Two-echelon vehicle routing problems

Citation for published version (APA):

Document license:
CC BY

DOI:
10.1016/j.ejor.2022.02.022

Document status and date:
Published: 01/02/2023

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Invited Review

Two-echelon vehicle routing problems: A literature review

Natasja Sluijk\textsuperscript{a,*}, Alexandre M. Florio\textsuperscript{b}, Joris Kinable\textsuperscript{a,b}, Nico dellaert\textsuperscript{a}, Tom Van Woensel\textsuperscript{a}

\textsuperscript{a} School of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, P.O. Box 513, Eindhoven 5600MB, The Netherlands

\textsuperscript{b} Surface Research, Middle Mile Product Tech, Amazon, USA

A R T I C L E   I N F O

Article history:
Received 28 May 2021
Accepted 9 February 2022
Available online 23 February 2022

Keywords:
Routing
Two-echelon vehicle routing
Literature review

A B S T R A C T

In the two-echelon vehicle routing problem (2E-VRP), the distribution network is split into two echelons. Different vehicles are operated on the first and second echelon to maintain economies of scale and adhere to any vehicle restrictions that may be present in either echelon. Intermediate facilities are located at the borders of the echelons to facilitate the consolidation and transshipment of goods between echelons. Examples of two-echelon distribution systems include express delivery, grocery and hypermarket products distribution, multi-modal freight transportation, city logistics, and e-commerce and home delivery services. In recent years, the body of literature on the 2E-VRP has expanded significantly. Over 60 research papers have appeared in the scientific literature so far, which underlines both the academic and practical relevance of 2E-VRPs. In this review, we structure and revise all literature on the 2E-VRP. Mathematical formulations, exact and heuristic solution methods, and benchmark datasets used to test and evaluate new algorithms are reviewed and discussed. This survey concludes with a selected list of open research areas on 2E-VRPs.

© 2022 The Author(s). Published by Elsevier B.V.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

1. Introduction

In the classical vehicle routing problem, vehicles perform deliveries by starting at the depot, visiting one or more customers, and returning to the depot. In practice, there may exist physical or legal vehicle restrictions near customer locations, such as insufficient parking space or vehicle weight limitations. To maintain economies of scale and adhere to such requirements, in many cases it is beneficial to divide the distribution network into two echelons. With this setup, different vehicles, called urban vehicles and city freighters (Crainic, Ricciardi, & Storchi, 2009), or simply first-echelon (FE) and second-echelon (SE) vehicles, are used for the pick-up and delivery operations at the depot and customers, respectively. Intermediate facilities named satellites facilitate the consolidation and transshipment of goods between the different vehicles employed on either echelon (Guastaroba, Speranza, & Vigo, 2016).

The problem of routing vehicles in a two-echelon distribution network is known as the two-echelon vehicle routing problem (2E-VRP). The first formal definition of a 2E-VRP is given in Crainic et al. (2009), where the authors study a rich variant of a 2E-VRP problem with multiple products and depots, time-dependencies and vehicle synchronization. Perboli, Tadei, & Vigo (2011) coined the term 2E-VRP and introduced a mathematical model for the 2E-VRP with a single depot. In the 2E-VRP, FE vehicles transport goods from depots (warehouses) to satellites located at the border between the echelons. At the satellites, goods are unloaded from the FE vehicles and loaded into smaller SE vehicles that satisfy the SE requirements. Next, the SE vehicles perform the freight deliveries from the satellites to the customers. Fig. 1 illustrates a two-echelon distribution network.

A practical example of a two-echelon network is found in the context of city logistics, where the aim is to find efficient and effective ways to transport goods in urban areas while taking into account the negative impact on safety and environment (Crainic, Gendreau, & Gendron, 2021; Dolati Neghabadi, Evrad Samuel, & Espinouse, 2019; Savelbergh & Van Woensel, 2016). Finding and coordinating consolidation opportunities in city logistics is becoming increasingly important, not at least because of the increasing number of zero and low emission zones in many cities (see, e.g., Lurkin, Hambuckers, & Van Woensel, 2021; TNL, 2021; Transport Environment, 2018). An often-used approach is to consolidate volumes outside the city using urban distribution centers, from which the freight is transported into the denser urban areas using clean (e.g., electric or cargo bikes) and highly utilized vehicles. By adding a layer of satellites closer to the inner-city areas, the urban distribution network is extended to improve efficiency both upstream (from the large warehouses) and downstream (towards the dif-
ferent customers) (Savelsbergh & Van Woensel, 2016). Other examples of two-echelon distribution networks arise in applications such as express delivery, grocery and supermarket products distribution, multi-modal freight transportation, and e-commerce and home delivery services (Belgin, Karaoglan, & Altıparmak, 2018; Perboli et al., 2011).

The most recent literature review on the 2E-VRP is by Cuda, Guastaroba, & Sprenza (2015), where, in addition to the 2E-VRP, two other classes of related problems are included: the two-echelon location-routing problem and the truck-and-trailer routing problem. The amount of research published since Cuda et al. (2015) is impressive: over 40 papers on the 2E-VRP alone appeared. The primary goal of this paper is to review all work on 2E-VRPs, including real-world-inspired problem variants that have been studied to bridge the gap between the canonical models and practical applications.

Our focus is on the broad class of 2E-VRPs, where only routing decisions have to be taken. Before proceeding, however, we briefly mention a few related problems. In the two-echelon location-routing problem, the locations of the depots and satellites are not determined a priori and must be picked when constructing the routes (see, e.g., Darvish, Archetti, Coelho, & Sprenza, 2019; Nguyen, Prins, & Prodhon, 2012). The two-echelon inventory-routing problem (2E-IRP) includes decisions on the management of inventory levels at satellites and customers (see, e.g., Guimarães, Coelho, Schenekemberg, & Scarpin, 2019; Rohmer, Claassen, & Laporte, 2019). When the 2E-IRP also includes decisions on production quantities, the resulting problem variant is referred to as the two-echelon production-routing problem (Schenekemberg, Scarpin, Pecora Jr, Guimarães, & Coelho, 2021). An overview of vehicle routing problems with more than two echelons is given in Dondo, Méndez, & Cerdá (2011). Two other related problems are the truck and trailer routing problem (TTRP) and the vehicle routing problem with drones (VRPD). In the TTRP, a truck may temporarily park its trailer along its truck-and-trailer route to visit customers that can only be visited by truck (see, e.g., Parragh & Cordeau, 2017; Rothenbächer, Drexl, & Irnich, 2018). In the VRPD, drones are launched from trucks to serve some customers, and then return to (possibly different) trucks to recharge their batteries and collect packages for the next route (see, e.g., Li, Chen, Wang, & Bai, 2021a; Rojas Viloria, Solano-Charris, Muñoz-Villamizar, & Montoya-Torres, 2021).

This literature review is structured as follows. Section 2 formally defines the 2E-VRP, presents the main mathematical formulations of the problem, and elaborates on the different practice-motivated extensions of the 2E-VRP. In Section 3, we discuss exact algorithms for the 2E-VRP and its variants. Section 4 reviews the large body of work on heuristics for 2E-VRPs. In Section 5.3, we give a systematic overview of the benchmark instances used in the literature. Additionally, we compare the performance of various solution methods (both exact and heuristic) on these instances. Section 6 describes a range of applications that have been addressed in 2E-VRP literature. Lastly, Section 7 identifies future research directions and concludes this review.

2. Problem description

In this section, we formally define the two-echelon capacitated vehicle routing problem (2E-CVRP), which is the simplest problem in the class of 2E-VRPs (Section 2.1). Next, we discuss two commonly-used mixed integer programming formulations for the 2E-CVRP, namely, the arc-based (ABF, Section 2.2) and route-based (RBF, Section 2.3) formulations. Over the years, the classical 2E-VRP is extended to include, among others, multiple depots, multiple products, heterogeneous vehicles, time windows, stochastic demand and travel times, and temporal synchronization. These important extensions clearly increase the computational complexity and pose substantial algorithmic challenges. In Section 2.4, we discuss these variants of the 2E-VRP. Finally, in Section 2.5, we review the different objective functions proposed for the 2E-VRP and its variants.

2.1. Problem statement

The 2E-CVRP is formally defined on a directed graph \( G = (V, A) \) where vertex set \( V = \{0\} \cup S \cup C \) consists of a central depot 0, a set of satellites \( S \) and a set of customers \( C \). Products located at the depot are transported by a homogeneous fleet of FE vehicles to satellites located at the boundary of the first and second echelon. In turn, smaller SE vehicles start from the satellites and transport the goods on their final leg to the customer. Sets \( K_1 \) and \( K_2 \) will be used to represent FE and SE vehicles, respectively.

The arc set \( A \) of graph \( G \) can be partitioned into two disjoint sets of FE arcs \( A_1 = \{(i, j)| i, j \in \{0\} \cup S, i \neq j\} \) connecting the depot and the satellites, and SE arcs \( A_2 = \{(i, j)| i \in S \cup C, i \neq j \} \) \( \{(i, j)| i, j \in S\} \) between the satellites and the customers. By definition, the FE arcs in \( A_1 \) and SE arcs in \( A_2 \) can only be traversed by vehicles in \( K_1 \) and \( K_2 \), respectively. With each arc \( (i, j) \in A \), a non-negative travel cost \( c_{ij} \) is associated. It is generally assumed that these travel costs are symmetric and satisfy the triangle inequality.

The FE and SE vehicles have capacities \( Q_1 \) and \( Q_2 \), respectively. The number of SE vehicles that can depart from a satellite \( s \in S \) is limited to \( m_s \). A non-negative handling fee \( h_i \) is charged for every unit transhipped at satellite \( s \in S \). Finally, each customer \( i \in C \) has a demand \( d_i \). This demand must be fulfilled by a single delivery of an SE vehicle. In general, it is assumed that the FE vehicles have a larger capacity than the SE vehicles (\( Q_1 > Q_2 \)) and that the demand of each customer is less than the SE vehicle capacity (\( d_i < Q_2 \) for \( i \in C \)).

The objective of the 2E-CVRP is to find a set of FE and SE routes that minimize the total transportation and handling cost.
Collectively, the routes must meet vehicle and satellite capacity constraints and satisfy all customer demand. Additionally, the total load delivered to a satellite should equal the total demand of the customers served from that satellite (load synchronization).

The two mathematical programming formulations ABF and RBF, frequently used to model the 2E-VRP and its variants (see Table 1), are provided in the next sections.

