wirtschafts
wissenschaften
htw saar

I. Correia I T. Melo

Dynamic facility location problem with modular capacity adjustments under uncertainty

Schriftenreihe Logistik der Fakultät für Wirtschaftswissenschaften der htw saar

Technical reports on Logistics of the Saarland Business School

Nr. 17 (2019)

logistik

ISSN 2193-7761

Dynamic facility location problem with modular capacity adjustments under uncertainty

I. Correia I T. Melo

Bericht/Technical Report 17 (2019)

Die Veröffentlichungen in der Berichtsreihe der Fakultät für Wirtschaftswissenschaften können bezogen werden über:

Hochschule für Technik und Wirtschaft des Saarlandes
Fakultät für Wirtschaftswissenschaften
Campus Rotenbühl
Waldhausweg 14
D-66123 Saarbrücken

Telefon: +49 (0) 681/ 5867 - 519
E-Mail: fakultaet-wiwi@htwsaar.de
Internet: www.htwsaar.de/wiwi

Dynamic facility location problem with modular capacity adjustments under uncertainty

Isabel Correia[a∗], Teresa Melo[b]

[a] *Centro de Matemática e Aplicações (CMA), Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, P 2829-516 Caparica, Portugal*
[b] *Business School, Saarland University of Applied Sciences, D 66123 Saarbrücken, Germany*

Abstract

We address a stochastic multi-period facility location problem with two customer segments, each having distinct service requirements. While customers in one segment receive preferred service, customers in the other segment accept delayed deliveries as long as lateness does not exceed a pre-specified threshold. In this case, late shipments incur additional tardiness penalty costs. The objective is to define a schedule for facility deployment and capacity scalability that satisfies all customer demands at minimum cost. Facilities can have their capacities adjusted over the planning horizon through incrementally increasing or reducing the number of modular units they hold. These two features, capacity expansion and capacity contraction, can help substantially improve the flexibility in responding to demand changes. Future customer demands are assumed to be unknown. We propose two different frameworks for the capacity scalability decisions and present a two-stage stochastic model for each one of them. When demand uncertainty is captured by a finite set of scenarios, each of which having some known probability of occurrence, we develop the extensive forms of the associated stochastic programs. Additional inequalities are derived to enhance the original formulations. An extensive computational study with randomly generated instances that are solved with a general-purpose optimization solver demonstrate the usefulness of the proposed enhancements. Specifically, a considerably larger number of instances can be solved to optimality in much shorter computing times. Useful insights are also provided on the impact of the two different frameworks for planning capacity adjustments on the network configuration and total cost.

Keywords: facility location, dynamic capacity adjustment, delivery lateness, stochastic programming, valid inequalities

∗Corresponding author. *E-mail address*: isc@fct.unl.pt

1 Introduction

Location analysis provides a framework for simultaneously finding sites for facilities and assigning spatially distributed demand points to these facilities in order to optimize some measurable criterion (typically, minimization of cost or maximization of profit). Large costs associated with property acquisition and facility construction make facility location projects long-term investments. Decision makers must select sites that will not simply perform well according to the current system state, but that will also continue to be effective as market conditions and customer service requirements evolve over time. To deal with these challenges, the decision process could also include the optimal time sequence for opening new facilities, in addition to the determination of the best locations (Nickel and Saldanha da Gama [43]). Hence, the ability to explicitly evaluate the impact of location decisions over the course of a multi-period planning horizon represents an important step towards modeling real-world problems. In a dynamic setting, capacity sizing decisions can also be considered in order to find the optimal timing for capacity acquisition, expansion, and contraction at the selected locations. This strategy increases the ability of firms to adapt to evolving economic contexts (e.g., fluctuations in demand).

Generally, (multi-period) facility location research relies upon the assumption that future conditions can be predicted with a reasonable degree of accuracy. This has given rise to the development of deterministic models, which are by far the most common. Such models have the advantage of often being amenable to mathematical analysis. The mathematical clarity offered by deterministic models can, however, come at a cost of decreased realism. In practice, many factors are uncertain, such as demands, prices, costs, and availability of resources, thereby motivating the development of stochastic models (Correia and Saldanha da Gama [16]). Ignoring uncertainty entails the risk that once the true nature of the uncertain parameters reveals itself, adjustments in the facility network configuration may be necessary that are time-consuming and require a prohibitively high expense. Although stochastic models often raise questions of tractability, the stochastic dimension is becoming increasingly present in facility location problems (Correia and Saldanha da Gama [16], Snyder [46]).

