The Two-Echelon Stochastic Multi-period Capacitated Location-Routing Problem

Imen Ben Mohamed, Walid Klibi, Ruslan Sadykov, Halil Şen, François Vanderbeck

The Two-Echelon Stochastic Multi-period Capacitated Location-Routing Problem

Imen Ben Mohamed[*1,2], Walid Klibi[†1], Ruslan Sadykov[‡3,2], Halil Şen[§4], and François Vanderbeck[¶5]

[1]The Center of Excellence in Supply Chain (CESIT), Kedge Business School, Bordeaux, France
[2]IMB, Université de Bordeaux, 351 cours de la Libération, 33405 Talence, France
[3]Inria Bordeaux – Sud-ouest, 200 avenue de la Vieille Tour, 33405 Talence, France
[4]Mapotempo, Bordeaux, France
[5]Atoptima, Bordeaux, France

7 July 2020

Abstract

Given the emergence of two-echelon distribution systems in several practical contexts, this paper tackles, at the strategic level, a distribution network design problem under uncertainty. This problem is characterized by the two-echelon stochastic multi-period capacitated location-routing problem (2E-SM-CLRP). In the first echelon, one has to decide the number and location of warehouse platforms as well as the intermediate distribution platforms for each period; while fixing the capacity of the links between them. In the second echelon, the goal is to construct vehicle routes that visit ship-to locations (SLs) from operating distribution platforms under a stochastic and time-varying demand and varying costs. This problem is modeled as a two-stage stochastic program with integer recourse, where the first-stage includes location and capacity decisions to be fixed at each period over the planning horizon, while routing decisions of the second echelon are determined in the recourse problem. We propose a logic-based Benders decomposition approach to solve this model. In the proposed approach, the location and capacity decisions are taken by solving the Benders master problem. After these first-stage decisions are fixed, the resulting sub-problem is a capacitated vehicle-routing problem with capacitated multiple depots (CVRP-CMD) that is solved by a branch-cut-and-price algorithm. Computational experiments show that instances of realistic size can be solved optimally within a reasonable time and provide relevant managerial insights on the design problem.

1 Introduction

At the strategic planning level of distribution networks, the design decisions involve the determination of the number, location and capacity of the distribution platforms required to satisfy the evolving demand of a customers base. It also determines the mission of these platforms in terms of the subset of customers they must supply and of the interconnection between the facilities. This design problem takes basis on the well-known Facility location problem (FLP) (Geoffrion

[*]imen.benmohamed@kedgebs.com
[†]walid.klibi@kedgebs.com
[‡]ruslan.sadykov@inria.fr
[§]halil.sen@mapotempo.com
[¶]fv@atoptima.com

and Graves, 1974; Klose and Drexl, 2005), and is also extensively tackled in the Location-Routing problem (LRP) that integrates transportation decisions (Nagy and Salhi, 2007; Prodhon and Prins, 2014). In recent years, the rise of online sales coupled with the consumers desire of speed, the increased attention to sustainability in urban logistics and the evolution of logistics assets capability, changed drastically the distribution landscape. Most of practitioners admit the limits of single echelon distribution networks and have nowadays turned their attention to two-echelon distribution structures to meet today's challenges. Several retailers realize that the location of their centralized warehouses, calibrated for logistics efficiency, is not necessarily optimized to provide fast deliveries, or to efficiently operate fulfillment and urban shipment services.

The growth of e-commerce and the emergence of omnichannel retailing have drastically favored a high proximity to customers' Ship-to locations (SLs), which impact the last-mile delivery cost (Chopra, 2018). As mentioned in Savelsbergh and Van Woensel (2016), the increase of direct-to-consumer deliveries and the emergence of new freight movement lead to additional fulfillment complexity and transportation complexity. For instance, Parcel delivery companies are experiencing important stretch of their distribution systems with the inclusion of a new echelon of urban fulfilment hubs (Winkenbach et al., 2016). In the same way, several retailers such as Walmart, JD.com or Amazon have adapted their distribution networks by adding an advanced echelon of distribution/fulfillment platforms, mostly in urban areas. According to Weinswig (2018), Walmart turned several Sam's Club stores into e-commerce fulfillment centers to support the rapid e-commerce growth. In link with that, urban last-mile delivery faces great challenges to satisfy growing customer needs for a punctual and a precise delivery location, either their home, office, car, smart locker or a store (Savelsbergh and Van Woensel, 2016). In several business contexts, this issue tends to increase the number of inventory locations and echelons in the distribution system. Crainic and Montreuil (2016) underline the current expansion to multi-tier urban distribution systems and proposed an interconnected logistics framework. City logistics is probably the most significant example of the shift from a one-echelon to a two-echelon distribution network setting (Crainic et al., 2016; Janjevic et al., 2019). This is achieved by creating peripheral distribution/consolidation centers dedicated to transferring and consolidating freight from back-level platforms/hubs. Given these challenges, it is legitimate for the distribution network designer to question to which extent current models are sufficient to provide good quality designs.

Current distribution network design models found in the literature have three main shortcomings when compared to the strategic needs of distribution businesses. First, most of related works rely mainly on a single echelon distribution structure (Melo et al., 2009), whereas several distribution systems are more suitable for a two-echelon network structure as highlighted with a number of examples above and in the reviews of Savelsbergh and Van Woensel (2016); Cuda et al. (2015). We refer hereafter in general to the two-echelon Location-Routing problem by 2E-LRP, and to its specific capacitated version by 2E-CLRP. When it comes to the 2E-LRP, the literature is still scarce. Deterministic versions of the 2E-LRP are introduced in Boccia et al. (2010); Contardo et al. (2012); Darvish et al. (2019).

Second, the dynamics of business operations nowadays calls for a decision process that integrates a more refined granularity of the operations than traditional high level of aggregation (Geoffrion and Graves, 1974). Once designed, the distribution network is in operation on a daily basis and has to deal with detailed customer locations and varying demand/orders. These operational details are found in routing models (VRPs) but unfortunately not often integrated when location decisions have to be taken (Klibi et al., 2016). Giving the complexity of VRPs, most 2E-LRPs rely on continuous approximations based on route length estimation (RLE), but this work show that, with recent advances, realistic size VRPs can be solved optimally within a reasonable time. Further, we note that most existent 2E-LRP and most LRP modeling approaches implicitly assume that location and routing decisions are made simultaneously for the planning horizon, without considering the hierarchical structure of the strategic problem. Such framework, introduced by Schneeweiss (2003), relies on the time lag and the top-down relation between location and routing decisions, as applied in Klibi et al. (2010); Ben Mohamed et al. (2020).

Third, the business environment is clearly dominated by a dynamic-stochastic setting. Ac-

cordingly, the traditional deterministic-static representation of the planning horizon is due to be replaced by a more realistic stochastic and multi-period characterization of the planning horizon. Since the 2E-CLRP is a strategic decision process that must be designed to last for several years, the horizon must be partitioned into a set of periods shaping the uncertainty and variability in time. The location and capacity decisions should be planned as a set of sequential decisions to be implemented at different design periods of the horizon (a year, for example) and promoting the structural adaptability of the network. Existing works focus on static and deterministic settings for the 2E-LRPs. A stochastic 2E-location model with RLE is proposed by Snoeck et al. (2018) and a multi-stage 2E-location-allocation model is introduced in Ben Mohamed et al. (2020). Including the features stressed above, gives rise to the two-echelon stochastic multiperiod capacitated location-routing problem, denoted by (2E-SM-CLRP). As far as we know, the 2E-SM-CLRP has not been addressed yet.

Figure 1 illustrates a typical 2E-CLRP partitioned into two capacitated distribution echelons: each echelon involves a specific location-assignment-transportation schema that must cope with the future demand. It aims to decide the number and location of warehousing/storage platforms (WPs) and distribution/fulfillment platforms (DPs), and on the capacity allocated from first echelon to second echelon platforms. It also decides the transportation activity between platforms. Direct routes with full truckloads transportation option, relying on consolidation policy, are considered in the first echelon between opened WPs and selected DPs. This is common in practice, where the distributor, using its own vehicle fleet or a contract carrier, considers a single-destination full trailer to deliver a regional/urban DP on a scheduled basis. In the second echelon, the transportation activity is shaped by multi-drop routes, as DPs are generally devoted to more fragmented urban services (Figure 1).

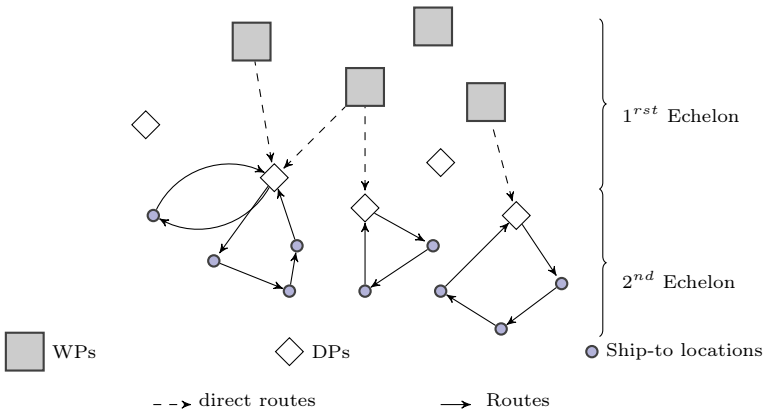


Figure 1: A potential two-echelon capacitated location-routing problem (2E-CLRP)

The contribution of this study is threefold. Firstly, we introduce the 2E-SM-CLRP under uncertain and time-varying demand and cost, a hierarchical decision problem that is characterized by a temporal hierarchy between design and transportation decisions. Secondly, we present a two-stage stochastic program with integer recourse for the 2E-SM-CLRP that captures the temporal hierarchy and the multi-stage setting of the decision process. A scenario-based approach is used to describe the uncertain SLs demands: it relies on a set of multi-period scenarios generated with a Monte-Carlo approach. In addition, 2E-SM-CLRP inherits complexity of the LRP (Laporte, 1988) and adds a combinatorial-stochastic structure, which makes it a very hard optimization problem. As a third contribution, we propose a first exact approach to solve this problem by means of a logic-based Benders decomposition approach (Hooker, 2019; Benders, 1962) and a sample average approximation (SAA) (Shapiro et al., 2009). It calls the branch-cut-and-price algorithm of Sadykov et al. (2020) to solve Benders sub-problems in a parallel computing scheme. Two families of Benders cuts are proposed to cut off infeasible solutions and help converging to an optimal solution of the 2E-SM-CLRP. The extensive computational experiments on a large set of instances, emphasize the performance of the proposed algorithm on solving medium-scale instances optimally, and on getting good feasible solutions – i.e., with less than 0.5% gap from

a lower bound – for larger instances with up to 50 SLs and 25 demand scenarios over a 5-year planning horizon. Finally, we share the insights we derived from the computational experiments about the impact of the stochastic and multi-period settings on the 2E-CLRP.

The reminder of this paper is organized as follows. Section 2 surveys the related work on the 2E-CLRP. Section 3 introduces the mathematical formulation of the 2E-SM-CLRP. Section 4 presents the proposed exact solution approach based on Benders decomposition. The computational results are presented and analyzed in Section 5. Section 6 provides conclusion and future research avenues.

2 Literature review

The literature on distribution network design problems is evolving towards the study of more complex and integrated models. Therein, strategic decisions and operational decisions are strongly interrelated (Crainic and Laporte, 1997; Ambrosino and Scutellà, 2005). More specifically, the strategic level involves location and capacity planning decisions and the operational level implies transportation decisions. The integration of the two decision levels into a location-routing problem (LRP), often leads to better network design solutions as introduced in Salhi and Rand (1989) and recently discussed in Shen and Qi (2007) and Klibi et al. (2016). Several variants of the LRPs are studied in the literature. They differ in terms of number of distribution echelons involved in the network, the number of echelons in which location decisions are made, the decisions on capacity planning, and the granularity of transportation integration under single / multi-period and deterministic / stochastic demand consideration. We distinguish between origin-destination flows, multi-drop routes and RLE formulas for transportation integration. Comprehensive surveys on distribution problems and classification scheme can be found in Prodhon and Prins (2014); Drexl and Schneider (2015) and Cuda et al. (2015). We summarize the main related studies in Tables 1 and 2 according to the aforementioned modeling options.

As highlighted in Table 1, most of works on LRPs rely on a single echelon distribution structure (Laporte and Dejax, 1989; Laporte et al., 1989; Albareda-Sambola et al., 2007; Contardo et al., 2014a,b) where explicit routes are integrated into the design level. First contributions to two-echelon distribution structure are from Jacobsen and Madsen (1980) and Madsen (1983) who introduce the two-echelon LRP (2E-LRP) in the context of newspapers distribution. The model determines the location of the second-echelon platforms at no cost and without capacity limitations. In the recent years with the high growth of e-commerce, 2E-LRPs have gained increasing attention in which the location decisions on both echelons may be questioned, and models consider platform costs and capacity limitations. Sterle (2010) formally introduces the 2E-CLRP in an urban context and proposes three mixed-integer programming formulations. The same model is studied by Boccia et al. (2010); Contardo et al. (2012) and Schwengerer et al. (2012) who focus on the development of exact and heuristic solution methods. Nguyen et al. (2012a,b) examine a 2E-CLRP including a single warehouse in the first echelon already located with unlimited capacity. Few recent works by Mirhedayatian et al. (2019) and Farham et al. (2020) study a variant of the 2E-CLRP with time windows. In Mirhedayatian et al. (2019), delivery and pickup operations are also incorporated.

Capacity planning decisions as a strategic/design decisions, that determine the capacity allocated to platforms, are included in some variants of distribution network design problems as in Correia et al. (2013); Georgiadis et al. (2011) and Ben Mohamed et al. (2020). These decisions are further investigated in capacity planning and dynamic distribution design problems (Pimentel et al., 2013; Jena et al., 2015).

Further, most of LRPs and 2E-LRPs modeling approaches implicitly assume that design and transportation decisions are made simultaneously, without considering the hierarchical structure of the strategic problem as applied in Klibi et al. (2010); Snoeck et al. (2018) and Ben Mohamed et al. (2020). Alternatively, some papers use a hierarchical approach to the two-echelon distribution structure, extending the facility location problem (FLP) (Daskin, 1995) to the two-echelon FLP (2E-FLP) (Geoffrion and Graves, 1974; Correia et al., 2013; Georgiadis et al., 2011). In

Table 1: Structure, decisions and context in distribution network studies

Article	# Distribution echelons	Design Decisions		Transportation integration			Urban context
		Location Echelons	Capacity	Flows	Routes	RLE	
Albareda-Sambola et al. (2007)	1	1			√		
Albareda-Sambola et al. (2012)	1	1			√		
Albareda-Sambola et al. (2013)	1	1		√			
Baldacci et al. (2011)	1	1			√		
Belenguer et al. (2011)	1	1			√		
Contardo et al. (2014b)	1	1			√		
Contardo et al. (2013, 2014a)	1	1			√		
Hemmelmayr et al. (2012)	1	1			√		
Klibi et al. (2010)	1	1			√		
Laporte and Dejax (1989)	1	1			√		
Laporte et al. (1989)	1	1			√		
Schneider and Löffler (2019)	1	1			√		
Shen (2007)	1	1				√	
Santoso et al. (2005)	1	1		√			
Schütz et al. (2009)	1	1		√			
Ben Mohamed et al. (2020)	2	1	√	√			
Correia et al. (2013)	2	2	√	√			
Georgiadis et al. (2011)	2	2	√	√			
Darvish et al. (2019)	2	1			√		
Jacobsen and Madsen (1980)	2	1			√		
Mirhedayatian et al. (2019)	2	1			√		√
Nguyen et al. (2012a,b)	2	1			√		√
Ambrosino and Scutellà (2005)	2	2			√		
Contardo et al. (2012)	2	2			√		
Farham et al. (2020)	2	2	√		√		√
Schwengerer et al. (2012)	2	2			√		
Sterle (2010)	2	2			√		√
Zhao et al. (2018)	2	2			√		√
Janjevic et al. (2019)	2	1				√	√
Merchán and Winkenbach (2018)	2	1				√	√
Snoeck et al. (2018)	2	1				√	√
Winkenbach et al. (2016)	2	2				√	√
Our work	2	2	√		√		√

these studies, transportation operations are highly aggregated as origin-destination flows. More recent works as Winkenbach et al. (2016); Merchán and Winkenbach (2018) and Janjevic et al. (2019) present a particular variant of 2E-CLRPs to design an urban distribution network in which RLE formulas are integrated into the model to approximate operational routing decisions. However, flows and RLE approximations may ignore the detailed customer locations and varying demands/orders that are better captured with daily explicit routes.