2.2. Arc-based formulation

The first arc-based formulation for the 2E-CVRP was originally proposed by Gonzalez-Feliu, Perboli, Tadei, & Vigo (2008). However, Jepsen, Spoorendonk, & Ropke (2013) showed that this formulation is invalid when there are more than two satellites in the solution and propose a different formulation. In the follow-
ing, we provide a formulation similar to the one in [Jepsen et al. (2013)]. The model uses four types of decision variables. For the first echelon, binary routing variables $x_{ijk} \in \{0, 1\}$ denote whether a vehicle $k \in K_1$ traverses arc $(i, j) \in A_1$. Integer variables $u_{sk} \in \mathbb{Z}^+$ record the position of satellite $s \in S$ in the route performed by vehicle $k \in K_1$ and are used to eliminate subtours in the first echelon. The amount of freight transported from the depot to satellite $s \in S$ by vehicle $k \in K$ is denoted by $w_{sk} \in \mathbb{R}^+$. Similar to the $x_{ijk}$ variables in the first echelon, in the second echelon, binary routing variables $y_{ij} \in \{0, 1\}$ denote whether a vehicle dispatched from satellite $s \in S$ uses arc $(i, j) \in A_2$. The load of this vehicle is tracked by continuous flow variables $f_{ij} \in \mathbb{R}^+$. Observe that the FE routing variables require a vehicle index to distinguish between different vehicle routes since each satellite can be visited by multiple FE vehicles. For the second echelon, this is unnecessary since each customer is visited by at most one vehicle. The 2E-CVRP is formulated as follows:

\[
\begin{align*}
\text{(ABF) minimize} & \quad \sum_{k \in K_1} \sum_{(i, j) \in A_1} c_{ij} x_{ijk} + \sum_{s \in S} \sum_{(i, j) \in A_2} c_{ij} y_{ij} + \sum_{k \in K} \sum_{s \in S} h_i w_{sk}, \\
\text{subject to} & \quad \sum_{(i, j) \in A_1} x_{ijk} = \sum_{(j, s) \in A_2} x_{ijk}, \quad i \in \{0\} \cup S, k \in K_1, \\
& \sum_{(i, j) \in A_1} x_{ijk} \leq 1, \quad i \in \{0\} \cup S, k \in K_1, \\
& u_{sk} = u_{sk} + |S||x_{ijk} - |S| - 1, \quad i, j \in S, i \neq j, k \in K_1, \\
& w_{sk} \leq Q_1 \sum_{(j, s) \in A_2} x_{ijk}, \quad s \in S, k \in K_1, \\
& w_{sk} \leq Q_1, \quad k \in K_1, \\
& \sum_{i \in S} y_{ij} = 1, \quad j \in C, \quad s \in S, \\
& \sum_{i \in S} y_{ij} \leq m_s, \quad s \in S, \\
& \sum_{s \in S} x_{ijk} = 1, \quad j \in C, \quad i \in C, \\
& \sum_{i \in C} f_{ij} - \sum_{s \in S} f_{ij} = d_i, \quad i \in C, \\
& f_{ij} \leq Q_1 y_{ij}, \quad (i, j) \in A_2, s \in S, \\
& f_{ij} \geq 0, \quad (i, j) \in A_2, s \in S, \\
& u_{sk} \in \mathbb{Z}^+, \quad s \in S, k \in K_1, \\
& w_{sk} \geq 0, \quad s \in S, k \in K_1, \\
& x_{ijk} \in \{0, 1\}, \quad (i, j) \in A_1, k \in K_1, \\
& y_{ij} \in \{0, 1\}, \quad (i, j) \in A_2, s \in S.
\end{align*}
\]

The objective function minimizes total routing and handling costs. Constraints (1) - (5) relate to the first echelon. Constraints (6) - (11) ensure flow conservation for each FE vehicle at each satellite. Constraints (12) impose that each FE vehicle visits each satellite at most once. Constraints (13) eliminate subtours. Constraints (14) impose that the flow from the depot to satellite $s \in S$ by vehicle $k \in K_1$ can only be positive if vehicle $k \in K_1$ is used. Constraints (15) are the capacity constraints for the FE vehicles. Constraints (16) link the first and second echelon and impose that the total flow from the depot to satellite $s \in S$ equals the total demand served from satellite $s \in S$. Constraints (17) - (21) model the SE routes. Constraints (22) - (26) ensure flow conservation at each SE node. Constraints (27) limit the number of SE vehicles leaving each satellite. Constraints (28) and Constraints (29) ensure that each customer is visited exactly once and the customer demands are met. Constraints (30) also forbid the presence of subtours. Constraints (31) are the capacity constraints for the SE vehicles. Finally, Constraints (32) - (36) define the variable domains.

### 2.3. Route-based formulation

Originally proposed by [Baldacci, Mingozzi, Roberti, & Calvo (2013)], the 2E-CVRP can also be modeled with a route-based formulation. Let $R_1$ and $R_2$ denote the set of feasible FE and SE routes, respectively. We denote the subsets of routes visiting satellite $s \in S$ by $R_1(s) \subset R_1$ and $R_2(s) \subset R_2$ respectively. Conversely, let $S(r) \subset S$ denote the set of satellites visited by route $r \in R_1 \cup R_2$. For each route in $r \in R_1 \cup R_2$, let $A(r)$ be the set of arcs traversed and $c_r$ its cost, i.e., $c_r = \sum_{(i, j) \in A(r)} c_{ij}$. Finally, with each SE route $r \in R_2$, we associate a binary parameter $a_r$ that is equal to 1 if route $r$ visits customer $i \in C$ and a load $q_i = \sum_{s \in S} q_{is} d_i$. Decision variable $x_{r} \in \{0, 1\}$ is equal to 1 if route $r \in R_1$ is selected. Similarly, $y_{r} \in \{0, 1\}$ equals 1 if route $r \in R_2$ is selected. Finally, let $f_{rs} \in \mathbb{R}^+$ be the amount of freight that route $r \in R_1$ delivers to satellite $s \in S$. The 2E-CVRP can be formulated as follows:

\[
\begin{align*}
\text{(RBF) minimize} & \quad \sum_{r \in R_1} c_{r} x_{r} + \sum_{r \in R_2} c_{r} y_{r} + \sum_{s \in S} \sum_{r \in R_1(s)} h_i f_{rs}, \\
\text{subject to} & \quad \sum_{r \in R_1} a_r y_{r} = 1, \quad i \in C, \\
& \sum_{s \in S(r)} f_{rs} \leq Q_1 x_{r}, \quad r \in R_1, \\
& \sum_{r \in R_1(s)} f_{rs} = \sum_{r \in R_2(s)} q_{is} y_{r}, \quad s \in S, \\
& \sum_{r \in R_1} x_{r} \leq |K_1|, \\
& \sum_{r \in R_2} y_{r} \leq |K_2|, \\
& \sum_{r \in R_1} y_{r} \leq m_s, \quad s \in S, \\
& \sum_{r \in R_2} y_{r} \leq m_s, \quad s \in S, \\
& f_{rs} \geq 0, \quad r \in R_1, s \in S(r), \\
& y_{r} \in \{0, 1\}, \quad r \in R_2.
\end{align*}
\]

The objective function again minimizes the total handling and transportation costs. Constraints (17) impose that each customer is visited exactly once. Constraints (18) ensure that the flow from the depot to satellite $s \in S$ in route $r \in R_1$ is only positive if route $r \in R_1$ is selected. Additionally, if route $r$ is selected, the total load
delivered along the route cannot exceed the FE vehicle capacity. Constraints (19) are the load synchronization constraints. They link the first- and second-echelon routing problems and impose that the total flow delivered to satellite \( s \in S \) equals the total customer demand served from satellite \( s \). Constraints (20) and (21) ensure that the number of selected routes does not exceed the number of available vehicles. Constraints (22) limit the number of SE routes starting from each satellite. Finally, Constraints (23) - (25) define the variable domains.

### 2.4. 2E-VRP variants

Over the years, a large number of unique 2E-VRP variants have appeared in the literature. These variants are obtained by incorporating aspects encountered in real-life vehicle routing applications, such as time windows, pick-up and delivery operations, and multiple commodities. Fig. 2 shows the resulting richness of the 2E-VRP problem class. For a detailed overview, we refer to Appendix (Table A1) in the Appendix. In the remainder of this section, we discuss each attribute in greater detail.

#### 2.4.1. Number of depots

A common extension of the traditional 2E-VRP is to include multiple depots, where each depot typically has its own set of FE vehicles. If the problem involves multiple commodities, one or more commodities can be supplied from these depots. Crainic, Errico, Rei, & Ricciardi (2016), Zhou, Baldacci, Vigo, & Wang (2018), Dellaert, Van Woensel, Crainic, & Dashy Saridarq (2021), Jia, Deng, Zhao, & Chen (2022), and Kancharla & Ramadurai (2019) assume that each commodity is supplied by a unique depot. In Jia et al. (2022), each customer requests two commodities from two distinct depots. In contrast, in Gu et al. (2021) each depot distributes multiple commodities, but supplies per depot are limited.

#### 2.4.2. Satellite capacity

To enable the implementation of a two-echelon distribution system with minimal changes to the existing network infrastructure, satellites are often assumed to have no storage capacity. In this way, existing locations such as parking lots or vehicle depots can be used as satellite locations. However, physical space at such locations might be limited or only accessible during certain hours of the day (Crainic, Ricciardi, & Storchi, 2004). This restricts the number of FE and SE vehicles that can visit a satellite. Li, Liu, Jian, & Lu (2018) study the 2E-VRP in the context of parcel delivery services where the satellites have limited storage capacity. Additional constraints are introduced to keep track of the satellites’ real-time transshipment capacity.

#### 2.4.3. Split deliveries to SE vehicles

At the satellites, freight is offloaded from the FE vehicles and loaded unto the SE vehicles (cross-docking). Without any restrictions on the allocation of inbound freight to the SE vehicles, a single SE vehicle may receive freight from multiple FE vehicles. To avoid the complex synchronization needed when an SE vehicle has to wait for multiple FE vehicles, some explicitly require that each SE vehicle is supplied by a single FE vehicle (see, e.g., Dellaert, Dashy Saridarq, Van Woensel, & Crainic, 2019; Marques, Sadykov, Deschamps, & Dupas, 2020b; Sluijk, Florio, Kinable, Del-laert, & Van Woensel, 2021).

#### 2.4.4. Vehicle fleet

Although most works assume homogeneous first and second echelon vehicles, a limited number of publications consider heterogeneous vehicles (Bevilaqua, Bevilaqua, & Yamanaka, 2019; Crainic et al., 2016; Kancharla & Ramadurai, 2019; Kergosien, Lenté, Bil-laut, & Perrin, 2013; Wang & Wen, 2020), e.g., vehicles with different capacities. In some publications, specific types of vehicles are assumed, for example, cargo-bicycles (Anderluh, Hemmelmayr, & Nolz, 2017; Mühlbauer & Fontaine, 2021) and electric vehicles (Agárdi, Kovács, & Bányai, 2019; Breunig, Baldacci, Hartl, & Vidal, 2019; Jie, Yang, Zhang, & Huang, 2019; Li, Liu, Chen, & Lin, 2020; Wu & Zhang, 2021). Electric vehicles in particular come with additional challenges, as these vehicles typically have range and capacity limitations. The range of electric vehicles can be increased by explicitly incorporating charging decisions into the problem definition, e.g., by extending the routing graph with additional nodes representing battery swapping and/or recharging stations (Agárdi et al., 2019; Breunig et al., 2019; Jie et al., 2019; Wu & Zhang, 2021). Alternatively, to account for the battery capacity limitations, one could impose an upper bound on the maximum duration and/or length of the routes assigned to electric vehicles (Li et al., 2020). Similar restrictions are applicable when designing routes for SE vehicles that transport perishable products but lack adequate cooling (Esmaili & Sahraeian, 2017).
2.4.5. Time windows

Braeckers, Ramaekers, & Van Nieuwenhuyse (2016) observe that almost 40% of the literature on the single-echelon VRP includes time windows. Therefore, a natural extension of the 2E-VRP is the inclusion of time windows (2E-VRPTW). In most works, a single hard time window is associated with each customer to specify the earliest and latest service times. An exception can be found in Wang & Wen (2020), where the authors associate with each customer both a hard and a soft time window. While the hard time windows define absolute earliest and latest possible service times, a soft time window inside a hard time window defines preferred service times. A penalty cost is incurred when a customer is serviced outside its soft time window. Li, Zhang, Lv, & Chang (2016b) specify time windows at both the satellites and the depot.

Time windows bring additional algorithmic challenges. For example, time windows at customers impose deadlines on the earliest and latest possible departures of SE vehicles from the satellites, and consequently on arrival and departure times of FE vehicles at the satellites and depots, respectively.

2.4.6. Synchronization

When time is an aspect of the problem definition, temporal synchronization of the vehicles is required. Examples of such 2E-VRP problems involve, for instance, time windows (Dellaert et al., 2019), time-dependent travel times (Soysal, Bloemhof-Ruwaard, & Bektas, 2015), or the delivery of perishable products (Esmaili & Sahraeian, 2017). A consequence of temporal synchronization is that some vehicles potentially incur waiting times at the customers or satellites. To reduce the time spent at the satellites, Anderluh et al. (2017) incorporate bounds on the dwell times at the satellites. In Jia et al. (2022), the arrival times of FE vehicles with loads for the same SE vehicle are synchronized to avoid long waiting times of the SE vehicles. Synchronization is also required when the SE vehicles can perform multiple trips (Grangier, Gendreau, Lehuédé, & Rousseau, 2016), as the SE vehicles must be resupplied prior to the start of a new trip.

