CIRRELT
1971 - 2021

CIRRELT-2021-28

Capacity Planning with Uncertainty on Contract Fulfillment

Teodor Gabriel Crainic
Guido Perboli
Walter Rei
Mariangela Rosano

July 2021

Bureau de Montréal
Université de Montréal
C.P. 6 128, succ. Centre-Ville
Montréal (Québec) H3C 3J7
Tél : 1-514-343-7575
Télécopie : 1-514-343-7121

Bureau de Québec
Université Laval,
2325, rue de la Terrasse
Pavillon Palasis-Prince, local 2415
Québec (Québec) G1V 0A6
Tél : 1-418-656-2073
Télécopie : 1-418-656-2624

Capacity Planning with Uncertainty on Contract Fulfillment[ɫ]

Teodor Gabriel Crainic[1,2,*], Guido Perboli[1,3], Walter Rei[1,2], Mariangela Rosano[3]

[1] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)

[2] Department of Management and Technology, Université du Québec à Montréal, Montréal, Canada

[3] ICT for City Logistics and Enterprises Center, Politecnico di Torino, Turin, Italy

Abstract. This paper focuses on the tactical planning problem faced by a shipper which seeks to secure transportation and warehousing capacity, such as containers, vehicles or space in a warehouse, of different sizes, costs, and characteristics, from a carrier or logistics provider, while facing different sources of uncertainty. The uncertainty can be related to the loads to be transported or stored, the cost and availability of ad-hoc capacity on the spot market in the future, and the availability of the contracted capacity in the future, when the shipper needs it. This last source of uncertainty on the capacity loss on the contracted capacity is particularly important in both long-haul transportation and urban distribution applications, but no optimization methodology has been proposed so far. We introduce the Stochastic Variable Cost and Size Bin Packing with Capacity Loss problem and model that directly address this issue, together with a metaheuristic to efficiently address it. We perform a set of extensive numerical experiments on instances related to long-haul transportation and urban distribution contexts, and derive managerial in sights on how such capacity planning should be performed.

Keywords: Capacity planning, stochastic programming, City Logistics, last-mile delivery, long-haul freight transportation, supply chain management

Acknowledgements. Partial funding for this work was provided by the Natural Sciences and Engineering Council of Canada (NSERC) through its Discovery Grants Program. While working on this paper, the first author was Adjunct Professor, Department of Computer Science and Operations Research, Université de Montréal, Québec, Canada, and the second author was the head of the *Urban Mobility and Logistics Systems* (UMLS) initiative of the interdepartmental *Center for Automotive Research and Sustainable mobility* (CARS) at Politecnico di Torino, Turin, Italy. While working to this paper, Prof. Guido Perboli was the representative in the Logistic Table in the Regional Council of Piedmont for Politecnico di Torino. The authors are grateful to Ms. Diane Larin for her invaluable linguistic and editorial assistance.

[ɫ]Revised version of the CIRRELT-2020-27

* Corresponding author: teodorgabriel.crainic@cirrelt.net

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2021

1 Introduction

Ensuring the reliability and flexibility of supply chains is a great challenge for managers, who are involved in various collaborations with several supply-chain partners and must perform complex planning processes on different decision levels, e.g., operational, tactical, and strategic. *Logistics capacity planning* constitutes an important component of those processes.

For the sake of simplicity of exposition but without loss of generality, we refer to the *shipper* as a retail firm, a producer or a supplier of goods, which requires capacity of various types in terms of size and cost, e.g., containers, ship or train slots, motor carrier trailers, or spaces in vans, rail cars or storage facilities, to store or transport its goods, e.g., raw materials, intermediate or final products packed in loads of various sizes, to respond to the demands of its own customers. We refer to the *carrier* as an external service provider (which could be a third-party logistics company) of transportation and warehousing services. Considering the regularity of the operations often conducted in supply chains and their cost-efficiency goals, the shipper often negotiates in advance a tactical plan-contract to secure the needed capacity to perform recurring activities (e.g., weekly or monthly) over a given planning horizon (e.g., one season or year). This tactical plan is beneficial for both sides, as the shipper benefits of a contract providing the estimated required capacity for the length of the planning horizon, and the carrier is guaranteed a regular volume of business.

The shipper faces significant uncertainty when negotiating, however. Indeed, not only the number and sizes of the loads the shipper will need to handle vary at each operation occurrence during the planning horizon, but the *availability of the contracted capacity at operation time* is also uncertain, as is the availability and characteristics, size and cost, of the *additional, ad-hoc capacity* the shipper would need to secure during operations to respond to the observed demand increase and loss of contracted capacity. The challenge is to account for these sources of uncertainty when selecting the units of capacity, of given types, to include in the contract, in order to minimize the total cost of the contracted capacity, plus the possible repetitive costs of handling the loss of capacity and securing the ad-hoc capacity every time the contract is to be used during operations.

This paper aims to introduce the decision-support method addressing these issues and challenges, to analyze the possible implications of contractual policies, and evaluate the effects of considering these sources of uncertainty explicitly in the capacity planning. In doing so, we address this topic from three points of view: from the transportation perspective, from a methodological point of view, and finally from a managerial one.

From a *transportation* perspective, the tactical capacity planning problem we address is relevant in many contexts characterizing the new generation of multi-stakeholder systems, e.g., synchromodal (Qu et al., 2019; Giusti et al., 2019; Perboli et al., 2017;

Giusti et al., 2018) and physical-internet-based (Ballot et al., 2014) inter-urban freight transport, data-based 3/4PL activities (Saglietto, 2013; Skender et al., 2017), and city logistics (Crainic and Montreuil, 2016; Crainic et al., 2021b). These recent paradigms in logistics and transportation require a continuously increasing amount of effort to coordinate stakeholders and provide more flexibility and better synchronization of operations (Ambra et al., 2019). Moreover, these contexts are affected by new business models and world-wide economic phenomena (e.g., growth of e-commerce, globalization of production and trade, and opening of broad free-trade economic zones). These trends result in contract logistics, which relies on service integrators and logistics service providers offering a wide range of modal and intermodal services as intermediaries between many and diverse shippers and carriers. These orchestrators coordinate stakeholders for increased efficiency and profitability for all, as illustrated by intermodal transport and logistics which combines the advantages of different transportation modes (Crainic et al., 2021a). Such coordination also brings, however, increased complexity in planning and management, the orchestrators capability of devising and implementing sophisticated plans being a critical success factor. Advance contracting of transportation, distribution, and warehousing capacity is an important piece of this capability and the methodology we introduce in this paper aims to support it.

We focus on the two complementary facets of freight transport in this paper, namely, urban distribution and long-haul transportation. The problem settings come from the rich literature on these topics synthesized in, e.g., Macharis and Bontekoning (2004); Ambra et al. (2019); Crainic (2003); Crainic and Kim (2007); Bektaş and Crainic (2008); Crainic and Speranza (2008); Crainic and Hewitt (2021); Bektaş et al. (2017); Crainic et al. (2021b), as well as from recent industrial and institutional collaborations of the authors, including work on 1) urban distribution in the metropolitan area of Turin, Italy, as part of the development of the new Logistics and Mobility Plan to be activated in 2022, through the collaboration of CARS@Polito (Automotive and mobility center of Politecnico di Torino), ICELab@Polito (ICT Center for City Logistics and Enterprises of Politecnico di Torino), and the Regional Government of Piedmont (Perboli et al., 2021; Brotcorne et al., 2019); 2) land-based long-haul freight transportation as part of Synchro-NET, the major European project for synchromodal long-haul corridor creation and operation (Synchro-NET Consortium, 2017; Perboli et al., 2017; Giusti et al., 2018), and 3) intermodal terminal optimization within long-haul freight transportation, as part of 5G-LOGINNOV, the European project for optimizing freight and traffic operations at ports and logistics hubs by 5G-enabled logistics corridors (5G-LOGINNOV Consortium, 2021; Porelli et al., 2021; Willenbrock et al., 2021). The world-wide severe shortage in maritime containers, particularly damaging on the trade routes from Asia, China, constitutes another relevant and timely problem setting and motivation. Indeed, the shortage of empty containers one observes currently, shortage which will not disappear any time soon, causes companies to wait for weeks to get capacity, skyrockets shipping costs (increases by more than 300% are observed), disrupts supply chains, and increases prices and delivery-delays for customers (e.g., La Presse, 2021; CNBC, 2021). This also lead to the creation of a new container booking service by Cainiao (China Smart Logistic Network, part of Alibaba Group, Reuters, 2021).

We introduce a *methodological perspective* to tackle the challenges of the transportation perspective, developing an operations research (OR)-based methodology to support decisions in addressing these capacity securing problems. As these problems involve the numbers and types of the capacity units one needs to contract for transportation or warehousing, the proposed model is based on the Bin Packing methodology, bins standing for the units of capacity, while items represent the freight loads one needs to handle.

The proposed methodology can be particularly useful to companies as support to decisions related to how much capacity to contract in advance and how much should be negotiated on a day-to-day basis. When surveying the literature, one observes that very few studies have addressed capacity planning problems under uncertainty in logistics applications. Furthermore, when this topic was addressed, the studies focused mainly on operational decisions, with very few exceptions dedicated to strategic and tactical planning applications (Crainic et al., 2014, 2016). Finally, to the best of our knowledge, no previous studies addressed jointly, within a single model and method, the issues discussed above, in particular, the different sources of uncertainty, which are relevant to contract building in both the long-haul transportation and the urban distribution contexts. In particular, the case where there is uncertainty on the availability of the contracted capacity, in addition to the ad-hoc and demand uncertainty, at the moment when operations are to be conducted is completely novel.

Last but not least, we consider the *managerial perspective, using the proposed methodology to bring managerial insights to the transportation perspective.* As already indicated, the logistics capacity planning problem represents a significant issue in supply chain management, especially when considering transportation and warehousing services, due to its impact on the performance of the firm in terms of service quality and costs (Crainic et al., 2016). Moreover, ignoring the uncertainty will generally result in decreasing the former while increasing the latter (Lium et al., 2009). Our experimental results provide the means to show that assessing and controlling the impact of uncertainty in such complex systems, by using appropriate OR-based methods and models, could support firms to achieve high-performance levels in both quality of service and economic efficiency and, thus, increase profits and gain competitive advantages in the long-run. Consequently, this paper aims to:

1. Present an integrated model that considers several uncertainty issues affecting capacity planning, extending the literature by considering the possibility that the contracted capacity turns out to be lower than planned at operations time. We model the problem as the *Stochastic Variable Cost and Size Bin Packing with Capacity Loss* problem, explicitly representing the uncertainty on the availability and volumes of the contracted capacity resources, the size and cost of extra capacity

one could secure during operations, and the number of sizes of the loads one will have to handle.

2. Overcome the computational limitations of standard solution methods, by proposing a particularly adapted progressive hedging-based metaheuristic.

3. Conduct an extensive set of computational experiments, using data that reflects the main issues involved in the problem for the urban distribution and the long-haul transportation contexts, to assess how various sources of uncertainty affect capacity planning (especially the random variability related to contracted capacity).

4. Perform a thorough analysis of the computational results and identify a series of managerial insights with respect to the structure of the contract choices given various urban distribution and the long-haul transportation characteristics and the expected information on the availability of the contracted capacity during future operations.

The remainder of the paper is organized as follows. We present the logistics capacity-planning problem we address in Section 2. We then present the two-stage stochastic formulation of the problem, and the metaheuristic solution approach to address it, in Sections 3 and 4, respectively. Section 5 is dedicated to the experimental plan and the analyses of the computational results with focus on the benefits of considering uncertainty in the capacity-planning process. The structure of the capacity plan under various problem settings and the derived managerial insights are the topic of Section 6. Finally, we provide the concluding remarks in Section 7.

2 Tactical Planning to Secure Capacity of Multiple Types under Uncertainty

This section introduces the logistics capacity planning problem addressed in this paper. Capacity planning is a challenging strategic/tactical decision, which is related to supply chain management. We consider, in particular, the tactical-planning problem of a decision-maker which needs to secure capacity, of different types, to meet its predicted demand over the next medium-term planning horizon. The decision-maker then negotiates medium-term contracts with service providers, to book in advance the capacity which will be used repeatedly to perform its activities for the duration of the planning horizon. The decision maker is different in different application contexts. We refer, e.g., to a shipper or forth/fifth-party logistics service provider securing capacity contracts with carriers for long-distance, regular shipments (Giusti et al., 2019), a wholesaler/retailer planning for transportation and storage capacity to support its procurement and sales processes

(Crainic et al., 2013), and the decision-platform of multi-stakeholder city logistics systems (Crainic et al., 2021b). Yet, the decision challenge and the general problem setting is the same in all cases. Consequently, in order to simplify the presentation but without loss of generality, we describe the problem within the context of the process of contract procurement between a shipper and a carrier. Given the time lag that usually exists between the signing of the tactical-level contracts and the actual logistics operations, the negotiations are performed under uncertainty, as discussed in the next section.

We first present the problem setting within two different contexts: *urban distribution* and *long-haul transportation*. We provide a compact description of the general problem in the third subsection. We finally enrich the presentation with a brief review of the literature on capacity planning directly related to the contexts at hand.

2.1 Urban distribution

Urban distribution refers to the overall process by which freight is transported both to and from dense urban environments. Such environments face increasing challenges of congestion and negative environmental impacts associated to transportation, freight transportation in particular. One also observes the continuous growth of e-commerce together with always higher customer desires to have their purchased goods delivered both fast and cheap. To answer these challenges and needs, many firms (e.g., the e-commerce giant platforms Alibaba, 2018; Amazon, 2018) are moving from a push cost-driven supply model to a time and cost pull-driven approach, that is, to demand-driven logistics. Simultaneously, private and public (e.g., transit authorities) carriers and service providers make coalitions for capacity sharing and integrated decision-making to consolidate freight and reduce the impact of freight transportation and logistics on the city. Multi-tier smart urban transportation, or *City Logistics*, systems are implementing these approaches (Crainic et al., 2009, 2021b).

The goal of such systems is to reduce the negative impacts (i.e., costs, congestion, noise, etc.) associated with the vehicles transporting freight in urban areas by more efficiently using their capacity (i.e., increasing the average vehicle fill rate and reducing the number of empty trips that are performed). City logistics is based on the application of two general principles: 1) the consolidation of loads originating from different shippers within the same vehicles and 2) the coordination of the distribution operations within the city. In this case, the use of multiple transportation tiers enables the system to utilize specifically adapted infrastructure and specialized fleets at each tier to better attain the overall goal that is pursued. While the first tier is generally the same in all contexts, most systems for medium-to-large urban areas involve two tiers, while three or more are part of the large-to-metropolis size urban areas.

The first tier includes a set of terminals, generally known as City Distribution Centers

(CDCs), which are usually located on the outskirts of the city, whose main function is to serve as the entry (exit) points and consolidation facilities for the inbound (outbound) freight. In the following, in an effort to simplify the exposition, we discuss the inbound case only; similar arguments can be evoked when considering the outbound freight. Long-haul transportation vehicles of various modes deliver their cargo at the CDCs, where the delivered loads are sorted and then consolidated into smaller urban vehicles. The connection between the first and the lower tiers takes place at transshipment facilities with no or low warehousing capabilities, called satellites and associated to the second tier of the system. The urban vehicles thus bring freight to satellites, where it is transshipped to city freighters, vehicles specifically adapted to perform distribution operations in dense urban zones. The city freighters deliver freight to their final destination within the city either directly (two-tier systems) or through a series of continuously smaller facilities (e.g., mini hub and lockers) and lower-capacity vehicles (e.g., drones and bicycles). Specific access and moving rules constrain activities to limit their negative impacts (e.g., urban trucks will move along specific paths that are chosen to efficiently reach satellites while minimizing congestion) and contribute toward the goals of economic, social, and environmental efficiency. Multi-tier systems are thus able to distribute freight in urban areas in a more efficient overall way, but the planning of such systems poses important challenges to managers at all decision levels (strategic, tactic, and operational).

As previously mentioned, the principle of consolidation is central to how multi-tier city logistics systems plan and operate. In all transportation tiers, loads are consolidated into vehicles, urban vehicles and city freighters, respectively, which are then used to move the freight within the city. These vehicles can be private or public (first-tier light rail, for example) but, often, they need to be contracted in advance, including the capacity of the transit vehicles whose future availability is uncertain due to variations in people transportation requirements. This justifies the need to plan in advance the required distribution capacity, while simultaneously taking into account the uncertainty on the shipments to be moved and their volumes, the possible capacity loss at operation time of the contracted resources, and the characteristics of the ad-hoc replacement solutions which could be available (Brotcorne et al., 2019).

Tactical capacity planning aims to ensure that such consolidation can be efficiently performed. Specifically, managers must secure the required numbers of vehicles of various types, which will be available at each tier to correctly perform the transportation operations. It should be noted that, the number of different vehicle types available for each tier is increasing, as are their characteristics and costs, e.g., various types of electric and, soon, hydrogen vans, electric bykes, drones, and lockers, without forgetting the autonomous versions of many of these vehicle types and the capacity offered by individuals under crowdsourcing operating principles (Crainic et al., 2021b; Perboli et al., 2018). These types and characteristics must be considered when preforming capacity planning. The incidents, e.g., accidents and mechanical failures, which occur regularly but randomly, result in booked vehicles not being available at the appropriate moment and thus, disrupted system operations and loads not delivered on time. Accounting for this uncertainty adds to the complexity of the capacity planning process but contributes to the flexibility and robustness of operations by contracting adequate levels of resources, while accounting for the ad-hoc capacity secured at operations time to hedge against unexpected variations. The optimization model that is proposed in the present paper provides this planning capability by explicitly integrating the possibility of random capacity loss of the contracted resources.

2.2 Long-haul transportation

Long-haul transportation is another context in which securing capacity for future operations is essential and capacity losses can randomly occur when this capacity is called upon during operations..

Globalization and the opening of broad free-trade economic zones have changed logistic chains dramatically. A higher volume of long-haul transportation operations are now required to be planned and performed by organizations everywhere. On the one hand, such operations have been reorganized around the use of bigger warehouses, and the movements of goods are now performed over longer distances involving different modes of transportation and larger vehicles (Perboli et al., 2017; Giusti et al., 2018). On the other hand, the “liberalization” of economies has increased the competition between firms and, in the process, the attention to controlling costs (especially transportation costs). In this context, Rodrigue and Notteboom (2013) discuss the concept of intermediacy in regional distribution and global logistics when organizing regular shipping between an origin and a destination at various market scales. Their study focuses, in particular, on containerized freight distribution in two major markets, North America and Europe. The authors point out that companies must take into account, when planning activities, the possibilities and limitations linked to the capacities of the nodes (e.g., seaports, intermodal terminals) and links (e.g., corridors) involved, capacities which have a great impact on the transportation network.

