

Logistics capacity planning: A stochastic bin packing formulation and a progressive hedging meta-heuristic

*Original*

Logistics capacity planning: A stochastic bin packing formulation and a progressive hedging meta-heuristic / Crainic, Teodor Gabriel; Gobbato, Luca; Perboli, Guido; Rei, Walter. - In: EUROPEAN JOURNAL OF OPERATIONAL RESEARCH. - ISSN 0377-2217. - STAMPA. - 253:2(2016), pp. 404-417. [10.1016/j.ejor.2016.02.040]

*Availability:*

This version is available at: 11583/2639867 since: 2017-12-04T09:51:11Z

*Publisher:*

Elsevier

*Published*

DOI:10.1016/j.ejor.2016.02.040

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)

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

**Teodor Gabriel Crainic**

Department of Management and Technology,  
School of Management, UQAM and CIRRELT, Montreal, Canada

**Luca Gobbato**

Department of Control and Computer Engineering,  
Politecnico di Torino, Turin, Italy

**Guido Perboli**

Department of Control and Computer Engineering,  
Politecnico di Torino, Turin, Italy

**Walter Rei**

Department of Management and Technology,  
School of Management, UQAM and CIRRELT, Montreal, Canada

January 20, 2016

## Abstract

We consider the logistics capacity planning problem arising in the context of supply-chain management. We address 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 introducing a new class of bin packing problems, the Stochastic Variable Cost and Size Bin Packing Problem. 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 and packing the items into the selected bins, performed each time the plan is applied and new information becomes known. We propose a new meta-heuristic based on progressive hedging ideas that includes 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

# 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 demand as well as on the availability of the means providing the required capacity.

Bin packing models offer important decision-support tools for transportation and logistics (Perboli et al., 2014a), and appear appropriate for the problem at hand. In bin packing problems, *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 dimensions and costs generating a rich set of problem settings and formulations (e.g., Crainic et al., 2011b; Baldi et al., 2012). The packing formulations currently proposed in the literature deal only partially, however, with the requirements of capacity planning. Most research focuses on operational issues, and very few address packing and uncertainty.

We aim to contribute to fill this gap and introduce the *Stochastic Variable Cost and Size Bin Packing Problem (SVCSBPP)*, a new class of stochastic bin packing formulations providing a modeling framework for the tactical capacity-planning problem with uncer-

tain demand, item characteristics, availability of extra bins in the future as well as their actual costs and dimensions. We propose a two-stage integer stochastic programming formulation. The first stage concerns the selection of bins, while the second stage concerns the acquisition of extra capacity (bins) and the assignment of loads (items) to bins selected in the first and second stages, once the actual demand information, i.e., number of characteristics of items, is revealed. To address realistically-sized instances, we also propose an efficient meta-heuristic based on Progressive Hedging principles (Rockafellar and Wets, 1991) for the *SVCSBPP*.

The contributions of this paper are:

1. A modeling framework for securing logistics capacity for a tactical-planning horizon, while explicitly considering uncertainty in both supply, i.e., the future availability and cost of various bin types, and demand as the number and dimension of items. This framework takes the form of a new class of stochastic bin packing problems, the *SVCSBPP*, which generalizes previous work (Crainic et al., 2014a) through a) the possibility that bins available in the future be different in cost and dimensions from those planned, and b) the explicit representation of the uncertainty on the future availability of bins in numbers and costs. This generalization contributes greatly to the practical relevance of the methodology, through the modeling of the spot market. It introduces, in particular, the possibility that no feasible solution exists due to insufficient availability of extra bins in the future, requiring additional care in formulating the two-stage model. Different from previous models, this generalization leads to a two-stage formulation that includes integer decision variables in the second stage to represent the selection of additional bins.
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. A particularly 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 significantly impairs the efficiency of a meta-heuristic implementing classical PH features. Several novel strategies are thus proposed, in particular a new definition of "consensus", the associated temporary global solution, and the heuristic penalty adjustments, together with bundle variable fixing.
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-based 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 briefly reviews the relevant literature. Section 4 presents the formulation of the problem. A lower bound and the meta-heuristic solution method are introduced in Sections 5 and 6, respectively. Section 7 presents the experimental plan and analyzes the computational results. Section 8 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 *supply capacity*, which 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, is then 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 actual *demand* to be serviced at each occurrence of the plan.

Demand is defined as a number of items of variable sizes to be packed into the available bins for shipment or storage. This disaggregated demand representation makes the problem harder to tackle, but contributes to its relevance as it models the fact that item-to-bin assignment decisions have to be taken to ensure the feasibility of the loading/shipment plan (i.e., the technological restriction due to the bin capacity). Then, variations in demand occurring at each application of the plan generally yield numbers and sizes of items that differ from the estimations used during negotiations, requiring ex-

tra capacity (additional bins) to be secured when the planned capacity is not sufficient. The additional capacity is to be obtained through the spot market, from the original 3PL or other sources, and it generally comes at a much higher cost than the fare negotiated initially. Moreover, while the 3PL generally ensures the planned capacity, the extra capacity may not be available when needed, and “penalty” costs are incurred due to either the corresponding shipping delay or the decision to ship using the more expensive less-than-container-load (*LCL*) mode. The availability and cost of the bins providing the extra capacity is thus 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 item demand and the spot-market bin availability and costs become known.

### 3 Literature Review

We present a brief review of the packing and stochastic programming literature directly related to the problem at hand. Notice that, generally, the packing literature focuses on operations, and that only recently has research started to address tactical and strategic problems (Perboli et al., 2014a). Research on stochastic packing is even less present in the literature.

The problem we address belongs to a class of rich packing problems motivated by the planning and management of modern logistics and transportation systems, and providing the means to consider additional loading constraints in multi-dimensional problems (Crainic et al., 2012) and the presence of multiple attributes (Crainic et al., 2011b; Baldi et al., 2012). The *VCSBPP* (Crainic et al., 2011b) generalized the classical Bin Packing problem (Martello and Toth, 1990) by considering unrestricted variations in the costs and dimension of both items and bins. Baldi et al. (2012) later proposed the *Generalized Bin Packing Problem*, encompassing the *VCSBPP* and the simple and multiple knapsack problems.

All the above formulations are deterministic, however. Traditionally, the literature on stochastic packing focused on the so-called on-line packing problem, where the uncertainty affects the arrival time of items (Seiden, 2002). Only recently were contributions proposed addressing packing problem settings found in strategic and tactical applications.

The *Multi-Handler Knapsack Problem under Uncertainty* (Perboli et al., 2014b) is a stochastic version of the knapsack problem, where the item profit depends on the handler managing. The authors proposed a two-stage formulation and a deterministic approximation based on multinomial Logit model. A similar approach was used for

the *Stochastic Generalized Bin Packing Problem* (Perboli et al., 2012), which considered profit oscillations due to the handling operations associated to bin loading.

Crainic et al. (2014a) proposed a stochastic extension of the VCSBPP (Crainic et al., 2011b) considering variations in the renting cost of the different bins, and in the size and number of items to be loaded. A state-of-the-art commercial solver was used to solve 360 instances. Results underlined the importance of the stochastic parameters on the solution structure and the limitation of this approach when the number of variables involved in the two-stage stochastic model increases. The *Stochastic Variable Cost and Size Bin Packing* problem defined in this paper generalizes the previous work. We introduce, in particular, the explicit representation of the uncertainty on the availability (existence and number) of bins of various characteristics.

The problems reviewed above were formulated as two-stage integer stochastic programs with recourse (Birge and Louveaux, 1997). We propose in Section 4 such a formulation for the SVCSBPP, where the first stage concerns the selection of bins, and the second (integer) stage concerns the acquisition of extra bins and the assignment of items to available bins, once the demand information is revealed.

Two-stage integer stochastic formulations separate 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) have shown that a state-of-the-art mix-integer programming solver struggles to solve even small instances of stochastic VCSBPP formulations. Meta-heuristics are thus required to successfully address larger problems.