2.4.7. Delivery options

In the standard 2E-VRP, customers are exclusively served by SE vehicles. In Anderluh et al. (2017), Anderluh, Nolz, Hemmelmayr, & Crainic (2021), Song, Gu, & Huang (2017), and Wang et al. (2018), the authors relax this requirement by allowing FE vehicles to serve some customers directly from a nearby depot. The advantage of this approach is that direct deliveries from the depot to the customers reduce the solution cost if customers are located near the depot and offer more flexibility when solving the problem, which is helpful, in particular, when dealing with time windows and/or synchronization constraints (Perboli et al., 2011). In Anderluh et al. (2017), customers located outside of the city center have to be visited by FE vehicles (direct deliveries). Anderluh et al. (2021) introduce a third set of customers that can be visited by both FE and SE vehicles, thereby creating additional flexibility in the construction of the routes. Lastly, Yu, Jodiawan, Hou, & Gunawan (2021) include the possibility of crowd-shipping, where existing on-road crowds are employed to serve customers and compensated for any detour they have to take to perform the assigned deliveries.

Zhou et al. (2018), Entohen, Jargalsaikhan, Roodbergen, uit het Broek, & Schrotenboer (2020) and Yu et al. (2021) consider delivery to a pick-up location as an alternative to home delivery. Orders delivered to a pick-up location are collected by the customers whenever it suits them. An additional decision variable is added to the formulations to determine whether a customer is serviced at home or at a pick-up location. The pick-up locations are visited by SE vehicles in Zhou et al. (2018) and by FE vehicles in Entohen et al. (2020) and Yu et al. (2021).

2.4.8. Pick-up and delivery operations

In practice, vehicles may perform both pick-ups and deliveries (Belgin et al., 2018; Li, Wang, Chen, & Bai, 2021b; Zhou, Qin, Zhang, & Li, 2022). For example, products for the same customer may be consolidated on trolleys (containers) to reduce the loading and unloading times of the vehicles. Then, at the customer locations, the trolleys are unloaded from the vehicle and empty trolleys from previous deliveries are loaded. Allowing vehicles to perform pick-up and delivery operations raises additional algorithmic challenges, since the capacity of a vehicle is no longer monotonically increasing nor decreasing along the route.

2.4.9. Stochasticity

Although most papers on 2E-VRPs assume all data inputs to be deterministic and known in advance, one typically has to account for uncertainty in travel times and customer demand in practice. Crainic et al. (2016) consider the issue of building a tactical plan for a two-level city logistics system with uncertain customer demands (2E-VRPSD) and provide a formal definition of the problem. Wang, Lan, & Zhao (2017a) and Liu, Tao, Hu, & Xie (2017) use a two-stage stochastic program with simple recourse to formulate the problem. First, a priori FE and SE routes are constructed. As the routes are executed, customer demands are realized. Whenever customer demand exceeds the remaining vehicle capacity, a replenishment trip to the satellite and/or depot is performed. Suijk et al. (2021) propose a chance constraint formulation in which probabilistic capacity constraints ensure, with high probability, that the total demand of the customers served along an SE route does not exceed the SE vehicle capacity. Finally, Anderluh, Larsen, Hemmelmayr, & Nolz (2019) study the impact of stochastic travel times on the solution cost.

2.4.10. Other

Liu, Luo, Qin, & Lim (2018) require groups of customers to be visited by SE vehicles departing from the same satellite. Grangier et al. (2016) and Marques, Sadykov, Deschamps, & Dupas (2020a), among others, consider the multi-trip 2E-VRP, where the same SE vehicle can be used for different SE routes starting from the same satellite.

2.5. Objective functions

Various objective functions have been considered for the 2E-VRP. They can be classified into four categories: “arc-based”, “vehicle-based”, “handling-based”, and “other”. Below, we briefly discuss each category. An overview of the occurrences of the different objectives in the literature is included in Table 1.

2.5.1. Arc-based costs

Nearly all papers include an arc-based cost in the objective function. Often, simply a travel cost per arc is given without specifying how the costs are computed. Sometimes, details on the derivation of the travel costs are included. For example, Yu et al. (2021) incur a compensation fee per employed occasional driver, where the fee depends on the extra distance traveled by the driver. Several papers base the travel costs on the fuel consumption (Kancharla & Ramadurai, 2019; Paul, Kumar, Rout, & Goswami, 2021; Soysal et al., 2015; Wang, Shao, & Zhou, 2017b; Wang & Wen, 2020), i.e., the travel cost is computed as a function of distance, speed, vehicle load, and fuel cost. Time-dependent fuel costs are considered in Soysal et al. (2015).

Similarly, climate-related and socio-economic externalities such as carbon emission and disturbance (noise and congestion) can be taken into account (Anderluh et al., 2017; Wang & Wen, 2020). Anderluh et al. (2017) define the values for these externalities as a function of the vehicle’s travel distance.
Finally, some papers minimize driver costs which scale proportional to the total travel time (Anderluh et al., 2017; Anderluh et al., 2021; Soysal et al., 2015; Wang et al., 2017b) or minimize total route duration (Li et al., 2020; Paul et al., 2021).

2.5.2. Vehicle-based costs

The second-most considered objective is the minimization of the number of FE and SE vehicles in the solution, which is categorized as the vehicle-based component in Table 1. In most papers, this is achieved by adding a fixed cost per vehicle to the objective function. Others add the minimization of the number of vehicles as a separate objective to the objective function and impose an ordering among the different objectives, resulting in a lexicographic objective function (Grangier et al., 2016; He & Li, 2019; Yu, Puchinger, & Sun, 2020).

2.5.3. Handling-based costs

Another frequently considered component is related to the handling operations at the satellites and involves the inclusion of a satellite-specific handling cost per unit of freight transshipped from the FE vehicles to the SE vehicles.

2.5.4. Other

To conclude this section, we discuss objective components that do not belong to any of the previous categories. Anderluh et al. (2017, 2021) include driver costs per unit of time used for servicing customers and the loading of vehicles. Another time-related objective is the minimization of waiting times of the vehicles at the satellites and/or customers (Anderluh et al., 2017; Anderluh et al., 2021; He & Li, 2019; Li et al., 2020; Li et al., 2021b). When deliveries to pick-up locations are possible, connection/assignment costs for every delivery to a pick-up location are included (Enthoven et al., 2020; Yu et al., 2021; Zhou et al., 2018). In Jie et al. (2019), an echelon-specific cost is incurred for each visit to a battery swapping station. Finally, Wang & Wen (2020) consider a multi-objective function with a component that maximizes customer satisfaction.

3. Exact algorithms for 2E-VRPs

Research on exact algorithms for 2E-VRPs has mostly focused on solving the canonical 2E-CVRP variant. In this section, we review the exact algorithms proposed for the 2E-CVRP (Sections 3.1 and 3.2) and other variants (Section 3.3).

3.1. Branch-and-cut algorithms for the 2E-CVRP

The first generation of 2E-CVRP algorithms is based on the original flow formulation by Gonzalez-Feliu et al. (2008), which, as demonstrated by Jepsen et al. (2013), may not produce correct upper bounds in instances with more than two satellites. Therefore, the algorithms based on that formulation are only suitable for the particular case of two satellites. The branch-and-cut method proposed in Gonzalez-Feliu et al. (2008) strengthens the flow model with subtour elimination constraints (SECs) on SE routes and with tighter reformulations of Constraints (11) (also present in the original flow model), which considers flow reduction due to already served customers. This particular method is only effective for solving instances with up to 22 customers and two satellites. Perboli, Tadei, & Mascio (2010) introduce new valid inequalities by adapting known (single-echelon) VRP inequalities to the two-echelon formulation. The improved algorithm is able to solve instances with up to 32 customers and two satellites. Finally, Perboli et al. (2011) improve the original branch-and-cut method with tighter SECs, which are derived by considering customer-vehicle assignments, and report comparable results.

The first model and branch-and-cut algorithm for the 2E-CVRP that is able to handle an arbitrary number of satellites was introduced by Jepsen et al. (2013). The model adapts the flow formulation from Gonzalez-Feliu et al. (2008) and is described in Section 2.2. The branch-and-cut algorithm, however, is based on a different, relaxed model. Jepsen et al. (2013) observe that, when the assignment of customers to satellites is given, the 2E-CVRP decomposes into a split delivery VRP (SDVRP; see, e.g., Archetti, Bianchessi, & Speranza, 2014) in the first echelon and up to |S| capacitated VRPs in the second echelon. The relaxed model follows by combining the relaxed SDVRP formulation from Belenguer, Martinez, & Mota (2000) with a model for the location routing problem (LRP; Contardo, Cordeau, & Gendron, 2013). The relaxed model reduces the symmetry of the flow-based formulation and improves its continuous relaxation bound. Since the relaxed model may allow infeasible 2E-CVRP solutions, an auxiliary feasibility problem must be solved for each candidate upper bound. A special branching scheme is activated when the solution to the relaxed model is not feasible for the 2E-CVRP. The algorithm by Jepsen et al. (2013) employs the CVRPSEP package by Lysgaard, Letchford, & Eglese (2004) to separate several families of valid inequalities, such as rounded and framed capacity cuts, multi-star and strengthened comb inequalities. Overall, the algorithm effectively solves instances with up to 50 nodes and five satellites.

We also note that exact algorithms for the related but less studied 2E-LRP could be applied to solve the 2E-VRP. The latter is a special case of the former when satellites are fixed and uncapped. In particular, the branch-and-cut method from Contardo, Himmelmayr, & Crainic (2012), which is based on a decomposition of the two-echelon problem into two LRPCs, is able to solve 2E-LRP instances with up to 50 customers. However, a direct comparison between this algorithm and other specialized 2E-CVRP algorithms on the same instances is not available.

3.2. Decomposition-based algorithms for the 2E-CVRP

The next generation of 2E-CVRP algorithms is based on decomposition methods. Baldacci et al. (2013) propose a customized algorithm for solving the route-based model described in Section 2.3. The method initially enumerates the set $R_c$ of feasible FE routes and computes lower and upper bounds on the 2E-CVRP. Next, it generates a set of candidate FE solutions to the problem, where a candidate FE solution is referred to as a configuration and consists of a set of FE routes. For each promising configuration, an upper bound for the 2E-CVRP is computed by solving a multi-depot VRP (MDVRP) for the second echelon with the method by Baldacci & Mingozzi (2009). Lower and upper bound functional are applied throughout the algorithm to discard configurations that cannot lead to optimal solutions, leaving only a few MDVRPs to solve to optimality. The method can solve 2E-CVRP instances with up to 100 customers and five satellites.

Although Baldacci et al. (2013) apply a customized approach successfully, route-based models for VRP variants are typically solved by branch-and-price-(and-cut). For a survey of the methodology, we refer to Costa, Contardo, & Desaulniers (2019). For the 2E-CVRP, Santos, Mateus, & da Cunha (2015) propose a route-based formulation and a branch-price-and-cut (BP&C) algorithm that separates, also with the CVRPSEP package, rounded capacity cuts, multi-star, and strengthened comb inequalities. This method improves the previous algorithm from Santos, da Cunha, & Mateus (2013) by employing a formulation that better addresses symmetry. The results, however, are not competitive with the method from Baldacci et al. (2013), since most instances with 50 customers could not be solved.

The most recent and cutting-edge BP&C algorithm for the 2E-CVRP is proposed by Marques et al. (2020b). This algorithm
specializes the generic VRP solver by Pessoa, Sadykov, Uchoa, & Vanderbeck (2020) to the 2E-CVRP. This solver implements various state-of-the-art components such as ng-route relaxation (Baldacci, Mingozzi, & Roberiti, 2011), limited memory rank-1 cuts (Pecin, Pessoa, Poggi, & Uchoa, 2017), path enumeration (Baldacci, Christofides, & Mingozzi, 2008) and bucket graph labeling (Sadykov, Uchoa, & Pessoa, 2021), and is competitive for solving several VRP variants including the extensively studied capacitated VRP and the VRP with time windows. Marques et al. (2020b) present a new route-based formulation with fewer variables than the route-based model from Baldacci et al. (2013) and exponentially more constraints, which can be separated in polynomial time. Even though the method enumerates FE routes (as all other decomposition methods), the new formulation allows, at least in theory, the dynamic generation of FE variables. Two new families of valid inequalities, named satellite supply inequalities (SSIs) and visited satellite inequalities (VSLs), are also introduced, alongside a heuristic method for the separation of SSIs (VSLs are separated by enumeration). Furthermore, a seven-criteria strong branching scheme specifically designed for the two-echelon setting is applied. Marques et al. (2020b) report optimal solutions to all previous literature instances with up to 200 customers and ten satellites, and also solutions to new benchmark instances with up to 300 customers and 15 satellites.

