trips

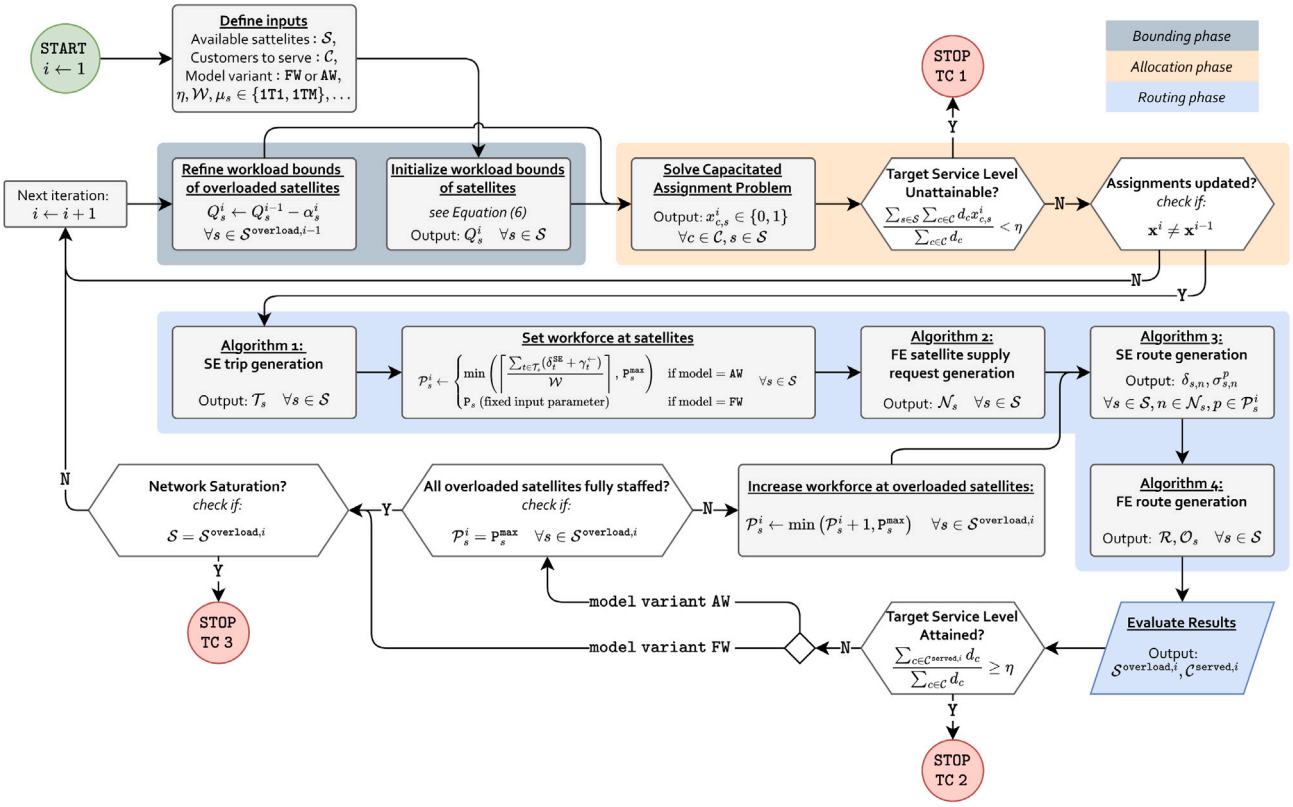


Fig. 4. Diagram of the iterative decomposition-based heuristic, detailing the adaptive feedback loop to refine workload bounds and reallocate customer demand.

from satellite s without delay (i.e., assuming no waiting due to 1T1 transshipments) is estimated as:

$$\mathcal{P}_s^i = \min\left(\left\lceil \frac{\sum_{t\in\mathcal{T}_s}(\delta_t^{\text{SE}} + \gamma_t^{\leftarrow})}{\mathcal{W}} \right\rceil, \mathrm{P}_s^{\max}\right) \quad \forall s \in \mathcal{S} \tag{7}$$

where $\mathcal{T}_s$ denotes the set of SE trips constructed from satellite s, δ_t^{SE} the trip duration excluding backhaul transshipment, and $\gamma_t^{\leftarrow}$ the time to unload collected goods upon SE trip completion. The backhaul transshipment time $\gamma_t^{\leftarrow}$ is excluded from δ_t^{SE} in Algorithm 1 to ensure correct synchronization during SE trip scheduling in Algorithm 3. After the routing phase, if allocated demand remains unserved, the framework checks whether overloaded satellites have reached their workforce limits. For each satellite that still has unserved demand and has not yet reached its personnel limit ($\mathcal{P}_s^i < \mathrm{P}_s^{\max}$), one additional worker is allocated, after which the routing phase is re-executed under the updated workforce level. This iterative prioritization of workforce expansion over immediate customer reassignment continues until either the personnel limit is reached or the demand can be feasibly served given the available workforce.

Upon termination, the framework reports the first solution found that meets the target service level, or the best-found solution if the target was not attained. This iterative methodology ensures a systematic exploration and refinement of the solution space, balancing assignment, routing, and workforce considerations while providing the flexibility needed in both parameter adaptation and service network design to address the complexities of capacitated two-echelon urban logistics.

4.4.3. Termination conditions

The iterative procedure requires stopping criteria to ensure termination. The framework therefore includes three termination conditions. These conditions are designed to identify when a fully feasible solution has been found, or when the procedure determines that no further improvement to the service level is possible under its progressive reallocation strategy. The specific conditions are defined as follows:

- **TC 1: Target Service Level Unattainable.** This condition is met when the capacitated assignment model allocates a share of customer demand to the dummy (unserved) node due to restrictive satellite workload bounds, making the configured target service level (η) unattainable. This check occurs after the allocation phase, allowing the procedure to terminate early if the maximum possible service level is already below the target η, thus avoiding unnecessary computation in the routing phase.
- **TC 2: Target Service Level Attained.** A feasible routing solution is found that successfully meets the configured target service level. Since the iterative mechanism typically reallocates customers to satellites further than their nearest one to resolve capacity constraints, continuing the process after finding a feasible solution would likely increase SE travel distances. Thus, the first feasible solution found is generally the most efficient in terms of distance.
- **TC 3: Network Saturation.** The iterative procedure converges to a state where every satellite is operationally overloaded, meaning the routing phase cannot find a feasible schedule for all assigned demand, and no further resources (like workforce or FEVs) can be added to increase the served demand. This indicates that no further customer reallocations can improve the service level, and the procedure terminates with the best-found service level.

The framework's termination conditions are adaptable to different analytical goals, configured by the target service level (η). For this study's main analysis, $\eta = 1.0$, making the objective to maximize service up to full coverage. Accordingly, the procedure terminates via TC 1 if this target becomes unattainable, via TC 2 once it is attained, or via TC 3 if the network saturates. However, the framework's flexibility

Table 3
Demand sets and problem scale analyses.

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Set 10
Customers	744	689	726	729	753	748	711	742	736	733
Total demand	1482	1401	1443	1466	1482	1500	1427	1433	1451	1458
SEV trips (at least)	297	281	289	294	297	300	286	287	291	292
FEV trips (at least)	60	57	58	59	60	60	58	58	59	59

enables the evaluation of strategic and operational trade-offs when targeting a lower service level (e.g., $\eta = 0.9$). Furthermore, the early termination check of `TC 1` can be disabled. This forces the framework to always proceed to the routing phase, ensuring that if the target is infeasible, the procedure converges to the best-found service level via `TC 3`, thus determining the maximum attainable service level and associated resource requirements under heuristic strategies for given system configuration constraints.

5. Case study

This section presents the generation of data sets and input parameters for a case study in Amsterdam, examining a multimodal logistics system that integrates road and waterborne transport via canals to serve HoReCa businesses in the city center. The case study data sets and input parameters were developed in collaboration with the municipality of Amsterdam. All relevant data sets are publicly accessible in the Mendeley Data repository (Bijvoet, 2025), comprising geospatial data on HoReCa and satellite locations, FE and SE networks, shortest path matrices, demand sets, and satellite service network designs. Specifically, operational parameters such as vehicle speeds, service times, and demand profiles were derived directly from expert knowledge and operational experiences gathered during the city's recent urban logistics pilots, thereby reflecting the specific characteristics of the Amsterdam canal and road network.

5.1. Demand profile for Amsterdam's HoReCa sector

We consider 1620 HoReCa businesses, $C_{\texttt{all}}$, in Amsterdam's city center. A single unit of delivery volume is defined as one cubic meter of goods, equivalent to the capacity of a standardized rolling container. In collaboration with the municipality of Amsterdam, each business is estimated to require about 2 to 3 deliveries per week, averaging 2 cubic meters per delivery. Deliveries are distributed uniformly across six operational days (Monday to Saturday), resulting in a daily demand probability of 45% per business, equivalent to 2.7 deliveries per week. On days with demand, businesses request 1, 2, or 3 cubic meters with equal probability, yielding a discrete demand distribution of [0, 1, 2, 3] cubic meters with associated probabilities [0.55, 0.15, 0.15, 0.15].

Based on this distribution, we generate 10 distinct demand sets that reflect the periodic demand pattern, each representing a single operational day with demand from a subset of businesses, $C \subseteq C_{\texttt{all}}$. Table 3 summarizes the scale and complexity of the problem instances tested in this study.

5.2. Multimodal transport network

The canal and road networks are interconnected through transshipment satellites, facilitating the transfer of goods between vessels and road vehicles, enabling multimodal operations. Each satellite is represented within both the FE canal network and the SE road network. For each candidate satellite location, two corresponding nodes are introduced: one positioned at the nearest point along a canal segment and the other at the nearest point on a road segment. These nodes partition the original canal and road segments into two sub-segments each. Similarly, HoReCa customer locations are integrated into the road network by inserting a node at the closest point along a road segment, thereby splitting that segment into two sub-segments. Fig. 5 depicts the resulting multimodal network, comprising the DC, all candidate satellite locations, and customer locations. The shortest path distances between all node pairs within both networks are precomputed using Dijkstra's algorithm (Dijkstra, 1959).

5.3. Vehicle, transshipment, and service characteristics

In the SE, two delivery modes are distinguished: walking and driving. Walking deliveries rely on a rolling container with a capacity of 1 cubic meter, a maximum service range of 100 m, and a travel speed of 1 m/s. Driving deliveries are carried out using light electric freight vehicles (LEFVs), each with a capacity of 5 cubic meters, an unrestricted service range covering all customers beyond the walking zone, and a travel speed of 5 m/s.

At the satellites, transshipments require 2 min per cubic meter for deliveries conducted by walking and 3 min per cubic meter for deliveries conducted by driving, reflecting the additional complexity of handling larger vehicles and loads. Similarly, at customer locations, the service time is 1 min per cubic meter for walking deliveries and 1.5 min per cubic meter for driving deliveries.

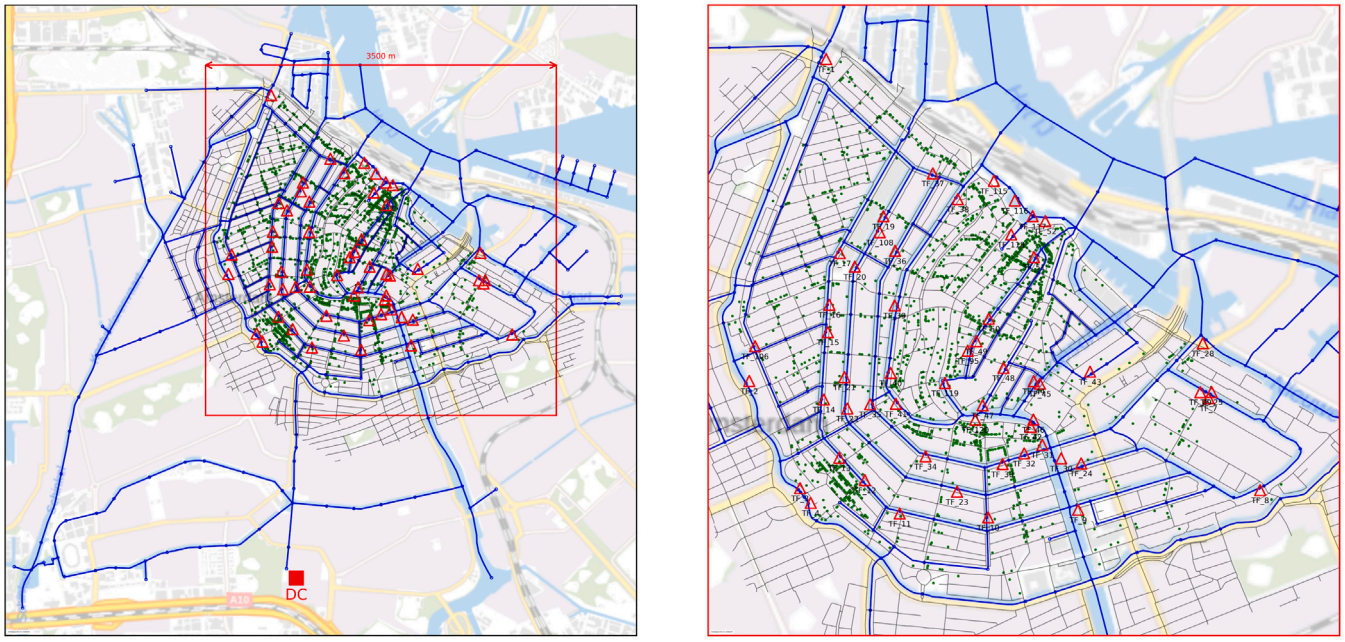
In the FE, a homogeneous fleet of FEVs (vessels) navigates the canal network. Each vessel has a capacity of 25 cubic meters and travels at a speed of 1.6 m/s. The loading time of a vessel at the DC is 1 min per cubic meter, resulting in 25 min to be fully loaded. This time also allows for potential recharging in the case of electric vessels.

We assume uniform per-unit transshipment and service times at satellites and customer locations, respectively. However, the methodology fully supports specifying these times on a per-satellite or per-customer basis to accommodate variations in handling operations. For example, differences in parking proximity or customer-specific unloading procedures can be explicitly modeled when relevant. Additionally, we consider the transshipment time for collected items at the end of SE and FE trips to be negligible, as these items are foldable, lightweight, and easily managed. Nonetheless, the methodology fully supports modeling nonzero transshipment times for backhaul flows when relevant.

5.4. Design of satellite service networks

Locating satellites is a strategic, long-term infrastructure decision that fundamentally shapes daily logistics operations. This study evaluates multiple satellite SNDs across various customer demand datasets. A total of 56 candidate satellite locations, denoted as $S_{\texttt{all}}$, are considered and depicted in Fig. 5. The problem of locating satellites for different SNDs is formulated as a p-median facility location problem (FLP), where p satellites are selected from the candidate set $S_{\texttt{all}}$ to minimize customer-satellite allocation costs. The mathematical formulation, adapted from Marín and Pelegrín (2019), uses a cost matrix that captures the shortest path distance on the Amsterdam road network between each customer $\forall c \in C_{\texttt{all}}$ and satellite $\forall s \in S_{\texttt{all}}$, with the objective of minimizing the total travel distance between the selected p satellites and their assigned customers.

Note that this parameter p denotes the number of satellites to locate and should not be confused with the number of personnel at a satellite. Furthermore, the term 'assigned customers' here refers exclusively to the network design problem and should not be confused with the customer-satellite assignments determined in the allocation



(a) Location of the DC (red square), canal network (blue lines), and road network (black lines).

(b) Candidate satellite locations (triangles) and HoReCa customer locations (green dots) in Amsterdam's city center

Fig. 5. Overview of the canal and road networks, candidate satellite locations, and HoReCa locations.

phase of the proposed methodology. This formulation focuses solely on establishing the satellite network structure.

The p-median FLP model is applied to generate several satellite SNDs by varying the number of available satellites. Specifically, p is varied from 1 to 15 in increments of one, and from 15 to 50 in increments of five, resulting in 22 distinct satellite SNDs: $p = [1, 2, \ldots, 14, 15, 20, 25, 30, 35, 40, 45, 50]$. Two examples of resulting SNDs are shown in Fig. 6. Satellites are labeled with the prefix 'TF' (transshipment facility), and their original identification numbers are preserved to remain consistent with municipal data sources.

We decouple the strategic location decision from the daily routing to bypass the computational intractability of solving a fully integrated Location-Routing Problem (LRP) under the highly constrained operational features of this study, such as multi-trip operations, strict vehicle occupancy limits, and load-dependent transshipment times requiring precise synchronization at storage-free satellites. Instead, we adopt a "generate-and-evaluate" approach: the p-median model generates strategic candidate configurations, and our operational model evaluates each configuration to determine its feasibility and resource requirements. It is important to note that the availability of a satellite location in the network design does not imply its actual use in daily operations for a given demand set, as this depends on the allocation and routing decisions. By testing a wide range of satellite densities, we capture the operational-strategic interactions ex-post, allowing us to systematically evaluate resource trade-offs.

5.5. Experimental layout: Design configurations and evaluation scope

To evaluate the trade-offs between infrastructure availability and operational feasibility for full service coverage, a 100% customer service level, we assess various system design configurations that differ along four dimensions.

Satellite SND: infrastructure investment and urban space allocation. We test 22 distinct satellite SNDs, as introduced in Section 5.4, generated by varying the number of located satellites p. These configurations determine the spatial distribution of transshipment capacity, which affects workload distribution, vehicle routing patterns, and overall travel and waiting times. Increasing the number of satellites generally improves SE efficiency by reducing distances and waiting times.

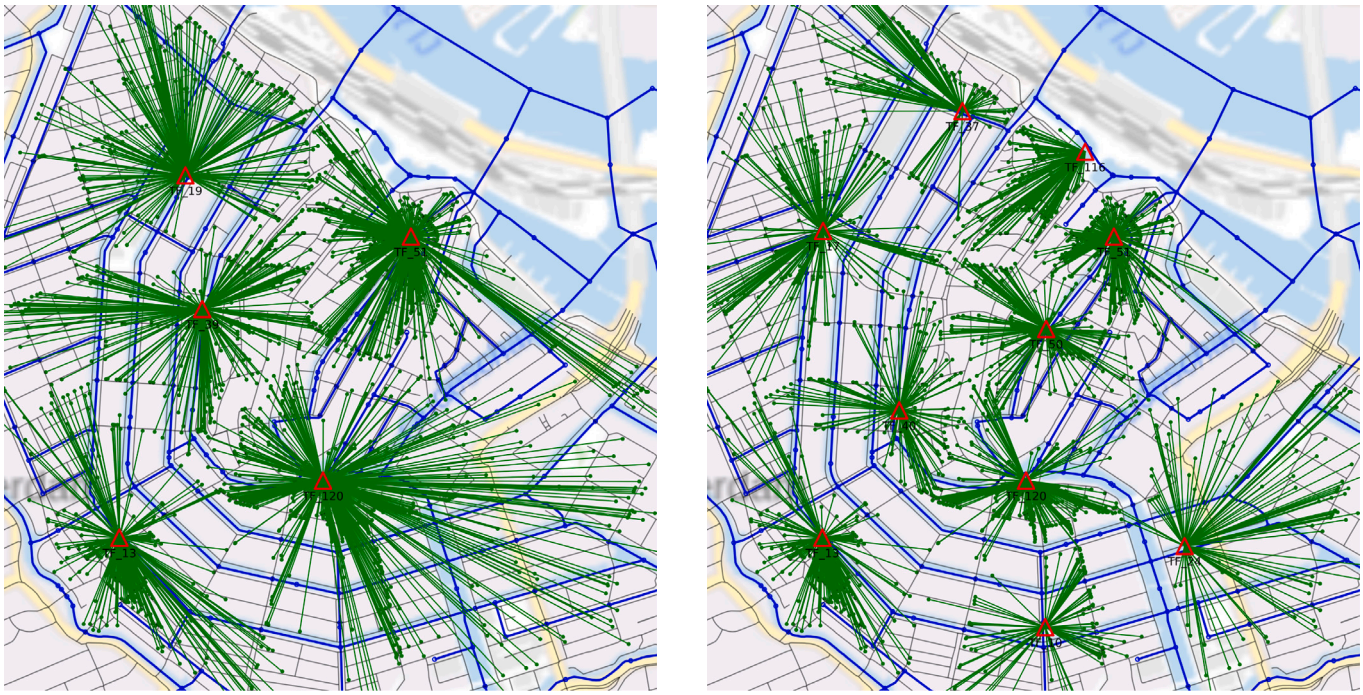
City access time window: regulatory policy for urban livability. Four city access time window spans $\mathcal{W} \in \{4, 8, 12, 16\}$ hours are evaluated, reflecting regulatory restrictions on urban freight movements, directly affecting vehicle scheduling flexibility.