Let us consider the case of a shipper (e.g., manufacturing firm, wholesaler, or retailer) acquiring resources, or products, from a set of suppliers located in distant regions, according to their specific global procurement process. In such a case, the shipper must secure in advance the required number of containers (for maritime or land-based modes) for the long-haul transport required to deliver the resources (or products) to its warehousing and distribution facilities. This advanced booking process is particularly important when the industry faces a shortage of resources, which is increasingly the case as illustrated by the container shortage evoked in the Introduction and the shortage of truck drivers in North America.

Crainic et al. (2013) illustrate such a case, presenting the specific situation of a North

American hardware and home-improvement wholesale-retail chain, which regularly imports a large variety of products from a set of suppliers located in South-East Asia. Consolidation is used in conjunction with intermodal shipping in this case. The products are first consolidated in containers, then moved by a liner containership from a port of origin in South-East Asia to a port of destination in North America, and then delivered to the firm's main distribution center by a combination of rail and motor-carrier services. To secure the regularity and quality of deliveries for the products it plans to buy over the next season, the firm must negotiate with a carrier or logistics service provider the required tactical capacity, i.e., to book the estimated required quantity and characteristics of containers, as well as of slots on maritime and rail transportation services. Several random changes were observed regarding the planned capacity. On the one hand, variations in the items and quantities purchased required securing additional capacity at often high prices. On the other hand, when the other customers of the same service provider had large volumes of freight to move, either the contracted containers were not all available, or only part of the capacity of some containers was available as the service provider consolidated freight from several customers into the same boxes. Consequently, stochastic capacity loss should again have been considered in the planning process.

2.3 Problem description

The tactical capacity planning problem addressed in this paper concerns a shipper which needs to secure capacity of different types from a carrier, to meet its uncertain demand. The capacity types could be transportation modes (e.g., ship or train slots, containers, space in cargo bikes or vans), specific carriers, or storage space within given facilities, each type having particular characteristics in terms of *unit cost* and *size*. The shipper negotiates this multi-type capacity in advance, and it will use it to perform its shipping or storage activities repeatedly, e.g., every day, week, or month, over a certain planning horizon, e.g., one semester, season, or year. The output of this negotiation is a medium-term contract, which includes the quantity, i.e., the number of units, of capacity of each type (this quantity is zero for non-relevant types given the demand) and the expected costs to use the contracted capacity, as well as to react to variations in supply and demand which could occur during operations. Indeed, given the time lag that usually exists between the signing of the contract and the logistics operations, as well as the hazards and risks associated to predicting future supply and demand levels, several sources of uncertainty are affecting the contract negotiation.

The first source of uncertainty is the *demand*, that is, the number of units, and the size of each unit, the shipper will need to transport or store at each occurrence of its activities over the planning horizon. Indeed, even in the most 'regular" context of operations, the demand fluctuates in time and what one observes at any given occurrence of activity is generally different from a single-value (also called point forecast) prediction of the number of units to transport or store and the size of each of those units. This may result in insufficient booked capacity available on the shipping day, compromising the fulfillment of the contract and generating additional costs to handle the situation. In this paper, we thus explicitly address these demand uncertainty issues and the strategies to secure additional, ad-hoc, capacity when needed. We also assume, without explicitly modeling, that the shipper deploys re-selling strategies of the surplus capacity when the observed overall demand is lower than estimated.

A second major source of uncertainty is the availability of the contracted capacity each time the shipper performs its activities and the contract is applied. In fact, due to unfavorable situations, e.g., mechanical failures, accidents, and delays, the contracted capacity may be entirely or partially unavailable at shipping time. This *capacity loss* fluctuates in time and, as its precise value cannot be predicted with certainty for any given moment of the planning horizon, it has to be assumed stochastic. Such loss of contracted capacity involves additional costs and decisions. On the one hand, goods which were supposed to be in the lost capacity need to be re-assigned to other units of capacity. We assume the associated cost is proportional to the total lost capacity. On the other hand, one needs to secure ad-hoc capacity through the spot market in order to proceed to the adjustment of the capacity-utilization plan by re-assigning shipments to contracted and ad-hoc capacity units. It is noteworthy that the number, size, and cost of the various types of capacity units which will be available in the future are also uncertain.

Capacity planning has been investigated, and identified as a major challenge in a number of supply-chain management settings, e.g., production and distribution. Thus, for example, Yuan and Ashayeri (2009) state that, insufficient capacity gradually leads to deteriorating delivery performance, consequently lowering revenue and market share. Yoon et al. (2016) highlight that access to freight transportation capacity has become a complex issue faced by logistics managers due to capacity shortages. Finally, according to Monczka et al. (2010), the planning of logistics capacity affects the distribution and operating costs of a company.

With a focus on the urban context, Bosona (2020) identifies in his review the available transport capacity as one of the major challenges of urban freight last-mile logistics, in particular related to the complexity of on-demand delivery platforms. Thus, e.g., Yildiz and Savelsbergh (2019) introduce service and planning of crowd-sourced transportation capacity in meal delivery in last-mile logistics planning.

Capacity expansion and its allocation in the supply chain has received considerable attention within the capacity planning literature (e.g., Luss, 1982; Singh et al., 2012; Birge, 2012; Liu and Papageorgiou, 2013). Singh et al. (2012) and Liu and Papageorgiou (2013) propose mixed integer programming models for the capacity expansion planning of global supply chains in the process industry. Birge (2012) considers capacity planning models to decide whether to install additional capacity at the production plant level.

The author takes into account the limited resources and demand uncertainty. Finally, Yuan and Ashayeri (2009) present an approach to combine system dynamics loops and control theory simulations to analyze the impacts of various factors on capacity expansion strategies.

Most of the research studies which have been conducted on this subject deal only partially with the requirements of capacity planning. Only a few have thus focused on stochastic capacity planning and the different sources of uncertainty involved. Indeed, several papers on this topic consider demand variability as the only source of uncertainty. For example, Pimentel et al. (2013) propose a mathematical model and solution approach to the Stochastic Capacity Planning and Dynamic Network Design problem under demand uncertainty. Ahmed et al. (2003) present a multi-stage capacity expansion problem with uncertain demand and cost parameters, while Aghezzaf (2005) discusses the capacity planning and warehouse location problem in supply chains operating under uncertainty on demand.

The papers by Crainic et al. (2016, 2014) propose first attempts to address capacity planning problem settings found in strategic and tactical applications. In particular, the authors present two versions of the Stochastic Variable Cost and Size Bin Packing Problem (SVCSBPP) in the long-haul transportation context. In these problems, the uncertainty related to the demand (i.e., loads to be transported) and the capacity availability on the spot market was explicitly considered. However, to the best of our knowledge, the uncertainty affecting the availability of booked capacity has not yet been considered in the literature. Moreover, there are no studies addressing all the above-presented issues in a single model, which can be applied and validated in both the long haul transportation and urban distribution applications.

We aim to fill this gap by 1) formalizing the tactical capacity planning problem under uncertainty on the loss of contracted capacity, available ad-hoc capacity, as well as the volume and characteristics of demand, which we identify as the *Stochastic Variable Cost and Size Bin Packing with Capacity Loss* (*SVCSBP-LS*) problem, and 2) proposing a new optimization model, which takes the form of a two-stage stochastic programming formulation (Birge and Louveaux, 1997). We formulate the model using the Bin Packing vocabulary and concepts, where capacity units are the *bins*, of various types, one has to select in order to load the *items*, of various sizes, representing the freight loads to transport or store. This model generalizes prior work on the Stochastic Variable Cost and Size Bin Packing problems (Crainic et al., 2016), which assumes that all the booked capacity is available at the shipping or storage date.

3 The Tactical Planning Model Formulation

This section is dedicated to the two-stage stochastic programming formulation we propose for the tactical capacity planning under uncertainty SVCSBP-LS. As indicated above, because the problem setting is found in many application fields and, thus, the proposed methodology is relevant in all those fields, we adopt the general vocabulary of Bin Packing problems. Thus, *items* represent the freight loads to be transported or stored, and *bins* stand for the capacity units of various transportation modes, e.g., containerships, rail wagons or container platforms (Crainic and Kim, 2007; Bektaş and Crainic, 2008; Kienzle et al., 2021), trucks, smart and modular containers (Ballot et al., 2014), space in cargo bikes, vans, or light-rail vehicles (Crainic et al., 2021b) in urban-distribution, and storage space in warehousing and distribution facilities.

The first stage concerns the tactical capacity planning decisions, i.e., the *a priori* selection of the bins, of various types, sizes, and fixed costs, to be secured to move or store the items for the duration of the planning horizon. The second stage refers to the operational decisions, i.e., the recourse actions one needs to take to adjust the plan once the actual demand, the list of items with their sizes, and the actual available size of the contracted capacity are observed. The recourse actions concern paying the cost involved in handling the items which should have gone into the lost capacity, securing the missing capacity through ad-hoc bins of various sizes and costs (at spot-market value, i.e., higher than the fares negotiated initially), and assigning the items to the available bins, either originally contracted, at possibly a smaller capacity, or currently acquired. These actions are carried out repeatedly over the planning horizon to cope with the fluctuation of supply and demand, here defined as *random events*, which affect the result of the first stage (i.e., booked capacity not sufficient or not available). The objective is to minimize the total expected cost for the planning horizon, computed as the sum of the tactical bin selection (first stage) and the expected cost of adjusting this plan to the observed information for all the time moments the plan-contract is applied.

Let T be the set of bin types known to be available at the first stage, defined by the size V^t and fixed cost f^t of the bins $t \in T$. Let also c^t be the cost to pay for the loss of a unit of capacity of a bin of type $t \in T$ selected in the first-stage. This cost is the additional expense required to react to the reduction of the available volume of first-stage bins, by rearranging the loads and assigning them to bins. Let $\mathcal{J}^t$ be the set of available bins of type t, with $\mathcal{J} = \bigcup_t \mathcal{J}^t$, the set of available bins at the first stage. Finally, let y_j^t be the first-stage capacity selection decision variable, equal to 1 if bin $j \in \mathcal{J}^t$ is selected, and 0, otherwise.

Let $\mathcal{T}$ be the set of bin types available at the second stage, with V^τ, the nominal volume of a bin of type $\tau \in \mathcal{T}$. Notice that $T \subseteq \mathcal{T}$, meaning that some (e.g., the types of the selected bins) or all types available at planning (first) stage are also available in the future, albeit with some capacity loss as defined in the following. Let Ω be the sample space of the random event, where $\omega \in \Omega$ defines a particular realization. The vector ξ contains the stochastic parameters defined in the model, and $\xi(\omega)$ represents a given realization of this random vector. We consider the following stochastic parameters in $\xi(\omega)$:

Items: $\mathcal{I}(\omega)$, Set of items, with $v_i(\omega)$, the volume of item $i \in \mathcal{I}(\omega)$;

Bins: $\mathcal{K}^\tau(\omega)$, Set of available bins of type $\tau \in \mathcal{T}$ at the second stage, with $\mathcal{K}(\omega) = \bigcup_\tau \mathcal{K}^\tau(\omega)$;

Bin sizes: $\mathcal{V}_j^t(\omega)$, Volume of second-stage bin $j \in \mathcal{J}^t$ of type $\tau \in \mathcal{T}$, with $\mathcal{V}_j^t(\omega) \leq V^\tau$ for the bins selected at the first stage ($j \in T \subseteq \mathcal{T}$);

Bin costs: $g^\tau(\omega)$, Unit cost of of second-stage (spot market) bin $j \in \mathcal{J}^t$ of type $\tau \in \mathcal{T}$.

The second-stage decision variables are

Bin selection: $z_k^\tau(\omega) = 1$, if bin $k \in \mathcal{K}^\tau(\omega)$ is selected in the second stage, 0 otherwise;

Item-to-bin assignment:
$x_{ij}(\omega) = 1$, if item $i \in \mathcal{I}(\omega)$ is packed in first-stage bin $j \in \mathcal{J}$, 0 otherwise;
$x_{ik}(\omega) = 1$ if item $i \in \mathcal{I}(\omega)$ is packed in second-stage bin $k \in \mathcal{K}(\omega)$, 0 otherwise.

The two-stage SVCSBP-LS model may then be formulated as:

$$\min_y \sum_{t \in T} \sum_{j \in \mathcal{J}^t} f^t y_j^t + E_\xi \left[Q\left(y, \xi(\omega)\right) \right] \tag{1}$$

$$\text{s.t.} \quad y_j^t \geq y_{j+1}^t, \qquad \forall t \in T, j = 1, \ldots, |\mathcal{J}^t| - 1, \tag{2}$$

$$y_j^t \in \{0,1\}, \qquad \forall t \in T, j \in \mathcal{J}^t. \tag{3}$$

where

$$Q(y, \xi(\omega)) = \min_{z(\omega), x(\omega)} \sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}^\tau(\omega)} g^\tau(\omega) z_k^\tau(\omega) + \sum_{t \in T} \sum_{j \in \mathcal{J}^t} c^t (V^t - \mathcal{V}_j^t(\omega)) y_j^t \tag{4}$$

$$\text{s.t.} \quad \sum_{j \in \mathcal{J}} x_{ij}(\omega) + \sum_{k \in \mathcal{K}(\omega)} x_{ik}(\omega) = 1, \qquad \forall i \in \mathcal{I}(\omega), \tag{5}$$

$$\sum_{i \in \mathcal{I}(\omega)} v_i(\omega) x_{ij}(\omega) \leq \mathcal{V}_j^t(\omega) y_j^t, \qquad \forall\, t \in T, j \in \mathcal{J}^t, \tag{6}$$

$$\sum_{i \in \mathcal{I}(\omega)} v_i(\omega) x_{ik}(\omega) \leq V^\tau z_k^\tau(\omega), \qquad \forall\, \tau \in \mathcal{T}, k \in \mathcal{K}^\tau(\omega), \tag{7}$$

$$x_{ij}(\omega) \in \{0,1\}, \qquad \forall\, i \in \mathcal{I}(\omega), j \in \mathcal{J}, \tag{8}$$

$$x_{ik}(\omega) \in \{0,1\}, \qquad \forall\, i \in \mathcal{I}(\omega), k \in \mathcal{K}(\omega), \tag{9}$$

$$z_k^\tau(\omega) \in \{0,1\}, \qquad \forall\, \tau \in \mathcal{T}, k \in \mathcal{K}^\tau(\omega). \tag{10}$$

The objective function (1) minimizes the sum of the total fixed cost of selecting capacity within the tactical capacity plan and the expected cost of addressing the negative impacts, over the planning horizon, of the non availability at operation time of the capacity contracted at negotiation time. This expected cost is computed over all possible future realizations of the loss of contracted capacity and the availability, size, and cost of ad-hoc capacity.

Packing problems usually present a strong symmetry in the solution space, and two solutions are considered symmetric (and equivalent) if they involve the same set of first-stage bins in different orders. However, when we consider the available capacity of first-stage bins as a source of uncertainty, this is no longer true. Indeed, each bin of type $t \in T$ may have a different volume, and we need to characterize it properly. We thus introduce constraint (2) to break the symmetry and ensure order in the selection of bins of type $t \in T$, i.e., bin $j \in \mathcal{J}^t$ can be selected at the first stage only if bin $j-1 \in \mathcal{J}^t$ has already been selected. Finally, constraint (3) imposes the integrality requirements on y.

In the second stage, the term $Q\left(y, \xi(\omega)\right)$ (4) details the expected cost, over the possible realizations of the random event, of the second stage of securing ad-hoc capacity and adjusting the plan, given the tactical capacity plan y and a realization $\xi(\omega)$ of the loss of capacity, the availability of ad-hoc capacity, and the list of items with their characteristics. Constraint (5) ensures that each item is packed in a single bin. Constraints (6) and (7) ensure that the total volume of items packed in each bin does not exceed its actual volume, for first and second-stage bins, respectively. Finally, constraints (8) to (10) impose the integrality requirements on all second-stage variables.

4 Progressive hedging-based metaheuristic

The SVCSBP-LS is a difficult stochastic combinatorial optimization problem to solve. It generalizes the SVCSBPP (Correia et al., 2008; Crainic et al., 2016). To overcome the computational limitations of standard solution methods, we propose a Progressive Hedging (PH)-based metaheuristic (Rockafellar and Wets, 1991), that is tailored for the SVCSBP-LS problem and its inherent complexity.

The proposed metaheuristic is applied by first defining a discretization of the sample space associated with the random event. This leads to the creation of a set of representative scenarios $\mathcal{S}$, each one providing the values of the considered stochastic parameters associated with a possible occurrence of the random event. The metaheuristic then proceeds by applying a scenario decomposition procedure, which produces $|\mathcal{S}|$ subproblems (one for each scenario included in $\mathcal{S}$). The algorithm then solves the problem by iteratively executing the following steps: (i) the single scenario subproblems are first heuristically solved to obtain local (or scenario-specific) solutions; (ii) a reference point, indicating the level of solution consensus among the subproblems, is obtained by calculating the weighted average over the local solutions found; (iii) the values of the fixed costs of the bin types in the objective function are then adjusted for all scenario subproblems to promote consensus among them with respect to the reference point (thus penalizing the dissimilarity observed among the local solutions).

It should be noted that the PH-based metaheuristic proposed in the present paper is based on the one originally developed by Crainic et al. (2016) for the simpler SVCSBPP problem variant. However, the uncertainty on the volume of every single bin makes the SVCSBP-LS a more complex problem to solve. Specifically, the uncertainty on the bin volumes may generate a huge number of bin types in the scenario subproblems (i.e., each bin may have a different volume, leading to single-bin bin types) that the metaheuristic must solve at each iteration performed. As in Crainic et al. (2016), each deterministic single scenario subproblem is solved using the heuristic developed by Crainic et al. (2011). This heuristic relies heavily on the concept of bin types, which are defined as distinct couples of values, i.e., the fixed cost and the volume of the bins. Therefore, to obtain an efficient PH method for the SVCSBP-LS, innovations were required to efficiently deal with the significant increase in the number of bin types.

A detailed description of the overall solution method is provided in A. In this section, we focus on the description of the different steps that compose the PH metaheuristic, summarized in Algorithm 1, while emphasizing the main contributions and enhancements that were applied to the original method to efficiently address the complexity of the problem at hand.

As previously indicated, the first step of the metaheuristic *builds a discretization of the stochastic problem* (Algorithm 1, lines 1 and 2). This entails reformulating the SVCSBP-LS two-stage model by discretizing the value space of the random variables through a set of representative scenarios $\mathcal{S}$, with p_s defining the probability of scenario $s \in \mathcal{S}$. The notation of the previous section is thus updated to account for the scenario definition. Therefore, $y_j^{ts} = 1$ if bin $j \in \mathcal{J}^t$ of type $t \in T$ is selected in the first stage under scenario $s \in \mathcal{S}$, and 0 otherwise. For $t \in T$, V^t and f^t refer to the volume and fixed cost associated with a bin of type t, respectively. Let c^t be the unit capacity-loss cost.