Further studies have addressed other variants of the two-echelon problems, defined as two-echelon vehicle-routing problems (2E-VRPs) (Crainic et al., 2004; Perboli et al., 2011). In this class of problems, the main focus is about the transportation capabilities of the network, without considering design decisions we stress here (i.e., platforms are used freely without inducing a setup cost). For a detailed review, the reader is referred to Cuda et al. (2015).

The complexities of real world applications lead to the incorporation of dynamic-stochastic setting into the problem. These features are addressed in one-echelon LRPs, as in (Laporte et al., 1989; Shen, 2007; Albareda-Sambola et al., 2007; Klibi et al., 2010). In Table 2, we focus on two-echelon distribution network studies under multi-period and stochastic demand characterization, and illustrate solution methods developed to tackle the problem. As pointed out in Table 2, most of studies with two-echelon distribution structure deal with deterministic-static setting. The few works considering stochastic and multi-period aspects concern specific variants of the 2E-CLRP. Snoeck et al. (2018) apply the RLE and segmentation approach and integrates stochastic demand in the 2E-CLRP in order to establish a more robust network with lower expected operational cost and reduced risk of performance decline. Ben Mohamed et al. (2020) introduce a stochastic multi-period version of the 2E location-allocation problem, where design decisions are determined on a multi-period setting and routes are represented by multi-period inter-facility flows. Others works as in Georgiadis et al. (2011) examine a multi-period stochastic two-echelon distribution network design model in which multi-period setting concerns only flow

Table 2: Multi-period and stochastic characterization and solution approaches in two-echelon distribution network studies

Article	Uncertainty		Multi-period			Solution approach
	Deterministic	SD	Location	Capacity	Transportation	
Georgiadis et al. (2011)		√			√	Numerical solver
Snoeck et al. (2018)		√				Numerical solver
Ben Mohamed et al. (2020)		√	√	√	√	Benders decomposition+SAA
Correia et al. (2013)	√		√	√	√	Numerical solver
Darvish et al. (2019)	√		√		√	Enhanced Parallel Branch-and-bound
Ambrosino and Scutellà (2005)	√				√	Lagrangean-based heuristic
Contardo et al. (2012)	√					Branch-and-cut & ALNS
Farham et al. (2020)	√					Branch-and-Price & Two heuristics based on hierarchical decisions decomposition
Jacobsen and Madsen (1980)	√					Three constructive heuristics
Janjevic et al. (2019)	√					Constructive heuristic and combined breadth-first and depth-first search
Merchán and Winkenbach (2018)	√					Numerical solver
Mirhedayatian et al. (2019)	√					decomposition-based heuristic
Nguyen et al. (2012a,b)	√					GRASP & MS-ILS
Schwengerer et al. (2012)	√					VNS
Sterle (2010)	√					Numerical solver
Winkenbach et al. (2016)	√					Numerical solver
Zhao et al. (2018)	√					Cooperative approximation heuristic based on Lagrangean relaxation and GTS
Our work		√	√	√	√	Logic-based Benders decomposition

SD: Stochastic demand; SAA: Sample Average Approximation; ALNS: Adaptive Large Neighborhood Search; GRASP: Greedy randomized Adaptive Search Procedure; VNS: Variable Neighborhood Search; MS-ILS: Multi-Start Iterated Local Search; GTS: Granular Tabu Search

decisions.

Accordingly, as far as we know, stochastic-multi-period setting with temporal hierarchy and explicit routes in 2E-CLRP has not been addressed yet. Moreover, the extant literature on two-echelon distribution studies shows several exact and heuristic solution methods used to solve these problems. But, further progress is still required to solve efficiently real size instances of 2E-CLRPs. Exact methods rely on the mathematical programs to optimize the problem. We distinguish numerical solver as applied in (Sterle, 2010; Georgiadis et al., 2011; Snoeck et al., 2018), and exact algorithms such as Branch-and-cut algorithm (Contardo et al., 2012; Darvish et al., 2019), Branch-and-price algorithm (Farham et al., 2020) and Benders decomposition (Ben Mohamed et al., 2020). The table also presents heuristics such as the Greedy Randomized Adaptive Search Procedure (GRASP) (Nguyen et al., 2012a), the Multi-Start Iterated Local Search (MS-ILS) (Nguyen et al., 2012b), the Adaptive Large Neighborhood Search (ALNS) (Contardo et al., 2012), and other constructive approximation heuristics based on problem decomposition (Zhao et al., 2018; Mirhedayatian et al., 2019; Farham et al., 2020). Therefore, the developed exact approaches are able to solve in reasonable computing time 2E-CLRPs instances containing up to 2 warehouses, 10 distribution centers, and 50 customers, while heuristic solutions solve problem instances with up to 5 warehouses, 20 distribution centers, and 200 customers (Cuda et al., 2015). The SAA is a sampling method introduced by (Shapiro, 2003) to limit the large number of scenarios in stochastic models, and is successfully applied to 2E- location-allocation (Ben Mohamed et al., 2020) and several one-echelon LRPs (Klibi et al., 2010; Santoso et al., 2005).

Thus, our review points to the literature's shortcomings in addressing the stochastic and multi-period variants of the 2E-CLRP, while integrating explicit routes and considering the hierarchical structure of the strategic problem. It also demonstrates the lack of exact methods which deal with the dynamic-stochastic version of this novel problem. Except the study presented by Klibi et al. (2010) who examine a variant of the one-echelon LRP with multi-period transportation level, there is no study on formulating and solving the 2E-CLRP under stochastic and multi-period features. Our study contributes to fill this gap.

3 Mathematical formulation

The two-echelon stochastic multi-period capacitated location-routing problem (2E-SM-CLRP) is a hierarchical decision problem. This stems from the temporal hierarchy between the design decisions and the operational transportation decisions. The design decisions – i.e, location and capacity – are made on yearly basis before the realization of the uncertain parameters, whereas the operational routing decisions are determined when uncertainty is revealed. To catch the temporal hierarchy, we formulate the 2E-SM-CLRP as a two-stage scenario-based stochastic model where the scenarios represent the realizations of uncertain SLs demands over the planning horizon. In the first-stage, design decisions on opening, operating and closing of WPs and DPs as well as on capacity allocated to links between platforms are taken here-and-now for the set of design periods considered. Then, SLs orders are revealed, and second-stage decisions are determined for every specific scenario. These decisions consist in building daily routes that visit SLs using a vehicle. Each route starts and finishes in an operating DP in the second echelon.

We consider a long-term planning horizon $\mathcal{T}$ that covers a set of successive design planning periods $\mathcal{T} = \{1, \ldots, T\}$. Such periods are defined in accordance with the evolution of the uncertain SLs demand over planning periods (typically years). Each planning period encompasses a set of operational periods represented generally in a discrete way by "typical" business days. Under uncertainty, the routing decisions depend on the actual realization of the demand at each period t. Thus, each realization defines a demand scenario ω representing a typical day of delivery which has a probability of occurrence $p(\omega)$. All potential scenarios characterize the set of demand scenarios Ω_t modeling the uncertainty behavior of the demand of SLs at period t. We denote by Ω the set of all scenarios: $\Omega = \cup_t \Omega_t$.

The 2E-SM-CLRP is defined on a graph with three disjoint sets of nodes — those representing the potential locations for warehouse platforms (WPs), $\mathcal{P} = \{p\}$, those for the potential locations for distribution platforms (DPs), $\mathcal{L} = \{l\}$, and those for the SLs, $\mathcal{J} = \{j\}$. WPs and DPs can be opened (started), kept operating or closed at any planning period with a time-varying fixed cost – i.e., respectively, fw^s_{pt}, fw_{pt}, fw^c_{pt} for WPs and f^s_{lt}, f_{lt}, f^c_{lt} for DPs. Each WP p (respectively DP l) has a limited capacity C_p (respectively C_l). Additionally, each SL j has an uncertain demand $d^{t\omega}_j$ at period t under scenario ω.

At the first echelon, we consider an undirected bipartite graph $\mathcal{G}^1 = (\mathcal{V}^1, \mathcal{E}^1)$, with the vertex set $\mathcal{V}^1 = \mathcal{P} \cup \mathcal{L}$, and the edge set $\mathcal{E}^1 = \{(p, l) : p \in \mathcal{P}, l \in \mathcal{L}\}$ representing the WPs and DPs, and the links in between. These links represent strategic assignment-transportation decisions to calibrate the DP throughput capacity based on transportation capabilities. To each link (p, l), we can assign one or several full truckloads, having each a capacity Q^1_{lp} and a fixed cost h_{lpt}. Moreover, multi-sourcing strategy is allowed in the first echelon – i.e., each DP can be supplied from more than one operating WP – while respecting the capacity limitations.

At the second echelon, an undirected graph $\mathcal{G}^2 = (\mathcal{V}^2, \mathcal{E}^2)$ is defined where $\mathcal{V}^2 = \mathcal{L} \cup \mathcal{J}$, and $\mathcal{E}^2 = \{(i, j) : i, j \in \mathcal{V}^2, i \text{ and } j \text{ not both in } \mathcal{L}\}$ representing the DPs and SLs and the links in between. No lateral transshipment between DPs is performed – i.e., there are no direct edges between DPs. A routing cost c^t_{ij} is associated with each edge $(i, j) \in \mathcal{E}^2$ at period t in the second echelon.

We consider an unlimited set $\mathcal{K}$ of identical vehicles with capacity Q^2 used to visit SLs in the second echelon, where $Q^2 < Q^1_{lp}$ for all $(p, l) \in \mathcal{E}^1$. If used, a fixed cost is paid for each vehicle for which we assign a route in the second echelon. We assume that this cost is already incorporated into the routing cost in the following way. The cost c^t_{ij} of each edge $(i, j) \in \mathcal{E}^2$ adjacent to a DP – i.e., those representing the departures from and arrivals at the DPs – is increased by the half of the fixed vehicle cost for each t.

The proposed model decides the opening, operating, and closing periods of each WP and DP, as well as the number of full truckloads assigned to each link $(p, l) \in \mathcal{E}^1$ thus defining the capacity allocated to DPs. In the second-stage, the goal is to build vehicle routes so that each SL is visited exactly once in each period and each scenario. The total demand delivered to SLs from each operating DP under each scenario is less than or equal to the capacity assigned to that DP from

WPs. Therefore, our model aims to minimize the total expected cost by minimizing the sum of the expected transportation cost and the design cost (facility location and capacity allocation).

Given a fixed first-stage design solution, the second-stage problem is the capacitated vehicle-routing problem with capacitated multiple depots (CVRP-CMD), modeled using a set partitioning formulation (SPF) (Toth and Vigo, 2014). Let us denote by $\mathcal{R}_l^{t\omega}$ the set of all feasible routes starting and ending at an operating DP l visiting at most once each SL and satisfying capacity constraints for period t under scenario $\omega \in \Omega_t$. We have $\mathcal{R} = \cup_{t\in\mathcal{T}} \cup_{\omega\in\Omega_t} \cup_{l\in\mathcal{L}} \mathcal{R}_l^{t\omega}$. Let ψ_{ij}^r denote the number of times edge (i,j) participates in route $r \in \mathcal{R}$, and ξ_j^r the number of times SL j is visited in route $r \in \mathcal{R}$ ($\xi_j^r = \sum_{i\in\mathcal{J}|(i,j)\in\mathcal{E}^2} \psi_{ij}^r$). Then $\sum_{j\in\mathcal{J}} d_j^{t\omega} \xi_j^r \leq Q^2$ should hold for every $r \in \mathcal{R}_l^{t\omega}$. Cost c_r^t of route $r \in \mathcal{R}_l^{t\omega}$ is calculated as $c_r^t = \sum_{(i,j)\in\mathcal{E}^2} \psi_{ij}^r c_{ij}^t$.

In order to strengthen the CVRP-CMD formulation, we introduce a cut that improves the lower bound obtained by the SPF. This cut imposes a valid lower bound $\Gamma_{t\omega}$ on the number of vehicles required to serve SL demand in period t under scenario ω. This lower bound is obtained through the solution of a bin packing problem (BPP) in which the aim is to pack a given set of items having different weights into a minimum number of equal-sized bins (Martello and Toth, 1990). In this case, the set of items represents the set of SLs $\mathcal{J}$ where each item has a weight $d_j^{t\omega}$ – i.e. the SL demand – and the bin capacity is the vehicle capacity Q^2. More precisely, $\Gamma_{t\omega}$ is the rounded up value of the linear relaxation of (BPP) obtained for period t and scenario ω.

The 2E-SM-CLRP is formulated as follows. Let y_{pt}^+, y_{pt}, and y_{pt}^- be binary variables which take value 1 if WP $p \in \mathcal{P}$ is selected for opening, operating and closing in period $t \in \mathcal{T}$. In a similar fashion, we define z_{lt}^+, z_{lt}, and z_{lt}^- for each DP $l \in \mathcal{L}$. Let x_{lpt} be an integer variable representing the number of full truckloads assigned from WP p to DP l in period t. Let $\lambda_{lr}^{t\omega}$ be a binary variable indicating whether a route $r \in \mathcal{R}_l^{t\omega}$ is selected in the optimal solution. The two-stage stochastic integer program with recourse then can be written as:

$$\min \sum_{t\in\mathcal{T}} \sum_{p\in\mathcal{P}} (fw_{pt}^s\, y_{pt}^+ + fw_{pt}\, y_{pt} + fw_{pt}^c\, y_{pt}^-) + \sum_{t\in\mathcal{T}} \sum_{l\in\mathcal{L}} (f_{lt}^s\, z_{lt}^+ + f_{lt}\, z_{lt} + f_{lt}^c\, z_{lt}^-) + \sum_{t\in\mathcal{T}} \sum_{p\in\mathcal{P}} \sum_{l\in\mathcal{L}} h_{lpt} x_{lpt} + \sum_{t\in\mathcal{T}} \mathbb{E}_{\Omega_t}[\phi_{t\omega}(x)] \tag{1}$$

$$\text{S.t.} \quad \sum_{l\in\mathcal{L}} Q_{lp}^1 x_{lpt} \leq C_p\, y_{pt} \qquad \forall p, t \tag{2}$$

$$\sum_{p\in\mathcal{P}} Q_{lp}^1 x_{lpt} \leq C_l\, z_{lt} \qquad \forall l, t \tag{3}$$

$$y_{pt} - y_{pt-1} \leq y_{pt}^+ \qquad \forall p \in \mathcal{P}, t \in \mathcal{T} \tag{4}$$

$$z_{lt} - z_{lt-1} \leq z_{lt}^+ \qquad \forall l \in \mathcal{L}, t \in \mathcal{T} \tag{5}$$

$$y_{pt-1} - y_{pt} \leq y_{pt}^- \qquad \forall p \in \mathcal{P}, t \in \mathcal{T} \tag{6}$$

$$z_{lt-1} - z_{lt} \leq z_{lt}^- \qquad \forall l \in \mathcal{L}, t \in \mathcal{T} \tag{7}$$

$$\sum_t y_{pt}^+ \leq 1 \qquad \forall p \in \mathcal{P} \tag{8}$$

$$\sum_t y_{pt}^- \leq 1 \qquad \forall p \in \mathcal{P} \tag{9}$$

$$\sum_t z_{lt}^+ \leq 1 \qquad \forall l \in \mathcal{L} \tag{10}$$

$$\sum_t z_{lt}^- \leq 1 \qquad \forall l \in \mathcal{L} \tag{11}$$

$$x_{lpt} \in \mathbb{N} \qquad \forall l \in \mathcal{L}, p \in \mathcal{P}, t \in \mathcal{T} \tag{12}$$

$$y_{pt}^+, y_{pt}, y_{pt}^- \in \{0,1\} \qquad \forall p \in \mathcal{P}, t \in \mathcal{T} \tag{13}$$

$$z_{lt}^+, z_{lt}, z_{lt}^- \in \{0,1\} \qquad \forall l \in \mathcal{L}, t \in \mathcal{T} \tag{14}$$

where $\phi^{t\omega}(x)$ is the solution of the recourse problem:

$$(\text{SPF}_{t\omega}) \quad \phi_{t\omega}(x) = \min \sum_{l\in\mathcal{L}} \sum_{r\in\mathcal{R}_l^{t\omega}} c_r^t \, \lambda_{lr}^{t\omega} \tag{15}$$

$$\text{S. t.} \quad \sum_{l\in\mathcal{L}} \sum_{r\in\mathcal{R}_l^{t\omega})} \xi_j^r \lambda_{lr}^{t\omega} = 1 \qquad \forall j \in \mathcal{J} \tag{16}$$

$$\sum_{r\in\mathcal{R}_l^{t\omega}} \left(\sum_{j\in\mathcal{J}} d_j^{t\omega} \xi_j^r \right) \lambda_{lr}^{t\omega} \leq \sum_{p\in\mathcal{P}} Q_{lp}^1 x_{lpt} \qquad \forall l \in \mathcal{L} \tag{17}$$

$$\sum_{l\in\mathcal{L}} \sum_{r\in\mathcal{R}_l^{t\omega}} \lambda_{lr}^{t\omega} \geq \Gamma_{t\omega} \tag{18}$$

$$\lambda_{lr}^{t\omega} \in \{0,1\} \qquad \forall l \in \mathcal{L}, r \in \mathcal{R}_l^{t\omega} \tag{19}$$

