



CIRRELT

Centre interuniversitaire de recherche
sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre
on Enterprise Networks, Logistics and Transportation

Logistics Capacity Planning: A Stochastic Bin Packing Formulation and a Progressive Hedging Meta-Heuristic

Teodor Gabriel Crainic
Luca Gobbato
Guido Perboli
Walter Rei

December 2014

CIRRELT-2014-66

Bureaux de Montréal :
Université de Montréal
Pavillon André-Aisenstadt
C.P. 6128, succursale Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
Pavillon Palais-Prince
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

Logistics Capacity Planning: A Stochastic Bin Packing Formulation and a Progressive Hedging Meta-Heuristic

Teodor Gabriel Crainic^{1,2,*}, Luca Gobbato^{1,3}, Guido Perboli¹, Walter Rei^{1,2}

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

² Department of Management and Technology, Université du Québec à Montréal, P.O. Box 8888, Station Centre-Ville, Montréal, Canada H3C 3P8

³ Department of Control and Computer Engineering, Politecnico di Torino, Corso Duca degli Abruzzi, 24 - I-10129 Torino, Italy

Abstract. In this paper, we consider the logistics capacity planning problem arising in the context of supply-chain management. We address, in particular, the tactical-planning problem of determining the quantity of capacity units, hereafter called bins, of different types to secure for the next period of activity, given the uncertainty on future needs in terms of demand for loads (items) to be moved or stored, and the availability and costs of capacity for these movements or storage activities. We propose a modeling framework for this problem introducing a new class of bin packing problems, the Stochastic Variable Cost and Size Bin Packing Problem (SVCSBPP). The resulting two-stage stochastic formulation with recourse assigns to the first stage the tactical capacity-planning decisions of selecting bins, while the second stage models the subsequent adjustments to the plan, securing extra bins, when needed, and packing the items into the selected bins each time the plan is applied and new information becomes known. We propose a new meta-heuristic to address this formulation. The meta-heuristic is based on progressive hedging ideas and includes several advanced strategies to accelerate the search and efficiently address the symmetry strongly present in the problem considered due to the presence of several equivalent bins of each type. Extensive computational results for a large set of instances support the claim of validity for the model, efficiency for the solution method proposed, and quality and robustness for the solutions obtained. The method is also used to explore the impact on the capacity plan and the recourse to spot-market capacity of a quite wide range of variations in the uncertain parameters and the economic environment of the firm.

Keywords: Logistics capacity planning, uncertainty, stochastic variable cost and size bin packing, stochastic programming, progressive hedging.

Acknowledgements. Partial funding for this project was provided by the Italian University and Research Ministry under the UrbeLOG project-Smart Cities and Communities. Partial funding was also provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), through its Discovery Grant program. We gratefully acknowledge the support of Fonds de recherche du Québec – Nature et technologies (FRQNT) through their infrastructure grants.

Results and views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect those of CIRRELT.

Les résultats et opinions contenus dans cette publication ne reflètent pas nécessairement la position du CIRRELT et n'engagent pas sa responsabilité.

* Corresponding author: Teodor-Gabriel.Crainic@cirrelt.ca

1 Introduction

The efficient consolidation of goods for movement and storage is essential for companies that aim to be competitive in the market and fulfill customer demand in a cost-efficient and flexible manner (Crainic et al., 2013b). Modern companies manufacture or acquire goods that are then consolidated into appropriate containers, which are in turn consolidated for long-distance, often international, transportation into ships or trains. Before reaching the final customer, the goods are stored for varying periods into a number of depots, warehouses or distribution centers. All these activities require that capacity, as containers, ship or train slots, motor carrier tractors, and warehousing space, be available when and where needed.

The planning of such *logistics capacity* directly affects the distribution and operating costs of the company, and therefore is a major challenge in supply chain management (Monczka et al., 2008). Logistics capacity planning as a tactical planning decision addresses the needs of the company for sufficient capacity to move and store its goods to meet demand in the next cycle of its activities. When tactical decisions are made, the company does not have detailed knowledge of many decision parameters such as the future demand and, thus, its real needs in terms of loads to be moved or stored, and the availability and costs of capacity. Such information becomes available, however, at the time for operational decisions, and it can be used to reconsider and update the capacity planning.

Our goal is providing models and methods to support decision making in logistics capacity planning. We focus, in particular, on the relation between a firm and its logistics capacity provider, and the tactical-planning problem of determining the quantity of capacity to secure for the next period of activity (from four to twelve months) given the uncertainty on future needs and availabilities. We propose a two-stage stochastic programming formulation based on the *Variable Cost and Size Bin Packing Problem (VCSBPP)* (Crainic et al., 2011b) to address this issue, together with an efficient meta-heuristic based on Progressive Hedging principles (Rockafellar and Wets, 1991).

Bin packing models offer important decision tools for transportation and logistics (Perboli et al., 2014a). In bin packing problems, various *items* must be packed into containers or other types of boxes and spaces, generally called *bins*, the particular characteristics of items and bins in terms of capacities and costs generating a rich set of problem settings and formulation. The packing formulations currently proposed in the literature deal only partially, however, with the requirements of capacity planning, as most research has explored operational issues only. Thus, most of the literature addressing bin packing and uncertainty, focuses on modeling uncertainty in item arrival, leading to the so-called on-line packing problem (Seiden, 2002). Studies focusing on stochastic versions of bin packing problems for strategic and tactical planning applications are few and recent. Perboli et al. (2012, 2014b) introduced stochastic costs and profits into a bin

packing and a knapsack problems. Crainic et al. (2014a) proposed a stochastic extension of the VCSBPP (Crainic et al., 2011b) for the capacity planning problem with demand uncertainty in terms of the number of items and their volumes.

These problems were formulated as two-stage integer stochastic programs with recourse (Birge and Louveaux, 1997). This approach separates the *a priori* planning decisions, taken under uncertainty before the activity cycle starts, and the adjustments performed at each period of operations once information becomes available. The former makes up the *first stage* of the model, while the *recourse* actions, defining the admissible adjustments to the plan, make up the *second stage*. Discretization methods are usually applied to approximate the probability distribution of the random parameters using a finite-dimension scenario tree (Mayer, 1998). Using the scenario tree approximation, one can reduce two-stage integer stochastic programs to deterministic linear programs with integer decision variables of generally very large dimensions, beyond the reach of exact methods. Indeed, Crainic et al. (2011a) and Crainic et al. (2014b) shown that a state-of-the-art MIP solver struggles to solve even small instances of stochastic VCSBPP formulations. Our goal is to successfully solve much larger problems. Therefore, heuristics are required.

Most solution approaches in the literature rely on decomposition techniques (Rosa and Ruszczyński, 1996; Hige and Sen, 1996). The *Progressive Hedging (PH)* algorithm, initially proposed by Rockafellar and Wets (1991) for stochastic convex programs, has emerged as an effective meta-heuristic for addressing stochastic programs with discrete decision variables (e.g., Crainic et al., 2011a; Carpentier et al., 2013; Crainic et al., 2014b; Veliz et al., 2014). It mitigates the computational difficulty associated with large problem instances by using augmented Lagrangian relaxation to decompose the stochastic program by scenario. It then iteratively solves penalized versions of the subproblems, gradually enforcing the consensus relative to the first-stage decision variables. PH enables the use of specialized heuristic methods for the subproblems, fast and accurate methods being crucial to the overall efficiency of the method, because subproblems are repeatedly solved.

The contribution of this paper is threefold.

1. A modeling framework for securing logistics capacity for a tactical-planning horizon, while explicitly considering uncertainty in both supply, availability and cost of various bin types, and demand as the number and dimension of items. In doing so, we introduce a new class of stochastic bin packing problems, namely the *Stochastic Variable Cost and Size Bin Packing Problem (SVCSBPP)*.
2. A new and effective meta-heuristic for the two-stage formulation of the SVCSBPP. The proposed solution method is based on PH ideas, uses the heuristic of Crainic et al. (2011b) to address the scenario subproblems, and includes several enhancements to deal with the particular characteristics of the problem. One par-

ticularly challenging characteristic is the symmetry inherent to packing problems, and strongly present in the SVCSBPP due to the presence of several equivalent bins of each type. The many equivalent but apparently different solutions in their choice of bins may significantly impair the efficiency of a PH-based meta-heuristic. Several strategies are thus proposed, including the definition of a new temporary global solution, heuristic penalty adjustments, and bundle fixing to reduce the bin selection at each iteration.

3. An extensive computational campaign using a large set of instances of various dimensions and combinations of different levels of variability in the stochastic parameters. The campaign includes the assessment of the efficiency of the meta-heuristic, the evaluation of the interest to explicitly consider uncertainty, an analysis of the impact of problem characteristics on capacity planning, the structure of solutions and the sensitivity to changes in data in particular, as well as an evaluation of the robustness of the decisions over a long time period under different economic and demand-evolution scenarios.

The numerical analysis indicates that the value of uncertainty is significant and that the proposed PH meta-heuristic finds good solutions fast. They also show that the proposed formulation is able to represent the impact of various levels of variation in the random parameters, and yields good-quality plans (solutions). Indeed, solutions are conservative and minimize the need to adjust the plan. Moreover, the capacity plan is robust (variations in the demand implying low operating costs) and relatively insensitive to changes. One may therefore conclude that SVCSBPP is a useful tool for logistics capacity planning under uncertainty, and that the proposed meta-heuristic is an appropriate and efficient solver for this class of formulations.

The paper is organized as follows. Section 2 details the logistics capacity planning problem we address in this paper. Section 3 presents the formulation of the problem. A lower bound and the meta-heuristic solution method are introduced in Sections 4 and 5, respectively. Section 6 presents the experimental plan and analyzes the computational results. Section 7 provides concluding remarks.

2 Logistics Capacity Planning

The tactical capacity planning problem we address is relevant in many contexts, e.g., manufacturing firms desiring to secure transportation capacity to bring in their resources or to distribute their products, wholesalers and retailers planning for transportation and storage capacity to support their procurement and sales processes, and logistics service providers securing capacity contracts with carriers for long-distance, regular shipments. Manufacturing and whole/retail distribution firms may negotiate directly with carriers

and owners/managers of storage space but, very often, they do business with a logistic service provider. 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 major retail firm and a third-party logistics (capacity) service provider (3PL).

In the contemporary economic and business environment, firms are engaged in a continuous procurement process (Aissaoui et al., 2007; Rizk et al., 2008), and engage in various collaborations with its supply-chain partners. Such inter-firm alliances yield several benefits, including a reduction in inefficiencies, total cost, and financial risks. The greatest advantages result from the outsourcing of logistics activities to a 3PL (Marasco, 2008). The overall supply-chain process may then be summarized as follows. The firm regularly orders products from suppliers in a given geographical region, according to current inventories, short-term forecast demand, estimated lead times, and specific procurement and inventory policies (Bertazzi and Speranza, 2005; Bertazzi et al., 2007). The suppliers are instructed to deliver their goods to a consolidation center (Crainic and Kim, 2007; Bertazzi and Speranza, 2012). The 3PL then consolidates the goods into containers and ensures their shipping, by consolidating these containers with those of other customers into the slots it secured on long-haul carriers, e.g., ships and trains (Chen et al., 2001; Crainic et al., 2013a), which generally operate according to a fixed schedule, e.g., twice a week.

A key factor in this process is the procurement of sufficient capacity, that we express in the following in terms of *bins*, at different locations in the network and for varying periods of time, to satisfy the demand. This entails negotiations with the 3PL to book the necessary capacity *a priori*, before operations start, at the best rate (Ford, 2001). The results of these negotiations often take the form of medium-term contracts specifying both the capacity to be used, the *quantity* and the *type* of bins, and the additional services to be performed (storage, transportation, bin operations, etc.) for a given planning horizon, e.g., a semester or a year. These contracts guarantee a regular volume of business to the 3PL (e.g., a fixed number of containers to ship every week for the next semester), which ensures a cost-effective service to the firm.