Transshipment modality: technological and spatial constraints at public docking areas. We distinguish between one-to-many (`1TM`) and one-to-one (`1T1`) transshipment modalities. Under `1TM`, an FEV may simultaneously serve multiple SEVs, whereas under `1T1` each FEV can only transship to a single SEV at a time. While the methodology supports different modalities per satellite, we restrict our evaluation to uniform cases where all satellites operate either under `1TM` or `1T1`.

Workforce availability at satellites: labor investments and operational flexibility. We consider two workforce management variants: the fixed workforce model (`FW`) and the adaptive workforce model (`AW`). Although the methodology allows specifying distinct workforce limits per satellite, we assume uniform availability across all satellites. The variant $\texttt{FW}_z$ denotes the fixed model, where the input parameter $z = \texttt{P}_s$ personnel is assigned to each satellite s, whereas $\texttt{AW}_z$ denotes the adaptive model, with an input limit of $z = \texttt{P}_s^{\max}$ personnel per satellite, allowing actual workforce allocations $\mathcal{P}_s$ to vary within this bound. Because a higher workforce per satellite under the `1T1` transshipment modality tends to increase waiting times, the maximum workforce is limited to four personnel per satellite for the `AW` model variant. In contrast, for `1TM` operations, the `AW` variant is evaluated with a maximum of five personnel per satellite, corresponding to the number of transshipments needed to fully unload an FEV into five LEFVs. The following workforce configurations are evaluated:

- For the `1TM` modality: $\texttt{FW}_1$, $\texttt{FW}_2$, $\texttt{FW}_3$, and $\texttt{AW}_3$, $\texttt{AW}_4$, $\texttt{AW}_5$.
- For the `1T1` modality: $\texttt{FW}_1$, $\texttt{FW}_2$, $\texttt{FW}_3$, and $\texttt{AW}_2$, $\texttt{AW}_3$, $\texttt{AW}_4$.

This configuration space enables a comprehensive assessment of the trade-offs involved in satellite network design, urban access management, and the coordination of workforce and transshipment resources within a multimodal freight system. Specifically, for the main



(a) Satellite SND 5 (b) Satellite SND 10

Fig. 6. Satellite SND examples.

case study, we evaluate a total of 10560 distinct problem instances, generated by combining 22 satellite SNDs, 4 time window spans, 2 transshipment modalities, 6 workforce variants, and 10 demand sets.

By systematically varying these four parameters, our evaluation framework quantifies the strategic trade-offs involved in designing and operating a multimodal city logistics network. For example, this allows us to determine whether restrictive urban access policies (shorter time windows) can be offset by higher labor investments (workforce), or if spatial constraints that limit operations to sequential transshipments necessitate a denser network of facilities (more satellites).

6. Computational results

This section presents the computational results of the Amsterdam case study, with the objective of identifying the minimum system requirements needed to guarantee full service coverage and derive insights on strategic and operational trade-offs. The analysis focuses on system design configurations that attain a 100% average service level (ASL) across all demand sets, corresponding to $\eta = 1.0$ for all instances. This allows us to determine the required number of satellites, fleet size, personnel, travel distances, and operational time across various network design and transshipment scenarios. Detailed results for all scenarios are provided in the Appendix C–F.

All experiments are executed on a server equipped with Intel Xeon Gold 5218 processors, allocating 1 CPU and 1 GB RAM per instance. The mathematical models for the capacitated assignment problem (allocation phase) and the *p*-median FLP used to generate satellite SNDs are solved using Gurobi 10.0.3 (Gurobi Optimization, 2022).

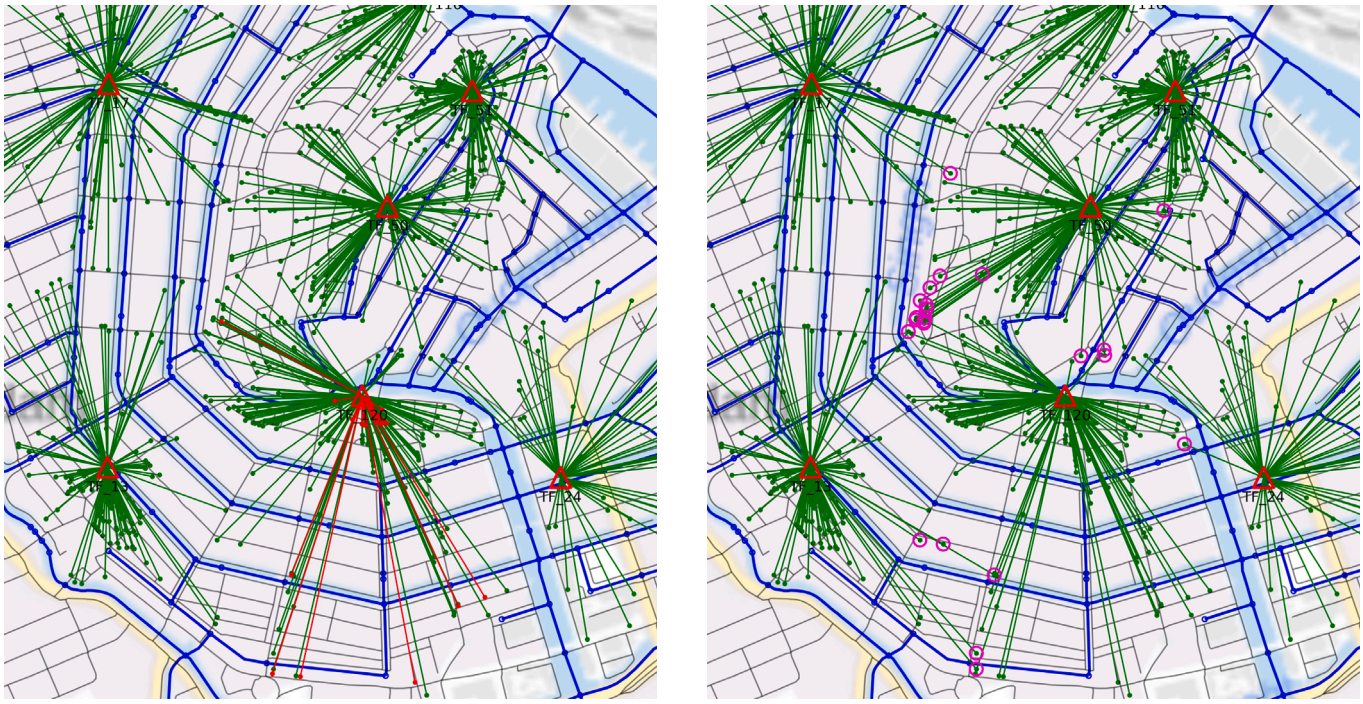
To update the satellite workload bounds between iterations (see Section 4.3), we use the following parameters in Eq. (5): $\alpha_0 = 1$, $\alpha_1 = 10$, $\alpha_2 = 50$ for step sizes, and $\Delta_0 = 50$, $\Delta_1 = 100$ for unserved demand thresholds. These values are chosen based on the demand profile and vehicle specifications of the case study. More specifically, depending on the level of excess workload, a satellite's workload bound may be reduced by one delivery unit, two full LEFV trips, or two full FE supply requests. This tiered strategy is designed to balance stability with convergence speed. For instance, if a satellite is overloaded by only 5 units, an aggressive reduction would force unnecessary customer reassignments to more distant satellites. Instead, our method applies a minimal correction ($\alpha_0 = 1$) to fine-tune the allocation. Conversely, if a satellite is overloaded by 150 units, a minimal correction would be ineffective. Therefore, a substantial reduction ($\alpha_2 = 50$) is applied to quickly guide the allocation toward a feasible state. This prevents both disruptive reallocations for minor issues and wasted computational effort on solutions that are known to be highly infeasible.

For the case study, the initialization of workload bounds defined in Eq. (6) is based on the maximum per-unit transshipment time rather than the minimum. This choice is appropriate in the Amsterdam context, where the vast majority of customers must be served by the SEV type with the highest handling time, the LEFV. Although such a choice could create overly restrictive initial bounds in other contexts, here it provided realistic initial bounds without limiting flexibility in the iterative refinement process.

6.1. Illustrative example of the operational dynamics

The proposed framework iteratively allocates customer demand across the satellite network and evaluates its feasibility by constructing synchronized FE and SE routes. Consequently, satellite and vehicle utilization are not predefined, but rather emerge dynamically from these interdependent scheduling decisions. To illustrate these complex operational dynamics, we visualize resource usage for a representative scenario, specifically focusing on customer allocation, satellite occupancy, and fleet deployment. This illustration is based on a service network design with seven available satellites, all operating under the `1TM` transshipment modality, a 12-hour time window, demand set 1, and workforce variants $\texttt{FW}_3$ and $\texttt{AW}_3$.

The allocation results are illustrated in Fig. 7, showing the customer-satellite assignments for the initial and final solutions of the procedure. The initial solution corresponds to the outcome of the first iteration of the method, as illustrated in Fig. 7(a), where red dots and lines indicate unserved customers. Notably, satellite TF_120 is unable to



(a) Initial solution (96.6% service level)

(b) Final solution (100% service level)

Fig. 7. Customer-satellite assignments for configuration: $\mathtt{FW}_3$, 1TM, $\mathcal{W}$ = 12 hours, SND 7, demand set 1.

serve all of its assigned customers, resulting in an overall service level of 96.6%. As the procedure progresses, the workload of this overloaded satellite is reduced in order to reallocate customers to other satellites, ultimately resulting in a final solution in which all demand is successfully served, as illustrated by Fig. 7(b). Customers whose assignments changed compared to the initial solution are encircled in magenta. These allocation changes primarily affect customers located in demand clusters near the boundaries between satellite service zones. Importantly, the reallocation process is not limited to initially unserved customers but also triggers broader shifts across multiple satellites, enabling a more balanced workload distribution and improving the overall service level across the system.

The final FE routes are depicted in Fig. 8, with the total demand served by each route indicated above it. The colored bars represent the service time of served satellite supply requests, indicating the duration each vessel occupies a satellite while personnel conduct SE trips to fulfill deliveries and pickups at customers. Each bar's label specifies the satellite supply request being served, with identical colors denoting requests associated with the same physical satellite. The red dotted vertical line marks the end of the operational time window, after which FEVs may still return to the DC. Several vessels remain idle at the DC before the time window closes, as they cannot complete additional supply requests within the allowed working hours.

When a vessel must wait to access a satellite due to space constraints, it remains at the DC and times its departure to arrive exactly when the satellite becomes available. For example, FEV 8 delays its departure so that it arrives at satellite TF_13 right after FEV 1 completes its service. After the initial 7 FE routes are constructed, additional FEVs experience waiting times to synchronize access to the satellites due to the limited number of only 7 available satellites in the SND. However, 11 more vessels are deployed in the system to ensure a continuous supply flow and enhance satellite utilization. While some FEVs return to the DC for replenishment, others occupy the satellites and perform transshipments to the SE, thereby increasing the total served demand. For example, satellite TF_13 is served by five different FE routes (FEV 1, 5, 6, 8, and 9), collectively fulfilling a total of 10 supply visits.

The satellite occupancy schedule is depicted in Fig. 9(a), showing the arrival and departure times of vessels for each served supply request. The total served demand per satellite is indicated above the diagram. In this scenario, most satellites are occupied by vessels for the majority of the permitted working hours. However, satellite TF_24 is utilized only for about half of this time, suggesting that its workload could potentially be managed with a smaller workforce, reducing from three to two personnel. Fewer personnel would increase the service time required for each supply request at satellite TF_24, and whether a feasible synchronized FE routing solution exists under this reduced workforce remains unknown.

The SE routes are illustrated in Fig. 9(b), showing that from each satellite, three personnel perform street-level deliveries to customers. Under 1TM operations, a vessel can simultaneously transfer goods to multiple personnel, thereby eliminating waiting times for transshipments. Blank spaces within an SE route indicate idle time spent waiting for vessel arrivals rather than delays caused by transshipments to other personnel. The figure shows that transshipments account for the majority of time in street-level operations and that most SE trips are conducted using an LEFV, as an SND with seven satellites limits the number of customers within a 100-meter walking range.

Fig. 10 illustrates dynamic workforce allocation for the same scenario evaluated under the $\mathtt{AW}_3$ variant (maximum three personnel per satellite). Under the $\mathtt{FW}_3$ variant, satellite TF_24 is utilized for only about half of the available time. The AW model identifies this surplus and reduces the local workforce to two personnel. Although this increases service times at TF_24, the $\mathtt{AW}_3$ variant still achieves a 100% service level with fewer overall personnel than $\mathtt{FW}_3$. To compensate for these longer service times, 19 FEVs are deployed under $\mathtt{AW}_3$ instead of 18 under $\mathtt{FW}_3$, highlighting the operational trade-off between workforce and the required vessel fleet size.

These results demonstrate the value of optimizing daily operations through dynamic workforce allocation within predefined limits to prevent allocating excess workforce, a challenge for manual planning given the unpredictability of interdependent resource requirements. The AW model better accounts for daily variations in geographical

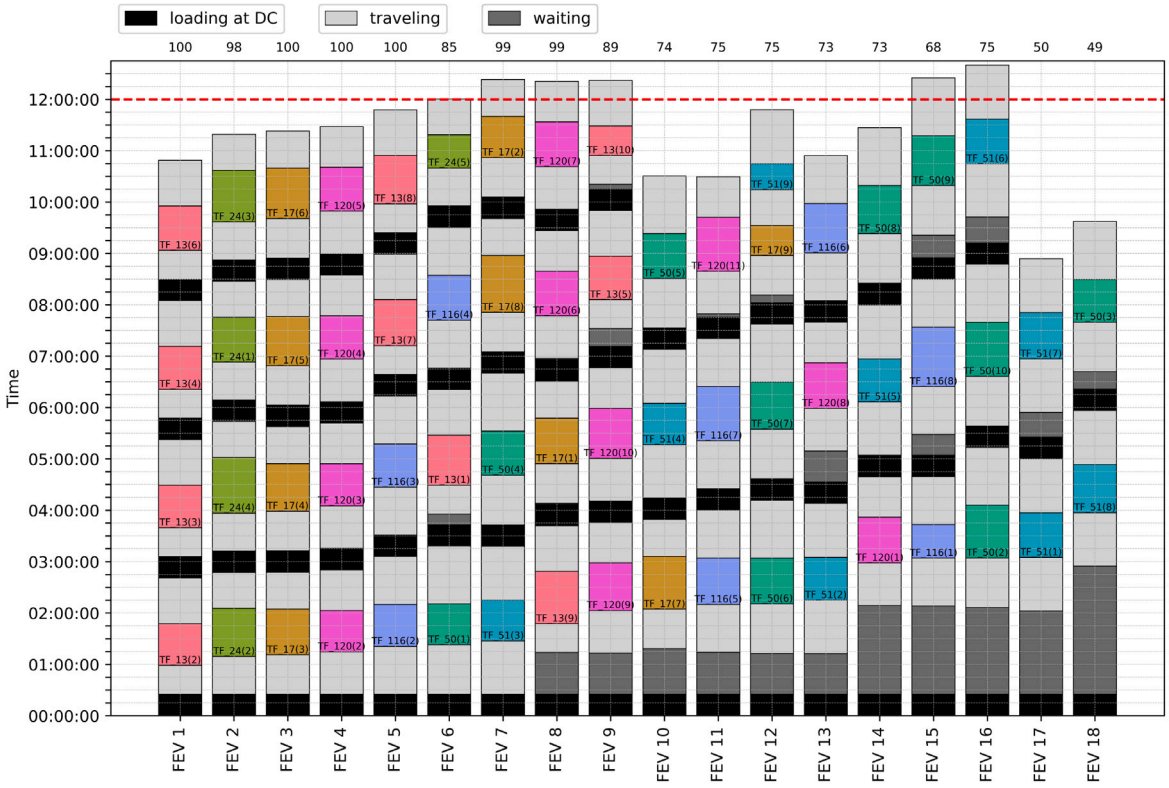
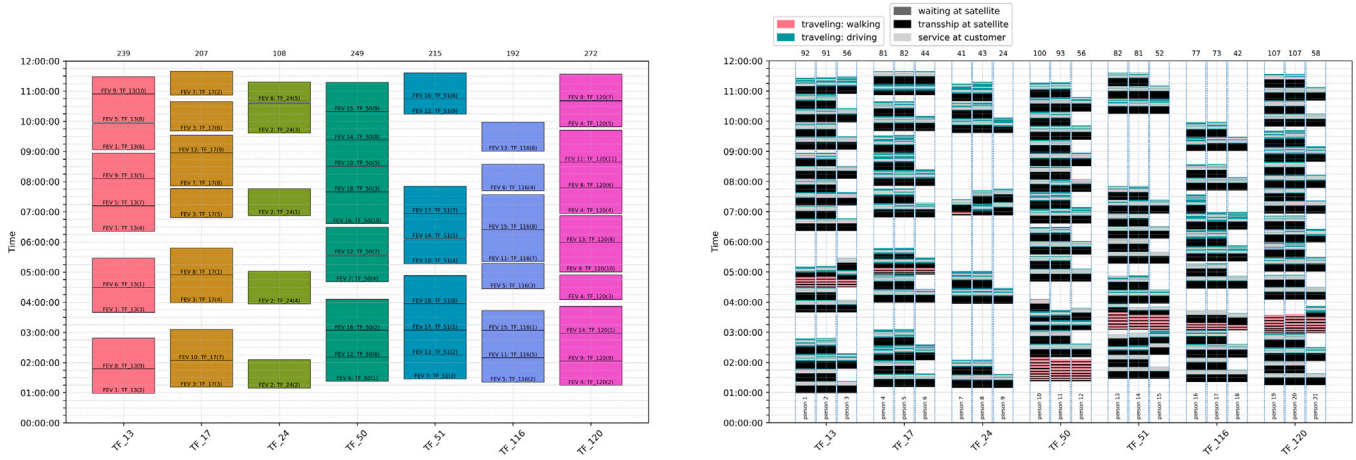


Fig. 8. FE routes of final solution (100% service level) for configuration: $\texttt{FW}_3$, 1TM, $\mathcal{W}$ = 12 hours, SND 7, demand set 1.