For the second stage, we then have the set of additional bins defined as $\mathcal{K}^s = \bigcup_\tau \mathcal{K}^{\tau s}$, where $\mathcal{K}^{\tau s}$ is the set of extra bins of type $\tau \in \mathcal{T}$ in scenario $s \in \mathcal{S}$, and $\mathcal{I}^s$ defines the set of items to pack under scenario $s \in \mathcal{S}$. Similarly, $g^{\tau s}$ is the cost associated with bins of type $\tau \in \mathcal{T}$ in scenario $s \in \mathcal{S}$, V^τ is the volume of bins of type $\tau \in \mathcal{T}$, $\mathcal{V}_j^{ts}$ defines the volume of first-stage bin $j \in \mathcal{J}^t$ under scenario $s \in \mathcal{S}$, and v_i^s is the volume of item $i \in \mathcal{I}^s$ in scenario $s \in \mathcal{S}$. Finally, variable $z_k^{\tau s}$ is equal to 1 if and only if extra bin $k \in \mathcal{K}^{\tau s}$ of type $\tau \in \mathcal{T}$ is selected in scenario $s \in \mathcal{S}$, while the binary variables x_{ij}^s and x_{ik}^s are the item-to-bin assignment variables for scenario $s \in \mathcal{S}$.

The SVCSBP-LS formulation (1)-(10) can now be approximated by the following deterministic model:

$$\min_{y,z,x} \sum_{s\in\mathcal{S}} p_s \left[\sum_{t\in T} \sum_{j\in\mathcal{J}^t} f^t y_j^{ts} + \sum_{\tau\in\mathcal{T}} \sum_{k\in\mathcal{K}^{\tau s}} g^{\tau s} z_k^{\tau s} + \sum_{t\in T} \sum_{j\in\mathcal{J}^t} c^t (V^t - \mathcal{V}_j^{ts}) y_j^{ts} \right] \tag{11}$$

$$\text{s.t.} \quad y_j^{ts} \geq y_{j+1}^{ts}, \qquad \forall\, t \in T, j = 1,\ldots,|\mathcal{J}^t|-1, s\in\mathcal{S}, \tag{12}$$

$$\sum_{j\in\mathcal{J}} x_{ij}^s + \sum_{k\in\mathcal{K}^s} x_{ik}^s = 1, \qquad \forall\, i \in \mathcal{I}^s, s\in\mathcal{S}, \tag{13}$$

$$\sum_{i\in\mathcal{I}^s} v_i^s x_{ij}^s \leq \mathcal{V}_j^{ts} y_j^{ts}, \qquad \forall\, t\in T, j\in\mathcal{J}^t, s\in\mathcal{S}, \tag{14}$$

$$\sum_{i\in\mathcal{I}^s} v_i^s x_{ik}^s \leq V^\tau z_k^{\tau s}, \qquad \forall\, \tau\in\mathcal{T}, k\in\mathcal{K}^{\tau s}, s\in\mathcal{S}, \tag{15}$$

$$y_j^{ts} = y_j^{ts'}, \qquad \forall\, t\in T, j\in\mathcal{J}^t, s, s'\in\mathcal{S}, \tag{16}$$

$$y_j^{ts} \in \{0,1\}, \qquad \forall\, t\in T, j\in\mathcal{J}^t, s\in\mathcal{S}, \tag{17}$$

$$z_k^{\tau s} \in \{0,1\}, \qquad \forall\, \tau\in\mathcal{T}, k\in\mathcal{K}^{\tau s}, s\in\mathcal{S}, \tag{18}$$

$$x_{ij}^s \in \{0,1\}, \qquad \forall\, i\in\mathcal{I}^s, j\in\mathcal{J}, s\in\mathcal{S}, \tag{19}$$

$$x_{ik}^s \in \{0,1\}, \qquad \forall\, i\in\mathcal{I}^s, k\in\mathcal{K}^s, s\in\mathcal{S}. \tag{20}$$

The objective function (11), and the constraints (12)–(15) and (17)–(20) have the same meaning as their counterparts in Section 3. One should note that it is the inclusion in (11) of the term that accounts for the capacity losses for the bins selected a priori and their related costs in the second stage that may cause a significant increase in the number of bin types in each scenario subproblem, once the problem is decomposed by scenario. Additionally, constraints (16) are the non-anticipativity requirements, which ensure that the first-stage decisions are not tailored to each scenario in $\mathcal{S}$. These constraints are necessary to guarantee that the model yields a single implementable capacity plan. At the same time, the presence of these constraints prevents the resulting model from being scenario separable.

In the second step, we then *apply the augmented Lagrangian-based scenario decomposition scheme*, originally proposed by Rockafellar and Wets (1991), to the resulting multi-scenario deterministic problem (Algorithm 1, lines 3 and 4). This is done by relaxing the non-anticipativity constraint (16) using an augmented Lagrangian strategy with the Lagrangian multipliers being defined as $\lambda_j^{ts}, \forall j \in \mathcal{J}^t, \forall t \in T$, and $\forall s \in \mathcal{S}$, and ρ_j^t being a penalty ratio associated with bin $j \in \mathcal{J}^t$ of type $t \in T$. Again, the detailed steps of the decomposition scheme are presented in A.

For the resulting subproblems, i.e., $\forall s \in \mathcal{S}$, let $\mathcal{B}^{\overline{\tau}s} = \mathcal{J}^{\overline{\tau}} \cup \mathcal{K}^{\overline{\tau}s}$ be the set of available bins of type $\overline{\tau} \in \overline{\mathcal{T}}$ (where $\overline{\mathcal{T}} = \mathcal{T} \cup T$) and $\mathcal{B}^s = \bigcup_{\overline{\tau}} \mathcal{B}^{\overline{\tau}s}$ be the whole set of bins available in the subproblem. For bin $b \in \mathcal{B}^{\overline{\tau}s}$, let $\mathcal{V}_b^{\overline{\tau}s}$ be the actual volume of the bin (for $b \in \mathcal{K}^{\overline{\tau}s}$, $\mathcal{V}_b^{\overline{\tau}s} = V^{\overline{\tau}}$) and let $f_b^{\overline{\tau}s}$ define its associated fixed cost. The related decision variables then become, $y_b^{\overline{\tau}s} = 1$ if bin $b \in \mathcal{B}^{\overline{\tau}s}$ of type $\overline{\tau} \in \overline{\mathcal{T}}$ is selected, 0 otherwise. Moreover, x_{ib}^s is equal to 1 if item $i \in \mathcal{I}^s$ is packed in bin $b \in \mathcal{B}^s$, 0 otherwise. The model (11)-(20) is thus decomposed into a series of deterministic VCSBPP subproblems (one for each scenario $s \in \mathcal{S}$) with modified fixed costs $f_b^{\overline{\tau}s}$ and additional constraints (see A, constraints (48)) that enforce an order in the selection of bins of each type $\overline{\tau} \in \overline{\mathcal{T}}$. When compared to the complete formulation, the resulting subproblems are much less complex to solve.

The algorithm then *builds a solution to the stochastic model* by performing the two phases as summarized in Algorithm 1, from line 5 to line 30. For a given iteration ν, we define $\lambda_b^{\overline{\tau}s\nu}$ and $\rho_b^{\overline{\tau}\nu}$ as the Lagrangian multiplier and the penalty ratio associated with bin $b \in \mathcal{B}^{\overline{\tau}s}$ for scenario $s \in \mathcal{S}$, respectively. Let $y_b^{\overline{\tau}s\nu}$, $\forall b \in \mathcal{B}^{\overline{\tau}s}, \overline{\tau} \in \overline{\mathcal{T}}$, define the local solution associated with subproblem $s \in \mathcal{S}$ at iteration ν. Furthermore, $\delta^{\overline{\tau}s\nu}$ is the total number of bins of type $\overline{\tau} \in \overline{\mathcal{T}}$ which can be derived from the capacity plan (i.e., local solution) for scenario subproblem $s \in \mathcal{S}$ at iteration ν. Using the subproblem solutions, the overall capacity plan (i.e., the reference point) is calculated thus producing the values $\bar{y}_b^{\overline{\tau}\nu}$. Equivalently, we define $\bar{\delta}^{\overline{\tau}\nu}$ to be the expected value, that is obtained from the subproblem solutions, for the total number of bins at the current iteration ν. Let $f^{\overline{\tau}s\nu}$ be the fixed cost of bin $b \in \mathcal{B}^{\overline{\tau}s}$ of type $\overline{\tau} \in \overline{\mathcal{T}}$ for scenario $s \in \mathcal{S}$ at iteration ν. The terms α and $\sigma_{\%}$ are two given constants such that $\alpha > 1$ and $0.5 \leq \sigma_{\%} \leq 1$. Finally, $\bar{\delta}_m^{\overline{\tau}\nu}$ and $\bar{\delta}_M^{\overline{\tau}\nu}$ are the lower and upper bounds, that represent the minimum and maximum number of bins of type $\overline{\tau}$ observed over all the solutions to the scenario subproblems at iteration ν.

At each iteration, the scenario subproblems are solved separately to obtain the local solutions (Algorithm 1, line 9). Each deterministic subproblem is solved using the best first increasing loading heuristic, originally proposed in Crainic et al. (2011). As mentioned previously, considering the uncertainty on bin volumes, one can observe a significant increase in the number of bin types in the scenario subproblems, with several bin types containing a single bin. Let us recall that the best first increasing loading heuristic relies on ordering the bins based on a merit function, which was defined as the ratio between the fixed cost and the volume of a bin (assuming that a single pair of values is defined for each bin). In the present problem setting, considering that the bins available in the first stage may have a different observed volume in the second stage, then the heuristic proposed in Crainic et al. (2011) needed to be modified. Therefore, we first introduced a lookup table enabling the first stage bin types defined in the scenario subproblems to be quickly identified (i.e., $\overline{\tau} \in \overline{\mathcal{T}} \rightarrow t \in T$). Second, we changed the sorting criterion that is used in the heuristic. Specifically, we apply a lexicographic sorting based on two criteria. The first criterion sorts the first stage bins according to a non-decreasing ratio of bin cost and bin volume, as expressed by the bin type to which the bin belongs (i.e., without the stochastic volume reduction). The second criterion then sorts the bins grouped by the same first criterion value by non-increasing order of the observed bin volume (i.e., explicitly considering the volume reductions). Based on this new ordering, the best first increasing loading heuristic is then applied as in Crainic et al. (2011).

Step 3 aims to reach the consensus for the first-stage variable values associated with the solutions obtained for the scenario subproblems. The consensus being defined here as the scenario solutions being similar in terms of the first-stage bin-selection decisions. A reference point is thus created through the aggregation of the subproblem solutions by applying the expected value operator (Algorithm 1, lines 10-12). This yields a temporary overall capacity plan, which is then used to identify the bins for which consensus may be achieved.

To induce consensus among the scenario solutions, the fixed costs of the bins are adjusted in the objective functions of the scenario subproblems. Two strategies are applied to update the fixed costs. The first is based on adjusting the Lagrangian multipliers to penalize the lack of consensus due to the differences in the values of first-stage variables (see Crainic et al., 2016, for details). In particular, the fixed costs of the bin types in each scenario subproblem are tuned according to the differences observed between the values of the bin-selection variables at the current iteration and the overall capacity plan (Algorithm 1, line 14). Thus, the fixed cost of a bin type is either increased, or reduced, depending on whether or not in the current scenario solution the bin type is overused, or underused, when compared to its usage in the overall capacity plan. These adjustments (Algorithm 1, lines 14 and 15) can be less effective when the differences observed between the subproblem solutions and the overall one are small, and thus when the overall solution is close to consensus. This may result in an unwarranted number of additional iterations performed to complete the search for a consensus solution.

To address this issue, we apply a second penalty-adjustment strategy, based on heuristic principles (Algorithm 1, lines 16-18). Therefore, when at least σ percent of the variables have reached consensus, we adjust in all the scenario subproblems the fixed cost $f^{\overline{\tau} s \nu}$ (see for details A (57)). In this way, we penalize the selection costs of bins of type $\overline{\tau}$ in scenario s at iteration ν when, at the previous iteration, the total number of bins of that type was larger than the number of bins of the same type in the corresponding reference solution. We thus discourage the adoption of those bins. If the opposite case is observed, then the cost adjustments will promote the use of the bins.

The search for consensus also involves the soft variable fixing scheme defined in Crainic et al. (2016) (Algorithm 1, line 19). As originally proposed, this scheme fixed part of the selection of the bins in all the scenario subproblems based on lower and upper bound values for the number of used bins of each type that were observed over all the scenario solutions. The best first increasing loading heuristic was then applied with these fixed selection decisions being enforced. As previously mentioned, given the uncertainty on the volume of the bins available in the first stage, the loading heuristic was modified to account for the sharp increase in the number of bin types in the scenario subproblems (i.e., the use of the lookup table and the two criteria lexicographic sorting approach). Thus, the soft variable fixing scheme is also updated to manage the assignment between the original bin type of every bin and the bin type in use in the heuristic solution of every single scenario subproblem. Specifically, the lookup table is again leveraged to efficiently identify the first stage bin types and their associated use in the scenario solutions obtained at each iteration ν of the PH-based metaheuristic.

Finally, it is important to note that Phase I can conclude without reaching a consensus solution. Consequently, Phase II is performed to produce an implementable solution to the SVCSBP-LS. The end of Phase I occurs either when consensus is achieved for all bin types except one, type $\overline{\tau}'$ for which $\bar{\delta}_m^{\overline{\tau}'} < \bar{\delta}_M^{\overline{\tau}'}$, or, when consensus is not achieved within a given maximum number of iterations (200 in our experiments). In the first case (Algorithm 1, line 25), given the efficiency of the item-to-bin heuristic, Phase II computes the final solution by iteratively examining the possible number of bins for $\overline{\tau}'$ within the interval $\left[\bar{\delta}_m^{\overline{\tau}'}, \bar{\delta}_M^{\overline{\tau}'}\right]$ (see Algorithm 1, line 26, and A). Otherwise, the final solution is obtained by solving exactly (using a commercial solver) a restricted SVCSBP-LS defined by fixing the first-stage variables for which consensus was reached (i.e., the same bins that are used in all the scenario solutions at the end of Phase I) (Algorithm 1, lines 28 and 29).

5 Experimental plan

We performed an extensive set of experiments with a threefold aim: 1) Analyze the new logistics capacity planning problem in the contexts of urban distribution and long-haul transportation, in particular, the relevance and impact of the capacity loss phenomenon we introduce and the corresponding uncertainty; 2) Measure the impact of uncertainty and the interest of building a stochastic programming model; 3) Study the relationship between the problem characteristics and parameters and the structure of the capacity plan, drawing managerial insights.

We begin by presenting the instance sets used to qualify our model and the solution procedure (Subsection 5.1). Subsection 5.2 then discusses the potential of considering uncertainty in the planning process, while Subsection 5.3 studies the issue from the point of view we introduce in this paper, the explicit consideration of the loss of capacity on contracted bins. Managerial insights are the object of Section 6.

Algorithm 1 PH-based metaheuristic for the SVCSBP-LS

Step 1: Build discretization of the stochastic problem
Generate a set of scenarios $\mathcal{S}$;
Step 2: Apply the Lagrangian-based scenario decomposition
Decompose the resulting deterministic model (11)–(20) by scenario using augmented Lagrangian relaxation
Step 3: Compute the solution to the stochastic model
Phase 1
$\nu \leftarrow 0$; $\lambda_b^{\overline{\tau}s\nu} \leftarrow 0$; $\rho_b^{\overline{\tau}\nu} \leftarrow f^{\overline{\tau}}/10$;
while Termination criteria not met **do**
  For all $s \in \mathcal{S}$, solve the corresponding VCSBPP subproblem $\rightarrow y_b^{\overline{\tau}s\nu}$;
  Compute temporary global solution
    $\bar{y}_b^{\overline{\tau}\nu} \leftarrow \sum_{s\in\mathcal{S}} p_s y_b^{\overline{\tau}s\nu}$
    $\bar{\delta}^{\overline{\tau}\nu} \leftarrow \sum_{s\in\mathcal{S}} p_s \delta^{\overline{\tau}s\nu}$
  Penalty adjustment
    $\lambda_b^{\overline{\tau}s\nu} = \lambda_b^{\overline{\tau}s(\nu-1)} + \rho_b^{\overline{\tau}(\nu-1)}(y_b^{\overline{\tau}s\nu} - \bar{y}_b^{\overline{\tau}\nu})$
    $\rho_b^{\overline{\tau}\nu} \leftarrow \alpha\rho_b^{\overline{\tau}(\nu-1)}$
  if consensus is at least $\sigma_{\%}$ **then**
    Adjust the fixed costs $f^{\overline{\tau}s\nu}$;
  end if
  Variable fixing
    $\bar{\delta}_m^{\overline{\tau}\nu} \leftarrow \min_{s\in\mathcal{S}} \delta^{\overline{\tau}s\nu}$ and $\bar{\delta}_M^{\overline{\tau}\nu} \leftarrow \max_{s\in\mathcal{S}} \delta^{\overline{\tau}s\nu}$
    Apply variable fixing;
  $\nu \leftarrow \nu + 1$
end while
Phase 2
if consensus not met for a single bin type $\overline{\tau}'$ ($\bar{\delta}_m^{\overline{\tau}'} < \bar{\delta}_M^{\overline{\tau}'}$) **then**
  Identify the consensus number of bins δ of type $\overline{\tau}'$ by enumerating $\delta \in \left[\bar{\delta}_m^{\overline{\tau}'}, \bar{\delta}_M^{\overline{\tau}'}\right]$ (and variable fixing)
else
  Fix consensus variables in model (11)–(20);
  Solve restricted (11)–(20) model using a commercial solver.
end if

5.1 Instance set

In this subsection, we provide a set of instances for the SVCSBP-LS and we present the instance generation process. Since, to the best of our knowledge, there is no prior study of the capacity planning problem with uncertainty on the actual volume of the contracted capacity, we generated new test instances for the SVCSBP-LS, based on previous work on bin packing problems (Monaci, 2002; Crainic et al., 2007, 2012, 2011, 2016; Gobbato, 2015).

Table 1 summarizes the parameters of the instances. Most parameters are self explanatory; a few require a bit of explanation.

The *bin availability* is assumed to be different at the time of the contract, the first stage, and when repeatedly executing the contract in the future, the second stage. We define the number of bins of each type $t \in T$ available at the first stage as the minimum number of bins of volume V^t needed to pack all items in the worst-case scenario. Three availability classes, *AV1 - AV3*, are defined for the second stage, representing different levels of variability. The first presents the largest variability, and its worst-case scenario may involve a limited number of extra bins. On the contrary, all the scenarios have the same availability of extra bins in the third class, equal to the first-stage availability. The second class stands for a middle-of-the-road situation.

