A two-echelon multi-trip vehicle routing problem with synchronization for an integrated water- and land-based transportation system

Karademir, Cigdem; Beirigo, Breno A.; Atasoy, Bilge

DOI
10.1016/j.ejor.2024.10.047

Publication date
2025

Document Version
Final published version

Published in
European Journal of Operational Research

European Journal of Operational Research 322 (2025) 480–499

Contents lists available at ScienceDirect

European Journal of Operational Research

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

A two-echelon multi-trip vehicle routing problem with synchronization for an integrated water- and land-based transportation system

Cigdem Karademir [a,*], Breno A. Beirigo [b], Bilge Atasoy [a]

[a] Dept. of Maritime & Transport Technology, Delft University of Technology, Delft, The Netherlands

[b] Dept. of Industrial Engineering & Business Information Systems, University of Twente, Enschede, The Netherlands

ARTICLE INFO

Keywords:
Two-echelon network
Vehicle routing problems
City logistics
Satellite synchronization
Logic-based Benders' decomposition

ABSTRACT

This study focuses on two-echelon synchronized logistics problems in the context of integrated water- and land-based transportation (IWLT) systems. The aim is to meet the increasing demand in city logistics as a result of the growth in transport activities, including parcel delivery, food delivery, and waste collection. We propose two models, a novel mixed integer linear joint model, and a logic-based Benders' decomposition (LBBD) model, for a two-echelon problem under realistic settings such as multi-trips, time windows, and synchronization at the satellites with no storage and limited resource capacities. The objective is to optimize transfers and satellite assignments, thereby reducing overall logistics costs for street vehicles and vessels. Computational experiments demonstrate that the LBBD model is more robust in terms of solution quality and solution time on average while the added value of the LBBD is more evident when solving large-scale instances with 100 customers, reducing the overall costs by 10.6% on average and significantly reducing the fleet costs on both networks. Furthermore, we assess the effect of changing cost parameters and satellite locations in the proposed IWLT system–analyzing system behavior and suggesting potential improvements–and evaluate several system alternatives in city logistics–consisting of different transportation network designs (single- and two-echelon), vehicle types, and operational constraints. On average, the proposed two-echelon IWLT system reduces the number of kilometers traveled by vehicles at street level by ranging from 20% to 30% compared to a typical single-echelon service design that relies solely on trucks.

1. Introduction

Freight activities in metropolitan areas have been increasing as a result of the growth in the need for parcel delivery, food delivery, and waste collection (Chevalier, 2021). Due to the greater preference for road infrastructure over more environmentally friendly options, the growth in logistics activities has been escalating the burden on the roads (Pfoser, 2022). The increased logistics movements and the increased number of trucks affect the quality of life in cities by contributing to congestion, emissions, and damage to the infrastructure. Logistics service providers (LSPs) are facing challenges to reduce congestion-related costs such as service delays, customer inconveniences, and traffic idling times. On the other hand, the authorities are looking for solutions to achieve emission-free cities by 2030 (EU, 2021). Nevertheless, overarching initiatives toward more sustainable and livable cities have not significantly contributed to a modal shift.

LSPs are increasingly exploring the implementation of innovative technologies like electric vehicles, autonomous vehicles, unmanned vessels, and drones in their logistics systems, to cut costs. These technologies are still limited in terms of storage space, driving range, or reliability, which limits their suitability to take up transport operations completely. Nevertheless, they can be combined with larger vehicles to supply capacity replenishment (Yu, Puchinger, & Sun, 2020). However, the economic benefits of such systems are still not very clear to the LSPs (Moolenburgh, Van Duin, Balm, Van Altenburg, & Van Amstel, 2020).

According to Sluijk, Florio, Kinable, Dellaert, and Van Woensel (2022), consolidating cargoes outside of cities via larger vehicles and coordinating them with smaller city freighters at urban transshipment facilities (satellites) can enhance efficiency in the logistics system. Adding another layer to the distribution system can lead us to economies of scale. They highlight the growing interest in such two-tier or two-echelon logistics systems in both academic and commercial applications. Crainic, Ricciardi, and Storchi (2009) introduce the city logistics concept to move toward integrated freight systems, particularly using two-echelon systems to meet the increasing demand in cities. To achieve this, they emphasize the importance of synchronization and coordination between fleets on different echelons.

* Corresponding author.
E-mail addresses: c.karademir@tudelft.nl (C. Karademir), b.alvesbeirigo@utwente.nl (B.A. Beirigo), b.atasoy@tudelft.nl (B. Atasoy).

https://doi.org/10.1016/j.ejor.2024.10.047
Received 11 July 2023; Accepted 31 October 2024
Available online 6 November 2024

The above-mentioned technological developments and inefficiencies in current road transportation require us to explore alternatives. Groothedde, Ruijgrok, and Tavasszy (2005) discuss the efficiency and reliability of intermodal systems for city logistics problems. They conclude that economies of scale can be achieved by advancing service network design methods to include coordination and synchronization costs in real settings. Mostly, the cost of transshipment operations in intermodal systems or two-echelon systems is overlooked in the literature by simplifying the transshipment capacities of satellites regarding the equipment, employee, and space at a time. These simplifications ignore possible delays and related costs (Côté, Guastaroba, & Speranza, 2017). For example, if multiple transfer requests overlap in time, especially when the satellite's resources are limited, delays can occur as the satellite needs to allocate its resources to handle each transfer. This can lead to queuing or prioritization issues, causing delays for specific transfers, which are not taken into account while deciding them.

To address the issues in city logistics, we study an integrated water- and land-based transportation (IWLT) system which aims at achieving a higher level of modal shift to take advantage of the growing worldwide applications over waterways (Janjevic & Ndiaye, 2014). In this system, light electric freight vehicles (LEFVs) serve the demand in cities, while vessels act as mobile depots whenever capacity replenishment is needed. Satellites are considered to have the capacity to tranship from a single vehicle to another vehicle at a time, providing one-to-one transfers between vessels operating over inland waterways and LEFVs operating as city freighters on streets. We model such a system as a two-echelon multi-trip vehicle routing problem with satellite synchronization (2E-MVRP-SS) considering unitary transshipment real-time capacities at the satellites with no storage and time windows at the customers. Unitary transshipment ensures that the vehicles at the satellite are unloaded and loaded one by one for an average transshipment duration, allowing a non-overlapping operations' sequence to eliminate congestion

The purpose of the proposed system is twofold: (i) alleviating congestion by reducing the burden on street vehicles with the integration of inland waterways and (ii) maximizing the utilization of new vehicle technologies to improve city logistics. The main contributions of our work are listed as follows.

- We provide a two-index compact formulation for a synchronized two-echelon system, 2E-MVRP-SS, with unitary transshipment capacities. To the best of our knowledge, the transshipment capacity of satellites, limited by both space and resources, is not addressed yet for such problems.
- We propose a logic-based Benders decomposition (LBBD) approach for the 2E-MVRP-SS to tackle the complexity of large-scale problems and show its superiority in terms of quality and solution time.
- We show how to adopt the proposed model under different scenarios regarding the service network design and operational costs. Furthermore, we provide managerial insights about the benefits and challenges of synchronized IWLT systems compared to the on-street alternatives.

The remainder of the paper is organized as follows. Section 2 provides a literature overview related to synchronized two-echelon transportation problems, focusing on important aspects. Then, in Section 3, we formulate 2E-MVRP-SS as a mixed integer linear programming (MILP) and discuss its applicability to different variants. Section 4 introduces an LBBD framework for solving 2E-MVRP-SS. In Section 5, we evaluate both the MILP formulation and the LBBD method, providing managerial insights for IWLT systems. Finally, Section 6 is devoted to the conclusions and further directions.

2. Literature review

Following the introduction of the two-echelon capacitated vehicle routing problem (2E-VRP) by Gonzalez-Feliu (2008), studies have demonstrated that two-echelon routing problems have become more prominent due to the increased freight movements in cities. Thus, many authors have researched two-echelon distribution systems tailored to city logistics to reduce the negative impacts of increased on-street movements on society, economy, and environment (Anderluh, Nolz, Hemmelmayr, & Crainic, 2021). These studies differ in terms of the synchronization degree between the vehicles using common resources during transshipment operations.

In single-echelon systems, vehicles transport goods directly from origin to destination without interacting with other vehicles. However, in two-echelon systems, services involve a combination of vehicles, creating a dependence between their operations for cargo-flow connectivity. This introduces complex decisions regarding transshipment synchronization at satellites. As a result, changing the route of one vehicle in this system can make other routes infeasible. The interdependence problem, as referred to by Drexl (2012), adds complexity compared to conventional solution methods. Resource synchronization at satellites significantly impacts operations and decisions at different echelons, influencing the degree of interdependence. Drexl (2012) defines the requirement such that the total utilization or consumption of a particular resource by all vehicles should not exceed a set limit at any given time.

Integrated vehicle routing problems are commonly used in the literature to describe two-echelon transportation systems. These problems arise when vehicle routing problems result from another optimization problem. According to Côté et al. (2017), solving "strongly interdependent problems" as integrated problems, considering joint decisions' feasibility and cost relations, brings benefits despite increased complexity. Integrated modeling bridges the gap between academia and the real world by reducing assumptions and unexpected costs resulting from simplifications.

Integrating inland waterways into urban freight transport offers a viable alternative to address congestion, environmental impact, and limited space challenges (Janjevic & Ndiaye, 2014). Besides transporting bulk materials for construction, there exist several applications for last-mile parcel and retail logistics using inland waterways, e.g., floating barges in combination with electric cargo bikes and LEFVs in Sweden, autonomous vessels with electric bikes in Germany, vessels loaded with electric cubicycles in Belgium, vessels with rolling containers in the Netherlands, ships and diesel trucks in France (Brauner, Kayle, & Pauwels, 2021). Recent studies have reflected this trend by focusing on route optimization and cost evaluation for new last-mile delivery systems, in which traditional delivery methods are replaced by alternatives such as electric vehicles (EVs) or cargo bikes (Divieso, Lima, & De Oliveira, 2021). However, most of the studies focus on case-specific operations at the last mile and ignore expensive transshipment operations in cost calculations. Thus, the economic gains of such integrated systems are not clear to all stakeholders. Accordingly, He and Haasis (2019) highlight the scarcity of research on the utilization of electric vehicles (EVs) requiring transshipment operations in integrated distribution systems. Caris, Limbourg, Macharis, Van Lier, and Cools (2014) emphasize the need for system-wide modeling to evaluate various design options for all stakeholders to determine risks and establish operational schemes to guide policies for the public and private sectors. With advancements in autonomous vehicles and the increased use of waterborne freight transport in cities, re-framing urban logistics problems to account for such a system-wide perspective is essential for assessing the required infrastructural investments and the extent of economic benefits. In this study, we propose a framework to provide managerial insights for the novel integrated systems to improve city logistics by coordination and synchronization compared to traditional logistics.

In order to assess the economic benefits of two-echelon systems, we present a comprehensive overview of the existing studies on synchronized systems, which are characterized by the transshipment capacity of satellites. Additionally, we focus on studies that aim to enhance the utilization of new technologies with limited capacities by exploring the concept of multi-trips. Interaction between vehicles due to their multiple use and limitations on these interactions due to the transshipment capacities increase the complexity of the decisions. To tackle complexity issues, we briefly summarize the LBBD approach, highlighting its effectiveness in integrating existing knowledge to solve integrated problems.

2.1. Unlimited transshipment capacity

Most studies have focused on the basic variant, 2E-VRP, where synchronization is required only for the flow of the items (Cattaruzza, Absi, Feillet, & González-Feliu, 2017). They only respect the capacities of the vehicles to supply the assigned flows, but ignore the satellites' capacities. All the goods are brought to the satellites without any time dependence. The satellites have unlimited resources and storage to process and store the freight at the satellite until the city freighters arrive.

For 2E-VRP, Jepsen, Spoorendonk, and Ropke (2013), Marques, Sadykov, Deschamps, and Dupas (2020) and Santos, da Cunha, and Mateus (2013) and more recently (Mhamedi, Andersson, Cherkesly, & Desaulniers, 2021) propose branch-and-cut or branch-and-price algorithms. Baldacci, Mingozzi, Roberti, and Calvo (2013) develop a bounding procedure combining dynamic programming using a decomposition approach to divide the problem into multi-depot capacitated VRPs. Marques et al. (2020) propose a branch and cut algorithm that first enumerates all solutions for supplying the satellites before optimizing city freighters' routes for each of the enumerated solutions. It outperforms the existing exact algorithms and solves problems with 200 customers and 10 satellites, indicating the potential of such a decomposition-based approach.

Time windows force the system to be semi-synchronized in time, only allowing departures of city freighters after the delivery at the satellites. It is also referred to as the basic variant with time dependence. Dellaert, Dashty Saridarq, Van Woensel, and Crainic (2019) propose a branch-and-price algorithm for a 2E-VRP with time windows and satellite synchronization (2E-VRP-SS), which can solve the problems with 100 nodes and 5 satellites to optimality. More recently, Dellaert, Van Woensel, Crainic, and Saridarq (2021) developed a decomposition-based exact solution approach for the 2E-VRP-SS. An adaptive large neighborhood is proposed by Li, Wang, Chen and Bai (2021) for 2E-VRP-SS with satellite bi-synchronization, that can solve instances with 4080 nodes and 34 satellites. The 2E-VRP with load synchronization is adequately studied in the literature with various exact and heuristic approaches and we refer to the recent survey by Sluijk et al. (2022) for more details.

2.2. Limited transshipment capacity

For city logistics due to the limited infrastructure, generally there exist dedicated spaces with limited storage, or public spaces such as parking lots or public transportation stops with no storage option. Moreover, most studies addressing synchronization in two-echelon settings assume that multiple transshipments can be performed, ignoring the synchronization of resources. Resource synchronization ensures the output rate of a satellite should not exceed its capacity in terms of total transshipped goods given the employee hours, equipment capacities and availability of the satellites at any time. In any of these cases, the vehicles operating different networks require semi or exact synchronization in space and time in addition to cargo flow synchronization (Li, Chen, Wang and Bai, 2021).

Li, Liu, Jian, and Lu (2018) consider a two-echelon distribution system with maximal transshipment capacity at satellites at any time to serve dedicated customers to the satellites. Capacity is defined as the maximum quantity of goods stored and processed at a time. They provide a non-tractable MILP formulation that becomes exhaustive to solve problems with 10 demand nodes within 4 h and solve large-scale problems by a large neighborhood search (LNS). A similar problem with maximal transshipment capacity is introduced for simultaneous pickup and delivery problems by Dumez et al. (2023). They provide a MILP formulation but it also becomes very expensive in terms of memory when the size of the demand nodes increases from 10 to 20. They show that doubling the satellite transshipment capacities provides more savings than doubling the number of satellites, indicating the effect of the resource synchronization is significant in terms of economic gains considering transfer operations. Escobar-Vargas, Crainic, and Contardo (2021) study the synchronization in multi-attribute two-echelon distribution systems with limited capacities, allowing storage for a limited duration. They propose a compact formulation by three-index vehicle flows and a time-space formulation that can solve problems with up to 10 customer nodes. They further integrate a dynamic discretization method to provide feasible solutions up to 50 nodes, suggesting the efficient use of the compact formulation for large-scale instances.

2.3. Multiple use of vehicles under no storage

There exists a limited number of studies for synchronized two-echelon settings considering the multi-trip nature of practical applications like those involving drones, bicycles, or LEFVs under limited satellite capacities. Crainic et al. (2009) introduce first general models and formulations for a 2E-MVRP-SS with time windows, multiple-depot, and heterogeneous vehicles. They stay at the conceptual phase by providing tactical and strategic level analysis for designing and solving such complex systems.

Grangier, Gendreau, Lehuédé, and Rousseau (2016) focus on a 2E-MVRP-SS with time windows and no storage, requiring a high degree of temporal and spatial synchronization. They assume that the satellites have the resources to operate an unlimited number of transshipments at any time. They suggest incorporating the process times of the transfers into travel times to and from satellites. However, in case of limited resources, it is not possible to know in advance how much time is needed to process given transfers. This simplification ignores the queuing problem at the satellites and underestimates the impact of lead times as well as the number of vehicles. They propose an intractable MILP with a three-index formulation and use an adaptive large neighborhood search (ALNS) to test Solomon's (1987) instances with 100 nodes.

Anderluh, Hemmelmayr, and Nolz (2017) focus on a 2E-MVRP-SS with no storage at the satellites to serve the customers assigned to the vans or cargo bikes. Using a greedy randomized adaptive search procedure (GRASP), they assess the impact of using bikes in combination with vans instead of using only vans for a real-life application of pharmacy wholesale and distributors of vegetable boxes in Vienna. In their Anderluh et al. (2021) study, they assess the effect of "gray zone" customers that can be served by direct and indirect shipments at any echelon to further improve the economic benefits of using lighter vehicles in city logistics.