The *Progressive Hedging (PH)* algorithm, initially proposed by Rockafellar and Wets (1991) for stochastic convex programs, has emerged as an effective meta-heuristic for combinatorial stochastic optimization models (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, as subproblems are repeatedly solved. We enhance the basic PH algorithm with a number of features to exploit the structure of the SVCSBPP. This very efficient meta-heuristic is

detailed in Section 6.

## 4 The SVCSBPP 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  $V^\tau$  and  $f^\tau$  respectively be the volume and fixed cost associated with bins of type  $\tau \in \mathcal{T}$ . We define  $\mathcal{J}^\tau$  to be the set of available bins of type  $\tau$  and  $\mathcal{J} = \cup_\tau \mathcal{J}^\tau$  to be the set of available bins at the first stage.

Let set  $\Omega$  be the sample space of the random event, where  $\omega \in \Omega$  defines a particular realization. Let vector  $\xi$  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

$$\min_y \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} f^\tau y_j^\tau + \mathbb{E}_\xi [Q(y, \xi(\omega))] \quad (4.1)$$

$$\text{s.t. } y_j^\tau \in \{0, 1\}, \quad \forall \tau \in \mathcal{T}, j \in \mathcal{J}^\tau, \quad (4.2)$$

where  $\mathbb{E}_\xi$  represents the expected value of  $Q(y, \xi(\omega))$ , the cost paid to adjust the plan at the second stage once information on demand and available supply becomes available, given the tactical capacity plan  $y$  and the vector  $\xi(\omega)$ . The objective function (4.1) then minimizes the sum of the total fixed cost of the tactical capacity plan and the expected cost associated with the plan adjustment, while constraints (4.2) impose the integrality requirements on  $y$ .

We consider several stochastic parameters in  $\xi(\omega)$ , in formulating  $Q(y, \xi(\omega))$ , representing the information revealed at the second stage that conditions the possible plan-adjustment decisions. On the demand side, let  $\mathcal{I}(\omega)$  be the set of items to be packed, with  $v_i(\omega)$ ,  $i \in \mathcal{I}(\omega)$ , the associated stochastic item volumes. Two actions have to be considered on the supply side to represent the two situations regarding the availability of extra capacity. The first concerns the possibility to acquire extra bins when needed, where  $\mathcal{K}^\tau(\omega)$  stands for the set of available bins of type  $\tau$  at the second stage,  $g^\tau(\omega)$ ,  $k \in \mathcal{K}^\tau(\omega)$  represent the associated stochastic costs, and  $\mathcal{K}(\omega) = \cup_\tau \mathcal{K}^\tau(\omega)$ . The second focuses on the case when the extra bins available in the second stage are insufficient to pack all the items, and additional capacity has to be secured by delaying the shipment or using a LCL mode. We model this situation through a set of virtual bins  $\bar{\mathcal{K}}(\mathcal{I}(\omega))$ , with  $|\bar{\mathcal{K}}(\mathcal{I}(\omega))| = |\mathcal{I}(\omega)|$ , where each bin  $k \in \bar{\mathcal{K}}(\mathcal{I}(\omega))$  corresponds to item  $k \in \mathcal{I}(\omega)$  and may be used to load it, i.e., its volume  $V^k = v_k(\omega)$ ,  $k \in \mathcal{I}(\omega)$ , at a cost  $g^k(\omega)$ ,  $\bar{\mathcal{K}}(\mathcal{I}(\omega))$ .

The second-stage variables then are  $z_k^\tau(\omega) = 1$  if bin  $k \in \mathcal{K}^\tau(\omega) \cup \bar{\mathcal{K}}(\mathcal{I}(\omega))$  is selected, 0 otherwise;  $x_{ij}(\omega) = 1$  if item  $i \in \mathcal{I}(\omega)$  is packed into bin  $j \in \mathcal{J}$ , 0 otherwise; and

$x_{ik}(\omega) = 1$  if item  $i \in \mathcal{I}(\omega)$  is packed into bin  $k \in \mathcal{K}(\omega) \cup \bar{\mathcal{K}}(\mathcal{I}(\omega))$ , 0 otherwise. The  $Q(y, \xi(\omega))$  formulation then becomes

$$Q(y, \xi(\omega)) = \min_{y, z, x} \sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}^\tau(\omega) \cup \bar{\mathcal{K}}(\mathcal{I}(\omega))} g^\tau(\omega) z_k^\tau(\omega) \quad (4.3)$$

$$\text{s.t.} \quad \sum_{j \in \mathcal{J}} x_{ij}(\omega) + \sum_{k \in \mathcal{K}(\omega) \cup \bar{\mathcal{K}}(\mathcal{I}(\omega))} x_{ik}(\omega) = 1, \quad \forall i \in \mathcal{I}(\omega), \quad (4.4)$$

$$\sum_{i \in \mathcal{I}(\omega)} v_i(\omega) x_{ij}(\omega) \leq V^\tau y_j^\tau, \quad \forall \tau \in \mathcal{T}, j \in \mathcal{J}, \quad (4.5)$$

$$\sum_{i \in \mathcal{I}(\omega)} v_i(\omega) x_{ik}(\omega) \leq V^\tau z_k^\tau(\omega), \quad \forall \tau \in \mathcal{T}, k \in \mathcal{K}^\tau(\omega), \quad (4.6)$$

$$x_{ij}(\omega) \in \{0, 1\}, \quad \forall i \in \mathcal{I}(\omega), j \in \mathcal{J}, \quad (4.7)$$

$$x_{ik}(\omega) \in \{0, 1\}, \quad \forall i \in \mathcal{I}(\omega), k \in \mathcal{K}^\tau(\omega) \cup \bar{\mathcal{K}}(\mathcal{I}(\omega)), \quad (4.8)$$

$$z_k^\tau(\omega) \in \{0, 1\}, \quad \forall \tau \in \mathcal{T}, k \in \mathcal{K}^\tau(\omega) \cup \bar{\mathcal{K}}(\mathcal{I}(\omega)). \quad (4.9)$$

The objective function (4.3) minimizes the cost associated with the extra bins, actual and virtual, selected at the second stage. Constraints (4.4) ensure that each item is packed into a single bin. Constraints (4.5) and (4.6) ensure that the total volume of items packed into each bin does not exceed the bin volume (such constraints are not required for the virtual bins as each holds exactly the corresponding item). Constraints (4.7) to (4.9) impose the integrality requirements on all second-stage variables.

## 5 A Lower Bound for the SVCSBPP

We present in this section a *lower bound* for the SVCSBPP formulation (4.1)–(4.9), which provides a way to measure the quality of the heuristic proposed in Section 6.

The lower bound is obtained by removing the item-to-bin assignment constraints (4.4), dropping the LCL option when an insufficient number of bins are available in the second stage, and aggregating the resulting individual bin feasibility constraints (4.5) and (4.6). The resulting *LB* formulation, (5.1)–(5.4), is a two-stage stochastic model with fixed recourse, which yields an optimal set of bins that includes those in the first-stage capacity plan and the second-stage extra bins, with a total capacity sufficient for the items considered:

$$\min_{y,z} \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} f_j^\tau y_j^\tau + \mathbb{E}_\xi \left[ \sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}^\tau(\omega)} g_k^\tau(\omega) z_k^\tau(\omega) \right] \quad (5.1)$$

$$\text{s.t.} \quad \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^\tau} V_j^\tau y_j^\tau + \sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}^\tau(\omega)} V_k^\tau z_k^\tau(\omega) \geq \sum_{i \in \mathcal{I}(\omega)} v_i(\omega), \forall \omega \in \Omega, \quad (5.2)$$

$$y_j^\tau \in \{0, 1\} \quad \forall \tau \in \mathcal{T}, j \in \mathcal{J}^\tau, \quad (5.3)$$

$$z_k^\tau(\omega) \in \{0, 1\} \quad \forall \tau \in \mathcal{T}, k \in \mathcal{K}^\tau(\omega), \forall \omega \in \Omega. \quad (5.4)$$

Note that, an optimal solution for LB does not guarantee feasibility for the problem (4.1)–(4.9) with respect to the loading of items into individual bins, even when sufficient capacity is available in the second stage. On the other hand, computing the lower bound for a given realization of the random event will signal, by returning an “unfeasible” solution, that insufficient bin capacity is available in the second stage and the penalty related to the LCL option will have to be paid.