In this paper, we consider a variant of the multi-period capacitated facility location problem in which some data are uncertain. We assume that there is a finite set of customers with demand for a single product family over a multi-period finite planning horizon, and a finite set of potential locations for the facilities that will offer the commodity. The goal is to choose facility locations, determine a schedule for facility opening and capacity adjustment (i.e., expansion and contraction), and to decide which customers to serve and from which facilities. A distinctive feature of our problem is that customers are sensitive to delivery lead times. This means that some customers require their demands to be met on time (i.e., these customers impose a zero delivery lead time), whereas the remaining customers accept delayed deliveries as long as lateness does not exceed a pre-specified threshold. This form of customer segmentation can be encountered in a variety of industries, such as retailing (Duran et al. [23], Li et al. [37]) and aftermarket services (Alvarez et al. [7], Hung et al. [31], Wang et al. [48]). Future demands are not known *a priori*. Instead, they are assumed to be described by a finite discrete random vector, following a probability distribution that is known in advance (e.g., estimated using historical data). Under the assumption that the decision maker follows a risk neutral strategy, the objective is to minimize the total expected cost. The latter includes fixed costs for opening and operating facilities, fixed costs for capacity scaling, and variable processing, distribution, and tardiness penalty costs. We note that our problem arises in the context of facility sizing decisions being reversible in the medium term. This is the case when space and equipment can be rented or leased. Therefore, the three planning levels are integrated through considering strategic decisions (involving facility location), tactical decisions (regarding capacity scalability), and operational decisions (pertaining to demand allocation). Accordingly, facility location and capacity sizing decisions can only be made at selected time periods, whereas demand allocation decisions can be made at any time period of the planning horizon.

A variant of our problem has been studied by Correia and Melo [15], who have shown that this is already a challenging NP-hard problem when all parameters are assumed to be known with certainty. Our work is the first attempt to understand the additional complexity that arises from incorporating stochasticity into the underlying difficult-to-solve deterministic location and capacity scalability problem in a multi-period setting. For this reason, we focus on demand uncertainty and defer the consideration of additional uncertain parameters. To the best of the authors' knowledge, this problem has not been previously addressed in the literature.

Different sources of uncertainty have been captured, either individually or simultaneously, in the literature dedicated to stochastic facility location, with demand uncertainty being the most often considered factor. Early contributions addressing uncertain demand include the works by Louveaux [38] for the p-median and the simple plant location problems, and Laporte et al. [36] for the capacitated facility location problem. Over the past decades, the stochastic dimension in facility location modeling has evolved, also towards other areas, such as hub location (Alumur et al. [6], Contreras et al. [14]) and covering facility location (Berman and Wang [12]), among others. Facility location also plays an important role in various application contexts and recent advances in stochastic modeling have been reported, e.g., in humanitarian logistics (Döyen et al. [22], Grass et al. [28], Kınay et al. [35]), supply chain network design (Govindan et al. [27], Mohamed et al. [42]), and planning health care services (Beraldi and Bruni [11]).

While static (or single-period) stochastic location problems have attracted increasing attention in the literature, there has been limited research on the dynamic (or multi-period) stochastic counterpart. Albareda-Sambola et al. [4] studied a multi-period facility location problem with various sources of uncertainty, such as demand, facility location and customer-assignment costs, minimum number of customers to be serviced, and minimum number of facilities to be established. Recently, Marín et al. [39] proposed a stochastic programming modeling framework for a general class of covering location problems and developed a Lagrangian relaxation-based heuristic to tackle large-scale instances. In the context of humanitarian logistics, Kim et al. [34] investigated the problem of finding optimal locations for drone facilities over a planning horizon taking into account that flight distance is uncertain due to battery consumption being affected by weather conditions. Khodaparasti et al. [33] addressed the problem of designing a network of nursing homes under a budget constraint on each time period. Chance-constraints are imposed that capture demand uncertainty. Marković et al. [40] developed a multi-stage stochastic program for the problem of locating law enforcement facilities on a road network to intercept stochastic vehicle flows that try to avoid these facilities. Since the facilities have a limited lifespan, the optimal timing for their deployment over the planning horizon needs to be determined. Mohamed et al. [42] proposed a multi-stage stochastic program to locate intermediate facilities in a two-echelon distribution network under uncertain demand. For tractability reasons, the original formulation is approximated by two-stage stochastic models. In some works (e.g., Baron et al. [10], Fattahi et al. [24], Georgiadis et al. [25]), facility location decisions are planned and implemented at the beginning of the planning horizon and remain unchanged thereafter, while customer allocation and other decisions are made in every time period taking into account demand uncertainty.