3.3. Exact algorithms for other 2E-VRP variants

Motivated by a practical situation arising after the redesign of a distribution network, Liu et al. (2018) propose a 2E-VRP where groups of customers must be visited by SE vehicles departing from the same satellite. A compact formulation is presented together with five families of valid inequalities that explore the structural characteristics of this variant. Even though the cuts are shown to strengthen the linear relaxation bound, the method is only effective for solving instances with up to 40 customers and four satellites.

Breunig et al. (2019) study a prototypical electric 2E-VRP, where the SE routes are performed by electric vehicles with a limited driving range. In this variant, SE routes must also comply with a battery capacity constraint, and recharging stops along SE routes may be scheduled to recover battery capacity. A multigraph reformulation is proposed, in which multiple paths are allowed between two given vertices in the second echelon. Each path corresponds to a possible way of traveling between two nodes, assuming a maximum of one recharging stop. This reformulation is solved with an adapted version of the 2E-CVRP algorithm by Baldacci et al. (2013). Overall, the algorithm is only effective for solving instances with 22 customers, but tight lower bounds are obtained on larger instances with up to 75 customers.

The 2E-VRP with time windows and multiple depots (2E-VRPTW) is first considered in Dellaert et al. (2019). In addition to a network flow formulation, two path-based formulations are proposed and solved by branch-and-price-based procedures. The best performing algorithm is based on the two-echelon two-path formulation (2E-2P), where first- and second-echelon decisions are decomposed. The algorithm starts by enumerating all possible FE routes. Then, it dynamically generates FE solutions and solves the SE problem for fixed FE routes by column generation. The algorithm is able to solve instances with up to five satellites and 100 customers. Dellaert et al. (2021) extend the method by Dellaert et al. (2019) to handle multiple commodities. Mhamdi, Andersson, Cherkesly, & Desaulniers (2022) develop a BP&C algorithm for the 2E-VRPTW that features new methodological components: dual values stabilization by dual inequalities (see Ben Amor, Desrosiers, & Valério de Carvalho, 2006), and the separation of rounded capacity cuts, subset row cuts, and VSls. The BP&C method is able to solve all instances solved by Dellareit et al. (2019) and 41 previously unsolved instances with up to five satellites and 100 customers.

Marques et al. (2020a) propose a mixed-integer formulation for the 2E-VRPTW with an exponential number of route variables and precedence constraints. These constraints ensure that the net flow at satellites is always nonnegative, that is, the total inflow of goods from depots is never less than the total outflow of goods from satellites, at any instant. They propose a separation algorithm to identify violated precedence constraints, and a BP&C method based on strong branching and enumeration of elementary routes.

We remark that the 2E-2P model considered in Dellaert et al. (2019) and Mhamedi et al. (2022) places an assumption not present in the classical 2E-VRP models considered in the works described in Sections 3.1 and 3.2, namely, that an SE vehicle only receives load from a single FE vehicle. As also pointed out by Mhamedi et al. (2022), this assumption certainly simplifies models and algorithms. In practice, however, this restriction also leads to simpler operations at satellites (see the example of swap containers in Mühlbauer & Fontaine, 2021). Such a simple operation does not require storage, sorting, waiting times, nor synchronization between FE vehicles’ arrivals. On the other hand, the model by Marques et al. (2020a) allows freight consolidation at satellites and, hence, is applicable to more general settings.

Finally, Sluijk et al. (2021) consider a 2E-VRP with stochastic customer demands (2E-VRPSD). The problem is modeled as a chance-constrained stochastic optimization problem, in which SE routes must comply with the vehicle capacity constraint with high probability. Two exact column generation procedures are proposed to compute lower bounds on the 2E-VRPSD. These procedures employ new techniques such as a multi-label pricing algorithm, feasibility bounds, and Monte Carlo simulation to determine efficiently whether SE routes are feasible. The method is effective for computing exact lower bounds on instances with up to 75 customers.

4. Heuristics for 2E-VRPs

In the previous section, we detailed the performance of various exact algorithms and showed that solving medium-sized 2E-VRP instances is already computationally challenging. In practice, the instances can be much larger. To solve these instances, we need to resort to approximate methods that prescribe good solutions within a reasonable amount of time. This explains the large body of research on heuristics for the 2E-VRP and its variants.

Fig. 3 provides a classification of heuristics designed for 2E-VRPs. We split the neighborhood search class into two subgroups: local and large neighborhood search methods. All heuristics require an initial solution to improve upon. In Section 4.1, we review the methods proposed for constructing initial solutions to the 2E-VRP. Regardless of the heuristic proposed, local search is often included to improve solutions further. Therefore, we proceed with an overview of the local search operators in Section 4.2. In Sections 4.3 - 4.6, we review articles that propose local neighborhood search heuristics, large neighborhood search heuristics, genetic algorithms, and path relinking heuristics, respectively.

4.1. Initial solution

A common pattern in all 2E-VRP heuristics is to generate an initial solution by splitting the 2E-VRP into two separate subproblems, one for each echelon, and then solve the VRP related to the second echelon first. Crainic, Mancini, Perboli, & Tadei (2008) propose two alternative methods for solving the routing problem of the second echelon. In the first approach, customers are greedily assigned to the nearest satellite while taking the satellite capacity, if any, into account. Next, independent capacitated VRPs (CVRPs)
are solved for each satellite with its assigned customers. In the second approach, a multi-depot CVRP is solved.

Crainic, Mancini, Perboli, & Tadei (2011), among others, employ a greedy randomized adaptive search procedure (GRASP) to assign customers to satellites. GRASP is an iterative randomized sampling technique that can be used to construct initial solutions (Feo & Resende, 1995). In Crainic et al. (2011), an assignment probability for each customer and satellite pair is derived from the distance between the two. Next, a roulette wheel mechanism is applied to assign each customer to a satellite based on these probabilities. In this way, multiple initial solutions can be generated.

Different heuristic methods are employed to solve the multiple CVRPs or multi-depot CVRP related to the second echelon, including nearest neighbor (Belgin et al., 2018), cheapest insertion (Liu et al., 2017), Clarke and Wright savings heuristic (Hemmelmayr, Cordeau, & Crainic, 2012), GRASP (Zeng, Xu, Xu, & Shao, 2014), and random insertion (Mühlbauer & Fontaine, 2021).

After solving the routing problem related to the second echelon, the demand of each satellite is computed as the total demand of the customers served by all SE routes departing from that satellite. Next, the FE routes are constructed. Given the relatively small number of depots and satellites, in many methods the FE routing problem is solved using exact methods (Amarouche, Guibadj, & Moukrim, 2018; Gu et al., 2021; Jie et al., 2019; Wang et al., 2017b). Other methods employ heuristics such as nearest neighbor (Belgin et al., 2018), cheapest insertion (Liu et al., 2017), Clarke and Wright savings heuristic (Hemmelmayr et al., 2012), and random insertion (He & Li, 2019), Zhou et al. (2018) and Mühlbauer & Fontaine (2021) build the FE solution by creating a giant tour visiting all satellites and using a split algorithm to obtain feasible FE routes.

Breunig, Schmid, Hartl, & Vidal (2016) propose a preprocessing step specifically for the situation where satellites have a demand that exceeds the capacity of a single FE vehicle. Any satellite with a demand greater than the FE vehicle capacity is served by direct (back-and-forth) trips from the nearest depot until the remaining demand is less than the FE vehicle capacity. Next, a cheapest insertion heuristic is applied to construct FE routes serving the remaining demand of the satellites.

### 4.2. Local search

Local search is often included in heuristics to intensify the search around promising solutions. It starts from a candidate solution and iteratively evaluates neighbor solutions, which are obtained by, for example, relocating a customer to a different position within a route. Nearly all articles employ a first-improvement policy, where the first improving solution is accepted as the next candidate solution. Only Belgin et al. (2018), Liu et al. (2017), and Li et al. (2018) employ a best-improvement policy, where all possible moves within the neighborhood are computed, and the solution with the best objective value is selected as the next candidate solution.
Table 2 provides an overview of the local search operators employed by 2E-VRP heuristics. The first four operators in the table are also commonly used by (single-echelon) VRP heuristics. The relocate and exchange operators are applied to single customers and groups of up to three consecutive customers. The 2-opt and 2-opt\(^*\) are intra-route and inter-route operators, respectively, and exchange links between two pairs of consecutive nodes. The four operators are mainly applied to SE routes only, but sometimes also to the FE routes (Enthoven et al., 2020; Liu et al., 2017; Mühlbauer & Fontaine, 2021).

Mühlbauer & Fontaine (2021) and Hemmelmayr et al. (2012) only apply operators on SE routes departing from the same satellite to ensure that the satellite demands do not change, thereby avoiding the need to recompute FE routes. Additionally, both articles only perform local search on solutions that are within a pre-specified threshold of the previously accepted solution and are therefore promising. Breunig et al. (2016), Breunig et al. (2019) and Zhou et al. (2018) only allow moves with the relocate operator within a limited neighborhood of the \(k\) closest nodes.

Zeng et al. (2014) propose two new operators specific for the 2E-VRP: satellite change and satellite swap. The satellite change operator replaces a satellite of an SE route with another satellite, while the satellite swap operator exchanges the satellites of two SE routes that are assigned to different satellites. Both operators may modify the demands of the satellites and consequently may require changes to the FE routes. Other local search procedures applied to the 2E-VRP include the reassignment of customers to their second nearest satellite (Crainic et al., 2011; 2013), a Lin-Kernighan local search procedure, (Bevilaqua et al., 2019), a split procedure (Hemmelmayr et al., 2012), and moves related to delivery options (Zhou et al., 2018). Breunig et al. (2016) investigate the performance of these operators. They indicate that removing the 2-opt operator has little impact on the solution quality, whereas removing the 2-opt\(^*\) operator has a more substantial impact, especially on larger instances. Additionally, Breunig et al. (2016) point out that the relocate operator is essential for instances with more than 50 customers.

4.3. Local neighborhood search methods

In this section, we discuss the following local neighborhood search methods: variable neighborhood descent (VND), variable neighborhood search (VNS), and tabu search. After a short description of each of these methods, we proceed with reviewing the corresponding articles. In local neighborhood search methods, local search operators are employed to explore the neighborhood of a solution. In VND, different neighborhoods are considered sequentially. Depending on the policy, the first-improving or the best-improving solution in the neighborhood is selected as the new solution, and the search continues until none of the neighborhoods produce a new solution, i.e., a local optimum is found. In VNS, an additional shaking step is included to escape these local optima. In this step, a random solution from the current neighborhood is selected. Next, VND is applied to obtain a local optimum. This procedure is repeated until a maximum number of iterations is reached. A drawback of the shaking step in VNS is that previously visited solutions could be revisited. Tabu search algorithms are particularly designed to avoid these cycles and employ a tabu list with attributes representing moves that are prohibited for a number of iterations; for example, in the 2E-VRP context, a combination of a satellite and a customer. For more details on these methods, we refer to Talbi (2009).

Belgin et al. (2018) propose a hybrid heuristic combining VND and local search. The heuristic starts with an initial solution and considers the neighborhoods sequentially. In each neighborhood, it selects the best solution as the next candidate solution. When the
termination criteria of the VND are met, the best solution is further improved by local search, using local search operators that are different from the ones used in the VND. Zeng et al. (2014) iteratively generate an initial solution with GRASP and search for improvements with VND. The procedure terminates when the maximum number of iterations is reached and the best solution found is returned.

Wang et al. (2017b) propose a matheuristic that combines the search abilities of VNS with integer programming. In this matheuristic, solutions to the SE routing problem are generated with a VNS algorithm. Upon termination of the VNS, an arc-and-route-based model is solved to find better SE solutions missed by the VNS algorithm and to construct least-cost FE routes. The arc-and-route-based model consists of an arc-based formulation for the first echelon and a route-based formulation for the second echelon. Wang et al. (2017b) show that solutions with larger travel distances are obtained when minimizing the drivers’ wage and fuel cost rather than total distance. By increasing the travel distance, congested areas can be avoided and travel times can be reduced, leading to lower drivers’ and fuel expenses. Similarly, Amarouche et al. (2018) propose a matheuristic that combines neighborhood search with a set covering formulation to obtain the best combination of routes.