We refer to the costs of bins selected in advance as *fixed costs* because they are fixed by the contract and thus represent the specific rates offered by the 3PL for bins of different sizes. The values of the fixed costs are in practice influenced by several factors, including bin size, bin type (e.g., thermal or refrigerated containers), physical handling operations required, the time period for which the bin is to be used, and the economic characteristics of the departure location (e.g., access rules and costs). The result of negotiations and the scope of the capacity planning problem then is a tactical plan defining these quantities for the firm at each location, given the proposed bin types and costs and an estimation of the demand over the planning horizon. The plan thus specifies the set of bins of particular volumes and fixed costs to be made available at each location to ship the estimated items.

Given the time lag that usually exists between the signing of the contracts and the logistics operations, the negotiations are performed under uncertainty, without all the necessary information concerning the *demand*, expressed as a number of items of variable sizes to be shipped or stored at the particular locations. At each application of the plan, variations in the demand may thus yield numbers and sizes of items that differ from the estimation used at negotiation. These variations then require further negotiations with the 3PL and an adjustment of the plan when the booked capacity is not sufficient. Extra capacity (additional bins available at the shipping date) must then be purchased, generally at a much higher cost (the so-called spot -market value) than the fare negotiated initially. Moreover, while the 3PL generally ensures the planned capacity for the firm, the extra capacity may not be available when needed. The extra capacity must therefore be considered stochastic as well.

The Stochastic Variable Cost and Size Bin Packing Problem thus consists in determining the most cost-effective tactical plan, including the costs of the bins secured within the negotiations and the possible cost of updating the plan and securing extra capacity when the actual demand and the spot-market bin availability become known.

We propose a two-stage stochastic programming formulation, where the first stage concerns the selection of bins, and the second stage concerns the acquisition of extra capacity when the actual demand information is revealed. The next section details the Stochastic Variable Cost and Size Bin Packing (*VCSBP*) formulation.

3 The VCSBPM Model

Let \mathcal{T} be the finite set of bin types, each type being defined by the volume and fixed cost associated with its bins. Let V^τ and f^τ respectively be the volume and fixed cost associated with bins of type $\tau \in \mathcal{T}$. We define \mathcal{J}^τ to be the set of available bins of type τ and $\mathcal{J} = \cup_\tau \mathcal{J}^\tau$ to be the set of available bins at the first stage.

Let set Ω be the sample space of the random event, where $\omega \in \Omega$ defines a particular realization. Let vector ξ contain the stochastic parameters defined in the model, and $\xi(\omega)$ be a given realization of this random vector. Let the first-stage variables be $y_j^\tau = 1$ if bin $j \in \mathcal{J}^\tau$ is selected and 0 otherwise. The two-stage model of the SVCSBPP may then be formulated as

$$\min_y \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} f^\tau y_j^\tau + E_\xi[Q(y, \xi(\omega))] \quad (3.1)$$

$$\text{s.t. } y_j^\tau \in \{0, 1\}, \quad \forall \tau \in \mathcal{T}, j \in \mathcal{J}^\tau, \quad (3.2)$$

where $Q(y, \xi(\omega))$ is the extra cost paid for the capacity that is added at the second stage, given the tactical capacity plan y and the vector $\xi(\omega)$. The objective function (3.1) then

minimizes the sum of the total fixed cost of the tactical capacity plan and the expected cost associated with the extra capacity added during operation, while constraints (3.2) impose the integrality requirements on y .

To formulate $Q(y, \xi(\omega))$, we consider the following stochastic parameters in $\xi(\omega)$, with $\mathcal{K}^\tau(\omega)$, the set of available bins of type τ at the second stage; $\mathcal{K}(\omega) = \cup_\tau \mathcal{K}^\tau$, the set of available bins at the second stage; $g^\tau(\omega)$, $k \in \mathcal{K}^\tau(\omega)$, the associated fixed costs; $\mathcal{I}(\omega)$, the set of items to be packed; and v_i , $i \in \mathcal{I}(\omega)$, the item volumes. The second-stage variables are defined as follows: $z_k^\tau = 1$ if bin $k \in \mathcal{K}^\tau(\omega)$ is selected, 0 otherwise; $x_{ij} = 1$ if item $i \in \mathcal{I}(\omega)$ is packed into bin $j \in \mathcal{J}(\omega)$, 0 otherwise; and $x_{ik} = 1$ if item $i \in \mathcal{I}(\omega)$ is packed into bin $k \in \mathcal{K}(\omega)$, 0 otherwise. We then define the function $Q(y, \xi(\omega))$ as

$$Q(y, \xi(\omega)) = \min_{y, z, x} \sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}^\tau(\omega)} g^\tau(\omega) z_k^\tau \quad (3.3)$$

$$\text{s.t.} \quad \sum_{j \in \mathcal{J}} x_{ij} + \sum_{k \in \mathcal{K}(\omega)} x_{ik} = 1, \quad \forall i \in \mathcal{I}(\omega) \quad (3.4)$$

$$\sum_{i \in \mathcal{I}(\omega)} v_i(\omega) x_{ij} \leq V^\tau y_j^\tau, \quad \forall \tau \in \mathcal{T}, j \in \mathcal{J}^\tau \quad (3.5)$$

$$\sum_{i \in \mathcal{I}(\omega)} v_i(\omega) x_{ik} \leq V^\tau z_k^\tau, \quad \forall \tau \in \mathcal{T}, k \in \mathcal{K}^\tau(\omega) \quad (3.6)$$

$$x_{ij} \in \{0, 1\}, \quad \forall i \in \mathcal{I}(\omega), j \in \mathcal{J} \quad (3.7)$$

$$x_{ik} \in \{0, 1\}, \quad \forall i \in \mathcal{I}(\omega), k \in \mathcal{K}^\tau(\omega) \quad (3.8)$$

$$z_k^\tau \in \{0, 1\}, \quad \forall \tau \in \mathcal{T}, k \in \mathcal{K}^\tau(\omega). \quad (3.9)$$

The objective function (3.3) minimizes the cost associated with the extra bins selected at the second stage. Constraints (3.4) ensure that each item is packed into a single bin. Constraints (3.5) and (3.6) ensure that the total volume of items packed into each bin does not exceed the bin volume. Finally, constraints (3.7) to (3.9) impose the integrality requirements on all second-stage variables.

4 A Lower Bound for the SVCSBPP

We present in this section a *lower bound* (*LB*) for the VCSBP formulation (3.1)–(3.9), which provides a way to measure the quality of the heuristic proposed in Section 5.

LB is obtained by removing the item-to-bin assignment constraints (3.4) and aggregating the individual bin feasibility constraints (3.5) and (3.6). The resulting formulation (4.1)–(4.4) is a two-stage stochastic model with fixed recourse, which yields an optimal set of bins, including both those in the capacity plan and extra bins, with a total ca-

capacity sufficient for the items considered (see constraints (4.2)). Thus, the LB does not guarantee feasibility for individual bins.

$$\min_{y,z} \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} f_j^\tau y_j^\tau + \mathbb{E}_\xi \left[\sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}^\tau(\omega)} g_k^\tau(\omega) z_k^\tau \right] \quad (4.1)$$

$$\text{s.t.} \quad \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} V_j^\tau y_j^\tau + \sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}^\tau(\omega)} V_k^\tau z_k^\tau \geq \sum_{i \in \mathcal{I}(\omega)} v_i(\omega), \quad (4.2)$$

$$y_j^\tau \in \{0, 1\} \quad \forall \tau \in \mathcal{T}, j \in \mathcal{J}^\tau, \quad (4.3)$$

$$z_k^\tau \in \{0, 1\} \quad \forall \tau \in \mathcal{T}, k \in \mathcal{K}^\tau(\omega). \quad (4.4)$$

Note that the LB formulation is independent of the number of items. This reduces the number of variables in the model and makes it possible to find optimal solutions in reasonable computing times (e.g., using commercial MIP solvers).

5 Progressive Hedging-based Meta-heuristic

Algorithm 1 displays the proposed meta-heuristic for the VCSBP formulation, inspired by the PH algorithm.

The method first applies a scenario decomposition (SD) technique based on the augmented Lagrangian relaxation, which separates the stochastic problem by scenario. Here, the Lagrangian multipliers are used to penalize a lack of implementability due to differences in the first-stage variable values among scenario subproblems. Section 5.1 describes how the SVCSBPP can be decomposed into deterministic VCSBPP subproblems with modified fixed costs. Then, the method proceeds in two phases. Phase 1 aims to obtain consensus among the subproblems. At each iteration, the subproblems are first solved separately. Their solutions are then aggregated into a temporary overall solution, as defined in Section 5.2.1. The search process is gradually guided toward scenario consensus by adjusting the Lagrangian multipliers and the subproblem penalties (Section 5.2.2), based on the deviations of the scenario solutions from the overall solution, and by a variable bundle-fixing strategy (Section 5.2.3). The search process continues until the consensus is achieved or one of the termination criteria is met (see Section 5.3). When consensus is not achieved in the first phase, Phase 2 solves the restricted SVCSBPP obtained by fixing the first-stage variables for which consensus has been reached, e.g., the bins used in all the scenario subproblems. Section 5.5 finally describes the parallel implementation of the algorithm solving the subproblems concurrently on multiple processors.

Algorithm 1 PH-based meta-heuristic for the SVCSBPP

Scenario decompositionGenerate a set of scenarios \mathcal{S} ;

Decompose the resulting deterministic model (5.1)–(5.9) by scenario using augmented Lagrangian relaxation;

Phase 1 $\nu \leftarrow 0$; $\lambda_j^{\tau s \nu} \leftarrow 0$; $\rho_j^{\tau \nu} \leftarrow f^\tau / 10$;**while** Termination criteria not met **do** **For all** $s \in \mathcal{S}$, solve the corresponding Variable Cost and Size Bin Packing subproblem $\rightarrow y_j^{\tau s \nu}$; **Compute temporary global solution**

$$\bar{y}_j^{\tau \nu} \leftarrow \sum_{s \in \mathcal{S}} p_s y_j^{\tau s \nu}$$

$$\bar{\delta}^{\tau \nu} \leftarrow \sum_{s \in \mathcal{S}} p_s \delta^{\tau s \nu}$$

Penalty adjustment

$$\lambda_j^{\tau s \nu} = \lambda_j^{\tau s \nu - 1} + \rho_j^{\tau(\nu-1)} (y_j^{\tau s \nu} - \bar{y}_j^{\tau \nu})$$

$$\rho_j^{\tau \nu} \leftarrow \alpha \rho_j^{\tau(\nu-1)}$$

if consensus is at least $\sigma\%$ **then** Adjust the fixed costs $f^{\tau s \nu}$ according to (5.26); **Bundle fixing**

$$\bar{\delta}_m^{\tau \nu} \leftarrow \min_{s \in \mathcal{S}} \delta^{\tau s \nu}$$

$$\bar{\delta}_M^{\tau \nu} \leftarrow \max_{s \in \mathcal{S}} \delta^{\tau s \nu}$$

Apply variable fixing;

 $\nu \leftarrow \nu + 1$ **Phase 2****if** consensus not met for a single bin type τ' ($\bar{\delta}_m^{\tau'} < \bar{\delta}_M^{\tau'}$) **then** Identify the consensus number of bins δ of type τ' by enumerating $\delta \in [\bar{\delta}_m^{\tau'}, \bar{\delta}_M^{\tau'}]$ (and variable fixing)**else**

Fix consensus variables in model (5.1)–(5.9);

Solve restricted (5.1)–(5.9) model using a MIP solver.