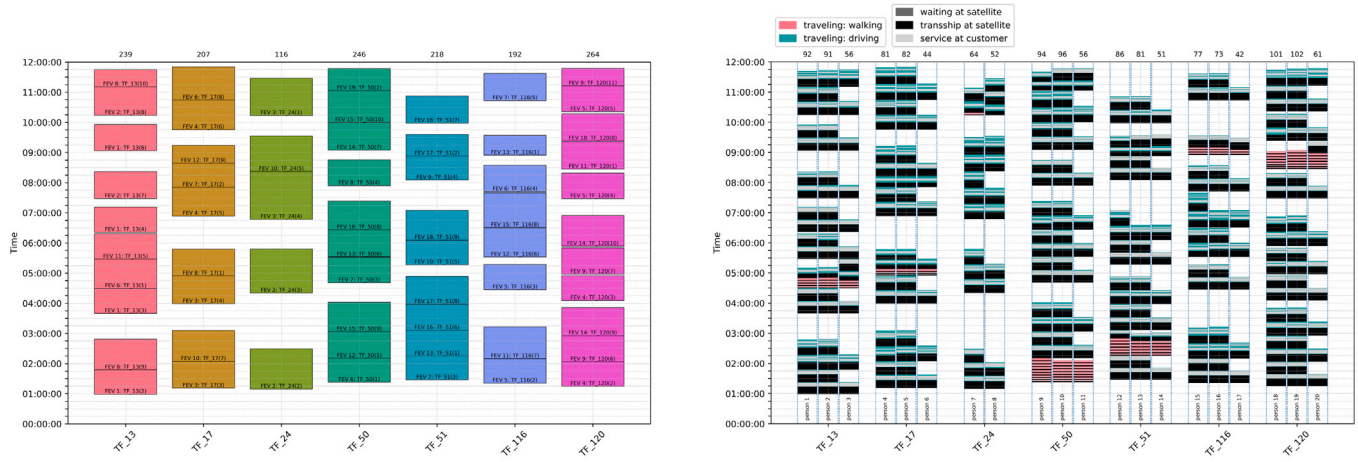
(a) Satellite occupancy (b) SE routes

Fig. 9. Satellite occupancy and SE routes of final solution (100% service level) for configuration: $\texttt{FW}_3$, 1TM, $\mathcal{W}$ = 12 hours, SND 7, demand set 1.

demand patterns through adaptive workforce allocation compared to the FW model. On the other hand, when these feasible boundaries and workforce requirements are well understood, manually setting workforce levels in the FW model can yield more efficient personnel utilization, as the FW model relies exclusively on reallocating satellite workload without increasing workforce. In contrast, the AW model does not immediately reallocate satellite workload in the presence of infeasibility but instead prioritizes incrementally increasing satellite workforce to resolve overloads. This reflects another general trade-off: reallocation leads to higher SE travel distances, while prioritizing workforce increases (within limits) over reallocation may result in requiring more personnel. Assessing such trade-offs is beyond the scope of this case study, instead, we aim to highlight the operational trade-offs and system design requirements for attaining 100% ASL.

6.2. Analysis of 1TM transshipment operations

This section presents an analysis of service design configurations under various operational settings and multiple satellite SNDs, with all satellites employing parallel 1TM transshipment operations. The evaluation focuses on system requirements and trade-offs among KPIs and model variants to attain the targeted 100% ASL. To evaluate the performance impact of multi-trip routing, a single-trip benchmark for FEVs was also conducted, with the results detailed in Appendix A.



(a) Satellite occupancy (b) SE routes

Fig. 10. Satellite occupancy and SE routes of final solution (100% service level) for configuration: $\texttt{AW}_3$, 1TM, $\mathcal{W}$ = 12 hours, SND 7, demand set 1.

6.2.1. System performance under 1TM operations for 8-hour time window

The attained service levels for various FW and AW model variants under an 8-hour time window across different satellite SNDs are shown in Fig. 11. Under 1TM transshipment operations, the corresponding minimum number of required satellites and associated daily resources are listed below (see Table C.2 in Appendix C for more details):

- $\texttt{FW}_1$: 25 satellites, 25 personnel, 44 vessels, and 425 street km
- $\texttt{FW}_2$: 14 satellites, 28 personnel, 33 vessels, and 420 street km
- $\texttt{FW}_3$: 10 satellites, 30 personnel, 29 vessels, and 400 street km
- $\texttt{AW}_3$: 10 satellites, 29 personnel, 30 vessels, and 398 street km
- $\texttt{AW}_4$: 9 satellites, 33 personnel, 29 vessels, and 420 street km
- $\texttt{AW}_5$: 6 satellites, 29 personnel, 26 vessels, and 478 street km

The minimum requirements demonstrate that increasing the available workforce at the satellites effectively mitigates workload imbalances, thereby reducing the number of required satellites. For instance, the $\texttt{FW}_1$ scenarios, where each satellite is limited to a single person to perform SE deliveries, rely on a denser satellite network to distribute the demand and achieve full coverage. In contrast, the $\texttt{AW}_5$ scenarios resolve these imbalances by dynamically deploying additional personnel at fewer satellites.

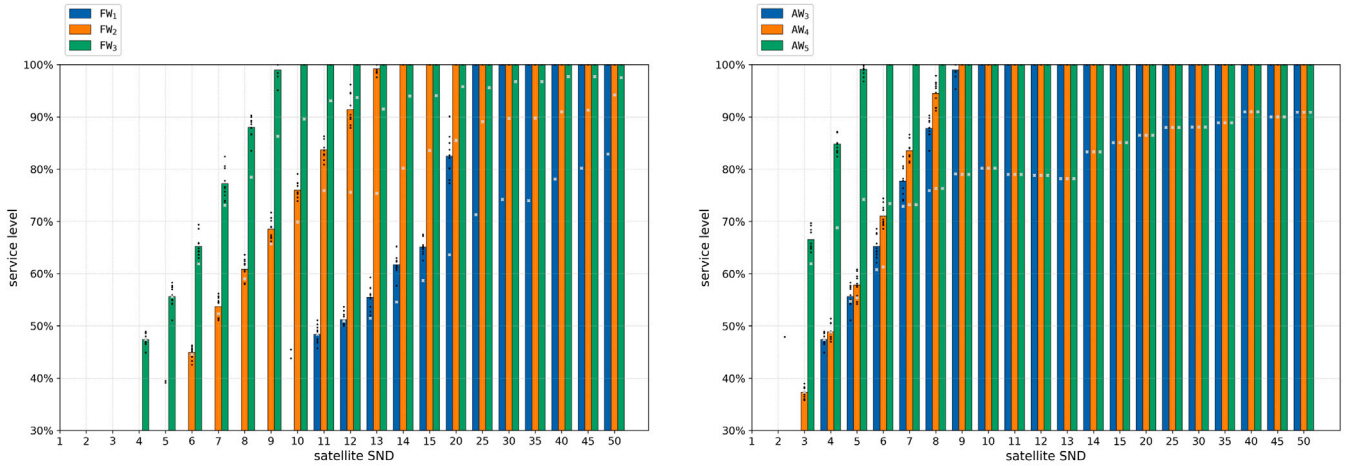
The daily system requirements for attaining 100% ASL with various model variants for an 8-hour time window under 1TM transshipment operations are shown in Fig. 12. Attaining full service coverage under 1TM transshipment constraints requires a minimum of 6 satellites across all evaluated configurations. The results indicate that the $\texttt{AW}_5$ model variant minimizes the required number of satellites. Additionally, both the FW and AW variants demonstrate that, for a given satellite SND, increasing the number of personnel reduces the number of vessels and street-level kilometers required. When comparing the two model variants, the AW variants consistently result in fewer street-level kilometers than the FW variants for the same SND. This can be explained by the fact that the workforce limit for the AW model is set to a maximum of five personnel per satellite, whereas the FW variants were only tested with up to three personnel per satellite. This lower workforce limit in the FW variants causes more reallocation of customers to balance satellite workloads. The FW model was not tested with a fixed workforce of five personnel per satellite due to excess workforce at satellites in low-demand areas causing substantial idle time for personnel. In contrast, the AW model mitigates this through workload-based personnel allocation. Overall, the FW and AW models employ distinct approaches to resolve infeasibilities arising from satellite overloads, as detailed at the end of Section 6.1, which complicates direct comparison between model variants.

Increasing the number of satellites reduces street-level travel distances, while water-level distances rise slightly due to additional inter-satellite travel for partial FEV loads. More available satellites also increase the share of deliveries made by walking, lowering energy consumption, emissions, and vehicle use. Moreover, a denser satellite network mitigates local congestion and disturbances in the vicinity of each satellite and enhances resilience to disruptions such as canal or road blockages, satellite failures, or maintenance. These benefits must be balanced against potential drawbacks, including a larger urban footprint occupying scarce public space in city centers, as well as higher implementation and maintenance costs.

6.2.2. Minimum system requirements under 1TM operations

The minimum system requirements are derived by analyzing the KPIs in isolation, identifying their respective minimum observed requirements necessary to attain 100% ASL (i.e., $\eta = 1.0$ for all instances). These minimum observed requirements, presented in Table 4, reflect the minimal values for satellites, vessels, and personnel across different model variants and time window spans. The minimum observed requirements per KPI correspond to the minimum values observed (per model variant and time window span) across Tables C.1–C.4 in Appendix C. While these minimum observed requirements may not be jointly attainable from a single service network configuration, they define the minimal resource requirements per KPI to attain 100% ASL.

The results show that with a 4-hour time window, the multi-trip capability of FEVs is largely ineffective, as vessel requirements closely match the single-trip benchmark (see Appendix A). This is primarily because loading and round-trip travel between the DC and satellites already takes approximately two hours, excluding service time, leaving insufficient time for most FEVs to complete multiple trips. An exception is observed in the $\texttt{AW}_5$ variant, where reduced satellite service times enable some FEVs to perform additional trips. Increasing the city access time window to 8 hours reveals the advantage of multi-trip FEV routing, with vessel requirements dropping to less than half of the single-trip benchmark. This FEV fleet size efficiency gain is accompanied by a modest increase in the required number of satellites and personnel, due to the scheduling complexity of coordinating multiple trips without violating space constraints that limit FEV occupation to one at a time, which slightly reduces utilization of available working hours at satellites. Overall, extending the city access time window consistently reduces the minimum system requirements necessary to attain 100% ASL.



(a) FW model variants (b) AW model variants

Fig. 11. Service level results for an 8-hour time window and 1TM operations. The bars show the final average service level across 10 demand sets, with black dots indicating individual results. The lightgray crosses (×) indicate the ASL of iteration 1, representing the cheapest initial customer-satellite allocation. Missing results indicate early termination by TC 1 in iteration 1.

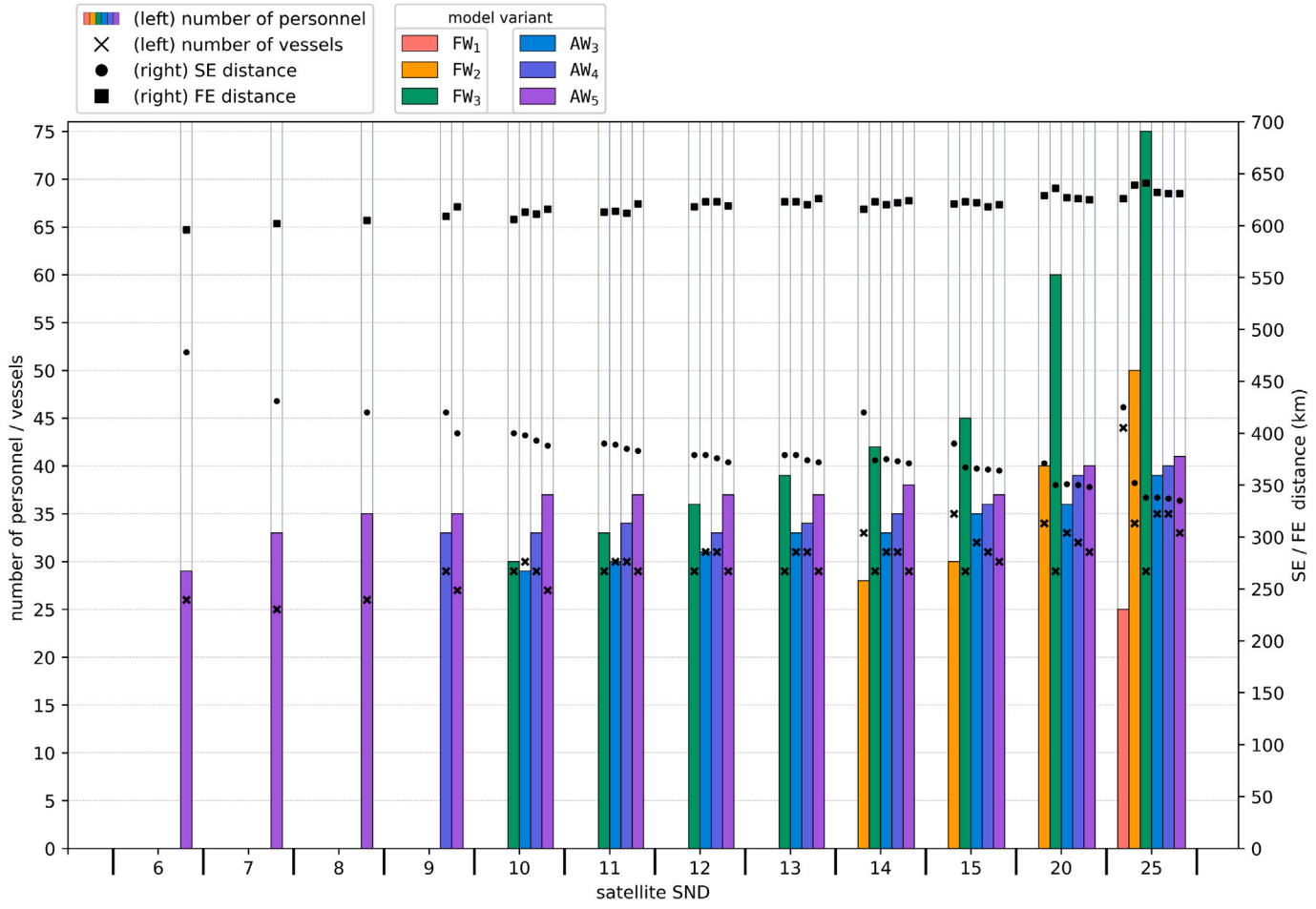


Fig. 12. Daily system requirements for attaining 100% ASL with FW and AW models under 1TM operations (8-hour time window, SND 6–25).

Table 4
Minimum observed requirements per KPI for attaining 100% ASL ($\eta = 1.0$), evaluated across 10 demand sets, for different FW and AW model variants with 1TM transshipments. These minimum requirements are evaluated individually per KPI and may not be jointly attainable within a single system configuration.

	Satellites						Vessels						Personnel					
	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5
$\mathcal{W}$ = 4 hours	✗	35	25	25	20	**13**	✗	63	63	63	62	**58**	✗	70	75	68	73	**64**
$\mathcal{W}$ = 8 hours	25	14	10	10	9	**6**	40	32	28	30	29	**25**	**25**	28	30	29	33	29
$\mathcal{W}$ = 12 hours	20	9	7	7	6	**4**	25	20	17	19	19	**16**	**20**	18	21	**20**	22	**20**
$\mathcal{W}$ = 16 hours	12	7	5	5	5	**3**	19	15	13	14	15	**12**	**12**	14	15	15	17	15

Table 5
Savings in minimum observed requirements per KPI for attaining at least 90% ASL with 1TM transshipment operations, compared to 100% ASL results (Table 4).

	Satellites						Vessels						Personnel					
	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5
$\mathcal{W}$ = 4 hours	✗	−5	−5	−5	0	−2	✗	−1	−2	−2	0	−5	✗	−10	−15	−8	0	−9
$\mathcal{W}$ = 8 hours	0	−2	−1	−1	−1	−1	0	−2	0	−1	−2	0	0	−4	−3	−2	−1	−4
$\mathcal{W}$ = 12 hours	−6	−1	−1	−1	0	−1	0	0	0	0	0	−1	−6	−2	−3	−2	0	−5
$\mathcal{W}$ = 16 hours	−1	−1	0	0	−1	0	0	0	0	0	−2	0	−1	−2	0	0	−1	0

6.2.3. Impact of relaxing the ASL target under 1TM operations

Beyond the trade-offs between KPIs, another consideration lies in the trade-off between targeted service levels and the corresponding system requirements under 1TM transshipment operations. While this study primarily evaluates configurations that ensure 100% ASL, practical implementations of the multimodal two-echelon system may tolerate lower service coverage to reduce operational costs. In such cases, unserved demand can be fulfilled through supplementary direct road-based deliveries from the DC.

To estimate the potential resource savings from relaxing the service level requirement, this section presents a post-hoc analysis of the primary experimental runs (which were configured with a target service level of $\eta = 1.0$). For this analysis, we identified system configurations that, while failing to attain full coverage, successfully served at least 90% of the demand. The reported savings therefore represent a conservative estimate, as the procedure did not leverage the early termination logic defined in TC 2 for a 90% target. Configuring the framework to explicitly terminate once $\eta = 0.9$ is met would likely yield even greater resource savings, as it would prevent the procedure from attempting to serve the final, most difficult-to-reach 10% of customer demand.

The savings in minimum system requirements from not serving all customers and maintaining at least 90% ASL under 1TM transshipment operations are shown in Table 5. These savings represent differences in the minimum observed requirements per KPI for identical model variants and time window spans compared to maintaining 100% ASL. As expected, required resources decrease as the ASL target is relaxed, although the relative savings diminish as the time window span increases.

6.3. Analysis of 1T1 transshipment operations

Fig. 13 illustrates the service levels attained under 1T1 transshipment operations within an 8-hour time window, comparing the FW and AW model variants across different satellite SNDs. Notably, guaranteeing 100% ASL under 1T1 operations necessitates a denser satellite network than under 1TM, reflecting the operational bottleneck created by sequential transshipments. Fig. 14 further details the daily system requirements under sequential transshipment operations, illustrating the trade-offs between physical resources and operational effort.

Service levels for FW_2 and FW_3 remain nearly identical across all satellite SNDs, as shown in Fig. 13(a), indicating minimal benefit from adding a third person per satellite under 1T1 operations. Similarly, increasing the workforce limit from two to four personnel yields little service level improvement for AW_2, AW_3, and AW_4, as illustrated in Fig. 13(b). In contrast, adding personnel under 1TM transshipment operations significantly improves service levels, since each person operates independently without waiting times caused by sequential transshipments. Under 1T1 constraints, this operational independence is lost. For instance, if four SE trips are assigned to a satellite with four personnel, the fourth person must wait for three prior transshipments after the FEV's arrival. By that time, the first person may have already completed their round trip, reducing the added value of the fourth person and leading to diminishing service level improvements when assigning more than two personnel per satellite.

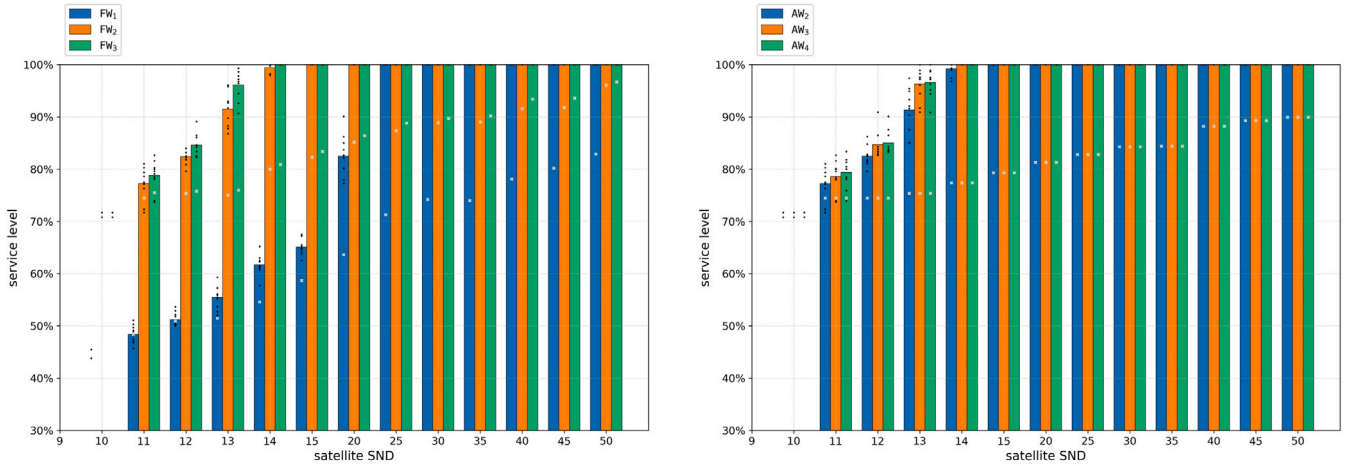
The minimum observed resource requirements for sequential transshipment operations are evaluated across multiple demand sets and time windows. Table 6 summarizes these requirements for the different model variants. These values correspond to the minimum results recorded across Tables D.1–D.4 in Appendix D. Similar to the parallel transshipment analysis, they represent the lowest resource levels necessary to attain 100% ASL. The inefficiency of assigning more than two personnel per satellite is evident in the FW_3, AW_3, and AW_4 variants, where additional personnel result in marginal reductions in the number of satellites or vessels required to attain 100% ASL.