The fixed costs of bins are assumed higher at the second stage from those at the time of contracting (built based on Correia et al., 2008), by a multiplying factor. Three values were used representing continuously increasing variations in the fixed costs.

Three parameters are used to represent the possible capacity loss on the contracted bins, from the global problem level to the individual bin-type level: 1) the percentage of *scenarios affected by capacity loss* (*SL*); 2) the probability that a *bin type* is affected by capacity loss (TL); 3) the percentage of the *overall capacity loss for all the bins of a certain type* selected in the first-stage (*BL*). Each parameter values represent an increasing level of potential capacity loss. The distributions used to generate these values are different for the two application cases. A *uniform* (U) capacity loss is assumed for long-haul transportation, reflecting the rather wide-spread inability to predict correctly the quality of the service that will be provided by carriers. The situation is different from urban distribution, and even more when city logistics systems are involved, as the relations with the service providers are generally smoother. We identify this type of capacity loss *localized* (L), i.e., only a few randomly-chosen first-stage bins lose their entire capacity and become unusable, while the others are unaffected. Localized capacity losses may be caused by mechanical failure of vehicles or other incidents, e.g., undelivered parcels during the previous operational day that were kept in the vehicle reducing the capacity for new demand to be loaded.

Characteristic	**Value - Parameters for all the problem settings**
Number of items	Uniformly distributed over $[100, 500]$
Item volume	*Small* (S): uniformly distributed over $[5, 10]$
	Medium (M): uniformly distributed over $[15, 25]$
	Big (B): uniformly distributed over $[20, 40]$
Bin types, with	Set *T3*: 3 bin types with volumes = 50, 100, 150
$\mathcal{T}$ is equal set T	Set *T5*: 5 bin types with volumes = 50, 80, 100, 120, 150
Bin availability 1st stage	$\|\mathcal{J}^t\|$ equal to $\lceil \frac{1}{V^t} \max_{s \in \mathcal{S}} \sum_{i \in \mathcal{I}^s} v_i^s \rceil$
Bin availability 2nd stage	Class 1 (*AV1*): $\|\mathcal{K}^{ts}\|$ uniformly distributed over $[0, \|\mathcal{J}^t\|]$
	Class 2 (*AV2*): $\|\mathcal{K}^{ts}\|$ uniformly distributed over $[\|\mathcal{J}^t\|/2, \|\mathcal{J}^t\|]$
	Class 3 (*AV3*): $\|\mathcal{K}^{ts}\|$ equal to $\|\mathcal{J}^t\|$
Bin costs 1st stage	$f^t = V^t(1 + \gamma^t,\ \gamma^t)$ uniformly distributed over $[-0.3, 0.3]$
Bin costs 2nd stage	$g^\tau = f^t(1 + \alpha)$, $\alpha \in \{0.3, 0.5, 0.7\}$
Capacity loss	*SL*: % of scenarios $= 20\%, 40\%, 60\%, 80\%$
	TL - Probability of a bin type $= 0.5, 0.75, 1$
	BL - % of overall loss for all 1st stage bins of a certain type$= 20\%, 30\%, 40\%, 50\%, 60\%, 70\%$
Unit capacity-loss cost	$c^t = \alpha^t f^t / V^t$, same α^t a for g^τ

Characteristic	**Value - Parameters specific to each problem setting**	
	Long-haul transportation	*Urban distribution*
Capacity loss type	Uniform (U)	Localized (L)

Table 1: Parameters of SVCSBP-LS instances

Finally, the unit additional due to capacity loss is set equal to the proportion of the overall loss of capacity among all first-stage bins of type t (BL).

Ten (10) random instances were generated for each combination of parameters, yielding a total of 51 840 instances. All the instances incorporate 100 scenarios. The size of the scenario trees to use in the experiments was tuned by analyzing the in-sample and out-of-sample stability conditions. Let us recall that in-sample stability refers to the requirement that the quality of the results obtained when solving a stochastic model using a fixed size for the scenario set remains stable for different samples of scenarios. As for out-of-sample stability, it refers to the requirement that when solving a stochastic model using a given size for the scenario set, one is guaranteed to closely approximate the true value of the stochastic model. According to Kaut et al. (2007), the stability requirements ensure the reliability and robustness of the solutions obtained when a different set of scenarios is considered.

Therefore, we first created a subset of instances for T3 and T5, based on different combinations of the parameters presented above to perform the stability testing. Then, we generated ten scenario trees for each instance, while also varying the cardinality of the scenario sets $|S| = \{10, 25, 50, 100, 150, 200\}$. The metaheuristic was then applied to solve all instances obtained. To assess the in-sample stability condition, we evaluated the solutions based on the scenario samples used to obtain them. Stability was reached when the standard deviation associated with the solution values obtained for the instances generated with the same size of the scenario set was judged to be low enough. As for the out-of-sample stability condition, it was evaluated on a different sample than the one that was used to find the solutions, see (Kaut et al., 2007). Thus, the following procedure was applied ten times for each instance generated: i) we solved a 200-scenario problem; ii) we solved the instance with the scenario trees of cardinality $|S| = \{10, 25, 50, 100, 150\}$; iii) we then evaluated each of the solutions obtained in step ii) within the 200-scenario context (this was done by fixing the first-stage decision variables and then solving the resulting second stage, the recourse, for the 200 scenarios); iv) we computed the relative gap of the objective-function value of this solution relative to that of the 200-scenario problem.

We observed that both the in-sample and out-of-sample stability conditions were reached with an accurate precision when generating trees with 100 scenarios. Thus, we used this tree dimension in the rest of the experiments. We do not report the detailed results, these were low in terms of both the observed computational times and the variability, while also being independent from the instance parameters. In the worst case, we achieved an average computation time of less than 5 seconds (which is quite low) and a variability of under 1% of standard deviation over all the instances.

5.2 Assessment of the SVCSBP-LS model

As stated in Section 2, much of the literature does not consider uncertainty in capacity planning problems. Then, the question we address is whether modeling uncertainty explicitly, through the two-stage SVCSBP-LS formulation with recourse, is beneficial compared to solving the deterministic variant of the problem only. Would the shipper gain by considering uncertainty, by the reduction in its overall expenses for the transportation and storage capacity plan? This would be important for the shipping industry where the marginal revenues are already low.

We use two classical and highly relevant measures in the literature (Birge, 1982). The *Expected Value of Perfect Information* (*EVPI*), representing the decision maker's willingness to pay for complete information about the future, and the *Value of the Stochastic Solution* (*VSS*) computing the difference between the solutions obtained by solving the deterministic problem with the expected value of the parameters (the expected value solution - *EEV*) and the stochastic SVCSBP-LS problem (*RP*).

In other words, the EVPI provides the value of having perfect information (i.e., the ability to perfectly predict what specific scenario would be observed), thus removing all uncertainty regarding the parameters that influence the capacity planning. As for the VSS, it measures the expected gain obtained by solving the stochastic model rather than its deterministic counterpart, i.e., where all random variables are replaced by their mean values (Maggioni and Wallace, 2012). In the present setting, one can interpret the VSS as the opportunity loss for the company if it uses a deterministic optimization model to perform the capacity planning. It thus shows the added value of estimating the future via the use of a scenario set that approximates how the values of the stochastic parameters may randomly vary and then applying the proposed metaheuristic to solve the resulting stochastic model and produce the capacity plan. Tables 2 and 3 display the average and maximum results for the two measures, respectively, computed as a percentage with respect to the RP for the two instance sets (Column 1), bin-availability class (Column 2), and value of the increase in the future bin cost and capacity loss α (Column 3). Results are displayed for each application type (Columns 4 and 5 for urban distribution, Columns 6 and 7 for long-haul transportation).

For the sake of brevity, we discuss the results of these stochastic programming measures at a macro level, analyzing how they vary in long-haul transportation and urban distribution, depending on the availability of second-stage bins and the extra cost due to loss of capacity. The interested reader may refer to B for more detailed results and analysis.

The results show high values for using a stochastic formulation in all cases, i.e., high values for the additional insight about the future. This value increases with the cost of future capacity and the decrease in the availability of future capacity. The higher un-

Set	Availability	α	Urban distribution		Long-haul transportation	
			$EVPI[\%]$	$EVPI[\%]_{max}$	$EVPI[\%]$	$EVPI[\%]_{max}$
T3	**AV1**	**0.3**	13.98	60.76	22.20	77.35
		0.5	18.65	48.07	25.47	75.24
		0.7	21.97	36.69	26.80	74.27
	AV2	**0.3**	9.05	13.85	10.19	20.23
		0.5	15.26	19.12	16.14	29.96
		0.7	19.34	23.71	19.82	35.28
	AV3	**0.3**	9.47	14.52	10.11	20.30
		0.5	15.79	20.38	16.18	29.43
		0.7	19.90	24.61	19.91	35.95
T5	**AV1**	**0.3**	12.13	15.71	13.28	54.26
		0.5	17.73	21.24	19.16	50.83
		0.7	21.44	25.11	22.74	47.86
	AV2	**0.3**	8.09	13.62	9.61	19.27
		0.5	15.23	21.60	16.45	31.17
		0.7	19.59	25.32	20.40	36.60
	AV3	**0.3**	8.97	13.66	9.48	20.72
		0.5	15.84	19.88	16.57	30.17
		0.7	20.20	25.57	21.07	37.25

Table 2: EVPI for SVCSBP-LS with different availability classes, values of α, and types of capacity loss

certainty of long-haul transportation is reflected in the higher information values. These results are confirmed by significant VSS values, double-digit gains in expected costs being obtained in most cases by using the stochastic SVCSBP-LS model. In both cases, the look-ahead capability offered by the stochastic formulation would mitigate the impacts of higher operation costs and missed or late deliveries due to loss of contracted capacity, and high costs for limited availability of ad-hoc capacity.

It should be further noted that trends can be more easily observed by considering the values of VSS[%] compared to VSS[%]$_{max}$. The reason being that VSS[%] is a global statistic that reports the average values over all the results obtained for the instances grouped within each category, while VSS[%]$_{max}$ reports the maximum observed value for a given instance category. Therefore, when analyzing the VSS[%], one observes that when the cost of the future capacity increases, the observed differences between the solutions obtained by solving the stochastic and the deterministic models tend to decrease. Specifically, for all instance categories, one observes the highest value of the VSS[%] when $\alpha = 0.3$ and the lowest value when $\alpha = 0.7$. While, in all cases, the VSS[%] values are always significant, this general trend is nonetheless observed.

Set	**Availability**	α	**Urban distribution**		**Long-haul transportation**	
			$VSS[\%]$	$VSS[\%]_{max}$	$VSS[\%]$	$VSS[\%]_{max}$
T3	**AV1**	**0.3**	11.29	23.53	13.79	33.85
		0.5	8.37	20.04	10.58	31.47
		0.7	5.63	15.49	8.47	56.65
	AV2	**0.3**	15.75	44.49	17.57	55.13
		0.5	12.20	30.92	13.95	55.03
		0.7	9.41	38.24	12.59	80.40
	AV3	**0.3**	15.67	43.82	17.02	62.07
		0.5	10.34	35.99	13.52	50.98
		0.7	8.08	29.52	11.90	80.83
T5	**AV1**	**0.3**	12.00	29.73	14.50	38.40
		0.5	7.79	22.61	12.02	49.84
		0.7	4.93	16.35	9.88	74.71
	AV2	**0.3**	14.12	45.00	16.17	58.54
		0.5	9.95	31.21	12.93	63.77
		0.7	6.70	22.28	11.40	88.36
	AV3	**0.3**	14.54	33.51	17.96	57.93
		0.5	9.07	34.95	14.67	48.97
		0.7	5.34	27.67	11.48	63.08

Table 3: VSS for SVCSBP-LS with different availability classes, values of α, and types of capacity loss

We now examine to what extent the first-stage decisions of the SVCSBP-LS and EV formulation differ. As highlighted in Crainic et al. (2016), the EV problem generally overestimates the future demand, that is, a total item volume larger than the actual volume, and a larger set of available bins in the future. Moreover, when the percentage of scenarios affected by capacity loss and the probability of bin types being affected by capacity loss are low, the EV formulation underestimates the reduction of available capacity, meaning that the total volume of first-stage capacity predicted to be available at operation times is larger than the actual available volume. This behavior can lead to two undesired situations. First, the EV solution may include a set of bins which are not suitable for the set of scenarios considered. The capacity plan is then more expensive than necessary even when the solution is feasible and implementable. Second, the EV solution may include insufficient capacity for certain situations (subset of scenarios) in which the actual availability of bins is limited, yielding an unfeasible capacity plan for those situations.

The importance of the problem and parameter setting was further emphasized as we observed about 10% infeasible instances when the variability in future bin availability and cost is high (AV1), while most instances were feasible in the other settings. Table 4 details this observation, showing that when uniform losses are expected (availability class A1), the number of infeasible instances grows considerably with the variability in availability and cost. The issue is particularly sensitive when only a limited number of bin types is available on the market (up to 30% for sets T5 but 98.75% pour T3). These observations highlight the need for considering uncertainty in capacity planning when the availability of bins may be limited in the future.

5.3 Capacity loss and uncertainty

As stated, the uncertainty on the availability of contracted capacity at operations time is not addressed in the literature. Thus, this subsection is dedicated to studying how considering the possible loss in the planned/contracted capacity as a stochastic parameter is valuable. We thus compare the results obtained by solving the SVCSBP-LS (Appendix B contains the complete result tables and analysis) to those of Crainic et al. (2016) where the possible capacity loss and its variability were not considered. It should be noted that, in both studies, the uncertainty related to the demand as well as to the availability and the costs of extra bins in the future are explicitly considered. Therefore, in the present paper, we model the capacity loss for contracted bins for the urban distribution and the long-haul transportation cases, while all other sources of uncertainty are the same as in Crainic et al. (2016).

The results obtained in both studies emphasize the usefulness of stochastic formulations to perform capacity planning. Furthermore, as previously observed, taking into account the uncertainty related to the capacity of the contracted bins significantly increases both the average and the maximum values of the EVPI and the VSS for all instances considered. When comparing these results to the ones obtained in Crainic

α	SL[%]	TL[%]	Set T3 - BL[%]			Set T5 - BL[%]		
			20-30	**40-50**	**60-70**	**20-30**	**40-50**	**60-70**
0.3	**20**	**50**	12.50	10.00	25.00	0.00	0.00	0.00
		75	10.00	20.00	47.50	0.00	0.00	0.00
		100	8.75	43.75	82.50	0.00	0.00	0.00
	40	**50**	12.50	12.50	32.50	0.00	0.00	0.00
		75	10.00	15.00	40.00	0.00	2.50	10.00
		100	6.25	25.00	77.50	0.00	5.00	22.50
	60-80	**50**	12.50	15.00	30.00	0.00	3.75	12.50
		75	10.00	17.50	12.50	1.25	16.25	28.75
		100	8.75	15.00	53.75	6.25	27.5	30.00
0.5	**20**	**50**	12.50	20.00	32.50	0.00	0.00	0.00
		75	10.00	22.50	70.00	0.00	0.00	0.00
		100	15.00	75.00	98.75	0.00	0.00	0.00
	40	**50**	15.00	20.00	32.50	0.00	0.00	0.00
		75	10.00	12.50	50.00	0.00	0.00	2.50
		100	8.75	52.50	98.75	0.00	0.00	25.00
	60-80	**50**	12.50	17.50	25.00	0.00	0.00	15.00
		75	10.00	12.50	35.00	0.00	13.75	26.25
		100	8.75	30.00	85.00	0.00	25.00	30.00
0.7	**20**	**50**	10.00	20.00	40.00	0.00	0.00	0.00
		75	5.00	35.00	100.00	0.00	0.00	0.00
		100	35.00	92.50	98.75	0.00	0.00	0.00
	40	**50**	10.00	20.00	30.00	0.00	0.00	0.00
		75	5.00	15.00	100.00	0.00	0.00	0.00
		100	10.00	77.50	100.00	0.00	0.00	5.00
	60-80	**50**	10.00	20.00	30.00	0.00	0.00	3.75
		75	5.00	15.00	55.00	0.00	2.50	23.75
		100	10.00	65.00	97.50	0.00	12.50	30.00

Table 4: % of infeasible instances for availability class AV1 in the long-haul transportation setting

et al. (2016), considering the localized capacity losses characterizing the instances of the urban distribution case, one observes that the VSS values are about 3 times higher than the ones reported in the prior study. The increase in the VSS values is even higher for the instances related to the long-haul transportation case (which are characterized by uniform capacity losses), i.e., the VSS values being 4 to 5 times higher in this case. We can therefore conclude that excluding this source of uncertainty from the stochastic model may lead to underestimate the capacity available at operations time and the additional costs one will have to support, and this, in both urban distribution and long-haul transportation contexts.

6 Managerial insights

Having established that incorporating the concept of capacity loss and uncertainty in capacity planning can provide the shipper with competitive advantage through better operations management and reduced costs, we now discuss the structure of the capacity planning solutions. We study, in particular, how solutions vary depending on the attributes of urban distribution and long-haul transportation problem settings, with emphasis on the expected available volumes of contracted bins at operations time.

We base our analysis on comparing the results of SVCSBP-LS and those of Crainic et al. (2016), where the loss of capacity was not considered, on the following performance indicators

- Average number of bin types contracted in the capacity plan N_t;
- Average percentage of the total capacity needed which is booked at the first stage Cap_{FS};
- Average percentage of the objective function value achieved at the first stage Obj_{FS};

computed for all combinations of instance sets, availability classes, and the other characteristics of the sets.

Table 5 summarizes the variation intervals means for the first three measures for each capacity-planning solution according to the number of bin types (Column1) and the availability of extra bins on the spot market (Column 2). When the parameters that determine the actual volumes of first-stage bins are equal, the resulting structures of the capacity-planning solutions are the same for availability classes AV2 and AV3 and we thus present the results of instances with availability classes AV2 and AV3 together. For further details and complete tables concerning the figures and results reported in this section, the interested reader may refer to Lerma (2018).