The objective function (1) minimizes the sum of the first-stage costs and the expected second-stage costs. The first-stage cost is the sum of the opening, operating, and closing costs of the WPs and DPs, and the capacity cost induced by the number of truckloads allocated to DPs from WPs. Constraints (2) and (3) guarantee the capacity restriction at operating WPs and DPs, respectively. Constraints (4) and (5) track the operating status of the WPs and DPs from one period to the next and mark the period of their opening. In a similar vein, constraints (6) and (7) manage the closing status of WPs and DPs, respectively. Constraints (8)–(11) specify that each platform is opened or closed at most once during the planning horizon. Constraints (12)–(14) describe the domain of the first-stage variables.

The second-stage objective function (15) minimizes the routing cost and the fixed cost for using vehicles. Constraints (16) ensure that each SL is served exactly once. Constraints (17) are the depot capacity inequalities. They guarantee that the demand satisfied from an operating DP will not exceed its throughput capacity. Inequalities (18) are the reinforcement cuts of the SPF on the CVRP-CMD imposing a valid lower bound $\Gamma_{t\omega}$ on the number of vehicles required to serve ship-to demand demand in period t under scenario ω.

4 Logic-based Benders decomposition approach

We solve the 2E-SM-CLRP using a logic-based Benders decomposition (LBBD) approach (Hooker, 2019). LBBD is a generalization of classical Benders decomposition (Benders, 1962) that allows the subproblem to be any optimization problem rather than specifically a linear or non-linear programming problem.

In order to use the LBBD approach, we slightly modify formulation (1)–(19). We introduce additional continuous variables

$$\theta_{t\omega} \geq 0 \qquad t \in \mathcal{T}, \omega \in \Omega_t. \tag{20}$$

We also replace the component $\sum_{t\in\mathcal{T}} \mathbb{E}_{\Omega_t}[\phi_{t\omega}(x)]$ in objective function (1) by $\sum_{t\in\mathcal{T}} \sum_{\omega\in\Omega_t} p(\omega)\theta_{t\omega}$, and we replace objective function (15) by constraints

$$\theta_{t\omega} \geq \sum_{l\in\mathcal{L}} \sum_{r\in\mathcal{R}_l^{t\omega}} c_r^t \, \lambda_{lr}^{t\omega} \qquad \forall t \in \mathcal{T}, \omega \in \Omega_t. \tag{21}$$

Thus, the resulting formulation of the 2E-SM-CLRP contains the modified objective function (1), and constraints (2)–(14), (16)–(19), (20), (21).

The Benders decomposition of this formulation consists in leaving variables x, y, z, θ in the master problem, and moving variables λ to the subproblem. The LBBD approach consists in solving the master problem (1)–(14), (20) on every iteration. Given an optimal solution $(\bar{x}, \bar{y}, \bar{z}, \bar{\theta})$ of the master problem, the subproblem is formulated which consists of objective function (15) and constraints (16)–(19), in which variables x are fixed to $\bar{x}$. If the value of an optimal solution $\bar{\lambda}$ of

the subproblem is not larger than $\sum_{t\in\mathcal{T}}\sum_{\omega\in\Omega_t} p(\omega)\theta_{t\omega}$, then the combined solution $(\bar{x},\bar{y},\bar{z},\bar{\theta},\bar{\lambda})$ is optimal for the 2E-SM-CLRP. Otherwise, Benders cuts are generated, which cut off the current solution $(\bar{x},\bar{y},\bar{z},\bar{\theta})$ but which are satisfied by any feasible solution of the 2E-SM-CLRP. After that, Benders cuts are added to the master problem, and the algorithm continues to the next iteration. An advantage of this LBBD approach is that the subproblem consists of independents sets of constraints: one set for each time period $t \in \mathcal{T}$ and each scenario $\omega \in \Omega_t$. Thus, the subproblem can be decomposed into smaller problems solved independently.

In each iteration k of the LBBD approach, the value v^k_{MP} of an optimal solution $(\bar{x}^k,\bar{y}^k,\bar{z}^k,\bar{\theta}^k)$ of the master problem provides a lower bound for the 2E-SM-CLRP. We define $v^k_{design} = \sum_{t\in\mathcal{T}}\sum_{p\in\mathcal{P}}(fw_{pt}\bar{y}^k_{pt}+fw^s_{pt}\bar{y}^{+k}_{pt}+fw^c_{pt}\bar{y}^{-k}_{pt})+\sum_{t\in\mathcal{T}}\sum_{l\in\mathcal{L}}(f_{lt}\bar{z}^k_{lt}+f^s_{lt}\bar{z}^{+k}_{lt}+f^c_{lt}\bar{z}^{-k}_{lt})+\sum_{t\in\mathcal{T}}\sum_{p\in\mathcal{P}}\sum_{l\in\mathcal{L}} h_{lpt}\bar{x}^k_{lpt}$ as the total design cost.

The problem-specific part of the LBBD approach is how the Benders cuts are generated. In order to have a practical approach, these cuts should cut off not only the current solution of the master, but also many other infeasible solutions, otherwise the number of iterations would be huge. We present below two families of Benders cuts we propose for our problem. The first family contains standard Benders optimality cuts obtained by solving the linear programming relaxation of the subproblem. These cuts are based on the linear programming duality. The second family contain problem-specific combinatorial cuts obtained by solving the subproblem to (near-)optimality.

4.1 Generating Benders optimality cuts

To generate Benders optimality cuts, for each $t \in \mathcal{T}$ and $\omega \in \Omega_t$, we solve the linear relaxation (SPLP$_{t\omega}$) of the set partitioning formulation (SPF$_{t\omega}$). Due to the exponential size of $\mathcal{R}^{t\omega}_l$, subproblems (SPLP$_{t\omega}$) are solved using column generation, which is an iterative approach. In every iteration, a subset of variables λ is considered, and a restricted set-partitioning linear program (RSPLP$_{t\omega}$) is solved. Let $(\tau^{t\omega}, \rho^{t\omega}, \iota^{t\omega})$ be an optimal dual solution of (RSPLP$_{t\omega}$), corresponding to constraints (16), (17), and (18). To determine whether this dual solution is optimal for (SPLP$_{t\omega}$), a pricing problem is solved. The pricing problem searches for a route $r \in \mathcal{R}^{t\omega}_l$, $l \in \mathcal{L}$, with a negative reduced cost. If such routes are found, the corresponding variables λ are added to (RSPLP$_{t\omega}$), and we pass to the next iteration. The pricing problem is decomposed into $|\mathcal{L}|$ problems, one for each DP. Reduced cost $\hat{c}^{t\omega}_{lr}$ of a route $r \in \mathcal{R}^{t\omega}_l$ is computed as

$$\hat{c}^{t\omega}_{lr} = c^t_r - \sum_{j\in\mathcal{J}} \tau^{t\omega}_j \xi^r_j + \left(\sum_{j\in\mathcal{J}} d^{t\omega}_j \xi^r_j\right) \rho^{t\omega}_l - \iota^{t\omega}. \tag{22}$$

By replacing ξ by ψ and removing the constant part in (22), the objective function of the pricing problem for (RSPLP$_{t\omega}$) and DP $l \in \mathcal{L}$ is

$$\min_{r\in\mathcal{R}^{t\omega}_l} \hat{c}^{t\omega}_{lr} = \min_{r\in\mathcal{R}^{t\omega}_l} \sum_{(i,j)\in\mathcal{E}^2} (c^t_{ij} - \tau^{t\omega}_j)\psi^r_{ij}. \tag{23}$$

Each pricing problem is a resource constrained shortest path problem (RCSP). The RCSP for DP l is defined on graph $\mathcal{G}^2$ from which all DPs except l are removed. The RCSP for DP l is then to find in this graph a circuit with minimum reduced cost starting and finishing at the node representing DP l.

To insure the feasibility of each (SPLP$_{t\omega}$), we add slack variables to constraints (16)–(18). We set the coefficients for these variables in the objective function large enough, so that the second-stage decisions are feasible for every period and every scenario in the final solution obtained for the 2E-SM-CLRP.

In our implementation, to solve (SPLP$_{t\omega}$), we use VRPSolver (Pessoa et al., 2019), in which the RCSPs are solved by the bucket graph based labeling algorithm (Sadykov et al., 2020).

Let $(\bar{\tau}^{t\omega k}, \bar{\rho}^{t\omega k}, \bar{\iota}^{t\omega k})$ be an optimal dual solution for $(\text{SPLP}_{t\omega})$ on iteration k of the LBBD approach. Then, by linear programming duality, the following Benders optimality cut is valid for the master problem:

$$\theta_{t\omega} + \sum_{l\in\mathcal{L}} \sum_{p\in\mathcal{P}} Q^1_{lp} \bar{\rho}^{t\omega k}_l x_{lpt} \geq \sum_{j\in\mathcal{J}} \bar{\tau}^{t\omega k}_j + \Gamma_{t\omega} \bar{\iota}^{t\omega k}. \tag{24}$$

4.2 Generating combinatorial Benders cuts

If all inequalities (24) are not violated on iteration k of the LBBD approach, for each period $t \in \mathcal{T}$ and each scenario $\omega \in \Omega_t$, we solve formulation $(\text{SPF}_{t\omega})$ to (near-)optimality. The minimum number of second-level routes $\Gamma_{t\omega}$ may be sometimes increased to $\bar{\Gamma}^k_{t\omega}$ based on current solution $\bar{x}^k$. A larger value may help to solve $(\text{SPF}_{t\omega})$ faster. We present the algorithm for such preprocessing in Appendix A.2.

To solve $(\text{SPF}_{t\omega})$, we use again VRPSolver (Pessoa et al., 2019), which implements the branch-cut-and-price (BCP) algorithm by Sadykov et al. (2020). The BCP algorithm is based on a combination of column generation, cut generation, enumeration and branching. Rounded Capacity Cuts (RCC) (Laporte and Nobert, 1983) and limited memory Rank-1 Cuts (R1C) (Jepsen et al., 2008; Pecin et al., 2017) are used to strengthen the linear relaxation of $(\text{SPF}_{t\omega})$. An enumeration procedure (Baldacci et al., 2008) based on route reduced costs is used to try to generate all routes which may take part in an optimal solution. If the number of such paths is small at some node of the branch-and-bound tree, the node is finished by a MIP solver instead of branching. Otherwise, three different branching strategies are used. They can be expressed as constraints over the following aggregated variables:

- branching on the number of vehicles used from a DP l: $\sum_{r\in\mathcal{R}^{t\omega}_l} \sum_{j\in\mathcal{J}} \frac{1}{2}\psi^r_{lj}\lambda^{t\omega}_{lr},\ \forall l \in \mathcal{L}$,
- branching on the assignment of a SL $j \in \mathcal{J}$ to an operating DP l: $\sum_{r\in\mathcal{R}^{t\omega}_l} \xi^j \lambda^{t\omega}_{lr},\ \forall j \in \mathcal{J},\ l \in \mathcal{L}$,
- branching on the edges of the original graph: $\sum_{l\in\mathcal{L}} \sum_{r\in\mathcal{R}^{t\omega}_l} \psi^r_{ij}\lambda^{t\omega}_{lr},\ (i,j) \in \mathcal{E}^2$.

Branch-cut-and-price algorithm is launched with a predefined time limit, and it finds values $\underline{\theta}^k_{t\omega}$ and $\hat{\theta}^k_{t\omega}$, which are lower and upper bounds on the optimal solution value of $(\text{SPF}_{t\omega})$. If $(\text{SPF}_{t\omega})$ is solved to optimality, both values coincide. After solving all problems $(\text{SPF}_{t\omega})$, $t \in \mathcal{T}$, $\omega \in \Omega_t$, upper bound on the solution value of the 2E-SM-CLRP can be computed as

$$v^k_{design} + \sum_{t\in\mathcal{T}} \sum_{\omega\in\Omega_t} p(\omega)\hat{\theta}^k_{t\omega}. \tag{25}$$

If the optimality gap – i.e., difference between lower bound v^k_{MP} and dual bound (25) – is sufficiently small, the algorithm is stopped. Otherwise, the following combinatorial Benders cuts are generated.

The intuition behind the combinatorial cuts is the following. Given period $t \in \mathcal{T}$, if for every $l \in \mathcal{L}$ the capacity of DP l induced by the first-stage decisions is not larger than $\sum_{p\in\mathcal{P}} Q^1_{lp}\bar{x}^k_{lpt}$, then for any scenario $\omega \in \Omega_t$ the value $\theta_{t\omega}$ of the second-stage decisions cannot be smaller than $\underline{\theta}^k_{t\omega}$. This idea can be represented as follows:

$$\forall t \in \mathcal{T}: \quad \begin{cases} \theta_{t\omega} \geq \underline{\theta}^k_{t\omega}, \forall \omega \in \Omega_t, & \text{if } \sum_{p\in\mathcal{P}} Q^1_{lp} x_{lpt} \leq \sum_{p\in\mathcal{P}} Q^1_{lp}\bar{x}^k_{lpt}, \forall l \in \mathcal{L}, \\ \theta_{t\omega} \geq 0, \forall \omega \in \Omega_t, & \text{otherwise.} \end{cases} \tag{26}$$

In order to linearize conditions (26), additional variables are introduced. We define a^k_t as a non-negative variable representing the maximum capacity increase of a DP in period t in

comparison with the fixed DP capacities of iteration k:

$$a_t^k = \max\left\{0,\ \max_{l\in\mathcal{L}}\left\{\sum_{p\in\mathcal{P}} Q_{lp}^1 x_{lpt} - \sum_{p\in\mathcal{P}} Q_{lp}^1 \bar{x}_{lpt}^k\right\}\right\} \tag{27}$$

Let also g_t^k be a binary variable equal to 1 if and only if the value of a_t^k is strictly positive. In addition, we define binary variables b_{lt}^k and b_{0t}^k involved in the linearization of expression (27). Let M' be a value which is larger than any possible value variables a can take – i.e., $M' \geq C_l$, $l \in \mathcal{L}$. The combinatorial Benders cuts are then formulated as:

$$a_t^k \leq M' g_t^k \qquad \forall t \in \mathcal{T} \tag{28}$$

$$g_t^k \leq a_t^k \qquad \forall t \tag{29}$$

$$\theta_{t\omega} \geq \underline{\theta}_{t\omega}^k (1 - g_t^k) \qquad \forall t \in \mathcal{T}, \omega \in \Omega_t \tag{30}$$

$$\sum_{p\in\mathcal{P}} Q_{lp}^1 x_{lpt} - \sum_{p\in\mathcal{P}} Q_{lp}^1 \bar{x}_{lpt}^k \leq a_t^k \qquad \forall t \in \mathcal{T}, l \in \mathcal{L} \tag{31}$$

$$a_t^k \leq \sum_{p\in\mathcal{P}} Q_{lp}^1 x_{lpt} - \sum_{p\in\mathcal{P}} Q_{lp}^1 \bar{x}_{lpt}^k + M'(1 - b_{lt}^k) \qquad \forall t \in \mathcal{T}, l \in \mathcal{L} \tag{32}$$

$$a_t^k \leq M'(1 - b_{0t}^k) \qquad \forall t \in \mathcal{T} \tag{33}$$

$$b_{0t}^k + \sum_l b_{lt}^k = 1 \qquad \forall t \in \mathcal{T} \tag{34}$$

$$a_t^k \in \mathbb{R}_+ \qquad \forall t \in \mathcal{T} \tag{35}$$

$$b_{lt}^k, b_{0t}^k, g_t^k \in \{0,1\} \qquad \forall t \in \mathcal{T}, l \in \mathcal{L} \tag{36}$$

Constraints (28) and (29) link the variables a and g. Constraints (30) impose lower bounds on variables θ according to the first condition in (26). As for constraints (31)–(34), they express a linearization of the definition (27) of variables a.