He and Li (2019) consider a 2E-MVRP-SS with dynamic satellites with no storage for a harvesting scheduling. The transshipments take place at the customer nodes by allowing vehicles to wait up to a maximum duration. This assumption simplifies the location problem and the cost of transshipment operations at customer sites since ensuring the availability of such spaces for city logistics is neither easy nor cheap. A memetic algorithm with a local search procedure is used to solve instances of up to 200 farmlands and 6 harvesters. They show that full synchronization increases the complexity but not necessarily the cost of the system for a given fleet of vehicles. However, pre-defined discretization of time, proposed as time windows for the transshipments, reduces the utilization of the satellite resources.

Our previous work, Karademir, Alves Beirigo, Negenborn, and Atasoy (2022), considers a 2E-MVRP-SS with unitary transshipments, and time windows, and proposes a MILP with a four-index-based formulation. The real-time capacity is defined only as a single transshipment operation between a vessel and a LEFV at a time without having option to store any good in between arrivals and departures of the vehicles. However, it also faces difficulties in solving problems with 10 customers and 4 satellites. This is due to the exponential number of choices available in two-echelon systems regarding the allocation of customers to the vehicles and to the satellites, and finally synchronizing the schedules of these vehicles at the allocated satellites. In this study, we enhance the proposed MILP using a new compact two-index formulation that reduces the number of binary variables to address the memory issues in solving problems up to 100 customer nodes with limited capacitated satellites to perform unitary transshipment operations.

2.4. Logic-based Benders' decomposition

Real-life applications often rely on upstream or downstream optimization problems. However, these problems are commonly treated separately, with a focus on solving them quickly by simplifying and making assumptions. This approach sacrifices optimal results in favor of reducing decision complexity.

There has been a growing body of literature exploring the application of LBBD to address integrated optimization problems. This approach involves breaking down complex problems into easier-to-solve problems in any form, typically consisting of a master problem for strategic decision-making and corresponding subproblem(s), leveraging existing knowledge in the literature. For instance, Raidl, Baumhauer, and Hu (2014) focus on a bi-level capacitated VRP and implement an LBBD by assigning the demand to the closest satellites first, and minimizing fleets for each satellite. Their proposed decomposition method is enhanced by a variable neighborhood search metaheuristic in order to tackle the scalability of the subproblems at the satellites with larger demand share. Roshanaei and Naderi (2021) re-formulate the integrated operating room planning and scheduling problem by decomposing the cost function to estimate the cost of strategic location decisions. Their proposed MILP outperforms the existing state-of-the-art branch-price-and-cut algorithm. They further show that when combined with a branch-check-and-cut method at every feasible master solution, an LBBD is more robust in terms of solution time and optimality gap compared to solving the master problem to optimality at every iteration. The LBBD method offers significant benefits for problems involving both assignment and task scheduling, especially when tasks cannot overlap due to resource constraints, such as in operation rooms or process planning problems. For example, Karamyar, Sadeghi, and Yazdi (2018) study a stochastic location–allocation and scheduling problem for a healthcare system and propose a simulated annealing method to find feasible solutions for locating new hospitals equipped with new machines at the master problem. Similarly, a multi-trip traveling repairman problem with drones is optimized using an LBBD by focusing on customer locations to launch the drones from a truck (Bruni, Khodaparasti, & Moshref-Javadi, 2022). Martínez, Adulyasak, and Jans (2022) focus on the cost of integrated process configuration decisions and solve related production planning problems. Typically, these models formulate the logic between strategic decisions and optimality using Big-M constraints leading to weak formulations but the LBBD method can exploit the relaxations of the feasible solutions to provide a tighter lower bound or better upper bounds (Rahmaniani, Crainic, Gendreau, & Rei, 2017).

The 2E-MVRP-SS with unitary transshipment capacity studied in this paper aims at jointly solving strongly interdependent problems to reduce the cost of integration. We propose an LBBD method to tackle the complexity of the problem based on a two-index compact formulation for solving large-scale instances. Instead of optimizing the resource allocation at the master problem by locating the satellites in space, we first ensure the feasibility of LEFV schedules to cover all the demand within the requested time windows. A subproblem is solved to locate the transshipment operations of these schedules at various satellite locations, considering resource availability for feasibility and cost evaluation. In other words, the master problem provides the temporal precedence graph of the operations, while the subproblem provides the temporal-spatial graph of the transfer operations for global optimality. Despite the increased complexity at the master level, it reduces the time to find feasible solutions for larger-scale instances.

3. Problem definition and formulation

In this section, we formally present the 2E-MVRP-SS for an IWLT system. Section 3.2 provides a mathematical formulation for only pickup services based on our previous work (Karademir et al., 2022), proposed for the waste collection problem in Amsterdam using an IWLT system. To avoid congestion at the satellites, Section 3.3 is devoted to modeling unitary transshipment constraints to respect satellite capacities, limiting the number of transfers to a maximum of one at a time. Moreover, to show the applicability of the proposed two-index formulation for different variants considering service types, satellite capacities, and charging requirements of the vehicles, we explicitly provide necessary modifications in the formulation for only delivery services in Section 3.4.

Instead of using the terms "first and second echelon" as in the literature, we refer to the two levels of operations in our study as the street level and water level. At the water level, vessels transport cargo between the central depot and the satellites. At the street level, LEFVs transport cargo between the satellites and pickup request locations as city freighters.

3.1. Problem statement

At the street level, there exists a fleet of K_s identical LEFVs with a capacity of Q_s units. All LEFVs are located at a main garage, g, in the city. They start and end their journeys at the garage while visiting a set of customer nodes, C, and one or more satellites in between to transfer the goods. Each customer node i requires q_i units of goods to be picked up by a single LEFV, and associated with a service duration of τ_i within a time window of $[a_i, b_i]$. t_{ij} denotes the shortest travel time between nodes i and j. LEFVs should meet with the vessels for transfer operations at a predefined set of satellites P over multiple times to transfer goods onto a vessel at a time safely. A transfer operation for unloading and loading goods requires U time units.

The water level fleet consists of K_w identical vessels with a capacity of Q_w units. These vessels are located at a central depot, d, that has sufficient space to store them along with the goods collected. The vessels depart from the central depot empty, stop at one or more satellites for transfer tasks, and then return to the central depot loaded. Satellites are public spaces that allow LEFVs and vessels to park while waiting for synchronization. Unlike most of the studies in the literature, there is no storage area where LEFVs can unload cargo before vessels arrive for related transfers. Instead, the synchronized system described in this study enables vehicles to function as temporary and secure storage places while they wait for the vessel at the transshipment location. Moreover, it is assumed that during a transshipment, each LEFV completely unloads its cargo onto a single vessel at the satellite locations.

The 2E-MVRP-SS seeks to minimize overall transportation costs at both levels by (i) routing LEFVs to serve all the city demand while assigning transfer tasks at satellites to them, and (ii) routing vessels to serve these transfer tasks. The transportation cost at each level consists of the fixed cost of the vehicles used (β_s for LEFVs and β_w for vessels) and the variable cost of the total traveled duration by all the vehicles (c^s for LEFVs and c^w for vessels). The fixed cost of using vehicles for

Table 1
Notation for the 2E-MVRP-SS model.

Sets and Indices	
g, d	Garage for LEFVs and central depot for vessels, respectively
C	Customer nodes indexed by i and j
C_s	Customer nodes and the garage g, $C \cup \{g\}$
C_w	Customer nodes and the central depot d, $C \cup \{d\}$
P	Satellites indexed by p
N	All nodes indexed by n, $C \cup \{g\} \cup \{d\} \cup P$
Parameters	
q_i	Demand at node $i \in C$
a_i	Earliest service time of node $i \in N$
b_i	Latest service time of node $i \in N$
τ_i	Service duration of node $i \in C$
U	Constant duration for a transfer task
t_{ij}	Shortest travel time from node $i \in N$ to $j \in N$
c^s/c^w	Cost of traveling a unit of time on the streets/water
Q_s/Q_w	Capacity of a LEFV/vessel
K_s/K_w	Number of available LEFVs/vessels
β_s/β_w	Fixed cost of a LEFV/vessel
M^s_{ij}	Sufficiently large number for constraint linearization, $M^s_{ij} = b_i + \tau_i + t^{max}_{ip} + U + t^{max}_{pj} - a_j$
M^w_{ij}	Sufficiently large number for constraint linearization, $M^w_{ij} = b_i + \tau_i + t^{max}_{ip} + U + t^{max}_{pr} - (a_j + \tau_j + t^{min}_{jp})$
Variables	
x_{ij}	(Binary) 1 if node $j \in C_s$ is visited immediately after node $i \in C_s$ by a LEFV, 0 otherwise
m_i	Total load on the LEFV after visiting node $i \in C$, $q_i \le m_i \le Q_s$
h_i	Service start time at node $i \in C$ with an LEFV, $\max\{a_g + t_{gi},\ a_i\} \le h_i \le \min\{b_g - t_{ig} - s_i - U,\ b_i\}$
ϕ_i	(Binary) 1 if there is a transfer task immediately after node $i \in C$ is served, 0 otherwise
v_{ip}	(Binary) 1 if satellite $p \in P$ is assigned to the transfer task $i \in C$ (if exists), 0 otherwise
y_{ij}	(Binary) 1 if the transfer task $j \in C_w$ is served immediately after the transfer task $i \in C_w$ by a vessel, 0 otherwise
u_i	Service start time of the transfer task $i \in C$ with a vessel and LEFV
l_i	Total load on the vessel after serving the transfer task $i \in C$
f^s_{ij}	Total travel time for an LEFV from node i to node j if it visits $i, j \in C_s$ consecutively
f^w_{ij}	Total travel time for a vessel from the transfer task $i \in C_w$ to the transfer task $j \in C_w$ if it serves tasks i, j consecutively

logistics typically includes expenses such as vehicle purchase or lease, insurance, maintenance, driver's salary, and fuel. Total travel duration is minimized since it is highly correlated to the total distance traveled and total fuel spent.

3.2. A two-index compact formulation for 2E-MVRP-SS with pickups

The decisions to be optimized are the routes for LEFVs, the best times to visit satellites for transfers, the best satellite assignments for the transfers, and optimal routes for vessels to serve the transfer tasks on time at the scheduled satellites. The x_{ij} variable gives the sequence of the pickup operations in which customers are assigned to a LEFV. ϕ_i determines the transshipment operation right after pickup operation at node i to unload collected items to a vessel before serving the next pickup at node j. Similarly, y_{ij} represents the sequence of transshipment operations assigned to a vessel. If there exists a transfer decision provided by $\phi_i = 1$, then v_{ip} decides whether the satellite p is assigned to the transfer task i.

The synchronization is ensured based on transfer task and satellite assignments, meaning that there is a transfer task at satellite p for the LEFV serving customer i and visiting the satellite after the service. v_{ip} allows us to create a time interval for the transfer task ϕ_i regarding the earliest arrival time to satellite p, and the latest time to leave the satellite for the next customer considering its time window since h_i gives the service start time at customer i. For temporal synchronization, the start time of transfer operation i, u_i must fall within this time interval and vehicles at different levels should wait for each other. For spatial synchronization, we adjust f^s_{ij}, the total travel time on the streets between nodes i and j, if a LEFV serves customers i and j consecutively. If there is no transfer decision after customer i (i.e., $\phi_i = 0$) then $f^s_{ij} = t_{ij}$. Otherwise, if $\phi_i = 1$, the LEFV should visit the satellite assigned to the transfer before going to the next customer j. Then, the travel time needs to incorporate that, i.e., $f^s_{ij} = t_{ip} + t_{pj}$, if the transfer is assigned to the satellite p. The same logic is used for deciding the travel time at water level, f^w_{ij}, if a vessel serves the transfer task i and j consecutively. Without taking explicit satellite assignments into account, the two-index routing formulation for water level allows us to reduce the number of binary variables required to represent each copy of a satellite for each customer node, allowing LEFVs to utilize a satellite more than once (Karademir, Beirigo, Negenborn, & Atasoy, 2021). Fig. 1 represents an IWLT network for pickups, multi-trips by LEFVs, and transfers at satellites executed in synchronization by vessels and LEFVs. All sets, parameters, and decision variables are presented in Table 1.

The objective function (1) minimizes the total logistics cost of both levels, $z^s(x, v)$ for the street level and $z^w(x, v, y)$ for the water level, by minimizing the fixed cost of the vehicles used and total traveling cost at both levels.

$$\min \overbrace{\sum_{i \in C} \beta_s x_{gi} + \sum_{i,j \in C_s} c^s f^s_{ij}}^{\text{Street Level Cost: } z^s(x,v)} + \overbrace{\sum_{i \in C} \beta_w y_{di} + \sum_{i,j \in C_w} c^w f^w_{ij}}^{\text{Water Level Cost: } z^w(x,v,y)} \tag{1}$$

subject to

Street Level Routing Problem

$$\sum_{j \in C_s} x_{ij} = \sum_{j \in C_s} x_{ji} = 1 \qquad \forall i \in C \tag{2}$$

$$\sum_{i \in C} x_{gi} = \sum_{i \in C} x_{ig} \le K_s \tag{3}$$

$$\phi_i \ge x_{ig} \qquad \forall i \in C \tag{4}$$

$$m_j - m_i \ge q_j - Q_s(1 - x_{ij} + \phi_i) \qquad \forall i, j \in C, i \ne j \tag{5}$$

$$h_i + \tau_i + f^s_{ij} + U\phi_i \le h_j + M^s_{ij}(1 - x_{ij}) \qquad \forall i \in C, j \in C_s, i \ne j \tag{6}$$

$$f^s_{ij} \ge t_{ij} x_{ij} + \min_{p \in P}\{t_{ip} + t_{pj} - t_{ij}\}(x_{ij} + \phi_i - 1) \qquad \forall i, j \in C_s, i \ne j \tag{7}$$

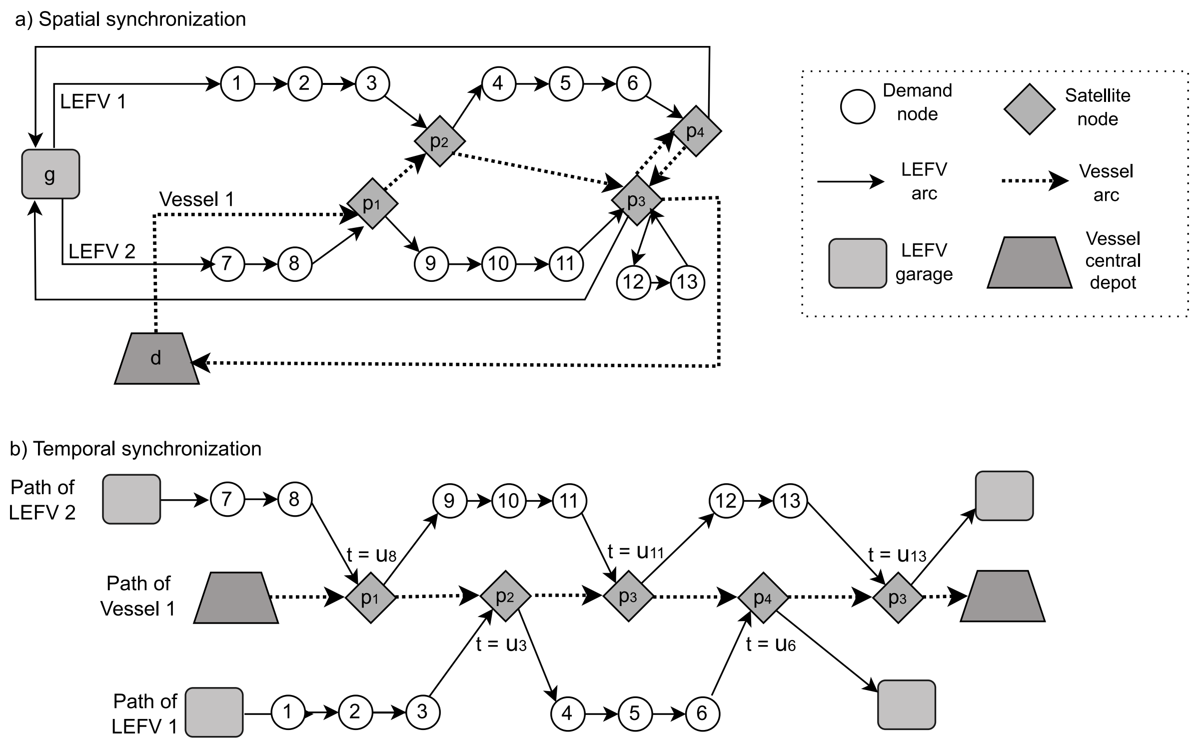


Fig. 1. Example of a small network and a feasible scheduling for the proposed IWLT system, where LEFVs collect goods at the customer level and visit satellites to transship all the goods onto a vessel. A satellite can be visited several times by the same or different vehicles, but at most a vessel and a LEFV perform a transfer operation at a time considering unitary transshipment. This lets vehicles have cycles in space (a) if necessary or efficient while temporal synchronization (b) guarantees that there is no cycle in the temporal-spatial graph of all operations. For a transfer decision after node i, spatial synchronization is ensured by assigning it to a satellite p in reach, v_{ip}. Furthermore, the temporal synchronization is ensured by the transshipment start time on a vessel, u_i, without violating time constraints of upstream or downstream operations of the vehicles interacting. Schedules of LEFVs: $\{x : g \to 1 \to 2 \to 3 \to 4 \to 5 \to 6 \to g,\}, \{x : g \to 7 \to 8 \to 9 \to 10 \to 11 \to 12 \to 13 \to g\}, \{\phi : 8, 3, 11, 6, 13\}$. Schedules of satellites: $\{v : 8 \to p_1,\ 3 \to p_2,\ 11 \to p_3,\ 6 \to p_4,\ 13 \to p_3\}$. Schedule of the vessel: $\{y : d \to 8 \to 3 \to 11 \to 6 \to 13 \to d\}$.