5.1 Scenario decomposition of the SVCSBPP

We first reformulate the SVCSBPP stochastic model (3.1)–(3.9) using scenario decomposition. 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 $y_j^{\tau s} = 1$ if bin $j \in \mathcal{J}^\tau$ of type $\tau \in \mathcal{T}$ is selected under scenario $s \in \mathcal{S}$ and 0 otherwise. For the second stage, define $\mathcal{K}^s = \cup_{\tau \in \mathcal{T}} \mathcal{K}^{\tau s}$, 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 $s \in \mathcal{S}$. Then, variable $z_k^{\tau s}$ equals 1 if and only if the extra bin $k \in \mathcal{K}^{\tau s}$ of type $\tau \in \mathcal{T}$ is selected in scenario $s \in \mathcal{S}$, while x_{ij}^s and x_{ik}^s are the item-to-bin assignment variables for scenario $s \in \mathcal{S}$. Given the probability p_s of each scenario $s \in \mathcal{S}$, the VCSBP formulation (3.1)–(3.9) can be approximated by the following deterministic model:

$$\min_{y,z,x} \sum_{s \in \mathcal{S}} p_s \left[\sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} f^\tau y_j^{\tau s} + \sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}^{\tau s}} g^{\tau s} z_k^{\tau s} \right] \quad (5.1)$$

$$\text{s.t.} \quad \sum_{j \in \mathcal{J}} x_{ij}^s + \sum_{k \in \mathcal{K}^s} x_{ik}^s = 1 \quad \forall i \in \mathcal{I}^s, s \in \mathcal{S}, \quad (5.2)$$

$$\sum_{i \in \mathcal{I}^s} v_i^s x_{ij}^s \leq V^\tau y_j^{\tau s} \quad \forall \tau \in \mathcal{T}, j \in \mathcal{J}^\tau, s \in \mathcal{S}, \quad (5.3)$$

$$\sum_{i \in \mathcal{I}^s} v_i^s x_{ik}^s \leq V^\tau z_k^{\tau s} \quad \forall \tau \in \mathcal{T}, k \in \mathcal{K}^{\tau s}, s \in \mathcal{S}, \quad (5.4)$$

$$y_j^{\tau s} = y_j^{\tau t} \quad \forall \tau \in \mathcal{T}, j \in \mathcal{J}^\tau, s, t \in \mathcal{S}, \quad (5.5)$$

$$y_j^{\tau s} \in \{0, 1\} \quad \forall \tau \in \mathcal{T}, j \in \mathcal{J}^\tau, s \in \mathcal{S}, \quad (5.6)$$

$$z_k^{\tau s} \in \{0, 1\} \quad \forall \tau \in \mathcal{T}, k \in \mathcal{K}^{\tau s}, s \in \mathcal{S}, \quad (5.7)$$

$$x_{ij}^s \in \{0, 1\} \quad \forall i \in \mathcal{I}^s, j \in \mathcal{J}, s \in \mathcal{S}, \quad (5.8)$$

$$x_{ik}^s \in \{0, 1\} \quad \forall i \in \mathcal{I}^s, k \in \mathcal{K}^s, s \in \mathcal{S}. \quad (5.9)$$

Constraints (5.5) are referred to as the nonanticipativity constraints, forcing all scenarios to yield the same first-stage decisions. The nonanticipativity constraints thus link the first- and second-stage variables, making the model not separable. To make the model separable and apply Lagrangian relaxation one rewrites the nonanticipativity constraints. Let $\bar{y}_j^\tau \in \{0, 1\}$, $\forall \tau \in \mathcal{T}, j \in \mathcal{J}^\tau$, be the *global capacity plan* (also identified as the *temporary global solution*, i.e., the set of bins currently selected at the first stage). The following constraints are equivalent to (5.5):

$$\bar{y}_j^\tau = y_j^{\tau s} \quad \tau \in \mathcal{T}, \quad j \in \mathcal{J}^\tau, \quad s \in \mathcal{S}, \quad (5.10)$$

$$\bar{y}_j^\tau \in \{0, 1\} \quad \tau \in \mathcal{T}, \quad j \in \mathcal{J}^\tau, \quad (5.11)$$

constraints (5.10) forcing the first-stage solution of each scenario to be equal to the global capacity plan.

Following the decomposition scheme proposed by Rockafellar and Wets (1991), we relax constraints (5.10) using an augmented Lagrangian strategy. We obtain the following objective function for the overall problem:

$$\min_{y,z} \sum_{s \in \mathcal{S}} p_s \left[\sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} f^\tau y_j^{\tau s} + \sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}^{\tau s}} g^{\tau s} z_k^{\tau s} + \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} \lambda_j^s (y_j^{\tau s} - \bar{y}_j^\tau) + \frac{1}{2} \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} \rho_j^\tau (y_j^{\tau s} - \bar{y}_j^\tau)^2 \right], \quad (5.12)$$

where $\lambda_j^s, \forall j \in \mathcal{J}$ and $\forall s \in \mathcal{S}$, define the Lagrangian multipliers for the relaxed constraints and ρ_j^τ is a penalty ratio associated with bin $j \in \mathcal{J}^\tau$ of type $\tau \in \mathcal{T}$. Given the binary requirements on $y_j^{\tau s}$ and \bar{y}_j^τ , the quadratic term becomes

$$\begin{aligned} \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} \rho_j^\tau (y_j^{\tau s} - \bar{y}_j^\tau)^2 &= \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} (\rho_j^\tau (y_j^{\tau s})^2 - 2\rho_j^\tau y_j^{\tau s} \bar{y}_j^\tau + \rho_j^\tau (\bar{y}_j^\tau)^2) = \\ &= \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} (\rho_j^\tau y_j^{\tau s} - 2\rho_j^\tau y_j^{\tau s} \bar{y}_j^\tau + \rho_j^\tau \bar{y}_j^\tau), \end{aligned} \quad (5.13)$$

and the objective function can therefore be formulated as

$$\begin{aligned} \min_{y,z} \sum_{s \in \mathcal{S}} p_s \left[\sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} \left(f^\tau + \lambda_j^s - \rho_j^\tau \bar{y}_j^\tau + \frac{\rho_j^\tau}{2} \right) y_j^{\tau s} + \sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}^{\tau s}} g^{\tau s} z_k^{\tau s} \right] \\ - \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} \lambda_j^s \bar{y}_j^\tau + \frac{\rho_j^\tau}{2} \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} \bar{y}_j^\tau. \end{aligned} \quad (5.14)$$

When the value of the overall plan $\bar{y}_j^\tau, \forall \tau \in \mathcal{T}, \forall j \in \mathcal{J}^\tau$, is fixed, e.g., to the expected value of the scenario solutions, the last two terms of the objective function (5.14) become constants. The formulation then decomposes according to the scenarios in \mathcal{S} , each scenario subproblem becoming

$$\min_{y,z,x} \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} \left(f^\tau + \lambda_j^s - \rho_j^\tau \bar{y}_j^\tau + \frac{\rho_j^\tau}{2} \right) y_j^{\tau s} + \sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}^{\tau s}} g^{\tau s} z_k^{\tau s} \quad (5.15)$$

subject to constraints (5.2) - (5.4) and (5.6) - (5.9).

Let $\mathcal{B}^{\tau s} = \mathcal{J}^\tau \cup \mathcal{K}^{\tau s}$ be the set of available bins of type τ in the subproblem of scenario $s \in \mathcal{S}$. Let f_b^τ define the fixed cost associated with bin $b \in \mathcal{B}^{\tau s}$, where

$$f_b^\tau = \begin{cases} f^\tau + \lambda_j^s - \rho_j^\tau \bar{y}_j^\tau + \frac{\rho_j^\tau}{2} & \tau \in \mathcal{T}, b, j \in \mathcal{J}^\tau, \\ g^{\tau s} & \tau \in \mathcal{T}, b \in \mathcal{K}^{\tau s}, \end{cases} \quad (5.16)$$

the values of the multipliers λ_j^s and ρ_j^τ being determined, at each iteration exogenously determined with respect to the subproblem.

Each scenario subproblem can then be reduced to a deterministic VCSBPP with modified fixed costs:

$$\min_{y,x} \sum_{\tau \in \mathcal{T}} \sum_{b \in \mathcal{B}^{\tau s}} f_b^{\tau} y_b^{\tau} \quad (5.17)$$

$$\text{s.t.} \quad \sum_{\tau \in \mathcal{T}} \sum_{b \in \mathcal{B}^{\tau s}} x_{ib}^s = 1 \quad \forall i \in \mathcal{I}^s, s \in \mathcal{S}, \quad (5.18)$$

$$\sum_{i \in \mathcal{I}^s} v_i^s x_{ib}^s \leq V^{\tau} y_b^{\tau} \quad \forall \tau \in \mathcal{T}, \forall b \in \mathcal{B}^{\tau s}, s \in \mathcal{S}, \quad (5.19)$$

$$y_b^{\tau} \in \{0, 1\} \quad \forall \tau \in \mathcal{T}, \forall b \in \mathcal{B}^{\tau s}, s \in \mathcal{S}, \quad (5.20)$$

$$x_{ib}^s \in \{0, 1\} \quad \forall \tau \in \mathcal{T}, \forall b \in \mathcal{B}^{\tau s}, \forall i \in \mathcal{I}^s, s \in \mathcal{S}, \quad (5.21)$$

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

It is time-consuming to solve a large VCSBPP to optimality, even using a state-of-the-art commercial MIP solver (Correia et al., 2008). Moreover, VCSBPP subproblems must be solved multiple times in the PH-based meta-heuristic. Among the effective algorithms developed for the VCSBPP (Crainic et al., 2007; Baldi et al., 2014; Crainic et al., 2011b), we choose the heuristic of Crainic et al. (2011b), because of its proved efficiency on instances with up to 1000 items. The heuristic implements an adapted best-first decreasing strategy that sorts items and bins by nonincreasing order of volume and unit cost, respectively. The heuristic then sequentially assigns each item to the best bin, i.e., the bin with the maximum free space once the item is assigned.

5.2 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. Section 5.2.1 describes how the overall plan is computed given the symmetry challenge of the bin packing formulation. Section 5.2.2 then introduces the strategies for the penalty adjustment when nonconsensus is observed, while Section 5.2.3 describes the techniques we propose to guide the search by bounding the number of bins that can be selected at the first stage.

5.2.1 Defining the overall capacity plan

Let ν be the iteration counter in the PH meta-heuristic. At each iteration, the algorithm solves subproblems (5.17)–(5.21), obtaining local solutions $y_j^{\tau s \nu}$, $\forall s \in \mathcal{S}$, $\forall \tau \in \mathcal{T}$, and $\forall j \in \mathcal{J}^{\tau}$. The subproblem solutions are then combined yielding the overall capacity plan

$\bar{y}_j^{\tau\nu}$ by using the expected value operator, as shown in Equation (5.22). The weight used for each component is the probability p_s associated with the corresponding scenario.

$$\bar{y}_j^{\tau\nu} = \sum_{s \in \mathcal{S}} p_s y_j^{\tau s\nu}, \quad \forall \tau \in \mathcal{T}, \forall j \in \mathcal{J}^\tau. \quad (5.22)$$

This definition does not take into account, however, the presence of a large number of equivalent solutions that is typical of packing problems (Baldi et al., 2012). In fact, packing problems present a strong symmetry in the solution space. Two solutions are considered symmetric (and equivalent) when they involve the same set of bins in different orders. Equation (5.22) concerns the use of the specific bin $j \in \mathcal{J}^\tau$ and is therefore dependent on the order of the bins in the solution. Thus, it provides misleading information on the consensus, and does not constitute a good yardstick to measure the progress toward consensus (loosely called "convergence") of the overall solution.