4.3 Overall algorithm

The complete description of the LBBD approach is given in Algorithm 1. We will now prove that our LBBD approach converges to an optimum solution of the 2E-SM-CLRP under certain conditions.

Proposition 1. *Algorithm 1 finds an optimum solution to the 2E-SM-CLRP after a finite number of iterations if $\epsilon = 0$ and second-stage integer problems ($SPF_{t\omega}$) are solved to optimality at every iteration.*

Proof. Proof. The validity of Benders optimality cuts follows from the strong duality of linear programming. The validity of combinatorial Benders cuts follows from the fact that they follow a linearization of conditions (26). The overall number of cuts is finite, as the number of different solutions $\bar{x}$ is finite. Therefore, there exists finite iteration k such that solution $\bar{x}^k$ is the same as solution $\bar{x}^{k'}$ at some previous iteration $k' < k$. The lower bound in iteration k is not smaller than $v_{design}^k + \sum_{t\in\mathcal{T}}\sum_{\omega\in\Omega_t} p(\omega)\bar{\theta}^k$. As all problems ($\text{SPF}_{t\omega}$) were solved to optimality in iteration k', upper bound is not larger than $v_{design}^{k'} + \sum_{t\in\mathcal{T}}\sum_{\omega\in\Omega_t} p(\omega)\underline{\theta}^{k'}$.

As $\bar{x}^{k'} = \bar{x}^k$, we have $v_{design}^k = v_{design}^{k'}$. Also from the construction of combinatorial cuts, we have $\bar{\theta}^k \geq \underline{\theta}^{k'}$ for all periods t and all scenarios ω. Thus, the lower and upper bounds match in iteration k, and feasible solution obtained in iteration k' is optimal for the 2E-SM-CLRP. □

The algorithm may stop before reaching predefined gap ϵ if the overall time limit is reached or if not all second-stage problems ($\text{SPF}_{t\omega}$) are solved to optimality. In this case, we are not

Algorithm 1 Logic-based Benders decomposition approach for the 2E-SM-CLRP

ϵ is set to the maximum optimality gap
$ub \leftarrow \infty$, $lb \leftarrow -\infty$, $k \leftarrow 0$
for all $t \in \mathcal{T}$, $\omega \in \Omega_t$ **do**
  Solve the linear relaxation of the bin packing problem (BPP) to obtain $\Gamma_{t\omega}$
end for
while $(ub - lb)/ub > \epsilon$ **do**
  Solve the (MP) to obtain solution $(\bar{x}^k, \bar{\theta}^k)$ of value v^k_{MP}
  $lb \leftarrow \max\{lb, v^k_{MP}\}$
  $newCut \leftarrow false$
  for all $t \in \mathcal{T}$, $\omega \in \Omega_t$ **do**
    Solve ($\text{SPLP}_{t\omega}$) by column generation to obtain dual solution $(\bar{\tau}^{t\omega}, \bar{\rho}^{t\omega}, \bar{\iota}^{t\omega})$ of value $\phi^{LP}_{t\omega}(\bar{x}^k)$
    if $\bar{\theta}^k_{t\omega} < \phi^{LP}_{t\omega}(\bar{x}^k)$ **then**
      Add the Benders optimality cut (24) to the (MP)
      $newCut \leftarrow true$
    end if
  end for
  if $newCut = false$ **then**
    for all $t \in \mathcal{T}$, $\omega \in \Omega_t$ **do**
      Obtain $\bar{\Gamma}^k_{t\omega}$ by preprocessing based on $\bar{x}^k$ (see Appendix A.2)
      Solve ($\text{SPF}_{t\omega}$) with $\bar{\Gamma}^k_{t\omega}$ by BCP algorithm to obtain lower bound $\underline{\theta}^k_{t\omega}$ and upper bound $\hat{\theta}^k_{t\omega}$
      if $\bar{\theta}^k_{t\omega} < \underline{\theta}^k_{t\omega}$ **then**
        Add combinatorial Benders cuts (28)–(36) to the (MP)
      end if
    end for
    $\phi(\bar{x}^k) \leftarrow \sum_{t\in\mathcal{T}} \sum_{\omega\in\Omega_t} p(\omega)\hat{\theta}_{t\omega}(\bar{x}^k)$
    $ub \leftarrow \min\{ub, v^k_{design} + \phi(\bar{x}^k)\}$
    if $\underline{\theta}^k_{t\omega} < \hat{\theta}^k_{t\omega}$ for some t, ω and no violated cuts (28)–(36) were added to the (MP) **then**
      stop
    end if
  end if
  $k \leftarrow k + 1$
end while

guaranteed to obtain a feasible solution, however, in our experiments a feasible solution were obtained for all tested instances.

5 Computational results

In this section, we present our experimental results. First, we present the instances used in the experiments. Then, we report and discuss the obtained results.

Our approach is implemented in the C++ and compiled with GCC 5.3.0. BaPCod package (Vanderbeck et al., 2018) is used to handle the branch-cut-and-price framework. The code from (Sadykov et al., 2020) is used to solve the resource constrained shortest path pricing problems. We use *CPLEX* 12.8.0 as the linear programming solver in column generation and as the integer programming solver for the set partitioning problem with enumerated columns as well as for the Benders master problem (MP). All tests are run on a cluster of 2 dodeca-core Haswell Intel Xeon E5-2680 v3 server running at 2.50 GHz with 128 GB RAM. The OpenMP API (OpenMP Architecture Review Board, 2000) is used to solve the $|\mathcal{T}| \times |\Omega_t|$ CVRP-CMD sub-problems in parallel scheme.

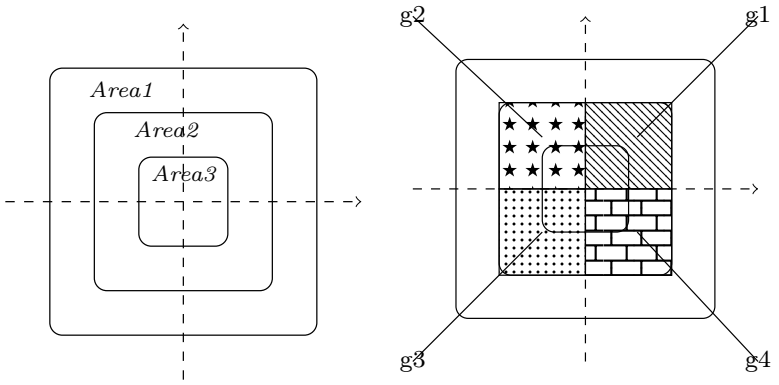
5.1 Test data

To test our approach, several 2E-SM-CLRP instances have been generated based on the following attributes: the problem size, the network characteristics, the demand process, the cost structure as well as the capacity dimension. Problems of nine different sizes are tested as shown in Table 3. We vary the number of DPs ($|\mathcal{L}|$) and the number of SLs ($|\mathcal{J}|$). The number of potential WPs ($|\mathcal{P}|$) is fixed. The sets of potential DPs given in instances with $|\mathcal{L}| = 8$ DPs and $|\mathcal{L}| = 12$ DPs are subsets of the large set with $|\mathcal{L}| = 16$ DPs. A 5-year planning horizon is considered and partitioned into 5 design periods – i.e., $|\mathcal{T}| = 5$.

Table 3: Test problems size

$\lvert\mathcal{P}\rvert$	4								
$\lvert\mathcal{L}\rvert$	8			12			16		
$\lvert\mathcal{J}\rvert$	15	20	50	15	20	50	15	20	50

Platforms and SLs are realistically scattered in a geographic area within a concentric square of increasing size as illustrated in Figure 2(a). We assume that the covered geographic territory is composed of three urban areas *Area1*, *Area2* and *Area3*, where *Area3* represents the central area. WPs are randomly located within *Area1*. 20% of the total number of DPs is located centrally in *Area3*, and 80% of DPs are in *Area2*. Two instance types are defined: I1 refers to concentric case where 80% of SLs are based in *Area3* and 20% in *Area2*; I2 corresponds to dispersed case where $\alpha = 60\%$ of SLs are based in *Area3* and 40% in *Area2*.



(a) The three urban area (b) ship-to locations dispersion zones

Figure 2: Representation of two-echelon urban area

Table 4: Demand processes

Demand process	Mean value	Trend
NIT	$\mu_{jt} = \mu_{jt-1}(1 + \delta_{z_j t})$	$\delta_{z_j t} \in [0, 0.4]$ such that $\sum_{z_j t} \delta_{z_j t} = 0.4$
NVT	$\mu_{jt} = \mu_{jt-1}(1 + \delta_{z_j t})$	$\delta_{z_j t} \in [0, 0.9]$ such that $\sum_{z_j t} \delta_{z_j t} = 0.9$
	$\mu_{jt} = \mu_{jt-1}(1 - \delta_{z_j t})$	$\delta_{z_j t} \in [0, 0.4]$
$\mu_{j0} \in [5, 25]$, $\frac{\sigma_{jt}}{\mu_{jt}} = 0.25$		

Euclidean distances between nodes are computed, and two unit costs are defined to compute platform costs and transportation costs. Higher unit cost is attributed to DPs based in *Area3* compared to the other areas. Two transportation costs configurations are tested: low transportation cost (LT) where the transportation cost represents 40% of the total network cost, and a high transportation cost (HT) where it represents 60%. These ratios are determined based on some preliminary tests. At the second echelon, vehicle capacity is fixed to $Q^2 = 75$, and

its fixed cost is $f = 600$ under the (LT) attribute and $f = 1300$ under (HT) attribute. The fixed WPs and DPs opening costs are generated, respectively in the ranges $fw_{pt} \in [14000, 25000]$ and $f_{lt} \in [7000, 10000]$ per location and period. The operating cost for both WPs and DPs is $fw^s_{pt} = 0.12fw_{pt}$, and the closing cost fw^c_{pt} is about $0.2fw_{pt}$. Additionally, we define two capacity configurations: a tight level (TC) where C_p and C_l are uniformly generated in the intervals $[600, 900]$ and $[220, 400]$ respectively; and a large level (LC) in which WPs and DPs capacities are uniformly generated in the intervals $[850, 1400]$ and $[550, 800]$, respectively. The truckload capacities Q^1_{lp} between WPs and DPs are generated in the interval $[150, 250]$. An inflation factor is considered to reflect the increase of the cost of capital and of truckloads on a periodic basis with $r = 0.005$.

Furthermore, we assume here, without loss of generality, that the demand scenario $d^{t\omega}_j$ of a SL j in period t follows the normal distribution with mean value μ_{jt} and standard deviation σ_{jt}. The SLs are partitioned into four zones $z_j \in \{g1, g2, g3, g4\}$ – see Figure 2(b) for illustration – each zone has its specific time-varying trend $\delta_{z_j t}$ linking demand mean values in time. Two demand processes are explored: a normal distribution with an increasing trend (NIT) and a normal distribution with a variable trend (NVT). In the latter, is applied a more pronounced increasing trend for periods $t = 1..3$ and then a decreasing trend for $t = 4, 5$. The ranges and ratios regarding the demand processes are given in Table 4. Combining all the elements above yields several problem instances, each instance is denoted by a problem size T-$|\mathcal{P}|/|\mathcal{L}|/|\mathcal{J}|$-, a scenario sample size N and a combination of SL dispersion (I1, I2), transportation attribute (LT, HT), capacity configuration (TC, LC) and the demand process (NIT, NVT).

5.2 Results

In this section, we evaluate the performance of our LBBD and provide an analysis of the design solutions produced by the algorithm. Further, we examine the sensitivity of the WP and DP location decisions to the uncertainty, and the behavior of the capacity decision in a multi-period and uncertain setting.

5.2.1 Algorithm performance.

In order to evaluate the performance of our approach, we solve the deterministic equivalent formulation (DEF) of the problem using a commercial solver (Cplex). To this end, in the DEF, we reformulate the CVRP-CMD as a three-index vehicle-flow formulation as introduced in Sterle (2010). Two instances of size 5-4/8/15-1- and 5-4/8/20-1- under attributes (I1, LT, LC, NIT) are tested. The results show that Cplex is not able to solve these two instances with one scenario to optimality and stops after 13 hours due to lack of memory with an optimality gap of 13% for 5-4/8/15-1- instance and 25% for 5-4/8/20-1- instance.

Then, the nine instances given in Table 3 are solved using the proposed LBBD. The results are reported in Table 5. The first two columns present the instance attribute and size associated with the sample size N. The third column *#Opt* presents the number of optimal solutions found under each scenario sample size N. The next two columns *#CombCuts* and *#OptCuts* provide the average number of iterations in which combinatorial Benders cuts (28)–(36) and Benders optimality cuts (24) are generated, respectively, and the column *#Iter* gives the required number of iterations for convergence. The column *Gap(%)* indicates the average optimality gap. The three next columns under the heading *Parallel computing time* give the average CPU time spent for solving the Benders master problem (MP), the $T \times N$ sub-problems and the total time needed to obtain an optimal solution of the problem under a parallel computing scheme for the $T \times N$ sub-problems, within the time limit (72 *hours*). The total time includes the MP solution time, the sub-problems solution time –i.e., cuts' generation time – and the BPP time.