$$x_{ij} \in \{0,1\} \qquad \forall i,j \in C_s, i \neq j \qquad (8)$$

$$\phi_i \in \{0,1\} \qquad \forall i \in C \qquad (9)$$

Synchronization Problem

$$\sum_{p\in P} v_{ip} = \phi_i \qquad \forall i \in C \qquad (10)$$

$$u_i \geq h_i + \tau_i + \sum_{p\in P} t_{ip} v_{ip} \qquad \forall i \in C \qquad (11)$$

$$u_i + \sum_{p\in P} (U + t_{pj}) v_{ip} \leq h_j + M^s_{ij}(1 - x_{ij}) \qquad \forall i \in C, j \in C_s, i \neq j \qquad (12)$$

$$a_d + \sum_{p\in P} t_{dp} v_{ip} \leq u_i \leq b_d - \sum_{p\in P} (U + t_{pd}) v_{ip} \qquad \forall i \in C \qquad (13)$$

$$f^s_{ij} \geq \sum_{p\in P} (t_{ip} + t_{pj})(x_{ij} + v_{ip} - 1) \qquad \forall i \in C, j \in C_s, i \neq j \qquad (14)$$

$$v_{ip} \in \{0,1\} \qquad \forall i \in C, p \in P \qquad (15)$$

Water Level Routing Problem

$$\sum_{j\in C_w} y_{ji} = \sum_{j\in C_w} y_{ij} = \sum_{p\in P} v_{ip} \qquad \forall i \in C \qquad (16)$$

$$\sum_{i\in C} y_{di} = \sum_{i\in C} y_{id} \leq K_w \qquad (17)$$

$$m_i \leq l_i \leq Q_w \qquad \forall i \in C \qquad (18)$$

$$l_j - l_i \geq m_j - Q_w(1 - y_{ij}) \qquad \forall i,j \in C, i \neq j \qquad (19)$$

$$u_i + U + f^w_{ij} \leq u_j + M^w_{ij}(1 - y_{ij}) \qquad \forall i,j \in C, i \neq j \qquad (20)$$

$$f^w_{di} \geq \sum_{p\in P} t_{dp}(y_{di} + v_{ip} - 1) \qquad \forall i \in C \qquad (21)$$

$$f^w_{id} \geq \sum_{p\in P} t_{pd}(y_{id} + v_{ip} - 1) \qquad \forall i \in C \qquad (22)$$

$$f^w_{ij} \geq t_{pr}(y_{ij} + v_{ip} + v_{jr} - 2) \qquad \forall i,j \in C, i \neq j, p, r \in P \qquad (23)$$

$$f^w_{ij} \geq 0 \qquad \forall i,j \in C_w, i \neq j \qquad (24)$$

$$y_{ij} \in \{0,1\} \qquad \forall i,j \in C_w, i \neq j \qquad (25)$$

Constraints (2)–(9) are related to the *street level routing problem*, formulated as a multi-depot vehicle routing problem with multi-trips and time windows where trips can visit any satellite to transfer and empty the load. Constraints (2) ensure that each customer is served exactly once by a LEFV while the constraint (3) indicates that the number of leaving and returning LEFVs must be equal and should not exceed the available fleet size of LEFVs. Constraints (4) impose a final transfer task for each LEFV to deliver the collected goods in the last trip to a vessel before returning to the garage. Constraints (5) are the capacity constraints considering direct flows and transfer decisions for load replenishment between customers i and j. Constraints (6) schedule the service start times of the customers assigned to a LEFV. If there exists any transfer task between nodes, the arrival time to the next node is delayed at least by the sum of the realized travel time and the duration of a transfer operation. Constraints (7) guarantee that the travel time realized from i to j must be positive if there is a LEFV visiting i and j subsequently and also define a lower bound on the travel if any transfer exists immediately after node i, assuming the closest satellite is visited. Lastly, constraints (8)–(9) are the variable domains for arc travel duration, routing LEFVs, and transfer task assignment.

Constraints (10)–(15) are related to the *synchronization problem* of the operations at both levels where the satellite assignments are optimized based on the trade-off between the cost of street and water levels. Constraints (10) assign a single satellite to a transfer task if it exists. Constraints (11) ensure that the transfer operation cannot start before the LEFV arrives at the assigned satellite while constraints (12) guarantee that LEFV cannot leave the satellite until the transfer operation is completed by delaying the arrival to the next node regarding the start time of the transfer. Besides temporal bounds imposed by LEFVs, constraints (13) ensure that the service start time for the transfer tasks should also respect the time window of the vessels, i.e., their daily operational hours. Constraints (14) update the travel time needed to visit the assigned satellite in between i and j if there is any. Satellite assignment decisions are binary as given by constraints (15).

Constraints (16)–(25) are related to the *water level routing problem* where all the scheduled transfer tasks should be assigned to the vessels and served to ensure load and temporal synchronization. Constraints (16) ensure that if there is a transfer task scheduled immediately after serving customer i, then the task should be served exactly once by a vessel. The constraint (17) limits the number of vessels used up to the fleet size. Constraints (18) ensure that the required capacity of a transfer task should be at least equal to the load of the LEFV after the last customer it served.

It eliminates partial transshipped loads to the vessels. However, the model is free to add more transfers if it is more efficient to split a transfer into two or more vessels to increase the vessels' capacity utilization. Constraints (19) guarantee that the capacity of each vessel is not exceeded at any point. Constraints (20) schedule the service start times of the transfer tasks assigned to a vessel and ensure that the delay between two subsequent tasks of a vessel must be as small as the sum of the required duration of a transfer task and the travel time that it should incur based on the satellite assignments of the tasks. Constraints (21)–(23) are the water level cost components to accurately calculate the realized travel time to serve the scheduled transfer tasks if there are any. They also ensure that there is sufficient time for vessels to be at the scheduled satellites for space synchronization. Constraints (24)–(25) are the domain of flow variables over water.

Notice that if there is no transfer task scheduled after a customer node i, then all constraints (10)–(25) related to synchronization and water level become redundant. For the majority of the cases in the optimal solution, the number of scheduled transfer tasks is way lower than the number of customers, and most of the constraints are redundant. Fewer transfer tasks reduce travel costs at both levels since the transfer tasks generated by the street level represent the problem to be solved by the water level. The only exception occurs if more transfers allow for further cost reduction by distributing demand more efficiently at the water level. Therefore, it is not necessarily true that global optimality is achieved at the minimum feasible number of transfers.

3.3. Modeling unitary transshipment capacity

In this study, satellites are public spaces without any necessity for infrastructures as opposed to typical operational hubs of different multimodal systems where goods can be sorted and stored. However, the real-time capacities of the satellites must be taken into account considering the limited spaces for multiple vehicles to park or maneuver, the lifting capacities (e.g., crane, rollers) and the labor-hour available at the vehicles. To address the synchronization issues in real-time capacities, it is assumed that at most one transfer task can be executed at a satellite at a time, meaning that any two scheduled transfer tasks cannot temporally overlap. It also means that a vessel cannot serve multiple LEFVs at the same time.

Let r_{ij} be the time difference between the service start time of the transfer tasks requested immediately after collecting node i and j. If they are allocated to the same satellite p, we guarantee that they are at least U units of time apart such that one is completed before the other begins by adding constraints:

$$|u_i - u_j| \geq U(v_{ip} + v_{jp} - 1) \qquad i, j \in C, i \neq j, p \in P, \qquad (26)$$

which can be linearized as:

$$u_i - u_j \leq r_{ij} \qquad i, j \in C, i < j \qquad (27)$$

$$u_j - u_i \leq r_{ij} \qquad i, j \in C, i < j \qquad (28)$$

$$r_{ij} = r_{ji} \qquad i, j \in C, i \neq j \qquad (29)$$

$$r_{ij} \geq U(v_{ip} + v_{jp} - 1) \qquad i, j \in C, i < j, p \in P. \qquad (30)$$

The capacity of the satellites can be increased to multiple transfers at a satellite. If there exists space and resources to execute more than a single transfer at a satellite at any point in time, a copy of the satellite can be added to the problem.

3.4. A compact formulation for 2E-MVRP-SS with deliveries

In the 2E-MVRP-SS with pickups (Section 3.2), LEFVs collect goods at customer locations and deliver them to vessels operating on inland waterways for the last mile to a central depot. In contrast, the 2E-MVRP-SS with deliveries considers the reverse flow: vessels transport goods from a central depot to satellites in order to transfer them to several LEFVs that perform the last mile to the customers. Both problems ensure satellite synchronization, where the different-echelon vehicles responsible for a transfer must be present at the selected satellite to realize the transfer operation. These vehicles may arrive earlier than the other but are strictly forbidden to leave before the transfer operation ends, which implies synchronization in time, space, and cargo flow. Therefore, the only difference between pickup and delivery problems lies in the direction of the transfers at the satellites. Performing deliveries requires loading packages to be delivered to the customers from vessels to LEFVs at the beginning of each trip, whereas performing pickups involves transferring collected items from LEFVs to vessels at the end of each trip.

The notation, sets, and parameters described in Section 3.2 are also valid for the delivery problem, but the following modifications apply:

- $\phi_i = 1$ if there is a transfer task immediately *before* serving customer i and 0 otherwise.
- m_i represents the total load on the LEFV immediately *after* visiting customer node i to deliver q_i, which is now bounded as $0 \leq m_i \leq Q_s - q_i$.

Additionally, modeling deliveries requires modifying the constraints related to load synchronization on both levels for forward flows. To do so, the constraints (5) and (18) are replaced with:

$$m_i \geq m_j + q_j - Q_s(1 - x_{ij} + \phi_j), \qquad \forall i, j \in C, i \neq j \qquad (31)$$

$$l_j \geq m_j + q_j, \qquad \forall j \in C \qquad (32)$$

For temporal synchronization, the constraints (6), (11) and (12) are replaced with:

$$h_i + \tau_i + f^s_{ij} + U\phi_j \leq h_j + M^s_{ij}(1 - x_{ij}), \qquad \forall i, j \in C, i \neq j \qquad (33)$$

$$u_j \geq h_i + \tau_i + \sum_{p \in S} t_{ip} v_{jp} - M^s_{ij}(1 - x_{ij}), \qquad \forall i, j \in C, i \neq j \qquad (34)$$

$$u_j + \sum_{p \in S} (U + t_{pj}) v_{jp} \leq h_j, \qquad \forall j \in C \qquad (35)$$

To ensure that LEFVs are loaded with delivery packages at a satellite before visiting any customer, the constraints (4) are replaced with:

$$\phi_i \geq x_{gi}, \qquad \forall i \in C \qquad (36)$$

Lastly, the cost definitions for street level considering transfer task assignments by constraints (7) and (14) are replaced with:

$$f^s_{ij} \geq t_{ij} x_{ij} + \min_{p \in P}\{t_{ip} + t_{pj} - t_{ij}\}(1 - x_{ij} + \phi_j), \qquad \forall i, j \in C_s, i \neq j \quad (37)$$

$$f^s_{ij} \geq \sum_{p \in S} (t_{ip} + t_{pj})(x_{ij} + v_{jp} - 1), \qquad \forall i \in C, j \in C_s, i \neq j \quad (38)$$

Table 2
Formulation modifications for different 2E-MVRP variants with satellite synchronization.

Use case (Variant)	Street level	Synchronization	Water level
Unlimited satellite capacity, e.g., Marques et al. (2020)	MVRP	Remove upper (lower) bound on transfer end (start) time for pickup (delivery)	LRP
Limited satellite capacity & dedicated customers, e.g., Li et al. (2018)	MVRP	Add flow balance constraints to ensure transfer task sequences do not exceed capacity at any time	VRP
Limited satellite capacity & re-charging at the satellites, e.g., Breunig et al. (2019)	e-MVRP	Delay transfer operations for charging decisions and limit the maximum amount of goods assigned to the satellites	LRP
Limited satellite capacity & simultaneous pickup-delivery, e.g., Dumez et al. (2023)	MVRPSPD	Track the satellites' used capacity and delay vehicle departures if exceeded	LRPSPD

Notes: VRP: Vehicle routing problem, MVRP: Multi-trip VRP, LRP: Location routing problem, e-MVRP: Electric MVRP, MVRPSPD: MVRP with simultaneous pickup and delivery, LRPSPD: LRP with simultaneous pickup and delivery.

3.5. Different variants through modular formulation

Ultimately, the proposed linear model for the synchronized two-echelon problem can be used for different variants considering different applications affecting the operations at one or both levels or transfers at the satellites. Thanks to our modular formulation, which is composed of three explicit problems, namely, the routing problems of street and water levels and the synchronization problem at the satellites, any change in the operations can be reflected in the related sub-module in the integrated problem. For example, service type change from pickups to deliveries given in this section only affects the flow of the goods at the street level and the start of the transfers at the satellites since each transfer is performed just before a trip to load necessary cargo on LEFVs. Another application might consider an electric 2E-VRP with charging options at the satellites (Breunig, Baldacci, Hartl, & Vidal, 2019). Table 2 provides an overview of potential variants that our formulation can handle upon necessary adaptations. Moreover, the proposed formulation can be extended to other transportation modes without losing its generality, e.g., vans-bikes, vans-trucks, bikes-trams. Vehicles operating on these modes, e.g., vans, trucks, and trams etc., can act as vessels and the transshipments operations can be executed at the satellites, e.g., parking lots for trucks and vans, and tram stations for trams. However, the changes related to the synchronization problem must be reflected accordingly considering the available satellites to perform transfers such as trucks, tram stations or warehouses, etc.

4. A logic-based benders decomposition approach

Benders Decomposition (BD) was proposed by Benders (1962) for tackling large-scale optimization problems, where the complexity tends to increase exponentially with the size of the problem. The essential idea is to decompose the problems into a master problem having the complicating variables and a linear subproblem by fixing those variables. The master problem is solved iteratively by generating feasibility and optimality cuts using duality information of fixed variables on the subproblem. The main drawback of BD is the limitation of having a linear problem (LP) structure for subproblems. Hooker and Ottosson (2003) proposes LBBD and generalizes the convergence mechanism of the classical BD to a wider variety of problems which can be decomposed into easier subproblems in the form of not only an LP but also a mixed integer problem (MIP) or constraint programming (CP). LBBD uses a more general inference dual to generate cuts derived from logical deductions.

$$\mathbb{J} = min\ \{z(x,v,y) \mid Constraints\ (2)\text{–}(25),\ x \in D_x,\ v \in D_v,\ y \in D_y\} \tag{39}$$

$$\mathbb{M} = min\{z(x,\phi) \mid Constraints\ (2)\text{–}(9),\ OPT(x,\phi),\ FEAS(x,\phi),\ x \in D_x,\ \phi \in D_\phi\} \tag{40}$$

$$\mathbb{S} = \mathbb{J}(\overline{x},\overline{\phi}) = min\{z(\overline{x},\overline{\phi},v,y) \mid Constraints\ (6)_{\overline{x},\overline{\phi}} \mid Constraints\ (10)\text{–}(25),\ v \in D_v,\ y \in D_y\} \tag{41}$$

The 2E-MVRP-SS is formulated in the form of $\mathbb{J}$ by (39) in terms of feasible regions provided in the MILP formulation in Section 3.2 and complicating binary variables, namely x for routing LEFVs, ϕ for transfers, v for satellite assignments of the transfers, and y for routing vessels. Constraints for the synchronization problem define the feasible region of the transfers considering all possible satellite assignments. Constraints for the water level routing problem represent the feasible region of the vessel routing problem to serve these transfers respecting temporal, spatial, and load synchronization requirements.