Note also 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 the optimal solution of LB in reasonable computing times (e.g., by using commercial MIP solvers). Moreover, computational results show (Section 7) that the lower bound obtained by solving the LB formulation is far better (tighter) than the solutions obtained by solving the continuous relaxation of the model (4.1)–(4.9).

## 6 Progressive Hedging-based Meta-heuristic

We propose a meta-heuristic inspired by the PH algorithm (Rockafellar and Wets, 1991) for the SVCSBPP formulation described in the previous section. The main algorithmic structure is known (e.g., Crainic et al., 2011a; Perboli et al., 2015). We enhanced it significantly, however, in a number of important algorithmic components, including the treatment of symmetry in the search for consensus, a bundle-fixing mechanisms to guide the search, a second phase to identify the final solution when consensus is not achieved during the initial search, and a parallel implementation of the algorithm solving the subproblems concurrently on multiple processors. We focus on these enhancements in the rest of the section, including structure of the parallel version of the meta-heuristic (Section 6.3).

As summarized in Algorithm 1, the proposed (sequential) meta-heuristic for the VCSBP first applies scenario decomposition, then proceeds in two phases. Phase 1 aims to induce consensus among the scenario subproblems, while Phase 2 computes the final solution when consensus is not reached within the prescribed computational conditions.

Sampling is first 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. Let  $p_s$  be the probability of each scenario  $s \in \mathcal{S}$ , and let the superscript  $s$  associate the decision variables of the previous section to each particular scenario  $s \in \mathcal{S}$ . Define  $\mathcal{I}^s$ , the set of items to pack under scenario  $s \in \mathcal{S}$ , and  $\mathcal{K}^s = \cup_{\tau} \mathcal{K}^{\tau s} \cup \bar{\mathcal{K}}(\mathcal{I}^s)$ , where  $\mathcal{K}^{\tau s}$  and  $\bar{\mathcal{K}}(\mathcal{I}^s)$  are the sets of extra bins of type  $\tau \in \mathcal{T}$  and of virtual bins corresponding to items to pack in scenario  $s \in \mathcal{S}$ , respectively. The SVCSBPP formulation (4.1)–(4.9) can then be approximated by the following deterministic model:

$$\min_{y,z,x} \sum_{s \in \mathcal{S}} p_s \left[ \sum_{\tau \in \mathcal{T}} \sum_{j \in \mathcal{J}^{\tau}} f^{\tau} y_j^{\tau s} + \sum_{\tau \in \mathcal{T}} \sum_{k \in \mathcal{K}^{\tau s}} g^{\tau s} z_k^{\tau s} \right] \quad (6.1)$$

$$\text{s.t.} \quad \sum_{j \in \mathcal{J}} x_{ij}^s + \sum_{k \in \mathcal{K}^s} x_{ik}^s = 1 \quad \forall i \in \mathcal{I}^s, s \in \mathcal{S}, \quad (6.2)$$

$$\sum_{i \in \mathcal{I}^s} v_i^s x_{ij}^s \leq V^{\tau} y_j^{\tau s} \quad \forall \tau \in \mathcal{T}, j \in \mathcal{J}^{\tau}, s \in \mathcal{S}, \quad (6.3)$$

$$\sum_{i \in \mathcal{I}^s} v_i^s x_{ik}^s \leq V^{\tau} z_k^{\tau s} \quad \forall \tau \in \mathcal{T}, k \in \mathcal{K}^{\tau s}, s \in \mathcal{S}, \quad (6.4)$$

$$y_j^{\tau s} = y_j^{\tau t} \quad \forall \tau \in \mathcal{T}, j \in \mathcal{J}^{\tau}, s, t \in \mathcal{S}, \quad (6.5)$$

$$y_j^{\tau s} \in \{0, 1\} \quad \forall \tau \in \mathcal{T}, j \in \mathcal{J}^{\tau}, s \in \mathcal{S}, \quad (6.6)$$

$$z_k^{\tau s} \in \{0, 1\} \quad \forall \tau \in \mathcal{T}, k \in \mathcal{K}^{\tau s}, s \in \mathcal{S}, \quad (6.7)$$

$$x_{ij}^s \in \{0, 1\} \quad \forall i \in \mathcal{I}^s, j \in \mathcal{J}, s \in \mathcal{S}, \quad (6.8)$$

$$x_{ik}^s \in \{0, 1\} \quad \forall i \in \mathcal{I}^s, k \in \mathcal{K}^s, s \in \mathcal{S}. \quad (6.9)$$

Constraints (6.5) are referred to as the nonanticipativity constraints, forcing all scenarios to yield the same first-stage decisions. They link the first- and second-stage variables and make the model not separable. 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. Then, constraints (6.10) - (6.11) are equivalent to (6.5).

$$\bar{y}_j^{\tau} = y_j^{\tau s} \quad \tau \in \mathcal{T}, \quad j \in \mathcal{J}^{\tau}, \quad s \in \mathcal{S}, \quad (6.10)$$

$$\bar{y}_j^{\tau} \in \{0, 1\} \quad \tau \in \mathcal{T}, \quad j \in \mathcal{J}^{\tau}. \quad (6.11)$$

The model is now separable and, following the decomposition scheme proposed by Rockafellar and Wets (1991), we relax the nonanticipativity constraints (6.10) using an augmented Lagrangian strategy with Lagrangian multipliers  $\lambda_j^s$ ,  $\forall j \in \mathcal{J}$  and  $\forall s \in \mathcal{S}$ , for the relaxed constraints, and a penalty ratio  $\rho_j^{\tau}$  associated with bin  $j \in \mathcal{J}^{\tau}$  of type  $\tau \in \mathcal{T}$ . This decomposition yields  $|\mathcal{S}|$  deterministic VCSBPP scenario subproblems with

modified fixed costs:

$$\min_{y,x} \sum_{\tau \in \mathcal{T}} \sum_{b \in \mathcal{B}^{\tau s}} f_b^\tau y_b^\tau \quad (6.12)$$

$$\text{s.t.} \quad \sum_{\tau \in \mathcal{T}} \sum_{b \in \mathcal{B}^{\tau s}} x_{ib}^s = 1 \quad \forall i \in \mathcal{I}^s, s \in \mathcal{S}, \quad (6.13)$$

$$\sum_{i \in \mathcal{I}^s} v_i^s x_{ib}^s \leq V^\tau y_b^\tau \quad \forall \tau \in \mathcal{T}, \forall b \in \mathcal{B}^{\tau s}, s \in \mathcal{S}, \quad (6.14)$$

$$y_b^\tau \in \{0, 1\} \quad \forall \tau \in \mathcal{T}, \forall b \in \mathcal{B}^{\tau s}, s \in \mathcal{S}, \quad (6.15)$$

$$x_{ib}^s \in \{0, 1\} \quad \forall \tau \in \mathcal{T}, \forall b \in \mathcal{B}^{\tau s}, \forall i \in \mathcal{I}^s, s \in \mathcal{S}, \quad (6.16)$$

where  $\mathcal{B}^{\tau s} = \mathcal{J}^\tau \cup \mathcal{K}^{\tau s}$  is the set of available bins of type  $\tau$  in the subproblem of scenario  $s \in \mathcal{S}$ , the associated decision variable  $y_b^\tau = 1$  if bin  $b \in \mathcal{B}^{\tau s}$  of type  $\tau \in \mathcal{T}$  is selected, 0 otherwise, and  $f_b^\tau$  defines the fixed cost associated with bin  $b \in \mathcal{B}^{\tau s}$

$$f_b^\tau = \begin{cases} f^\tau + \lambda_j^s - \rho_j^\tau \bar{y}_j^\tau + \frac{\rho_j^\tau}{2} & \tau \in \mathcal{T}, b, j \in \mathcal{J}^\tau, \\ g^{\tau s} & \tau \in \mathcal{T}, b \in \mathcal{K}^{\tau s}, \end{cases} \quad (6.17)$$

the values of the multipliers  $\lambda_j^s$  and  $\rho_j^\tau$  being determined at each iteration exogenously with respect to the subproblem. Hence, the Lagrangian multipliers are used to penalize a lack of implementability due to differences in the first-stage variable values among scenario subproblems.