Table 5 shows that the LBBD is able to solve most instances within a reasonable time with an $\epsilon = 0.05\%$ optimality. It solves all instances of 15 and 20 SLs, in less than 30 minutes average time with parallel computing. Note that for instances with 15 SLs, solving the BPP takes the

Table 5: Average results per problem-instance

Instance	N	#Opt	#CombCuts	#OptCuts	#Iter	Gap(%)	Parallel computing time			Sequential computing time
							MP	Sub-problems	total	
5-4/8/15-	5	24/24	2	6.3	8.3	0.00	4s	50s	2m47s	3m48s
	10	24/24	1.9	6.3	8.1	0.01	5s	50s	4m6s	6m55s
	15	24/24	1.9	6.3	8.2	0.00	6s	1m4s	5m51s	8m39s
	25	24/24	2	6.7	8.7	0.00	9s	1m27s	9m21s	13m49s
5-4/12/15-	5	15/15	1.9	6.8	8.7	0.00	8s	47s	2m32	8m20s
	10	15/15	2	7.4	9.4	0.00	12s	59s	4m22s	11m32s
	15	15/15	2.1	7.3	9.3	0.00	15s	1m8s	6m5s	18m2s
	25	15/15	2.1	7.5	9.6	0.00	25s	1m28s	9m44s	24m7s
5-4/16/15-	5	12/12	2.25	7.3	9.6	0.00	12s	2m35s	4m24s	11m39s
	10	12/12	2.3	7	9.3	0.00	20s	1m	4m34s	11m29s
	15	12/12	2.3	7.25	9.5	0.00	24s	1m11s	6m24s	17m29s
	25	12/12	1.9	7	8.9	0.01	27s	1m32s	9m50s	38m21s
5-4/8/20-	5	24/24	2.1	9	11.1	0.00	6s	1m19s	3m39s	36m7s
	10	24/24	2.1	8.7	10.7	0.00	7s	2m40s	7m14s	1h2m
	15	24/24	1.9	8.8	10.7	0.01	10s	3m23s	10m15s	1h46m
	25	24/24	2.1	9	11.1	0.00	16s	5m48s	17m14s	1h59m
5-4/12/20-	5	15/15	2.1	9.3	11.5	0.00	12s	2m28s	4m55s	39m9s
	10	15/15	1.9	8.5	10.5	0.01	16s	1m51s	6m37s	1h25m
	15	15/15	2	8.7	10.7	0.00	22s	5m26s	12m29s	2h11m
	25	15/15	1.9	7.9	9.9	0.00	30s	14m21s	26m3s	1h34m
5-4/16/20-	5	12/12	1.9	8.8	10.75	0.00	16s	1m47s	4m20s	29m2s
	10	12/12	2.2	9.7	11.8	0.01	37s	2m59s	8m7s	1h10m
	15	12/12	2.5	9.9	12.4	0.00	1m3s	11m47s	18m33s	2h49m
	25	12/12	2.25	9.67	11.91	0.01	1m34s	18m9s	30m56s	5h25m
5-4/8/50-	15	9/15	7.4	20.4	27.8	0.03	39m13s	16h54s	17h54m	
5-4/12/50-	15	8/15	7.1	26.5	33.6	0.03	2h50m	13h3m	16h14m	
5-4/16/50-	15	9/12	4.8	19.7	24.7	0.02	35m48s	7h34m	8h52m	

most computational time (2 to 7 minutes) compared to the MP+sub-problems time. This is essentially due to the sequential computing of the BPP for each period and each scenario. The sequential computing time demonstrates that the parallel computing reduces the running time drastically, mainly for moderate-size instances with 20 SLs, as the sequential solving approach could take from 4 to 10 times longer for such instances, reaching an average running time of 5 hours 25 minutes – whereas they are solved in about 30 minutes on average with parallelization. Also worth noting is that most computing time (about 70%) is spent by the branch-cut-and-price to get an integer solution to the $T \times N$ CVRP-CMD sub-problems (see Table 13 in Appendix B.2 for details). This justifies the use of a parallel computing scheme to solve the $T \times N$ sub-problems.

Furthermore, when the larger instances with 50 SLs are concerned, the LBBD is able to solve optimally 26 out of 42 instances within an average time of 14 hours 20 minutes. The average optimality gap in the remaining 16 instances is generally below 0.5% which underlines the efficiency of the developed LBBD approach. These instances are not solvable with the sequential approach within the maximum allocated time, as solving the $T \times N$ sub-problems is too time consuming. Moreover, Table 5 emphasizes the inherent complexity of the stochastic setting where computing time grows as the size of the scenarios sample N increases. Additionally, the Benders master problem (MP) takes only few seconds to be solved in most of the instances with 15 and 20 SLs and grows up to two hours in the instances (5-4/12/50-.) using 15 scenarios. The solution time of MP is also highly correlated with the number of potential DPs in the considered instance and the number of generated cuts. Adding many cuts at each iteration leads to a larger MP, and this increases its solution time. Our approach converges to an optimal solution in few iterations: an average of 10 iterations for instances with 15 and 20 SLs, and around 30 iterations for larger instances. Two-thirds of iterations involves Benders optimality cuts to introduce valid inequalities on the second-stage costs while less than one-third of which is taken by the Benders combinatorial cuts.

We then assess the performance of the solution approach regarding the combination of problem

Table 6: Detailed results for 5-4/8/15-25-, 5-4/8/20-25- and 5-4/8/50-15-

		5-4/8/15-25-				5-4/8/20-25-				5-4/8/50-15-			
		ub	lb	Gap(%)	Time	ub	lb	Gap(%)	Time	ub	lb	Gap(%)	Time
(I2, LT, TC, NIT)	1	74486.0	74486.0	0.0	10m36s	82611.2	82611.2	0.0	12m52	205012.0	204953.0	0.03	3h49m
	2	72344.1	72344.1	0.0	8m55s	96468.5	96468.5	0.0	13m19s	188933.0	188856.0	0.04	53h50m
	3	84334.7	84334.7	0.0	8m58s	89290.4	89290.4	0.0	33m53s	200191.0	200116.0	0.04	18h9m
average		77054.9	77054.9	0.0	9m30s	89456.7	89456.7	0.0	20m2s	194562.0	194486.0	0.04	36h
(I1, LT, TC, NIT)	1	74387.2	74387.2	0.0	9m44s	91276.5	91276.5	0.0	14m43s	203537.8	203367.7	0.08	-
	2	73465.3	73465.3	0.0	8m48s	84255.3	84255.3	0.0	13m40s	187485.0	187396.0	0.05	56h16m
	3	82051.9	82051.9	0.0	9m30s	89212.5	89212.5	0.0	28m4s	199104.0	199045.0	0.03	6h53m
average		76634.8	76634.8	0.0	9m21	88248.1	88248.1	0.0	18m49s	196708.9	196602.9	0.05	31h34m
(I1, HT, TC, NIT)	1	104470.0	104470.0	0.0	10m21s	128554.0	128554.0	0.0	15m54s	286317.3	286003.7	0.11	-
	2	103992.0	103992.0	0.0	9m24s	123425.0	123425.0	0.0	13m52s	268163.6	267670.7	0.18	-
	3	113685.0	113685.0	0.0	9m56s	126692.0	126692.0	0.0	21m18s	283578.6	282502.6	0.38	-
average		107382.3	107382.3	0.0	9m54s	126223.7	126223.7	0.0	17m2s	279353.2	278725.7	0.22	-
(I1, LT, LC, NIT)	1	84275.7	84275.7	0.0	8m54s	96880.2	96880.2	0.0	12m41s	190453.0	190453.0	0.0	35m10s
	2	82808.3	82808.3	0.0	8m30s	100817.0	100817.0	0.0	13m41s	188179.0	188179.0	0.0	34m5s
	3	86280.8	86280.8	0.0	9m34s	93665.0	93665.0	0.0	20m36s	196985.0	196898.0	0.04	1h16m
average		84454.9	84454.9	0.0	9m	97120.7	97120.7	0.0	15m40s	191872.3	191843.3	0.01	48m25s
(I1, LT, TC, NVT)	1	75856.1	75856.1	0.0	9m4s	107509.0	107465.0	0.04	13m22s	189925.8	189346.1	0.31	-
	2	57568.5	57568.5	0.0	8m46s	96637.7	96637.7	0.0	16m56s	218641.0	218553.0	0.04	19h43m
	3	82664.3	82664.3	0.0	9m13s	98871.2	98871.2	0.0	13m41s	208253.5	208028.5	0.11	-
average		72029.6	72029.6	0.0	9m1s	101006.0	100991.3	0.01	14m40s	205606.8	205309.2	0.15	19h43m

attributes: SL dispersion, transportation cost, capacity configuration and demand process, as defined in Section 5.1. Table 6 shows the results obtained for instances with 8 DPs and 15, 20, 50 SLs, respectively, in terms of the best upper bound (ub), the best lower bound (lb), the optimality gap (%), and the computation time. For consistency purpose, for each instance, three random instantiation of the input parameters is made as indicated in the second column.

We observe that instances under large capacity attribute (LC) are easier to solve than the other attributes, and are all solved to optimality. On the other side, the tight capacity attribute (TC) makes data sets more difficult to solve. This is consistent with former results on capacitated LRPs and VRPs. Furthermore, the instances with network configuration I2 are solved more efficiently compared to I1-based instances. This is mainly due to the fact that in I2, where SLs are more dispersed in the urban area (Figure 2(b)), the routing sub-problems are easier to solve. Indeed, two instances out of three with 50 SLs are optimally solved under I1, whereas all three instances are solved to optimality under I2. Cost attribute also impacts the complexity of the problem. In fact, for 5-4/8/50-15-(I1,.,TC,NIT), two instances out of three are optimally solved under LT attribute. However, optimality is not reached under the HT. This is mainly due to the increase of transportation costs under HT, which makes location-routing cost trade-offs more contrasting. A difference in solvability is also observed with the demand processes. Instances under NVT process are more difficult to solve compared to those under NIT, which is due to the augmented variability of the demand process in the former. As seen in Table 6 for instance, one instance sized 5-4/8/50-15-(I1,LT,TC,.) is optimally solved under NVT process compared to two under NIT. Therefore, the results confirm the efficiency of the developed LBBD approach to solve several medium to large instances with up to 25 demand scenarios of the 2E-SM-CLRP.

5.2.2 Locational analysis.

We next look at the design decisions produced by our model. The results are presented in Table 7 for the different problem sizes with 20 SLs. The results for instances with 15 and 50 SLs are given in Tables 16 and 17 in Appendix B.2. These tables present the DP opening decisions and their operating periods: value 0 refers to DPs kept closed and a value in the range $[1, 5]$ corresponds to the DP opening period t. A star (*) symbol indicates if the DP has been closed before the end of the horizon and the closing period is given at the last row. The second row provides the number of potential DPs for each combination of attributes in row 1. The first column corresponds to

the list of potential DPs, where are underlined DPs located in the central *Area3* and colored – red, blue, brown and green – the platforms located in the same zone in reference to Figure 2(b).

Table 7: Location decisions and their operating periods for 5-4/./20-25-(.,.,.,.)-3

	(I1,LT,TC,NIT)			(I1,HT,TC,NIT)			(I1,LT,TC,NVT)			(I1,LT,LC,NIT)		
$\|\mathcal{L}\|$	8	12	16	8	12	16	8	12	16	8	12	16
l1	0	0	0	0	0	0	0	0	0	0	0	0
l2	0	0	0	0	0	0	0	0	0	0	0	0
l3	0	0	0	0	0	0	1	1	0	0	0	0
l4	1	0	1	1	0	0	0	0	1	0	0	0
l5	0	2	0	0	2	0	0	0	0	1	0	0
l6	1	0	0	1	0	0	1	0	0	0	0	0
l7	0	0	0	0	0	0	2*	1	0	0	1	0
l8	0	0	0	0	0	0	0	0	0	0	0	0
l9		0	0		0	0		0	0		0	0
l10		0	0		0	0		0	0		0	0
l11		0	0		0	0		0	0		0	0
l12		1	0		1	0		0	0		0	0
l13			0			0			0			0
l14			0			1			0			1
l15			1			1			1			0
l16			0			0			0			0
Closed at t=							5					

From these results, it is confirmed that the number of opened DPs increases as the SL size grows, since it impacts the total demand level of the network. Under tight capacity attribute (TC), two to three DPs are opened for instances with 20 SLs (up to 6 with 50 SLs). Nevertheless, under large capacity (LC), the number of opened DPs is smaller compared to the ones under TC as DPs can, in the former case, accommodate more inbound flows from WPs. Regarding WPs, the opened number and their locations are quite stable in the instances (2 WPs with 20 SLs and 3 WPs at most with 50 SLs). Additionally, the location of both WPs and DPs and the throughput capacity levels are correlated with the demand process, the SLs dispersion, and costs. Interestingly, these latter are also impacted by DPs' zones as we can observe a change in the location decisions as well as in the DPs' zones for instance 5-4/12/20-25-(I1,LT,TC,.) with both NIT and NVT demand processes. Clearly, the obtained results highlight the sensitivity of the strategic location decisions because in several cases, the design structure varies between high and low transportation cost, which underlines the importance of integrating the routing sub-problem at the design level. For instance, we observe in Table 7 with 5-4/16/20-25-(I1,LT,TC,NIT) that the network opened DPs 4 and 15, whereas with 5-4/16/20-25-(I1,HT,TC,NIT), DPs 14 and 15 are opened. We notice that only in few instances the centralized DPs (*Area3*) which are underlined in these tables – were opened, which is due to their higher fixed costs. In addition, the SL dispersion – i.e., I1 vs I2 – impacts the DP location decisions, mainly under (.,LT,TC,NIT) attributes, as illustrated in Table 8. For example, for the instance size 5-4/12/20-25-, DPs 5 and 11 are opened under I2 attribute instead of 5 and 12 under I1. Even though the DP 5 is opened under both attributes settings, the timing is not the same – i.e., opening period $t = 2$ for I1 and $t = 1$ for I2. These results are congruent with recent findings in two-echelon urban networks that favor the multiplicity of surrounding locations and their interconnectivity rather than centralisation-prone practices.

5.2.3 Multi-period design behavior

Furthermore, a key finding is the network hedging capabilities in terms of location and capacity decisions in response to the demand uncertainty along time. For instance, under the (I1,LT,TC,NIT) configuration, the problem size 5-4/12/20-25- fixes DP 12 from the first design period, and further at period $t = 2$ it opens DP 5. The same instance under the NVT demand process fixes all the opening locations from the first design period. The instance 5-4/8/20-25-(I1,LT,TC,NVT) also attests to the necessity of the multi-period design flexibility proposed in this work. We see that DP 7 is opened at $t = 2$ to meet the increasing trend of the demand at periods one to three. This latter is then closed at $t = 5$ as the demand decreases at periods four and five. The results obtained justify the modeling efforts to include a multi-period design

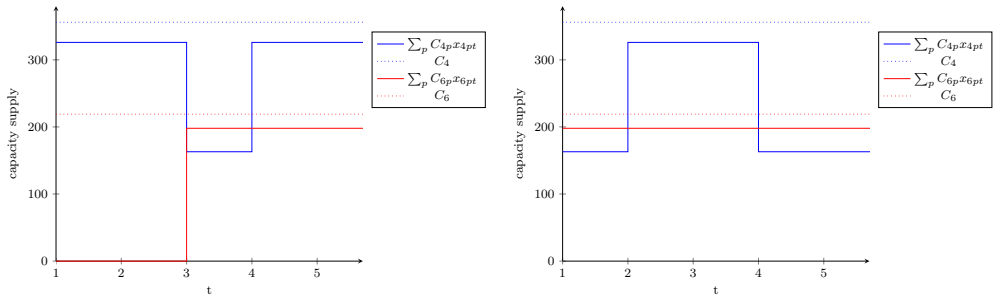
Table 8: Location decisions and their operating periods for the SL dispersion attribute (I1 vs I2)

	5-4/./20-25-				5-4/./50-15-			
	(I1,LT,TC,NIT)		(I2,LT,TC,NIT)		(I1,LT,TC,NIT)		(I2,LT,TC,NIT)	
$\|\mathcal{L}\|$	8	12	8	12	8	12	8	12
l1	0	0	0	0	0	0	0	0
l2	0	0	0	0	0	0	0	0
l3	0	0	0	0	4	0	3	0
l4	1	0	1	0	1	1	1	1
l5	0	2	0	1	1	1	1	1
l6	1	0	1	0	2	0	1	0
l7	0	0	0	0	1	1	1	1
l8	0	0	0	0	0	0	0	1
l9		0		0		0		0
l10		0		0		4		0
l11		0		1		0		0
l12		1		0		1		4
Closed at t=								

setting since we noticed that opening decisions adequately capture the variations in demand level. In addition, we notice a clear difference in the location of opened DPs when comparing solutions from the two demand processes. To emphasize this result, one can closely observe the instance 5-4/12/20-25-(I1,LT,TC,.) where there is no identical DP for NIT vs NVT. As for 5-4/16/15-25-(I1,LT,TC,.), we obtain 100% identical DPs for NIT vs NVT, but they differ in opening and closing periods (see Table 16). Therefore, the obtained results confirm the high variability in solutions in terms of location when different capacity levels, costs and SLs dispersion are considered, and the hedging capabilities of the 2E-SM-CLRP to uncertainty. Our results validate our modeling approach by adequately capturing the stochastic multi-period demand process through our two-stage stochastic formulation.