The complexity of the problem increases quadratically to construct the synchronization and water level problems' regions. However, all the constraints and complicating variables related to these problems are redundant except the binding constraints defined by the optimal solution (if it exists). To reduce the complexity and have a tractable algorithm, we propose decomposing the problem $\mathbb{J}$ into a master problem and a subproblem. The master problem, $\mathbb{M}$ by (40), is to solve the street level problem over the region provided by constraints set defined by considering x and ϕ decisions. To ensure feasibility and the optimality of the street level decisions, logical Benders cuts are added to the problem derived from the solution of the subproblem $\mathbb{S}$ by (41), which solves the 2E-MVRP-SS for a given feasible solution to the master problem.

4.1. Master problem ($\mathbb{M}$)

The master problem $\mathbb{M}$ is a relaxed street level problem for all feasible x and ϕ decisions constrained by (2)–(9). The actual cost of street level movements depends on the total travel time adjusted by satellite assignments for transfer tasks, $z^s(x,v)$ to achieve global optimality. $\mathbb{M}$ relaxes these assignments into $z^s(x,\phi)$ by assuming that transfers are assigned to the closest satellite and served on time upon the arrivals of LEFVs by (7). It implies no waiting time for LEFVs in the best-case scenario under unlimited resources at satellites. Accordingly, the total traveling cost at the street level is formulated by (42) and (43) to bound the travel cost of the routing decisions, x, without and with satellite visit, ϕ. On the other hand, the best feasible water level cost is the fleet cost of having the minimum number of vessels to store all the demand, $K_w^l = \lceil \frac{\sum_{i \in C} q_i}{Q_w} \rceil$, and traveling cost of those vessels to the closest satellite from the central depot. Therefore, the lowest bound on the cost of water level movements to serve a given set of transfers is formulated by (44). With these assumptions, $\mathbb{M}$ does not eliminate any feasible solution to the 2E-MVRP-SS, and reformulates the objective of the synchronized two-echelon problem by decomposing the cost function into a linear part, denoted as f_- by (45), and a nonlinear part, denoted as f_+ by (46). The objective, $z_{\mathbb{M}}$, is to minimize the integrated cost based on routing and transfer decisions by (47) without

$$\mathbb{S}(\overline{x},\overline{\phi}) : \min \; z(E_-, E_+) \tag{50}$$

s.t.

$$h_i + \tau_i + t_{ij} \le h_j \qquad \forall (i,j) \in E_- \tag{51}$$

$$h_i + \tau_i + t_{ij} + \epsilon_{ij} + U \le h_j + M^s_{ij}(1 - \sum_{p\in P} v_{ip}) \qquad \forall (i,j) \in E_+ \tag{52}$$

$$\sum_{p\in P} v_{ip} = 1 \qquad \forall (i,j) \in E_+ \tag{53}$$

$$Constraints\ (11) - -(13) \qquad \forall (i,j) \in E_+ \tag{54}$$

$$\epsilon_{ij} \ge \sum_{p\in P} \{t_{ip} + t_{pj}\} v_{ip} \qquad \forall (i,j) \in E_+ \tag{55}$$

$$Constraints\ (15) - (25) \& (44) \tag{56}$$

$$z^{UB} \ge z^s_{\mathbb{S}} \ge \sum_{x_{gi}\in E_-} (\beta_s + c^s t_{gi}) + \sum_{(i,j)\in E_-} c^s t_{ij} + \sum_{(i,j)\in E_+} c^s \epsilon_{ij} + z^w(s) \tag{57}$$

$$f_+ \ge \sum_{(i,j)\in E_+} c^s \epsilon_{ij} + z^w \tag{58}$$

Fig. 2. The subproblem for a given $\mathbb{M}$ feasible solution, $\{\overline{x},\overline{\phi}\} := E_- \cup E_+$. E_- is the set of (i,j) pairs representing the consecutive customer visits from i to j without a transfer in between, where $x_{ij} = 1$ and $\phi_i = 0$ in the solution (fixed cost of the solution). E_+ is the set of (i,j) pairs with a transfer, where $x_{ij} = 1$ and $\phi_i = 1$ (variable cost of the solution).

losing the generality. It reduces the problem to a multi-depot MVRPTW (MDMVRPTW), where street vehicles visit any of the satellites as many times as needed for capacity replenishment.

$$\psi_{ij} \ge c^s t_{ij}(x_{ij} - \phi_i), \qquad \forall i \in C, j \in C_s, i \ne j \tag{42}$$

$$\epsilon_{ij} \ge c^s \min_{p\in P}\{t_{ip} + t_{pj}\}(x_{ij} + \phi_i - 1), \qquad \forall i \in C, j \in C_s, i \ne j \tag{43}$$

$$z^w \ge K^l_w \left(\beta_w + \min_{p\in P}\{t_{dp} + t_{pd}\}\right) \tag{44}$$

$$f_- \ge \sum_{i\in C} (\beta_s + c^s t_{gi}) x_{gi} + \sum_{i\in C, j\in C_s} \psi_{ij} \tag{45}$$

$$f_+ \ge z^w + \sum_{i\in C, j\in C_s} \epsilon_{ij} \tag{46}$$

$$z_{\mathbb{M}} \ge f_- + f_+ \tag{47}$$

$$\psi_{ij}, \epsilon_{ij}, f_-, f_+ \ge 0 \qquad \forall i,j \in C_s, i \ne j \tag{48}$$

$$Logic - based\ Benders\ cuts \tag{49}$$

$\mathbb{M}$ provides feasible solutions with lower bounds. A feasible solution, $\{\overline{x},\overline{\phi}\}$, is fed to the subproblem representing the variables for the used arcs and transfer points only. The solution is checked against feasibility and optimality by solving the related $\mathbb{S}(\overline{\overline{x}},\overline{\phi})$ to optimize the costs over a two-echelon synchronized setting. The proof of optimality or infeasibility is added to $\mathbb{M}$ before accepting a solution as the incumbent solution via logic-based Benders cuts (49) deducted from the corresponding subproblem provided in Section 4.2. These cuts improve the bound on the nonlinear part of the cost, f_+, due to the interdependence problem in the two-echelon problems.

4.2. Subproblem ($\mathbb{S}$)

The synchronization and optimization subproblem, $\mathbb{S}$, is solved for a given feasible solution with $\{\overline{x},\overline{\phi}\}$ to optimize transfer-satellite assignments (v) and routing of vessels (y) to serve the transfers in the set of $\overline{\phi}$ in synchronization with LEFVs considering the routing decisions in $\overline{x}$. A feasible schedule for LEFVs defines the temporal precedence relationships between the operations, the subproblem defines the satellite visits on the arcs with transfer decisions.

The subproblem $\mathbb{S}(\overline{\overline{x}},\overline{\phi})$ is relatively easier to solve since the constraints related to the synchronization and water level problems are constructed only for the given set of transshipment decisions on the chosen arcs, provided Fig. 2. It is formulated as a MILP to minimize the total cost with respect to complicating satellite assignments and water level routing decisions. The problem is to locate all the transshipment operations given on the selected arcs, by (53) and route the vessels by complying the temporal precedence constraints on all arcs, without (E_-) and with (E_+) transshipment, by (51) and (52). $\mathbb{S}$ is used as a proof of infeasibility as well as optimality by using the feasible solutions that provide better or at least as good lower bounds as $\mathbb{M}$ at linear relaxation.

4.3. Branch and check: a generalization of LBBD

Both classical BD and LBBD approaches first solve the master problem to optimality and then solve the subproblem(s) using the fixed master-optimal solution. In this study, the master problem is formulated as MDMVRPTW with the assumption of the best satellite assignment for LEFVs. The subproblem is defined as an assignment and scheduling optimization problem.

The master problem relaxes all the resource constraints and provides solutions that are feasible on the demand side ensuring a feasible service by the LEFVs. On the other hand, the subproblem evaluates the feasibility of the supply side by the vessels and the cost of the proposed solutions for global optimality. However, it does not necessarily match with the minimum feasible cost schedule for $\mathbb{M}$. Optimizing $\mathbb{M}$ at each iteration improves the lower bound but prolongs improving upper bound by discarding intermediate solutions (Fragkogios, Qiu, Saharidis, & Pardalos, 2024). However, these discarded solutions may be revisited in subsequent iterations if optimality is not achieved.

The multi-trip VRP is challenging to solve considering the timing aspects of the trips when scheduling the vehicles in the existence of time windows (Pan, Zhang, & Lim, 2021). The state-of-the-art model for the single depot MVRPTW literature solves problems up to 50 nodes by using column generation, column enumeration, and cutting plane for single depot case (Paradiso, Roberti, Laganá, & Dullaert, 2020). Similarly, Huang, Li, Zhu, and Qin (2021) propose a column generation method for MVRPTW with a limited number of transshipment operations at the depot at a time. They show that resource capacity at the depots complicates the problem further by limiting the scale of the problems to 50 customers for providing feasible solutions. Considering the complexity of the master problem, we propose a branch and check (B&C) method, a generalization of LBBD (Thorsteinsson, 2001), to solve $\mathbb{M}$ with the use of subproblems until a predefined termination criterion. As a single search tree, it incorporates solving subproblems into branch and bound (B&B) process at every feasible solution as a proof of feasibility to cut off infeasible or sub-optimal solution, instead of solving $\mathbb{M}$

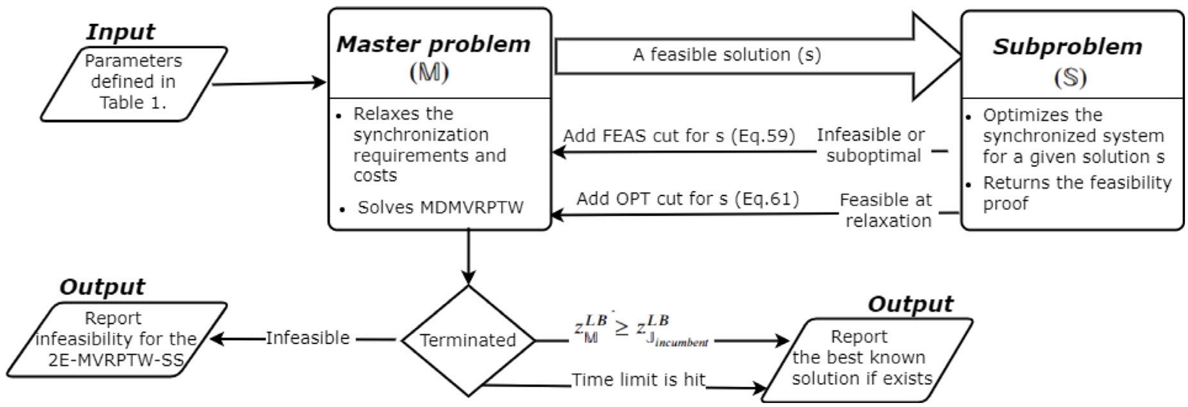


Fig. 3. The flowchart for the proposed LBBD proposed as B&C. $z_{\mathbb{M}}^{LB}$ and $z_{\mathbb{J}_{incumbent}}^{LB}$ are the best lower bounds of the master problem and the synchronized two-echelon problem respectively.

to optimality at once. B&C is also referred as branch-and-Benders-cut (B&BC) (Rahmaniani et al., 2017).

Geoffrion (1972) shows that using the optimal multiplier values corresponding to various trials ensures termination in a finite number of steps if a positive optimality gap is allowed, *epsilon*-optimality. Instead of using a predefined gap, we add optimality cuts to improve the lower bound of $\mathbb{M}$ up to the one of the corresponding relaxed subproblem. The method terminates if the lower bounds of the best-known solution and $\mathbb{M}$ are equal, the time limit is achieved or the problem is infeasible. The lower bounds are derived from the Lagrangian dual of the synchronized two-echelon problem improving the bounds provided by linear relaxation of $\mathbb{M}$. Frangioni (2005) reminds the effective use of a continuous relaxation for Lagrangian approaches within algorithms by using the dual values of the "easy" satellite assignment constraints by (10) in the form of $Ex = b$. The proposed B&C aims at finding feasible solutions faster to maintain improved upper bounds in a single B&B tree and exponentially converges to the optimality if the duality gap is zero. Otherwise, it terminates earlier with a positive gap. While the joint MILP depends on continuous relaxations, $\mathbb{M}$ exploits the Lagrangian relaxations for feasible street level solutions. The flow chart for the proposed LBBD is given in Fig. 3.

4.3.1. Benders feasibility cuts

A feasible solution to $\mathbb{M}$ is considered infeasible for $\mathbb{S}$ due to two reasons. Firstly, it might be infeasible under available temporal capacities at the satellites limited to unitary transshipments. $\mathbb{M}$ relaxes this limitation assuming each transfer is served upon arrival of LEFVs. Scheduling the transfers at each satellite might cause waiting times for LEFVs and the delay might lead to temporal infeasibility. Secondly, a given solution with a promising relaxed objective value might not yield a better solution than the incumbent solution found up to that point. The model needs to cut off the solution. If any of the two cases occur, (59) is added to $\mathbb{M}$ implying a change in the κth master feasible solution $E^{\kappa} := E_{-} \cup E_{+}$. This should be at least a single change in excluded routing decisions or the removal of at least one transfer decision included in the current solution.

$$FEAS(E^{\kappa}) : \sum_{(i,j)\notin E^{\kappa}} x_{ij} + \sum_{i\notin \overline{\phi}} \phi_i + \sum_{(i,j)\in \overline{\phi}} (1-\phi_i) \geq 1 \tag{59}$$

Theorem 1. *The cut proposed in (59) must remove the current solution from the master problem space.*

Proof. The combinatorial cut given by (59) eliminates a single unique solution by enforcing the master problem to make at least one change in the solution, i.e., either the routing should change or the transfer set. It discards the current solution without providing any bound information or search direction. It acts as "naive" feasibility cuts. If the subproblem is a linear optimization problem, then this cut is sufficient for the optimality convergence (Ahat, Ekim, & Taşkın, 2018). For nonlinear subproblems, it provides *no-good* optimality cut that might or might not improve the lower bound (Martínez et al., 2022). This cut is used as a feasibility cut for an infeasible solution and as an optimality cut for a sub-optimal solution. Cutting off a sub-optimal solution using a "naive cut" might not improve the lower bound but reduces the burden on the solver compared to using dual information for each feasible solution.

4.3.2. The proposed Lagrangian bounding approach

In classical BD, a master feasible solution improves the upper and lower bounds using the duality of the convex subproblems. However, the strong duality is not applicable for solving the 2E-MVRP-SS due to the non-convex nature of the underlying subproblems. Therefore, a combinatorial logical cut must be derived to provide a valid lower bound if applicable. Otherwise, it should not remove any globally optimal solution.

Suppose $\mathbb{S}$ provides a feasible solution to the synchronized two-echelon problem fixed at a master feasible solution, E^{κ}. $\mathbb{S}$ only includes the complicating decisions set by E^{κ} and excludes the decisions not in E^{κ}. The relaxed solution $\mathbb{S}$ still can be used to obtain Lagrangian multipliers for the decisions in E^{κ}_{+} constrained to the routing decisions in E^{κ}.