No capacity loss							
Set	*Availability*	$Cap_{FS}range$	$Cap_{FS}mean$	$Obj_{FS}range$	$Obj_{FS}mean$	$N_t range$	$N_t mean$
T3	AV1	71.82%-83.96%	78.50%	63.38%-72.87%	68.56%	1.10-1.20	1.13
	AV2+AV3	60.81%-81.58%	72.91%	52.64%-70.86%	62.76%	1.00-1.10	1.03
T5	AV1	67.12%-83.61%	76.15%	59.21%-73.17%	66.84%	1.33-1.44	1.37
	AV2+AV3	65.62%-83.14%	74.53%	56.53%-72.56%	64.58%	1.00-1.20	1.03
Uncertain capacity loss - long-haul transportation							
Set	*Availability*	$Cap_{FS}range$	$Cap_{FS}mean$	$Obj_{FS}range$	$Obj_{FS}mean$	$N_t range$	$N_t mean$
T3	AV1	61.19%-82.85%	64.81%	48.45%-73.41%	60.39%	1.20-3.00	1.98
	AV2+AV3	0%-78.62%	42.89%	0%-68.40%	34.99%	0-1.70	0.93
T5	AV1	6.17%-81.12%	49.70%	4.87%-70.99%	40.33 %	0.30-3.00	1.60
	AV2+AV3	0%-81.25%	44.35%	0%-71.71%	36.42%	0-1.90	1.00
Uncertain capacity loss - urban distribution							
Set	*Availability*	$Cap_{FS}range$	$Cap_{FS}mean$	$Obj_{FS}range$	$Obj_{FS}mean$	$N_t range$	$N_t mean$
T3	AV1	66.38%-84.02%	74.39%	59.17%-75.32%	65.92 %	1.00-3.00	1.87
T5	AV1	55.25%-81.62%	72.01%	48.15%-72.14%	63.41%	1.30-3.90	2.18

Table 5: Comparative performance of capacity-planning solutions

When the capacity loss on contracted bins is not accounted for, the shipper books the capacity sufficient to limit the adjustments and costs when the actual demand becomes known. As observed previously (Crainic et al., 2016; Lerma, 2018), this plan tends in this case to mostly include bins of the same type, with only one or two bins of different types. This relates to the cost-orientation of the shipper who uses standardized bins tailored by the carrier to the shipper' needs to avoid the higher loading/unloading and handling costs generated by non-standardized loading schemes. Indeed, results in Table 5 show that, when the availability of second-stage bins is limited, the average number of bin types, N_t, increases slightly, reaching the maximum values of 1.20 for set T3 and 1.44 for set T5. Most capacity is booked (Cap_{FS} around 79%) and paid for (Obj_{FS} around 69%) at contracting time. It is worth noting, however, the large variance of all values.

We now turn to examining to what extent and how the structure of the capacity plan changes when the shipping company takes into account the uncertain nature of capacity loss of contracted bins. The percentage of the total capacity needed which is booked at the first stage, Cap_{FS}, characterizes the capacity plan and its variation is a good indication of the structural changes brought by varying the problem definition. Table 6 displays the average Cap_{FS} values for long-haul transportation and urban distribution contexts for each set of bin types (Column 1), bin availability class in the second stage (Column 2), and capacity-lost cost (Column 3).

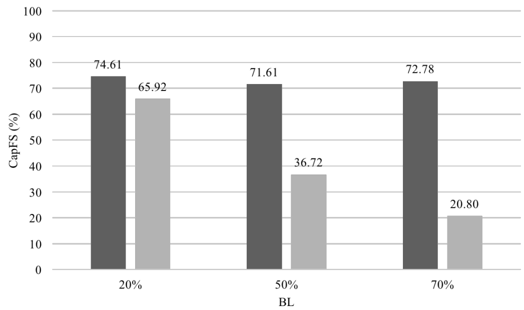
The results show the sensitivity of the capacity plan to the application context, the availability of extra bins on the spot market, the way capacity is lost and modeled, and the cost of the capacity loss. They thus illustrate the impact of these factors on the managerial decisions concerning how much capacity to contract. The sensitivity and impact are particularly strong in the long-haul transportation context where the capacity the shipper should contract in the first stage changes dramatically with the changes in problem parameters. In particular, when freight demand rises, the supply falls, and the cost of the spot market rates rise, the shipper may suffer from the higher second-stage costs and the methodology proposes to book in advance most of the required capacity. The costs of extra bins and capacity loss at operation time, modeled through the parameter α, impacts strongly the creation of safety buffers in the long-haul transportation context. Thus, the percentage of capacity contracted initially, Cap_{FS}, doubles when $alpha$ increases from 0.3 to 0.7. The situation is different in the urban distribution context, where the shipper should contract more of less the same high-value capacity in all cases. Notice that the percentage of capacity contracted initially is the same for all settings when the possibility of capacity loss is higher, irrespective of the number of bin types.

			Long-haul transportation	**Urban distribution**
Set	*Availability*	α	Cap_{FS}	Cap_{FS}
T3	AV1	0.3	35.44%	70.09%
	AV1	0.5	53.36%	78.44%
	AV1	0.7	69.03%	86.78%
	AV2+AV3	0.3	26.49%	61.24%
	AV2+AV3	0.5	42.01%	69.30%
	AV2+AV3	0.7	53.50%	75.01%
T5	AV1	0.3	30.32%	61.96%
	AV1	0.5	45.71%	70.43%
	AV1	0.7	57.01%	75.64%
	AV2+AV3	0.3	26.99%	62.98%
	AV2+AV3	0.5	42.91%	69.99%
	AV2+AV3	0.7	53.66%	76.16%

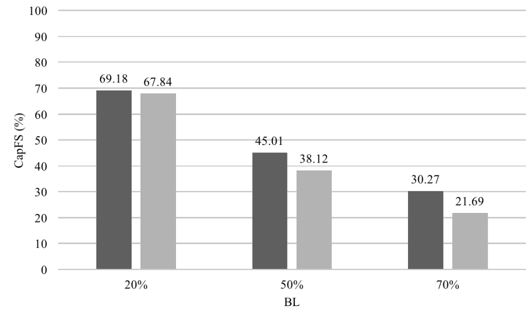
Table 6: Variation of Cap_{FS}, % of contracted capacity during the 1st stage, with problem parameters

Figure 1 depicts the average values of the percentage of the capacity which is booked at the first stage, Cap_{FS}, and the average number of bin types contracted in the capacity plan, N_t, for the long-haul transportation context (where the capacity loss is uniformly distributed) for the sets T3 and T5, the availability classes AV1, dark gray, and AV2, light gray, and three levels of BL, the % of overall capacity loss for the contracted bins (low = 20%, medium = 50%, and high = 70%). The figure illustrates further the need in this case to book most of the capacity needed in the first stage, irrespective of the possibility of capacity loss at operation time. Moreover, the capacity plan includes several bin types, nearly in all the instances we addressed, the number increasing with the level of possible capacity loss. In practice, such a capacity plan would require, however, that attention be paid to the loading/unloading requirements of the different bin types; the complexity of such requirements should be reflected in the bin type cost.

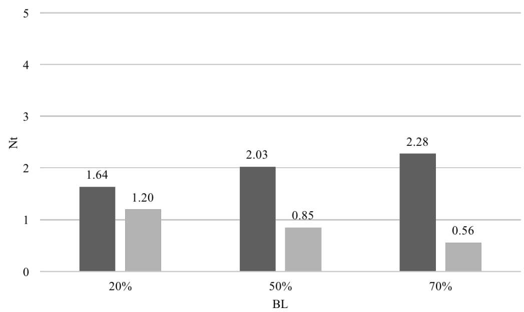
Some cases are of particular interest. First, when there are only three types of bins and the availability of the second-stage bins is limited (AV1), the structure of the capacity-



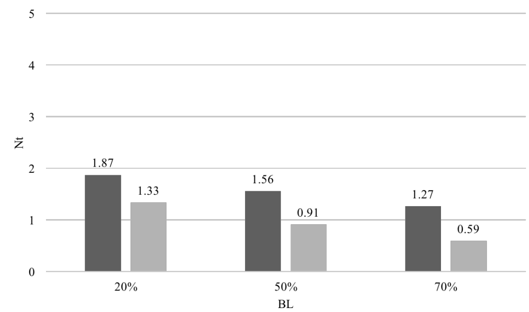
(a) *Cap_{FS} for instance set T3.*



(b) *Cap_{FS} for instance set T5.*



(c) *N_t for instance set T3.*



(d) *N_t for instance set T5.*

Figure 1: Average values of Cap_{FS} and N_t for capacity-loss levels, long-haul transportation setting, availability classes AV1 (dark gray) and AV2 (light gray)

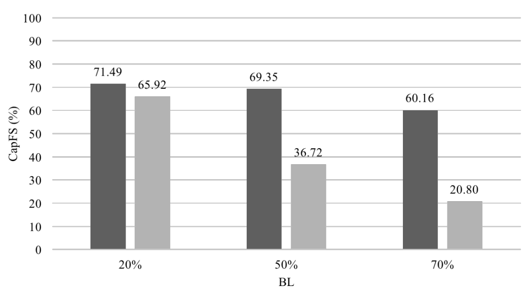
planning solution is always the same, regardless of the likelihood and amplitude of capacity loss and the plan books in advance almost all the capacity needed for the planning horizon.

The second case worthy of interest is when the probability of loosing a large amount of capacity is high. Given the risk of a limited availability of extra bins in the future and the obligation to satisfy the demand, the plan leads in this case to increase the percentage of capacity booked in advance, even though the cost of bins and capacity loss is higher. As illustrated in Figure 1 parts a and b, this increase is much more significant when the number of bin types is low.

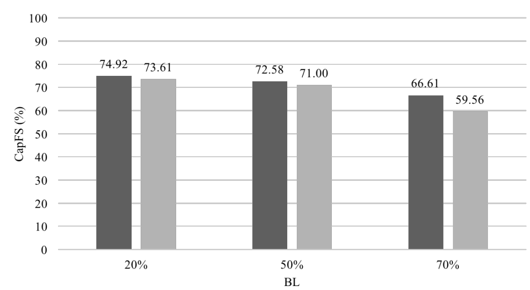
The third case concerns the availability of bins in the future as represented through the classes AV1 - AV3. When the predicted level of availability is high, as in class AV2, the capacity plan that is based mainly on the premium cost of extra bins and capacity loss (parameter α), and varies considerably depending on the value of the predicted capacity loss for the contracted bins (parameter BL). The percentage of capacity contracted (first stage) increases with the premium cost and decreases as the BL increases. The latter behaviour corresponds to the realization that there is little value in booking in advance capacity that one will loose for the most part when it will be necessary to use it.

Finally, in the long-haul transportation context, the average number of bin types selected when the contract is established (N_t) increases with α and is particularly sensitive when the number of bin types is relatively low and the predicted future availability is highly uncertain (class AV1). When the latter is not a concern, most of the bins included in the capacity plan are of the same type (the value of N_t is always below 1.9), irrespective of the variations in the other problem parameters.

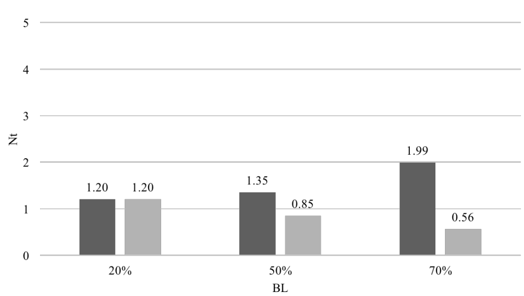
We now turn to the urban distribution context, where the capacity loss is "localized", that is, it is assumed more predictable and less wide-spread than the long-haul context, with only a few contracted bins losing their entire capacity,while the others remain unaffected. The results are nearly the same for all availability classes in this context, and thus we display the results for the availability class AV1 only in Table 5. To complete those figures, Figure 2 depicts the average values of the percentage of the capacity booked at the first stage, Cap_{FS}, and the average number of bin types contracted in the capacity plan, N_t, for the sets T3 and T5, the availability classes AV1 (dark gray) and AV2 (light gray), and three levels of BL, the % of overall capacity loss for the contracted bins (low = 20%, medium = 50%, and high = 70%).



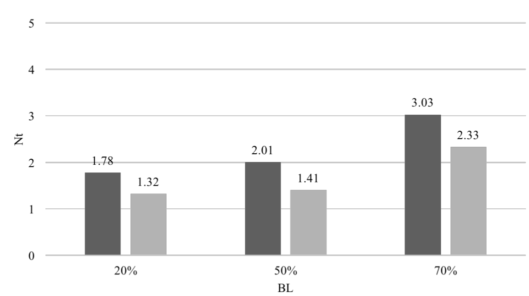
(a) Cap_{FS} *in instance set T3.*



(b) Cap_{FS} *in instance set T5.*



(c) N_t *in instance set T3.*



(d) N_t *in instance set T5.*

Figure 2: Average Cap_{FS} and N_t for capacity-loss levels, urban distribution setting, availability classes AV1 (dark gray) and AV2 (light gray)

It is noticeable that an increase of accurate information about the capacity loss in the urban distribution, compared to the long-haul transportation, allows the shipper to book in advance the same capacity as for the case with no capacity loss, but with the greater managerial flexibility of being able to select among a larger set of bin types. Thus, the structure of the capacity plan, reflected in Cap_{FS}, the percentage of total capacity contracted at the first stage, does not change in any significant manner with the variation of most parameters. The values observed (Figure 2) for the average number of bin types selected in the capacity plan, N_t, also support this observation, raising from an average of 1.87 when three bin types are available to 2.18 when five types are available (results for the volatile class AV1). Obviously, this number increases with the level of capacity loss (given by the BL parameter). This flexibility would prove beneficial given the availability of new transportation modes, e.g., cargo-bikes and light rail, for city logistics systems.

7 Conclusions

We focused in this paper on the tactical logistics capacity-planning problem arising in the supply-chain management context, which is relevant in both the long-haul transportation and urban distribution contexts. We addressed the planning problem faced by a shipper negotiating with a carrier a tactical plan-contract to secure the capacity, of various types in terms of size and cost, needed to perform recurring storage or transport activities of goods, packed in loads of various sizes, to respond to the demands of its own customers. over a given medium-term planning horizon. The contract negotiation is undertaken in an uncertain environment.

We introduced, for the first time in the literature, the issue of the availability of the contracted capacity when needed at operations time. We explicitly addressed and modeled the uncertainty related to the loss of contracted capacity, simultaneously with the uncertainty in demand, i.e., the number and sizes of the loads the shipper handles at each operation occurrence during the planning horizon, and the availability and cost of future capacity to be used in an ad-hoc (spot) manner when needed. We thus introduced the Stochastic Variable Cost and Size Bin Packing with Capacity Loss, SVCSBP-LS, problem, generalizing several bin packing problems of the literature.

We proposed a two-stage stochastic formulation with recourse to address the SVCSBP-LS, where the first stage is dedicated to selecting the capacity units of each type to include in the tactical capacity plan, while the second stage concerns the adjustments to the plan through acquisition of ad-hoc capacity on the spot market and the assignment of loads to the available capacity units, following the revelation of new information on the loads to handle, the loss of contracted capacity, and the characteristics and costs of capacity units available on the spot market. The SVCSBP-LS formulation minimizes the total cost of the contracted capacity, plus the expected costs of handling the loss of capacity and securing the ad-hoc capacity over the repetitions of activities during the planning horizon. We then proposed an efficient progressive-hedging-based metaheuristic adapted to the complexity of the SVCSBP-LS.

The proposed model and solution method have been validated for both the long-haul transportation and urban distribution contexts, through an extensive experimental campaign on a large set of instances. These two contexts not only qualify the methodology for two broad and important application areas, but also provide a rich experimental ground through differences in their physical and operational characteristics.

Computational results highlight the need to consider explicitly the uncertainty in capacity-planning applications, as well as the usefulness of building a stochastic programming model integrating the uncertainty on the actual volume of contracted capacity which is expected to be available during operations. Indeed, the benefits of using the stochastic programming SVCSBP-LS model, compared to solving deterministic formulations assuming perfect knowledge of the future, are significant. Not only the deterministic formulation yields infeasible capacity plans in several relevant situations, but the numerical analysis shows that the stochastic formulation results in improved operations management (prediction of the capacity needed) and economic benefits in terms of lower operating costs.

The solution method also provided the means to explore the different behaviors of the model in the urban distribution and long-haul transportation settings. Managerial insights were drawn, specific to each context, concerning the impact on the structure of the capacity plan of a wide range of variations in the uncertain parameters describing the context in which the firms operate, including the probability of the reduction of contracted capacity, the type and scope of the capacity loss, and the cost of replacing the lost capacity.

It is noteworthy that, when uncertainty on future availability of contracted and ad-hoc capacity is high and wide spread, it is advisable to book most capacity in advance; in fact, book more than expected to be needed when there is a high risk of capacity loss. On the contrary, when there is a high probability of losing a large amount of the contracted capacity but the availability of ad-hoc resources is not an issue, then, very little capacity should be booked in advance. The shipper should rather wait until the shipping date to purchase the necessary capacity at that moment's price. Finally, when the potential loss of capacity is highly localized, i.e., the loss concerns a few types of capacity only which might, however, be entirely missing, the shipper should contract the capacity in advance paying particular attention to the corresponding resource types.

Many interesting developments are still ahead regarding the tactical capacity planning problem under uncertainty. The generalization of the problem to address other important issues, such as the selection, and associated contracting, of a limited set of carriers among several service providers proposing different contract cost, capacity types, availability, and costs, appears of prime interest. Considering the service-quality rating of various carriers would nicely enrich this generalization. Extending the range of uncertainty issues considered to, e.g., the hazard types generating the loss of capacity or the shortage of ad-hoc resources, and the correlations which may occur among the future availability of contracted and ad-hoc capacity given the type of carrier and disturbing events, constitutes a challenging and important research and development avenue. The continuous development of efficient solution methods, for increasingly complex problem settings considered and the associated model formulations, is a necessary and challenging R&D field.

We also believe that the methodology proposed in this paper and the research stream evoked above are particularly relevant to the planning of resilient supply chains which have to adjust and operate in rapidly evolving contexts, as was observed during the Covid-19 pandemic and the recovery which started even before the pandemic is fully controlled. The benefits to decision-support science would come from advances in modeling uncertainty and tactical planning in complex situations and efficiently addressing the corresponding formulations. The benefits to transportation would follow from, on the one hand, the need to evaluate and understand in more depth, and model adequately, the various sources of uncertainty and hazards which characterize the application context, and, on the other hand, the increase in managerial agility with respect to decision making at planning and operation levels. We plan to contribute to these areas in the near future, in particular in the context of the forthcoming developments related to the Italian Recovery and Resilience Plan (part of the European recovery plan) (Minister of Economy and Finance, 2021).

Acknowledgments

Partial funding for this work was provided by the Canadian Natural Sciences and Engineering Research Council (NSERC) through its Discovery Grants Program. While working on this paper, the first author was Adjunct Professor, Department of Computer Science and Operations Research, Université de Montréal, Montréal, Canada, and the second author was the head of the *Urban Mobility and Logistics Systems* (UMLS) initiative of the interdepartmental *Center for Automotive Research and Sustainable mobility* (CARS) at Politecnico di Torino, Turin, Italy. While working to this paper, Prof. Guido Perboli was the representative in the Logistic Table in the Regional Council of Piedmont for Politecnico di Torino. The authors are grateful to Ms. Diane Larin for her invaluable linguistic and editorial assistance.