To complement the above analysis, we also investigate closely the evolution of the capacity decisions in the multi-period and uncertain setting. In some cases, opening/closing decisions over the periods are quite stable because adjustments to meet the demand is made through the capacity decisions, as reported in Figure 3. It depicts the capacity decisions produced by the model and contrast it to the evolution of each demand process along the planning horizon. This figure provide the results of instance 5-4/8/15-25-(I1,LT,TC,.) for both the NIT and NVT demand processes which produces design solutions with two DPs. Each solid line corresponds to the capacity available at each opened DP in period t – i.e., $\sum_p C_{lp} x_{lpt}$ – and each dotted line its predetermined capacity C_l where each color represent a separate opened DP. We use the same color if the same DP is opened under both demand processes. Figure 3 clearly illustrates the impact of the multi-period modeling approach where the capacity decisions for each opened DP are precisely adapted periodically. Even if the two demand processes of the instance 5-4/8/15-25-(I1,LT,TC,.) produce the same DP location decisions, the capacity decisions behave differently under each demand process to follow the time-varying demand process. This means that the two-stage model for the 2E-SM-CLRP mimics the behavior of a dynamic capacity expansion model with the inclusion of multi-period capacity decisions at the strategic level. The same behavior is observed in other instances, reported in results Table 18 and Figures 4 and 5 in Appendix B.2.

Furthermore, one can confirm the value of the design solutions produced by the multi-period modeling approach for its hedging capabilities. Table 9 provides the best upper and lower bounds from the static design modeling approach, the optimality gap(%) as well as an evaluation of the cost loss with respect to the multi-period approach – i.e., $\frac{\left(ub^{static} - ub^{multi-period}\right)}{ub^{multi-period}} \times 100$. In Table 9, we observe that in almost all cases the static design setting provides higher expected cost. The cost loss increases with the problem size and it reaches more than 4%. The largest losses are observed under the NVT process as it presents more variability. This is because the static model anticipates the DP openings and the required capacity level at the first design period (no further changes are allowed), which result in an over-estimation of the capacity allocated.



(a) NIT process (b) NVT process

Figure 3: Multi-period Capacity-allocation decisions versus the a priori capacity C_l for 5-4/8/15-25-(I1,LT,TC,.)

Table 9: Comparison between static and multi-period modeling approach

		Static modeling approach			Cost loss (%)
		ub	lb	Gap(%)	
5-4/./20-25-	(I1,LT,TC,NIT)	86442.6	86438.67	0.01	0.55
	(I1,HT,TC,NIT)	124334	124322	0.01	0.57
	(I1,LT,TC,NVT)	99324.78	99324.8	0	1.8
	Average	103367.13	103361.82	0	0.97
5-4/./50-15-	(I1,LT,TC,NIT)	196726.37	196542.65	0.09	2.64
	(I1,HT,TC,NIT)	282394.07	282277.50	0.04	2.78
	(I1,LT,TC,NVT)	209313.98	209228.03	0.04	3.94
	Average	229478.14	229349.4	0.06	3.12

5.2.4 Sensitivity analysis

This last subsection presents sensitivity analyses of the results to the capacity ratio and to the platforms positioning using instance attributes (I1, LT, TC, NIT) and structure 5-4/8/.-, with sizes 15, 20 and 50.

Capacity ratio behavior. We introduce here a capacity factor to scale the capacity between WPs and DPs and solve the selected instances with different values of the ratio while keeping the same locations. The results are presented in Table 10 in which we provide the opened platforms for both WPs and DPs and DPs throughput over the 5 design periods. The opening period for a given platform is indicated between, parentheses if deferred in time. The interval [1.5; 3.5] refers to the baseline which mean that in the extreme case the capacity factor between WPs and DPs is of 1.5. Table 10 shows how an increasing capacity factor impacts the decisions on the opened platforms and the throughput capacity. We notice a stability in the number of opened WPs between the baseline and the two cases with accentuated factors. When contrasted with the baseline, the two expanded cases are also stable in the location of WPs for all the instance sizes. For the second echelon, the solutions show also a stability in terms of the number of opened DPs, but a significant sensitivity of the location and the capacity level. It is clear that DPs related decisions are a trade-off between first-echelon WPs capacity and openings, and second-echelon SLs demands. These results further underline the importance to model jointly the two-echelons of distribution systems, which is congruent to findings in Ben Mohamed et al. (2020).

Platforms Position. We test here the sensitivity of our results to the location of WPs. To do so, we changed the position of 50% of potential WPs to *Area2* at a higher cost and replicated our tests with various capacity ratios. The results are described in Table 11 where underlined WPs are those located in *Area2*. From these results, one can observe a high stability in the two-echelon design structure, where in the majority of the cases similarities in WPs and DPs openings and capacity levels are observed (contrasted with Table 10). In almost all cases, for a given capacity ratio, we obtain the same design with WPs in and outside *Area2*. WPs within *Area2* are essentially opened in instances with 50 SLs where more than one WP is needed. The

Table 10: Solutions for instances 5-4/8/.-.-(I1, LT, TC, NIT) according to capacity ratio

Instance	Indicators	Capacity factor		
		[1.5; 3.5]	[2; 3.5]	[2.7; 4]
	WPs	p3	p2	p2
./15-25-	DPs	l3, l8	l3, l7	l4, l5
	DPs throughput	(385, 385, 385, 385, 385)	(380, 380, 380, 380, 380)	(364, 364, 364, 364, 364)
	WPs	p3	p2	p2
./20-25-	DPs	l3, l7(t=3), l8	l3, l6	l4, l5, l6(t=2)
	DPs throughput	(385, 385, 556, 556, 556)	(506, 506, 506, 506, 506)	(364, 526, 526, 526, 526)
⋆	WPs	p3, p4	p2, p4	p2, p4
./50-15-	DPs	l1, l3, l4(t=5), l5, l7, l8(t=3)	l3, l5, l6, l7(t=3), l8	l1, l3(t=3), l4, l5, l6, l8
	DPs throughput	(926, 1011, 1086, 1181, 1200)	(928, 1060, 1126, 1157, 1244)	(983, 1006, 1188, 1196, 1204)

⋆ The best feasible solution is reported

Table 11: Solutions for instances 5-4/8/.-.-(I1, LT, TC, NIT) according to capacity ratio with 50% of WPs in *Area2*

Instance	Indicators	Capacity factor		
		[1.5; 3.5]	[2; 3.5]	[2.7; 4]
	WPs	p3	p2	p2
./15-25-	DPs	l3, l8	l3, l7	l4, l5
	DPs throughput	(385, 385, 385, 385, 385)	(380, 380, 380, 380, 380)	(364, 364, 364, 364, 364)
	WPs	p2	p2	p2
./20-25-	DPs	l3, l6	l3, l6	l4, l5, l6(t=2)
	DPs throughput	(506, 506, 506, 506, 506)	(506, 506, 506, 506, 506)	(364, 526, 526, 526, 526)
⋆	WPs	p2, p3	p2, p4	p2, p4
./50-15-	DPs	l1, l3, l4, l6, l7(t=5), l8(t=3)	l3, l5, l6, l7(t=3), l8	l1, l3, l4, l5, l6, l8(t=3)
	DPs throughput	(948, 1007, 1103, 1153, 1268)	(959, 1011, 1096, 1157, 1258)	(949, 1026, 1096, 1158, 1204)

⋆ The best feasible solution is reported

specific case observed in the instance with 20 SLs under the ratio [1.5; 3.5] in which WP 2 is chosen instead of WP 3, is explained by fixed costs tradeoffs. From these complementary experiments, one can further underline that the design decisions are driven by the location, the capacity and the costs, and that the location of nearby WPs does not impact the two-echelon posture of the distribution network.

6 Conclusion

In this paper, we introduced the two-echelon stochastic multi-period capacitated location-routing problem (2E-SM-CLRP). The problem is characterized as a hierarchical decision process involving a design level in which network location and capacity allocation decision are taken, and an operational level dealing with transportation decisions of the second echelon. A stochastic multi-period characterization of the planning horizon is considered, shaping the evolution of the uncertain SL demand and costs. This problem is formulated as a two-stage stochastic integer program with recourse and solved by an exact logic-based Benders decomposition (LBDD) algorithm. In the first-stage, location (WPs and DPs) and capacity assignment decisions are fixed by solving the Benders master problem. The resulting sub-problem is a capacitated vehicle-routing problem with capacitated multiple depots (CVRP-CMD) which is further decomposed by period and scenario and then solved using the state-of-the-art branch-cut-and-price algorithm. Two families of Benders cuts are proposed to cut off infeasible solutions and to help converging to the optimal solution of the 2E-SM-CLRP. To the best of our knowledge, this is the first exact method that is proposed for a stochastic multi-period location-routing problem.

The proposed method is able to solve optimally instances containing up to 50 SLs and 25 demand scenarios under a 5-year planning horizon, and provides good lower bounds for the

instances that it cannot solve to optimality within the time limit of three days. Extensive computational experiments provide relevant managerial insights regarding the impact of uncertainty on the 2E-SM-CLRP, in addition to the effectiveness of multi-period modeling setting and its flexibility to adapt its hedging capabilities over time. Moreover, the obtained results validate our proposed two-stage stochastic modeling approach to capture the essence of the dynamic setting in the 2E-SM-CLRP. They also shown how design decisions are driven by the location, the capacity and the costs, and confirmed the necessity of the two-echelon posture of the distribution network.

Although the LBBD provides good solutions, it might be worthwhile improving the proposed algorithm to reduce the running time and tackle larger instances. This can be done through the incorporation of several algorithmic features such as the scenario group cuts, pareto-optimal cuts and the use of a heuristic procedure. As for the routing sub-problems, in some instances with 50 SLs, the CVRP-CMD cannot be solved in an hour, whereas all standard multi-depot CVRP literature instances of similar size are solved much faster. This points out a much larger complexity of the CVRP-CMD involved in the 2E-SM-CLRP, with respect to the uncapacitated variant. Therefore, it would be interesting to develop new valid inequalities adapted to the CVRP-CMD to strengthen its set-partitioning formulation.

Future research could focus on additional features encountered in practice such as synchronization constraints at intermediate distribution platforms, and constraints limiting route length. Moreover, since routing decisions are often used to anticipate the decision of the operational level, an interesting research perspective of this work would be to examine the route approximation formulæ instead of explicitly computing the vehicle routes. This may speed up the decision process significantly.

Acknowledgements

Experiments presented in this paper were carried out using the PlaFRIM (Federative Platform for Research in Computer Science and Mathematics), created under the Inria PlaFRIM development action with support from Bordeaux INP, LABRI and IMB and other entities: Conseil Régional d'Aquitaine, Université de Bordeaux, CNRS and ANR in accordance to the "Programme d'Investissements d'Avenir" (see www.plafrim.fr/en/home).

HS gratefully acknowledges the financial support of the FMJH 'Program Gaspard Monge for optimization and operations research and their interactions with data science', and EDF.

References

M. Albareda-Sambola, E. Fernández, and G. Laporte. Heuristic and lower bound for a stochastic location-routing problem. *European Journal of Operational Research*, 179(3):940–955, 2007.

M. Albareda-Sambola, E. Fernández, and S. Nickel. Multiperiod location-routing with decoupled time scales. *European Journal of Operational Research*, 217(2):248–258, 2012.

M. Albareda-Sambola, A. Alonso-Ayuso, L. F. Escudero, E. Fernández, and C. Pizarro. Fix-and-relax-coordination for a multi-period location–allocation problem under uncertainty. *Computers & Operations Research*, 40(12):2878–2892, 2013.

D. Ambrosino and M. G. Scutellà. Distribution network design: new problems and related models. *European Journal of Operational Research*, 165(3):610–624, 2005.

R. Baldacci, N. Christofides, and A. Mingozzi. An exact algorithm for the vehicle routing problem based on the set partitioning formulation with additional cuts. *Mathematical Programming*, 115(2):351–385, 2008.

R. Baldacci, A. Mingozzi, and R. W. Calvo. An exact method for the capacitated location-routing problem. *Operations Research*, 59(5):1284–1296, 2011.

J. M. Belenguer, E. Benavent, C. Prins, C. Prodhon, and R. W. Calvo. A branch-and-cut method for the capacitated location-routing problem. *Computers & Operations Research*, 38(6):931–941, 2011.

I. Ben Mohamed, W. Klibi, and F. Vanderbeck. Designing a two-echelon distribution network under demand uncertainty. *European Journal of Operational Research*, 280(1):102–123, 2020.

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

M. Boccia, T. G. Crainic, A. Sforza, and C. Sterle. A metaheuristic for a two echelon location-routing problem. In *In International Symposium on Experimental Algorithms*, pages 288–301. Springer, 2010.

Sunil Chopra. The evolution of omni-channel retailing and its impact on supply chains. *Transportation Research Procedia*, 30:4–13, 2018.

C. Contardo, V. C Hemmelmayr, and T. G. Crainic. Lower and upper bounds for the two-echelon capacitated location-routing problem. *Computers & Operations Research*, 39(12):3185–3199, 2012.

C. Contardo, J-F. Cordeau, and B. Gendron. A computational comparison of flow formulations for the capacitated location-routing problem. *Discrete Optimization*, 10(4):263–295, 2013.

C. Contardo, J-F. Cordeau, and B. Gendron. An exact algorithm based on cut-and-column generation for the capacitated location-routing problem. *INFORMS Journal on Computing*, 26 (1):88–102, 2014a.

C. Contardo, J-F. Cordeau, and B. Gendron. A grasp+ ilp-based metaheuristic for the capacitated location-routing problem. *Journal of Heuristics*, 20(1):1–38, 2014b.

I. Correia, T. Melo, and F. Saldanha da Gama. Comparing classical performance measures for a multi-period, two-echelon supply chain network design problem with sizing decisions. *Computers & Industrial Engineering*, 64(1):366–380, 2013.

T. G. Crainic and G. Laporte. Planning models for freight transportation. *European Journal of Operational Research*, 97(3):409–438, 1997.

T. G. Crainic, N. Ricciardi, and G. Storchi. Advanced freight transportation systems for congested urban areas. *Transportation Research Part C: Emerging Technologies*, 12(2):119–137, 2004.

Teodor Gabriel Crainic and Benoit Montreuil. Physical internet enabled hyperconnected city logistics. *Transportation Research Procedia*, 12:383–398, 2016.

Teodor Gabriel Crainic, Fausto Errico, Walter Rei, and Nicoletta Ricciardi. Modeling demand uncertainty in two-tier city logistics tactical planning. *Transportation Science*, 50(2):559–578, 2016.

R. Cuda, G. Guastaroba, and M. G. Speranza. A survey on two-echelon routing problems. *Computers & Operations Research*, 55:185–199, 2015.

M. Darvish, C. Archetti, L. C. Coelho, and M. S. Speranza. Flexible two-echelon location routing problem. *European Journal of Operational Research*, 277(3):1124–1136, 2019.

M. S. Daskin. *Network and Discrete Location: Models, Algorithms and Applications*. Wiley-Interscience Publication, New York, 1995.

M. Drexl and M. Schneider. A survey of variants and extensions of the location-routing problem. *European Journal of Operational Research*, 241(2):283–308, 2015.

M. S. Farham, C. Iyigun, and H. Süral. The two-echelon location-routing problem with time windows: Formulation, branch-and-price, and clustering. Technical Report, 2020.

A. M. Geoffrion and G. W. Graves. Multicommodity distribution system design by benders decomposition. *Management Science*, 20(5):822–844, 1974.

M. C. Georgiadis, P. Tsiakis, P. Longinidis, and M. K. Sofioglou. Optimal design of supply chain networks under uncertain transient demand variations. *Omega*, 39(3):254–272, 2011.