Dual inference using Lagrangian optimization

The street level and water level problems regarding the transfer decisions in a feasible solution are connected via the synchronization problem formulated linearly in Section 3.2. It is clear that $\mathbb{S}$ is equivalent to $\mathbb{J}$ for any given solution. Satellite-transfer assignment constraints (53) enforce space synchronization for transfers to achieve global optimality that minimizes the total logistic costs at both levels. If they are relaxed in $\mathbb{S}$, it solves the synchronized two-echelon problem by allocating the transfers to the satellites while solving a minimum cost flow problem for feasible allocations. Additionally, the cost of water level routing built by (22)–(24) provides equal or improved bounds on the integrated cost considering the quadratic assignment problem of the transfer decisions. If any of them is removed, it simply removes the related arc from the solution by breaking binding temporal constraints on the interacting vehicles. There is no reason to assign transfers to satellites and to vessels or assign any cost on both levels. The optimal dual values of these constraints, λ^{κ}_{ij} for each decision in E^{κ}_{+} where $x_{i,j} = 1$ and $\phi_i = 1$, provided by the subproblem relaxation give us the maximum improvements in the relaxed objective if both are removed from the solution (Fisher, 2004). Therefore at the optimal solution, the Lagrangian relaxation of $\mathbb{J}$ with respect to the complicating satellite assignment constraints can be written as:

$$z^{*}_{\mathbb{J}} = z^{*}_{\mathbb{S}} \geq z^{*}_{\mathbb{M}} \geq f^{*}_{-} + f^{*}_{+} + \sum_{(i,j)\in E^{*}_{+}} \lambda^{*}_{ij}(x_{ij} + \phi_i - 2) \tag{60}$$

Equivalently, it can be written for the feasible relaxation of the subproblem corresponding to the solution E^{κ} as follows:

$$z_{\mathbb{M}} \geq \sum_{(g,i)\in E^{\kappa}_{-}} (\beta_s + c^s t_{gi}) x_{gi} + \sum_{(i,j)\in E^{\kappa}_{-}} \psi_{ij} + f_{+}^{E_0^{\kappa}} + \sum_{(i,j)\in E^{\kappa}_{+}} \lambda_{ij}^{\kappa_0} (x_{ij} + \phi_i - 2) \quad (61)$$

Theorem 2. *The cut proposed in Eq. (61) provides improved bounds for the synchronized two-echelon problem compared to the bounds from continuous relaxation of* $\mathbb{M}$.

Proof. The complicating constraints for the synchronized problem are the satellite assignments, which are relaxed in $\mathbb{M}$ all together with the constraints in the synchronization and water level problems. The lower bound provided by the relaxation of $\mathbb{S}$ for a feasible solution, might be good for the feasible solutions that also exclude the same decisions. Exclusion is a decision given by the B&B process to improve the lower or upper bound for $\mathbb{M}$. It might not be necessarily the best decision for optimality. When this is the case, $\mathbb{M}$ should search for removing at least one decision included in the current solution to achieve a different feasible solution. Otherwise, it should not affect the global lower bound. The lower bound provided by $\mathbb{S}$ is valid for a unique solution and does not affect the lower bound for globally feasible solutions by providing a feasible relaxation to the synchronized two-echelon problem.

Eq. (61) does not provide a better lower bound than the global lower bound if excluded decisions for any feasible solution to $\mathbb{M}$ differ at least by one. When this is not the case, it is a valid cut because the only possibility for a different feasible solution is to have the same routing and change transfer decisions in the current solution. The linear part ensures a lower bound based on shadow prices of the decisions on the arcs with transshipments.

4.3.3. Solution time accelerating strategies

To accelerate the LBBD method, $\mathbb{M}$ is first solved to obtain initial solutions without checking the feasibility. Then, these solutions are tested using $\mathbb{S}$ to provide an upper bound for the model. In this way, expensive subproblems are eliminated for the solutions in the beginning. Additionally, the subproblems are solved with a time limit to minimize the time to achieve optimality for intermediate solutions. However, the relaxations of the subproblems are solved optimally to obtain dual values and the integrated lower bound of the solutions. A master feasible solution might be feasible for the subproblem relaxation. However, it might not improve the incumbent solution by yielding a sub-optimal solution for the synchronized two-echelon problem. In the case of sub-optimal solutions during B&C, optimality cuts by (61) are added to the $\mathbb{M}$ as lazy constraints for the master feasible solutions providing a feasible subproblem relaxation to improve the lower bound. However, feasibility cut by (59) is also added to the model to cut off the solution.

5. Computational results

In this section, we introduce the data set used in the experimental study and present the results. The problems are solved by a commercial solver, Gurobi 9.12. The master problems are solved using lazy callbacks for each feasible solution as documented in the Gurobi manual (Gurobi Optimization, 2021). The test are conducted using 4 CPUs on Intel(R) Xeon(R) Gold 5218 with 2.30 GHz clock speed. For the solution methods, the time limit is 4 h (14 400 s) regarding the extent and the scale of the problems solved in this study. Furthermore, the time limit to optimize the subproblems is set to 50 s for the instances up to 30 demand nodes while it is set as 100 s for the instances with 100 nodes. The same time limits are used for initial solution generation at the beginning of the proposed LBBD approach.

5.1. Test instances

Grangier et al. (2016) introduce and solve the 2E-MVRP-SS with no storage option using a customized ALNS by rescheduling operations after an insertion to prevent time window violations, with a linear complexity in the size of the route. The linearity depends on the assumption that many transshipment operations can be handled at the satellites at any time and these operations happen instantly, without causing delays due to the loading/unloading of the goods. However, it removes the resource synchronization problem at the satellites and only accounts for the delays in visiting the satellites. The 2E-MVRP-SS with unitary transshipment is proposed to schedule transshipment operations to the available satellites with no storage ensuring that the real-time capacities are not violated. Since it has not been studied in the literature, we first describe the test instances used in the experiments.

Network: The instances are generated by modifying Solomon's (1987) VRPTW instances for geographical configuration of the demand. Solomon class "2" instances are considered with relatively wider time windows and long scheduling horizons to let LEFVs have multi-trips.

To generate problems of varying sizes, we utilize the initial $|C|$ nodes from the given Solomon instances. For each size, three types of demand distribution are considered. The "C" type comprises 8 instances where customers are concentrated in clusters, while the "R" type consists of 11 instances with customers randomly located throughout the area. As a mix of R and C, the third type "RC" includes 8 instances where customers are either clustered or randomly positioned.

To ensure the feasibility of the instances for 2E-MVRP-SS with unitary transshipment, the time windows for satellites and the central depot are the same as the garage defining the working hours of all the vehicles, and set to the latest possible return time from any customer in the demand set. The return time assumes the latest possible service start time at a customer and the furthest satellite to visit to perform the last transshipment before vessels returning to the depot, d. This limits the feasible number of the transshipment operations in time at the satellites after the daily customer service is completed. It prevents SL vehicles queue at the cheapest satellites and wait long enough for the vessels to minimize the SL logistics costs.

For the water level network, the satellites and the vessel central depot are located outside the city while general practice is to locate them in urban areas. Keeping the transfer operations away from the public is primarily motivated by our previous study on waste collection (Karademir et al., 2022), and we adopt similar strategy. This assumption reduces the concerns and related costs about inconveniences caused by transfer operations such as noise, congestion and reduced mobility due to the lack of space in the city. In total 4 satellites and a vessel central depot, d, are located for all problems tested in this study ranging from 10 to 100 customer nodes for different scenarios. The satellites are positioned at the midpoint of every side of the map that covers all of the demand nodes. The assumed locations of the satellites are depicted in Fig. 4 in Section 5.4 for different demand distributions. We further examine this assumption by using several rules based on the proximity to the centroid of the demand network.

Vehicles: To better observe multiple trips and transfer tasks, the capacity of a LEFV is set to 50 units, and the capacity of a vessel is set to 250 units. The distance and travel duration along any arc are defined to be equal to the Euclidean distance between the nodes on that arc. For all computations, the fleet size of each level is assumed to be unlimited to observe at least a feasible solution. Lastly, U is assumed to be twice the average service duration of the customers, rounded up.

Objective: We use a lexicographic objective to prioritize the fleet size over the travel cost, meaning that any solution with fewer vehicles is superior to any other with more vehicles at any level. However, the model is highly dependent on the relative costs of the levels to minimize the total cost on one level further over the cost of the other level. Estimating the fixed cost of using vehicles for logistics and determining the optimal number of vehicles for a single-echelon

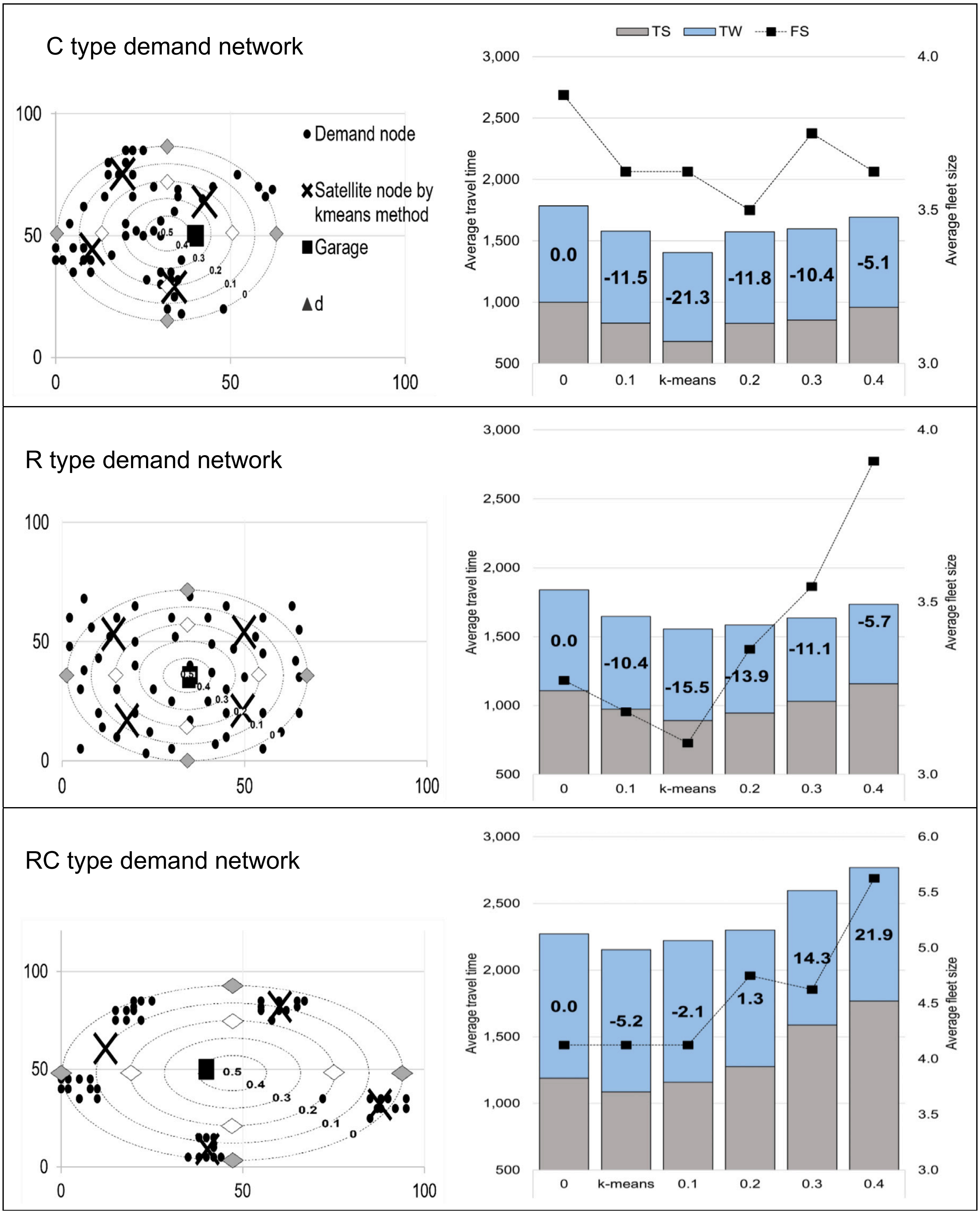


Fig. 4. Comparison of different satellite locations across clustered (C), random (R), and random clustered (RC) customer distributions with 50 customers and 4 satellites ($|C| = 50$, $|P| = 4$). For each type, on the left, the satellite location scenarios with different proximity values are illustrated as diamonds, and k-means centers are shown as crosses. On the right, the results are ordered by the proximity values, and the k-means scenario is placed at the intersection of two consecutive proximity scenarios closest in total costs for each distribution type. The percentage savings are reported for each scenario in terms of total travel time with respect to the base scenario (0), shown as gray diamonds for the demand network of each type on the left. Time limit is 4 h for the LBBD method which solves all the problems feasibly. WLS = 1:10.

VRP can be challenging. The larger the fixed cost, the less the model focuses on improving the travel cost. It might lead to early termination due to the smaller gaps considering large fixed cost values for the vehicles. The smaller the fixed cost, the more the model works on improving the travel cost. However, the minimum travel cost does not always guarantee the minimum number of vehicles. Therefore, it is important to choose the parameters wisely to reflect the priorities of the stakeholders. To reduce the effect of the cost parameters on the experiments, we consider a reference value for each cost parameter. Then, the parameters are multiplied by the importance ratio defined by the user based on the purpose of the experiments.

For the travel costs on both networks, we assume $\underline{c} = 1$ implying that navigating on the streets per unit of time is equal to the one on the waterways. For the fixed costs, the maximum travel cost of serving

Table 3
Comparative results on test instances.

Size	|C| = 10, |P| = 4						|C| = 20, |P| = 4						|C| = 30, |P| = 4					
Method	Joint MILP			LBBD			Joint MILP			LBBD			Joint MILP			LBBD		
Instance	BK	Gap %	Inc. Time (s)	BK	Imp. %	Inc. Time (s)	BK	Gap %	Inc. Time (s)	BK	Imp. %	Inc. Time (s)	BK	Gap %	Inc. Time (s)	BK	Imp. %	Inc. Time (s)
C201	365.5	0	0	365.5	0.0	1	697.7	0	4	697.7	0.0	2	1011.9	0	1313	1012.5	0.1	113
C202	327.5	0	18	327.5	0.0	12	678.7	22	1 690	678.7	0.0	241	876.4	32	9 620	**854.9**	−2.4	4 235
C203	327.5	0	19	327.5	0.0	14	682.7	54	8 395	**676.3**	−0.9	1649	1047.2	52	10 406	**843.6**	−19.4	6 006
C204	304.3	0	78	304.3	0.0	2	683.5	57	2 261	**668.1**	−2.3	380	902.9	62	11 985	**810.8**	−10.2	2 663
C205	365.5	0	1	365.5	0.0	1	685.2	0	69	685.2	0.0	99	869.6	10	4735	869.6	0.0	194
C206	365.5	0	2	365.5	0.0	9	680.8	0	137	680.8	0.0	133	853.7	10	4 212	**847.5**	−0.7	3 081
C207	365.5	0	1	365.5	0.0	0	681.8	23	1 258	**678.8**	−0.4	428	836.9	35	7 128	843.3	0.8	687
C208	358.4	0	6	358.4	0.0	2	678.1	0	2 457	680.3	0.3	14	845.5	42	5 856	**835.7**	−1.2	324
R201	420.1	0	0	420.1	0.0	0	821.7	0	102	821.7	0.0	50	1048.8	2	6 757	1048.8	0.0	143
R202	387.8	0	22	387.8	0.0	11	757.0	19	11 478	**756.7**	0.0	243	1046.8	49	13 800	**1003.0**	−4.2	5 179
R203	387.8	0	23	387.8	0.0	12	638.4	37	12 783	**635.8**	−0.4	1710	999.9	54	10 185	**919.8**	−8.0	6 090
R204	351.4	0	61	351.4	0.0	1	605.7	42	5 713	605.7	0.0	3036	929.3	55	9 735	**851.1**	−8.4	2 015
R205	376.4	0	2	376.4	0.0	1	642.2	0	1 990	676.3	5.3	529	1017.6	40	6 300	**968.5**	−4.8	2 401
R206	327.6	0	32	327.6	0.0	1	612.9	31	7405	**603.8**	−1.5	517	950.1	50	9 933	**914.0**	−3.8	6 695
R207	327.6	0	31	327.6	0.0	1	601.7	36	8 972	605.4	0.6	2747	969.9	55	8 144	**914.5**	−5.7	6 204
R208	327.6	0	54	327.6	0.0	1	549.9	37	11 009	549.9	0.0	809	931.2	56	8 706	**884.6**	−5.0	3 532
R209	362.1	0	4	362.1	0.0	1	626.3	6	10 974	626.3	0.0	99	1071.7	53	7 514	**933.9**	−12.9	4 303
R210	353.4	0	9	353.4	0.0	7	647.1	11	3 954	**632.8**	−2.2	3158	1024.6	36	3;302	**937.3**	−8.5	2 760
R211	344.5	0	15	344.5	0.0	2	669.1	48	10 782	**584.4**	−12.7	1316	964.9	57	13 067	**900.3**	−6.7	1 618
RC201	392.8	0	1	392.8	0.0	0	981.2	0	27	981.2	0.0	168	1712.7	11	2 321	**1665.5**	−2.8	1 896
RC202	356.5	0	55	356.5	0.0	8	906.4	44	5 928	906.4	0.0	751	1641.9	60	10 662	**1413.4**	−13.9	1 984
RC203	356.5	0	51	356.5	0.0	7	754.3	56	8 413	767.6	1.8	2947	1589.6	70	3 490	**1379.3**	−13.2	4 536
RC204	320.2	0	86	320.2	0.0	1	642.4	50	10 375	642.4	0.0	3763	1336.6	71	13 249	**1273.9**	−4.7	1 927
RC205	440.6	0	71	440.6	0.0	24	935.5	15	1 098	938.4	0.3	570	1639.7	39	3 705	**1597.7**	−2.6	2 571
RC206	363.1	0	5	363.1	0.0	1	885.2	28	1 149	885.2	0.0	837	1395.6	26	13 448	**1384.6**	−0.8	442
RC207	364.4	0	4	364.4	0.0	1	885.0	26	4 242	888.7	0.4	13	1705.9	64	4 718	**1602.3**	−6.1	12 293
RC208	318.7	0	210	318.7	0.0	2	881.6	65	4 166	**858.7**	−2.6	1478	1353.9	71	8 681	**1293.0**	−4.5	6 252
Averages	*357.7*	*0.0*	*31.9*	*357.7*	*0.0*	*4.5*	*722.7*	*26.2*	*5067.8*	*719.0*	*−0.5*	*1025.3*	*1132.4*	*43.0*	*7887.8*	*1066.8*	*−5.5*	*3338.6*
Run time (s)			*197*			*11*			*10 554*			*9192*			*13 346*			*13 365*
Initial solution				*357.8*						*747.7*						*1197.0*		

BK: The best-known solutions to the methods, *Inc. Time*: The time it takes to find the best-known solution, *Gap %*: The percent gap reported by the solver for the joint method, *Imp. %*: The percent improvement of the *BK* of LBBD compared to the *BK* of the joint MILP. Time limit is 14 400 s for both models.

a customer requires a visit to the customer by a LEFV and another visit to a satellite by the LEFV and a vessel for transferring the collected items. In other words, a customer service costs two vehicles in a two-echelon setting. The maximum value guarantees that fewer vehicles are preferred over travel costs since the smallest fleet does not necessarily always achieve the smallest travel costs.

$$\underline{\beta}(c^s, c^w) = \max_{i\in C,\, p\in S}\{c^s(t_{gi} + t_{ip} + t_{pg}) + c^w(t_{dp} + t_{pd})\} \tag{62}$$

Since the fixed cost is the worst possible route for a LEFV and a vessel to serve, it is always larger or equal to the serving in an existing route if possible. The fixed costs of the vehicles help the model decide whether it is cheaper to serve a customer or a transfer in an existing route or use a new vehicle to serve it.