Tabu search is implemented by Liu et al. (2017) to solve the 2E-VRP with stochastic demands. Two types of attribute sets (moves) are included in the tabu list. The first set consists of pairs of satellites and FE vehicles, whereas the second set contains triplets of customers, satellites, and SE vehicles. Note that the second set is specific to the 2E-VRP since it takes the customer-satellite assignment into account. Local search operators are applied to both FE and SE routes. The stochastic nature of customer demand complicates the evaluation of a neighborhood change. Instead of computing the cost exactly, it is estimated through Monte Carlo simulation. The algorithm’s performance is compared to a variation of the proposed algorithm based on expected demand values. Liu et al. (2017) show that the former outperforms the latter on all small instances and most of the larger instances.

Zhou et al. (2022) propose a variable neighborhood tabu search algorithm to solve the 2E-VRP with time windows and simultaneous pick-up and delivery. They consider two local search operators that can be applied to the FE delivery routes, the SE pick-up and delivery routes, and the FE pick-up routes. To facilitate the exploration of the search space, infeasible solutions are allowed and a penalty is incurred for the violation of time windows and vehicle capacities. In each iteration of the algorithm, two neighborhood operators are selected at random to generate neighborhood solutions. The best neighborhood solution is selected as the new solution and the changes from the previous to the new solution are tabu for a number of iterations. Their experiments demonstrate that the inclusion of satellite time windows (derived from the time windows of the customers assigned to the satellite) accelerates the algorithm.

4.4. Large neighborhood search methods

In large neighborhood search (LNS), new solutions are obtained by partially destroying an existing feasible solution through a so-called destroy operator and repairing it with a repair operator. At each iteration, one pair of destroy and repair operators is selected with a roulette wheel mechanism. In adaptive LNS (ALNS), the probability of each operator being selected depends on its performance in previous iterations (Ropke & Pisinger, 2006). Most operators are applied to the SE routes only but also have an impact on the FE routes. After an iteration of (A)LNS, the FE routes are often reconstructed using the methods discussed in Section 4.1. In the remainder of this section, we discuss the destroy and repair operators, how they are selected, and how well they perform. We also examine the different acceptance criteria for new solutions.

4.4.1. Destroy operators

Table 3 provides an overview of the destroy and repair operators. The first three destroy operators were originally proposed for the single-echelon VRP (Ropke & Pisinger, 2006). For a given integer q, these operators remove q customers at random; select one customer and remove the q − 1 closest customers; or remove the q customers with the highest removal gain, where the gain associated with each customer is the difference between the objective values of the solutions with and without the customer.

Hemmelmayr et al. (2012) proposes four destroy operators specific to the 2E-VRP. The random route removal operator removes a random SE route from the solution. The satellite removal operator randomly selects a used satellite in the solution, and “closes” it; that is, the satellite cannot be reopened again by the repair operator in the same iteration. The satellite opening operator randomly selects an unused satellite and removes the q customers closest to this satellite from their current routes to encourage the use of the satellite by the repair operator. Finally, the satellite swap operator replaces a used satellite with an unused satellite selected with the roulette wheel mechanism. The probability of each satellite being selected is inversely proportional to the distance to the removed satellite. In this way, the satellite is replaced by a nearby satellite and the impact on the solution is reduced. Nevertheless, the satellite operators change the solutions substantially. To ensure that a new combination of satellites is fully explored, the satellite operators are only allowed to be used after x iterations without improvement (Hemmelmayr et al., 2012) or after x iterations since a previous change to the satellites (Breunig et al., 2016). In Enthoven et al. (2020) and Yu et al. (2021), the satellite operators are also applied to the covering locations.

Instead of including an unused satellite with the satellite opening operator, Breunig et al. (2016) and Breunig et al. (2019) allow all currently unused satellites to be used by the repair operators. Grangier et al. (2016) consider the multi-trip 2E-VRP and propose destroy operators specifically related to the multi-trips. Amarouche et al. (2018) vary the number of customers to be removed with the node removal operators to intensify the search around promising solutions and increase diversification as the search converges towards a local optimum. Other unique destroy operators applied to the 2E-VRP include biased node removal (Breunig et al., 2016), minimum quantity removal (Enthoven et al., 2020), least used vehicle removal (Enthoven et al., 2020), route redistribution (Hemmelmayr et al., 2012), related order removal (Li et al., 2021b), node neighborhood removal (Yu et al., 2021), perturbed worst and related node removals (Yu et al., 2021), and removal of routes containing just a single customer (Breunig et al., 2019; Breunig et al., 2016; Paul et al., 2021).

4.4.2. Repair operators

The repair operators in Table 3 that were originally proposed for the single-echelon VRP do not require any modifications before they can be applied to the SE routes. Variants of existing repair operators and new repair operators have been applied, including the greedy insertion perturbation operator (Enthoven et al., 2020; Hemmelmayr et al., 2012; Kancharla & Ramadurai, 2019; Yu et al., 2021), greedy insertion regret operator (Enthoven et al., 2020), first feasible position operator (Yu et al., 2020), best insertion operator (Amarouche et al., 2018; Anderlhub et al., 2021), greedy insertion forbidden operator (Hemmelmayr et al., 2012), and repack insertion operator (Mühlbauer & Fontaine, 2021).
4.4.3. Selection of destroy and repair operators

Different procedures have been implemented to select destroy and repair operators in each iteration of the LNS algorithm. Breunig et al. (2016) sequentially apply the node- and route-related destroy operators in each iteration. In contrast, the satellite-related destroy operators are only invoked when there has not been a change to the set of satellites visited in the solution for a number of iterations. Yu et al. (2020) employ only one destroy and one repair operator. The remaining articles on LNS select (some of) the operators at random, where each operator has an equal probability of being selected.

In ALNS, the selection of operators is based on each operator’s performance in previous iterations. A score is associated with each operator and is either updated with some fixed value $\sigma$ regardless of the quality of the new solution (Hemmelmaier et al., 2012) or with values $\sigma_1$, $\sigma_2$, or $\sigma_3$ in case a new global best solution is found, the current solution is improved while the global solution is not, or the solution is accepted without improving the solution value, respectively (Enthoven et al., 2020; Jia et al., 2022; Kancharla & Ramadurai, 2019). The probability (weight) of selecting a destroy (repair) operator is set equal to the operator’s score divided by the total score of all destroy (repair) operators. Enthoven et al. (2020), Li et al. (2021), and Yu et al. (2021) take a reaction factor into account when updating the weight of an operator to control how fast the weight adjustment procedure reacts to changes in the scores. Additionally, Li et al. (2021) define various values for $\sigma_1$ and $\sigma_2$ depending on the percentage gap between the new best solution and the current best solution and the gap between the new solution and the current solution, respectively.

4.4.4. Performance of destroy and repair operators

Common metrics used to analyze the operators’ performance are counting the average number of times that applying the operator produced a better solution and/or comparing the objective values obtained when solving with and without the operator. Hemmelmaier et al. (2012) investigate the performance of the operators and conclude that all operators contribute to obtaining new solutions. However, some of them are more successful than others. Hemmelmaier et al. (2012) identify the regret insertion and route removal as the most successful repair and destroy operators, respectively. Breunig et al. (2016) point out that some operators perform better on smaller instances, while others perform better on the larger instances.

Focusing on the 2E-VRP specific operators, we recall that the satellite operators are included for diversification purposes and strongly impact the solution quality. Numerical results in Breunig et al. (2019) show that the elimination of certain satellite choices through the satellite removal operator at different phases of LNS is essential to obtain structurally different solutions. At the same time, opening satellites is necessary to control the frequency of exploration of different satellite configurations. Thus, the satellite operators proposed by Hemmelmaier et al. (2012) are important to be included when solving the 2E-VRP with (A)LNS.

4.4.5. Acceptance criteria

New solutions are always accepted when they improve upon the current solution. If a solution is non-improving, its acceptance depends on the selection criterion. The most frequently used criterion is based on simulated annealing (SA), where the acceptance probability depends on the quality of the solution and an initial temperature (Enthoven et al., 2020). Mühlbauer & Fontaine (2021) propose a parallelized LNS with different fixed temperatures. The algorithm starts with the creation of $N$ initial feasible solutions. Each feasible solution initiates a tree with $B$ branches, and each branch applies LNS for a given number of iterations. Instead of continuously decreasing the temperature for the SA crite-
rion, the $B$ branches starting from the same solution are assigned equidistant temperature values ranging from zero up to some maximum value. The temperature is held constant for all iterations within the branch. The information among the different branches is shared by randomly reassigning the $N \cdot B$ solutions obtained into $N$ groups of $B$ solutions, and the best solution in each group is selected as the starting solution for the next iteration. Results show that this sharing of information is beneficial for larger instances.

Other methods only allow solutions that improve upon the current solution. Local optima are escaped by restarting the algorithm with a new initial solution when a maximum number of iterations is reached (Breunig et al., 2019; Breunig et al., 2016). Another possible criterion is to accept only solutions that are no more than a prespecified percentage away from the best solution found (Hemmelmayer et al., 2012; Kancharla & Ramadurai, 2019; Wang et al., 2017b).

4.5. Genetic algorithms

Genetic algorithms (GA) mimic the natural evolution of biological organisms (Talbi, 2009). First, an initial pool of solutions is generated. Each solution is encoded into a chromosome, and together, they form a so-called population. Next, chromosomes are selected to enter the evolution process. New solutions (offsprings) are generated from pairs of chromosomes (parents) through genetic crossover and mutation of the offspring, resulting in a new population. The evolution process is repeated until a stopping criterion is met. We identified nine articles that use GA to solve the 2E-VRP. In the following, we compare the definitions of the chromosomes and fitness functions, the parents’ selections, and the creations of offspring.

Wang et al. (2017a) propose a GA where each chromosome represents a unique SE solution, and each gene in a chromosome represents an SE route. The gene (route) length and the chromosome length depend on the number of customers visited in each SE route and the number of SE routes used in the solution. The FE routes are created using the procedure proposed by Breunig et al. (2016) as detailed in Section 4.1. The remaining articles define chromosomes representing both FE and SE routes. Zhou et al. (2018), among others, use separate chromosomes for the FE and SE routes and an additional chromosome to link the two levels.

The initial population consists of solutions generated with construction heuristics. Fitness functions are defined to compute a score for each solution in the population. Similarly to the fitness functions specified for single-echelon VRPs, they consist of the objective value and a penalty term. The penalty term is included to take into account violations of constraints in the crossover and mutation operations. With the exception of the GAs by He & Li (2019) and Agárdi et al. (2019), all other algorithms allow infeasible solutions and include penalties to deal with temporary infeasibilities.

Parent selection criteria considered for the 2E-VRP are random, tournament, roulette wheel, or best fitness selection. After the chromosomes for reproduction have been selected, crossover operators are applied to pairs of chromosomes (parents) to generate new offspring. Different crossover operators are applied, including the order-based crossover operator (Agárdi et al., 2019; He & Li, 2019), the route copy operator (Bevilaqua et al., 2019; Wang et al., 2017a), the cycle crossover operator (Agárdi et al., 2019), the partially mapped crossover operator (Agárdi et al., 2019; Wang et al., 2018), the 2-point crossover operator (Sahraeian & Esmaeili, 2018; Wang & Wen, 2020), and a crossover operator-specific for the 2E-VRP with delivery options (Zhou et al., 2018).

After the crossover operators, mutation operators are applied to some of the offspring to increase the diversity of the population. In most articles, an additional local search procedure is considered to further improve solutions. The resulting offspring are added to the population. Zhou et al. (2018) and Bevilaqua et al. (2019) use a multi-population strategy to guarantee exploration. Stopping criteria standard to GAs have been imposed, such as a maximum number of iterations without improvement, or a maximum computation time (Talbi, 2009).

4.6. Path relinking

Two papers use path relinking to find solutions to the 2E-VRP. The main idea behind path relinking is to generate and explore the trajectory in the search space connecting a starting solution and a target solution. Both papers develop initial solutions with GRASP. Anderluh et al. (2017) specify that the combination of the greedy memoryless procedure of GRASP with path relinking allows the generation of good solutions.