References

5G-LOGINNOV Consortium (2021). 5G-LOGINNOV Home Page. https://5g-loginnov.eu/. Last access: 31/03/2021.

Aghezzaf, E. (2005). Capacity planning and warehouse location in supply chains with uncertain demands. *Journal of the Operational Research Society*, 56(4):453–462.

Ahmed, S., King, A. J., and Parija, G. (2003). A multi-stage stochastic integer programming approach for capacity expansion under uncertainty. *Journal of Global Optimization*, 25(1):3–24.

Alibaba (2018). Alibaba Web Site. http://www.alibaba.com. Last access: 08/09/2019.

Amazon (2018). Amazon Web Site. https://www.amazon.com. Last access: 08/09/2019.

Ambra, T., Caris, A., and Macharis, C. (2019). Towards freight transport system unification: reviewing and combining the advancements in the physical internet and synchromodal transport research. *International Journal of Production Research*, 57(6):1606–1623.

Ballot, E., Montreuil, B., and Meller, R. D. (2014). *The Physical Internet*. La Documentation Française.

Bektaş, T. and Crainic, T.G. (2008). A Brief Overview of Intermodal Transportation. In Taylor, G.D., editor, *Logistics Engineering Handbook*, chapter 28, pages 1–16. Taylor and Francis Group, Boca Raton, FL, USA.

Bektaş, T., Crainic, T.G., and Van Woensel, T. (2017). From managing urban freight to smart city logistics networks. In Gakis, K. and Pardalos, P., editors, *Networks Design and Optimization for Smart Cities*, volume 8 of *Series on Computers and Operations Research*, pages 143–188. World Scientific Publishing.

Birge, J. and Louveaux, F. (1997). *Introduction to Stochastic Programming*. Springer Series in Operations Research and Financial Engineering. Springer.

Birge, J. R. (1982). The value of the stochastic solution in stochastic linear programs with fixed recourse. *Mathematical Programming*, 24(1):314–325.

Birge, J. R. (2012). Option methods for incorporating risk into linear capacity planning models. *Manufacturing & Service Operations Management*, 2(1):19–31.

Bosona, T. (2020). Urban freight last mile logistics—challenges and opportunities to improve sustainability: A literature review. *Sustainability (Switzerland)*, 12(21):8769.

Brotcorne, L., Perboli, G., Rosano, M., and Wei, Q. (2019). A Managerial Analysis of Urban Parcel Delivery: A Lean Business Approach. *Sustainability*, 11(12):34–39.

CNBC (2021). Aggressive fight over containers is causing shipping costs to rocket by 300%. https://www.cnbc.com/2021/01/22/shipping-container-shortage-is-causing-shipping-costs-to-rise.html. Last access: 11/06/2021.

Correia, I., Gouveia, L., and Saldanha-da-Gama, F. (2008). Solving the variable size bin packing problem with discretized formulations. *COR*, 35:2103–2113.

Crainic, T. G., Giusti, R., Manerba, D., and Tadei, R. (2021a). The synchronized location-transshipment problem. *Transportation Research Procedia*, 52:43–50. 23rd EURO Working Group on Transportation Meeting, EWGT 2020, 16-18 September 2020, Paphos, Cyprus.

Crainic, T.G. (2003). Long-Haul Freight Transportation. In Hall, R.W., editor, *Handbook of Transportation Science*, pages 451–516. Kluwer Academic Publishers, Norwell, MA, second edition.

Crainic, T.G., Gobbato, L., Perboli, G., and Rei, W. (2016). Logistics capacity planning: a stochastic bin packing formulation and a progressive hedging meta-heuristic. *European Journal of Operational Research*, 253:404–417.

Crainic, T.G., Gobbato, L., Perboli, G., Rei, W., Watson, J., and Woodruff, D. (2014). Bin packing problems with uncertainty on item characteristics: an application to capacity planning in logistics. *Procedia - Social and Behavioral Sciences*, 111:654–662.

Crainic, T.G. and Hewitt, M. (2021). Service Network Design. In Crainic, T.G., Gendreau, M., and Gendron, B., editors, *Network Design with Applications in Transportation and Logistics*, chapter 12, pages 347–382. Springer, Boston.

Crainic, T.G. and Kim, K.H. (2007). Intermodal Transportation. In Barnhart, C. and Laporte, G., editors, *Transportation*, volume 14 of *Handbooks in Operations Research and Management Science*, chapter 8, pages 467–537. North-Holland, Amsterdam.

Crainic, T.G., Marcotte, S., Rei, W., and Takouda, P. M. (2013). Proactive order consolidation in global sourcing. In Bookbinder, J. H., editor, *Handbook of Global Logistics: Transportation in International Supply Chains*, pages 501–530, New York, NY. Springer New York.

Crainic, T.G. and Montreuil, B. (2016). Physical Internet Enabled Hyperconnected City Logistics. In Taniguchi, E. and Thompson, R. G., editors, *9th International Conference on City Logistics, June 2015, Tenerife, Spain*, volume 12 of *Transportation Research Procedia*, pages 383–398. Elsevier.

Crainic, T.G., Perboli, G., Pezzuto, M., and Tadei, R. (2007). New bin packing fast lower bounds. *COR*, 34:3439–3457.

Crainic, T.G., Perboli, G., Rei, W., and Tadei, R. (2011). Efficient lower bounds and heuristics for the variable cost and size bin packing problem. *COR*, 38:1474–1482.

Crainic, T.G., Perboli, G., and Ricciardi, N. (2021b). City Logistics. In Crainic, T.G., Gendreau, M., and Gendron, B., editors, *Network Design with Applications in Transportation and Logistics*, chapter 16, pages 507–537. Springer, Boston.

Crainic, T.G., Ricciardi, N., and Storchi, G. (2009). Models for Evaluating and Planning City Logistics Transportation Systems. *Transportation Science*, 43(4):432–454.

Crainic, T.G. and Speranza, M.G. (2008). Logistics and the Internet. In Taylor, G.D., editor, *Logistics Engineering Handbook*, chapter 24, pages 1–14. Taylor and Francis Group, Boca Raton, FL, USA.

Crainic, T.G., Tadei, R., Perboli, G., and Baldi, M. M. (2012). The generalized bin packing problem: Models and bounds. In *Odysseus 2012, the 5th International Workshop on Freight Transportation and Logistics, Mykonos (Greece), May 21–25, 2012*, pages 347–350.

Giusti, R., Manerba, D., Bruno, G., and Tadei, R. (2019). Synchromodal logistics: An overview of critical success factors, enabling technologies, and open research issues. *Transportation Research Part E: Logistics and Transportation Review*, 129:92–110.

Giusti, R., Manerba, D., Perboli, G., Tadei, R., and Yuan, S. (2018). A New Open-source System for Strategic Freight Logistics Planning: The SYNCHRO-NET Optimization Tools. In *Transportation Research Procedia*, volume 30, pages 245–254.

Gobbato, L. (2015). *Stochastic programming for City Logistics: new models and methods.* PhD thesis, Politecnico di Torino.

Kaut, M., Vladimirou, H., Wallace, S. W., and Zenios, S. A. (2007). Stability analysis of portfolio management with conditional value-at-risk. *Quantitative Finance*, 7(4):397–409.

Kienzle, J., Crainic, T.G., Frejinger, E., and Bisaillon, S. (2021). The Intermodal Railroad Blocking and Railcar Fleet Management Planning Problem. Technical Report CIRRELT-2021, Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et les transports, Université de Montréal, Montréal, QC, Canada.

La Presse (2021). High costs to secure empty containers for long-haul maritime transport - "Des coûts comme ça, je n'ai jamais vu ça" . https://www.lapresse.ca/affaires/2021-06-11/pme/des-couts-comme-ca-je-n-ai-jamais-vu-ca.php. Last access: 11/06/2021.

Lerma, V. (2018). Stochastic bin packing models for capacity planning in logistic application. Models and policy simulations. Publication CIRRELT-2018-25, Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport, Université de Montréal, Montréal, QC, Canada.

Liu, S. and Papageorgiou, L. G. (2013). Multiobjective optimisation of production, distribution and capacity planning of global supply chains in the process industry. *Omega*, 41(2):369–382. Management science and environmental issues.

Lium, A.-G., Crainic, T.G., and Wallace, S. W. (2009). A study of demand stochasticity in service network design. *Transportation Science*, 43(2):144–157.

Luss, H. (1982). Operations research and capacity expansion problems: A survey. *European Journal of Operational Research*, 30(5):803–1022.

Macharis, C. and Bontekoning, Y.M. (2004). Opportunities for OR in Intermodal Freight Transport Research: A Review. *European Journal of Operational Research*, 153(2):400–416.

Maggioni, F. and Wallace, S. W. (2012). Analyzing the quality of the expected value solution in stochastic programming. *Annals of Operations Research*, 200:37–54.

Minister of Economy and Finance (2021). The Recovery and Resilience Plan Next Generation Italy. https://www.mef.gov.it/en/focus/The-Recovery-and-Resilience-Plan-Next-Generation-Italia/. Last access: 31/03/2021.

Monaci, M. (2002). *Algorithms for packing and scheduling problems*. PhD thesis, Universitá di Bologna, Bologna, Italy.

Monczka, R. M., Handfield, R. B., Giunipero, L. C., Patterson, J. L., and Waters, D. (2010). *Purchasing and supply chain management*. South-Western Cengage Learning.

Morganti, E., Seidel, S., Blanquart, C., Dablanc, L., and Lenz, B. (2014). The impact of e-commerce on final deliveries: Alternative parcel delivery services in france and germany. *Transportation Research Procedia*, 4:178 – 190.

Perboli, G., Brotcorne, L., Bruni, M. E., and Rosano, M. (2021). A new model for Last-Mile Delivery and Satellite Depots management: The impact of the on-demand economy. *Transportation Research Part E: Logistics and Transportation Review*, 145:102–184.

Perboli, G., Musso, S., Rosano, M., Tadei, R., and Godel, M. (2017). Synchro-modality and slow steaming: New business perspectives in freight transportation. *Sustainability (Switzerland)*, 9(10):1843.

Perboli, G. and Rosano, M. (2019). Parcel delivery in urban areas: Opportunities and threats for the mix of traditional and green business models. *Transportation Research Part C: Emerging Technologies*, 99:19 – 36.

Perboli, G., Rosano, M., Saint-Guillain, M., and Rizzo, P. (2018). A simulation-optimization framework for City Logistics. an application on multimodal last-mile delivery. *IET Intelligent Transport Systems*, 12(4):262–269.

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

Porelli, A., Hadjidimitriou, N. S., Rosano, M., and Musso, S. (2021). Enhancing port's competitiveness thanks to 5g enabled applications and services. In *3rd IEEE International Workshop on Smart & Sustainable Mobility & Logistics in Smart Cities - 2021*

IEEE Annual Computer Software and Applications Conference (COMPSAC), Virtual event, July 12–16, 2021, pages 1–6. forthcoming.

Qu, W., Rezaei, J., Maknoon, Y., and Tavasszy, L. (2019). Hinterland freight transportation replanning model under the framework of synchromodality. *Transportation Research Part E: Logistics and Transportation Review*, 131:308–328.

Reuters (2021). Alibaba's Cainiao launches container booking service, cites global cargo woes. https://www.reuters.com/article/us-alibaba-cainiao/alibabas-cainiao-launches-container-booking-service-cites-global-cargo-woes-idUSKBN29J07Y. Last access: 11/06/2021.

Rockafellar, R. T. and Wets, R. J.-B. (1991). Scenarios and policy aggregation in optimization under uncertainty. *Math. Oper. Res.*, 16(1):119–147.

Rodrigue, J.-P. and Notteboom, T. (2013). Containerized freight distribution in north america and europe. In Bookbinder, J. H., editor, *Handbook of Global Logistics: Transportation in International Supply Chains*, pages 219–246. Springer New York, New York, NY.

Saglietto, I. (2013). Towards a Classification of Fourth Party Logistics (4PL). *Universal Journal of Industrial and Business Management*, 1(3):104–116.

Singh, G., Sier, D., Ernst, A. T., Gavriliouk, O., Oyston, R., Giles, T., and Welgama, P. (2012). A mixed integer programming model for long term capacity expansion planning: A case study from the hunter valley coal chain. *European Journal of Operational Research*, 220(1):210–224.

Skender, H.P., Mircović, P.A., and Prudky, I. (2017). The Role of the 4PL Model in a Contemporary Supply Chain. *Scientific Journal of Maritime Research*, 31:96–101.

Synchro-NET Consortium (2017). Synchro-NET Home Page. https://www.synchronet.eu/. Last access: 31/03/2021.

Willenbrock, R., Basaras, P., Sterle, J., Hadjidimitriou, N. S., and Catana, E. (2021). 5G LOGistics value chain INNOVation. In *IEEE 5G Summit – IEEE 5G for CAM, Virtual event, May 11—12, 2021*, pages 1–4. forthcoming.

Yildiz, B. and Savelsbergh, M. (2019). Service and capacity planning in crowd-sourced delivery. *Transportation Research Part C: Emerging Technologies*, 100:177–199.

Yoon, J., Yildiz, H., and Talluri, S. (2016). Risk management strategies in transportation capacity decisions: An analytical approach. *Journal of Business Logistics*, 37(4):364–381.

Yuan, X. and Ashayeri, J. (2009). Capacity expansion decision in supply chains: A control theory application. *International Journal of Computer Integrated Manufacturing*, 22(4):356–373.

A PH-based meta-heuristic for the SVCSBP-LS

Build the discretization of the stochastic problem

We first rewrite the SVCSBP-LS stochastic (1)-(10) model using a discretized form. Sampling is applied to obtain a set of representative scenarios, namely the set $\mathcal{S}$, and these are used to approximate the expected cost associated with the second stage. For the first stage, let $y_j^{ts} = 1$ if bin $j \in \mathcal{J}^t$ of type $t \in T$ is selected under scenario $s \in \mathcal{S}$ and 0 otherwise. For the second stage, define $\mathcal{K}^s = \bigcup_\tau \mathcal{K}^{\tau s}$, where $\mathcal{K}^{\tau s}$ is the set of extra bins of type $\tau \in \mathcal{T}$ in scenario $s \in \mathcal{S}$, and let $\mathcal{I}^s$ be the set of items to pack under scenario $s \in \mathcal{S}$. Let $g^{\tau s}$ be the cost associated with bins of type $\tau \in \mathcal{T}$ in scenario $s \in \mathcal{S}$, $\mathcal{V}_j^{ts}$ be the actual volume of first-stage bin $j \in \mathcal{J}^t$ under scenario $s \in \mathcal{S}$, and v_i^s be the volume of item $i \in \mathcal{I}^s$ in scenario $s \in \mathcal{S}$. Then, variable $z_k^{\tau s}$ is equal to 1 if and only if extra bin $k \in \mathcal{K}^{\tau s}$ of type $\tau \in \mathcal{T}$ is selected in scenario $s \in \mathcal{S}$, and x_{ij}^s and x_{ik}^s are item-to-bin assignment variable for scenario $s \in \mathcal{S}$.

Given the probability p_s of each scenario $s \in \mathcal{S}$, the SVCSBP-LS problem (1)-(10) can be approximated by the following equivalent deterministic model:

$$\min_{y,z,x} \sum_{s \in \mathcal{S}} p_s \left[\sum_{t \in T} \sum_{j \in \mathcal{J}^t} f^t y_j^{ts} + \sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}^{\tau s}} g^{\tau s} z_k^{\tau s} + \sum_{t \in T} \sum_{j \in \mathcal{J}^t} c^t (V^t - \mathcal{V}_j^{ts}) y_j^{ts} \right] \tag{21}$$

$$\begin{aligned}
\text{s.t.} \quad & y_j^{ts} \geq y_{j+1}^{ts}, && \forall\, t \in T, j = 1, \dots, |\mathcal{J}^t| - 1, s \in \mathcal{S}, && (22)\\
& \sum_{j \in \mathcal{J}} x_{ij}^s + \sum_{k \in \mathcal{K}^s} x_{ik}^s = 1, && \forall\, i \in \mathcal{I}^s, s \in \mathcal{S}, && (23)\\
& \sum_{i \in \mathcal{I}^s} v_i^s x_{ij}^s \leq \mathcal{V}_j^{ts} y_j^{ts}, && \forall\, t \in T, j \in \mathcal{J}^t, s \in \mathcal{S}, && (24)\\
& \sum_{i \in \mathcal{I}^s} v_i^s x_{ik}^s \leq V^\tau z_k^{\tau s}, && \forall\, \tau \in \mathcal{T}, k \in \mathcal{K}^{\tau s}, s \in \mathcal{S}, && (25)\\
& y_j^{ts} = y_j^{ts'}, && \forall\, t \in T, j \in \mathcal{J}^t, s, s' \in \mathcal{S}, && (26)\\
& y_j^{ts} \in \{0,1\}, && \forall\, t \in T, j \in \mathcal{J}^t, s \in \mathcal{S}, && (27)\\
& z_k^{\tau s} \in \{0,1\}, && \forall\, \tau \in \mathcal{T}, k \in \mathcal{K}^{\tau s}, s \in \mathcal{S}, && (28)\\
& x_{ij}^s \in \{0,1\}, && \forall\, i \in \mathcal{I}^s, j \in \mathcal{J}, s \in \mathcal{S}, && (29)\\
& x_{ik}^s \in \{0,1\}, && \forall\, i \in \mathcal{I}^s, k \in \mathcal{K}^s, s \in \mathcal{S}. && (30)
\end{aligned}$$

Constraints (26) are referred as the non-anticipativity constraints. They ensure that the first-stage decisions are not tailored to the scenarios considered in $\mathcal{S}$. Indeed, all the scenario solutions must be equal to produce a single implementable plan. Thus, the non-anticipativity constraints link the first-stage variables to the second-stage variables, so the model is not separable.

To apply Lagrangean relaxation and make the model separable, we need a different expression of the non-anticipativity constraints. Let $\bar{y}_j^t \in \{0,1\}, \forall\, t \in T, j \in \mathcal{J}^t$, be the *global capacity plan*, i.e., the set of bins selected at the first stage. The following constraints are equivalent to (26):

$$\bar{y}_j^t = y_j^{ts}, \qquad \forall\, t \in T,\ j \in \mathcal{J}^t, s \in \mathcal{S}, \tag{31}$$

$$\bar{y}_j^t \in \{0,1\}, \qquad \forall\, t \in T,\ j \in \mathcal{J}^t. \tag{32}$$

Constraints (31) force the first-stage solution of each scenario to be equal to the global capacity plan. Constraints (32) are simply the integrality conditions on the selection of the bins. With this formulation of the non-anticipativity constraints, when we apply Lagrangean relaxation to (31), we can penalize individually the difference between the scenario solution and the global solution of each bin in the plan.