To deal with the symmetry of the solution space, we define an overall solution based on the number of bins in the capacity plan. Let $\delta^{\tau s\nu} = \sum_{j \in \mathcal{J}^\tau} y_j^{\tau s\nu}$ be the total number of bins of type $\tau \in \mathcal{T}$ in the capacity plan for scenario subproblem $s \in \mathcal{S}$ at iteration ν . Equivalently to (5.22), using the expected value operator on $\delta^{\tau s\nu}$, $\forall s \in \mathcal{S}$, we can define the overall capacity plan for each bin type $\tau \in \mathcal{T}$ as

$$\bar{\delta}^{\tau\nu} = \sum_{s \in \mathcal{S}} p_s \delta^{\tau s\nu} = \sum_{s \in \mathcal{S}} p_s \sum_{j \in \mathcal{J}^\tau} y_j^{\tau s\nu} = \sum_{j \in \mathcal{J}^\tau} \sum_{s \in \mathcal{S}} p_s y_j^{\tau s\nu} = \sum_{j \in \mathcal{J}^\tau} \bar{y}_j^{\tau\nu}. \quad (5.23)$$

Equation (5.23) breaks the symmetry of the solutions (the order of the bins in the solution does not change the value of $\delta^{\tau s\nu}$) and can be used to define the stopping criterion. Thus, we consider consensus to be achieved when the values of $\delta^{\tau s\nu}$, $\forall s \in \mathcal{S}$, are equal to $\bar{\delta}^{\tau\nu}$.

It is important to note that (5.22) and (5.23) 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 SVCSBPP, 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, (5.22) and (5.23) are used as reference solutions with the goal of helping the search identify bins for which consensus is possible. Both measures are used for penalty adjustment (Section 5.2.2), while (5.23) is also used in the bounding strategy (Section 5.2.3).

5.2.2 Penalty adjustment strategies

To induce consensus among the scenario subproblems, we adjust the penalties in the objective function at each iteration to penalize 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 (e.g., variable $y_j^{\tau s \nu}$), it operates on the fixed costs by changing the Lagrangian multipliers. For a given iteration ν , let $\lambda_j^{s \nu}$ be the Lagrangian multiplier associated with bin $j \in \mathcal{J}^\tau$ for scenario $s \in \mathcal{S}$, and let $\rho_j^{\tau \nu}$ be the penalty deriving from the quadratic term. Note that the value of $\rho_j^{\tau \nu}$ is variable-specific. This approach outperforms scalar ρ strategies and guarantees faster convergence of the algorithm (Watson and Woodruff, 2011). At each iteration, we update the values $\lambda_j^{s \nu}$ and $\rho_j^{\tau \nu}$, $\forall j \in \mathcal{J}$ and $\forall s \in \mathcal{S}$, as follows:

$$\lambda_j^{s \nu} = \lambda_j^{s(\nu-1)} + \rho_j^{\tau(\nu-1)}(y_j^{\tau s \nu} - \bar{y}_j^{\tau \nu}) \quad (5.24)$$

$$\rho_j^{\tau \nu} \leftarrow \alpha \rho_j^{\tau(\nu-1)}, \quad (5.25)$$

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

We initialize $\lambda_j^{s 0} = 0$ for each scenario $s \in \mathcal{S}$. Equation (5.24) 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 $\rho_j^{\tau 0}$ 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 $\rho_j^{\tau 0}$ proportional to the fixed cost associated with the bin-selection variable: $\rho_j^{\tau 0} = \max(1, f^\tau/10)$, $\forall j \in \mathcal{J}^\tau$ and $\forall \tau \in \mathcal{T}$. The value of $\rho_j^{\tau \nu}$ increases according to (5.25) 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 (5.24) and (5.25) lose their effectiveness, requiring several iterations to reach consensus.

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

$$f^{\tau s \nu} = \begin{cases} f^{\tau s(\nu-1)} \cdot M & \text{if } \delta^{\tau s(\nu-1)} > \bar{\delta}^{\tau(\nu-1)} \\ f^{\tau s(\nu-1)} \cdot \frac{1}{M} & \text{if } \delta^{\tau s(\nu-1)} < \bar{\delta}^{\tau(\nu-1)} \\ f^{\tau s(\nu-1)} & \text{otherwise.} \end{cases} \quad (5.26)$$

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 (5.26) is the following: if $\delta^{\tau s(\nu-1)} > \bar{\delta}^{\tau(\nu-1)}$, this means that in the previous iteration the number of bins of a given bin type τ in scenario s was larger than the number of bins in the reference solution $\bar{\delta}^{\tau(\nu-1)}$. Thus, the use of bins of type τ is penalized by increasing the fixed cost by M . On the other hand, if $\delta^{\tau s(\nu-1)} < \bar{\delta}^{\tau(\nu-1)}$, we promote bins of type τ by reducing the fixed cost by $1/M$.

5.2.3 Bundle fixing

We introduce a variable-fixing strategy to guide the search process. Because there are multiple equivalent solutions, it might not be efficient to fix a single bin-selection variable $\bar{y}_j^{\tau\nu}$. We instead restrict the number of bins of each type that can be used, specifying lower and upper bounds. We call this strategy *bundle fixing*.

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

$$\bar{\delta}_m^{\tau\nu} \leftarrow \min_{s \in S} \delta^{\tau s \nu}, \quad (5.27)$$

$$\bar{\delta}_M^{\tau\nu} \leftarrow \max_{s \in S} \delta^{\tau s \nu}. \quad (5.28)$$

At each iteration, the bundle strategy applies two bounds to reduce the number of decision variables in the subproblems. The lower bound $\bar{\delta}_m^{\tau\nu}$ determines the set of compulsory bins that must be used in each subproblem by setting the decision variables $y_j^{\tau s(\nu+1)} = 1$ for $j = 1, \dots, \bar{\delta}_m^{\tau\nu}$. The upper bound $\bar{\delta}_M^{\tau\nu}$ is an estimate of the maximum number of bins of type τ available in the next iteration. This is performed by setting decision variables $y_j^{\tau s(\nu+1)} = 0$ for $j = \bar{\delta}_M^{\tau\nu} + 1, \dots, \|\mathcal{J}^\tau\|$.

5.3 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 (5.27) and (5.28), as consensus is reached when $\bar{\delta}_m^{\tau\nu} = \bar{\delta}_M^{\tau\nu}, \forall \tau \in \mathcal{T}$. 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 τ' , for which $\bar{\delta}_m^{\tau'} < \bar{\delta}_M^{\tau'}$.

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.

5.4 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 the first case, there is only one bin type τ' with $\bar{\delta}_m^{\tau'} < \bar{\delta}_M^{\tau'}$, 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 τ' (a consensus solution is always possible because $\bar{\delta}_M^{\tau'}$ is feasible in all scenarios):

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

Set the number of bins of type τ' to δ ;

Solve all the scenario subproblems with the VCSBPP 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 (5.1)–(5.9) 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 bundle fixing, and the resulting MIP is solved exactly.

5.5 Parallel implementation

Given the straightforward parallelization of the PH algorithm, we developed a synchronous master–slave implementation for the proposed meta-heuristic. The subproblems are assigned to a number of slaves. The master implements the meta-heuristic. It sends the values of the global information, overall solution and multiplier values, and collects the solutions from the slaves waiting until all the scenario subproblems have been solved. The master then proceeds to calculate the new overall solution. If consensus is not reached, the master updates the penalties of each subproblem and starts a new iteration. The parallelization reduces the computational time for each iteration and thus speeds up convergence. The quality of the solutions is not affected by the parallel execution, as the search process follows the same dynamics as in the sequential case.

It should be noted that when the computational time for the subproblems is unbalanced, the parallel algorithm is less efficient. To mitigate this effect, the master checks the current load of each slave (the number of subproblems remaining) and, if necessary, it can reassign subproblems among the slaves.

The complete list of duties of the master and slaves is:

- **Master**

1. Creates the pool of scenario subproblems;
2. Assigns each slave an equal number of subproblems;
3. Checks the load of the slaves and adjusts the assignments;
4. Computes the global solutions, computes the bounds, and updates the penalties.

- **Slave**

1. Solves the assigned subproblems;
2. Saves the solutions in a pool accessible by the master.

6 Computational Results

We performed an extensive set of experiments. The goals of the experimental campaign were to 1) analyze the performance of the proposed PH-based meta-heuristic by comparing it to a state-of-the-art commercial MIP solver; 2) measure the impact of uncertainty and determine whether building a stochastic programming model is really useful, and 3) explore the potential of the proposed model and algorithm by performing a number of analyzes of the structure, sensitivity and robustness of the logistics capacity plan under various problem settings.

We start by introducing the test instances generated for the numerical experiments (Section 6.1). We then proceed, Section 6.2, to analyze the performance of the proposed meta-heuristic by comparing its results (objective values and computational times), those of the direct solution of the multi-scenario deterministic problem (called *RP* in the following) (5.1)–(5.9), and the lower bound (LB) (4.1)–(4.4). We also analyze the efficiency of the parallel implementation by studying the scalability to 16 threads. Section 6.3 explores the benefits of using a two-stage model with recourse by computing the classical measures of the expected value of perfect information, *EVPI*, and value of the stochastic solution *VSS* (Birge, 1982; Maggioni and Wallace, 2012).

The potential of the proposed model and solution method is explored in three steps. Section 6.4 studies the structure of the capacity plan with respect to the attributes of the problem setting (bin-type usage and bin fill ratios); The sensitivity of the capacity planning to lower and upper limits on the number bins within each type one may used is analyzed in Section 6.5. Finally, we examine in Section 6.6 the robustness and reliability of the capacity planning decisions, through a Monte Carlo simulation under variations in the economic conditions (e.g., higher or lower demand).

The RP and LB formulations were solved using CPLEX version 12.5 (ILOG Inc., 2012) with maximum running time of 24 h. The PH-based meta-heuristic was coded in C++. It was stopped wither on reaching consensus or reaching 200 iterations. Phase 2 solves the restricted MIP using CPLEX with a time limit of one hour. Experiments were performed on a computer with 16 AMD Bulldozer cores at 2.3 GHz and 64 GB of RAM, at the high-performance computing cluster of Politecnico di Torino (DAUIN, 2014) .

6.1 Instances

We generated a new set of instances, denoted \mathbf{T} , starting from existing instances for different bin packing problem variants (Monaci, 2002; Crainic et al., 2007, 2011b, 2014a). The goal of the generation process was to allow exploring the structure of the capacity planning solutions for different configurations of bin types and items, as well as measuring the effect of different levels of uncertainty in the demand and the extra capacity. Instances are characterized by the number of bin types, the availability and the cost of the bins in the first and second stages, and the number and volume of the items. These characteristics are:

- **Number of bin types.** We consider instances with 3 (T3), 5 (T5), and 10 (T10) bin types. The bin volumes are:
 - 50, 100, 150 for T3;
 - 50, 80, 100, 120, 150 for T5;
 - 50, 60, 70, 80, 100, 110, 120, 130, 140, 150 for T10.
- **Availability of bins.** The number of bins of type $\tau \in \mathcal{T}$ available in the first stage, $\|\mathcal{J}^\tau\|$, is the minimum number of bins of volume V^τ needed to pack all items in the worst-case scenario (the scenario with the most items). This number is

$$\|\mathcal{J}^\tau\| = \left\lceil \frac{1}{V^\tau} \max_{s \in \mathcal{S}} \sum_{i \in \mathcal{I}^s} v_i^s \right\rceil. \quad (6.1)$$

To ensure a large variability at the second stage, $\|\mathcal{K}^{\tau s}\|$, the number of bins of type $\tau \in \mathcal{T}$ in scenario $s \in \mathcal{S}$, is uniformly distributed in the range $[0, \|\mathcal{J}^\tau\|]$. These rules mean that the worst-case scenario may involve a limited number of extra bins.