For all experiments, the reference values are calculated for each instance. Then, the user chooses a *water level significance (WLS or α)*, the relative importance of water level logistic costs compared to the street level cost. The values for the water level is updated as the multiplication of the reference values, $\left[c^s, c^w, \beta_s, \beta_w\right] = \left[1, \alpha, \underline{\beta}(1,\alpha), \alpha\underline{\beta}(1,\alpha)\right]$. The instances and models can be accessed online for future uses of the problems discussed and solved in this paper.[1]

5.2. Performance of the proposed methods

This study proposes two models to solve 2E-MVRP-SS with unitary transsipments: the joint MILP described in Section 3.2 and the LBBD method outlined in Section 4.3. Table 3 provides an overview of the proposed methods for analyzing the value of using a decomposed approach for highly complicated and integrated problems. We conduct tests on various problem sizes for each instance, considering 10, 20, or 30 demand nodes ($|C|$) and 4 satellites ($|P|$). *WLS* is set as 1:10 for the cost parameters, prioritizing the minimization of street-level logistic costs due to congestion-related issues addressed by the 2E-MVRP-SS in this study. First, the results for the joint MILP are presented. Then, the results for LBBD are provided, with the percent improvement of the solution quality compared to the joint MILP. Lastly, we provide averages across all instances within the same size for both methods.

The first set of instances with 10 demand nodes is easily solved by both methods. This confirms that the proposed LBBD can find optimal solutions for all instances with different time window structures, geographical distributions and cargo loads at the demand nodes. When the size of the demand network increases from 10 to 20, both methods face difficulties in solving the problems due to the complexity of the joint MILP and the inherent weaknesses of the underlying MDVRPTW, i.e., the master problem. However, LBBD manages to provide near-optimal solutions, improving the best-known solutions by 0.5% on average. The maximum optimality gap, 5.3%, is for the instance $R205$, a randomized network with tight time windows, where the Lagrangian bound is weaker for the global optimality. The performance of LBBD becomes particularly evident when analyzing the last set of instances, which consist of 30 demand nodes. LBBD provides an improved solution for 23 out of 27 instances, with improvement ranging from 0.7% to 19.4% while the joint method only improves two of them, $C201$ and $C207$. To maintain the best feasible lower bound, the LBBD ignores the solutions that do not improve the global lower bound but still can improve the upper bound. On average, it takes less than half the time compared to the joint model to achieve 5.5% improvement for all problems.

Initial solutions for 10-node problems are the optimal solutions for MDMVRPTW, the minimum cost schedules for the SL problem. However, for instances $R202$ and $R203$, it does not guarantee the global optimality. Furthermore, the initial solutions provide feasible solutions within a 0.2%, 3.4% and 5.7% gap compared to the joint MILP for the instances with 10, 20 and 30 demand nodes, respectively. Using the subproblems for checking up on the feasibility of the SL solutions indicates the potential of the proposed decomposition approach to provide better upper bounds within shorter times. These solutions are further improved by the LBBD using the Lagrangian lower bounds and consecutively analyzing more solutions to the SL problem toward the global optimality.

The proposed LBBD method offers an effective approach for tackling complex mixed integer combinatorial problems. It leverages existing knowledge to solve simpler problems at the master and subproblem levels.

5.3. Impact of cost coefficients

In designing an IWLT system, a system designer must assess the feasibility and cost-effectiveness of intermodal transportation, which

[1] https://github.com/cigdemkarademir/2echelon-synchronization.

Table 4
Performance with respect to the cost parameters and scale of the problem, |P| = 4.

|C|	WLS	Method	BK	Inc. time	Gap/Imp.%	Average travel distance			Fleets		Used satellites
						Total	Streets	Waterways	SL	WL	
10	1:10	Joint MILP	357.7	31.9	0.0	381.9	207.9	174.1	1.00	1.00	2.8
		LBBD	357.7	4.5	0.0	381.9	207.9	174.1	1.00	1.00	2.8
	1:5	Joint MILP	403.5	40.7	0.0	364.3	210.9	153.4	1.00	1.00	2.3
		LBBD	403.5	7.1	0.0	364.3	210.9	153.4	1.00	1.00	2.3
	1:1	Joint MILP	851.5	205.0	0.0	334.1	247.4	86.7	1.00	1.00	1.0
		LBBD	854.5	19.2	0.4	337.1	247.7	89.4	1.00	1.00	1.0
	10:1	Joint MILP	19,330.8	101.4	0.0	343.8	262.7	81.0	1.00	1.00	1.0
		LBBD	19,337.4	29.4	0.0	346.8	265.3	81.4	1.00	1.00	1.0
20	1:10	Joint MILP	722.7	5067.8	26.2	785.8	404.9	380.8	1.59	2.00	3.8
		LBBD	719.0	1025.7	−0.5	786.8	401.9	385.0	1.59	2.00	3.9
	1:5	Joint MILP	833.0	5910.1	26.6	773.6	407.8	365.8	1.59	2.00	3.6
		LBBD	833.0	2218.0	0.0	782.0	397.9	384.1	1.63	2.00	3.6
	1:1	Joint MILP	1961.4	7031.2	22.5	717.2	448.1	269.1	1.63	2.00	2.3
		LBBD	1939.8	2741.4	−1.1	723.6	447.6	276.0	1.56	2.00	2.3
	10:1	Joint MILP	51,167.3	6819.3	5.0	781.1	538.6	242.4	1.68	2.00	1.6
		LBBD	50,903.9	3198.9	−0.5	734.8	485.0	249.7	1.56	2.00	1.6
50	1:10	Joint MILP	1989.3	12,007.9	58.0	2017.3	1117.4	899.9	3.81	3.89	4.0
		LBBD	1941.8	5107.3	−2.4	1951.0	1099.4	851.6	3.70	3.59	4.0
	1:5	Joint MILP	2361.9	12,404.6	57.1	2051.6	1165.4	886.1	4.04	3.85	4.0
		LBBD	2250.5	4867.4	−4.7	1924.2	1101.4	822.8	3.89	3.59	3.9
	1:1	Joint MILP	5477.0	12,269.1	45.2	2096.5	1312.9	783.7	4.26	3.63	3.5
		LBBD	5021.8	5615.6	−8.3	1918.1	1173.1	745.0	3.70	3.59	3.7
	10:1	Joint MILP	124,438.2	13,248.9	12.5	2436.7	1703.7	733.0	5.04	3.59	3.1
		LBBD	119,488.7	8436.3	−4.0	2089.9	1428.7	661.2	3.67	3.59	2.8

BK: The best-known solutions to the methods, *Inc. Time*: The time it takes to find the best-known solution, *Gap/Imp.%*: The percent gap reported by the solver for the joint method/The percent improvement for the *BK* of LBBD compared to the *BK* of the joint MILP. Time limit is 4 h for both models. 27 instances in the test data are averaged for each scenario.

involves transferring goods from roads to waterways. Therefore, in this section, we analyze the relative importance of logistics costs on water level and street level.

To reflect the implications of multi-level hierarchical objectives in synchronized environments, an analysis of the relative importance of different levels' logistics costs is conducted from a methodological perspective. Different scenarios are created by adjusting the cost ratio between water level (WL) and street level (SL) cost parameters, which is referred to as *WLS* in Section 5.1. The scenarios include:

i. SL costs are significantly higher than WL costs (WLS = 1:10).
ii. SL costs are five times higher than WL costs (WLS = 1:5).
iii. A balanced scenario where WL and SL costs are equal (WLS = 1:1).
iv. WL costs are significantly higher than SL costs (WLS = 10:1).

Considering significant changes in the cost scenarios, the proposed methods are tested on instances with 10, 20, and 50 demand nodes for the convergence and the solution quality. The objective is twofold: to analyze the trade-offs associated with IWLT systems and to verify the LBBD method under the laid out experimental setting in Section 5.1. Hence, we present the average results of the methods in Table 4, first for the overall solution quality and then for each objective component including the number of used satellites in the solutions.

The results of 10-node scenarios verify that the LBBD method converges to solutions within 1% of the optimal values found by the joint MILP, across various cost configurations. For problems with 20 and 50 nodes, the LBBD method provides better solutions on average and reaches those solutions much faster than the joint MILP. LBBD is especially superior for the problems with 50 nodes, improving the total cost by 4.9% on average and achieving better metrics on both levels regarding the total travel and fleets. The results of the scenario with 20 nodes and WLS set to 1:5 indicate that the multiple solutions with the same objective value result in different schedules. Regarding fleet size minimization, LBBD can reduce street vehicles more than the joint model when WL significance increases. Across all scenarios, the decomposed model achieves the lower bounds for the vessels in terms of cargo load while the joint MILP struggles with fleet optimization on both levels when the size increases to 50 nodes

The master problem in LBBD is intentionally formulated in a simplified way to reduce the complexity of optimizing street-level operations to improve city logistics while seeking global optimality. It ignores the water level and synchronization problems, focusing solely on spatial synchronization costs for a given solution. Consequently, the convergence requires LBBD to explore all feasible routes for street-level operations while considering only the cheapest spatial synchronization. In contrast, the joint model can utilize the relationships between transfers and satellite assignments to aid in proving optimality. Nonetheless, LBBD remains a stronger method for finding feasible solutions within shorter computational times compared to the joint model.

Among all the analyzed scenarios, the balanced scenario with equal costs, WLS = 1:1, achieves the minimum average travel distance in total. The balanced scenario provides a lower bound on the cost of the IWLT system where fleets are first minimized, then the travel cost. This indicates that generating cost scenarios based on *WLS* can minimize a lexicographic objective of an IWLT system if the relative cost parameters are available. The LSPs can use it as a guide to prioritize different components of the objective. If the objective is to improve the logistics costs on the streets, they can reduce the *WLS* value meaning that WL costs are less significant than SL costs. Conversely, they can increase *WLS* value to prioritize the logistics cost over waterways for further reduction up to its lower bound.

The value of the *WLS* parameter further highlights the significance of integrated modeling for achieving global optimality. Traditional approaches often involve separate optimization problems that oversimplify the limitations of upstream or downstream processes. In the balanced scenario, as expected, the street or waterway distances are not minimized to their best possible values. Better solutions occur in different scenarios when the model significantly favors one component over the other. Disregarding the integration costs and solving problems independently at each level can result in asynchronous schedules for upstream or downstream operations.

Another observation is that significantly expensive WL operations lead the solution to locate the minimum number of vessels across all scenarios at the satellites without any visit or with very few visits between satellites. It converges to a two-echelon location routing problem (2E-LRP) where the most important decisions are to locate the satellites to visit. The average number of satellites used in WLS = 10:1 scenario is 1.0, 1.6, and 2.8 for the problems with 10, 20, and 50 nodes, respectively, choosing the same or as few satellites as possible for all transshipment operations compared to the number of the vessels used in the solutions. When SL is expensive, it uses almost all available satellites to further reduce the cost on the streets. These observations suggest the need for improvement in the relaxed cost formulation of the master problem considering explicit satellite assignments, particularly regarding spatial synchronization when the cost parameters prioritize the water level problem. Therefore, it is important to design the solution methods regarding the cost parameters and their effect on the system to understand the economic benefits of the IWLT systems better.

5.4. Impact of satellite locations

In this study, we assume that satellites are positioned on the outskirts of cities to minimize infrastructure investments and inconveniences associated with transfer operations. According to Crainic, Perboli, Mancini, and Tadei (2010), the maximum benefits of two-echelon systems are achieved when the satellites are situated between the central depot and the customers. Increasing the proximity of satellites to customers reduces the distances traveled to visit the satellites. However, it is still challenging to define the closeness in highly integrated and synchronized systems.

To analyze the extent of the benefits of the proposed IWLT system, we test the satellite locations by using different proximity values and *k-means* clustering. Proximity values indicate the distance between the satellites and the edges of the service area, as shown in Fig. 4. The values allow us to proportionally move the satellites along the radius toward the center and analyze the savings associated with locating transshipment operations closer to the city center, where the garage is generally located. We use the k-means algorithm to minimize the sum of distances between demand nodes and satellites by locating the satellites at the centroids of 4 clusters. The use of k-means is to assess the economic gains associated with placing satellites closer to the demand points, without considering the cost of the satellites, inconveniences related to transshipment operations, or the feasibility of reaching the waterways. The instances are solved by the LBBD method based on its performance on the cost analysis.

Fig. 4 presents a summary of the results obtained from the satellite location scenarios, provided with the percentage increase in the total distance costs compared to the base scenario, where proximity equals to 0. The results are categorized based on the demand distribution types to analyze the effectiveness of location policies and verify the LBBD method considering different service networks. For each scenario, the LBBD method achieves the lower bound on the required number of vessels, which are 4, 3, and 4 for C, R, and RC type problems, respectively. Overall, the k-means algorithm for satellite locations outperforms proximity-based locations by minimizing the travel time between satellite and customer visits on the streets. It reduces the total travel time more than any proximity value, achieving 21.3, 15.5, and 5.2% improvements on average for C, R, and RC type problems, respectively. On the other hand, the integrated systems, regardless of the demand distribution, work better when satellites are located at the outer rings instead of the inner rings. On average, across all problems, total travel time increased by 15.3, 6.6, 6.8, 14.0, and 21.2% for proximity values 0 to 0.4 compared to k-means. With the same number of vessels at the lower bounds, k-means achieves the best average street-level fleet at 3.61, but fleet requirements vary with travel time (3.76, 3.64, 3.87, 3.97, 4.39) for proximity-based locations. Among them, the proximity scenario with value 0.1 optimizes the problems better on average in terms of fleet size and overall travel time compared to k-means, likely due to its geographical superiority in capturing the spatial demand centroids for the problems solved in this study.

The costs of the proximity measures are proportional to the distances between the corresponding located satellites and k-means centers. For C-type problems, these centers lie between 0.1 and 0.3, where the best SL travel is achieved by k-means and the best fleet is achieved by the proximity value at 0.2. It also represents the closest scenario with respect to the k-means centers by locating the satellites in the middle ranges of these centers. The presence of randomness in the locations of demand nodes complicates the problem of locating satellites. In the existence of time windows and synchronization, ignoring the temporal distance between demand points might lead to several visits to the satellites at different points of the day. For R-type problems, where there are no clear centers in the demand network, the savings become marginally proportional to the savings in the SL distances. However, these changes affect the fleet size requirements by delaying the arrival time at the customers with time windows. On the other hand, RC-type problems present the demand networks where randomly distributed clusters exist. It becomes more significant to locate the satellites closer to k-means centers, lie between proximity values at 0 and 0.1. These proximity scenarios contain the k-means centers, and differ slightly in total distances. However, centralizing the satellites at 0.4 proximity value for these problems only deteriorates the logistics costs without improving the trip lengths.