## 6.1 Phase 1 - Searching for consensus among subproblems

Phase 1 aims to obtain ‘‘consensus’’ among the subproblems, defined as scenario solutions being similar with regard to the first-stage decisions with the overall capacity plan and, thus, being similar among themselves.

At each iteration, the subproblems are first addressed separately by applying a heuristic. Indeed, 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), and 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.

**Compute temporary global solution.** The subproblem solutions are used to build a temporary global solution, the overall capacity plan. We developed a particular strategy

for this step of the meta-heuristic to address the symmetry challenge of the bin packing formulation.

Let  $\nu$  be the iteration counter in the PH meta-heuristic, and  $y_j^{\tau s\nu}$ ,  $\forall s \in \mathcal{S}$ ,  $\forall \tau \in \mathcal{T}$ , and  $\forall j \in \mathcal{J}^\tau$ , the local solutions of the scenario subproblems. Two temporary global solutions are computed. The first is based on the usage of the individual bins and computation of the classical expectation yielding

$$\bar{y}_j^{\tau\nu} = \sum_{s \in \mathcal{S}} p_s y_j^{\tau s\nu}, \quad \forall \tau \in \mathcal{T}, \forall j \in \mathcal{J}^\tau. \quad (6.18)$$

This definition does not take into account the strong symmetry in the solution space typical of packing problems (Baldi et al., 2012), where two solutions are considered symmetric (and equivalent) when they involve the same set of bins in different orders. As Equation (6.18) considers consensus with respect to the use of each specific bin  $j \in \mathcal{J}^\tau$ , the computation of  $y_j^{\tau s\nu}$  is strongly dependent on the order of the bins in the solution.

To address this issue, we define a second 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  $\nu$ . We then define the overall capacity plan for each bin type  $\tau \in \mathcal{T}$  as

$$\bar{\delta}^{\tau\nu} = \sum_{s \in \mathcal{S}} p_s \delta^{\tau s\nu} = \sum_{s \in \mathcal{S}} p_s \sum_{j \in \mathcal{J}^\tau} y_j^{\tau s\nu} = \sum_{j \in \mathcal{J}^\tau} \sum_{s \in \mathcal{S}} p_s y_j^{\tau s\nu} = \sum_{j \in \mathcal{J}^\tau} \bar{y}_j^{\tau\nu}. \quad (6.19)$$

Equation (6.19) 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  $\bar{\delta}^{\tau\nu}$ .

The two measures yielded by Equations (6.18) and (6.19) are used as reference solutions for penalty adjustments with the goal of helping the search identify bins for which consensus is possible. Equation ((6.19)) is also used in the bounding strategy described below.

**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 (Rockafellar and Wets, 1991) operates on the fixed costs by changing the Lagrangian multipliers using information on the individual bin selection (variables  $y_j^{\tau s\nu}$ ). At each iteration  $\nu$ , we update the values of the multipliers  $\lambda_j^{s\nu}$  and  $\rho_j^{\tau\nu}$ ,  $\forall j \in \mathcal{J}$ ,  $\forall s \in \mathcal{S}$ , (defining variable-specific  $\rho_j^{\tau\nu}$  multipliers enhances the convergence behavior of

the method, Watson and Woodruff, 2011):

$$\lambda_j^{s\nu} = \lambda_j^{s(\nu-1)} + \rho_j^{\tau(\nu-1)}(y_j^{\tau s\nu} - \bar{y}_j^{\tau\nu}) \quad (6.20)$$

$$\rho_j^{\tau\nu} \leftarrow \alpha \rho_j^{\tau(\nu-1)}, \quad (6.21)$$

where  $\lambda_j^{s0} = 0$ ,  $s \in \mathcal{S}$ ,  $\alpha > 1$  is a given constant, and  $\rho_j^{\tau0} = \max(1, f^\tau/10)$ ,  $\forall j \in \mathcal{J}^\tau$ ,  $\forall \tau \in \mathcal{T}$ , is set proportional to the fixed cost associated with the bin-selection variable, ensuring  $\rho_j^{\tau\nu} \rightarrow \infty$  as the number of iterations  $\nu$  increases.

When the overall solution is close to consensus, the small differences between the subproblem solutions and the overall solution often make the adjustments (6.20) and (6.21) less effective, resulting in an increase in the number of iterations needed to reach consensus. The heuristic strategy of the second penalty adjustment aims to address this issue and accelerate the search by directly tuning the fixed costs of bins of the same type. Thus, when at least  $\sigma\%$  of the variables have reached consensus, we perturb every subproblem  $s \in \mathcal{S}$  by changing  $f^{\tau s\nu}$ , the fixed cost of bin  $j \in \mathcal{J}^\tau$  of type  $\tau \in \mathcal{T}$  for scenario  $s \in \mathcal{S}$  at iteration  $\nu$ :

$$f^{\tau s\nu} = \begin{cases} f^{\tau s(\nu-1)} \cdot M & \text{if } \delta^{\tau s(\nu-1)} > \bar{\delta}^{\tau(\nu-1)} \\ f^{\tau s(\nu-1)} \cdot \frac{1}{M} & \text{if } \delta^{\tau s(\nu-1)} < \bar{\delta}^{\tau(\nu-1)} \\ f^{\tau s(\nu-1)} & \text{otherwise,} \end{cases} \quad (6.22)$$

where,  $f^{\tau s0} = f^\tau$ ,  $M$  is a constant greater than 1, and  $\sigma\%$  is a constant such that  $0.5 \leq \sigma\% \leq 1$  ( $\sigma\% = 0.75$  and  $M = 1.1$  in the current implementation). We thus discourage (penalize their selection costs) the use of bins of type  $\tau$  in scenario  $s$  at iteration  $\nu$  when, at the previous iteration,  $\delta^{\tau s(\nu-1)}$ , the total number of bins of that type, was larger than the number of bins of the same type in the corresponding reference solution  $\bar{\delta}^{\tau(\nu-1)}$ . Symmetrically, the use of bins of type  $\tau$  is promoted (selection fixed cost reduction) in the opposite case.

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

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

$$\bar{\delta}_m^{\tau\nu} \leftarrow \min_{s \in \mathcal{S}} \delta^{\tau s\nu}, \quad (6.23)$$

$$\bar{\delta}_M^{\tau\nu} \leftarrow \max_{s \in \mathcal{S}} \delta^{\tau s\nu}. \quad (6.24)$$

The bundle strategy applies the 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  $\tau$  available in the next iteration. This is performed by setting decision variables  $y_j^{\tau s(\nu+1)} = 0$  for  $j = \bar{\delta}_M^{\tau\nu} + 1, \dots, \|\mathcal{J}^\tau\|$ .

## 6.2 Termination criteria and Phase 2

We implement two stopping criteria for the search phase of the proposed meta-heuristic, based on the level of consensus reached and the number of iterations.

Recall that, the level of consensus is measured through equations (6.23) and (6.24), and 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 stop the search, and proceed to Phase 2, either when consensus is achieved for all bin types except one, type  $\tau'$  for which  $\bar{\delta}_m^{\tau'} < \bar{\delta}_M^{\tau'}$ , or when consensus is not achieved within a given maximum number of iterations (200 in our experiments).