- **Fixed cost of bins.** For the set of bins available in the first stage the fixed cost is

$$f^\tau = V^\tau(1 + \gamma), \quad (6.2)$$

where γ is uniformly distributed in the range $[-0.3, 0.3]$. According to Correia et al. (2008) this range replicates realistic situations. The fixed cost for extra bins $g^{\tau s}$

at the second stage is the original fixed cost f^τ multiplied by a factor $(1 + \alpha_\tau^s)$, inversely proportional to the availability of extra bins of type τ in scenario $s \in \mathcal{S}$, where

$$\alpha_\tau^s = 1 - \frac{\|\mathcal{K}^{\tau s}\|}{\sum_{\tau \in \mathcal{T}} \|\mathcal{K}^{\tau s}\|} \cdot \beta, \quad (6.3)$$

and $\beta \in U[0, 0.5]$. Thus, the maximum increase in the fixed cost is 50%.

- **Number of items** at the second stage is uniformly distributed in the range $[25, 100]$ for T3 and T5, $[100, 500]$ for T10.
- **Volume of items.** An item is defined as:
 - Small (S): volume in the range $[5, 10]$;
 - Medium (M): volume in the range $[15, 25]$;
 - Big (B): volume in the range $[20, 40]$.

These categories are then combined into four volume-spread classes reflecting different realistic settings:

- SP1: high percentage of small items (S=60%, M=20%, B=20%);
- SP2: high percentage of medium items (S=20%, M=60%, B=20%);
- SP3: high percentage of big items (S=20%, M=20%, B=60%);
- SP4: no restrictions on the maximum number of items in each category.

For each combination of the parameters defined above, we generated 10 instances, yielding a total of 120 instances. A total of $s = 1, \dots, 100$ scenarios were generated. This number of scenarios guarantees in-sample stability (Kaut et al., 2007), which means that the value of the optimal decisions of the first-stage variables does not change when a different set of scenarios is considered.

6.2 PH meta-heuristic performance analysis

We started by examining computational performance of the MIP solver of CPLEX. The quality of the solutions is measured by the relative optimality gap defined as the difference between the best known solution UB^{CP} (i.e., the incumbent solution) and the best bound LB^{CP} :

$$\Delta^{CP} = \frac{UB^{CP} - LB^{CP}}{LB^{CB}} \cdot 100. \quad (6.4)$$

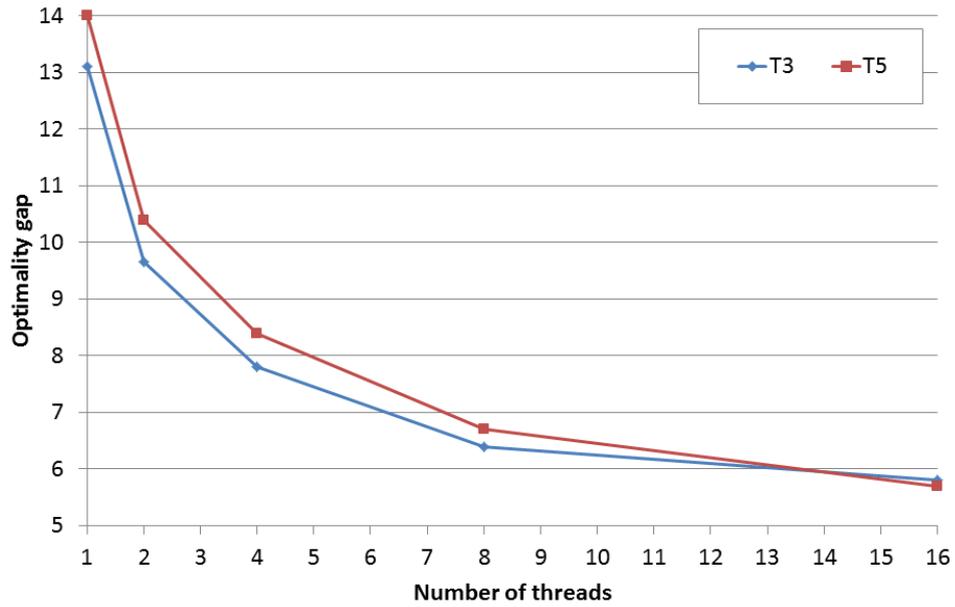


Figure 1: CPLEX average relative optimality gap (%) with respect to the number of parallel threads

Initial experiments on the T3 and T5 instance sets showed Δ^{CP} figures greater than 10% after 24 hours of execution. This underlined the difficulty of even very good commercial MIP solvers such as CPLEX to solve even limited-size problem instances. The high degree of symmetry inherent to the problem contributes to this difficulty.

We therefore used CPLEX in parallel mode. Figure 1 shows the trend of the average relative optimality gap for T3 and T5 instances with respect to the number of parallel threads. The gap decreases rapidly as the number of threads increases. With 16 parallel threads, CPLEX has an average optimality gap below 6%. The results of parallel CPLEX with 16 threads are therefore used in the comparisons of this section.

Table 1 details the comparison of CPLEX and the PH meta-heuristic for the T3 and T5 instances; CPLEX was unable to solve T10 with a reasonable optimality gap. We report, for each combination of instance set (Column 1) and item spread class (Column 2), the average optimality gaps Δ^{CP} of CPLEX (Column 3), the average running times t^{CP} of CPLEX (Column 4), the average running times t^{PH} of the meta-heuristic (Column 6), and, in Column 5, the average gaps Δ_{UB}^{PH} of the meta-heuristic computed as

$$\Delta_{UB}^{PH} = \frac{UB^{PH} - UB^{CP}}{UB^{CP}}, \quad (6.5)$$

where UB^{PH} is the objective value of the meta-heuristic. Note that the value of Δ_{UB}^{PH} may be positive or negative. When Δ_{UB}^{PH} is negative, the meta-heuristic solution is better than that of CPLEX.

Set	Spread	CPLEX		PH	
		Δ^{CP}	t^{CP} [s]	Δ_{UB}^{PH}	t^{PH} [s]
T3	SP1	5.38	86400	-1.19	1.21
	SP2	5.74	86400	-1.39	2.38
	SP3	6.38	86400	-1.06	0.40
	SP4	5.77	86400	-0.86	0.48
T5	SP1	4.99	86400	-1.17	2.08
	SP2	5.23	86400	-1.38	0.75
	SP3	6.44	86400	-1.98	0.62
	SP4	5.20	86400	-1.36	2.70

Table 1: Comparison of CPLEX and PH-based meta-heuristic

As stated before, CPLEX cannot solve the SVCSBPP to optimality within a reasonable time limit for any instance considered. The average optimality gap after 24 hours is always greater than 5%, with a maximum value of 6.44%. The PH meta-heuristic is accurate and effective on all T3 and T5 instances. It always converges quickly to better solutions than those obtained by CPLEX. The gap Δ_{UB}^{PH} is negative for all instances with an average improvement in the solution values between 0.86% and 1.98%. The PH meta-heuristic always reaches a consensus solution in less than 3s. This performance is directly related to the efficiency of the heuristic solver for the VCSBPP, which is able to solve deterministic subproblems in negligible computational times.

Despite the use of concurrent computation, CPLEX is not able to identify a good feasible solution for T10 instances in 24h of computing time. These instances can easily involve more than 10 million variables. They require a huge amount of memory for the branch and bound (e.g., more than 64 GB of memory already occupied after 2 h of computation). Memory thus rapidly becomes a bottleneck, and CPLEX cannot run for the assigned time. Thus, to validate the PH meta-heuristic on the remaining instances, we compare its solutions with those obtained by solving the LB (4.1)–(4.4). Recall that the LB does not consider item-to-bin assignments, which drastically reduces the number of variables in the model. The reduced model can thus be solved to optimality by CPLEX with a limited computational effort. The time depends on the number of bins involved, so CPLEX requires only a few seconds to solve LB for sets T3, T5, and it needs at most 120s for set T10.

Table 2 reports, for each instance set (Column 1) and item spread class (Column 2), the percentage gap Δ_{LB}^{PH} between the objective function value obtained by the PH meta-heuristic and LB (Column 3), and the average and maximum percentage increase in the objective function value (Columns 4 and 5) resulting from the use of the LB solutions as the capacity plan. On average, the gaps are smaller than 2%, underlining the accuracy of the PH meta-heuristic. We further qualified this accuracy by evaluating the quality of the LB solutions by fixing the bin selections defined by LB and computing the recourse

Set	Spread	Δ_{LB}^{PH}	C^{LB}	C_{max}^{LB}
T3	SP1	0.88	0.21	1.45
	SP2	0.72	0.22	1.79
	SP3	0.68	0.05	0.45
	SP4	1.21	0.00	0.00
T5	SP1	1.57	0.09	0.63
	SP2	1.45	0.00	0.00
	SP3	1.03	0.06	0.47
	SP4	2.03	0.10	0.77
T10	SP1	0.81	0.04	0.31
	SP2	0.79	0.03	0.18
	SP3	0.76	0.02	0.19
	SP4	0.86	0.01	0.06

Table 2: Comparison of LB and PH solutions

cost. For instances T3, T5, and T10, the average increase in the objective function was smaller than 0.25%. No increase, $C^{LB} = 0$, was observed in a number of cases, signaling that the capacity plans defined by LB and the PH meta-heuristic match exactly.

We conclude this section by analyzing the scalability of the parallel implementation of the meta-heuristic with up to 16 threads. The computational wall-clock times t^{PH} for each instance set (Column 1) and item spread class (Column 2) are reported in Table 3 for 1, 2, 4, 8 and 16 threads. The average speedups for the three instance sets, defined as the ratio of the sequential running time to the parallel running time, are displayed in Figure 2.

Computational times are negligible for T3 and T5 instances, while becoming more significant for the larger instances, the computational effort increasing with the number of items and especially with the number of bin types (e.g., instances in T10). Yet, the maximum sequential time of the meta-heuristic, of the order of 600 wall-clock seconds, is still several orders of magnitude smaller than the time required by CPLEX.

The scaling of the master-slave parallel implementation is quite linear. The speedups are of the order of 4 with 16 parallel threads for T3 and T5 instances, while they reach almost 14 (80% of the linear speedup) for T10 instances. Recalling that the VCSBP-Heuristic is very fast on all instances (which results in a balanced workload for slaves), the results indicate that for parallel computation is not really required for the small instances (and certainly not 16 threads; 4 are giving a 50% speedup), most time being spent on exchanging information between the master and the slaves. Parallel computing becomes appropriate as dimensions grow, which indicates that large problem instance can be addressed efficiently with the proposed methodology.