6.4. Comparison of 1T1 and 1TM transshipment operations

To illustrate the impact of sequential transshipment constraints, Table 7 presents the increase in minimum resource requirements under 1T1 operations compared to 1TM across various model variants and time windows. The model variant AW_2 was not evaluated under 1TM operations and is therefore excluded from the direct comparison. The comparison highlights the pronounced impact of sequential transshipment constraints under restrictive time windows. The additional resource requirements for satellites, vessels, and personnel are highest for the 4-hour time window and diminish as the service window expands to 8, 12, and 16 hours. This trend indicates that the efficiency gains of parallel transshipment operations are most significant when operational time is limited. Consequently, adopting 1TM strategies can substantially reduce system requirements under short service windows, enabling more effective utilization of satellite time across the network.

6.5. Street-level operations and transshipment time sensitivity

Street-level operations consist of transshipment time at satellites, service time at customers, travel time, and waiting time due to sequential transshipments. To analyze the distribution of time across these activities, each customer is assigned to the nearest available satellite, without considering constraints on satellite workload, personnel availability, or FEV availability. This abstraction represents a best-case, unrestricted scenario for SE operations and enables a configuration-independent analysis, offering insight into typical SE activity durations. On average, the total duration of street-level operations is composed of:



(a) FW model variants (b) AW model variants

Fig. 13. Service level results for an 8-hour time window and 1T1 operations. The bars show the final average service level across 10 demand sets, with black dots indicating individual results. The lightgray crosses (×) indicate the ASL of iteration 1, representing the cheapest initial customer-satellite allocation. Missing results indicate early termination by TC 1 in iteration 1.

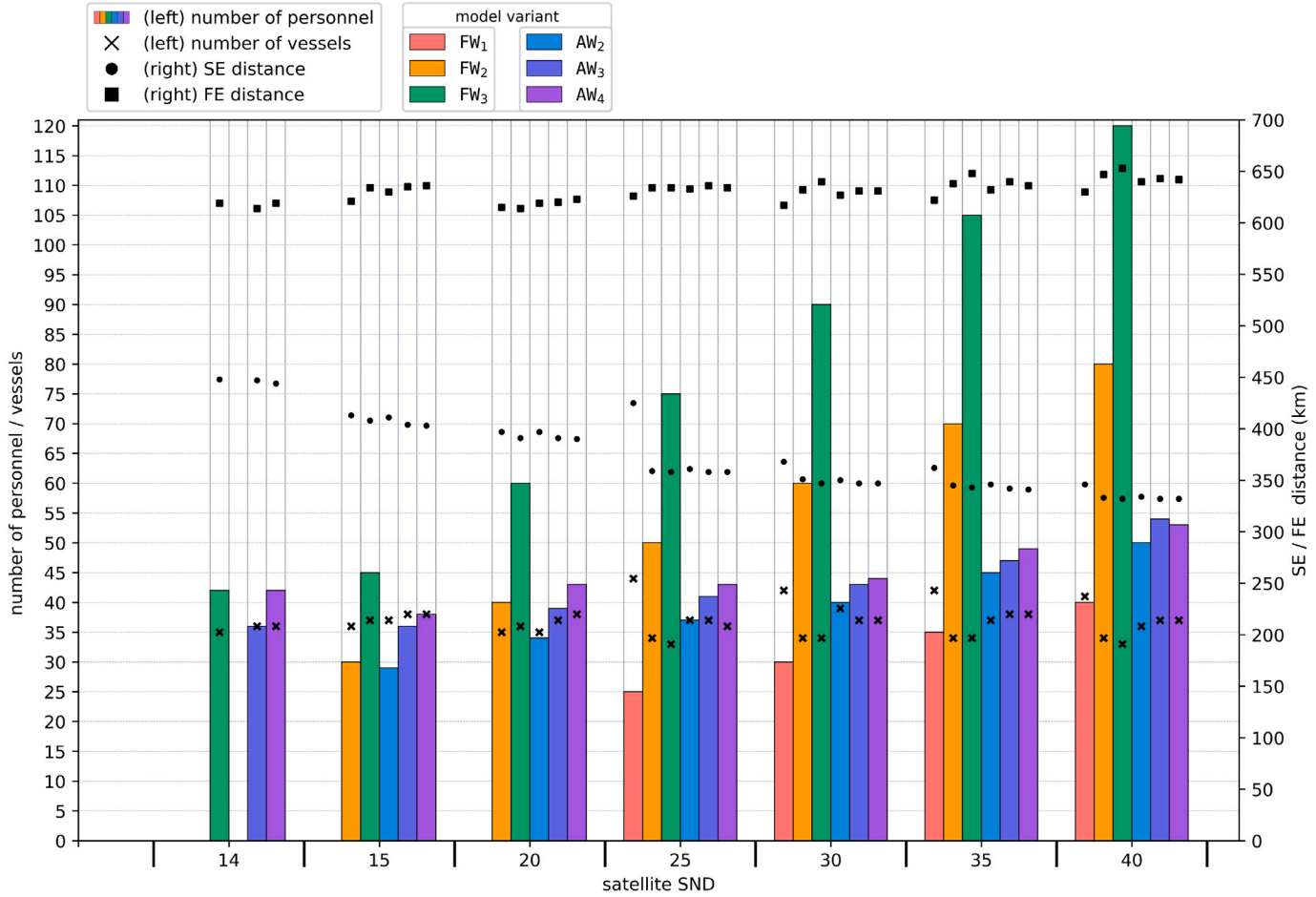


Fig. 14. Daily system requirements for attaining 100% ASL with FW and AW models under 1T1 operations (8-hour time window, SND 14–40).

Table 6
Minimum observed requirements per KPI for attaining 100% ASL ($\eta = 1.0$), evaluated across 10 demand sets, for different FW and AW model variants with 1T1 transshipments. These minimum requirements are evaluated individually per KPI and may not be jointly attainable within a single system configuration.

	Satellites						Vessels						Personnel					
	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4
$\mathcal{W}$ = 4 hours	✗	**35**	**35**	**35**	**35**	**35**	✗	64	**63**	65	64	64	✗	70	105	**67**	82	87
$\mathcal{W}$ = 8 hours	25	15	**14**	15	**14**	**14**	40	**33**	**33**	35	36	36	**25**	30	42	29	36	38
$\mathcal{W}$ = 12 hours	20	**10**	**10**	**10**	**10**	**10**	25	**21**	**21**	22	22	23	**20**	**20**	30	**20**	25	26
$\mathcal{W}$ = 16 hours	12	8	**7**	8	**7**	**7**	19	16	**15**	17	16	17	**12**	16	21	16	17	18

Table 7
Increased minimum requirements per KPI to attain 100% ASL with 1T1 (Table 6) compared to 1TM transshipments (Table 4).

	Satellites						Vessels						Personnel					
	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4
$\mathcal{W}$ = 4 hours	✗	0	+10	–	+10	+15	✗	+1	0	–	+1	+2	✗	0	+30	–	+14	+14
$\mathcal{W}$ = 8 hours	0	+1	+4	–	+4	+5	0	+1	+5	–	+6	+7	0	+2	+12	–	+7	+5
$\mathcal{W}$ = 12 hours	0	+1	+3	–	+3	+4	0	+1	+4	–	+3	+4	0	+2	+9	–	+5	+4
$\mathcal{W}$ = 16 hours	0	+1	+2	–	+2	+2	0	+1	+2	–	+2	+2	0	+2	+6	–	+2	+1

- 53.3% transshipment time at satellites (71 hours)
- 26.7% service time at customer locations (35 hours)
- 20.0% travel time between customers and satellites (25 hours)

In this unrestricted scenario for SE operations, the total operational time averages 156 hours with one available satellite and decreases to 128 hours with 50 available satellites, primarily due to reduced travel time. The proportion of trips performed by walking increases from 0.4% with one satellite to 14.5% with 50 satellites, due to more customers being within walking range as the density of satellites in the network increases. To gain general insights into waiting time caused by 1T1 transshipment operations, we consider model variants FW_2, FW_3, and FW_4, disregarding FEV availability. Across the 22 satellite SNDs, total waiting time for personnel varies within a range of 17–26 hours for FW_2, 68–77 hours for FW_3, and 114–119 hours for FW_4. This waiting time adds a substantial operational burden for SE activities and can significantly affect service coverage in configurations with limited available working hours. Specifically, depending on the satellite SND, waiting time for personnel due to 1T1 operations increases the total required operational time in the SE (in addition to transshipment, travel, and customer service time) by a range of 10.8–20.4% for FW_2, 43.9–59.3% for FW_3, and 73.2–92.0% for FW_4.

Given the substantial share of transshipment time in street-level operations and the waiting time induced by the 1T1 transshipment constraint, we conduct a sensitivity analysis to evaluate the impact of varying transshipment time (TT) on system requirements. Starting from baseline values (TT = 100%) of 120 s/unit for walking and 180 s/unit for driving, as used in the main case study, these transshipment times are reduced to 75%, 50%, and 25%.

The sensitivity analysis quantifies how system requirements change as handling efficiency improves. Table 8 summarizes minimum observed requirements necessary to attain 100% ASL for an 8-hour time window under 1T1 transshipment operations. As discussed in Section 6.3, assigning more than two personnel per satellite under the 1T1 constraint with TT = 100% is inefficient due to significant waiting times, yielding minimal service level gains. However, as transshipment times decrease, these inefficiencies diminish, and configurations with three or more personnel become more effective, lowering system requirements to attain 100% ASL. Additional detailed results for this sensitivity analysis are provided in Tables E.1–E.3 in Appendix E.

6.6. Analysis of computational runtime

The computational runtime exhibits substantial variability across scenarios, shaped by workforce availability and model variant, satellite SND, city access time window, and specific demand set characteristics. As reported in Tables F.1–F.4 in Appendix F, runtimes are presented exclusively for scenarios that attained 100% ASL, since scenarios failing to meet this target terminate prematurely and not representative. The reported runtimes reflect the full execution of the methodology, including detailed intermediate result tracking, comprehensive logging, and the generation of visual KPIs for analysis.

Under identical configurations of model variant, satellite SND, and time window, the computational time can vary significantly between demand sets due to differences in the spatial distribution of customer locations. Quantitatively, for scenarios under 1TM transshipment operations that attained 100% ASL, the average runtime was 1102 s, with values ranging from 18 to 10037 s. In the more constrained 1T1 setting, which necessitates more complex scheduling, the average runtime increased to 1479 s, with values ranging from 20 to 8087 s.

Problem complexity increases as satellite density decreases or time window constraints tighten. Under such conditions, a limited number of satellites must accommodate a larger set of customers within tighter spatial, temporal, and workforce constraints, complicating the search for feasible assignments and synchronized schedules. In real-world urban logistics, the number and location of satellite facilities are often limited by public space availability, financial costs, and regulatory restrictions. Consequently, satellite assignment becomes not only an operational but also a strategic planning decision, requiring careful trade-offs between feasibility and service quality. Even well-designed systems may struggle to meet service targets under these constraints, highlighting the importance of robust design and planning methodologies that explicitly account for real-world limitations.

7. Conclusions and future research directions

This study addresses the complex operational challenges of multi-modal city logistics by formulating a highly constrained multi-trip 2E-VRP variant that integrates exact spatial and temporal synchronization for transshipments at storage-free satellites. The problem incorporates strict vehicle occupancy limits, load-dependent transshipment times, multi-trip routes on both echelons, workforce allocation, distinct transshipment modalities, and urban access restrictions. To evaluate system performance under these conditions, we develop a flexible iterative decomposition-based heuristic. This modular framework integrates a capacitated assignment model with routing heuristics via an adaptive workload-bounding feedback loop. It operates through an allocation phase that assigns customers to satellites, a routing phase that constructs fully synchronized schedules to evaluate these assignments, and a bounding phase that tightens satellite workload bounds in response to observed operational infeasibilities. This iterative mechanism ensures that downstream routing constraints actively shape upstream customer-to-satellite assignments to find a feasible operational routing schedule

Table 8
Minimum observed requirements for 100% ASL under varying transshipment times (TT) under 1T1 operations (8-hour window).

	Satellites						Vessels						Personnel					
	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4
TT = 100%	25	15	**14**	15	**14**	**14**	40	**33**	**33**	35	36	36	**25**	30	42	29	36	38
TT = 75%	25	13	**11**	13	**11**	**11**	38	31	**30**	33	32	32	**25**	26	33	**25**	32	34
TT = 50%	20	11	**9**	10	**9**	**9**	35	29	**26**	30	29	30	**20**	22	27	**20**	25	26
TT = 25%	15	9	7	9	7	**6**	33	25	**23**	29	26	26	**15**	18	21	18	20	21

that attains the configured target service level (η) while heuristically minimizing resource consumption.

The proposed methodology is demonstrated through a large-scale case study of a multimodal distribution system in Amsterdam, serving HoReCa businesses via integrated canal and road networks. To derive strategic insights from this operational model, we conduct a comprehensive scenario analysis. We systematically evaluate 22 candidate satellite network designs against varying city access time windows, transshipment modalities, and workforce limits across multiple daily demand sets. By determining the minimum operational metrics necessary to guarantee full service coverage, this approach effectively quantifies the strategic trade-offs between physical infrastructure, vehicle fleets, and labor. For instance, under an 8-hour city access time window, the results demonstrate that enabling parallel transshipments and adaptive workforce allocation can drastically reduce the minimum required infrastructure from 25 down to 6 satellites, and the vessel fleet from 44 down to 26, compared to sequential transshipments and fixed workforce allocation.

From a policy and planning perspective, the Amsterdam case study results offer several managerial insights for similar dense urban environments. First, extending urban access time windows (e.g., from 4 to 8 hours) facilitates multi-trip operations, which substantially reduces the required vessel fleet size. Second, the choice of transshipment modality strongly influences the required urban footprint. When operations are restricted to sequential transshipments, adding personnel yields diminishing returns, requiring a denser network of satellites. In such cases, investing in handling efficiency to reduce transshipment times can effectively mitigate waiting inefficiencies. Conversely, facilitating parallel transshipments allows a smaller number of satellites to handle the same demand, especially with an adaptive workforce. Third, sparse satellite networks force customer reallocation to more distant facilities, leading to longer street-level travel distances, thereby compromising the intended reduction of urban road traffic. In contrast, while denser satellite networks require more urban space, they effectively shift travel from the road network to the waterways and significantly increase the feasibility of zero-emission walking deliveries. Finally, relaxing the target service level (e.g., to 90%) and fulfilling residual demand via alternative channels provides substantial infrastructure savings, particularly under restrictive access windows. Ultimately, these findings indicate that regulatory access policies, infrastructure investments, and service targets should be evaluated jointly.

Several limitations of the proposed methodology and experimental scope should be noted. First, the model currently assumes a single distribution center, while larger networks may need to rely on multiple consolidation hubs. Second, because the employed routing and scheduling method is heuristic, the resource requirements reported in this study represent conservative upper bounds instead of mathematical lower bounds. Furthermore, this means the reported service levels are best-found results for a given set of resources, not a guaranteed maximum. However, for strategic capacity planning, these conservative estimates serve as a practical baseline that guarantees operational feasibility under the highly complex and restrictive physical constraints of real-world city logistics.

Building on the strategic insights and heuristic limitations identified in this study, we propose three promising directions for future research, namely expanding the strategic network design, enhancing operational realism, and advancing the underlying solution methodologies:

- **Integrated location-routing for depots and satellites:** A relevant extension is to integrate both depot and satellite location decisions directly into the current decomposition framework. This could be achieved by reformulating the allocation phase as a multi-echelon location–allocation model to determine which depots and satellites to open. The iterative feedback loop could then guide these facility selection decisions based on downstream routing feasibility, enabling trade-off analysis between infrastructure investment and routing efficiency.
- **Infrastructure investments at satellites:** Evaluating the trade-offs of targeted infrastructure investments to relax strict spatial synchronization requirements offers a promising extension. For example, future models could assess how combinations of temporary storage capacity (e.g., floating pontoons) and expanded vessel docking areas impact transshipment throughput, overall operational feasibility, and system resource requirements.
- **Free-floating fleets and dynamic allocation:** Building on the benefits of adaptive workforce allocation, future studies could explore decoupling street vehicles and personnel from specific satellites to operate as a free-floating fleet. Evaluating this approach would highlight the trade-offs between increased spatial efficiency and the added algorithmic complexity of synchronizing first- and second-echelon routes for transshipments.
- **Stochasticity and network resilience:** While the current framework utilizes deterministic parameters, real-world urban logistics are subject to traffic congestion and service delays. Future research could introduce stochastic, time-dependent travel and handling times, as well as broader segment attributes such as costs and emissions, to assess the resilience of the generated schedules and support the development of online re-planning strategies for network disruptions.
- **Advanced multi-objective methods:** While the proposed methodology prioritizes finding a feasible solution that meets the target service level η, future research could employ a combination of advanced methods to further optimize the solution's efficiency. The current routing and scheduling heuristics could be replaced with more powerful advanced metaheuristics or exact solvers for key subproblems. Subsequently, to improve performance on secondary objectives, a two-staged approach could be adopted. For example, once a feasible schedule meeting η is generated, post-processing optimization techniques like local search could be applied to minimize travel distance or fleet size. For a more systematic analysis of these trade-offs, the framework could incorporate formal multi-objective optimization methods, such as an epsilon-constraint approach, to find a set of Pareto-optimal solutions for a given service level target.

CRediT authorship contribution statement

Bas Bijvoet: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Conceptualization. **Cigdem Karademir:** Writing – review & editing, Writing – original draft, Validation, Methodology, Conceptualization. **Bilge Atasoy:** Writing – review & editing, Validation, Methodology, Funding acquisition, Conceptualization.

Funding

This research is supported by the project 'Sustainable Transportation and Logistics over Water: Electrification, Automation and Optimization (TRiLOGy)' of the Dutch Research Council (NWO) and partly funded by the European Research Council (ERC) project 'Adaptive Transport Systems with Holistic Representation of Supply and Demand' (ADAPT-OR, No. 101117675).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, the European Research Council, the Dutch Research Council (NWO), the National Road Traffic Data Portal (NDW), or the Municipality of Amsterdam. None of these institutions can be held responsible for the content of this publication. We thank ir. M.W. (Marcel) Ludema, dr. T. (Thomas) Meindertsma, and T. (Thomas) Vernooij from the Municipality of Amsterdam for their valuable input in designing and setting up the case study as well as validating the results.

Appendix A. Single-trip benchmark analysis

To quantify the operational benefits of enabling multi-trip (MT) operations for first-echelon vehicles (FEVs), we evaluate a benchmark scenario in which each FEV is restricted to a single trip (ST). Compared to the proposed MT model, the ST benchmark imposes a restrictive operational policy that bypasses the scheduling complexity of coordinating multiple FE trips within a single route. To provide a representative comparison, this benchmark is evaluated exclusively for configurations utilizing an 8-hour city access time window.

Table A.1 and Table A.2 present the minimum observed requirements for attaining 100% ASL for the ST benchmark compared to the proposed MT model for an 8-hour time window. The results demonstrate that restricting FEVs to a single trip significantly increases the required vessel fleet size across all configurations. Because a single trip leaves a substantial portion of the 8-hour time window unutilized, the multi-trip capability allows a much smaller fleet of FEVs to iteratively resupply the satellites, significantly improving vehicle utilization.

Appendix B. Algorithms

See Algorithms 1–4 and Function 1.

Appendix C. Results for 1TM transshipment operations

See Tables C.1–C.4.

Appendix D. Results for 1T1 transshipments operations

See Tables D.1–D.4.

Appendix E. Result for sensitivity analysis on transshipment time

See Tables E.1–E.3.