In the first case, given the efficiency of the item-to-bin heuristic, Phase 2 computes the final solution by iteratively examining the possible number of bins for  $\tau'$  (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  $\tau'$  to  $\delta$ ;

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.

Consensus is less close when Phase 2 is invoked because the maximum number of iterations is reached. Phase 2 of the meta-heuristic then builds a restricted version of the formulation (6.1)–(6.9) 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.

## 6.3 Parallel implementation

We developed a synchronous master–slave implementation for the proposed meta-heuristic.

---

**Algorithm 1** PH-based meta-heuristic for the SVCSBPP

---

**Scenario decomposition**

Generate a set of scenarios  $\mathcal{S}$ ;

Decompose the resulting deterministic model (6.1)–(6.9) 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 (6.22);

**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  $\tau'$  ( $\bar{\delta}_m^{\tau'} < \bar{\delta}_M^{\tau'}$ ) **then**

    Identify the consensus number of bins  $\delta$  of type  $\tau'$  by enumerating  $\delta \in [\bar{\delta}_m^{\tau'}, \bar{\delta}_M^{\tau'}]$  (and variable fixing)

**else**

    Fix consensus variables in model (6.1)–(6.9);

    Solve restricted (6.1)–(6.9) model using a MIP solver.

---

The master implements the meta-heuristic, assigning the tasks of solving the subproblems to a number of slave processors. (Note that the master may reassign subproblems among slaves when the workload of some appears larger than that of the others.) The master thus sends the overall solution and multiplier values to the slaves, and later collects their solutions, waiting until all the scenario subproblems have been solved. It then proceeds to calculate the new overall solution, update penalties and, if consensus is not reached, either starts a new iteration or proceeds to the second phase.

The parallelization reduces the computational time for each iteration and thus speeds up the algorithm. 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.

## 7 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 analyses 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 7.1), and by studying the stability of stochastic solutions (Section 7.2). We then proceed, Section 7.3, 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) (6.1)–(6.9), and the lower bound (LB) (5.1)–(5.4). We also analyze the efficiency of the parallel implementation by studying the scalability to 16 threads. Section 7.4 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 7.5 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 use is analyzed in Section 7.6. Finally, we examine in Section 7.7 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 either 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).

## 7.1 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:

- **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  $V^\tau$  needed to pack all items in the worst-case scenario (the scenario with the most items). This number is

$$\|\mathcal{J}^\tau\| = \left\lceil \frac{1}{V^\tau} \max_{s \in \mathcal{S}} \sum_{i \in \mathcal{I}^s} v_i^s \right\rceil. \quad (7.1)$$

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

$$f^\tau = V^\tau(1 + \gamma), \quad (7.2)$$

where  $\gamma$  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  $f^\tau$  multiplied by a factor  $(1 + \alpha_\tau^s)$ ,

inversely proportional to the availability of extra bins of type  $\tau$  in scenario  $s \in \mathcal{S}$ , where

$$\alpha_{\tau}^s = 1 - \frac{\|\mathcal{K}^{\tau s}\|}{\sum_{\tau \in \mathcal{T}} \|\mathcal{K}^{\tau s}\|} \cdot \beta, \quad (7.3)$$

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.

We generated (sampling) ten instances for each combination of the parameters defined above, yielding a total of 120 instances. A total of  $s = 1, \dots, 100$  scenarios were used in the experiments, which satisfies the stability conditions (Kaut et al., 2007) as shown in the next subsection.

## 7.2 Analysis of the size of the scenario tree

We performed in-sample and out-of-sample analyses (Kaut et al., 2007; Kaut and Wallace, 2007) in order to select the dimension of the scenario tree to use in the experiments described later in the section. The analysis was conducted on a subset of twenty-four instances, eight for each size T3, T5 and T10, considering different combination of the parameters mentioned in Section 7.1.

Ten scenario trees were generated for each instance and each cardinality  $|S| = 10, 25, 50, 100, 150, 200$ . Tables 1 and 2 display the aggregated results, by problem size (rows) and scenario-tree cardinality (columns), of the application of the meta-heuristic to these instances. Table 1 displays the average objective-function values for the solutions

	$S = 10$	$S = 25$	$S = 50$	$S = 75$	$S = 100$	$S = 150$	$S = 200$
T3	652.58	645.98	639.23	635.97	634.373	632.89	631.9
T5	799.27	790.83	780.24	776.98	774.547	772.01	771.34
T10	4798.58	4666.21	4620.45	4590.14	4570.03	4546.54	4550.73

Table 1: Average solution value for varying instance size and number of scenarios

	$S = 10$	$S = 25$	$S = 50$	$S = 75$	$S = 100$	$S = 150$	$S = 200$
T3	4.61%	3.49%	1.28%	1.01%	0.73%	0.41%	0.16%
T5	5.14%	3.35%	2.74%	1.34%	0.93%	0.52%	0.34%
T10	4.93%	3.30%	1.63%	1.39%	1.09%	0.55%	0.32%

Table 2: Standard deviation of solution values for varying instance size and number of scenarios

obtained through this experimentation, while Table 2 shows the corresponding standard deviations.

It is interesting to note out of this in-sample analysis that the standard deviation does not appear to be influenced by the size of the instance measured in number of bins (e.g., from three bins for T3 to ten bins for T10) of instances. This points to a stability within each scenario-tree dimension. Overall, increasing the number of scenarios appears beneficial. We adopted a compromise dimension of 100 scenarios providing in-sample stability of some 1% with respect to the standard deviation of the objective-function value.

For the out-of-sample analysis, we performed ten times the following procedure for each of the twenty-four instances (using the same scenario trees generated above): 1) solve a 200-scenario problem; 2) solve the instance with scenario trees of cardinality  $|S| = 10, 25, 50, 100, 150$ ; 3) evaluate each of the solutions obtained at step 2 within the 200-scenario context, by fixing their first-stage decision variables and solving the second stage, the recourse, of the 200 scenarios problem with the fixed first-stage variables; 4) compute the relative gap of the objective-function value of this solution relative to that of the 200-scenario problem (step 1). The results are displayed in Figure 1, where cardinality of the scenario tree makes up the horizontal axis, while the vertical one displays the average values of these relative gap (in percentage) over 10 runs for all instances.

We observe fundamentally the same phenomena for the out-of-sample stability analysis as previously for the in-sample one. According to the numerical results, both in-sample and out-of-sample stability are thus achieved with a sufficiently accurate precision (under 1% of standard deviation and relative gap, respectively) with a 100 scenario tree. We used this tree dimension in the rest of the experimental campaign.