Set	Spread	t_1^{PH}	t_2^{PH}	t_4^{PH}	t_8^{PH}	t_{16}^{PH}
T3	SP1	1.21	0.87	0.73	0.50	0.30
	SP2	2.36	1.73	1.41	0.97	0.59
	SP3	0.41	0.29	0.25	0.17	0.10
	SP4	0.49	0.35	0.29	0.20	0.12
T5	SP1	2.08	1.48	1.15	0.83	0.50
	SP2	0.75	0.53	0.42	0.30	0.18
	SP3	0.62	0.44	0.35	0.25	0.15
	SP4	2.70	1.93	1.50	1.08	0.65
T10	SP1	625.17	350.10	187.08	89.02	45.62
	SP2	580.59	325.52	174.53	85.05	42.56
	SP3	432.48	241.24	122.87	74.25	31.52
	SP4	524.87	293.35	155.24	75.85	38.47

Table 3: Average computational times for the meta-heuristic with respect to parallel threads

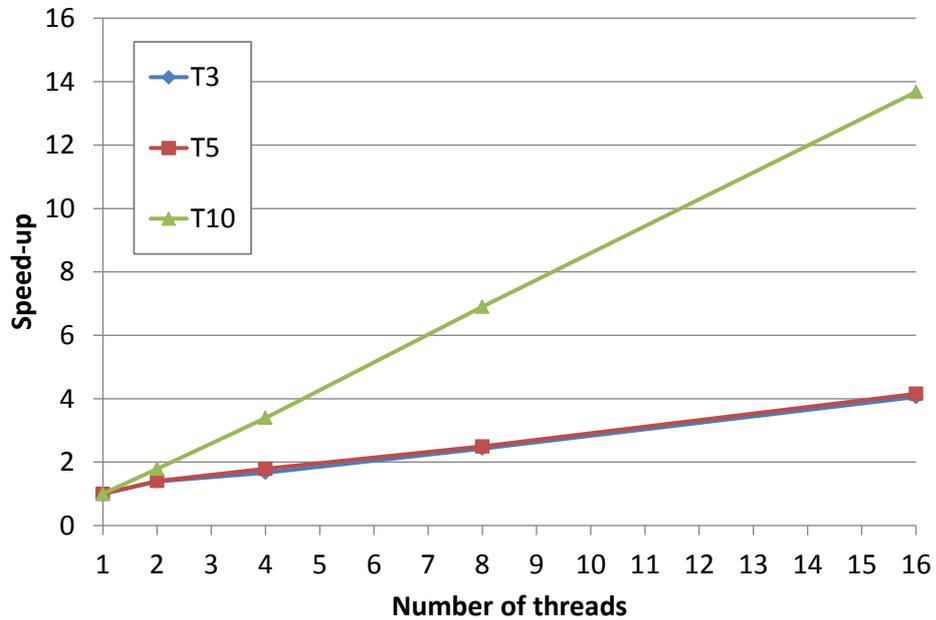


Figure 2: Speed-up of PH algorithm for different numbers of parallel threads

6.3 Benefit of modeling the uncertainty

The question addressed in this section is whether modeling uncertainty explicitly through the SVCSBPP two-stage model with recourse is warranted compared to solving some deterministic variant of the problem. We address this question by comparing the results of the stochastic formulation RP, to those of the *wait and see*, WS , and *expected value*, EV , problems, respectively. The well-known stochastic programming measures, the *expected value of perfect information*, $EVPI$, and the *value of the stochastic solution*, VSS (Birge, 1982; Maggioni and Wallace, 2012).

The $EVPI$ is defined as the difference between the objective values of the stochastic solution and that of the wait-and-see problem

$$EVPI := RP - WS, \quad (6.6)$$

the latter problem being defined by assuming the realizations of all the random parameters are known at the first stage.

The VSS indicates the expected gain from solving the stochastic model rather than its deterministic counterpart in which the random parameters are replaced with their expected values:

$$VSS := EEV - RP, \quad (6.7)$$

where, in our case, EEV denotes the solution value of EV , the stochastic model with the first-stage decision variables fixed at the optimal values obtained by using the expected values of the coefficients.

We present aggregated results in Table 4, grouped by instance set (Column 1) and spread class (Column 2). The table reports, the average and maximum $EVPI$ percentages (Columns 3 and 4), computed as $EVPI/RP \cdot 100$, and the average and maximum VSS percentages (Columns 5 and 6) computed as $VSS/RP \cdot 100$.

The $EVPI$ percentage is around 10%, showing the benefit of having information about the future in advance. It increases with the problem dimensions to a maximum value above 15%. The average and maximum values of VSS also increase as the size of the instance increases. The gap between the expected-value solution and the stochastic solution is significant for all sets considered. Even for small instances, the maximum VSS reaches 11%, emphasizing the potential loss incurred by following the capacity plan of the deterministic solution. The most critical item spread class is SP4, representing the case when no information is available about the category distribution of the items to be shipped. It is therefore the most representative instance class for the shipping freight over long distances, and it displays a maximum $EVPI$ of 15.20%, with a maximum VSS always greater than 7% and reaching 17.70%. We therefore conclude that using the stochastic model is appropriate and beneficial.

Set	Spread	$EVPI$ [%]	$EVPI_{max}$ [%]	VSS [%]	VSS_{max} [%]
T3	SP1	5.66	6.97	5.63	6.89
	SP2	4.35	7.08	5.28	7.66
	SP3	5.42	7.10	4.21	6.94
	SP4	6.18	6.67	8.10	11.25
T5	SP1	5.72	7.01	5.98	6.92
	SP2	6.28	7.29	4.18	6.07
	SP3	5.51	7.38	4.47	5.76
	SP4	5.35	8.10	12.34	14.17
T10	SP1	11.73	14.52	8.09	15.95
	SP2	11.52	12.62	6.66	9.37
	SP3	10.87	12.51	7.25	12.00
	SP4	11.17	15.20	9.67	17.70

Table 4: EVPI and VSS comparison

We complete this section by examining to what extent the first-stage decisions differ between the stochastic and the EV formulations. On average, the EV problem overestimates the demand to be loaded (the total volume of the items is larger than the actual volume) and the availability of extra bins (a larger set of bins is available for the recourse action). This can lead to two situations. On the one hand, EV may plan to use a larger set of bins that is not actually required, given the set of scenarios considered. The capacity plan is then more expensive (i.e., the cost increases by 17.70%), but the solution is feasible. ON the other hand, EV may plan an insufficient capacity for a subset of scenarios in which the actual availability of bins is very limited. The capacity plan is infeasible for these scenarios. This case is rare, occurring for only 2% of the instances. In both cases, the results clearly show the need to explicitly consider uncertainty in capacity-planning applications.

6.4 Solution analysis

We focus on the structure of the capacity-planning solutions and in particular on the use of bins in the plan. Table 5 reports, for each combination of instance set (Column 1) and item spread class (Column 2), the average number of bin types $N_{\mathcal{T}}$ used in the capacity plan (Column 3), the percentage of the objective function value achieved in the first stage Obj_{FS} (Column 4), and the average fill level of the bins (as a percentage of capacity) at the two stages, f_{FS} and f_{SS} respectively (Columns 5 and 6).

It is noticeable that only a few bin types are used in the plan. The number is close to 1 for T3 instances, and it is between 2 and 3 for instances in sets T5 and T10, which have five and ten bin types, respectively. Furthermore, the number of bins is not evenly

Set	Spread	$N_{\mathcal{T}}$	Obj_{FS} [%]	f_{FS} [%]	f_{SS} [%]
T3	SP1	1.4	80.59	91.94	86.28
	SP2	1.2	78.10	93.25	87.27
	SP3	1	81.91	94.47	79.47
	SP4	1.3	86.02	95.68	85.78
T5	SP1	2	77.10	95.38	87.29
	SP2	1.8	80.57	97.34	85.35
	SP3	1.5	82.00	95.64	81.46
	SP4	2.1	81.60	94.66	85.68
T10	SP1	2.5	80.99	94.76	83.32
	SP2	2.1	79.88	94.23	84.35
	SP3	2	78.90	95.58	84.62
	SP4	2.3	81.63	95.17	82.49

Table 5: Structure of solutions

distributed among the types. Almost all of the bins included in the capacity plan are of the same type; only one or two bins are of different types. Concerning the spread classes, it is interesting to note that, when the stochastic problem considers a demand with a high percentage of small items (classes SP1 and SP4), $N_{\mathcal{T}}$ is maximum. In fact, small items may be loaded into any bins (large items cannot be placed into bins with limited capacities), and they are usually used to fill near-empty bins. Thus, it becomes attractive for the model to mix bin types.

The percentage of the objective function coming from the first stage planning decisions is of the order of 80%. This indicates interest for tactical decisions to be conservative: the company should book sufficient capacity in advance to limit the adjustments necessary when the actual demand becomes known. This also indicates that 10%-20% of plan adjustment (recourse) is a good compromise between cost reduction and uncertainty management.

Finally, the fill levels of the bins selected during the first and second stages are over than 90% and 70%, respectively. This indicates the effectiveness of the capacity plan, which uses the contracted bins at almost full capacity and requires only limited adjustments at the second stage.

6.5 Sensitivity analysis

We now turn to the sensitivity of the problem and the proposed logistics capacity plan to limits on the availability of bins coming, e.g., from contractual policies imposed by the 3PL. We analyze the issue through two policies. The first policy ($P1$) imposes a

minimum number of bins for each bin type τ included in the capacity plan. The second policy (*P2*) reduces the number of available bins of each type, which requires the plan to combine different types of bins in the most efficient way.

Set	$N_{\mathcal{T}}$	Policy	m (%)	$N'_{\mathcal{T}}$	Δ'_{Obj}
T3	1.2	P1	10	1.2	0.90
		P1	20	1.2	1.46
		P1	30	1.2	2.35
		P2	50	1.2	0.83
		P2	75	1.2	0.00
		P2	90	1.2	0.00
T5	1.8	P1	10	1.8	0.18
		P1	20	1.8	0.21
		P1	30	1.8	0.34
		P2	50	2.0	0.86
		P2	75	1.8	0.00
		P2	90	1.8	0.00
T10	2.2	P1	10	2.0	0.14
		P1	20	1.8	0.25
		P1	30	1.5	0.37
		P2	50	2.2	0.00
		P2	75	2.2	0.00
		P2	90	2.2	0.00

Table 6: Sensitivity of capacity planning to bin availability

Let $u^{\tau} = \|\mathcal{J}^{\tau}\|$ be the number of bins of type τ that can be selected for the capacity plan. We define the minimum number of bins to be $m \cdot u^{\tau}$, with $m = \{10\%, 20\%, 30\%\}$. Similarly, the maximum availability is reduced to $m \cdot u^{\tau}$ with $m = \{0.50, 0.75, 0.80\}$. To impose these policies, we introduce into the formulation the additional constraints (6.8) for the first policy, and (6.9) for the second policy:

$$\sum_{j \in \mathcal{J}^{\tau}} y_j^{\tau s} \geq m \cdot u^{\tau} \quad \forall \tau \in \mathcal{T}, \quad s \in \mathcal{S} \quad (6.8)$$

$$\sum_{j \in \mathcal{J}^{\tau}} y_j^{\tau s} \leq m \cdot u^{\tau} \quad \forall \tau \in \mathcal{T}, \quad s \in \mathcal{S}. \quad (6.9)$$

We compare the results of the PH meta-heuristic for the two policies and those of the original formulation. Table 6 reports the average results of the sensitivity analysis for each set of instances (Column 1), the original number of bin types used in the plan $N_{\mathcal{T}}$ (Column 2), the policy type (Column 3), and the associated value of factor m (Column 4). For each policy, we show the changes in the solutions with respect to the number of bin types used ($N'_{\mathcal{T}}$, Column 5) and the extra cost for the policy (Δ'_{Obj} , Column 6).

We note that the instances with a limited availability of bin types (T3) are more sensitive to policy P1. For these sets, P1 penalizes combinations of bin types, causing an increase in the cost of the first-stage capacity plan or in the recourse action contracting more bins. This cost increases as the number of bin types available decreases, e.g., it is approximately 2% for instances T3. Similarly, P1 tends to penalize the use of bin types for T5 and T10, but the impact on the objective function is limited. The wide choice of bins that characterizes these instances allows them to efficiently implement the policy, reducing the cost associated with the second stage.