5.5. Service design alternatives

This section presents and evaluates various design alternatives that have been considered in the literature and in practice. The aim is to obtain managerial insights for the possible implementation of IWLT systems for city logistics.

5.5.1. Alternative systems

Alternative systems are first classified according to the service network design of interest, analyzing systems with or without an integrated synchronized fleet as a secondary echelon. Next, we evaluate the impact of adopting new technologies in city logistics in comparison to the trucks primarily used by LSPs. For a fair comparison, we do not allow multiple trips for large vehicles (e.g., trucks) in the proposed alternative systems either as the multi-trip aspect of the vessels is left out of the scope in this study for the sake of simplicity.

Single-echelon systems: City freight vehicles are the sole resource of these services, operating between pickup demand points in cities and transferring freight goods at a central depot. The proposed model is modified to change the service network to single echelon by locating satellites at the central depot. We assess two vehicle type choices operating on the roads regarding their sizes and fuel types.

- *Only trucks:* Most of the current practices rely on large fossil-fueled vehicles, but LSPs now need to explore other options to comply with regulations. Therefore, one might consider the system with heavy trucks as a benchmark to analyze the trade-offs between current practices and alternatives.

  Only trucks system assumes the fleet is composed of cargo trucks as the SL freighters have limited access in cities due to the restrictions. Trucks serve many customers in a single trip and are

Table 5
Service design alternatives.

Alternative system	Single-echelon systems		Synchronized two-echelon systems	
	Only trucks	Only LEFVs	IWLT-Stationary	IWLT-Flexible
SL freighters (capacity)	Trucks (250)	LEFVs (50)	LEFVs (50)	LEFVS (50)
WL freighters (capacity)	–	–	Barges (250)	Vessels (250)
Sailing between satellites	–	–	**X**	✓
Location of satellites	Central depot (d)	Central depot (d)	Proximity = 0.1	Proximity = 0.1
Optimization problem	VRP	MVRP	2E-LRP-SS	2E-MVRP-SS

not allowed to perform multiple trips. To prevent multiple trips, the number of vehicles on the street is limited to the number of transfers.

- *Only LEFVs:* New technologies are promising to be viable and cost-efficient vehicles for LSPs to reach customers in the existence of restrictions. Therefore, one might consider the system with LEFVs to analyze the effect of new technologies with limited capacities. *Only LEFVs* system assumes the fleet is composed of LEFVs that are five times lighter than trucks in size. They can perform multiple trips but need to visit the central depot whenever necessary.

Two-echelon systems: LEFVs and vessels constitute the primary and secondary resources of the services. Together, they perform the first and last mile of the logistics service respectively within a synchronized IWLT system. We evaluate two vehicle type choices operating on inland waterways in terms of their operational costs for achieving spatial synchronization. The proposed two-echelon systems can also be applicable to other modes of transport for the integration of transshipment activities with existing transportation services. This could be achieved by adjusting the locations and time intervals of satellites to accommodate different types of vehicles such as trams, trucks, or trains. While the number of satellites increases the complexity for the joint MILP, it only affects the subproblem of the proposed LBBD.

- *IWLT-Stationary:* Traveling over waterways between satellites might be expensive, inefficiently slow, or restricted due to network capacity or safety reasons during the day. Then, one might consider a stationary vessel system that uses satellites as fixed depots and vessels as temporary safe storage spaces.
  The IWLT-Stationary system assumes that vessels are delivered to the best possible satellites before the operations and located there until the end of operations. We consider barges as the vessel type that can act well as storage spaces and be moved efficiently by tugboats. The problem minimizes the number of satellites used at least once for transshipment operations.
  The proposed model is modified to prevent any movements between satellites other than moving the barges to the satellites, by only letting positive f^s_{ij} from/to the central depot. Due to the capacity limitations requiring unitary transshipment at the satellites, we also set the fleet size to the lower bound to contain all the cargo load. In the cases of limited number of satellites compared to demand, multiple barges might be necessary at the satellites to handle the workload.
- *IWLT-Flexible:* The more flexible a water level transportation operates, the more an IWLT system achieves to address issues related to congested cities. Traveling over waterways might become easy and cheap enough thanks to the advancements in autonomous sailing and the high level of water network accessibility in populated cities such as Amsterdam, Brussels, New York, etc. (Janjevic & Ndiaye, 2014). Then, one might consider a flexible vessel system operating over water between satellites as a viable alternative to reduce the global cost of the system.
  The *IWLT-Flexible* system assumes that vessels are large electric vehicles that can visit any satellite at any time, as proposed in this study. Vessels act as mobile depots in contrast to the stationary system and sail between satellites.

5.5.2. Modeling and testing the alternative systems

The proposed methods can solve all the alternative systems by changing the service network and limitations accordingly. In each alternative scenario, all vehicles respect the latest return time defined in Section 5.1. For single-echelon networks, the transshipment operations are performed at the central depot by trucks or LEFVs only. Therefore, satellites are located at the central depot, where non-overlapping transshipment operations are executed for U units of time. However, there is no prioritization problem at the satellites due to the unitary transshipment if the number of satellites is more than or equal to the fleet size. Therefore, we model *Only trucks* and *Only LEFVs* systems as a VRP and MVRP respectively for the single echelon systems. For the two-echelon networks, it eliminates any movement between the satellites for the *IWLT-Stationary* case. Synchronized unitary transshipment operations ensure that the barges will be replaced when needed and the goods are stored at the barges one by one. Therefore, we model *IWLT-Stationary* and *IWLT-Flexible* systems as multi-trip 2E-LRP-SS and 2E-MVRP-SS for the two-echelon systems. Overall, these methods enable the evaluation of different design alternatives and assist decision-makers in optimizing service network design and operations with multiple objectives in mind.

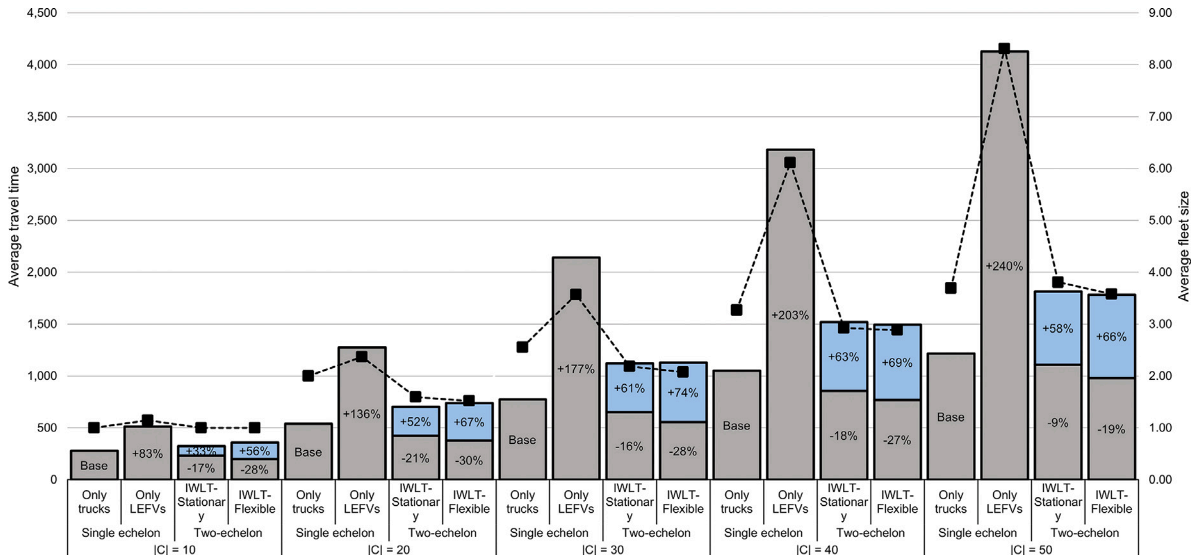
Test instances consist of different number of demand nodes, ranging from 10 to 50, and 4 satellites are located at the 0.1 proximity value for two-echelon systems, and at the depot for single-echelon systems. We assume that the larger vehicles (trucks, barges, sailing vessels) are five times bigger in storage capacity than LEFVs. Furthermore, WLS is set to 1:10 for the two-echelon systems to minimize the SL costs as much as possible. Table 5 summarizes the settings for the proposed alternative systems.

Based on the performances of the methods discussed in Section 5.2 and verification of the proposed LBBD on changing cost parameters, the alternatives are successively solved using the proposed LBBD as outlined for two-echelon systems. For single-echelon systems, spatial synchronization is included at the master level precisely considering the depot for satellite assignments. Hence, the cost of the master problem is always the same as the subproblem for a given solution if feasible. Infeasibility occurs due to the ignored synchronization constraints of the transfers at the satellites located at the depot. To address this, the subproblem is employed only to assess the feasibility of transfers at the depot and eliminate infeasible solutions accordingly.

5.5.3. Evaluation of the alternative systems

For the LSPs, the first goal of assessing different alternative designs is to switch from large trucks to economically viable lighter vehicles. Therefore, *Only trucks* system is compared to the three alternatives as the base scenario. The average results on the test instances are presented in Fig. 5. The data table in the figure provides the detailed costs on the streets for all alternatives and on the waterways if applicable. Additionally, we provide increases or reductions in traveled time on both levels as a percentage change compared to the base scenario. The values for the water level of the two-echelon systems represent the relative ratio of the traveled time on waterways to the traveled time on the streets of the base scenario, providing the extra logistics needed to be performed on the waterways.

Electrified fleets offer several advantages, including low operating, environmental, maintenance, and repair costs, excluding battery



	Only trucks	Only LEFVs	IWLT-Stationary	IWLT-Flexible	Only trucks	Only LEFVs	IWLT-Stationary	IWLT-Flexible	Only trucks	Only LEFVs	IWLT-Stationary	IWLT-Flexible	Only trucks	Only LEFVs	IWLT-Stationary	IWLT-Flexible	Only trucks	Only LEFVs	IWLT-Stationary	IWLT-Flexible
TW			92	156			278	359			468	573			664	723			707	803
TS	279	512	232	202	539	1,274	423	379	774	2,140	652	556	1,049	3,181	856	770	1,215	4,127	1,108	981
FW			1.00	1.00			2.00	2.00			2.56	2.56			3.27	3.27			3.58	3.58
FS	1.00	1.15	1.00	1.00	2.00	2.37	1.59	1.52	2.56	3.57	2.19	2.07	3.27	6.12	2.92	2.88	3.69	8.31	3.81	3.58

Fig. 5. Comparison of benchmark systems. For the number of customers $|C|$ ranging from 10 to 50, the "Only LEFVs", "IWLT-Stationary", and "IWLT-Flexible" systems are compared with the baseline "Only trucks". TS and TW represent average travel times, while FS and FW represent average fleet sizes on streets and waterways, respectively. The on-street movement was reduced across all two-echelon scenarios at the expense of increased total travel duration over the two levels. Time limit is 4 h for the LBBD method. 27 instances in the test data are averaged for each alternative system, up to 30 demand nodes. 26 instances are averaged for the scenarios with 40 and 50 demand nodes, where the instance $RC205$ becomes infeasible for single echelon systems under the test settings.

change. However, the high investment costs associated with acquiring electric vehicles hinder service providers from adopting these promising alternatives (Carrese, Colombaroni, & Fusco, 2021), hence they prioritize smaller and cheaper fleets. On a positive note, policymakers have been actively working on subsidies for electric vehicles in commercial applications. Besides promoting emission-free cities, the policies also favor fewer vehicles in urban areas. Comparing different scenarios, the *Only LEFVs* benchmark requires a larger fleet of vehicles on the streets in all instances. In contrast, the two-echelon systems achieve the same service level with fewer vehicles. Additionally, the *IWLT-Flexible* system further reduces the number of LEFVs needed. Furthermore, overall, *Only LEFVs* systems cause increases in total travel time by approximately 4 times of the two-echelon systems The integration allows LSPs to initially invest less in new vehicle technologies and support these vehicles with more conventional larger vehicles in less restricted zones.

The results presented in Fig. 5 highlight the significant advantages of IWLT systems in terms of total traveled time on the streets. When considering the system where only LEFVs are utilized without integration, the total traveled time on the streets increases by 83, 136, 177, 203, and 240% for varying problem sizes. In contrast, two-echelon benchmarks enable LEFVs to focus on providing the primary service in cities, while vessels are responsible for capacity replenishment. The *IWLT-Stationary* system achieves reductions of 17, 21, 16, 18, and 9% in the traveled time on the streets, while the *IWLT-Flexible* system enables even greater reductions of 28, 30, 28, 27, and 19%, respectively. By reducing vehicle kilometers driven, we can alleviate traffic congestion through mode shift from roads to waterways. This not only leads to cost savings but also enhances road safety and improves overall transportation efficiency.

While IWLT systems offer benefits such as fewer LEFVs and reduced vehicle kilometers on the streets, they require the coordination and deployment of vessels to support the replenishment and supply of LEFVs. This introduces additional logistical complexities and costs associated with managing vessel operations. However, IWLT can mitigate some of these challenges. I.e., *IWLT-Stationary* systems simplify the logistical aspects strategically locating barges at optimal satellites at the expense of flexibility. On the other hand, *IWLT-Flexible* systems allow vessels to dynamically navigate between any pair of satellites for adaptable supply chain operations at the expense of employing a flexible and reliable vessel fleet. As expected due to the consolidation opportunity of using trucks, moving from larger vehicles to lighter electric vehicles increases the total traveled time for the alternative benchmarks. However, the increase is more than double for *Only LEFVs* systems on average. For *IWLT-Stationary* systems, total travel time, as a sum of both levels, are 16, 31, 45, 45, and 49% more, whereas *IWLT-Flexible* systems require 28, 37, 46, 42, and 47% more in total compared to the base system. The increase in flexible systems compared to the stationary system is mainly intended by the objective function to reduce street movements further. Moreover, decreasing vehicle kilometers not only contributes to more economical freight transportation but also helps to address environmental concerns related to emissions per kilometer.

The summarized indicators suggest that IWLT systems can be a viable option in designing city logistics transportation systems to meet the increasing demand under different limitations in terms of access, fuel, and size of the vehicles in urban areas. While the integration reduces the costs of the transition with the help of coordination and synchronization, the flexibility in navigating over waterways improves the overall costs further compared to the stationary barges. LSPs can further improve the logistics costs of the IWLT systems by locating satellites within the service region using more customized clustering approaches. *IWLT-Flexible* systems can achieve more by letting vessels perform multiple trips on water level considering their flexibility in navigating, which is not included in this paper.

5.6. Results on large-scale instances

Up to now, several analyses have been provided regarding the proposed methods as well as design choices on the instances with up to 50 demand nodes. In this section, we evaluate the proposed models on large-scale instances to show their applicability to real-life problems where the real time capacities of the satellites are limiting the feasibility and the cost of the system.

The average results are summarized in Table 6. The column "zIP" presents the average of the best-known solutions to the methods for

Table 6
Comparison of the proposed models on large-scale instances. |C| = 100, and |P| = 4.

Type	Method	zIP	Travel time			Fleets			Transshipment operations		
			Total	SL	WL	SL	WL	Deviation from vessel LB	Transshipment per LEFV	Transshipment per vessel	Transshipment per satellite
C	Joint MILP	4271.3	4723.6	2212.4	2511.2	8.00	8.75	0.75	5.3	4.8	10.6
	LBBD	3857.2	4039.6	1958.5	2081.1	7.50	8.00	0.00	5.5	5.2	10.3
	Savings (%)	9.7	14.5	11.5	17.1	6.25	8.57				
R	Joint MILP	4158.9	4443.4	2182.8	2260.6	9.27	9.55	3.55	4.3	4.1	9.9
	LBBD	3720.9	3415.3	1997.7	1417.6	8.64	6.09	0.09	4.4	6.2	9.5
	Savings (%)	10.5	23.1	8.5	37.3	6.86	36.19				
RC	Joint MILP	5432.5	5452.9	2840.2	2612.7	9.50	9.25	2.25	4.8	4.9	11.4
	LBBD	4810.6	4451.6	1525.1	1970.6	8.75	7.88	0.88	5.0	5.6	11.0
	Savings (%)	11.4	18.4	46.3	24.6	7.89	14.86				
Average savings (%)		10.6	18.7	22.1	26.3	7.0	19.9				

Time limit is 4 h for both models. WLS = 1:10, and the proximity = 0.1. The average gaps reported by the Gurobi solver for the Joint MILP are 60.1%, 74.6%, and 77.0% for C, R, and RC type problems, respectively, averaged over 8, 11, and 8 instances in the test data.

the solution quality. For travel time, the averages are shown as travel time of the best-known solutions of the methods at both levels, SL and WL, respectively. The average fleets are provided for SL and WL, along with the fleet size deviation of vessels from the lower bound based on the total cargo load. Lastly, we report the average numbers of transshipment operations per LEFV, vessel and satellite. To compare the models, the average percentage savings are calculated for the LBBD model with respect to the joint MILP. Increasing the scale increases the number of minimum transfers to schedule at available satellites. These values are (37.0, 30.9, and 33.3) for the C, R, and RC type problems, respectively. For LBBD, the master problem is to find at least 30 trips, and the subproblem is to schedule transshipment operations for these trips at the resource-constrained satellites.