Crainic, Mancini, Perboli, & Tadei (2013) take the initial solution as the starting solution and the best solution obtained so far as the target solution. Anderluh et al. (2017) perform the path relinking between the initial solution and the most diverse solution in the pool of solutions generated thus far. The diversity of a solution is measured by comparing the number of FE and SE vehicles used, the number of synchronizations, and the absolute difference in the number of vehicles visiting each satellite. Next, a sequence of neighboring solutions from the starting solution $s$ to the target solution $s^*$ is generated, and the best solution in the sequence is returned (Talbi, 2009).

Crainic et al. (2013) apply the path relinking procedure to the assignment of customers to a satellite. The procedure iterates over customers that are assigned to different satellites in $s$ and $s^*$ and iteratively changes $s$ by reassigning customers to the same satellite as they are assigned to in $s^*$ and solving the corresponding CVRP subproblems, one for each satellite and one for the first echelon. Anderluh et al. (2017) apply path relinking to the SE routes. A pool of SE solutions is kept for this procedure, where the SE routes in a solution are concatenated to form a giant tour. The path between two solutions is created by iteratively selecting and swapping customers in $s$ that have a different position in $s^*$. Suppose that customer $i$ is positioned at different locations $j$ and $k$ in solutions $s$ and $s^*$ respectively. The path relinking procedure swaps the nodes located at index $j$ (customer $i$) and index $k$ in solution $s$. Once there are no customers with different positions, the satellite locations are aligned using the same procedure. The resulting giant tour is split into multiple SE routes, and the corresponding FE routes are generated with a nearest neighbor heuristic.

5. Comparison of solution methods

In Sections 3 and 4, we discussed various exact and heuristic solution approaches for the 2E-VRP. To identify the state-of-the-art algorithms, we compare the results obtained on benchmark instances. In Section 5.1, we summarize the sets of benchmark instances. Solution characteristics are given in Section 5.2. Finally, in Section 5.3 we identify the best performing algorithms.

5.1. Benchmark instances

Numerous benchmark instances for the 2E-VRP are available. Table 4 displays an overview of 2E-VRP benchmark sets, as well as their characteristics. We show for each set the number of instances, the number of depots, satellites, and customers in the instances, any additional settings, and the paper in which it was proposed, respectively.

Sets 1, 2a, 2b, 3a, and 3b are derived from the classical VRP instances proposed by Christofides & Eilon (1969). Additional satellite
locations are generated at random to convert the instances to 2E-VRP instances. All instances in set 1 are small and already solved to optimality in Perboli et al. (2011). Therefore, we omit the discussion of these instances in the remainder of this section. Breunig et al. (2016) observed some inconsistencies in sets 2b and 3b as they were considered in Hemmelmayr et al. (2012) and available in the OR-library (Beasley, 2018), and introduced sets 2c and 3c to resolve the differences. Sets 2b and 2c contain different satellite locations, whereas the difference between sets 3b and 3c is the location of the depot. In set 3b, the depot is placed at the southwest corner of the customer area. In contrast, in set 3c, the depot is positioned near the center of the customer area.

The instances in sets 4a and 4b contain three customer distribution patterns (random, centroids, and quadrants) and three satellite location patterns (random, sliced, and forbidden zone). Instances representing downtown and suburban zones in a large city are obtained with centroid customer distributions. To simulate cities with a natural border such as a river or valley, customers are clustered into zones (quadrants). As for the satellite locations, a more even distribution of the satellites on the ring around the customers is achieved with a sliced pattern. Forbidden satellite zones are included in some of the instances to represent the practical case where the presence of sea or mountains may forbid the location of satellites. For more details on the instances, we refer to Crainic, Mancini, Perboli, & Tadei (2010). The number of SE vehicles leaving a satellite $m_s$ is bounded in set 4a and unbounded in set 4b.

Sets 5, 6, and 7 contain larger instances in terms of the number of customers and satellites. The depot is located at the border of the customers’ area in all three sets. In set 6b, different handling costs per unit transshipped at each satellite are incurred.

The instances in sets 8a and 8b are developed for the 2E-VRP with multiple depots and time windows. Each instance represents a circular urban area that is split up into four areas. Starting in the inner city and moving outwards, the areas contain (1) customers, (2) customers and satellites, (3) customers, and (4) depots. Similar to Crainic et al. (2010), a sliced pattern is considered for the location of the satellites and depots. In set 8b, multiple commodities are introduced and a depot is assigned to each commodity.

To facilitate future research, we include a data archive with all 2E-VRP benchmark instances in the online appendix to this paper.

5.2. Solution characteristics

We start this section by presenting an analysis of the optimal solutions obtained on the various instance sets. Next, we discuss the impact of the network topology on solutions.

In Table 5, the number of instances solved to optimality are given for each instance set, as well as some additional solution statistics. It can be observed that all instances in sets 2a, 2b, 3a, 3b, and 4a up to 6b are solved to optimality. Sets 2c and 3c are introduced in Breunig et al. (2016) and the reported solutions on those instances have not been proven optimal. Comparing the FE and SE vehicle utilizations, we observe that higher capacity utilization is achieved with the SE vehicles. The average number of customers visited in a SE route ranges from 3 up to 10. On average, fewer customers are visited in the SE routes in the solutions to the instances in sets 8a and 8b, which is inherent to the ratio between the SE vehicle capacity (Q2) and the average customer demand ($\bar{d}$) in the instances.

The location of the depot, satellites, and customers, and the travel costs of the arcs have an impact on the solutions. Crainic et al. (2010) conclude that using a two-echelon network instead of a single-echelon network results in solutions with lower costs, in particular when the depot is located outside of the customer area and the satellites are located between the depot and the customers. Belgin et al. (2018) motivate this reduction and state that in a two-echelon distribution system, larger vehicles are used on the first echelon, enabling the consolidation of loads of different SE vehicles into one FE vehicle. Consequently, the total distance traveled in a 2E-VRP solution is lower than the total distance traveled in a single-echelon VRP solution. The presence of forbidden zones may result in satellites being located in areas with low demand, leading to additional SE routing costs (Crainic et al., 2010).

Crainic, Mancini, Perboli, & Tadei (2012) emphasize that the cost of traversing an arc does not only depend on the distance but also the typology of the arc (highway, city center street, etc.) and the type of vehicle covering it. They proceed with an analysis of the impact of different distance-based costs, costs with fixed tolls, and different travel times on the reported solutions to the instances in set 4a in Crainic et al. (2010). The results show that the solutions are not impacted by distance-based cost variations, but change in case of high tolls on arcs between customers. The impact of customer locations on the difficulty of the instances is studied in Breunig et al. (2016). They conclude that instances with customers located in centroids are more challenging to solve than instances where customers are located at random or in quadrants.

5.3. Comparison of the algorithms

We identify the best performing exact algorithms by comparing their reported computational results on benchmark instances. Tables 6 and 7 summarize the results for the single-depot 2E-VRP
Table 5
Characteristics of optimal solutions.

<table>
<thead>
<tr>
<th>Set</th>
<th>Solved</th>
<th>( f_{re}^a )</th>
<th>( f_{se}^b )</th>
<th>Vehicle utilization</th>
<th>Number of customers in SE route</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>FE</td>
<td>SE</td>
<td>Minimum</td>
<td>Average</td>
</tr>
<tr>
<td>2a</td>
<td>12/12</td>
<td>0.27</td>
<td>0.92</td>
<td>0.93</td>
<td>3</td>
</tr>
<tr>
<td>2b</td>
<td>9/9</td>
<td>0.09</td>
<td>0.75</td>
<td>0.90</td>
<td>8</td>
</tr>
<tr>
<td>2c</td>
<td>0/9</td>
<td>0.11</td>
<td>0.56</td>
<td>0.97</td>
<td>8</td>
</tr>
<tr>
<td>3a</td>
<td>12/12</td>
<td>0.22</td>
<td>0.92</td>
<td>0.93</td>
<td>3</td>
</tr>
<tr>
<td>3b</td>
<td>0/6</td>
<td>0.18</td>
<td>0.83</td>
<td>0.97</td>
<td>6</td>
</tr>
<tr>
<td>3c</td>
<td>6/6</td>
<td>0.14</td>
<td>0.75</td>
<td>0.97</td>
<td>6</td>
</tr>
<tr>
<td>4a</td>
<td>54/54</td>
<td>0.44</td>
<td>0.80</td>
<td>0.87</td>
<td>1</td>
</tr>
<tr>
<td>4b</td>
<td>54/54</td>
<td>0.39</td>
<td>0.63</td>
<td>0.87</td>
<td>1</td>
</tr>
<tr>
<td>5a</td>
<td>18/18</td>
<td>0.39</td>
<td>0.72</td>
<td>1.00</td>
<td>1</td>
</tr>
<tr>
<td>6a</td>
<td>27/27</td>
<td>0.26</td>
<td>0.79</td>
<td>0.71</td>
<td>1</td>
</tr>
<tr>
<td>6b</td>
<td>27/27</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7a</td>
<td>23/51</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8a</td>
<td>223/240</td>
<td>0.12</td>
<td>0.79</td>
<td>0.81</td>
<td>0.95</td>
</tr>
<tr>
<td>8b</td>
<td>50/60</td>
<td>0.30</td>
<td>0.70</td>
<td>0.66</td>
<td>0.91</td>
</tr>
</tbody>
</table>

\[\text{a} \text{ Fraction of FE costs to total cost, } \text{b} \text{ Fraction of satellites visited, } \text{c} \text{ Ratio between SE vehicle capacity and average customer demand, } \text{d} \text{ Based on the optimal solutions obtained from the webpage provided in Breung et al. (2016), } \text{e} \text{ Based on 6 optimal solutions, } \text{f} \text{ Based on 18 optimal solutions, } \text{g} \text{ No detailed solutions available, } \text{h} \text{ Based on 223/240 optimal solutions to the instances with up to 50 customers obtained from Mhamede et al. (2022).} \]

Table 6
Results exact methods - 2E-VRP.

<table>
<thead>
<tr>
<th>Set</th>
<th>#</th>
<th>BMRC (2013)[i]</th>
<th>JSR (2013)[i]</th>
<th>SCM (2013)[i]</th>
<th>SMC (2015)[a]</th>
<th>MSDD (2020)[d]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Solved Time(s)</td>
<td>Solved Time(s)</td>
<td>Solved Time(s)</td>
<td>Solved Time(s)</td>
<td>Solved Time(s)</td>
</tr>
<tr>
<td>2a</td>
<td>12</td>
<td>12</td>
<td>223</td>
<td>12</td>
<td>115</td>
<td>6</td>
</tr>
<tr>
<td>2b</td>
<td>9</td>
<td>9</td>
<td>90</td>
<td>6</td>
<td>4096</td>
<td>2</td>
</tr>
<tr>
<td>2c</td>
<td>6</td>
<td>12</td>
<td>46</td>
<td>11</td>
<td>920</td>
<td>5</td>
</tr>
<tr>
<td>3a</td>
<td>12</td>
<td>6</td>
<td>403</td>
<td>3</td>
<td>5251</td>
<td>0</td>
</tr>
<tr>
<td>3b</td>
<td>6</td>
<td>54</td>
<td>50</td>
<td>15</td>
<td>7955</td>
<td>0</td>
</tr>
<tr>
<td>3c</td>
<td>5</td>
<td>18</td>
<td>3</td>
<td>17,515</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>4a</td>
<td>27</td>
<td>27</td>
<td>19995</td>
<td>27</td>
<td>5144</td>
<td>27</td>
</tr>
</tbody>
</table>

\[\text{i} \text{ Baldacci et al. (2013), Jepsen et al. (2013), Santos et al. (2013), Santos et al. (2015), and Marques et al. (2020b), } \text{i} \text{ 3/6 instances with 100 customers and 5 satellites, } \text{i} \text{ Obtained from Santos et al. (2015), } \text{i} \text{ Baldacci et al. (2008) imposed an upper bound on the computation time for solving the reduced formulation and the number of routes created (10h), } \text{i} \text{ Shared memory for solving multiple instances in parallel, } \text{i} \text{ 36,000 seconds (10 hours) for sets 4a-6b and 216,000 seconds (60 hours) for set 7.} \]

Table 7
Results exact methods - 2E-VRPTW.