Appendix F. Results for computational runtime

See Tables F.1–F.4.

Table A.1
Minimum requirements for 100% ASL: ST benchmark vs. MT model with 1TM operations and 8-hour window.

	Satellites						Vessels						Personnel					
	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5
ST benchmark	25	13	9	9	8	**5**	63	62	62	62	**61**	**61**	**25**	26	27	27	30	**25**
MT model	25	14	10	10	9	**6**	40	32	28	30	29	**25**	**25**	28	30	29	33	29

Table A.2
Minimum requirements for 100% ASL: ST benchmark vs. MT model with 1T1 operations and 8-hour window.

	Satellites						Vessels						Personnel					
	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4
ST benchmark	25	14	**13**	14	**13**	**13**	63	64	63	63	63	**62**	**25**	**25**	39	27	35	37
MT model	25	15	**14**	15	**14**	**14**	40	**33**	**33**	35	36	36	**25**	30	42	29	36	38

Algorithm 1: SE trip generation

Input:
Sets of customer assigned (C_s) to each satellite $s \in S$, customer demand d_c for all $c \in C$,
travel times $\tau^k_{a,b}$ between any two locations a, b for all $k \in \mathcal{K}$, SEV capacity Q^k and service range ρ^k_s for all $k \in \mathcal{K}$, and $s \in S$

Output:
SE trip sets $\mathcal{T}_s$ for all satellites $s \in S$

Initialize $t \leftarrow 0$ *// Initialize trip index*
foreach $s \in S$ **do**
  Initialize $\mathcal{C}^k_s \leftarrow \emptyset$ *// Initialize set of customers assigned to satellite s and SEV type k*
  foreach $c \in \mathcal{C}_s$ **do** *// Assign customer c to SEV type with smallest service range*
    foreach $k \in \texttt{sorted}(\mathcal{K})$ by ascending ρ^k_s **do**
      if $\left(\rho^k_s \geq \tau^k_{s,c}\right)$ **and** $\left(\rho^k_s \geq \tau^k_{c,s}\right)$ **then**
        for $\texttt{visit} \in \left\{1, \ldots, \left\lfloor \frac{d_c}{Q^k} \right\rfloor\right\}$ **do** *// Split deliveries to customer needed if $d_c > Q_k$*
          $\mathcal{C}^k_s \leftarrow \mathcal{C}^k_s \cup \left\{c, Q^k\right\}$ *// Assign full-capacity visit of Q^k for customer c to $\mathcal{C}^k_s$*
        if $\left(d_c \mod Q^k\right) > 0$ **then**
          $\mathcal{C}^k_s \leftarrow \mathcal{C}^k_s \cup \left\{c, \left(d_c \mod Q^k\right)\right\}$ *// Add partial or original-demand visit*
        break
  Initialize $\mathcal{T}_s \leftarrow \emptyset$ *// Initialize SE trip set for satellite s*
  foreach $k \in \mathcal{K}$ **do**
    Initialize $\mathcal{U} \leftarrow \mathcal{C}^k_s$ *// Unvisited customers (incl split deliveries) for satellite s and SEV type k*
    while $\mathcal{U} \neq \emptyset$ **do**
      // Construct new SE trip
      Update $t \leftarrow t + 1$ *// Increment SE trip index*
      Initialize $\sigma^{\text{SE}}_t \leftarrow [s]$, $\ell \leftarrow s$ *// Initialize stop sequence and current location at satellite s*
      Initialize $\delta^{\text{SE}}_t \leftarrow 0$, $q^{\text{SE}}_t \leftarrow 0$ *// Initialize duration and load of SE trip t*
      $\texttt{add\_stop} \leftarrow \texttt{False}$ *// Initialize flag*
      while $\texttt{add\_stop}$ **do**
        foreach $(c, d_c) \in \texttt{sorted}(\mathcal{U})$ by ascending $\tau^k_{\ell,c}$ **do** *// Sort by travel time from current location ℓ*
          if $q^{\text{SE}}_t + d_c \leq Q^k$ **then**
            Update $\sigma^{\text{SE}}_t \leftarrow \sigma^{\text{SE}}_t \cup \{c\}$ *// Add customer c to SE trip t*
            Update $\mathcal{U} \leftarrow \mathcal{U} \setminus \{c\}$ *// Remove c from unassigned customers*
            Update $\delta^{\text{SE}}_t \leftarrow \delta^{\text{SE}}_t + \tau^k_{\ell,c}$ *// Update trip duration with travel time to c*
            Update $q^{\text{SE}}_t \leftarrow q^{\text{SE}}_t + d_c$, $\ell \leftarrow c$ *// Update trip load and current location*
            Update $\delta^{\text{SE}}_t \leftarrow \delta^{\text{SE}}_t + \left(\theta^{k\rightarrow}_c + \theta^{k\leftarrow}_c\right) \cdot d_c$ *// Service time at c (delivery and pickup)*
            if $q^{\text{SE}}_t = Q^k$ **or** $\mathcal{U} = \emptyset$ **then**
              $\texttt{add\_stop} \leftarrow \texttt{False}$ *// Flag that trip must end*
            break
      // Return to the satellite
      Update $\sigma^{\text{SE}}_t \leftarrow \sigma^{\text{SE}}_t \cup \{s\}$ *// Add satellite s as stop to SE trip t*
      Update $\delta^{\text{SE}}_t \leftarrow \delta^{\text{SE}}_t + \tau^k_{\ell,s}$ *// Update trip duration with travel time to s*
      Determine $\gamma^{\rightarrow}_t \leftarrow \theta^{k\rightarrow}_s \cdot q^{\text{SE}}_t$ *// Determine delivery transshipment time for trip t*
      Determine $\gamma^{\leftarrow}_t \leftarrow \theta^{k\leftarrow}_s \cdot q^{\text{SE}}_t$ *// Determine backhaul transshipment time for trip t*
      Update $\delta^{\text{SE}}_t \leftarrow \delta^{\text{SE}}_t + \gamma^{\rightarrow}_t$ *// Update trip duration by adding delivery transshipment time only*
      Update $\mathcal{T}_s \leftarrow \mathcal{T}_s \cup \sigma^{\text{SE}}_t$ *// Add current trip t to the set of SE trips from satellite s*
Return $\mathcal{T}_s$ for all $s \in S$

Algorithm 2: FE satellite supply request generation

Input:
Sets of SE trips $\mathcal{T}_s$ for all satellites $s \in \mathcal{S}$, trip loads q_t^{SE}, and total durations δ_t^{SE} for each trip t
Output:
Supply request sets $\mathcal{N}_s$ for all satellites $s \in \mathcal{S}$
Set of SE trips $\mathcal{T}_{s,n}$ and demands $d_{s,n}$ per supply request (s,n) for all $s \in \mathcal{S}, n \in \mathcal{N}_s$

foreach $s \in \mathcal{S}$ **do**
  Initialize $\mathcal{N}_s \leftarrow \emptyset$ *// Initialize set of supply requests for satellite s*
  Initialize $n \leftarrow 0$ *// Initialize supply request counter*
  Initialize $\mathcal{U} \leftarrow \mathcal{T}_s$ *// Unassigned SE trips from satellite s (all SEV types)*
  Sort $\mathcal{U}$ by ascending δ_t^{SE} *// Prioritize supply of quicker SE trips*
  while $\mathcal{U} \neq \emptyset$ **do**
    Update $n \leftarrow n+1$
    Initialize $\mathcal{T}_{s,n} \leftarrow \emptyset$ *// Initialize set for n-th supply request of satellite s*
    Initialize $d_{s,n} \leftarrow 0$ *// Initialize demand for n-th supply request of satellite s*
    foreach $t \in \mathtt{sorted}(\mathcal{U})$ **do**
      if $d_{s,n} + q_t^{\mathrm{SE}} \leq Q^{\mathrm{FEV}}$ **then**
        Update $\mathcal{T}_{s,n} \leftarrow \mathcal{T}_{s,n} \cup \{t\}$ *// Assign SE trip t to supply request (s,n)*
        Update $d_{s,n} \leftarrow d_{s,n} + q_t^{\mathrm{SE}}$ *// Update demand of supply request (s,n)*
        Update $\mathcal{U} \leftarrow \mathcal{U} \setminus \{t\}$ *// Remove SE trip t from unassigned trips*
        if $d_{s,n} = Q^{\mathrm{FEV}}$ **then**
          break *// Supply request (s,n) is full*
    Update $\mathcal{N}_s \leftarrow \mathcal{N}_s \cup \mathcal{T}_{s,n}$ *// Add supply request (s,n) to set for satellite s*
Return $\mathcal{N}_s$, $\mathcal{T}_{s,n}$, and $d_{s,n}$ for all $s \in \mathcal{S}$, $n \in \mathcal{N}_s$

Function 1: Compute waiting time and earliest arrival for FEV (used in Algorithm 4)

Input:
Supply request (s,n), scheduled supply requests at satellite s: $\mathcal{O}_s$, arrival time at the DC for the current trip of FE route r: $\mathcal{A}_r^{\mathrm{DC}}$, total duration of current trip t (so far): δ_t^{FE}, travel time from current location ℓ to satellite s: $\tau_{\ell,s}^{\mathrm{FEV}}$, service duration of request (s,n): $\delta_{s,n}$
Output: (both outputs are provisional and utilized by Algorithm 4)
Waiting time $\mathtt{w}_{s,n}$ for FE route r to serve (s,n) as the next stop in the current trip t, earliest possible arrival time $\mathtt{EPA}_{s,n}$ at satellite s when serving (s,n)

Determine $\mathtt{EPA\_NO\_WAIT} \leftarrow \mathcal{A}_r^{\mathrm{DC}} + \delta_t^{\mathrm{FE}} + \tau_{\ell,s}^{\mathrm{FEV}}$
// EPA_NO_WAIT: earliest possible arrival time for the FEV if it immediately travels to satellite s for serving (s,n) without waiting.
if $\mathcal{O}_s = \emptyset$ **then**
  // Satellite s is currently unoccupied (no scheduled supply requests). Hence, upon arrival, the FEV can serve request (s,n) immediately without waiting. Note that this waiting time is not yet realized. The actual supply request to be served is determined later based on feasibility (capacity and time) and scheduling priority, as detailed in Algorithm 4.
  $\mathtt{w}_{s,n} = 0$
  $\mathtt{EPA}_{s,n} \leftarrow \mathtt{EPA\_NO\_WAIT}$
  Return $\mathtt{w}_{s,n}, \mathtt{EPA}_{s,n}$
else
  Initialize $\mathtt{EPA\_CANDIDATE} \leftarrow \mathtt{EPA\_NO\_WAIT}$
  Sort $\mathcal{O}_s$ by ascending $\mathtt{arr\_time}$
  // Now, $\mathcal{O}_s = \left\{\{arr\_time_\lambda, dep\_time_\lambda\}\right\}_{\lambda=1}^{|\mathcal{O}_s|}$ with $\mathtt{arr\_time}_1 \leq arr\_time_2 < \cdots < arr\_time_{|\mathcal{O}_s|}$
  // Iterate $\mathcal{O}_s$ to find the earliest available time slot to insert request (s,n) without overlap.
  for $\lambda \leftarrow 1$ **to** $|\mathcal{O}_s|$ **do**
    if $\mathtt{EPA\_CANDIDATE} + \delta_{s,n} < \mathtt{arr\_time}_\lambda$ **then**
      // Request (s,n) can be completed before scheduled request λ, which start at $\mathtt{arr\_time}_\lambda$, without violating space constraints at satellite s.
      break
    Update $\mathtt{EPA\_CANDIDATE} \leftarrow \max\left(\mathtt{EPA\_CANDIDATE}, \mathtt{dep\_time}_\lambda\right)$ *// Ensure arrival follows scheduled request completion*
  $\mathtt{w}_{s,n} \leftarrow \mathtt{EPA\_CANDIDATE} - \mathtt{EPA\_NO\_WAIT}$ *// Set the required (possibly zero) provisional waiting time.*
  $\mathtt{EPA}_{s,n} \leftarrow \mathtt{EPA\_CANDIDATE}$
  Return $\mathtt{w}_{s,n}, \mathtt{EPA}_{s,n}$

Algorithm 3: SE route generation

Input: Set of satellites S, Set of SE trips $\mathcal{T}_{s,n}$ for all satellites $s \in S$ and supply requests $n \in \mathcal{N}_s$, Number of persons $\mathcal{P}_s$ available at each satellite s, Trip durations δ_t^{SE} and delivery transshipment times $\gamma_t^{\rightarrow}$ for each SE trip t, Transshipment mode $\mu_s \in \{1\texttt{TM}, 1\texttt{T1}\}$ at satellite $s \in S$

Output: Supply request service times $\delta_{s,n}$, Sequence of SE trip $\sigma_{s,n}^p$ for all $s \in S, n \in \mathcal{N}_s, p \in \{1, \dots, \mathcal{P}_s\}$

foreach $s \in S$ **do**
  foreach $n \in \mathcal{N}_s$ **do**
    if $\mu_s = 1\texttt{TM}$ **then**
      Sort $\mathcal{T}_{s,n}$ by ascending δ_t^{SE} *// $\mathcal{T}_{s,n} = \{t_1, t_2, \dots\}$, sorted SE trips for request (s,n)*
    else if $\mu_s = 1\texttt{T1}$ **then**
      Sort $\mathcal{T}_{s,n}$ by descending δ_t^{SE} *// $\mathcal{T}_{s,n} = \{t_1, t_2, \dots\}$, sorted SE trips for request (s,n)*
      Initialize $\texttt{ready}^{\text{FEV}} \leftarrow 0$ *// Helper variable to compute waiting time*
      Initialize $\texttt{ready}^p \leftarrow 0$ **foreach** $p \in \{1, \dots, \mathcal{P}_s\}$ *// Helper variable to compute waiting time*
    Initialize $\sigma_{s,n}^p \leftarrow [\,]$ **foreach** $p \in \{1, \dots, \mathcal{P}_s\}$ *// Initialize trip sequence for person p for (s,n)*
    Initialize $p \leftarrow 1$ *// Initialize person index*
    foreach $t \in \texttt{sorted}(\mathcal{T}_{s,n})$ **do**
      if $\mu_s = 1\texttt{TM}$ **or** $\mathcal{P}_s = 1$ **then**
        Update $\delta_t^{\text{SE}} \leftarrow \delta_t^{\text{SE}} + \gamma_t^{\leftarrow}$ *// Update trip duration for backhaul transshipment*
      else if $\mu_s = 1\text{T1}$ **and** $\mathcal{P}_s > 1$ **then** *// If 1T1 and multiple persons*
        if $|\sigma_{s,n}^p| = 0$ **then**
          Set $\omega_t^{\rightarrow} \leftarrow \max(0,\ \texttt{ready}^{\text{FEV}} - \texttt{ready}^p)$ *// Person waits if FEV not ready for transshipment*
          Update $\delta_t^{\text{SE}} \leftarrow \delta_t^{\text{SE}} + \omega_t^{\rightarrow}$ *// Include waiting time before delivery transshipment for t*
        else if $|\sigma_{s,n}^p| \neq 0$ **then** *// Person p has already been assigned a trip for (s,n)*
          // Calculate waiting time for the backhaul transshipment of person p's previous trip. The variable $\texttt{ready}^p$ now indicates the time when person p returns to the satellite for the previous trip but still needs to transship collected goods to the FEV.
          Let $t_{\text{prev}} \leftarrow \sigma_{s,n}^p[-1]$ *// Get the previous trip assigned to person p*
          Set $\Omega_{t_{\text{prev}}}^{\leftarrow} \leftarrow \max(0,\ \texttt{ready}^p - \texttt{ready}^{\text{FEV}})$ *// FEV waits if person not ready for transshipment*
          Set $\omega_{t_{\text{prev}}}^{\leftarrow} \leftarrow \max(0,\ \texttt{ready}^{\text{FEV}} - \texttt{ready}^p)$ *// Person waits if FEV not ready for transshipment*
          Update $\delta_{t_{\text{prev}}}^{\text{SE}} \leftarrow \delta_{t_{\text{prev}}}^{\text{SE}} + \omega_{t_{\text{prev}}}^{\leftarrow} + \gamma_{t_{\text{prev}}}^{\leftarrow}$ *// Update trip duration for backhaul transshipment*
          Update $\texttt{ready}^{\text{FEV}} \leftarrow \texttt{ready}^{\text{FEV}} + \Omega_{t_{\text{prev}}}^{\leftarrow} + \gamma_{t_{\text{prev}}}^{\leftarrow}$ *// Update FEV ready time*
          Update $\texttt{ready}^p \leftarrow \texttt{ready}^p + \omega_{t_{\text{prev}}}^{\leftarrow} + \gamma_{t_{\text{prev}}}^{\leftarrow}$ *// Update person p ready time*
        // FEV and person p are now aligned and delivery transshipment can occur for trip t
        Update $\texttt{ready}^{\text{FEV}} \leftarrow \texttt{ready}^{\text{FEV}} + \gamma_t^{\rightarrow}$ *// Update FEV ready time*
        Update $\texttt{ready}^p \leftarrow \texttt{ready}^p + \delta_t^{\text{SE}}$ *// Update person p ready time (excl backhaul transshipment)*
      // Assign SE trip t to the person
      Update $\sigma_{s,n}^p \leftarrow \sigma_{s,n}^p \cup \sigma_t^{\text{SE}}$ *// Append trip t to the trip sequence of person p for (s,n)*
      Update $p \leftarrow \arg\min_{p \in \{1,\dots,\mathcal{P}_s\}} \texttt{ready}^p$ *// Select person with earliest ready time for next trip*
    if $\mu_s = 1\text{T1}$ **and** $\mathcal{P}_s > 1$ **then**
      // Process the final backhaul transshipment for each person
      Let $\texttt{persons\_by\_ready} \leftarrow \texttt{sorted}\{(p,\ \texttt{ready}^p)$ **for all** $p \in \{1, \dots, \mathcal{P}_s\}\}$ by ascending $\texttt{ready}^p$
      foreach $(p, \texttt{ready}^p) \in \texttt{persons\_by\_ready}$ **do**
        Let $t_{\texttt{last}} \leftarrow \sigma_{s,n}^p[-1]$ *// Get the last trip assigned to person p*
        Set $\Omega_{t_{\text{prev}}}^{\leftarrow} \leftarrow \max(0,\ \texttt{ready}^p - \texttt{ready}^{\text{FEV}})$ *// FEV waits if person is not yet ready for transshipment*
        Set $\omega_{t_{\texttt{last}}}^{\leftarrow} \max(0,\ \texttt{ready}^{\text{FEV}} - \texttt{ready}^p)$ *// Person waits if FEV is not ready for transshipment*
        Update $\delta_{t_{\texttt{last}}}^{\text{SE}} \leftarrow \delta_{t_{\texttt{last}}}^{\text{SE}} + \omega_{t_{\texttt{last}}}^{\leftarrow} + \gamma_{t_{\texttt{last}}}^{\leftarrow}$ *// Update trip duration for backhaul transshipment*
        Update $\texttt{ready}^{\text{FEV}} \leftarrow \texttt{ready}^{\text{FEV}} + \Omega_{t_{\text{prev}}}^{\leftarrow} + \gamma_{t_{\texttt{last}}}^{\leftarrow}$ *// Update FEV ready time*
    Set $\delta_{s,n} \leftarrow \max_{p \in \mathcal{P}_s} \left(\sum_{t \in \sigma_{s,n}^p} \delta_t^{\text{SE}} \right)$ *// Compute the service time $\delta_{s,n}$ for the supply request*
Return $\delta_{s,n}$ and $\sigma_{s,n}^p$ for all $s \in S,\ n \in \mathcal{N}_s,\ p \in \{1, \dots, \mathcal{P}_s\}$

Algorithm 4: FE route generation

Input: Set of supply requests $\mathcal{N}_s$ and service durations $\delta_{s,n}$ for all $s \in \mathcal{S}, n \in \mathcal{N}_s$
Output: Set $\mathcal{R}$ of FE routes and satellite occupancy $\mathcal{O}_s$ for all $s \in \mathcal{S}$