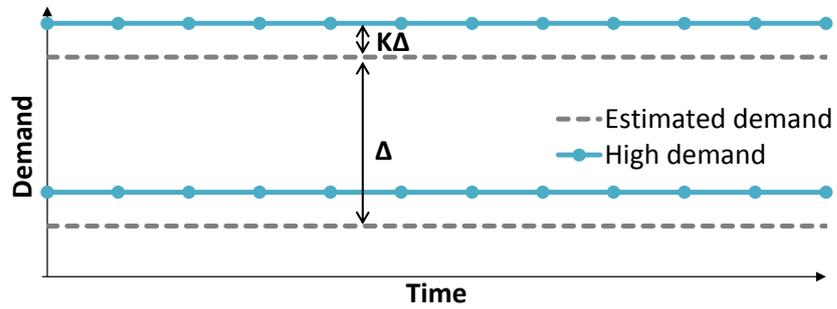
P2 reduces the set of bins available at planning time. Thus, it is necessary to combine the available bins for each type (replacing the large bins that cannot be selected with more smaller bins). The policy becomes significant only in extreme cases in which the original availability of the bins is halved (e.g., $m = 50\%$). The solutions then include more bins, resulting in a higher fixed cost. The effect of this policy is very limited, however, with an increase in objective function of 0.86%. In the worst case, when the plan does not meet the demand, the recourse selects extra capacity at a premium cost.

6.6 Long-term behavior of capacity plan

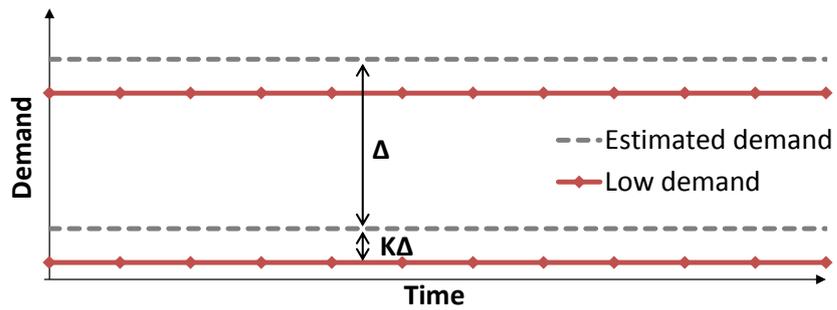
The goal of this analysis is to evaluate the robustness and reliability of the capacity planning decisions when the demand moves significantly away from the distribution and the estimation (reflected in the set of scenarios \mathcal{S}). This situation can be observed when economic conditions change significantly resulting in higher and lower demands.

We performed Monte Carlo simulations (on a 10% subset of the instances) for a relatively long period (1 year) for four demand scenarios:

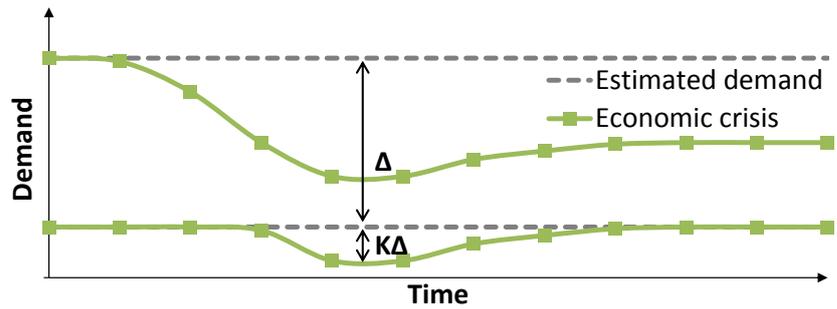
1. **High demand** (Figure 3a illustrates the range of the demand in this scenario relative to the initial one): The average demand is higher than the demand of the scenarios in \mathcal{S} . This case aims to measure the cost of overusing the spot market to obtain the extra capacity.
2. **Low demand** (Figure 3b): The average demand is lower than the demand of the scenarios in \mathcal{S} . This case focuses on the cost of not using part of the planned capacity.
3. **Economic crisis** (Figure 3c): The demand decreases rapidly to a value below the estimated demand. It then stabilizes below the average demand of the scenarios in \mathcal{S} .
4. **Economic recovery** (Figure 3d): The demand is initially below the average estimated demand, but it increases rapidly to a peak above the maximum estimation. It then stabilizes.



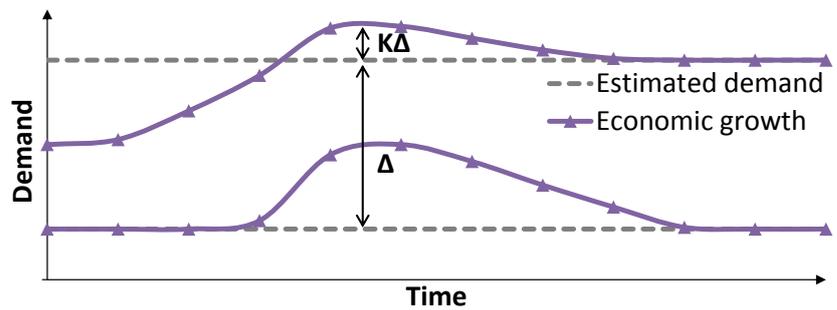
(a) Distribution of high-demand scenario



(b) Distribution of low-demand scenario



(c) Distribution of economic-crisis scenario



(d) Distribution of economic-growth scenario

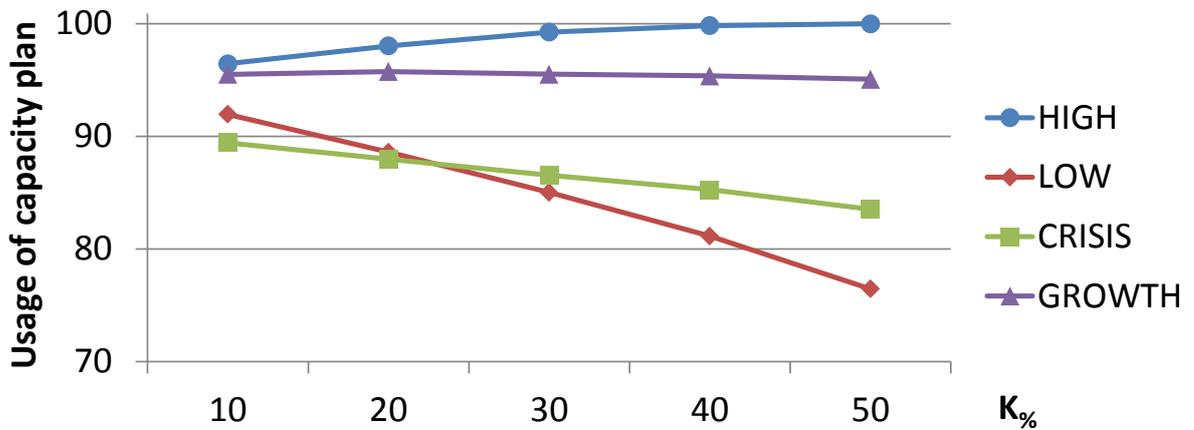
Figure 3: Demand scenarios for the Monte Carlo simulation

The overall simulation process was:

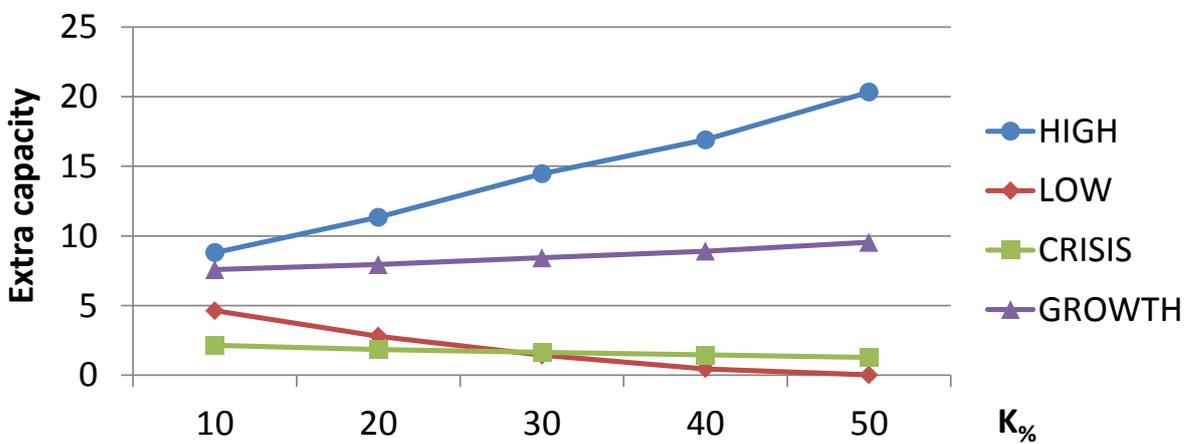
- Given an instance and the associated set of scenarios \mathcal{S} , find the capacity-planning solution using the PH meta-heuristic.
- Create a new set of scenarios \mathcal{S}' :
 - Identify the maximum and minimum demand in \mathcal{S} and compute the difference Δ between them.
 - Compute $K\Delta$ for the current demand scenario, where $K = \{10, 20, 30, 40, 50\}\%$ is an offset factor that defines how far the demand distribution has to be from that estimated in set \mathcal{S} (Figure 3).
 - Define 365 scenarios respecting the trends of the demand and the characteristics of the instance (number of bin types and spread class of items).
- For each scenario $s \in \mathcal{S}'$, build a VCSBPP with the bin set formed by the bins included in the capacity-planning solution and the bins at premium cost defined in the scenario, and the demand in terms of the items to be loaded. The resulting VCSBPP is then solved by the meta-heuristic.
- Given the solution, compute the expected value of the total cost and statistics related to the use of the planned bins and of the extra capacity required to cover higher demand, and the corresponding cost.
- Repeat 10 times and average the results.

Figure 4 shows the results of the Monte Carlo simulation for an illustrative instance from set T10 (the other instances produced similar results). The figure reports, for each demand scenario, the usage percentage of the logistics capacity plan, defined as the ratio of the number of planned (selected in the first stage) bins actually used in the solution to the total number of bins selected in the plan (Figure 4a), the percentage of extra bins in the solution defined as the ratio of the number of extra bins to the total number of bins in the solution (Figure 4b), and the percentage of the objective function value associated with the capacity plan (Figure 4c), with respect to the offset factor K .

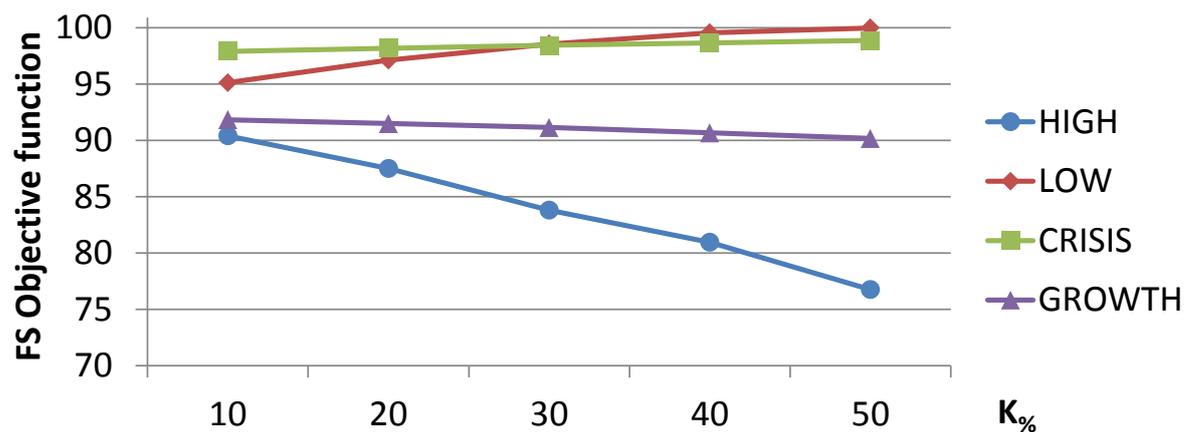
Despite the high variability of the stochastic parameters in instances T10, the capacity-planning decisions appear valid. Increasing the offset factor K increases or decreases the percentage of the planned capacity actually used according to the demand trend in different demand scenarios. Thus, for the LOW and CRISIS scenarios, the use of the planned bins decreases, to an average of 77% for T10 instances in the worst case. In contrast, for the HIGH and GROWTH scenarios, the number of planned bins actually used in the solution is always greater than 90%. The reduction in the planned capacity used in the LOW and CRISIS scenarios parallels a decrease in the need for extra capacity. This is particularly evident for T10 instances, for which no extra capacity is needed when the



(a) Percentage of planned (first-stage) bins actually used in the solution



(b) Percentage of extra bins in the solution



(c) Objective function value associated with the capacity plan

Figure 4: Monte Carlo simulation results for an illustrative T10 instance

demand is minimum (LOW scenario and $K = 50\%$). Conversely, when the demand is underestimated, it may be necessary to purchase additional capacity; more than 20% for the HIGH scenario (some 7 bins for T10 instances) and between 10% and 15% for the GROWTH scenario (some 4 bins for T10 instances). This limited use of the spot market further supports the claim of validity of the solutions found.