Both models can provide feasible solutions to the problems while LBBD provides significant savings compared to the joint MILP, namely an improvement of 10.6% is achieved in the objective on average across all instances. Even though the subproblems become larger, LBBD optimizes the WL fleet size closer to the lower bound in terms of cargo load. Besides, it provides better upper bounds for the other objectives by reducing SL fleet more and improving the total travel time by 18.7%. Additionally, the average savings in travel time for WL are significantly larger than those for SL, especially when randomization exists in the demand distribution by scheduling transshipment efficiently for WL services. Lastly, it improves the utilization of the resources allocated to the transshipment operations. It reduces the number of transfers with smaller fleets leading to more trips served by LEFVs and vessels on average while reducing the time spent at each satellite.

6. Conclusions

This study focuses on two-echelon synchronized logistic problems for integrated water- and land-based transportation systems. We have proposed two models, the joint model and the LBBD model, for solving the 2E-MVRP-SS, a novel rich variant of a two-echelon vehicle routing problem arising in city logistics. The models' performances are compared using instances of varying sizes. For smaller instances, the joint model obtains the best-known solution, indicating optimal or near-optimal performance. However, as the demand network size increases, the LBBD model outperforms the joint model almost across all instances. It finds improved solutions for a larger number of instances, suggesting its effectiveness in exploring the solution space. Additionally, on average, the LBBD model reduces the computational time required to find the best-known solution, indicating improved efficiency compared to the joint model.

Besides comparing the two models, we analyze the impact of cost parameters and the locations of satellites on the performance of the proposed IWLT system. Cost analysis involves adjusting the cost ratio between water level and street level cost parameters to create different scenarios. For different cost scenarios, on average, LBBD succeeds in finding better solutions compared to the joint model in less computational time. For satellite locations, we have explored proximity-based and clustering-based approaches, finding that the k-means algorithm provides the most savings in total traveled distances for clustered demand networks but offers limited improvements for randomized networks. Further benefits can be achieved by considering both spatial and temporal distances in satellite location decisions, particularly in scenarios with randomness in demand geographical distribution.

Additionally, different system alternatives are evaluated in the context of LSPs and their goal to transition from large trucks to economically viable lighter vehicles in urban areas. Comparisons are made between the *Only trucks* system and three alternatives, showing that IWLT systems significantly reduce total travel distances compared to the system without any integration. We have shown that flexible IWLT systems achieve even greater reductions in street travel distances ranging from 20% to 30% on average, providing cost savings, improved road safety, and transportation efficiency. Although IWLT systems introduce additional complexities and costs related to vessel operations, they offer viable solutions for urban logistics transportation under various limitations. Moreover, experiments on the large-scale instances show that the proposed LBBD can improve the costs by 10.6% while providing substantial reductions in fleet size by 7.0% and 19.9% for the SL and WL, respectively, as well as a reduction of 18.7% in total travel time on both networks.

In this study, we have proposed novel formulations for 2E-MVRP-SS with unitary transshipment capacities and demonstrated that the proposed LBBD method is an effective approach for solving such complex mixed integer combinatorial problems. It leverages existing knowledge to solve complicated problems iteratively in simple forms. The simple and compact formulation of the system can be used to consider different real-life settings such as service type, storage options at the satellites, and vehicle charging considerations. Further research can be devoted to enhancing the relaxed cost formulation of the master problem to potentially address the computational time issue and improve the performance of the LBBD model in scenarios where water level operations are of greater importance.

The outcomes also indicate that LBBD has the potential to facilitate the development of further heuristics enabling better resolutions for large-scale instances. Instead of using a single search tree in B&B, future studies can exploit the underlying decomposition structures to develop metaheuristics to explore the solutions using diversification strategies. Similarly, the underlying lagrangian bounding framework can provide the basis for designing subgradient optimization methods.

Recent studies take into account the minimization of total waiting times at the customers or at the satellites (Sluijk et al., 2022). Anderluh et al. (2017) suggest a method to decrease long waiting times by imposing bounds on the waiting time at any satellite specific to vehicle types. We allow vehicles to wait at no cost, thereby enabling the evaluation of services with minimal fleets and vehicle kilometers. Inconveniences related to parking the vehicles at the satellites or customers can be overcome by dedicating areas by the service designers. However, to fully understand the impact of the waiting times, further work is needed to collect and incorporate the cost parameters for unit waiting times on both networks regarding city regulations and also vehicle types (i.e., bikes, barges).

CRediT authorship contribution statement

Cigdem Karademir: Writing – review & editing, Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Breno A. Beirigo:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Bilge Atasoy:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Acknowledgments

This research is supported by the project "Sustainable Transportation and Logistics over Water: Electrification, Automation and Optimization (TRiLOGy)" of the Netherlands Organization for Scientific Research (NWO), domain Science (ENW). We specifically thank our collaborators from the Municipality of Amsterdam - Thomas Vernooij, Marcel Ludema and Thomas Meindertsma - for their inputs for potential applications of such two-echelon systems.

References

Ahat, B., Ekim, T., & Taşkın, Z. C. (2018). Integer programming formulations and benders decomposition for the maximum induced matching problem. *INFORMS Journal on Computing*, *30*(1), 43–56.

Anderluh, A., Hemmelmayr, V. C., & Nolz, P. C. (2017). Synchronizing vans and cargo bikes in a city distribution network. *Central European Journal of Operations Research*, *25*(2), 345–376.

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 Journal of Operational Research*, *289*(3), 940–958.

Baldacci, R., Mingozzi, A., Roberti, R., & Calvo, R. W. (2013). An exact algorithm for the two-echelon capacitated vehicle routing problem. *Operations Research*, *61*(2), 298–314.

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

Brauner, T., Kayle, M., & Pauwels, T. (2021). Market review on city freight distribution using inland waterways, publication in the framework of AVATAR, a project co-funded by the interreg North Sea Region programme 2014–2020 (ERDF).

Breunig, U., Baldacci, R., Hartl, R. F., & Vidal, T. (2019). The electric two-echelon vehicle routing problem. *Computers & Operations Research*, *103*, 198–210.

Bruni, M., Khodaparasti, S., & Moshref-Javadi, M. (2022). A logic-based benders decomposition method for the multi-trip traveling repairman problem with drones. *Computers & Operations Research*, *145*, Article 105845.

Caris, A., Limbourg, S., Macharis, C., Van Lier, T., & Cools, M. (2014). Integration of inland waterway transport in the intermodal supply chain: a taxonomy of research challenges. *Journal of Transport Geography*, *41*, 126–136.

Carrese, F., Colombaroni, C., & Fusco, G. (2021). Accessibility analysis for urban freight transport with electric vehicles. *Transportation Research Procedia*, *52*, 3–10.

Cattaruzza, D., Absi, N., Feillet, D., & González-Feliu, J. (2017). Vehicle routing problems for city logistics. *EURO Journal on Transportation and Logistics*, *6*(1), 51–79.

Chevalier, S. (2021). Global retail e-commerce market size 2014–2023. https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/. (Accessed 08 October 2021).

Côté, J.-F., Guastaroba, G., & Speranza, M. G. (2017). The value of integrating loading and routing. *European Journal of Operational Research*, *257*(1), 89–105.

Crainic, T. G., Perboli, G., Mancini, S., & Tadei, R. (2010). Two-echelon vehicle routing problem: a satellite location analysis. *Procedia-Social and Behavioral Sciences*, *2*(3), 5944–5955.

Crainic, T. G., Ricciardi, N., & Storchi, G. (2009). Models for evaluating and planning city logistics systems. *Transportation Science*, *43*(4), 432–454.

Dellaert, N., Dashty Saridarq, F., Van Woensel, T., & Crainic, T. G. (2019). Branch-and-price–based algorithms for the two-echelon vehicle routing problem with time windows. *Transportation Science*, *53*(2), 463–479.

Dellaert, N., Van Woensel, T., Crainic, T. G., & Saridarq, F. D. (2021). A multi-commodity two-echelon capacitated vehicle routing problem with time windows: Model formulations and solution approach. *Computers & Operations Research*, *127*, Article 105154.

Divieso, E., Lima, O. F., & De Oliveira, H. C. (2021). The use of waterways for urban logistics. *Theoretical and Empirical Researches in Urban Management*, *16*(1), 62–85.

Drexl, M. (2012). Synchronization in vehicle routing—a survey of VRPs with multiple synchronization constraints. *Transportation Science*, *46*(3), 297–316.

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 Journal of Operational Research*, *305*(1), 64–84.

Escobar-Vargas, D., Crainic, T. G., & Contardo, C. (2021). *Synchronization in two-echelon distribution systems: Models, algorithms, and sensitivity analyses*. Bureau de Montreal, Université de Montreal.

EU (2021). EU mission: Climate-neutral and smart cities. URL: https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe/eu-missions-horizon-europe/climate-neutral-and-smart-cities_en. June 9, 2023.

Fisher, M. L. (2004). The Lagrangian relaxation method for solving integer programming problems. *Management Science*, *50*(12_supplement), 1861–1871.

Fragkogios, A., Qiu, Y., Saharidis, G. K., & Pardalos, P. M. (2024). An accelerated benders decomposition algorithm for the solution of the multi-trip time-dependent vehicle routing problem with time windows. *European Journal of Operational Research*, *317*(2), 500–514.

Frangioni, A. (2005). About Lagrangian methods in integer optimization. *Annals of Operations Research*, *139*, 163–193.

Geoffrion, A. M. (1972). Generalized benders decomposition. *Journal of Optimization Theory and Applications*, *10*, 237–260.

Gonzalez-Feliu, J. (2008). *Models and methods for the city logistics: the two-echelon capacitated vehicle routing problem* (Ph.D. thesis), Politecnico di Torino.

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 Journal of Operational Research*, *254*(1), 80–91.

Groothedde, B., Ruijgrok, C., & Tavasszy, L. (2005). Towards collaborative, intermodal hub networks: A case study in the fast moving consumer goods market. *Transportation Research Part E: Logistics and Transportation Review*, *41*(6), 567–583.

Gurobi Optimization, L. (2021). Gurobi optimizer reference manual.

He, Z., & Haasis, H.-D. (2019). Integration of urban freight innovations: Sustainable inner-urban intermodal transportation in the retail/postal industry. *Sustainability*, *11*(6), 1749.

He, P., & Li, J. (2019). The two-echelon multi-trip vehicle routing problem with dynamic satellites for crop harvesting and transportation. *Applied Soft Computing*, *77*, 387–398.

Hooker, J. N., & Ottosson, G. (2003). Logic-based benders decomposition. *Mathematical Programming*, *96*(1), 33–60.

Huang, N., Li, J., Zhu, W., & Qin, H. (2021). The multi-trip vehicle routing problem with time windows and unloading queue at depot. *Transportation Research Part E: Logistics and Transportation Review*, *152*, Article 102370.

Janjevic, M., & Ndiaye, A. B. (2014). Inland waterways transport for city logistics: a review of experiences and the role of local public authorities. *Urban Transport XX*, *138*, 279–290.

Jepsen, M., Spoorendonk, S., & Ropke, S. (2013). A branch-and-cut algorithm for the symmetric two-echelon capacitated vehicle routing problem. *Transportation Science*, *47*(1), 23–37.

Karademir, C., Alves Beirigo, B., Negenborn, R., & Atasoy, B. (2022). Two-echelon multi-trip vehicle routing problem with synchronization for an integrated water-and land-based transportation system. In *HEART 2022: 10th symposium of the European association for research in transportation*.

Karademir, C., Beirigo, B. A., Negenborn, R., & Atasoy, B. (2021). Multi-trip vehicle routing problem with time windows for waste collection in amsterdam. In *Annual meeting of the transportation research board 2021 (virtual)*.

Karamyar, F., Sadeghi, J., & Yazdi, M. M. (2018). A benders decomposition for the location-allocation and scheduling model in a healthcare system regarding robust optimization. *Neural Computing and Applications*, *29*(10), 873–886.

Li, H., Chen, J., Wang, F., & Bai, M. (2021). Ground-vehicle and unmanned-aerial-vehicle routing problems from two-echelon scheme perspective: A review. *European Journal of Operational Research*.

Li, H., Liu, Y., Jian, X., & Lu, Y. (2018). The two-echelon distribution system considering the real-time transshipment capacity varying. *Transportation Research, Part B (Methodological)*, *110*, 239–260.

Li, H., Wang, H., Chen, J., & Bai, M. (2021). Two-echelon vehicle routing problem with satellite bi-synchronization. *European Journal of Operational Research*, *288*(3), 775–793.

Marques, G., Sadykov, R., Deschamps, J.-C., & Dupas, R. (2020). An improved branch-cut-and-price algorithm for the two-echelon capacitated vehicle routing problem. *Computers & Operations Research*, *114*, Article 104833.

Martínez, K. P., Adulyasak, Y., & Jans, R. (2022). Logic-based benders decomposition for integrated process configuration and production planning problems. *INFORMS Journal on Computing*.

Mhamedi, T., Andersson, H., Cherkesly, M., & Desaulniers, G. (2021). A branch-price-and-cut algorithm for the two-echelon vehicle routing problem with time windows. *Transportation Science*.

Moolenburgh, E., Van Duin, J., Balm, S., Van Altenburg, M., & Van Amstel, W. P. (2020). Logistics concepts for light electric freight vehicles: A multiple case study from the netherlands. *Transportation Research Procedia*, *46*, 301–308.

Pan, B., Zhang, Z., & Lim, A. (2021). Multi-trip time-dependent vehicle routing problem with time windows. *European Journal of Operational Research*, *291*(1), 218–231.

Paradiso, R., Roberti, R., Laganá, D., & Dullaert, W. (2020). An exact solution framework for multitrip vehicle-routing problems with time windows. *Operations Research*, *68*(1), 180–198.

Pfoser, S. (2022). Developing user-centered measures to increase the share of multimodal freight transport. *Research in Transportation Business & Management*, *43*, Article 100729.

Rahmaniani, R., Crainic, T. G., Gendreau, M., & Rei, W. (2017). The benders decomposition algorithm: A literature review. *European Journal of Operational Research*, *259*(3), 801–817.

Raidl, G. R., Baumhauer, T., & Hu, B. (2014). Speeding up logic-based benders' decomposition by a metaheuristic for a bi-level capacitated vehicle routing problem. In *International workshop on hybrid metaheuristics* (pp. 183–197). Springer.

Roshanaei, V., & Naderi, B. (2021). Solving integrated operating room planning and scheduling: Logic-based benders decomposition versus branch-price-and-cut. *European Journal of Operational Research*, *293*(1), 65–78.

Santos, F. A., da Cunha, A. S., & Mateus, G. R. (2013). Branch-and-price algorithms for the two-echelon capacitated vehicle routing problem. *Optimization Letters*, *7*(7), 1537–1547.

Sluijk, N., Florio, A. M., Kinable, J., Dellaert, N., & Van Woensel, T. (2022). Two-echelon vehicle routing problems: A literature review. *European Journal of Operational Research*.

Solomon, M. M. (1987). Algorithms for the vehicle routing and scheduling problems with time window constraints. *Operations Research*, *35*(2), 254–265.

Thorsteinsson, E. S. (2001). Branch-and-check: A hybrid framework integrating mixed integer programming and constraint logic programming. In *Principles and practice of constraint programming—CP 2001: 7th international conference, CP 2001 paphos, cyprus, November 26–December 1, 2001 proceedings 7* (pp. 16–30). Springer.

Yu, S., Puchinger, J., & Sun, S. (2020). Two-echelon urban deliveries using autonomous vehicles. *Transportation Research Part E: Logistics and Transportation Review*, *141*, Article 102018.