Initialize $\mathcal{R} = \emptyset$, $r \leftarrow 0$, $t \leftarrow 0$ // Initialize set of FE routes $\mathcal{R}$, FE route index r, and FE trip index t
Initialize $\mathcal{O}_s = \emptyset$ **foreach** $s \in \mathcal{S}$ // Initialize set of satellite occupancy (arrival and departure times)
Initialize $\mathcal{U} \leftarrow \bigcup_{s \in \mathcal{S}} \{(s,n) \mid n \in \mathcal{N}_s\}$ // Initialize set of unserved supply requests of all satellite
Initialize current_served_demand $\leftarrow 0$ // Initialize current total served demand
add_FE_route $\leftarrow$ True // Initialize flag
while add_FE_route **do**
  Update $r \leftarrow r + 1$ // Update FE route index (additional FEV is required)
  Initialize $\sigma_r^{\text{FE}} \leftarrow [\,]$ // Initialize sequence of FE trips for FE route r
  Initialize $\mathcal{A}_r^{\text{DC}} \leftarrow 0$ // Track return time to DC for starting the next trip in route r (updated at line 50)
  add_new_trip $\leftarrow$ True // Initialize flag
  while add_new_trip **do**
    Update $t \leftarrow t + 1$ // Advance to next trip index for FE route r
    Initialize $q_t^{\text{FE}} \leftarrow 0$ // Initialize load of trip t for FE route r
    Initialize $\sigma_t^{\text{FE}} \leftarrow [\text{DC}]$ // Initialize stop sequence for trip t of FE route r: start at DC
    Initialize $\delta_t^{\text{FE}} \leftarrow \theta_{\text{DC}}^{\text{FEV}\rightarrow} \cdot Q^{\text{FEV}}$ // Initialize duration of trip t for FE route r assuming full loading
    Initialize $\ell \leftarrow \text{DC}$ // Initialize current location for FE route r at DC
    add_stop $\leftarrow$ True // Initialize flag for trip t of FE route r
    while add_stop **do**
      foreach $(s,n) \in \mathcal{U}$ **do**
        // Compute provisional waiting time and feasibility to serve (s,n) next
        Determine $\mathtt{w}_{s,n}$ and $\mathtt{EPA}_{s,n}$ using Function 1
        Determine $f_{s,n}^{\text{cap}} \leftarrow$ True **if** $q_t^{\text{FE}} + d_{s,n} \le Q^{\text{FEV}}$ **else** False // Capacity feasibility check
        Determine $f_{s,n}^{\text{time}} \leftarrow$ True **if** $\mathcal{A}_r^{\text{DC}} + \delta_t^{\text{FE}} + \mathtt{w}_{s,n} + \delta_{s,n} \le \mathcal{W}$ **else** False // Time feasibility check
      // Apply hierarchical sorting to determine scheduling priority
      Sort $\mathcal{U}$ by ascending $\mathtt{EPA}_{s,n}$, break ties by descending $d_{s,n}$, then ascending $\delta_{s,n}$
      stop_added $\leftarrow$ False // Initialize flag for trip t of FE route r
      foreach $(s,n) \in$ sorted$(\mathcal{U})$ **do**
        if $f_{s,n}^{\text{cap}} =$ True **and** $f_{s,n}^{\text{cap}} =$ True **then**
          Update $\sigma_t^{\text{FE}} \leftarrow \sigma_t^{\text{FE}} \cup \{(s,n)\}$ // Append request (s,n) to the stop sequence of trip t
          Update $\mathcal{U} \leftarrow \mathcal{U} \setminus \{(s,n)\}$ // Remove (s,n) from unserved requests
          Set $\Omega_{s,n} \leftarrow \mathtt{w}_{s,n}$ // Set FEV waiting time before departing to serve request (s,n)
          arr_time $\leftarrow \mathcal{A}_r^{\text{DC}} + \delta_t^{\text{FE}} + \Omega_{s,n} + \tau_{\ell,s}^{\text{FEV}}$ // Set arrival time at satellite s
          dep_time $\leftarrow$ arr_time $+ \delta_{s,n}$ // Set departure time from s
          Update $\mathcal{O}_s \leftarrow \mathcal{O}_s \cup \{arr_time, dep_time\}$ // Update satellite occupancy
          Update $\delta_t^{\text{FE}} \leftarrow \delta_t^{\text{FE}} + \tau_{\ell,s}^{\text{FEV}} + \Omega_{s,n} + \delta_{s,n}$ // Update trip duration
          Update $\ell \leftarrow s$, $q_t^{\text{FE}} \leftarrow q_t^{\text{FE}} + d_{s,n}$
          Update current_served_demand $\leftarrow$ current_served_demand $+ d_{s,n}$ // Update served demand
          stop_added $\leftarrow$ True
          break
      add_stop $\leftarrow$ False **if** (stop_added $=$ False **or** $q_t^{\text{FE}} = Q^{\text{FEV}}$)
      add_stop $\leftarrow$ add_new_trip $\leftarrow$ False **if** $(f_{s,n}^{\text{time}} =$ False **foreach** $(s,n) \in \mathcal{U})$
      add_stop $\leftarrow$ add_new_trip $\leftarrow$ add_FE_route $\leftarrow$ False **if** $(\mathcal{U} = \emptyset)$
      // Optional: terminate if the target service level (η) is attained
      add_stop $\leftarrow$ add_new_trip $\leftarrow$ add_FE_route $\leftarrow$ False **if** $\left(\frac{\text{current\_served\_demand}}{\text{total\_demand}} \ge \eta\right)$
    // Return to the DC
    Update $\sigma_t^{\text{FE}} \leftarrow \sigma_t^{\text{FE}} \cup \{\text{DC}\}$ // Append DC as final stop in trip t of FE route r
    Update $\delta_t^{\text{FE}} \leftarrow \delta_t^{\text{FE}} + \tau_{\ell,\text{DC}}^{\text{FEV}} + \theta_{\text{DC}}^{\text{FEV}\leftarrow} \cdot Q^{\text{FEV}}$ // Update total duration of trip FE t
    Update $\sigma_r^{\text{FE}} \leftarrow \sigma_r^{\text{FE}} \cup \sigma_t^{\text{FE}}$ **if** $(|\sigma_t^{\text{FE}}| > 2)$ // Append t to route r if any request served (beyond DC–DC)
    Update $\ell \leftarrow \text{DC}$ // Update current location
    Update $\mathcal{A}_r^{\text{DC}} \leftarrow \mathcal{A}_r^{\text{DC}} + \delta_t^{\text{FE}}$ // Update DC arrival time for next trip in route r
  Update $\mathcal{R} \leftarrow \mathcal{R} \cup \sigma_r^{\text{FE}}$ **if** $(|\sigma_r^{\text{FE}}| > 0)$ // Add FE route r if at least one trip was constructed
  add_FE_route $\leftarrow$ False **if** $(|\sigma_r^{\text{FE}}| = 0)$ // Terminate: no feasible trip for FE route r
Return $\mathcal{R}$ and $\mathcal{O}_s$ for all $s \in \mathcal{S}$

Table C.1
Results 1TM transshipments and 4-hour time window span.

	Service level (average) [%]						Vessels (maximum)						Personnel (maximum)						SE distance (average) [km]						FE distance (average) [km]					
SND	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5
1-4	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
5	✗	✗	✗	✗	✗	48.0	✗	✗	✗	✗	✗	27	✗	✗	✗	✗	✗	25	✗	✗	✗	✗	✗	163	✗	✗	✗	✗	✗	279
6	✗	✗	✗	✗	30.8	55.3	✗	✗	✗	✗	19	31	✗	✗	✗	✗	24	30	✗	✗	✗	✗	120	212	✗	✗	✗	✗	176	314
7	✗	✗	✗	✗	36.0	64.3	✗	✗	✗	✗	22	36	✗	✗	✗	✗	28	35	✗	✗	✗	✗	114	233	✗	✗	✗	✗	212	374
8	✗	✗	40.9	40.9	41.1	73.6	✗	✗	25	25	26	41	✗	✗	24	24	32	40	✗	✗	129	129	140	288	✗	✗	237	237	239	422
9	✗	✗	45.4	45.4	46.2	81.0	✗	✗	28	28	29	47	✗	✗	27	27	36	45	✗	✗	142	142	162	316	✗	✗	272	272	279	478
10	✗	✗	49.6	49.6	51.6	86.9	✗	✗	31	31	32	50	✗	✗	30	30	40	50	✗	✗	167	167	183	354	✗	✗	295	295	303	507
11	✗	✗	54.0	54.0	56.4	93.4	✗	✗	34	34	35	53	✗	✗	33	33	44	55	✗	✗	183	183	199	393	✗	✗	329	329	336	552
12	✗	40.2	58.9	59.0	62.5	98.2	✗	25	37	37	38	57	✗	24	36	36	48	60	✗	137	208	209	248	439	✗	241	359	359	373	586
13	✗	43.5	62.8	63.1	67.6	100	✗	27	40	40	41	58	✗	26	39	39	52	64	✗	162	249	248	297	442	✗	258	375	376	395	599
14	✗	46.3	66.9	68.0	72.7	100	✗	29	41	43	44	58	✗	28	42	42	56	66	✗	168	261	268	316	408	✗	271	393	396	421	599
15	✗	50.4	72.7	73.4	78.0	100	✗	32	45	46	47	61	✗	30	45	45	60	69	✗	183	281	286	329	384	✗	295	428	432	453	612
20	✗	64.5	93.6	94.2	100	100	✗	42	61	61	62	60	✗	40	60	60	73	73	✗	308	484	491	493	371	✗	380	563	569	588	614
25	41.6	80.5	100	100	100	100	27	54	64	65	63	63	25	50	75	68	73	74	181	384	405	407	386	351	242	491	613	616	608	623
30	53.2	99.7	100	100	100	100	40	62	63	65	65	64	30	60	90	69	76	75	244	465	368	368	363	340	335	566	615	616	619	625
35	55.4	100	100	100	100	100	38	64	63	65	64	65	35	70	105	72	79	79	271	432	360	361	358	335	334	577	621	625	623	633
40	65.8	100	100	100	100	100	59	63	63	65	64	65	40	80	120	76	83	85	313	373	346	346	342	329	469	605	634	630	628	642
45	81.4	100	100	100	100	100	72	63	63	66	65	65	45	90	135	82	85	87	395	367	341	341	339	326	627	615	640	641	635	640
50	89.2	100	100	100	100	100	72	64	63	63	64	63	50	100	150	87	91	94	441	356	334	334	332	323	590	618	631	624	630	631

Table C.2
Results 1TM transshipments and 8-hour time window span.

	Service level (average) [%]						Vessels (maximum)						Personnel (maximum)						SE distance (average) [km]						FE distance (average) [km]					
SND	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5
1–2	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
3	✗	✗	✗	✗	37.3	66.5	✗	✗	✗	✗	10	16	✗	✗	✗	✗	12	15	✗	✗	✗	✗	120	284	✗	✗	✗	✗	208	369
4	✗	✗	47.4	47.4	48.9	84.8	✗	✗	13	13	14	21	✗	✗	12	12	16	20	✗	✗	160	160	174	400	✗	✗	268	268	280	481
5	✗	✗	55.6	55.6	57.8	99.1	✗	✗	17	17	17	25	✗	✗	15	15	20	25	✗	✗	200	201	217	509	✗	✗	328	330	344	588
6	✗	44.9	65.2	65.2	71.0	100	✗	15	19	19	21	26	✗	12	18	18	24	29	✗	171	279	277	325	478	✗	265	378	379	410	596
7	✗	53.7	77.2	77.7	83.5	100	✗	18	23	23	25	25	✗	14	21	21	28	33	✗	184	309	310	342	431	✗	323	458	462	498	602
8	✗	60.8	88.0	87.8	94.5	100	✗	20	25	25	27	26	✗	16	24	24	32	35	✗	221	376	374	421	420	✗	359	522	519	558	605
9	✗	68.5	99.0	99.0	100	100	✗	23	29	29	29	27	✗	18	27	27	33	35	✗	251	432	433	420	400	✗	417	603	603	609	618
10	✗	76.0	100	100	100	100	✗	25	29	30	29	27	✗	20	30	29	33	37	✗	294	400	398	393	388	✗	461	606	613	611	616
11	48.4	83.7	100	100	100	100	20	28	29	30	30	29	11	22	33	30	34	37	161	330	390	389	385	383	301	515	613	614	612	621
12	51.2	91.4	100	100	100	100	22	30	29	31	31	29	12	24	36	31	33	37	175	400	379	379	376	372	324	566	618	623	623	619
13	55.5	99.2	100	100	100	100	25	34	29	31	31	29	13	26	39	33	34	37	211	463	379	379	374	372	346	608	623	623	620	626
14	61.7	100	100	100	100	100	27	33	29	31	31	29	14	28	42	33	35	38	238	420	374	375	373	371	385	616	623	620	622	624
15	65.1	100	100	100	100	100	28	35	29	32	31	30	15	30	45	35	36	37	241	390	367	366	365	364	403	621	623	622	618	620
20	82.5	100	100	100	100	100	38	34	29	33	32	31	20	40	60	36	39	40	396	371	350	351	350	348	516	629	636	627	626	625
25	100	100	100	100	100	100	44	34	29	35	35	33	25	50	75	39	40	41	425	352	338	338	337	335	626	639	641	632	631	631
30	100	100	100	100	100	100	42	32	28	36	35	33	30	60	90	44	43	45	368	343	332	332	331	330	617	639	646	635	635	635
35	100	100	100	100	100	100	42	33	29	36	35	34	35	70	105	47	48	50	362	337	326	326	326	325	622	646	649	643	636	638
40	100	100	100	100	100	100	41	33	29	35	36	33	40	80	120	51	53	54	346	330	324	323	323	323	630	656	660	642	645	644
45	100	100	100	100	100	100	40	33	28	36	35	35	45	90	135	56	57	59	342	327	322	322	322	321	635	655	658	645	647	647
50	100	100	100	100	100	100	40	32	28	36	36	35	50	100	150	60	58	61	335	323	319	320	319	318	636	650	658	637	638	636

Table C.3
Results 1TM transshipments and 12-hour time window span.

	Service level (average) [%]						Vessels (maximum)						Personnel (maximum)						SE distance (average) [km]						FE distance (average) [km]					
SND	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5
1	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
2	✗	✗	✗	✗	36.6	66.3	✗	✗	✗	✗	7	11	✗	✗	✗	✗	8	10	✗	✗	✗	✗	134	321	✗	✗	✗	✗	216	389
3	✗	✗	53.7	53.7	53.9	93.2	✗	✗	10	10	10	15	✗	✗	9	9	12	15	✗	✗	198	198	200	519	✗	✗	302	302	303	524
4	✗	47.8	68.1	68.1	71.8	100	✗	11	13	13	13	16	✗	8	12	12	16	20	✗	161	278	278	311	501	✗	276	391	391	415	589
5	✗	59.0	84.6	84.5	89.5	100	✗	13	16	16	17	17	✗	10	15	15	20	23	✗	224	388	385	437	479	✗	352	507	506	536	611
6	✗	68.4	98.9	98.9	100	100	✗	15	19	19	19	18	✗	12	18	18	23	23	✗	301	522	517	506	458	✗	406	583	583	590	609
7	43.8	81.4	100	100	100	100	12	19	18	19	19	18	7	14	21	20	22	25	138	329	433	435	433	430	272	488	600	607	602	598
8	49.8	93.6	100	100	100	100	13	20	18	19	20	19	8	16	24	21	23	23	168	418	421	421	421	417	307	564	602	609	609	605
9	55.7	100	100	100	100	100	15	23	19	21	21	20	9	18	27	22	23	23	189	428	400	399	400	400	349	614	612	618	615	615
10	63.0	100	100	100	100	100	17	22	17	20	20	20	10	20	30	23	23	23	229	394	388	388	388	388	400	618	613	613	613	613
11	70.9	100	100	100	100	100	20	21	17	21	21	21	11	22	33	24	24	24	262	389	383	382	383	383	452	626	612	618	616	616
12	75.8	100	100	100	100	100	20	21	17	21	20	20	12	24	36	24	24	24	287	377	372	372	372	372	480	630	616	618	618	618
13	83.0	100	100	100	100	100	23	21	17	21	22	22	13	26	39	25	26	26	351	378	371	371	372	372	520	628	617	621	621	621
14	92.1	100	100	100	100	100	25	21	18	22	22	22	14	28	42	26	26	26	387	374	371	371	371	371	571	628	616	620	620	620
15	99.2	100	100	100	100	100	27	21	17	21	21	21	15	30	45	26	26	26	410	365	364	364	364	364	623	624	615	622	622	622
20	100	100	100	100	100	100	27	21	18	22	22	22	20	40	60	28	28	28	392	349	348	348	348	348	631	632	629	629	629	629
25	100	100	100	100	100	100	26	20	17	25	25	25	25	50	75	33	33	33	360	336	335	335	335	335	634	643	634	633	633	633
30	100	100	100	100	100	100	26	21	17	23	23	23	30	60	90	35	35	35	347	331	330	330	330	330	635	644	631	639	639	639
35	100	100	100	100	100	100	25	21	17	23	23	23	35	70	105	40	40	40	341	325	325	325	325	325	642	648	634	641	641	641
40	100	100	100	100	100	100	26	20	18	24	24	24	40	80	120	45	45	45	330	323	323	323	323	323	652	656	644	646	646	646
45	100	100	100	100	100	100	26	20	18	23	23	23	45	90	135	50	50	50	328	321	321	321	321	321	655	658	645	649	649	649
50	100	100	100	100	100	100	26	20	17	24	24	24	50	100	150	54	54	54	323	319	318	318	318	318	649	653	638	644	644	644

Table C.4
Results 1TM transshipments and 16-hour time window span.

	Service level (average) [%]						Vessels (maximum)						Personnel (maximum)						SE distance (average) [km]						FE distance (average) [km]					
SND	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5
1	✗	✗	✗	✗	✗	47.3	✗	✗	✗	✗	✗	6	✗	✗	✗	✗	✗	5	✗	✗	✗	✗	✗	265	✗	✗	✗	✗	✗	273
2	✗	✗	47.0	47.0	48.4	84.4	✗	✗	7	7	7	10	✗	✗	6	6	8	10	✗	✗	199	199	206	495	✗	✗	278	278	286	495
3	✗	46.6	66.7	66.7	68.3	100	✗	8	10	10	10	12	✗	6	9	9	12	15	✗	166	274	274	290	549	✗	266	376	376	388	575
4	✗	63.9	88.2	88.8	92.9	100	✗	10	13	13	13	12	✗	8	12	12	16	18	✗	253	409	416	459	500	✗	369	510	513	539	593
5	41.6	78.0	100	100	100	100	9	13	14	14	15	13	5	10	15	15	18	18	141	350	502	498	490	480	265	466	609	609	610	614
6	48.2	93.1	100	100	100	100	10	15	14	15	15	14	6	12	18	15	17	19	179	481	469	471	469	458	296	543	604	605	606	611
7	58.5	100	100	100	100	100	12	16	13	15	15	15	7	14	21	17	17	18	209	443	429	430	430	430	367	602	597	602	600	600
8	65.3	100	100	100	100	100	13	16	13	15	15	15	8	16	24	19	18	18	245	431	416	417	417	417	406	607	601	604	604	604
9	76.4	100	100	100	100	100	15	15	13	16	16	16	9	18	27	19	19	19	292	403	400	400	400	400	482	616	613	615	615	615
10	86.4	100	100	100	100	100	18	15	13	17	17	17	10	20	30	18	18	18	338	388	388	388	388	388	538	614	610	615	615	615
11	97.0	100	100	100	100	100	20	15	13	17	17	17	11	22	33	18	18	18	395	383	383	383	383	382	608	621	613	618	618	619
12	100	100	100	100	100	100	20	16	13	18	18	18	12	24	36	19	19	19	420	372	372	372	372	373	628	623	617	620	620	621
13	100	100	100	100	100	100	20	15	13	18	18	17	13	26	39	20	20	20	411	372	372	372	372	372	627	626	617	626	626	625
14	100	100	100	100	100	100	21	15	13	17	17	18	14	28	42	20	20	21	384	371	371	371	371	371	627	621	617	627	627	625
15	100	100	100	100	100	100	20	15	13	18	18	18	15	30	45	20	20	20	374	364	364	364	364	364	628	622	617	624	624	624
20	100	100	100	100	100	100	20	15	13	18	18	18	20	40	60	24	24	24	357	348	348	348	348	348	634	628	623	629	629	628
25	100	100	100	100	100	100	19	15	13	18	18	18	25	50	75	28	28	28	343	335	335	335	335	336	636	630	626	632	632	635
30	100	100	100	100	100	100	20	15	13	18	18	18	30	60	90	32	32	33	334	330	330	330	330	330	638	636	637	639	639	640
35	100	100	100	100	100	100	19	15	13	18	18	18	35	70	105	37	37	37	329	325	325	325	325	325	644	640	637	643	643	644
40	100	100	100	100	100	100	19	15	13	18	18	18	40	80	120	42	42	42	326	323	323	323	323	323	655	643	641	652	652	650
45	100	100	100	100	100	100	19	15	13	18	18	18	45	90	135	47	47	47	324	321	321	321	321	321	660	649	643	653	653	657
50	100	100	100	100	100	100	19	15	13	18	18	18	50	100	150	51	51	51	321	318	318	318	318	319	648	645	633	650	650	647

Table D.1
Results 1T1 transshipments and 4-hour time window span.