V. C Hemmelmayr, J-F. Cordeau, and T. G. Crainic. An adaptive large neighborhood search heuristic for two-echelon vehicle routing problems arising in city logistics. *Computers & Operations Research*, 39(12):3215–3228, 2012.

John N. Hooker. Logic-based Benders decomposition for large-scale optimization. In Jesús M. Velásquez-Bermúdez, Marzieh Khakifirooz, and Mahdi Fathi, editors, *Large Scale Optimization in Supply Chains and Smart Manufacturing: Theory and Applications*, pages 1–26. Springer International Publishing, Cham, 2019.

S. Jacobsen and O. Madsen. A comparative study of heuristics for a two-level location-routing problem. *European Journal of Operational Research*, 6:378–387, 1980.

M. Janjevic, M. Winkenbach, and D. Merchán. Integrating collection-and-delivery points in the strategic design of urban last-mile e-commerce distribution networks. *Transportation Research Part E: Logistics and Transportation Review*, 131:37–67, 2019.

S. D. Jena, J-F. Cordeau, and B. Gendron. Dynamic facility location with generalized modular capacities. *Transportation Science*, 49(3):484–499, 2015.

M. Jepsen, B. Petersen, S. Spoorendonk, and D. Pisinger. Subset-row inequalities applied to the vehicle-routing problem with time windows. *Operations Research*, 56(2):497–511, 2008.

W. Klibi, F. Lasalle, A. Martel, and S. Ichoua. The stochastic multiperiod location transportation problem. *Transportation Science*, 44(2):221–237, 2010.

W. Klibi, A. Martel, and A. Guitouni. The impact of operations anticipations on the quality of stochastic location-allocation models. *Omega*, 62:19–33, 2016.

A. Klose and A. Drexl. Facility location models for distribution system design. *European Journal of Operational Research*, 162(1):4–29, 2005.

G. Laporte. Location-routing problems. In B. L. Golden and A. A. Assad, editors, *Vehicle Routing: Methods and Studies*, pages 163–198. North-Holland, Amsterdam, 1988.

G. Laporte and PJ. Dejax. Dynamic location-routing problems. *Journal of the Operational Research Society*, 40(5):471–482, 1989.

G. Laporte and Y. Nobert. A branch and bound algorithm for the capacitated vehicle routing problem. *Operations Research Spektrum*, 5(2):77–85, 1983.

G. Laporte, F. Louveaux, and H. Mercure. Models and exact solutions for a class of stochastic location-routing problems. *European Journal of Operational Research*, 39(1):71–78, 1989.

O. Madsen. Methods for solving combined two level location-routing problems of realistic dimensions. *European Journal of Operational Research*, 12:295–301, 1983.

S. Martello and P. Toth. Lower bounds and reduction procedures for the bin packing problem. *Discrete Applied Mathematics*, 28(1):59–70, 1990.

M. T. Melo, S. Nickel, and F. Saldanha da Gama. Facility location and supply chain management–a review. *European Journal of Operational Research*, 196(2):401–412, 2009.

D. Merchán and M. Winkenbach. High-resolution last-mile network design. *City Logistics 3: Towards Sustainable and Liveable Cities*, pages 201–214, 2018.

S. M. Mirhedayatian, T. G. Crainic, M. Guajardo, and S. W. Wallace. A two-echelon location-routing problem with synchronisation. *Journal of the Operational Research Society*, pages 1–16, 2019.

G. Nagy and S. Salhi. Location-routing: Issues, models and methods. *European Journal of Operational Research*, 177(2):649–672, 2007.

V. P. Nguyen, C. Prins, and C. Prodhon. Solving the two-echelon location routing problem by a grasp reinforced by a learning process and path relinking. *European Journal of Operational Research*, 216(1):113–126, 2012a.

V. P. Nguyen, C. Prins, and C. Prodhon. A multi-start iterated local search with tabu list and path relinking for the two-echelon location-routing problem. *Engineering Applications of Artificial Intelligence*, 25(1):56–71, 2012b.

OpenMP Architecture Review Board. The OpenMP API specification for parallel programming. https://www.openmp.org/, 2000.

D. Pecin, A. Pessoa, M. Poggi, and E. Uchoa. Improved branch-cut-and-price for capacitated vehicle routing. *Mathematical Programming Computation*, 9(1):61–100, 2017.

G. Perboli, R. Tadei, and D. Vigo. The two-echelon capacitated vehicle routing problem: models and math-based heuristics. *Transportation Science*, 45:364–380, 2011.

Artur Pessoa, Ruslan Sadykov, Eduardo Uchoa, and François Vanderbeck. A generic exact solver for vehicle routing and related problems. In Andrea Lodi and Viswanath Nagarajan, editors, *Integer Programming and Combinatorial Optimization*, volume 11480 of *Lecture Notes in Computer Science*, pages 354–369, Cham, 2019. Springer International Publishing.

B. S. Pimentel, G. R. Mateus, and F. A. Almeida. Stochastic capacity planning and dynamic network design. *International Journal of Production Economics*, 145(1):139–149, 2013.

C. Prodhon and C. Prins. A survey of recent research on location-routing problems. *European Journal of Operational Research*, 238(1):1–17, 2014.

R. Sadykov, E. Uchoa, and A. Pessoa. A bucket graph based labeling algorithm with application to vehicle routing. *Transportation Science*, accepted, 2020.

S. Salhi and G. K Rand. The effect of ignoring routes when locating depots. *European Journal of Operational Research*, 39(2):150–156, 1989.

T. Santoso, S. Ahmed, M. Goetschalckx, and A. Shapiro. A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, 167(1):96–115, 2005.

Martin Savelsbergh and Tom Van Woensel. 50th anniversary invited article—city logistics: Challenges and opportunities. *Transportation Science*, 50(2):579–590, 2016.

Christoph Schneeweiss. *Distributed decision making*. Springer Science & Business Media, 2003.

M. Schneider and M. Löffler. Large composite neighborhoods for the capacitated location-routing problem. *Transportation Science*, 53(1):301–318, 2019.

P. Schütz, A. Tomasgard, and S. Ahmed. Supply chain design under uncertainty using sample average approximation and dual decomposition. *European Journal of Operational Research*, 199(2):409–419, 2009.

M. Schwengerer, S. Pirkwieser, and G. R. Raidl. A variable neighborhood search approach for the two-echelon location-routing problem. In *European Conference on Evolutionary Computation in Combinatorial Optimization*, pages 13–24. Springer, 2012.

A. Shapiro. Monte carlo sampling methods. *Handbooks in Operations Research and Management Science*, 10:353–425, 2003.

A. Shapiro, D. Dentcheva, and A. Ruszczynski. Lectures on stochastic programming: modeling and theory. *The Society for Industrial and Applied Mathematics and the Mathematical Programming Society, Philadelphia, USA*, 2009.

Z. Shen. Integrated supply chain design models: a survey and future research directions. *Journal of Industrial and Management Optimization*, 3(1):1, 2007.

Z-J. M. Shen and L. Qi. Incorporating inventory and routing costs in strategic location models. *European journal of operational research*, 179(2):372–389, 2007.

A. Snoeck, M. Winkenbach, and E.E Mascarino. Establishing a robust urban logistics network at femsa through stochastic multi-echelon location routing. *City Logistics 2: Modeling and Planning Initiatives*, pages 59–78, 2018.

C. Sterle. *Location-Routing models and methods for Freight Distribution and Infomobility in City Logistics*. PhD thesis, Università degli Studi di Napoli "Federico II", 2010.

P. Toth and D. Vigo. *Vehicle Routing: Problems, Methods, and Applications*. SIAM, 2 edition, 2014.

F. Vanderbeck, R. Sadykov, and I. Tahiri. Bapcod–a generic branch-and-price code. https://realopt.bordeaux.inria.fr/?pageid=2, 2018.

D. Weinswig. Sam's club is closing one-tenth of its stores and converting 12 to e-commerce hubs: Proof that store closures have become respectable? https://www.linkedin.com/pulse/sams-club-closing-one-tenth-its-stores-converting-12-hubs-weinswig/, 2018. Accessed: 2018-03-05.

M. Winkenbach, P. R. Kleindorfer, and S. Spinler. Enabling urban logistics services at la poste through multi-echelon location-routing. *Transportation Science*, 50(2):520–540, 2016.

Q. Zhao, W. Wang, and R. De Souza. A heterogeneous fleet two-echelon capacitated location-routing model for joint delivery arising in city logistics. *International Journal of Production Research*, 56(15):5062–5080, 2018.

A Complement to the solution method

A.1 Bin packing problem

In this section, we give the description of the bin packing problem formulated for each period and each scenario, denoted as ($\text{BPP}_{t\omega}$), and present its mathematical formulation. Then, we briefly describe the column generation algorithm used to solve the linear relaxation of the ($\text{BPP}_{t\omega}$).

Consider a large set of bins (i.e. vehicles) with capacity q and a set of $|\mathcal{J}|$ items (i.e. customers) with weights $d_{j\omega t}$ to pack into bins. The objective is to find the minimum number of bins required to pack the set of items so that the capacity of the bins is not exceeded. Let $\mathcal{B}$ be the family of all the subsets of items which fit into one bin, i.e., the solutions to a subproblem. We define the parameter x_j^B that takes 1 if item $j \in \mathcal{J}$ is in set $B \in \mathcal{B}$. Let λ^B be the binary variable taking value 1 if the corresponding subset of items B is selected to fill one bin. The set covering reformulation is:

$$\min \sum_{B \in \mathcal{B}} \lambda^B \tag{37}$$

$$\text{S. t.} \quad \sum_{B \in \mathcal{B}} x_j^B \lambda^B \geq 1 \qquad j = 1, ..., \mathcal{J} \tag{38}$$

$$\lambda_k \in \{0, 1\} \tag{39}$$

The linear relaxation of (37)-(39) is solved by column generation to provide a lower bound Γ. This lower bound is obtained by iteratively solving:

- the restricted master problem (RMP) which is the linear relaxation of (37)-(39) with a restricted number of variables;
- and the pricing problem which determines whether there exists a variable λ^B to be added to (RMP) in order to improve its current solution; this refers to solve a knapsack problem to get the set $B \in \mathcal{B}$, satisfying capacity constraints, and yielding to the minimum reduced cost column for (RMP).

Let π_j be the dual variable associated to constraints (38), the pricing problem for the (BPP$_{t\omega}$) is written as:

$$\max_{j=1..\mathcal{J}} \pi_j z_j \tag{40}$$

$$\text{S. t.} \quad \sum_{j=1..\mathcal{J}} d_{j\omega t} z_j \leq q \tag{41}$$

$$z_j \in \{0,1\} \qquad j \in \mathcal{J} \tag{42}$$

A column generated by solving the knapsack problem (40)-(42) will terminate column generation procedure in case its reduced cost $1 - \sum_j \pi_j z_j$ turns negative.

We apply the MIP solver *CPLEX* to the pricing formulation (40)-(42). Then, the linear relaxation of (37)-(39) is terminated by LP solver *CPLEX*. This leads to the lower bound Γ.

A.2 Preprocessing function for improving the minimum number of second-level vehicles

Algorithm 2 details the procedure developed to improve the minimum number of vehicles required in the CVRP-CMD.

B Results

B.1 Parameters tuning

The Benders decomposition algorithm terminates when one of the following criteria is met: (i) the optimality gap between the upper and lower bounds is below an $\epsilon = 0.0005$ threshold – i.e., $(ub - lb)/ub < \epsilon$ – or (ii) the maximum time limit of 72 hours is reached. A time limit of 50 minutes is considered for each CVRP-CMD.

An important parameter to calibrate for stochastic models is the number N of scenarios to include in the optimization phase. Under a scenario-based optimization approach, generating the adequate set of scenarios Ω could be complex due to the high enumeration issue induced by continuous normal distribution (Shapiro et al., 2009). Assessing their probabilities also entails a tremendous effort. A combination of the Monte Carlo sampling methods (Shapiro, 2003) and the sample average approximation technique (SAA) (Shapiro et al., 2009) helps in finding a good trade-off in terms of the scenario probability estimation and the sufficient number of scenarios to consider in the model. The SAA consists in generating for each time period t, before the optimization procedure, an independent sample of N equiprobable scenarios $\Omega_t^N \subset \Omega_t$ from the initial probability distribution, which removes the need to explicitly compute the scenario demand probabilities $p(\omega)$. The quality of the solution obtained with this approach improves as the scenario sample size N increases. However, one would choose N taking into account the trade-off between the quality of the obtained design and the computational effort needed to solve

Algorithm 2 Preprocessing function for $\bar{\Gamma}^k_{t\omega}$

INPUT: t, ω, $d^{t\omega}$, Q^2, $\bar{x}^k_t$
Initialize $\gamma_t = [0 \ \forall \ l \in \mathcal{L}]$ and $rx_t = [0 \ \forall \ l \in \mathcal{L}]$
Compute the total capacity assignment at t, $C = \sum_{l \in \mathcal{L}} \bar{x}_{tl}$
Compute the total demand under t, ω, $D = \sum_{j \in \mathcal{J}} d^{t\omega}_j$
$id \leftarrow 0$, $cumul \leftarrow 0$
for all $l \in \mathcal{L}$ **do**
  if $\bar{x}_{tl} > 0$ **then**
    Compute $\gamma_{tl} = \lceil \frac{\bar{x}_{tl} + (D - C)}{Q^2} \rceil$
  end if
end for
for all $l \in \mathcal{L}$ **do**
  if $\bar{x}_{tl} - \gamma_{tl} * Q^2 > 0$ **then**
    Compute $rx_{tl} \leftarrow \bar{x}_{tl} - \gamma_{tl} * Q^2$
  end if
end for
Compute $\bar{\Gamma}^k_{t\omega} = \sum_{l \in \mathcal{L}} \gamma_{tl}$
Compute the unsatisfied demand $rD = D - \bar{\Gamma}^k_{t\omega} * Q^2$
if $rD > 0$ **then**
  Sort rx_t in descending order
  while $cumul < rD$ **do**
    $cumul \leftarrow cumul + rx_{tl}$
    $id \leftarrow id + 1$
    $\bar{\Gamma}^k_{t\omega} \leftarrow \bar{\Gamma}^k_{t\omega} + 1$
  end while
end if

the problem. Thus, to determine the best value of N, solving the problem with M independent samples of demand repeatedly can be more efficient. This leads to a maximum of M different design decisions – i.e. location and capacity allocation. It is worth to note that some samples may provide identical design decisions. The average value of the M expected costs based on the N scenarios gives a statistical lower bound. Then, we evaluate the obtained designs based on the expected daily routing cost. We fix the first-stage decisions according to each of these different designs and solve the resulting problem for $N' = |\Omega^{N'}_t| \gg N$ independent scenarios to get an upper bound on the optimal solution of the problem. Finally, a statistical optimality gap is computed for each obtained design from these lower and upper bounds. For more details, interested reader is referred to (Santoso et al., 2005) and (Schütz et al., 2009). We notice that an external recourse option is added here, at a high cost, in order to guarantee the feasibility of all scenarios.

To apply the SAA technique, we solved $M = 10$ demand samples and used sample sizes of $N = 5, 10, 15$ and 25 scenarios for each time period t. The best feasible solution of each SAA sample is then stored as a candidate solution for valuation in the reference sample. The size of the reference sample per period t is set to $N' = |\Omega^{N'}_t| = 120$ scenarios. The average gap values for problem sizes 5-4/8/15-N- and 5-4/8/20-N- under (I1,LT,TC,NIT) instance configuration using the different values of N are summarized in Table 12.