Following the decomposition scheme proposed by Rockafellar and Wets (1991), we relax constraints (31) using an augmented Lagrangean strategy. We thus obtain the following objective function for the overall problem:

$$\begin{aligned}\min_{y,z,x} \sum_{s\in\mathcal{S}} p_s \Bigg[&\sum_{t\in T}\sum_{j\in\mathcal{J}^t} f^t y_j^{ts} + \sum_{\tau\in\mathcal{T}}\sum_{k\in\mathcal{K}^{\tau s}} g^{\tau s} z_k^{\tau s} + \sum_{t\in T}\sum_{j\in\mathcal{J}^t} c^t (V^t - \mathcal{V}_j^{ts}) y_j^{ts} + \\ &+ \sum_{t\in T}\sum_{j\in\mathcal{J}^t} \lambda_j^{ts}(y_j^{ts} - \bar{y}_j^t) + \frac{1}{2}\sum_{t\in T}\sum_{j\in\mathcal{J}^t} \rho_j^t (y_j^{ts} - \bar{y}_j^t)^2 \Bigg]\end{aligned} \tag{33}$$

where $\lambda_j^{ts}, \forall j \in \mathcal{J}^t, \forall t \in T$, and $\forall s \in \mathcal{S}$, define the Lagrangean multipliers for the relaxed constraints and ρ_j^t is a penalty ratio associated with bin $j \in \mathcal{J}^t$ of type $t \in T$. Within function 33, let us consider the quadratic term. Given the binary requirements of y_j^{ts} and $\bar{y}_j^t$, the term becomes:

$$\sum_{t\in T}\sum_{j\in\mathcal{J}^t} \rho_j^t \left(y_j^{ts} - \bar{y}_j^t\right)^2 = \sum_{t\in T}\sum_{j\in\mathcal{J}^t} \left(\rho_j^t (y_j^{ts})^2 - 2\rho_j^t y_j^{ts} \bar{y}_j^t + \rho_j^t (\bar{y}_j^t)^2\right) = \tag{34}$$

$$= \sum_{t\in T}\sum_{j\in\mathcal{J}^t} \left(\rho_j^t y_j^{ts} - 2\rho_j^t y_j^{ts} \bar{y}_j^t + \rho_j^t \bar{y}_j^t\right). \tag{35}$$

Therefore, the objective function can be formulated as follows:

$$\begin{aligned}\min_{y,z,x} \sum_{s\in\mathcal{S}} p_s \Bigg[&\sum_{t\in T}\sum_{j\in\mathcal{J}^t} \left(f^t + c^{ts}(V^t - \mathcal{V}_j^{ts}) + \lambda_j^{ts} - \rho_j^t \bar{y}_j^t + \frac{\rho_j^t}{2} \right) y_j^{ts} + \\ &+ \sum_{\tau\in\mathcal{T}}\sum_{k\in\mathcal{K}^{\tau s}} g^{\tau s} z_k^{\tau s} - \sum_{t\in T}\sum_{j\in\mathcal{J}^t} \lambda_j^{ts} \bar{y}_j^t + \frac{1}{2}\sum_{t\in T}\sum_{j\in\mathcal{J}^t} \rho_j^t \bar{y}_j^t \Bigg].\end{aligned} \tag{36}$$

Given constraints (22)-(30) and the objective function (36), the relaxed problem is not separable by scenario. However, if the overall plan $\bar{y}_j^t, \forall t \in T$ and $\forall j \in \mathcal{J}^t$, is fixed to a given value vector (i.e., the expected value of the scenario solutions), then the model decomposes according to the scenarios in $\mathcal{S}$ and the scenario subproblems can be expressed as follows:

$$\min_{y,z,x} \quad \sum_{t \in T} \sum_{j \in \mathcal{J}^t} \left(f^t + c^{ts}(V^t - \mathcal{V}_j^{ts}) + \lambda_j^{ts} - \rho_j^t \bar{y}_j^t + \frac{\rho_j^t}{2} \right) y_j^{ts} + \sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}^{\tau s}} g^{\tau s} z_k^{\tau s} \tag{37}$$

$$\begin{aligned}
\text{s.t.} \quad & y_j^{ts} \geq y_{j+1}^{ts}, & \forall\, t \in T, j = 1, \ldots, |\mathcal{J}^t| - 1, s \in \mathcal{S}, \quad & (38)\\
& \sum_{j \in \mathcal{J}} x_{ij}^s + \sum_{k \in \mathcal{K}^s} x_{ik}^s = 1, & \forall\, i \in \mathcal{I}^s, s \in \mathcal{S}, \quad & (39)\\
& \sum_{i \in \mathcal{I}^s} v_i^s x_{ij}^s \leq \mathcal{V}_j^{ts} y_j^{ts}, & \forall\, t \in T, j \in \mathcal{J}^t, s \in \mathcal{S}, \quad & (40)\\
& \sum_{i \in \mathcal{I}^s} v_i^s x_{ik}^s \leq V^\tau z_k^{\tau s}, & \forall\, \tau \in \mathcal{T}, k \in \mathcal{K}^{\tau s}, s \in \mathcal{S}, \quad & (41)\\
& y_j^{ts} \in \{0, 1\}, & \forall\, t \in T, j \in \mathcal{J}^t, s \in \mathcal{S}, \quad & (42)\\
& z_k^{\tau s} \in \{0, 1\}, & \forall\, \tau \in \mathcal{T}, k \in \mathcal{K}^{\tau s}, s \in \mathcal{S}, \quad & (43)\\
& x_{ij}^s \in \{0, 1\}, & \forall\, i \in \mathcal{I}^s, j \in \mathcal{J}, s \in \mathcal{S}, \quad & (44)\\
& x_{ik}^s \in \{0, 1\}, & \forall\, i \in \mathcal{I}^s, k \in \mathcal{K}^s, s \in \mathcal{S}. \quad & (45)
\end{aligned}$$

Furthermore, by noting that λ_j^{ts} and ρ_j^t are exogenous constants for the model (37)-(45), we can reformulate each scenario subproblem as follows. We define $\overline{\mathcal{T}} = T \cup \mathcal{T}$ to be the overall set of bin types. For each scenario s, let $\mathcal{B}^{\overline{\tau}s} = \mathcal{J}^{\overline{\tau}} \cup \mathcal{K}^{\overline{\tau}s}$ be the set of available bins of type $\overline{\tau}$ in the subproblem and $\mathcal{B}^s = \bigcup_{\overline{\tau}} \mathcal{B}^{\overline{\tau}s}$ be the whole set of bins available in the subproblem. For $b \in \mathcal{B}^{\overline{\tau}s}$, let $\mathcal{V}_b^{\overline{\tau}s}$ be the actual volume of bin b (for $b \in \mathcal{K}^{\overline{\tau}s}$, $\mathcal{V}_b^{\overline{\tau}s} = V^{\overline{\tau}}$) and let $f_b^{\overline{\tau}s}$ define the fixed cost associated with bin b. The value of $f_b^{\overline{\tau}s}$ is given by

$$f_b^{\overline{\tau}s} = \begin{cases} f^{\overline{\tau}} + c^{\overline{\tau}s}(V^{\overline{\tau}} - \mathcal{V}_b^{\overline{\tau}s}) + \lambda_b^{\overline{\tau}s} - \rho_b^{\overline{\tau}} \bar{y}_b^{\overline{\tau}} + \frac{\rho_b^{\overline{\tau}}}{2} & \overline{\tau} \in \overline{\mathcal{T}}, b \in \mathcal{J}^{\overline{\tau}} \\ g^{\overline{\tau}s} & \overline{\tau} \in \overline{\mathcal{T}}, b \in \mathcal{K}^{\overline{\tau}s}. \end{cases} \tag{46}$$

Thus, each scenario subproblem can be reduced to a deterministic VCSBPP with modified fixed costs and an additional constraint that ensures an order in the selection of bins of type $\overline{\tau} \in \overline{\mathcal{T}}$:

$$\min_{y,x} \quad \sum_{\overline{\tau}\in\overline{\mathcal{T}}} \sum_{b\in\mathcal{B}^{\overline{\tau}s}} f_b^{\overline{\tau}s} y_b^{\overline{\tau}s} \tag{47}$$

$$\text{s.t.} \quad y_b^{\overline{\tau}s} \geq y_{b+1}^{\overline{\tau}s}, \qquad \forall\, \overline{\tau} \in \overline{\mathcal{T}}, b = 1, \ldots, |\mathcal{B}^{\overline{\tau}s}| - 1, \tag{48}$$

$$\sum_{b\in\mathcal{B}^s} x_{ib}^s = 1, \qquad \forall\, i \in \mathcal{I}^s, \tag{49}$$

$$\sum_{i\in\mathcal{I}^s} v_i^s x_{ib}^s \leq \mathcal{V}_b^{\overline{\tau}s} y_b^{\overline{\tau}s}, \qquad \forall\, \overline{\tau} \in \overline{\mathcal{T}}, b \in \mathcal{B}^{\overline{\tau}s}, \tag{50}$$

$$y_b^{\overline{\tau}s} \in \{0, 1\}, \qquad \forall \overline{\tau} \in \overline{\mathcal{T}}, b \in \mathcal{B}^{\overline{\tau}s}, \tag{51}$$

$$x_{ib}^s \in \{0, 1\}, \qquad \forall\, i \in \mathcal{I}^s, b \in \mathcal{B}^s, \tag{52}$$

where $y_b^{\overline{\tau}s} = 1$ if bin $b \in \mathcal{B}^{\overline{\tau}s}$ of type $\overline{\tau} \in \overline{\mathcal{T}}$ is selected, 0 otherwise.

Phase 1 of the meta-heuristic

Obtaining consensus among subproblems. At each iteration of the meta-heuristic, the solutions of the scenario subproblems are used to build a temporary global solution (the overall capacity plan). *Consensus* is then defined as scenario solutions being similar with regard to the first-stage decisions with the overall capacity plan and, thus, being similar among themselves. This section describes how the overall plan is computed. Moreover, we introduce strategies for the penalty adjustment when nonconsensus is observed and techniques to guide the search process by bounding the number of bins that can be selected at the first stage.

Defining the overall capacity plan. Let ν be the iteration counter in the PH algorithm. At each iteration, the algorithm solves subproblems (47)–(52), obtaining local solutions $y_b^{\overline{\tau}s\nu}$ $y_j^{\tau s\nu}$, $\forall s \in \mathcal{S}$, $\forall \overline{\tau} \in \overline{\mathcal{T}}$, and $\forall b \in \mathcal{B}^{\overline{\tau}s}$. The subproblem solutions are then combined in the overall capacity plan $\bar{y}_b^{\overline{\tau}\nu}$ by using the expected value operator, as shown in Equation (53). The weight used for each component is the probability p_s associated with the corresponding scenario.

$$\bar{y}_b^{\overline{\tau}\nu} = \sum_{s\in\mathcal{S}} p_s y_b^{\overline{\tau}s\nu}, \quad \forall \overline{\tau} \in \overline{\mathcal{T}}, \forall b \in \mathcal{B}^{\overline{\tau}}. \tag{53}$$

Moreover, we define an overall solution based on the number of bins in the capacity plan. Let $\delta^{\overline{\tau}s\nu} = \sum_{b\in\mathcal{B}^{\overline{\tau}}} y_b^{\overline{\tau}s\nu}$ be the total number of bins of type $\overline{\tau} \in \overline{\mathcal{T}}$ in the capacity plan for scenario subproblem $s \in \mathcal{S}$ at iteration ν. Equivalently to (53), using the expected value operator on $\delta^{\overline{\tau}s\nu}$ $\forall s \in \mathcal{S}$, we can define the overall capacity plan for each bin type $\overline{\tau} \in \mathcal{T}$ as

$$\bar{\delta}^{\overline{\tau}\nu} = \sum_{s\in\mathcal{S}} p_s \delta^{\overline{\tau}s\nu} = \sum_{s\in\mathcal{S}} p_s \sum_{b\in\mathcal{B}^{\overline{\tau}}} y_b^{\overline{\tau}s\nu} = \sum_{b\in\mathcal{J}^{\overline{\tau}}} \sum_{s\in\mathcal{S}} p_s y_b^{\overline{\tau}s\nu} = \sum_{b\in\mathcal{B}^{\overline{\tau}}} \bar{y}_b^{\overline{\tau}\nu}. \tag{54}$$

Equation (54) can be used to define the stopping criterion. Thus, we consider consensus to be achieved when the values of $\delta^{\overline{\tau}s\nu}$, $\forall s \in \mathcal{S}$, are equal to $\bar{\delta}^{\overline{\tau}\nu}$.

It is important to note that (53) and (54) do not necessarily produce a feasible capacity plan. When consensus is not achieved, the overall solution may not satisfy the integrality constraints on the first-stage decision variables. For nonconvex problems such as the SVCSBP-LS using the expected value as an aggregation operator does not guarantee that the algorithm converges to the optimal solution. Moreover, it cannot ensure that a good (feasible) solution will be obtained for the stochastic problem. Therefore, (53) and (54) are used as reference solutions with the goal of helping the search process of the PH algorithm to identify bins for which consensus is possible. Both are used in the penalty adjustment, while (54) is also used in the bounding strategy.

Penalty adjustment strategies. To promote consensus among the scenario subproblems, we adjust the fixed costs of bin types in the objective function at each iteration to penalize a lack of implementability and dissimilarity between local solutions and the overall solution. We propose two different strategies for these adjustments, both working at the local level in the sense that they affect every scenario subproblem separately.

The first strategy was originally proposed by Rockafellar and Wets (1991). Using information on the bin selection (i.e., variable $y_b^{\overline{\tau}s\nu}$), it operates on the fixed costs by changing the Lagrangean multipliers. For a given iteration ν, let $\lambda_b^{\overline{\tau}s\nu}$ be the Lagrangean multiplier associated with bin $b \in \mathcal{B}^{\overline{\tau}s}$ for scenario $s \in \mathcal{S}$, and let $\rho_b^{\overline{\tau}\nu}$ be the penalty deriving from the quadratic term. Note that the value of $\rho_b^{\overline{\tau}\nu}$ is variable-specific. At each iteration, we update the values $\lambda_b^{\overline{\tau}s\nu}$ and $\rho_b^{\overline{\tau}\nu}$, $\forall b \in \mathcal{B}^{\overline{\tau}s}$ and $\forall s \in \mathcal{S}$, as follows:

$$\lambda_b^{\overline{\tau}s\nu} = \lambda_b^{\overline{\tau}s(\nu-1)} + \rho_b^{\overline{\tau}(\nu-1)}(y_b^{\overline{\tau}s\nu} - \bar{y}_b^{\overline{\tau}\nu}) \tag{55}$$

$$\rho_b^{\overline{\tau}\nu} \leftarrow \alpha \rho_b^{\overline{\tau}(\nu-1)}, \tag{56}$$

where $\alpha > 1$ is a given constant and $\rho_b^{\overline{\tau}0}$ is fixed to a positive value to ensure that $\rho_b^{\overline{\tau}\nu} \to \infty$ as the number of iterations ν increases.

We initialize $\lambda_b^{s0} = 0$ for each scenario $s \in \mathcal{S}$. Equation (55) can then reduce, increase, or maintain this contribution according to the difference between the value of the bin-selection variables in the subproblem solutions and the overall capacity plan. The initial choice of $\rho_b^{\overline{\tau}0}$ is important. An inaccurate choice may cause premature convergence to a solution that is far from optimal or cause slow convergence of the search process. To avoid this, we set $\rho_b^{\overline{\tau}0}$ proportional to the fixed cost associated with the bin-selection variable: $\rho_b^{\overline{\tau}0} = max(1, f^{\overline{\tau}}/10)$, $\forall b \in \mathcal{B}^{\overline{\tau}s}$ and $\forall \overline{\tau} \in \overline{\mathcal{T}}$. The value of $\rho_b^{\overline{\tau}0}$ increases according to (56) as the number of iteration grows.

The second penalty adjustment is a heuristic strategy, which directly tunes the fixed costs of bins of the same type. The goal of this strategy is to accelerate the search process when the overall solution is close to consensus. When consensus is close, the difference between the subproblem solution and the overall solution may be small, and adjustments (55) and (56) lose their effectiveness, requiring several iterations to reach consensus.

Let $f^{\overline{\tau}s\nu}$ be the fixed cost of bin $b \in \mathcal{B}^{\overline{\tau}s}$ of type $\overline{\tau} \in \overline{\mathcal{T}}$ for scenario $s \in \mathcal{S}$ at iteration ν. At the beginning of the algorithm ($\nu = 0$), we impose $f^{\overline{\tau}s0} = f^{\tau}$. Then, when at least $\sigma_{\%}$ of the variables have reached consensus, we perturb every subproblem by changing $f^{\overline{\tau}s\nu}$ as follows:

$$f^{\overline{\tau}s\nu} = \begin{cases} f^{\overline{\tau}s(\nu-1)} \cdot M & \text{if } \delta^{\overline{\tau}s(\nu-1)} > \overline{\delta}^{\overline{\tau}(\nu-1)} \\ f^{\overline{\tau}s(\nu-1)} \cdot \frac{1}{M} & \text{if } \delta^{\overline{\tau}s(\nu-1)} < \overline{\delta}^{\overline{\tau}(\nu-1)} \\ f^{\overline{\tau}s(\nu-1)} & otherwise. \end{cases} \tag{57}$$

Here M is a constant greater than 1, while $\sigma_{\%}$ is a constant such that $0.5 \leq \sigma_{\%} \leq 1$. The current implementation of this heuristic strategy uses $\sigma_{\%} = 0.75$ and $M = 1.1$. The rationale for (57) is the following: if $\delta^{\overline{\tau}s(\nu-1)} > \overline{\delta}^{\overline{\tau}(\nu-1)}$, this means that in the previous iteration the number of bins of a given bin type $\overline{\tau}$ in scenario s was larger than the number of bins in the reference solution $\overline{\delta}^{\overline{\tau}(\nu-1)}$. Thus, the use of bins of type $\overline{\tau}$ is penalized by increasing the fixed cost by M. On the other hand, if $\delta^{\overline{\tau}s(\nu-1)} < \overline{\delta}^{\overline{\tau}(\nu-1)}$, we promote bins of type $\overline{\tau}$ by reducing the fixed cost by $1/M$.

Bundle fixing. To guide the search process, we introduce a variable-fixing strategy called bundle fixing.

We restrict the number of bins of each type that can be used, specifying lower and upper bounds. It should be noticed that it is equivalent to fix single bin-selection variables, since all bins of a certain type $\overline{\tau}$ are ordered and constraint 48 ensures that the selection of bins follows this order.