<table>
<thead>
<tr>
<th>[D]</th>
<th>[S]</th>
<th>[C]</th>
<th>#</th>
<th>Solved Time(s)</th>
<th>DDWC (2019)[i]</th>
<th>MACD (2022)</th>
<th>MSDD (2020)</th>
<th>DWCD (2021)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>15</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>&lt;1</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>15</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>&lt;1</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>15</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>&lt;1</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>30</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>6</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>30</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>6</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>30</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>6</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>50</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>14</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>50</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>14</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>50</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>14</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>100</td>
<td>20</td>
<td>6</td>
<td>20</td>
<td>8</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>100</td>
<td>20</td>
<td>3</td>
<td>11</td>
<td>20</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>100</td>
<td>20</td>
<td>2</td>
<td>15</td>
<td>20</td>
<td>19</td>
<td>20</td>
</tr>
</tbody>
</table>

\[\text{i} \text{ Delhae et al. (2019), Mhamede et al. (2022), Marques et al. (2020a), and Delhae et al. (2019), } \text{i} \text{ Only 17/20 instances considered, } \text{i} \text{ Considered 5/20 instances, } \text{i} \text{ Shared memory for solving multiple instances in parallel.} \]

879
and the multi-depot 2E-VRPTW, respectively. The reported average computation times include the time spent on unsolved instances. Details regarding the computer hardware used to conduct experiments are included in the tables. Although a one-to-one comparison of the different computational results is complicated by the fact that the experiments are conducted on different machines, it is nevertheless evident from Table 6 that the best results are obtained by Marques et al. (2020b). It outperforms the other exact methods on sets 4a up to 6b in terms of number of instances solved and computation time. To solve the instances in set 7, Marques et al. (2020b) set the time limit to 60 hours and solve 23 of the 51 instances to optimality. Marques et al. (2020a) report the most competitive results on the multi-depot 2E-VRPTW instances with a single commodity (set 8a). Only four of the instances with 100 customers could not be solved to optimality.

In Table 8, we compare the performance of the nine best heuristics on the 2E-VRP benchmark instances. Again, details regarding the computer hardware used to conduct experiments are included in the table. The first two rows of the table state the papers and the corresponding solution methods. We observe that all solution methods belong to the class of single-solution neighborhood search-based heuristics. Columns “Gap” and “t” show the best optimality gap (in %) and the average computation time (in seconds) reported over five runs. The optimality gaps are computed with respect to the optimal or best-known solutions. With the exception of Jie et al. (2019) (JYZH), Entenhoven et al. (2020) (EJRBS), and Yu et al. (2021) (YHG), all heuristics find the optimal solutions to the instances in sets 2a up to 3b. For the former three heuristics, it holds that they were designed for a variant of the 2E-VRP: the 2E-VRP with covering options (Entenhoven et al., 2020; Yu et al., 2021), and the 2E-VRP with electric vehicles (Jie et al., 2019). Varying optimality gaps and computation times are reported on the remaining instance sets. For instance, on sets 4a and 4b, Mühlbauer & Fontaine (2021) report shorter computation times and higher optimality gaps. To conclude, none of the heuristics outperforms the other heuristics on all instance sets.

6. Two-echelon vehicle routing in practice

The first application of routing in a two-echelon distribution system can be traced back to Jacobsen & Madsen (1980). They consider the problem of distributing newspapers in western Denmark. Additional to the first and second echelon routing, decisions on the number and locations of the transfer points (satellites) have to be taken. Crainic et al. (2004) propose the use of satellites near cities to achieve higher average load factors and fewer empty trips in the city center. A preliminary evaluation of the proposed satellite organization of urban freight transportation is performed based on data from the city of Rome (Italy). It shows that satellites may indeed contribute towards a better city environment at a reasonable cost. Anderlüh et al. (2017, 2019) and Anderlüh et al. (2021) consider a distribution system with bikes on the second echelon and propose a real-world instance based on Vienna (Austria). The instance contains 100 pharmacies selected at random to represent the customers, and one pharmacy wholesaler’s location is used as the depot. The satellites and the bike depot are placed at potential locations (parking places) at the border of the city center and in the city center, respectively.

A logistics problem for a hospital complex in Tours (France) is studied in Kergosien et al. (2013). Hospital units are located in different parts of the city. The first echelon involves routing vehicles between different hospital units to deliver medicines, clean linen, meals, various supplies, patient files, and pick up waste and dirty linen. In the second echelon, routes for pick-ups and deliveries to different buildings within a unit are constructed.

Gu et al. (2021) present a study on instances generated from a case study proposed by local authorities of the French department of Isère. The case study aims to increase the volume of fresh food products sold through short and local supply chains. Two customers sets are included: school canteens and supermarkets, with fresh fruit and vegetables as commodities. The instances include over 100 school canteens, over 150 supermarkets, up to 65 suppliers (depos), and five distribution centers (satellites).

Mühlbauer & Fontaine (2021) study the use of swap containers in a parcel delivery network to streamline operations at the satellites. At the depot, a dedicated swap container is loaded for each SE route. At the satellites, the swap containers with the packages are directly transferred from the FE vehicles (vans) to the SE vehicles (cargo-bikes). Additionally, asymmetric travel durations mimic real-life traffic situations such as one-way streets, different waiting times for left and right turns, and asymmetric congestion. Their real-life problem instance covers two distinct districts in Munich (Germany). Customer locations are determined by randomly selecting peoples’ addresses from a phone book to have the clustering of customers in the city reflected in the instance.

Next to the standard at-home delivery, a package can be delivered at a nearby pick-up point. With this service, customers can pick up their packages at a convenient time and preferred place. Zhou et al. (2018) and Entenhoven et al. (2020) consider the application of package delivery and incorporate additional delivery options for last mile distribution. Zhou et al. (2018) include a case study on the city of Chongqing (China). Alewijnse & Hübl (2021) consider a unique mode of transport for the second echelon and investigate how the city canal network of Amsterdam can be utilized for last mile parcel delivery. Finally, Li et al. (2020, 2018) allow the temporary storage capacities of satellites to vary throughout the day and design instances using data from an express delivery company operating in Beijing, Shanghai, and Guangzhou (China).

Wang & Wen (2020) study the distribution process of a cold chain logistics company in China, which is responsible for transferring dairy products. The impact of restricted route choices in a 2E-VRP model on the urban area of Beijing (China) is analyzed in Huang, Savelbergh, & Zhao (2018). Liu et al. (2018) discuss a practical example where the distribution network of a Chinese company is organized according to the many administrative divisions of the country. Instead of having a large number of small warehouses for each division, the company prefers to operate only a few large warehouses (satellites). To reduce the impact on the existing management system and organizational structure, the distribution network should be such that each warehouse replenishes all stores belonging to the same division.

Several papers use the supply chain of supermarkets, retailers, and wholesalers as their application domain. Belgin et al. (2018) and Soysal et al. (2015) present case studies on supermarket chains in Turkey and The Netherlands, respectively. Bevilaqua et al. (2019) implement their solutions approaches on the distribution network of a Brazilian wholesale company. Ramirez-Villamil, Jaegler, & Montoya-Torres (2021) design their instances based on data of a French delivery company in Paris (France).

Finally, Agárdi et al. (2019), Breunig et al. (2019), and Jie et al. (2019) focus on the design of sustainable 2E-VRP networks which rely on electric vehicles. They extend the transportation network by incorporating stations for recharging electric vehicles or swapping their batteries.

7. Conclusions and open research areas

We presented an overview of the literature on the 2E-VRP and showed that early work is focused on the single-depot 2E-VRP. Recently, various challenging extensions of the 2E-VRP appeared,
### Table 8
Results heurisitic solution approaches - 2E-VRP

<table>
<thead>
<tr>
<th>Paper</th>
<th>Method</th>
<th>ALNS</th>
<th>GRASP + VND</th>
<th>LNS</th>
<th>VNS + IP</th>
<th>LNS + SC</th>
<th>ALNS</th>
<th>ALNS</th>
<th>PLNS</th>
<th>ALNS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set</td>
<td>#</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2a 6</td>
<td>21</td>
<td>2</td>
<td>0.00</td>
<td>35</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
<td>60</td>
<td>0.00</td>
<td>1</td>
</tr>
<tr>
<td>2b 6</td>
<td>50</td>
<td>2</td>
<td>0.00</td>
<td>147</td>
<td>0.00</td>
<td>2</td>
<td>0.00</td>
<td>60</td>
<td>0.00</td>
<td>18</td>
</tr>
<tr>
<td>3a 6</td>
<td>50</td>
<td>2</td>
<td>0.00</td>
<td>150</td>
<td>0.00</td>
<td>1</td>
<td>0.00</td>
<td>60</td>
<td>0.00</td>
<td>13</td>
</tr>
<tr>
<td>4a 6</td>
<td>50</td>
<td>2</td>
<td>0.00</td>
<td>78</td>
<td>0.00</td>
<td>1</td>
<td>0.00</td>
<td>60</td>
<td>0.00</td>
<td>6</td>
</tr>
<tr>
<td>3b 6</td>
<td>50</td>
<td>16</td>
<td>0.00</td>
<td>161</td>
<td>0.00</td>
<td>16</td>
<td>0.00</td>
<td>60</td>
<td>0.00</td>
<td>37</td>
</tr>
<tr>
<td>3c 6</td>
<td>50</td>
<td>2</td>
<td>0.00</td>
<td>70</td>
<td>0.00</td>
<td>27</td>
<td>0.00</td>
<td>60</td>
<td>0.00</td>
<td>28</td>
</tr>
<tr>
<td>4b 6</td>
<td>50</td>
<td>2</td>
<td>0.00</td>
<td>48</td>
<td>0.00</td>
<td>30</td>
<td>0.00</td>
<td>60</td>
<td>0.04</td>
<td>36</td>
</tr>
<tr>
<td>5 6</td>
<td>100</td>
<td>5</td>
<td>0.15</td>
<td>900</td>
<td>0.21</td>
<td>953</td>
<td>0.07</td>
<td>900</td>
<td>0.31</td>
<td>1452</td>
</tr>
<tr>
<td>6 6</td>
<td>100</td>
<td>10</td>
<td>0.17</td>
<td>900</td>
<td>0.31</td>
<td>1452</td>
<td>0.52</td>
<td>557</td>
<td>0.52</td>
<td>10308</td>
</tr>
<tr>
<td>7 6</td>
<td>100</td>
<td>5</td>
<td>0.00</td>
<td>700</td>
<td>0.32</td>
<td>1500</td>
<td>0.90</td>
<td>900</td>
<td>2.12</td>
<td>14691</td>
</tr>
<tr>
<td>8 6</td>
<td>100</td>
<td>10</td>
<td>0.16</td>
<td>500</td>
<td>0.32</td>
<td>1500</td>
<td>0.90</td>
<td>900</td>
<td>2.12</td>
<td>14691</td>
</tr>
<tr>
<td>9 6</td>
<td>100</td>
<td>5</td>
<td>0.00</td>
<td>700</td>
<td>0.01</td>
<td>730</td>
<td>0.06</td>
<td>235</td>
<td>3.15</td>
<td>418</td>
</tr>
<tr>
<td>10 6</td>
<td>100</td>
<td>5</td>
<td>0.15</td>
<td>900</td>
<td>0.05</td>
<td>1216</td>
<td>0.07</td>
<td>560</td>
<td>6.32</td>
<td>1370</td>
</tr>
<tr>
<td>11 6</td>
<td>100</td>
<td>4</td>
<td>0.00</td>
<td>60</td>
<td>0.00</td>
<td>141</td>
<td>0.01</td>
<td>33</td>
<td>0.29</td>
<td>325</td>
</tr>
<tr>
<td>12 6</td>
<td>100</td>
<td>4</td>
<td>0.00</td>
<td>900</td>
<td>0.00</td>
<td>578</td>
<td>0.00</td>
<td>363</td>
<td>0.28</td>
<td>509</td>
</tr>
<tr>
<td>13 6</td>
<td>100</td>
<td>4</td>
<td>0.20</td>
<td>900</td>
<td>0.06</td>
<td>1322</td>
<td>0.03</td>
<td>712</td>
<td>1.68</td>
<td>1332</td>
</tr>
</tbody>
</table>

**N. Simaji, A. Amiri, I. Khandelal et al.**

**European Journal of Operational Research** (2021) 865–886
including, multiple depots, time windows, pick-up and delivery operations, a heterogeneous fleet of FE and SE vehicles, and the inclusion of alternative delivery options. A large number of both exact and heuristic methods are developed for the 2E-VRP and its variants. Presently, the state-of-the-art exact algorithm for the 2E-VRP is the BP&C method by Marques et al. (2020b), which is capable of solving instances with up to 300 customers and 15 satellites. To obtain reasonable solutions in a short amount of time, heuristics have been developed, including variable neighborhood descent, variable neighborhood search, tabu search, (adaptive) large neighborhood search, genetic and path relinking algorithms. The best-performing heuristics for the single-depot 2E-VRP are the neighborhood search-based approaches. However, currently, there is no single heuristic that outperforms all others.