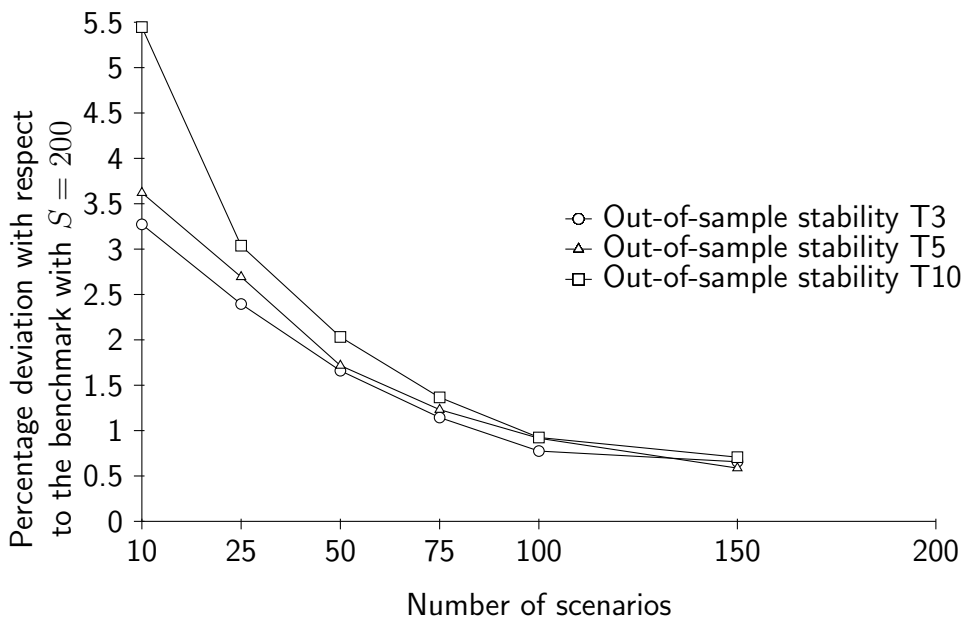


Figure 1: Out-of-sample stability with respect to the size of the scenario tree

### 7.3 Meta-heuristic performance analysis

We started by examining the computational performance of the MIP solver of CPLEX for the problem at hand. Initial experiments on the T3 and T5 instance sets showed a relative optimality gap (comparing the best solution and the lower bound computed by the solver) over 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. The gap decreases rapidly as the number of threads increases. With 16 parallel threads, CPLEX has an average optimality gap below 6% after 24 hours of wall-clock computing time. We therefore used these results in the comparisons of this section where, unless stated otherwise, all references to CPLEX and the PH-based meta-heuristic mean their parallel versions on 16 threads.

Table 3 details the comparison of CPLEX and the PH meta-heuristic, in their parallel version on 16 threads, for the T3 and T5 instances. We report, for each combination of instance set (Column 1) and item spread class (Column 2), the average optimality gaps  $\Delta^{CP}$  and average wall-clock running times  $t^{CP}$  of parallel CPLEX (Columns 3 and 4, respectively), the average gaps  $\Delta_{UB}^{PH} = (UB^{PH} - UB^{CP})/UB^{CP} * 100$  and average wall-clock running times of the parallel meta-heuristic (Columns 5 and 6, respectively), where  $UB^{CP}$  and  $UB^{PH}$  are the objective values of the two methods, respectively. Note that the value of  $\Delta_{UB}^{PH}$  may be positive or negative. When  $\Delta_{UB}^{PH}$  is negative, the meta-heuristic solution is better than that of CPLEX.

Set	Spread	CPLEX		PH	
		$\Delta^{CP}$	$t^{CP}$ [s]	$\Delta_{UB}^{PH}$	$t^{PH}$ [s]
T3	SP1	5.38	86400	-1.19	0.30
	SP2	5.74	86400	-1.39	0.59
	SP3	6.38	86400	-1.06	0.10
	SP4	5.77	86400	-0.86	0.12
T5	SP1	4.99	86400	-1.17	0.50
	SP2	5.23	86400	-1.38	0.18
	SP3	6.44	86400	-1.98	0.15
	SP4	5.20	86400	-1.36	0.65

Table 3: Comparison of parallel CPLEX and PH-based meta-heuristic

The results show that CPLEX cannot solve the SVCSBPP to optimality within a reasonable time limit for any instance considered. The average optimality gap after 24 hours for T3 and T5 instances is always greater than 5%, with a maximum value of 6.44%. Performance is worst as the problem size increases. Moreover, CPLEX was not able to identify in 24 h of wall-clock computing time a good feasible solution for T10 instances, which may easily involve more than 10 million variables. Memory (e.g., more than 64 GB of memory already occupied after 2 h of computations) rapidly becomes a bottleneck.

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 parallel CPLEX. The gap  $\Delta_{UB}^{PH}$  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 3 s. 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.

Because CPLEX results were not available for T10 instances, we compare the results of the meta-heuristic and those obtained by solving the lower bound (5.1)–(5.4) (using parallelCPLEX as well). Recall that the lower bound 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 required only a few seconds to compute the bound for sets T3, and T5, and at most 120 s for set T10.

Table 4 reports, for each instance set (Column 1) and item spread class (Column 2), the relative gap (in percentage)  $\Delta_{LB}^{PH}$  between the objective function value obtained by the PH meta-heuristic and the lower bound (Column 3), as well as the average and maximum percentage increase in the objective function value (Columns 4 and 5) resulting from the use of the lower-bound solutions as the capacity plan. On average, the gaps

Set	Spread	$\Delta_{LB}^{PH}$	$C^{LB}$	$C_{max}^{LB}$
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 4: Comparison of LB and PH solutions

are smaller than 2%, underlining the accuracy of the PH meta-heuristic. We further qualified this accuracy by evaluating the quality of the lower bound solutions by fixing the bin selections defined by the bound and computing the recourse 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 the lower bound and the PH meta-heuristic match exactly.

We conclude this section with a short analysis of the computational efficiency of the sequential meta-heuristic and the scalability of the parallel the meta-heuristic (the results are displayed in the Appendix). We analyzed the computational wall-clock times for each instance set and item spread class for 1, 2, 4, 8, and 16 threads.

Sequential computational times are negligible for T3 and T5 instances (i.e., maximum of 2.36 wall-clock seconds), 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 (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 VCSBPP heuristic is very fast on all instances (which results in a balanced workload for slaves), the results indicate that parallel computation is not really required for the small instances (and certainly not 16 threads; 4 are already 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 instances

can be addressed efficiently with the proposed methodology.

## 7.4 Benefit of modeling the uncertainty

The question addressed in this section is whether modeling the 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.

We use 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  $EVPI = RP - WS$  is defined as the difference between the objective values of the stochastic solution and that of the wait-and-see problem defined assuming the realizations of all the random parameters are known at the first stage. The  $VSS = EEV - RP$  indicates the expected gain from solving the stochastic model rather than its deterministic counterpart, with  $EEV$  denoting the solution value of the stochastic model with the first-stage decision variables fixed at the optimal values obtained by solving the deterministic problem with the random parameters fixed at their expected values..

Table 5 displays aggregated results 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  $EVPI/RP * 100$ , and the average and maximum  $VSS$  percentages (Columns 5 and 6) computed as  $VSS/RP * 100$ .

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 5:  $EVPI$  and  $VSS$  comparison

Both measures emphasize the interest of addressing the stochastic formulation rather than a deterministic approximation. The EVPI percentage is around 10%, increasing with the problem dimensions to a maximum value above 15%. This underlines the benefit of having information about the future in advance. The average and maximum values of VSS also increase as the instance size increases, the gap between the expected-value solution and the stochastic solution being 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 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.

We complete this analysis with a comments on the increasing trend of the *EVPI* and *VSS*. It is indeed common to see these measures increasing with the problem size (when they are meaningful). What is particular about the problem we address is that there are two sources of uncertainty, the item number and volumes, on one hand, and the availability and cost of the extra bins, on the other hand, which combine to make uncertainty a core characteristic of the problem. One thus notices that the relative gap to the optimal solution is sensitive to both the potential number of bin types and the number of items to be packed, moving from 5% for instances with 3 bin types to a mean of 11% for the 10-bin-types instances. One also notices on examining the structure of EV formulations that they yield different first-stage decision, generally overestimating the the future demand to be loaded (in terms of total item volume) and the availability of extra bins (a larger set of bins is assumed available for the recourse action). This yields a more expensive plan in many cases (by about 17.70%), and very large LCL costs in about 2% of the cases. This behavior strongly suggests that the usage of mean values is a bad estimator of the uncertainty. On the other hand, the gaps are sufficiently large from an economic point of view to justify the usage of the full stochastic recourse problem (Perboli et al., 2014a, 2011).

## 7.5 Solution analysis

We focus on the structure of the capacity-planning solutions and in particular on the use of bins in the plan. Table 6 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

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 6: Structure of solutions

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 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.