The limited use of the spot market is also indicated by the objective function value related to the capacity plan, corresponding to the total fixed cost of the bins that are selected in the plan (first stage). This cost increases as demand decreases in the demand scenarios. Indeed, when the demand is overestimated (scenarios LOW and CRISIS) the capacity plan is generally adequate, and only a few adjustments are required for a small set of scenarios. This means that most of the cost is due to the capacity plan. For the instances considered, the range is 75% to 100%, values that indicate the robustness of the solution even when the estimation errors are considerable.

The fill level of the planned and extra bins is always between 99% and 80% even for the LOW and CRISIS scenarios. This result emphasizes the robustness of the capacity-planning decisions that need limited and targeted adjustments, through the spot market, to meet the demand.

To conclude, the sensitivity of the methodology to the various demand scenarios considered, we note that the GROWTH scenario has a limited impact on the results with variations of a few percentage points, while the most significant impacts follow from the “extreme” scenarios LOW and HIGH. Moreover, it is interesting to note that the CRISIS and LOW scenarios exchange roles for values of K greater than 30%. For small values of K , CRISIS is the worst-case scenario because of the decrease in the demand. For high values of K , LOW has the most significant impact because of the error in the estimation of the demand, which is constant over time (for each scenario in \mathcal{S}').

7 Conclusions and Future Directions

We considered the logistics capacity planning problem arising in the context of supply-chain management. We focused, in particular, on the relation between a firm and its logistics capacity provider, and the tactical-planning problem of determining the quantity of capacity units - the bins - of different types to secure for the next period of activity, given the uncertainty on future needs in terms of demand for items to be moved or stored, and the availability and costs of capacity for these movements or storage activities.

We proposed a modeling framework for this problem introducing a new class of bin packing problems, namely the Stochastic Variable Cost and Size Bin Packing Problem (SVCSBPP). The model takes the form of a two-stage stochastic formulation with recourse that assigns to the first stage the tactical capacity-planning decisions of selecting capacity

units of various types, while the second stage models the subsequent adjustments to the plan, securing extra units of capacity, when needed, and packing the loads into the capacity units, each time the plan is applied and new information becomes known.

We also proposed an effective and accurate meta-heuristic based on progressive hedging ideas and several advanced strategies to accelerate the search and efficiently address the symmetry inherent to packing problems and strongly present in the SVCSBPP due to the presence of several equivalent bins of each type. Extensive computational results for a large set of instances support the claim of validity for the model and efficiency for the solution method proposed. The capacity-planning solutions account efficiently for the uncertainty in supply and demand, with high filling rates of the bins and relatively little utilization of capacity acquired on the spot market. Solutions are high quality, when compared to the optimum solution (when available) or a good lower bound, and robust with respect to changes in the demand distribution. The proposed stochastic model and PH-based meta-heuristic have thus been shown to deliver significant performance in terms of computational efficiency, solution economic value (cost reduction), and operations management (prediction of the capacity needed by the firm).

Future work will focus on different policies arising in contract procurement (e.g., accounting for future loss of planned capacity units during operations due, for example, to delivery delays and damage) and their impact on the methodology and the capacity-planning solutions. In the longer term, the problem will be generalized to address, e.g., the selection of a limited set of service providers among several proposing different contracts, and the network-wide logistics capacity planning issues.

Acknowledgments

Partial funding for this project was provided by the Italian University and Research Ministry under the UrbeLOG project-Smart Cities and Communities. Partial funding was also provided by the Natural Sciences and Engineering Council of Canada (NSERC) through its Discovery Grant program. We gratefully acknowledge the support of Fonds de recherche du Québec through their infrastructure grants.

References

- Aissaoui, N., Haouari, M., and Hassini, E. (2007). Supplier selection and order lot sizing modeling: A review. *Computers & Operations Research*, 34(12):3516–3540.
- Baldi, M. M., Crainic, T. G., Perboli, G., and Tadei, R. (2012). The generalized bin packing problem. *Transportation Research Part E*, 48(6):1205–1220.
- Baldi, M. M., Crainic, T. G., Perboli, G., and Tadei, R. (2014). Branch-and-price and beam search algorithms for the variable cost and size bin packing problem with optional items. *Annals of Operations Research*, 222(CIRRELT-2012-01):125–141.
- Bertazzi, L., Chan, L. M. A., and Speranza, M. G. (2007). Analysis of practical policies for a single link distribution system. *Naval Research Logistics (NRL)*, 54(5):497–509.
- Bertazzi, L. and Speranza, M. G. (2005). Worst-case analysis of the full load policy in the single link problem. *International Journal of Production Economics*, 93–94:217–224. Proceedings of the Twelfth International Symposium on Inventories.
- Bertazzi, L. and Speranza, M. G. (2012). Inventory routing problems: An introduction. *EURO Journal on Transportation and Logistics*, 1(4):307–326.
- Birge, J. (1982). The value of the stochastic solution in stochastic linear programs with fixed recourse. *Mathematical Programming*, 24(1):314–325.
- Birge, J. and Louveaux, F. (1997). *Introduction to Stochastic Programming*. Springer Series in Operations Research and Financial Engineering. Springer.
- Carpentier, P.-L., Gendreau, M., and Bastin, F. (2013). Long-term management of a hydroelectric multireservoir system under uncertainty using the progressive hedging algorithm. *Water Resources Research*, 49(5):2812–2827.
- Chen, F., Federgruen, A., and Zheng, Y.-S. (2001). Coordination mechanisms for a distribution system with one supplier and multiple retailers. *Management Science*, 47(5):693–708.
- Correia, I., Gouveia, L., and Saldanha-da-Gama, F. (2008). Solving the variable size bin packing problem with discretized formulations. *Computers & Operations Research*, 35:2103–2113.
- Crainic, T. G., Fu, X., Gendreau, M., Rei, W., and Wallace, S. W. (2011a). Progressive hedging-based metaheuristics for stochastic network design. *Networks*, 58(2):114–124.
- Crainic, T. G., Gobbato, L., Perboli, G., Rei, W., Watson, J. P., and Woodruff, D. L. (2013a). Bin packing problems with uncertainty on item characteristics: An application to capacity planning in logistics. Technical Report CIRRELT-2013-31, CIRRELT.

- Crainic, T. G., Gobbato, L., Perboli, G., Rei, W., Watson, J. P., and Woodruff, D. L. (2014a). 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., Hewitt, M., and Rei, W. (2014b). Scenario grouping in a progressive hedging-based meta-heuristic for stochastic network design. *Computers & Operations Research*, 43:90–99.
- 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. (2013b). Proactive order consolidation in global sourcing. In Bookbinder, J. H., editor, *Handbook of Global Logistics*, volume 181 of *International Series in Operations Research & Management Science*, pages 501–530. Springer New York.
- Crainic, T. G., Perboli, G., Pezzuto, M., and Tadei, R. (2007). New bin packing fast lower bounds. *Computers & Operations Research*, 34:3439–3457.
- Crainic, T. G., Perboli, G., Rei, W., and Tadei, R. (2011b). Efficient lower bounds and heuristics for the variable cost and size bin packing problem. *Computers & Operations Research*, 38:1474–1482.
- Crainic, T. G., Tadei, R., Perboli, G., and Baldi, M. M. (2012). The generalized bin packing problem: Models and bounds. In *Odyssey 2012, the 5th International Workshop on Freight Transportation and Logistics, Mykonos (Greece), May 21–25, 2012*.
- DAUIN (2014). HPC cluster at Politecnico di Torino, <http://www.dauin-hpc.polito.it>.
- Ford, D. J. (2001). Inbound freight consolidation: A simulation model to evaluate consolidation rules. Master’s thesis, Massachusetts Institute of Technology.
- Higle, J. and Sen, S. (1996). *Stochastic Decomposition: A Statistical Method for Large Scale Stochastic Linear Programming*. Nonconvex Optimization and Its Applications. Springer.
- ILOG Inc. (2012). *IBM ILOG CPLEX v12.1 User’s Manual*.
- Kaut, M., Vladimirov, H., Wallace, S., and Zenios, S. (2007). Stability analysis of portfolio management with conditional value-at-risk. *Quantitative Finance*, 7(4):397–409.
- Maggioni, F. and Wallace, S. (2012). Analyzing the quality of the expected value solution in stochastic programming. *Annals of Operations Research*, 200:37–54.

- Marasco, A. (2008). Third-party logistics: A literature review. *International Journal of Production Economics*, 113(1):127–147. Research and Applications in E-Commerce and Third-Party Logistics Management Special Section on Meta-standards in Operations Management: Cross-disciplinary perspectives.
- Mayer, J. (1998). *Stochastic Linear Programming Algorithms: A Comparison Based on a Model Management System*. Optimization Theory and Applications. Taylor & Francis.
- Monaci, M. (2002). *Algorithms for packing and scheduling problems*. PhD thesis, Università di Bologna, Bologna, Italy.
- Monczka, R., Handfield, R., Giunipero, L., and Patterson, J. (2008). *Purchasing and Supply Chain Management*. Cengage Learning.
- Perboli, G., Gobbato, L., and Perfetti, F. (2014a). Packing problems in transportation and supply chain: New problems and trends. *Procedia - Social and Behavioral Sciences*, 111:672–681.
- Perboli, G., Tadei, R., and Baldi, M. (2012). The stochastic generalized bin packing problem. *Discrete Applied Mathematics*, 160:1291–1297.
- Perboli, G., Tadei, R., and Gobbato, L. (2014b). The multi-handler knapsack problem under uncertainty. *European Journal of Operational Research*, 236(3):1000–1007.
- Rizk, N., Martel, A., and D’Amours, S. (2008). Synchronized production-distribution planning in a single-plant multi-destination network. *Journal of the Operational Research Society*, 59(1):90–104.
- Rockafellar, R. T. and Wets, R. J.-B. (1991). Scenarios and policy aggregation in optimization under uncertainty. *Math. Oper. Res.*, 16(1):119–147.
- Rosa, C. H. and Ruszczyński, A. (1996). On augmented Lagrangian decomposition methods for multistage stochastic programs. *Computers & Operations Research*, 64(1):289–309.
- Seiden, S. (2002). On the online bin packing problem. *Journal of the ACM*, 49(5):640–671.
- Veliz, F., Watson, J.-P., Weintraub, A., Wets, R.-B., and Woodruff, D. (2014). Stochastic optimization models in forest planning: A progressive hedging solution approach. *Annals of Operations Research*, pages 1–16.
- Watson, J.-P. and Woodruff, D. (2011). Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. *Computational Management Science*, 8(4):355–370.