	Service level (average) [%]						Vessels (maximum)						Personnel (maximum)						SE distance (average) [km]						FE distance (average) [km]					
SND	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4
1–20	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
25	41.6	71.1	76.2	71.3	76.4	77.4	27	50	50	50	50	50	25	50	75	50	75	100	181	325	375	325	370	373	242	464	472	464	470	472
30	53.2	92.4	99.4	92.9	99.4	99.7	40	61	65	62	65	65	30	60	90	60	83	92	244	441	481	445	483	477	335	549	578	549	580	577
35	55.4	100	100	100	100	100	38	64	63	65	64	64	35	70	105	67	82	90	271	458	444	459	445	442	334	587	580	589	581	577
40	65.8	100	100	100	100	100	59	65	66	66	65	65	40	80	120	71	82	87	313	385	382	387	383	381	469	613	613	617	609	610
45	81.4	100	100	100	100	100	72	68	66	67	67	66	45	90	135	76	86	92	395	378	374	378	376	375	627	628	625	629	626	626
50	89.2	100	100	100	100	100	72	66	66	67	67	66	50	100	150	79	87	92	441	361	361	361	361	359	590	632	632	635	635	632

Table D.2
Results 1T1 transshipments and 8-hour time window span.

	Service level (average) [%]						Vessels (maximum)						Personnel (maximum)						SE distance (average) [km]						FE distance (average) [km]					
SND	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4
1–10	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
11	48.4	77.2	78.8	77.2	78.6	79.4	20	28	28	28	28	28	11	22	33	22	33	44	161	305	316	305	314	317	301	472	481	472	479	484
12	51.2	82.4	84.6	82.5	84.7	85.0	22	28	29	29	30	30	12	24	36	24	36	47	175	354	376	351	376	380	324	509	520	512	519	522
13	55.5	91.5	96.1	91.3	96.3	96.6	25	33	34	33	36	34	13	26	39	26	38	47	211	427	462	423	464	465	346	564	600	562	597	602
14	61.7	99.4	100	99.2	100	100	27	36	35	36	36	36	14	28	42	28	36	42	238	452	448	451	447	444	385	617	619	608	614	619
15	65.1	100	100	100	100	100	28	36	37	37	38	38	15	30	45	29	36	38	241	413	408	411	404	403	403	621	634	630	635	636
20	82.5	100	100	100	100	100	38	35	36	35	37	38	20	40	60	34	39	43	396	397	391	397	391	390	516	615	614	619	620	623
25	100	100	100	100	100	100	44	34	33	37	37	36	25	50	75	37	41	43	425	359	358	361	358	358	626	634	634	633	636	634
30	100	100	100	100	100	100	42	34	34	39	37	37	30	60	90	40	43	44	368	351	347	350	347	347	617	632	640	627	631	631
35	100	100	100	100	100	100	42	34	34	37	38	38	35	70	105	45	47	49	362	345	343	346	342	341	622	638	648	632	640	636
40	100	100	100	100	100	100	41	34	33	36	37	37	40	80	120	50	54	53	346	333	332	334	332	332	630	647	653	640	643	642
45	100	100	100	100	100	100	40	33	33	38	37	38	45	90	135	54	56	56	342	329	328	329	328	329	635	648	652	642	638	645
50	100	100	100	100	100	100	40	34	33	37	37	37	50	100	150	59	59	61	335	324	323	323	323	323	636	642	649	638	637	636

Table D.3
Results 1T1 transshipments and 12-hour time window span.

	Service level (average) [%]						Vessels (maximum)						Personnel (maximum)						SE distance (average) [km]						FE distance (average) [km]					
SND	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4
1–6	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
7	43.8	75.0	78.0	75.0	78.2	78.3	12	17	17	17	17	17	7	14	21	14	21	28	138	308	325	308	323	325	272	448	465	448	464	466
8	49.8	86.6	88.6	86.7	89.2	89.8	13	21	19	21	20	21	8	16	24	16	24	30	168	378	386	378	395	392	307	522	529	521	534	536
9	55.7	98.2	99.6	98.0	99.4	99.5	15	22	22	22	22	22	9	18	27	18	25	27	189	438	445	438	443	442	349	600	601	598	604	604
10	63.0	100	100	100	100	100	17	23	22	23	22	23	10	20	30	20	25	26	229	410	407	414	406	405	400	617	616	617	611	616
11	70.9	100	100	100	100	100	20	23	22	23	23	23	11	22	33	22	25	27	262	395	393	395	394	393	452	623	622	625	618	617
12	75.8	100	100	100	100	100	20	23	22	23	23	23	12	24	36	21	26	29	287	384	383	383	383	382	480	631	622	624	622	619
13	83.0	100	100	100	100	100	23	23	23	23	24	23	13	26	39	23	28	30	351	382	381	384	383	381	520	630	622	628	623	621
14	92.1	100	100	100	100	100	25	23	23	22	23	23	14	28	42	24	27	27	387	378	377	378	376	377	571	627	627	621	621	622
15	99.2	100	100	100	100	100	27	22	21	24	24	24	15	30	45	25	26	26	410	369	368	369	368	368	623	624	622	620	624	623
20	100	100	100	100	100	100	27	22	22	25	24	24	20	40	60	28	28	32	392	351	350	352	351	351	631	635	631	631	633	633
25	100	100	100	100	100	100	26	22	21	24	23	24	25	50	75	30	33	34	360	338	337	340	338	338	634	639	634	638	636	635
30	100	100	100	100	100	100	26	22	21	24	24	24	30	60	90	35	36	37	347	332	332	332	332	332	635	641	640	639	638	641
35	100	100	100	100	100	100	25	22	21	24	24	24	35	70	105	40	41	41	341	327	326	328	326	326	642	648	644	641	642	644
40	100	100	100	100	100	100	26	22	21	24	24	25	40	80	120	45	47	47	330	325	325	324	324	324	652	659	655	648	647	649
45	100	100	100	100	100	100	26	22	21	25	25	25	45	90	135	50	50	50	328	323	323	322	322	322	655	657	656	654	654	652
50	100	100	100	100	100	100	26	21	21	24	24	25	50	100	150	52	54	55	323	320	320	320	319	319	649	646	644	639	640	645

Table D.4
Results 1T1 transshipments and 16-hour time window span.

	Service level (average) [%]						Vessels (maximum)						Personnel (maximum)						SE distance (average) [km]						FE distance (average) [km]					
SND	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4
1–4	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
5	41.6	71.3	72.8	71.3	72.9	73.1	9	11	11	11	11	11	5	10	15	10	15	20	141	313	324	313	323	326	265	431	436	431	438	438
6	48.2	86.9	89.8	86.8	90.2	90.3	10	14	14	14	14	14	6	12	18	12	18	24	179	444	463	440	472	469	296	508	519	511	525	528
7	58.5	99.7	100	99.8	100	100	12	16	16	16	16	17	7	14	21	14	20	22	209	455	453	454	453	453	367	596	601	593	603	604
8	65.3	100	100	100	100	100	13	18	16	17	17	17	8	16	24	16	20	22	245	436	433	436	433	433	406	607	607	605	609	607
9	76.4	100	100	100	100	100	15	17	17	17	17	17	9	18	27	17	19	20	292	403	404	405	403	403	482	612	610	616	612	612
10	86.4	100	100	100	100	100	18	17	16	17	17	17	10	20	30	18	19	20	338	388	388	389	388	387	538	611	610	616	614	613
11	97.0	100	100	100	100	100	20	16	16	18	18	18	11	22	33	17	17	18	395	383	382	383	383	383	608	620	618	620	620	621
12	100	100	100	100	100	100	20	16	16	18	18	18	12	24	36	19	20	20	420	372	372	373	372	372	628	619	620	623	622	622
13	100	100	100	100	100	100	20	16	16	17	18	18	13	26	39	19	20	21	411	372	371	372	371	371	627	621	619	622	622	623
14	100	100	100	100	100	100	21	16	16	18	18	18	14	28	42	20	21	22	384	371	371	371	371	371	627	623	620	625	624	624
15	100	100	100	100	100	100	20	16	17	18	18	18	15	30	45	20	20	21	374	364	364	364	364	364	628	618	619	624	625	624
20	100	100	100	100	100	100	20	16	16	19	19	19	20	40	60	23	23	23	357	349	349	348	348	348	634	631	629	632	632	632
25	100	100	100	100	100	100	19	16	15	18	18	18	25	50	75	27	28	28	343	335	336	336	336	336	636	634	633	640	639	638
30	100	100	100	100	100	100	20	16	15	19	19	19	30	60	90	32	32	32	334	330	330	330	330	330	638	643	639	640	640	640
35	100	100	100	100	100	100	19	16	16	19	19	19	35	70	105	37	38	38	329	325	325	325	325	325	644	646	643	645	644	644
40	100	100	100	100	100	100	19	16	15	18	18	18	40	80	120	42	42	42	326	323	323	323	323	323	655	652	650	655	655	655
45	100	100	100	100	100	100	19	16	16	18	18	18	45	90	135	47	48	48	324	321	321	321	321	321	660	650	650	653	652	652
50	100	100	100	100	100	100	19	16	15	18	18	18	50	100	150	51	51	51	321	318	318	318	318	318	648	641	640	648	650	650

Table E.1
Results for the sensitivity analysis at 75% transshipment time for 1T1 transshipments with an 8-hour time window span.

	Service level (average) [%]						Vessels (maximum)						Personnel (maximum)						SE distance (average) [km]						FE distance (average) [km]					
SND	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4
1–7	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
8	38.6	66.4	72.0	66.4	72.0	72.5	16	23	23	23	23	23	8	16	24	16	24	32	121	255	285	255	285	289	230	396	429	396	429	431
9	44.6	74.5	81.4	74.5	81.5	81.9	18	24	25	24	25	25	9	18	27	18	27	35	138	286	329	286	328	329	275	462	496	462	496	499
10	49.0	84.2	91.1	84.1	91.0	92.1	20	28	28	27	29	29	10	20	30	20	30	37	164	338	381	337	378	380	297	515	549	514	550	555
11	54.9	93.6	100	93.7	100	100	22	31	34	31	35	33	11	22	33	22	32	34	187	386	419	391	418	416	336	582	614	581	611	612
12	60.3	99.6	100	99.5	100	100	24	33	32	33	32	32	12	24	36	24	32	36	215	437	408	432	407	406	388	627	618	625	619	621
13	64.2	100	100	100	100	100	26	33	32	33	32	32	13	26	39	25	34	36	256	432	406	433	408	405	394	616	623	618	617	620
14	68.3	100	100	100	100	100	28	35	31	33	32	33	14	28	42	27	34	34	263	398	388	400	386	387	410	622	621	622	625	623
15	76.4	100	100	100	100	100	30	33	32	33	33	32	15	30	45	28	34	36	298	379	373	379	374	373	473	623	622	622	625	622
20	97.8	100	100	100	100	100	40	33	32	33	34	33	20	40	60	32	35	39	498	365	359	365	359	357	618	628	627	625	630	627
25	100	100	100	100	100	100	41	32	31	36	36	36	25	50	75	35	38	39	397	350	345	349	344	343	609	643	634	634	630	633
30	100	100	100	100	100	100	39	32	31	36	35	35	30	60	90	37	41	42	365	340	335	338	335	335	615	647	641	634	631	632
35	100	100	100	100	100	100	38	32	32	36	36	36	35	70	105	42	45	46	358	335	330	332	330	330	617	652	650	637	638	637
40	100	100	100	100	100	100	40	32	31	36	36	36	40	80	120	46	49	49	344	328	327	328	327	326	629	657	660	643	640	641
45	100	100	100	100	100	100	38	32	30	35	36	36	45	90	135	52	55	56	338	325	324	325	324	323	634	663	662	646	644	642
50	100	100	100	100	100	100	38	31	31	35	35	35	50	100	150	57	57	58	333	322	321	322	321	321	631	653	654	641	640	637

Table E.2
Results for the sensitivity analysis at 50% transshipment time for 1T1 transshipments with an 8-hour time window span.

	Service level (average) [%]						Vessels (maximum)						Personnel (maximum)						SE distance (average) [km]						FE distance (average) [km]					
SND	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4
1–5	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
6	35.1	58.5	72.6	58.5	73.0	75.4	12	17	21	17	20	21	6	12	18	12	18	24	125	243	348	243	346	365	216	351	417	351	424	437
7	41.8	70.1	86.2	70.1	86.4	89.6	15	21	24	21	24	26	7	14	21	14	21	28	135	261	367	261	370	395	254	423	508	423	510	536
8	47.8	79.8	98.6	79.8	98.5	99.4	16	23	28	23	28	28	8	16	24	16	24	29	166	326	455	326	457	454	289	478	589	478	588	591
9	53.5	91.1	100	91.1	100	100	20	28	29	28	29	30	9	18	27	18	26	29	188	375	416	375	415	411	334	554	621	552	620	617
10	60.0	99.9	100	100	100	100	22	30	28	30	31	31	10	20	30	20	25	27	216	412	392	411	391	391	371	608	618	607	617	615
11	65.9	100	100	100	100	100	24	30	29	30	30	31	11	22	33	22	26	27	239	395	384	395	385	384	416	620	621	619	623	622
12	70.9	100	100	100	100	100	26	31	29	32	30	31	12	24	36	22	26	26	275	383	375	381	375	375	444	625	624	633	624	627
13	74.9	100	100	100	100	100	29	30	28	31	31	30	13	26	39	23	26	27	311	385	374	386	375	374	461	630	621	627	624	627
14	81.9	100	100	100	100	100	29	30	28	32	31	31	14	28	42	24	27	28	348	378	372	378	371	372	501	627	618	630	623	626
15	89.3	100	100	100	100	100	33	29	27	31	31	31	15	30	45	24	26	26	373	369	366	368	364	365	547	628	620	624	619	619
20	100	100	100	100	100	100	37	30	27	33	32	32	20	40	60	29	31	33	439	352	350	351	351	350	621	633	631	628	630	630
25	100	100	100	100	100	100	36	29	26	33	34	33	25	50	75	31	33	34	372	339	337	338	337	336	631	643	635	637	635	635
30	100	100	100	100	100	100	36	30	27	33	32	33	30	60	90	36	37	38	357	333	331	333	331	330	630	642	644	636	634	635
35	100	100	100	100	100	100	36	30	26	33	32	33	35	70	105	41	42	42	350	328	326	327	326	326	639	648	640	642	641	642
40	100	100	100	100	100	100	36	30	28	33	32	32	40	80	120	45	47	47	336	325	323	324	323	323	645	657	646	646	646	646
45	100	100	100	100	100	100	35	29	26	33	32	32	45	90	135	50	51	52	333	323	321	322	321	321	645	659	650	642	642	642
50	100	100	100	100	100	100	35	29	26	33	33	33	50	100	150	55	53	53	326	320	318	320	317	318	637	654	646	637	638	638

Table E.3
Results for the sensitivity analysis at 25% transshipment time for 1T1 transshipments with an 8-hour time window span.

	Service level (average) [%]						Vessels (maximum)						Personnel (maximum)						SE distance (average) [km]						FE distance (average) [km]					
SND	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4
1–2	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
3	24.0	40.0	51.8	40.0	51.8	58.0	8	11	13	11	13	14	3	6	9	6	9	12	55	119	185	119	185	224	136	225	290	225	290	323
4	31.6	52.4	68.8	52.4	68.8	78.0	11	14	18	14	18	20	4	8	12	8	12	16	94	196	276	196	276	341	187	305	394	305	395	448
5	38.6	63.2	84.2	63.2	84.3	93.5	13	18	23	18	23	24	5	10	15	10	15	20	126	243	380	243	385	457	232	381	505	381	505	558
6	45.0	72.5	98.5	73.2	98.6	100	16	20	26	22	27	27	6	12	18	12	18	23	175	324	516	331	518	492	268	427	580	431	579	594
7	53.0	88.8	100	89.3	100	100	18	25	26	25	26	26	7	14	21	14	20	22	189	374	435	379	436	432	320	534	606	542	604	604
8	60.5	99.9	100	99.8	100	100	20	28	25	28	27	26	8	16	24	16	21	23	226	453	423	453	424	421	362	603	607	602	615	607
9	68.1	100	100	100	100	100	23	28	25	29	28	28	9	18	27	18	22	23	258	406	399	407	400	399	425	624	616	623	623	622
10	75.0	100	100	100	100	100	25	28	25	30	28	28	10	20	30	19	23	23	292	389	388	390	388	388	457	622	611	618	621	621
11	81.3	100	100	100	100	100	27	28	24	30	28	28	11	22	33	20	21	21	321	382	383	385	383	383	504	624	615	624	625	624
12	87.3	100	100	100	100	100	31	27	24	29	30	30	12	24	36	20	22	22	378	374	372	374	372	372	550	625	617	629	624	624
13	94.3	100	100	100	100	100	33	28	24	29	29	29	13	26	39	21	22	23	426	373	371	373	372	372	586	622	616	627	624	623
14	99.8	100	100	100	100	100	35	28	24	30	29	29	14	28	42	22	24	24	438	371	371	371	371	371	623	624	616	631	629	629
15	100	100	100	100	100	100	34	27	24	31	30	30	15	30	45	23	25	25	399	364	364	364	364	364	630	624	615	632	631	631
20	100	100	100	100	100	100	34	27	24	31	30	30	20	40	60	26	26	26	378	349	348	349	348	348	635	630	625	634	634	634
25	100	100	100	100	100	100	34	26	23	31	31	31	25	50	75	28	29	29	354	336	335	335	335	335	642	630	625	641	640	640
30	100	100	100	100	100	100	33	26	23	32	31	31	30	60	90	34	33	33	344	330	330	330	330	330	641	636	630	643	644	644
35	100	100	100	100	100	100	34	26	23	32	31	31	35	70	105	38	39	39	338	325	325	326	325	325	654	636	634	648	651	651
40	100	100	100	100	100	100	33	25	23	32	32	32	40	80	120	43	44	44	329	323	323	323	323	323	656	645	638	653	655	655
45	100	100	100	100	100	100	33	26	23	32	32	32	45	90	135	47	49	49	327	321	321	321	321	321	659	650	642	652	651	651
50	100	100	100	100	100	100	33	25	23	32	32	32	50	100	150	51	52	52	323	318	318	318	318	318	644	642	633	647	649	649

Table F.1
Average computational runtime (seconds) of the demand sets for 1TM configurations (runtime is only given for scenarios with 100% ASL).