## 7.6 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  $\tau$  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	$m$ (%)	$N'_{\mathcal{T}}$	$\Delta'_{Obj}$
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 7: Sensitivity of capacity planning to bin availability

Let  $u^{\tau} = \|\mathcal{J}^{\tau}\|$  be the number of bins of type  $\tau$  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 (7.4) for the first policy, and (7.5) for the second policy:

$$\sum_{j \in \mathcal{J}^{\tau}} y_j^{\tau s} \geq m \cdot u^{\tau} \quad \forall \tau \in \mathcal{T}, \quad s \in \mathcal{S} \quad (7.4)$$

$$\sum_{j \in \mathcal{J}^{\tau}} y_j^{\tau s} \leq m \cdot u^{\tau} \quad \forall \tau \in \mathcal{T}, \quad s \in \mathcal{S}. \quad (7.5)$$

We compare the results of the PH meta-heuristic for the two policies and those of the original formulation. Table 7 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 ( $\Delta'_{Obj}$ , 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 a larger number of 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.

## 7.7 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 or lower demands.

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

1. **High demand** (Figure 2a illustrates the range of the demand in this situation 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 extra capacity.
2. **Low demand** (Figure 2b): 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.

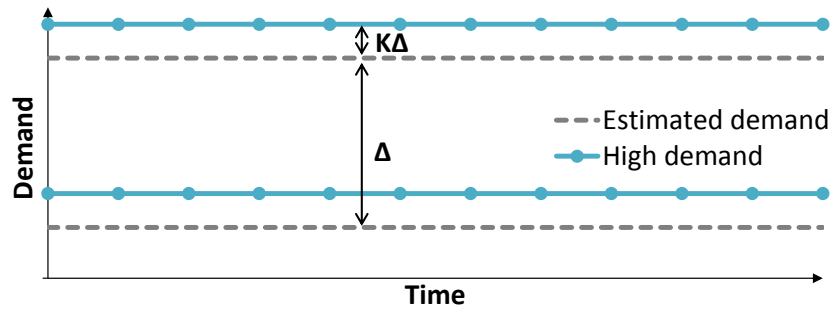
3. **Economic crisis** (Figure 2c): The demand decreases rapidly to a value below the estimated demand. It then stabilizes below the average demand of the scenarios in  $\mathcal{S}$ .
4. **Economic recovery** (Figure 2d): The demand is initially below the average estimated demand, but it increases rapidly to a peak above the maximum estimation. It then stabilizes.

The overall simulation process for a given demand situation was:

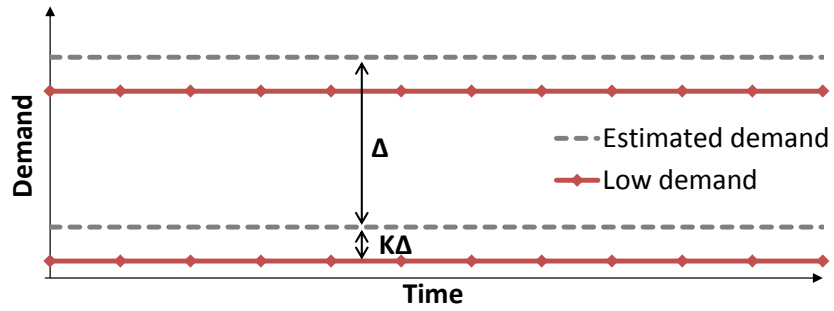
- Given an instance and the associated set of scenarios  $\mathcal{S}$ , find the capacity-planning solution using the PH-based meta-heuristic.
- Create a new set of scenarios  $\mathcal{S}'$ , with  $|\mathcal{S}'| = 365$ :
  - Identify the maximum and minimum demand in  $\mathcal{S}$  and compute the difference  $\Delta$  between them.
  - Compute  $K\Delta$  for the current demand situation, 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 2).
  - Generate (same generation procedure used previously) 365 new scenarios respecting the trends of the demand in the current situation 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 compute statistical measures of the results.

We are thus using a 3650-scenario sample for the Monte-Carlo simulation for each demand situation, that is, we simulate 10 years (repetitions) with 365 days (scenarios) each. We performed a number of tests varying the size of the sample and these strengthened our confidence that the sample size is sufficient to be representative.

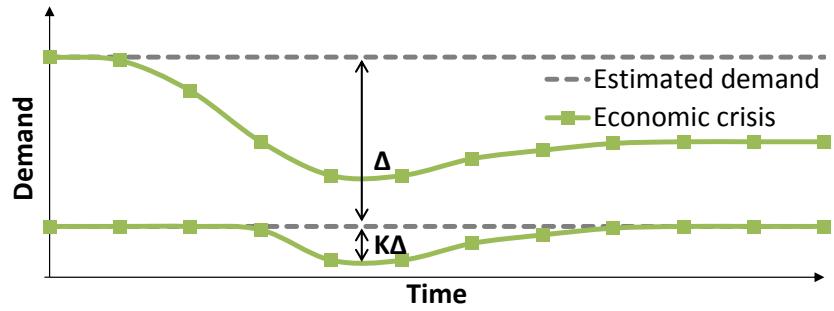
Figure 3 shows the results of the Monte Carlo simulation for an illustrative instance from set T10 (the other instances produced similar results). The figure reports, for each demand scenario, the usage percentage of the logistics capacity plan, defined as the ratio of the number of planned (selected in the first stage) bins actually used in the solution to



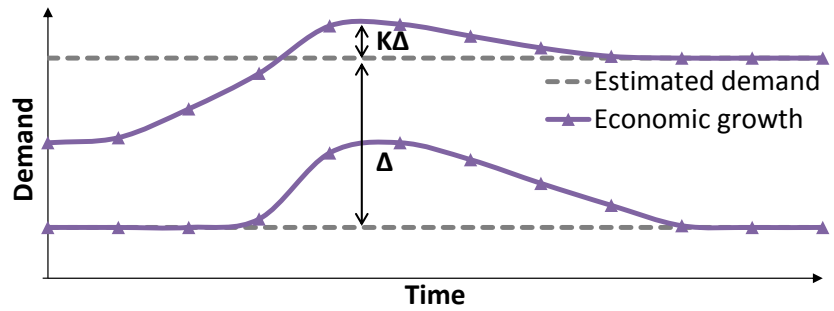
(a) Distribution of high-demand situation