Let $\overline{\delta}_m^{\overline{\tau}\nu}$ and $\overline{\delta}_M^{\overline{\tau}\nu}$ be the minimum and maximum number of bins of type $\overline{\tau}$ involved in the overall solution at iteration ν:

$$\overline{\delta}_m^{\overline{\tau}\nu} \leftarrow \min_{s \in \mathcal{S}} \delta^{\overline{\tau}s\nu}, \tag{58}$$

$$\overline{\delta}_M^{\overline{\tau}\nu} \leftarrow \max_{s \in \mathcal{S}} \delta^{\overline{\tau}s\nu}. \tag{59}$$

At each iteration, the bundle strategy applies two bounds as follows. The lower bound $\overline{\delta}_m^{\overline{\tau}\nu}$ determines a set of compulsory bins that must be used in each subproblem;

to implement this we set the decision variables $y_b^{\overline{\tau}s(\nu+1)}$ to one for $b = 1, ..., \bar{\delta}_m^{\overline{\tau}\nu}$. The upper bound $\bar{\delta}_M^{\overline{\tau}\nu}$ is an estimate of the maximum number of bins of type $\overline{\tau}$ available in the next iteration; this reduces the number of decision variables in the subproblems. To implement this we remove decision variables $y_b^{\overline{\tau}s(\nu+1)}$ for $b = \bar{\delta}_M^{\overline{\tau}\nu} + 1, ..., \|\mathcal{B}^{\overline{\tau}}\|$.

Termination criteria. There are to date no theoretical results on the convergence of the PH algorithm for integer problems. Thus, we implement three stopping criteria for the search phase of the proposed meta-heuristic, based on the level of consensus reached and the number of iterations.

The level of consensus is measured through equations 58 and 59, as consensus is reached when $\bar{\delta}_m^{\overline{\tau}\nu} = \bar{\delta}_M^{\overline{\tau}\nu}, \forall \overline{\tau} \in \overline{\mathcal{T}}$. To speed up the algorithm, we actually stop the search, and proceed to Phase 2, as soon as consensus has been reached for all the bin types except one, type $\overline{\tau}'$, for which $\bar{\delta}_m^{\overline{\tau}'} < \bar{\delta}_M^{\overline{\tau}'}$.

When neither of the preceding conditions has been reached within a maximum number of iterations (200 in our experiments), the search is stopped and the meta-heuristic proceeds to the Phase 2.

Phase 2 of the meta-heuristic Phase 2 is thus invoked either when consensus is not achieved within a given maximum number of iterations, or the search was stopped when all but one bin type were in consensus.

In this case, there is only one bin type $\overline{\tau}'$ with $\bar{\delta}_m^{\overline{\tau}'} < \bar{\delta}_M^{\overline{\tau}'}$, that is, not in consensus. Given the efficiency of the item-to-bin heuristic, Phase 2 computes the final solution by iteratively examining the possible number of bins for $\overline{\tau}'$ (a consensus solution is always possible because $\bar{\delta}_M^{\overline{\tau}'}$ is feasible in all scenarios):

For all $\delta \in \left[\bar{\delta}_m^{\overline{\tau}'}, \bar{\delta}_M^{\overline{\tau}'}\right]$ **do**

- Set the number of bins of type $\overline{\tau}'$ to δ;
- Solve all the scenario subproblems with the heuristic;
- Check the feasibility of the solutions;
- Update the overall solution if a better solution has been found;

Produce the consensus solution.

When the maximum number of iterations is reached, consensus is less close. Phase 2 of the meta-heuristic then builds a restricted version of the formulation (21)–(30) by fixing the bin-selection first-stage variables for which consensus has been achieved, together with the associated item-to-bin assignment variables. The range of the bin types not in consensus is reduced through soft variable-fixing strategy. The resulting restricted MIP problem is solved exactly with a commercial solver.

B Analysis of $EVPI$ and VSS

In this appendix, we evaluate how the values of the $EVPI$ and VSS change depending on the parameters that characterize the actual volume reduction of first-stage bins (i.e., SL, TL and BL).

B.1 Expected value of perfect information

Table 7 reports the average and maximum percentages $EVPI$, showing how different parameters such as the level of the volume reduction, the percentage of scenarios affected by capacity losses and the probability that a bins type has a capacity reduction, affect the $EVPI$. Indeed, the above mentioned aspect is highlighted by the reduction of the average percentage $EVPI$ with an increase of SL and TL. For example, Table 7 highlights that when SL and TL are respectively equal to 20% and 50%, and BL is between 60% and 70%, the average and maximum percentages of $EVPI$ are 17% and 32% for instances with three bin types (set T3), and 16% and 25% for instance with five bin types (set T5).

Finally, the considerable risk of not being able to pack all items affects the decisions of the shipper, whose willingness to pay for the complete information about the future depends on the availability of bin types. Indeed, when the shipper can include in its capacity plan a wide range of bin types (in terms of volumes and types), its decisions are not affected by the availability of the second-stage bins, regardless of the context (long-haul transportation or urban distribution). In this case, at the shipping day, most likely it will be able to pack all the items using different configurations of bins or split them in different bin types. This aspect emerges by the results obtained considering the instances in T5 (see Tables 7 and 9).

On the contrary, the knowledge of the future becomes particularly important when the shipper can use a lower number of bin types and their availability could be limited at the shipping day. In this case, the risk of not being able to pack all items is high and the shipper may not be able to switch to other carrier who supply more capacity, with the consequent risk of unshipped products that turn into a loss of revenues. As highlighted in Table 8, this aspect is particularly relevant in the long-haul transportation, where considering three types of bins (set T3), the impacts of the number of scenarios affected by the uncertainty and the probability that a bin types has a capacity reduction depend on the availability of second-stage bins. For example, the average percentage of $EVPI$ reaches 46% when SL, TL, and BL are equal to 80%, 100%, and 70%, respectively (see Table 8).

B.1.1 Value of the Stochastic Solution (VSS)

In this section, we focus on the VSS. Tables from 10 to 12 report the average and maximum percentages VSS, showing how different parameters such as the level of the volume reduction, the percentage of scenarios affected by capacity losses and the probability that a bins type has a capacity reduction, affect the VSS.

As stated in the Section B.1, in the urban distribution, where the losses are localized, the stochastic approach is more valuable when there is a low probability of losing a large number of entire bins, which is for the example the case of unavailability of vans, when they are modeled as bins. Indeed, given the atomization of parcel flows (Morganti et al., 2014) and the high performance levels required by the contractual schemes in terms of number of delivery per day (Perboli and Rosano, 2019), an event that disrupts the regularity of operations and makes capacity fully unavailable, could have a huge impact on the service and profitability levels.

Indeed, in this case, Table 10 shows that the average VSS decreases as SL increases for both sets T3 and T5. The maximum values of VSS are reached when SL is equal to 20% and BL is 70%. In this case, the average and maximum percentages of VSS are 15% and 44% for T3 and 14% and 45% for T5.

In the case of the long-haul transportation (Table 11), when we consider instance set T3, when the availability of second-stage bins is limited and a considerable amount of capacity is likely to be lost in first-stage bins, the stochastic problem is not worth solving from a pure cost point of view, while the eventual infeasibility may be the real issue. In this case, the experimental tests revealed that when SL and TL are low, VSS increases as BL increases. On the contrary, when all the parameters have high values, VSS drops sharply. In particular, when we consider the availability class AV1 and SL, TL and BL are respectively equal to 80%, 75%, and 70%, and the average VSS percentage falls to 0%.

As in instance set T3, and even in instance set T5 (see Table 12), when SL and TL are low, the value of VSS increases as BL increases. In particular, the average percentage of VSS reaches 22% when SL, TL, and BL are respectively equal to 20%, 100%, and 70%, while the maximum percentage of VSS reaches 88%, with SL, TL, and BL respectively equaling 40%, 100%, and 70%. On the contrary, when SL and TL are high, the value of VSS decreases as BL increases and falls to 2% when SL, TL, and BL are respectively equal to 80%, 75%, and 70%.

SL[%]	TL[%]	BL[%]	Set T3 $EVPI[\%]$	Set T3 $EVPI[\%]_{max}$	Set T5 $EVPI[\%]$	Set T5 $EVPI[\%]_{max}$
20	**50**	**20-30**	16.26	26.62	15.77	23.91
		40-50	16.24	28.09	16.00	23.63
		60-70	16.90	31.99	16.15	24.90
	75	**20-30**	16.30	26.65	15.94	23.90
		40-50	16.33	28.47	15.95	23.31
		60-70	17.06	33.43	16.08	25.32
	100	**20-30**	16.45	26.13	15.93	23.84
		40-50	16.22	28.67	15.80	23.30
		60-70	17.00	33.75	16.00	24.48
40	**50**	**20-30**	16.33	26.54	15.98	23.97
		40-50	16.10	29.08	15.79	23.23
		60-70	17.05	33.93	16.16	25.57
	75	**20-30**	16.32	26.66	15.91	23.94
		40-50	15.99	29.99	15.61	23.16
		60-70	16.66	34.84	15.91	24.34
	100	**20-30**	16.25	26.35	15.92	23.60
		40-50	15.68	30.20	15.20	22.81
		60-70	16.03	35.49	15.40	23.29
60	**50**	**20-30**	16.23	26.70	16.07	23.75
		40-50	15.92	30.10	15.47	23.38
		60-70	16.52	48.07	15.99	24.46
	75	**20-30**	16.22	26.46	15.88	23.85
		40-50	15.47	28.95	15.06	22.81
		60-70	15.75	60.76	15.23	23.27
	100	**20-30**	16.11	26.50	15.79	23.56
		40-50	15.08	28.00	14.54	22.73
		60-70	14.59	50.01	14.21	21.93
80	**50**	**20-30**	16.23	26.55	15.95	23.79
		40-50	15.48	29.18	15.20	23.10
		60-70	15.84	36.60	15.47	23.81
	75	**20-30**	16.05	26.47	15.73	23.83
		40-50	15.04	30.29	14.46	22.61
		60-70	14.76	51.06	14.20	23.17
	100	**20-30**	15.87	25.68	15.48	23.50
		40-50	14.28	26.51	13.90	22.61
		60-70	12.99	28.73	12.81	21.73

Table 7: The impact of SL, TL and BL on $EVPI$ in the urban distribution setting.

SL[%]	**TL[%]**	**BL [%]**	**AV1**		**AV2-AV3**	
			$EVPI[\%]$	$EVPI_{[}\%]max$	$EVPI[\%]$	$EVPI_{[}\%]max$
		20-30	18.12	25.59	15.61	24.23
	50	**40-50**	20.27	34.30	17.28	27.36
		60-70	25.21	63.69	19.79	29.42
20		**20-30**	17.73	24.95	14.93	23.54
	75	**40-50**	20.67	38.63	16.66	24.65
		60-70	26.03	63.14	19.80	29.13
		20-30	16.58	24.60	13.68	22.41
	100	**40-50**	19.01	37.97	15.02	21.84
		60-70	27.83	61.04	18.59	25.83
		20-30	18.30	25.66	15.67	26.48
	50	**40-50**	23.20	40.98	18.45	29.05
		60-70	30.92	63.53	22.30	32.52
40		**20-30**	17.01	24.10	14.31	22.53
	75	**40-50**	22.16	40.79	17.31	26.21
		60-70	32.49	64.40	21.30	31.79
		20-30	14.75	22.89	11.56	21.09
	100	**40-50**	19.43	38.07	13.42	20.23
		60-70	35.60	65.67	16.50	25.29
		20-30	18.29	29.49	15.44	26.84
	50	**40-50**	25.35	50.96	19.16	31.13
		60-70	34.46	74.27	23.01	35.04
60		**20-30**	17.09	35.24	13.29	22.41
	75	**40-50**	24.91	52.88	16.71	27.18
		60-70	40.30	76.84	19.65	30.42
		20-30	13.85	33.34	8.91	19.27
	100	**40-50**	22.10	57.15	9.56	16.96
		60-70	43.53	77.34	10.39	18.53
		20-30	18.54	34.95	15.02	27.57
	50	**40-50**	27.77	56.97	19.09	32.17
		60-70	37.34	75.52	22.34	35.95
80		**20-30**	16.58	34.37	12.09	22.42
	75	**40-50**	25.22	53.94	14.95	25.70
		60-70	42.52	75.24	16.00	28.90
		20-30	12.09	31.38	6.27	16.55
	100	**40-50**	22.20	56.19	3.97	9.24
		60-70	46.16	77.35	4.14	8.15

Table 8: The impact of SL, TL and BL on $EVPI$ for instance set T3 in the long-haul transportation setting.

SL[%]	**TL**[%]	**BL** [%]	$EVPI[\%]$	$EVPI[\%]max$
20	**50**	**20-30**	16.05	24.71
		40-50	17.93	27.18
		60-70	20.51	31.33
	75	**20-30**	15.39	23.33
		40-50	17.65	27.78
		60-70	20.70	30.36
	100	**20-30**	13.87	22.11
		40-50	15.19	21.97
		60-70	18.51	27.49
40	**50**	**20-30**	16.28	25.73
		40-50	19.47	30.75
		60-70	23.47	35.77
	75	**20-30**	15.03	25.03
		40-50	18.53	28.08
		60-70	22.84	33.54
	100	**20-30**	11.83	19.62
		40-50	13.75	20.66
		60-70	17.41	26.68
60	**50**	**20-30**	16.31	27.08
		40-50	20.49	32.35
		60-70	24.92	36.79
	75	**20-30**	14.29	24.45
		40-50	18.46	31.79
		60-70	22.14	34.89
	100	**20-30**	9.42	18.44
		40-50	10.34	17.39
		60-70	12.04	30.31
80	**50**	**20-30**	16.14	28.30
		40-50	21.23	34.39
		60-70	25.04	43.02
	75	**20-30**	13.32	26.15
		40-50	17.44	32.46
		60-70	19.51	50.83
	100	**20-30**	7.07	16.82
		40-50	5.33	27.16
		60-70	6.61	54.26

Table 9: The impact of SL, TL and BL on $EVPI$ for instance set T5 in the long-haul transportation setting.

SL[%]	BL[%]	Set T3		Set T5	
		VSS[%]	VSS_{max}[%]	VSS[%]	VSS_{max}[%]
20	**20**	9.31	23.61	8.24	22.61
	30	9.21	24.83	7.98	23.48
	40	9.83	28.23	8.25	27.30
	50	12.24	32.97	10.26	32.69
	60	14.03	38.92	12.74	38.68
	70	15.49	44.49	13.82	45.00
40	**20**	9.24	23.61	7.88	22.61
	30	8.86	24.80	7.78	23.15
	40	9.31	29.54	7.95	27.27
	50	11.65	36.34	9.84	34.15
	60	13.11	40.16	11.20	39.31
	70	13.02	42.68	11.28	40.68
60	**20**	9.14	23.61	7.71	22.61
	30	8.71	23.27	7.68	22.44
	40	8.93	22.39	8.04	22.74
	50	11.12	25.71	9.55	23.97
	60	12.49	31.96	10.61	30.59
	70	12.08	38.24	10.42	27.85
80	**20**	9.09	23.61	7.75	22.61
	30	8.70	22.63	7.74	22.44
	40	9.03	22.36	8.09	22.83
	50	11.10	25.71	9.61	24.03
	60	12.39	32.66	10.45	31.22
	70	11.80	30.58	10.29	27.41

Table 10: The impact of SL, TL and BL on VSS in the urban distribution setting.

SL[%]	TL[%]	BL[%]	AV1		AV2-AV3	
			$VSS[\%]$	$VSS[\%]max$	$VSS[\%]$	$VSS[\%]max$
20	**50**	**20-30**	6.41	21.52	10.42	25.14
		40-50	8.32	23.05	13.19	29.33
		60-70	11.65	26.89	17.51	35.35
	75	**20-30**	6.19	19.88	10.41	27.02
		40-50	11.16	23.43	15.84	30.57
		60-70	12.00	25.33	19.47	41.04
	100	**20-30**	8.10	23.15	12.01	26.96
		40-50	12.08	22.68	17.08	32.72
		60-70	13.15	27.41	22.58	43.59
40	**50**	**20-30**	8.31	21.81	11.98	27.75
		40-50	12.02	27.09	17.47	36.46
		60-70	13.14	33.15	22.21	43.40
	75	**20-30**	10.28	22.39	13.57	27.58
		40-50	12.27	25.54	18.86	39.70
		60-70	14.24	31.47	21.90	57.95
	100	**20-30**	10.24	23.03	14.58	32.92
		40-50	11.75	23.87	20.65	47.63
		60-70	15.74	32.00	15.03	62.07
60	**50**	**20-30**	9.92	27.21	14.00	33.47
		40-50	12.41	24.71	19.64	40.68
		60-70	16.17	36.68	19.85	80.83
	75	**20-30**	10.52	24.11	14.74	32.86
		40-50	12.32	26.85	18.68	53.22
		60-70	23.70	31.38	8.00	78.57
	100	**20-30**	10.67	24.83	15.64	36.01
		40-50	11.27	26.44	14.18	62.05
		60-70	0.25	0.75	3.33	8.33
80	**50**	**20-30**	10.18	23.70	15.01	32.25
		40-50	12.44	28.37	19.10	56.13
		60-70	28.13	37.55	8.22	80.40
	75	**20-30**	10.36	23.92	14.96	35.19
		40-50	16.85	56.65	13.11	74.73
		60-70	0.00	0.00	1.95	30.88
	100	**20-30**	9.46	23.52	14.78	44.59
		40-50	8.45	20.00	6.46	52.34
		60-70	-	-	2.92	6.81

Table 11: The impact of SL, TL and BL on VSS for instance set T3 in the long-haul transportation setting.

SL[%]	**TL**[%]	**BL** [%]	VSS[%]	$VSS_[\%]max$
20	**50**	**20-30**	8.03	24.23
		40-50	10.12	35.71
		60-70	15.79	41.69
	75	**20-30**	7.94	24.87
		40-50	14.92	37.66
		60-70	19.66	43.39
	100	**20-30**	11.16	35.40
		40-50	17.18	39.62
		60-70	21.95	43.35
40	**50**	**20-30**	10.16	33.00
		40-50	16.80	38.14
		60-70	19.90	43.09
	75	**20-30**	13.56	35.41
		40-50	17.88	37.49
		60-70	20.48	58.54
	100	**20-30**	14.40	35.04
		40-50	18.47	45.31
		60-70	10.66	88.36
60	**50**	**20-30**	13.39	34.73
		40-50	18.10	38.91
		60-70	19.28	57.73
	75	**20-30**	14.51	35.28
		40-50	18.04	78.50
		60-70	7.17	84.46
	100	**20-30**	14.30	34.40
		40-50	11.27	64.27
		60-70	3.11	23.86
80	**50**	**20-30**	14.66	35.38
		40-50	17.49	74.47
		60-70	10.46	85.88
	75	**20-30**	14.41	35.69
		40-50	11.61	63.08
		60-70	1.88	19.62
	100	**20-30**	13.71	55.57
		40-50	4.46	51.44
		60-70	2.68	6.56

Table 12: The impact of SL, TL and BL on VSS for instance set T5 in the long-haul transportation setting.