Table 12: Average statistical optimality gap values for (I1,LT,TC,NIT) instances

$\text{Gap}_{N,120}$ (in %)				
Problem size	Sample size (N)			
	5	10	15	25
5-4/8/15-N-	2.11	0.85	0.39	-0.14
5-4/8/20-N-	2.55	1.07	1.14	0.34

Table 12 shows that the optimality gap improves as the sample size N increases and converges to 0%. Samples of $N = 15$ and 25 scenarios provide satisfactory results, optimality gap is generally less than 1% for both instances. Moreover, we note that the decisions produced with alternative samples ($M = 10$) present a high similarity in terms of the opened DPs and the inbound allocation. However, the solution time increases considerably with the sample size. Accordingly, the sample size of $N = 15$ is retained as the best trade-off to use in the experiments for instances with 50 ship-to locations and $N = 25$ is selected for instances with 15 and 20 ship-to locations. Recall that when N scenarios are used in the SAA model, $5 \times N$ instances are then sampled from the probability distribution as the planning horizon includes 5 periods.

B.2 Complementary results

Tables 13, 14 and 15 present complementary results on the algorithm performance described in subsection 5.2.1.

Table 13 provides detailed results of the sequential optimization approach, while Tables 14 and 15 detail the results obtained for instances with 12 and 16 DPs and 15, 20, 50 SLs, respectively.

Tables 16 and 17 illustrate design decisions obtained with 15 and 50 SLs. They complete the analysis in subsection 5.2.2.

Table 13: Average results under attributes (.,.,.,.) with sequential approach

Instance	N	#Opt	#CombCuts	#OptCuts	#Iter	Gap(%)	Sequential computing time		
							MP	CVRP-CMD	total
5-4/8/15-	5	24/24	2	6.4	8.4	0.0	30s	2m	3m48s
	10	24/24	1.8	6.3	8.1	0.0	40s	3m33s	6m55s
	15	24/24	2	6.2	8.2	0.0	1m	3m40s	8m39s
	25	24/24	2	6.7	8.7	0.0	1m52s	5m29s	13m49s
5-4/12/15-	5	15/15	2	6.8	8.8	0.0	2m47s	4m6s	8m20s
	10	15/15	1.8	7.4	9.2	0.0	2m50s	6m	11m32s
	15	15/15	2	7.2	9.2	0.0	3m	11m	18m2s
	25	15/15	2	7.6	9.6	0.0	2m50s	13m38s	24m8s
5-4/16/15-	5	12/12	2.2	7.4	9.7	0.0	2m16s	7m	11m39s
	10	12/12	2.2	7.1	9.3	0.0	2m34s	6m41s	11m29s
	15	12/12	2.3	7.2	9.5	0.01	2m28s	10m7s	17m29s
	25	12/12	2	7	9	0.01	8m30s	22m39s	38m21s
5-4/8/20-	5	24/24	2.1	8.9	11	0.0	2m	33m	36m
	10	24/24	2	8.7	10.7	0.0	2m30s	55m54s	1h2m
	15	24/24	2	8.8	10.8	0.0	2m27s	1h38m	1h46m
	25	24/24	2.1	9	11.1	0.0	4m20s	1h45m	2h
5-4/12/20-	5	15/15	2	9.4	11.4	0.01	1m50s	35m32s	39m9s
	10	15/15	2.2	8.6	10.8	0.0	2m56s	1h19m	1h25m
	15	15/15	2.6	8.6	11.2	0.01	7m7s	1h57m	2h11m
	25	15/15	2	8.2	10.2	0.0	4m27s	1h19m	1h34m
5-4/16/20-	5	12/12	2	8.9	10.9	0.0	4m38s	22m32s	29m2s
	10	12/12	2.1	9.8	11.9	0.0	4m2s	1h2m	1h10m
	15	12/12	2.5	9.8	12.3	0.0	16m30s	2h26m	2h49m
	25	12/12	2.4	9.9	12.3	0.01	1h16m	3h58m	5h25m

Table 14: Detailed results for 5-4/12/15-25-, 5-4/12/20-25- and 5-4/12/50-15-

		5-4/12/15-25-				5-4/12/20-25-				5-4/12/50-15-			
		ub	lb	Gap(%)	Time	ub	lb	Gap(%)	Time	ub	lb	Gap(%)	Time
(I2, LT, TC, NIT)	1	71146.4	71146.4	0.00	8m52s	76597.5	76597.5	0.00	1h50m	192216.0	192136.0	0.04	11h9m
	2	75246.1	75246.1	0.00	9m36s	90515.9	90489.4	0.03	13m5s	183250.0	183165.0	0.05	57h16m
	3	82505.5	82505.5	0.00	10m12s	84316.8	84316.8	0.00	1h2m	190410.9	190228.3	0.10	-
average		76299.3	76299.3	0.00	9m33s	83810.1	83801.2	0.01	1h2m	188625.6	188509.8	0.06	34h12m
(I1, LT, TC, NIT)	1	69862.2	69862.2	0.00	9m51s	79842.3	79842.3	0.00	17m30s	188827.0	188539.1	0.15	-
	2	75536.5	75536.5	0.00	9m39s	88416.3	88416.3	0.00	15m41s	182033.0	181944.0	0.05	5h36m
	3	76998.7	76998.7	0.00	9m48s	82822.0	82822.0	0.00	15m33s	188881.3	188686.4	0.10	-
average		74132.5	74132.5	0.00	9m46s	83693.5	83693.5	0.00	16m14s	186580.4	186389.8	0.10	5h36m
(I1, HT, TC, NIT)	1	98256.0	98256.0	0.00	9m35s	117079.0	117079.0	0.00	19m21s	275113.3	272759.1	0.86	-
	2	105654.0	105654.0	0.00	10m28s	126107.0	126098.0	0.01	14m36s	263067.0	262956.0	0.04	17h43m
	3	109252.0	109251.0	0.00	10m28s	119896.0	119896.0	0.00	23m55s	272392.7	269664.1	1.00	-
average		104387.3	104387.0	0.00	10m44s	121027.3	121024.3	0.00	19m17s	270191.0	268459.7	0.63	17h43m
(I1, LT, LC, NIT)	1	85410.0	85410.0	0.00	10m1s	92902.8	92902.8	0.00	13m31s	190269.0	190259.0	0.01	36m44s
	2	85285.8	85285.8	0.00	9m19s	93485.5	93485.5	0.00	13m23s	180619.0	180619.0	0.00	34m56s
	3	83886.4	83886.4	0.00	9m42s	89401.7	89401.7	0.00	13m15s	190580.0	190551.0	0.02	36m55s
average		84860.7	84860.7	0.00	9m40s	91930.0	91930.0	0.00	13m23s	187156.0	187143.0	0.01	36m12s
(I1, LT, TC, NVT)	1	71174.0	71174.0	0.00	9m31s	96847.2	96847.2	0.00	13m32s	181685.2	181146.7	0.30	-
	2	56085.6	56085.6	0.00	9m4s	92481.1	92481.1	0.00	20m45s	214547.0	214475.0	0.03	36h21m
	3	77253.9	77253.9	0.00	9m47s	92953.2	92953.2	0.00	23m41s	194353.6	194160.2	0.10	-
average		68171.2	68171.2	0.00	9m27s	94093.8	94093.8	0.00	19m19s	196861.9	196594.0	0.14	36h21m

Table 5 and Figures 4 and 5 provide further results on the multi-period behavior as discussed in subsection 5.2.1.

In Figure 5 is shaped the obtained capacity decisions – i.e., $\sum_p C_{lp} x_{lpt}$ – for instance 5-4/8/20-25-(I1,LT,TC,.) under both demand processes using the static setting. Each opened DP is represented with a different color in the figure. For comparison purpose, we also draw with a dash-dotted line the capacity obtained using the multi-period approach in Figure 5 (Figure 4 for multi-period approach separately. From Figure 5(a), under the NIT process, we notice that both static and multi-period modeling settings converge to the same location openings, but therein

Table 15: Detailed results for 5-4/16/15-25-, 5-4/16/20-25- and 5-4/16/50-15-

		5-4/16/15-25-				5-4/16/20-25-				5-4/16/50-15-			
		ub	lb	Gap(%)	Time	ub	lb	Gap(%)	Time	ub	lb	Gap(%)	Time
(I1, LT, TC, NIT)	1	68308.6	68308.6	0.00	10m31s	77457.4	77457.4	0.00	14m40s	176344.0	176273.0	0.04	4h10m
	2	68268.8	68268.8	0.00	9m31s	83721.4	83721.4	0.00	45m22s	174587.0	174570.0	0.01	6h30m
	3	69865.8	69865.8	0.00	9m43s	84059.2	84059.2	0.00	31m36s	182674.0	182632.0	0.02	19h38m
average		68814.4	68814.4	0.00	9m55s	81746.0	81746.0	0.00	30m32s	177868.3	177825.0	0.02	10h6m
(I1, HT, TC, NIT)	1	97063.5	97063.5	0.00	10m47s	115242.0	115242.0	0.00	14m35s	261513.0	261384.0	0.05	6h28m
	2	97640.0	97640.0	0.00	11m3s	122079.0	122025.0	0.04	1h55m	254798.0	254725.0	0.03	29h1m
	3	102520.0	102491.0	0.03	9m57s	120723.0	120723.0	0.00	48m17s	267690.8	265847.9	0.69	-
average		99074.5	99064.8	0.01	10m35s	119348.0	119330.0	0.01	59m22s	261333.9	260652.3	0.26	17h44m
(I1, LT, LC, NIT)	1	84768.3	84768.3	0.00	9m16s	92933.6	92933.6	0.00	14m26s	187444.0	187444.0	0.00	43m30s
	2	68358.8	68358.8	0.00	9m18s	79351.7	79351.7	0.00	15m57s	151987.0	151987.0	0.00	39m28s
	3	69606.2	69606.2	0.00	10m12s	75279.5	75279.5	0.00	22m14s	158576.0	158522.0	0.03	52m5s
average		74244.4	74244.4	0.00	9m35s	82521.6	82521.6	0.00	17m32s	166002.3	165984.3	0.01	45m1s
(I1, LT, TC, NVT)	1	69772.4	69748.9	0.03	9m10s	93171.1	93171.1	0.00	13m58s	169987.0	169949.0	0.02	11h46m
	2	52241.9	52241.9	0.00	9m18s	86603.9	86560.8	0.01	18m1s	206768.8	206072.7	0.34	-
	3	69025.6	69025.6	0.00	9m14s	85880.8	85880.8	0.00	16m55s	192274.7	191670.8	0.31	-
average		63680.0	63672.1	0.01	9m14s	90003.9	90003.9	0.00	16m18s	189676.8	189230.8	0.22	11h46m

Table 16: Location decisions and their operating periods for 5-4/./15-25-(.,.,.,.)-3

	(I1,LT,TC,NIT)			(I1,HT,TC,NIT)			(I1,LT,TC,NVT)			(I1,LT,LC,NIT)		
$\|\mathcal{L}\|$	8	12	16	8	12	16	8	12	16	8	12	16
l1	0	0	0	0	0	0	0	0	0	0	0	0
l2	0	0	0	0	0	0	0	0	0	0	0	0
l3	0	0	0	0	0	0	0	0	0	0	0	0
l4	1	0	1	1	0	1	1	0	1	0	0	0
l5	0	0	0	0	0	0	0	0	0	1	0	0
l6	3	0	0	3	0	0	1	0	0	0	0	0
l7	0	4	0	0	4	0	0	2	0	0	1	0
l8	0	0	0	0	0	0	0	0	0	0	0	0
l9		0	0		0	0		0	0		0	0
l10		0	0		0	0		0	0		0	0
l11		0	0		0	0		0	0		0	0
l12		1	0		1	0		1	0		0	0
l13			0			0			0			0
l14			0			0			0			1
l15			0			0			0			0
l16			3			3			2*			0
Closet at t=									4			

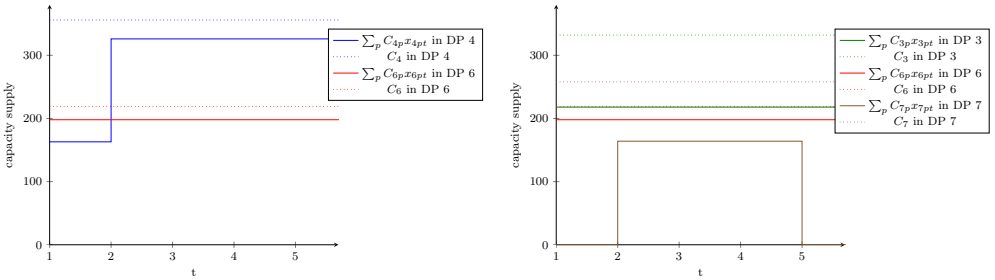
Table 17: Location decisions and their operating periods for 5-4/./50-15-(.,.,.,.)-3

	(I1,LT,TC,NIT)			(I1,HT,TC,NIT)			(I1,LT,TC,NVT)			(I1,LT,LC,NIT)		
$\|\mathcal{L}\|$	8	12	16	8	12	16	8	12	16	8	12	16
l1	0	0	0	0	0	0	0	0	0	0	0	0
l2	0	0	0	0	0	0	1	0	0	0	0	0
l3	4	0	0	2	1	0	0	1	0	0	0	0
l4	1	1	1	1	0	1	1	2*	1	0	1	1
l5	1	1	0	1	2	0	1*	1	0	1	0	0
l6	2	0	0	4	0	0	1$^{\diamond}$	0	0	0	0	0
l7	1	1	2	1	1	1	3	1	2	0	1	1
l8	0	0	0	0	0	0	0	0	0	1	0	0
l9		0	0		0	0		0	0		0	0
l10		4	0		0	0		0	0		0	0
l11		0	1		4	1		0	3		0	0
l12		1	0		1	0		1	1		0	0
l13			0			0			0			0
l14			0			0			0			0
l15			4			4			1*			0
l16			1			1			0			0
Closed at t=							*5,$^{\diamond}$4		*4	*5		

the static setting fixes all its capacity at the first period for DP 4 contrary to the multi-period approach. Under the NVT process, both approaches have only DP 3 in common fixing the same level of capacity from the beginning of the planning horizon. As mentioned, the static approach

Table 18: Location decisions under static modeling approach for 5-4/./20-25-(.,.,.,.)-3

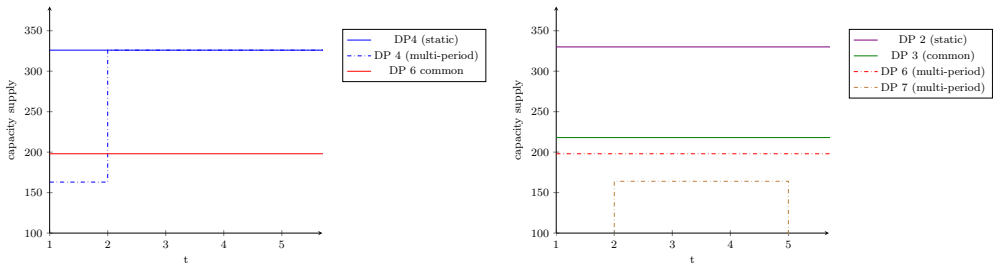
	(I1,LL,TC,NIT)		(I1,HL,TC,NIT)		(I1,LL,TC,NVT)	
$\|\mathcal{L}\|$	8	12	8	12	8	12
l1	0	0	0	0	0	0
l2	0	0	0	0	1	0
l3	0	0	0	0	1	1
l4	1	0	1	0	0	0
l5	0	0	0	0	0	0
l6	1	0	1	0	0	0
l7	0	1	0	1	0	1
l8	0	0	0	0	0	0
l9		0		0		0
l10		0		0		0
l11		0		0		0
l12		1		1		0



(a) NIT process (b) NVT process

Figure 4: Capacity-allocation decisions from multi-period modeling approach versus the apriori capacity C_l for 5-4/8/20-25-(I1,LT,TC,.)

over-estimates its capacity level to hedge against the variability of the demand. Looking closely at the aforementioned example, the system allocates in total about 2620 units of capacity with the static setting versus 2457 for the multi-period case under the NIT process (resp, 2740 vs 2572 under NVT).



(a) NIT process (b) NVT process

Figure 5: Capacity-allocation decisions from static modeling approach versus the a priori capacity C_l for 5-4/8/20-25-(I1,LT,TC,.)