(b) Distribution of low-demand situation

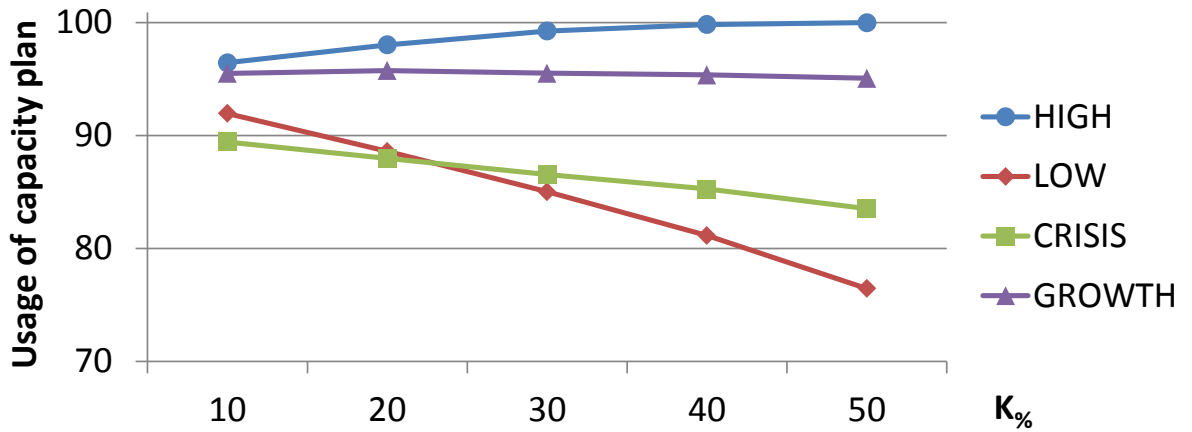


(c) Distribution of economic-crisis situation

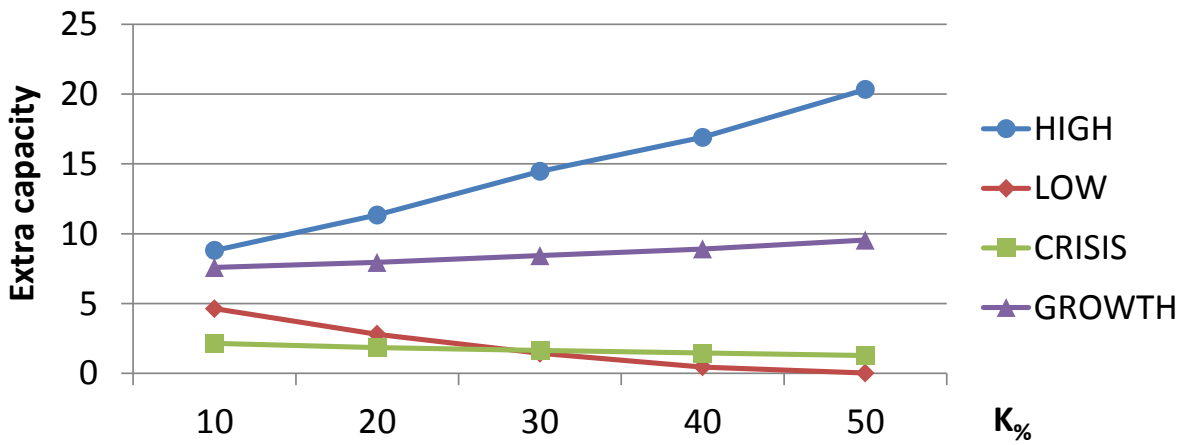


(d) Distribution of economic-growth situation

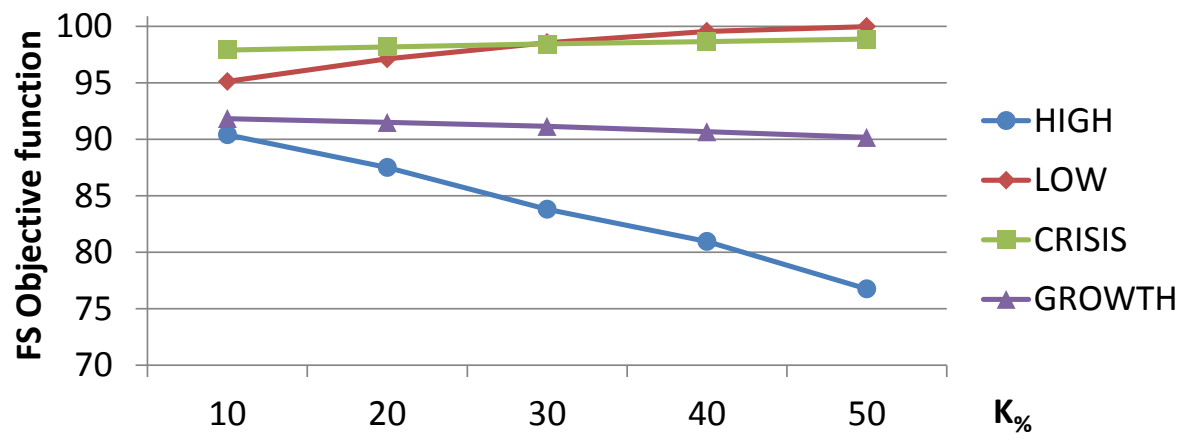
Figure 2: Demand situations for the Monte Carlo simulation



(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 3: Monte Carlo simulation results for an illustrative T10 instance

the total number of bins selected in the plan (Figure 3a), 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 3b), and the percentage of the objective function value associated with the capacity plan (Figure 3c), 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 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 relative weight of the first-stage cost ranges between 75% and 100%. These values emphasize 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 analysis 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  $\mathcal{S}'$ ).

## 8 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 bins (capacity units) of various types, while the second stage models the subsequent adjustments to the plan performed each time the plan is applied and new information becomes known. The adjustment secures, when needed, the extra units of capacity required to pack all the received items.

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*, 222:125–141.
- Bertazzi, L., Chan, L. M. A., and Speranza, M. G. (2007). Analysis of practical policies for a single link distribution system. *Naval Research Logistics*, 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.
- 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, Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation - CIRRELT, Montréal, QC, Canada.
- 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., Perboli, G., and Tadei, R. (2012). Recent advances in multi-dimensional packing problems. In Volosencu, C., editor, *New Technologies - Trends, Innovations and Research*. ISBN: 978-953-51-0480-3, pages 91–110. InTech.
- 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.
- 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.
- Kaut, M. and Wallace, S. W. (2007). Evaluation of scenario generation methods for stochastic programming. *Pacific Journal of Optimization*, 3:257–271.
- 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.
- Martello, S. and Toth, P. (1990). *Knapsack Problems - Algorithms and computer implementations*. John Wiley & Sons, Chichester, UK.
- 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., Crainic, T. G., and Tadei, R. (2011). An efficient metaheuristic for multi-dimensional multi-container packing. In *Automation Science and Engineering (CASE), 2011 IEEE Conference on Automation Science and Engineering*, pages 563–568.
- Perboli, G., Gobbato, L., and Maggioni, F. (2015). A progressive hedging method for the multi-path travelling salesman problem with stochastic travel times. *IMA Journal of Management Mathematics*, page dpv024.
- 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.
- 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.

## A Behavior of the Parallel Meta-heuristic

We present results with respect to the the scalability of the parallel implementation of the meta-heuristic with up to 16 threads.

The wall-clock times  $t^{PH}$  for each instance set (Column 1) and item spread class (Column 2) are reported in Table 8 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 4.

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 8: Average computational times for the meta-heuristic with different numbers of threads

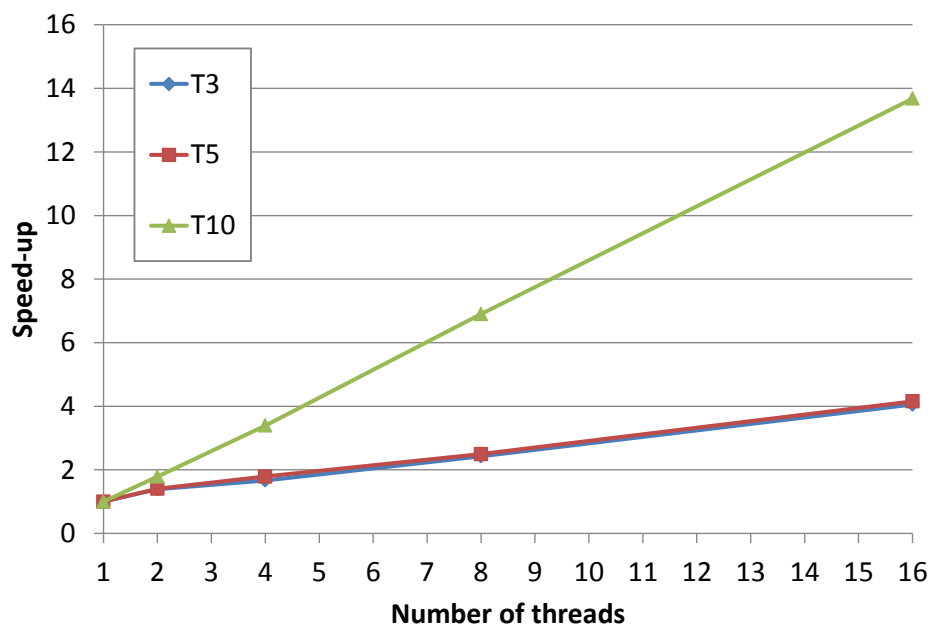


Figure 4: Speed-up of PH algorithm for different numbers of parallel threads