	$\mathcal{W}$ = 4 hours						$\mathcal{W}$ = 8 hours						$\mathcal{W}$ = 12 hours						$\mathcal{W}$ = 16 hours					
SND	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5
1–2	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
3	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	1316
4	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	968	✗	✗	✗	✗	✗	138
5	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	159	✗	✗	1995	2607	2552	91
6	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	3870	✗	✗	✗	✗	4733	94	✗	✗	952	2598	1352	94
7	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	202	✗	✗	1272	3059	1644	88	✗	1805	236	570	68	61
8	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	396	✗	✗	926	2658	1141	68	✗	1311	131	446	55	50
9	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	4151	93	✗	1600	513	1059	147	58	✗	1185	22	61	62	57
10	✗	✗	✗	✗	✗	✗	✗	✗	1478	3176	1683	76	✗	1361	93	54	54	49	✗	240	21	55	55	51
11	✗	✗	✗	✗	✗	✗	✗	✗	1438	3010	2163	75	✗	1159	22	731	57	54	✗	236	21	59	59	51
12	✗	✗	✗	✗	✗	✗	✗	✗	1478	2956	1940	77	✗	1067	191	931	69	64	2152	83	21	46	47	43
13	✗	✗	✗	✗	✗	5963	✗	✗	1229	3212	1510	83	✗	1237	103	572	58	54	2040	157	21	53	54	50
14	✗	✗	✗	✗	✗	5388	✗	1674	1069	3031	1133	76	✗	985	22	58	57	54	1653	21	21	40	40	47
15	✗	✗	✗	✗	✗	5115	✗	1678	1038	2299	871	77	✗	746	22	60	60	55	1502	87	21	35	36	36
20	✗	✗	✗	✗	6382	5014	✗	1296	826	2732	978	81	2306	825	23	52	51	47	1472	156	22	48	49	45
25	✗	✗	2450	5176	5702	5124	2942	1298	1008	2537	667	81	1422	693	24	55	54	50	1196	92	24	46	47	41
30	✗	✗	1782	4342	4522	4310	2159	1259	774	2585	592	77	1383	579	27	59	59	54	1204	112	25	44	45	42
35	✗	4894	1873	4545	4762	4779	2297	1324	866	2236	712	89	1342	672	29	64	63	58	1330	27	27	47	49	45
40	✗	2462	1964	4118	4881	4641	2178	1465	609	2544	209	90	1591	160	31	83	82	76	1138	29	29	48	52	47
45	✗	2866	2099	4958	4937	4401	2286	1520	838	2090	316	96	1476	548	33	61	60	55	1166	31	31	56	62	56
50	✗	2351	2220	5768	5814	4529	2414	1356	730	2673	342	98	1707	506	33	59	58	53	1346	31	31	39	40	35

Table F.2
Maximum computational runtime (seconds) of the demand sets for 1TM configurations (runtime is only given for scenarios with 100% ASL).

	$\mathcal{W}$ = 4 hours						$\mathcal{W}$ = 8 hours						$\mathcal{W}$ = 12 hours						$\mathcal{W}$ = 16 hours					
SND	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5	FW_1	FW_2	FW_3	AW_3	AW_4	AW_5
1–2	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
3	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	5159
4	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	4529	✗	✗	✗	✗	✗	292
5	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	273	✗	✗	3124	3512	4763	114
6	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	5640	✗	✗	✗	✗	7096	116	✗	✗	1388	5479	2337	132
7	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	434	✗	✗	1956	5791	4107	118	✗	3154	1204	2236	130	114
8	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	1601	✗	✗	1630	4681	3951	85	✗	1847	1099	2199	73	67
9	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	5220	173	✗	3000	1361	2559	594	83	✗	1543	23	88	89	83
10	✗	✗	✗	✗	✗	✗	✗	✗	2347	4809	3240	87	✗	1815	740	67	67	61	✗	776	22	107	107	99
11	✗	✗	✗	✗	✗	✗	✗	✗	2147	4244	4414	107	✗	1432	23	2749	86	81	✗	756	22	85	85	80
12	✗	✗	✗	✗	✗	✗	✗	✗	2089	3876	3812	112	✗	1950	921	2701	86	80	3975	643	22	64	64	59
13	✗	✗	✗	✗	✗	7097	✗	✗	2154	5421	2563	122	✗	1448	832	1879	88	81	2809	743	22	83	86	79
14	✗	✗	✗	✗	✗	7828	✗	2648	2296	5309	3065	110	✗	1572	23	87	86	81	3096	22	22	43	44	74
15	✗	✗	✗	✗	✗	7113	✗	2419	2249	3146	2554	87	✗	1014	23	89	88	82	2342	680	22	43	43	61
20	✗	✗	✗	✗	10 037	6261	✗	2129	1613	5838	3021	126	3996	1143	24	69	67	62	2338	705	22	67	68	62
25	✗	✗	5968	6152	6916	7496	3598	1853	1673	4214	1231	118	2530	1156	25	73	72	67	1750	716	24	69	71	44
30	✗	✗	2316	6361	5311	7283	3498	1719	1679	3729	1369	99	2339	1365	28	103	102	94	1925	902	25	50	51	48
35	✗	8009	2257	7320	5664	7286	3293	1981	1205	6388	1900	137	2212	1422	31	114	110	102	2012	28	28	53	56	51
40	✗	2872	3083	5810	5466	6508	3135	1955	1469	5358	931	114	2121	500	32	149	146	136	1457	30	30	56	60	55
45	✗	3783	3670	7445	6094	6445	3051	2741	1690	3694	896	123	2284	1311	35	98	96	87	1391	32	32	83	91	84
50	✗	3285	2902	7689	7273	6864	3247	1858	1498	4287	1044	124	2324	1585	35	100	98	90	1791	32	32	60	63	59

Table F.3
Average computational runtime (seconds) of the demand sets for 1T1 configurations (runtime is only given for scenarios with 100% ASL).

	$\mathcal{W}$ = 4 hours						$\mathcal{W}$ = 8 hours						$\mathcal{W}$ = 12 hours						$\mathcal{W}$ = 16 hours					
SND	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4
1–6	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
7	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	1879	✗	2609	3549
8	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	1505	1330	1813	1994	2881
9	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	928	983	1550	1308	1648
10	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	1278	1385	2349	2873	3595	✗	448	299	312	177	166
11	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	1149	1157	2163	2543	3212	✗	472	421	286	268	298
12	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	1035	1221	2133	2656	3432	2151	492	465	197	201	227
13	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	1170	1226	2443	2888	3435	2007	564	491	318	218	224
14	✗	✗	✗	✗	✗	✗	✗	✗	1442	✗	3060	4087	✗	1106	1229	1945	2287	2786	1569	323	401	142	122	151
15	✗	✗	✗	✗	✗	✗	✗	1835	1727	2578	3009	3963	✗	1056	1013	2016	1831	2340	1475	186	317	109	93	73
20	✗	✗	✗	✗	✗	✗	✗	1915	2033	2920	3703	5346	2399	868	745	1668	1407	1592	1460	529	385	78	62	69
25	✗	✗	✗	✗	✗	✗	2956	1127	1098	2533	2589	3177	1341	767	700	1300	1347	1605	1137	213	165	184	76	80
30	✗	✗	✗	✗	✗	✗	2213	1169	1101	2353	2222	3153	1296	926	794	1633	1112	1358	1215	342	267	64	54	56
35	✗	3357	2463	4529	4276	5479	2258	1328	1252	2590	2498	3277	1327	971	758	1413	1433	1576	1342	421	165	106	132	165
40	✗	1743	1695	2589	3326	3501	2120	1284	1492	2533	2684	3519	1460	1087	953	1081	954	1211	1081	201	192	60	59	62
45	✗	2025	1486	2638	3226	3622	2237	1337	1404	2549	2979	3677	1428	1111	987	1061	921	986	1282	398	128	67	59	58
50	✗	2517	1795	2618	3086	3562	2328	1148	882	2247	2246	2604	1560	1008	957	1221	968	998	1412	33	33	55	43	41

Table F.4
Maximum computational runtime (seconds) of the demand sets for 1T1 configurations (runtime is only given for scenarios with 100% ASL).

	$\mathcal{W}$ = 4 hours						$\mathcal{W}$ = 8 hours						$\mathcal{W}$ = 12 hours						$\mathcal{W}$ = 16 hours					
SND	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4	FW_1	FW_2	FW_3	AW_2	AW_3	AW_4
1–6	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
7	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	2454	✗	3813	5178
8	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	2345	1912	3215	3737	4914
9	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	1593	1418	2233	2484	2940
10	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	1779	2458	3356	5424	4710	✗	817	929	1088	727	542
11	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	1702	1591	2729	3800	4681	✗	1040	1089	1036	1126	1451
12	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	1231	1859	2790	3266	4452	3726	1126	1081	875	722	1015
13	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	1547	1978	3097	4274	4561	2726	1234	1103	1160	919	900
14	✗	✗	✗	✗	✗	✗	✗	✗	1892	✗	4919	6638	✗	1463	2005	3008	3593	4034	2911	858	857	715	635	935
15	✗	✗	✗	✗	✗	✗	✗	3209	2405	3438	3589	5258	✗	2012	1356	2859	2972	3189	2355	868	759	371	382	158
20	✗	✗	✗	✗	✗	✗	✗	2811	3288	3548	4708	6919	4707	1528	1470	2475	2918	2992	2340	1213	1081	192	229	295
25	✗	✗	✗	✗	✗	✗	3415	1981	1957	3580	3786	4435	2311	1200	1142	2394	2449	3153	1599	708	782	1196	194	259
30	✗	✗	✗	✗	✗	✗	3611	1505	1705	2829	2816	3801	2150	1919	1451	2778	1734	2306	1951	1045	850	140	102	125
35	✗	5088	4291	6465	5421	8087	2945	1683	1738	3493	3202	4289	2205	1458	1206	2369	2557	2262	2014	1038	767	400	715	997
40	✗	3050	2535	3795	5333	4949	3086	1560	2153	3349	3742	4379	1937	1540	1371	2039	2294	2933	1450	925	868	144	141	167
45	✗	3205	2416	3367	5119	5095	2800	1979	2552	3288	3580	4479	2228	1538	1292	2436	2208	3064	1614	981	1023	176	88	87
50	✗	4305	3133	5946	3680	5741	3366	1891	1386	3463	2743	4300	2033	1589	1434	2489	2350	2174	1859	34	34	134	63	59

References

Agnimo, V., Ouhimmou, M., Paquet, M., Montecinos, J., 2023. Integrated strategic and tactical design of multi-echelon city distribution systems with vehicles synchronization: A case of the greater montréal area. Comput. Ind. Eng. 183, 109458. http://dx.doi.org/10.1016/j.cie.2023.109458.

Anderluh, A., Hemmelmayr, V.C., Nolz, P.C., 2017. Synchronizing vans and cargo bikes in a city distribution network. Central Eur. J. Oper. Res. 25 (2), 345–376. http://dx.doi.org/10.1007/s10100-016-0441-z.

Anderluh, A., Nolz, P.C., Hemmelmayr, V.C., Crainic, T.G., 2021. Multi-objective optimization of a two-echelon vehicle routing problem with vehicle synchronization and 'grey zone' customers arising in urban logistics. European J. Oper. Res. 289 (3), 940–958. http://dx.doi.org/10.1016/j.ejor.2019.07.049.

Barratt, M., 2004. Understanding the meaning of collaboration in the supply chain. Supply Chain Manag. 9 (1), 30–42. http://dx.doi.org/10.1108/13598540410517566.

Bijvoet, B., 2025. Amsterdam multimodal two-echelon vehicle routing case study: Canal-road network, HoReCa demand, and satellite location scenarios. http://dx.doi.org/10.17632/skyry266wz.2.

Browne, M., Allen, J., Nemoto, T., Patier, D., Visser, J., 2012. Reducing Social and Environmental Impacts of Urban Freight Transport: A Review of Some Major Cities. Procedia-Soc. Behav. Sci. 39, 19–33. http://dx.doi.org/10.1016/j.sbspro.2012.03.088.

Cattaruzza, D., Absi, N., Feillet, D., González-Feliu, J., 2017. Vehicle routing problems for city logistics. EURO J. Transp. Logist. 6 (1), 51–79. http://dx.doi.org/10.1007/s13676-014-0074-0.

Cleophas, C., Cottrill, C., Ehmke, J.F., Tierney, K., 2019. Collaborative urban transportation: Recent advances in theory and practice. European J. Oper. Res. 273 (3), 801–816. http://dx.doi.org/10.1016/j.ejor.2018.04.037.

Cruijssen, F., Dullaert, W., Fleuren, H., 2007. Horizontal cooperation in transport and logistics: A literature review. Transp. J. 46 (3), 22–39. http://dx.doi.org/10.2307/20713677.

Cuda, R., Guastaroba, G., Speranza, M.G., 2015. A survey on two-echelon routing problems. Comput. Oper. Res. 55, 185–199. http://dx.doi.org/10.1016/j.cor.2014.06.008.

Dijkstra, E.W., 1959. A note on two problems in connexion with graphs. Numer. Math. 1, 269–271. http://dx.doi.org/10.1007/BF01386390.

Dumez, D., Tilk, C., Irnich, S., Lehuédé, F., Olkis, K., Péton, O., 2023. A matheuristic for a 2-echelon vehicle routing problem with capacitated satellites and reverse flows. European J. Oper. Res. 305 (1), 64–84. http://dx.doi.org/10.1016/j.ejor.2022.05.022.

European Commission, 2017. European urban mobility – Policy context. Publications Office, Directorate-General for Mobility and Transport. https://data.europa.eu/doi/10.2832/827766.

Gansterer, M., Hartl, R.F., 2018. Collaborative vehicle routing: A survey. European J. Oper. Res. 268 (1), 1–12. http://dx.doi.org/10.1016/j.ejor.2017.10.023.

Gemeente Amsterdam, 2019. Programma smart mobility 2019–2025. https://openresearch.amsterdam/nl/page/48272/programma-smart-mobility-2019-2025.

Gemeente Amsterdam, 2020a. Amsterdam maakt ruimte: Agenda amsterdam autoluw. https://openresearch.amsterdam/nl/page/49784/amsterdam-autoluw—amsterdam-maakt-ruimte.

Gemeente Amsterdam, 2020b. Marktstrategie programma bruggen en kademuren. https://assets.amsterdam.nl/publish/pages/958051/marktstrategie_programma_bruggen_en_kademuren_jun2020.pdf.

Gemeente Amsterdam, 2020c. Nota varen | deel 2: Ligplaatsen en op- en afstappen passagiersvaart, verduurzaming, vignetbeleid en stimuleren van (goederen)transport over water. https://assets.amsterdam.nl/publish/pages/1018385/nota_varen_deel_2_def.pdf.

Gemeente Amsterdam, 2020d. Programmaplan bruggen en kademuren. https://openresearch.amsterdam/nl/page/56187/programmaplan-bruggen-en-kademuren.

Gemeente Amsterdam, 2021a. Hubsvisie amsterdam. https://openresearch.amsterdam/nl/page/100040/hubvisie-amsterdam.

Gemeente Amsterdam, 2021b. Omgevingsvisie amsterdam 2050: Een menselijke metropool. https://openresearch.amsterdam/nl/page/72317/omgevingsvisie-amsterdam-2050-versie-2021.

Gemeente Amsterdam, 2022a. Gefundeerd herstellen: Actieplan programma bruggen en kademuren 2023–2026. https://openresearch.amsterdam/nl/page/91392/actieplan-programma-bruggen-en-kademuren-2023-2026.

Gemeente Amsterdam, 2022b. Logistieke strategie: Samen werken aan balans tussen bevoorrading en leefbaarheid. https://openresearch.amsterdam/nl/page/81240/logistieke-strategie.

Grangier, P., Gendreau, M., Lehuédé, F., Rousseau, L.-M., 2016. An adaptive large neighborhood search for the two-echelon multiple-trip vehicle routing problem with satellite synchronization. European J. Oper. Res. 254 (1), 80–91. http://dx.doi.org/10.1016/j.ejor.2016.03.040.

Gurobi Optimization, L., 2022. Gurobi optimizer reference manual, 2023.

Karademir, C., Beirigo, B.A., Atasoy, B., 2025. A two-echelon multi-trip vehicle routing problem with synchronization for an integrated water-and land-based transportation system. European J. Oper. Res. 322 (2), 480–499. http://dx.doi.org/10.1016/j.ejor.2024.10.047.

Lehmann, J., Winkenbach, M., 2024. A matheuristic for the two-echelon multi-trip vehicle routing problem with mixed pickup and delivery demand and time windows. Transp. Res. C: Emerg. Technol. 160, 104522. http://dx.doi.org/10.1016/j.trc.2024.104522.

Li, H., Liu, Y., Chen, K., Lin, Q., 2020. The two-echelon city logistics system with on-street satellites. Comput. Ind. Eng. 139, 105577. http://dx.doi.org/10.1016/j.cie.2018.12.024.

Li, H., Liu, Y., Jian, X., Lu, Y., 2018. The two-echelon distribution system considering the real-time transshipment capacity varying. Transp. Res. B: Methodol. 110, 239–260. http://dx.doi.org/10.1016/j.trb.2018.02.015.

Li, H., Wang, H., Chen, J., Bai, M., 2021. Two-echelon vehicle routing problem with satellite bi-synchronization. European J. Oper. Res. 288 (3), 775–793. http://dx.doi.org/10.1016/j.ejor.2020.06.019.

Marín, A., Pelegrín, M., 2019. p-median problems. In: Laporte, G., Nickel, S., Saldanha da Gama, F. (Eds.), Location Science. Springer International Publishing, Cham, pp. 25–50. http://dx.doi.org/10.1007/978-3-030-32177-2_2.

Marques, G., Sadykov, R., Dupas, R., Deschamps, J.-C., 2022. A branch-cut-and-price approach for the single-trip and multi-trip two-echelon vehicle routing problem with time windows. Transp. Sci. 56 (6), 1598–1617. http://dx.doi.org/10.1287/trsc.2022.1136.

Masson, R., Lehuédé, F., Péton, O., 2013. Efficient feasibility testing for request insertion in the pickup and delivery problem with transfers. Oper. Res. Lett. 41 (3), 211–215. http://dx.doi.org/10.1287/trsc.1110.0400.

Perboli, G., Tadei, R., Vigo, D., 2011. The two-echelon capacitated vehicle routing problem: Models and math-based heuristics. Transp. Sci. 45 (3), 364–380. http://dx.doi.org/10.1287/trsc.1110.0368.

Savelsbergh, M., Van Woensel, T., 2016. 50Th anniversary invited article—City logistics: Challenges and opportunities. Transp. Sci. 50 (2), 579–590. http://dx.doi.org/10.1287/trsc.2016.0675.

Sluijk, N., Florio, A.M., Kinable, J., Dellaert, N., Van Woensel, T., 2023. Two-echelon vehicle routing problems: A literature review. European J. Oper. Res. 304 (3), 865–886. http://dx.doi.org/10.1016/j.ejor.2022.02.022.

Soares, R., Marques, A., Amorim, P., Parragh, S.N., 2024. Synchronisation in vehicle routing: classification schema, modelling framework and literature review. European J. Oper. Res. 313 (3), 817–840. http://dx.doi.org/10.1016/j.ejor.2023.04.007.

United Nations, 2019. World urbanization prospects: The 2018 revision (ST/ESA/SER.A/420). Department of Economic and Social Affairs, Population Division. https://www.un-ilibrary.org/content/books/9789210043144.

Zamal, M.A., Schrotenboer, A.H., Van Woensel, T., 2025. The two-echelon vehicle routing problem with pickups, deliveries, and deadlines. Comput. Oper. Res. 179, 107016. http://dx.doi.org/10.1016/j.cor.2025.107016.

Zhou, H., Qin, H., Zhang, Z., Li, J., 2022. Two-echelon vehicle routing problem with time windows and simultaneous pickup and delivery. Soft Comput. 26 (7), 3345–3360. http://dx.doi.org/10.1007/s00500-021-06712-2.