We conclude this review with suggestions for future research. Spatial load synchronization is adequately covered in the existing 2E-VRP research. However, in practice, spatial load synchronization alone is insufficient. The timing aspect is essential, i.e., the operations should happen in a meaningful sequence. For example, an SE vehicle can only depart from a satellite after it has received its load from the FE vehicle(s). Often, this issue is not addressed explicitly, and it is assumed that the SE vehicles only start after the FE vehicles visited the satellite. Thus, temporal synchronization is a crucial aspect of two-echelon distribution systems that, by and large, has been neglected in recent 2E-VRP literature. Long waiting times of FE and SE vehicles at the satellites can be avoided by imposing additional constraints on the vehicles’ waiting times at the satellites (see, e.g., Anderluh et al., 2017). When the transshipment area is small, additional temporal capacity constraints are required to limit the number of vehicles that can visit a satellite simultaneously. Including constraints for temporal synchronization, waiting times, and temporal capacity is the first step towards more practical models for 2E-VRPs.

Additional research opportunities involving the satellite operations exist. In the classical 2E-VRP, split deliveries at the satellites are allowed, i.e., a SE vehicle may receive its load from multiple FE vehicles. However, several recent papers explicitly prohibit split deliveries as they assume that each SE route is associated with a single FE route. Although this assumption simplifies operations at the satellite and reduces the need for synchronization, the impact of this assumption is unclear and should be quantified. Another common assumption is that satellites have no storage capacity. A limited amount of research is conducted on the benefits brought by satellites with temporary storage capabilities (Li et al., 2018). When temporary storage is possible, for example, in the trailer of a truck parked at the satellite for the duration of the delivery period (Marujo et al., 2018), it is not required that an FE and SE vehicle are simultaneously present at a satellite. Temporal capacity constraints are necessary to prevent storage capacity violations at the satellites.

Analysis of the optimal solutions showed that some of the satellites in the instances are not visited by any FE and SE vehicles, and others are visited by many vehicles. A more uniform utilization of the available satellites could potentially result in a network that is easier to manage. Furthermore, additional research on the option of direct deliveries from the depot to customers located in areas that are accessible with FE vehicles is required. This option offers more flexibility when solving the problem. It can potentially reduce costs, but the impact of direct deliveries on the solution cost has not yet been quantified.

Given the relatively short distances between customers and satellites and the smaller capacity of the SE vehicles, the SE routes are likely to be short in length. Explicitly limiting the duration and/or distance of an SE route has additional benefits when electric vehicles and/or cargo bikes are used to execute the routes. For example, if the SE routes are sufficiently short, the planning of recharging points of electric vehicles along the SE routes is no longer required. Instead, recharging facilities can be arranged at the satellite, e.g., using parking spots with power outlets for electric vehicles. With the recharging possibilities and the relatively short duration of the SE routes, SE vehicles can be used for multiple SE routes, which can be modeled as a multi-trip 2E-VRP (Grangier et al., 2016; Marques et al., 2020a). Hence, another promising research direction is the 2E-VRP with multiple trips on the second echelon.

In practice, one typically has to account for uncertainty in the travel times and/or customer demand. Both impact the synchronization decisions at the satellite. Stochastic travel times (STT) on the second echelon, together with time windows at customers, may require an early departure of SE vehicles from their satellites to compensate for potential delays. Consequently, earlier FE vehicles’ arrival times at the satellites are required to supply the SE vehicles on time. Stochastic travel times further impact the assignment of customers to satellites, resulting in different satellite demands and, therefore, a different routing problem on the first echelon. Additional research on the applicability of the existing 2E-VRP methodology to the 2E-VRP with STT is needed. Stochasticity in customer demand is often captured with a recourse policy. In the classical single-echelon VRP, recourse policies involve a detour of the vehicle to the depot to restock. For the 2E-VRP, this translates to a detour of an SE vehicle to a satellite. If the satellites have no storage capacity, other vehicles (possibly FE vehicles) have to visit the satellite (again) to restock the SE vehicle, which is an expensive operation. Thus, designing cost-efficient recourse policies for the 2E-VRP with stochastic demand is a challenge that remains to be investigated.

Acknowledgements

This work was funded by The Dutch Research Council (NWO) DAREFUL project under grant 629.002.211. The authors thank the editors and two anonymous referees for their comments, which led to significant improvements in this paper.

Supplementary materials

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2022.02.022.

Appendix A

Table A.1
<table>
<thead>
<tr>
<th>Reference</th>
<th>Number of depots</th>
<th>Number of commodities</th>
<th>Satellite capacity</th>
<th>Split delivery to SE vehicles</th>
<th>Fleet type</th>
<th>Time windows</th>
<th>Synchronization</th>
<th>Maximum route length</th>
<th>Multi-trip</th>
<th>Pickup and delivery</th>
<th>Delivery options</th>
<th>Electric fleet</th>
<th>Stochasticity</th>
<th>Case study</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agárdi et al. (2019)</td>
<td>S</td>
<td>S</td>
<td>Q&lt;sub&gt;e&lt;/sub&gt;</td>
<td>✓</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alewijnse &amp; Hübl</td>
<td>M</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>H</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Amarouche et al.</td>
<td>S</td>
<td>S</td>
<td>m&lt;sub&gt;s&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Anderluh et al. (2017)</td>
<td>S</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Anderluh et al. (2019)</td>
<td>S</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Baldacci et al. (2013)</td>
<td>S</td>
<td>S</td>
<td>m&lt;sub&gt;n&lt;/sub&gt;, Q&lt;sub&gt;n&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Belgin et al. (2018)</td>
<td>M</td>
<td>S</td>
<td>Q&lt;sub&gt;e&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Bevilaqua et al. (2019)</td>
<td>S</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Breunig et al. (2016)</td>
<td>S</td>
<td>S</td>
<td>m&lt;sub&gt;n&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Breunig et al. (2019)</td>
<td>S</td>
<td>S</td>
<td>m&lt;sub&gt;n&lt;/sub&gt;, Q&lt;sub&gt;n&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Crainic et al. (2009)</td>
<td>M</td>
<td>S</td>
<td>m&lt;sub&gt;n&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Crainic et al. (2010)</td>
<td>S</td>
<td>M</td>
<td>m&lt;sub&gt;n&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Crainic et al. (2011)</td>
<td>S</td>
<td>S</td>
<td>m&lt;sub&gt;n&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Crainic et al. (2013)</td>
<td>S</td>
<td>S</td>
<td>m&lt;sub&gt;n&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Crainic et al. (2016)</td>
<td>M</td>
<td>M</td>
<td>m&lt;sub&gt;n&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Dellaert et al. (2019)</td>
<td>M</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Dellaert et al. (2021)</td>
<td>M</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Esmali &amp; Sahraei (2020)</td>
<td>S</td>
<td>S</td>
<td>m&lt;sub&gt;n&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Enthoven et al. (2020)</td>
<td>S</td>
<td>S</td>
<td>m&lt;sub&gt;n&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Granger et al. (2016)</td>
<td>S</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Gu et al. (2021)</td>
<td>M</td>
<td>M</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>He &amp; Li (2019)</td>
<td>S</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Hemmelmayr et al. (2012)</td>
<td>S</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Huang et al. (2018)</td>
<td>S</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Huang et al. (2021)</td>
<td>S</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Jepsen et al. (2013)</td>
<td>S</td>
<td>S</td>
<td>m&lt;sub&gt;n&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Jia et al. (2022)</td>
<td>M</td>
<td>M</td>
<td>Q&lt;sub&gt;e&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Jie et al. (2019)</td>
<td>S</td>
<td>S</td>
<td>m&lt;sub&gt;n&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Kancharla &amp; Ramadurai (2019)</td>
<td>M</td>
<td>M</td>
<td>m&lt;sub&gt;n&lt;/sub&gt;</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

(continued on next page)
Table A1 (continued)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Number of depots</th>
<th>Number of commodities</th>
<th>Satellite capacitya</th>
<th>Split delivery to SE vehicles</th>
<th>Fleet typeb</th>
<th>Time windows</th>
<th>Synchronizationc</th>
<th>Maximum route lengthd</th>
<th>Multi-trip</th>
<th>Pickup and delivery</th>
<th>Delivery options£</th>
<th>Electric fleet</th>
<th>Stochasticityf</th>
<th>Case study</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kergosien et al. (2013)</td>
<td>M</td>
<td>M</td>
<td>×</td>
<td>✓</td>
<td>H</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>[S] = 1</td>
</tr>
<tr>
<td>Li et al. (2016b)</td>
<td>S</td>
<td>S</td>
<td>×</td>
<td>✓</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Li et al. (2016a)</td>
<td>S</td>
<td>S</td>
<td>×</td>
<td>✓</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Li et al. (2018)</td>
<td>S</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Li et al. (2020)</td>
<td>S</td>
<td>S</td>
<td>Qₖ</td>
<td>✓</td>
<td>H</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Li et al. (2021b)</td>
<td>S</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>H</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Liu et al. (2017)</td>
<td>S</td>
<td>S</td>
<td>Qₖ</td>
<td>✓</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Liu et al. (2018)</td>
<td>S</td>
<td>S</td>
<td>Qₖ</td>
<td>✓</td>
<td>S</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Marques et al. (2020a)</td>
<td>M</td>
<td>S</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Marques et al. (2020b)</td>
<td>S</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Storage at satellites</td>
</tr>
<tr>
<td>Mhamedi et al. (2022)</td>
<td>M</td>
<td>S</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Mühlbauer &amp; Fontaine (2021)</td>
<td>S</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Paul et al. (2021)</td>
<td>M</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Perboli et al. (2011)</td>
<td>S</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Ramirez-Villamil et al. (2021)</td>
<td>S</td>
<td>S</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Sahraei &amp; Esmaeili (2018)</td>
<td>S</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Santos et al. (2013)</td>
<td>S</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Santos et al. (2015)</td>
<td>S</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Stek &amp; Wikarek (2014)</td>
<td>S</td>
<td>S</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Sluijk et al. (2021)</td>
<td>M</td>
<td>S</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Song et al. (2017)</td>
<td>M</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Soysal et al. (2015)</td>
<td>S</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Wang et al. (2017b)</td>
<td>S</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Wang et al. (2017a)</td>
<td>S</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Wang et al. (2018)</td>
<td>M</td>
<td>S</td>
<td>Qₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Wang et al. (2021)</td>
<td>S</td>
<td>S</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Wu &amp; Zhang (2021)</td>
<td>S</td>
<td>S</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Yan et al. (2019)</td>
<td>S</td>
<td>S</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Yu et al. (2020)</td>
<td>S</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Yu et al. (2021)</td>
<td>S</td>
<td>S</td>
<td>mₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Zeng et al. (2014)</td>
<td>S</td>
<td>S</td>
<td>mₖ, Qₖ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Zhou et al. (2022)</td>
<td>S</td>
<td>S</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

*a* Expressed in terms of number of SE vehicles (mₖ) and/or capacity (Qₖ); *b* Same (homogeneous) or heterogeneous; *c* Temporal synchronization (T); *d* Maximum route duration and/or distance; *e* Direct delivery from the depot (D), delivery to pick-up point (P), and/or delivery through occasional drivers (OD); *f* Stochastic demand (D) or stochastic travel times (TT).
References