Capacity planning decisions are often incorporated into facility location problems (Owen and Daskin [45], Verter and Dincer [47]). In particular, the gradual adjustment of capacities over a planning horizon is an important driver for improving flexibility within a facility network. While excess capacity results in the loss of capital investment and opportunity costs, capacity shortage leads to delivery performance deterioration, and as a consequence, lowers the revenue and market share of a firm. Ideally, the provision of capacity should be in line with actual demand needs. Depending on the application context, different strategies can be adopted along the planning horizon, such as capacity expansion (Correia et al. [17], Cortinhal et al. [19]), capacity expansion combined with capacity reduction (Antunes and Peeters [8], Wilhelm et al. [49]), relocation of capacities to different locations (Melo et al. [41]), and temporary facility closing and reopening (Dias et al. [21], Jena et al. [32]). Among these strategies, capacity expansion is the predominant policy in a stochastic setting. Early contributions in this area focus on determining a schedule for capacity acquisition and expansion for a set of facilities already operating at fixed locations (Ahmed and Garcia [2], Ahmed et al. [3]). Uncertain parameters include demands and investment costs. Furthermore, any continuous amount of capacity can be installed. In contrast, Alonso-Ayuso et al. [5] assume a finite set of capacity expansion levels in a multi-commodity production system comprising suppliers, plants, and demand markets. For a multi-echelon supply chain network, Aghezzaf [1] propose a decomposition-based algorithm to determine a facility location and capacity expansion schedule that is robust to the uncertain realization of customer demand. In the problem studied by Hernández et al. [30], the location and size of new prisons in Chile are to be determined along with a schedule for capacity upgrading at both existing and new facilities under stochastic demand. Recently, Correia et al. [18] proposed a stochastic modeling framework for a multi-period hub location problem with uncertain demand. In addition to finding the location of hub facilities and setting their initial capacities, an expansion plan is also determined. This entails deciding on the time period and the number of additional modular units to be installed in a hub facility. Even though modular capacities are relevant in many contexts, e.g. to represent modular equipment (Delmelle et al. [20], Gourdin and Klopfenstein [26]), they have been hardly considered in the literature. This feature is captured by our work.

Uncertainty can be formalized in many different ways (Correia and Saldanha da Gama [16]), among which the probabilistic interpretation of randomness seems to be the prevailing approach in stochastic facility location research. Often, this leads to capturing uncertainty by a finite set of scenarios with known probabilities, where each scenario specifies the complete realization of all uncertain parameters. Due to the distinct nature of the decisions to be made in a facility location problem, either a two-stage or a multi-stage stochastic modeling framework is often adopted (Correia and Saldanha da Gama [16], Govindan et al. [27], Snyder [46]). In this paper, we follow a scenario-based approach and propose two alternative modeling strategies, both involving two-stage stochastic programming formulations. In both models, the first-stage problem determines the opening schedule for facilities and their initial capacities over the entire planning horizon before uncertainty about demand is revealed. Additionally, operational decisions are captured in the second-stage problem by specifying the commodity flow between the selected locations and the customers. Hence, distribution decisions are made after the demand uncertainty has been resolved. The proposed stochastic programming models differ in the way capacity scalability decisions are treated. In the first model, we follow the approach usually taken in the literature (Correia and Saldanha da Gama [16], Govindan et al. [27], Snyder [46]), and assume that these decisions have a strategic nature. Hence, we include them in the set of *ex ante* or first-stage decisions. In contrast, in the second model, decisions on the timing and sizing of capacity adjustments are assumed to have a tactical nature. In this case, they become *ex post* or second-stage decisions. To the best of our knowledge, only one other study has addressed both capacity expansion and capacity reduction in a stochastic facility location setting, namely Zhuge et al. [51]. These authors propose a two-stage stochastic program for the problem of locating distribution centers (DCs) in a two-echelon network. First-stage decisions include identifying the locations where to open DCs in the first time period and selecting their initial capacities from a set of three options (small, medium, and large capacity levels). The second-stage decisions define the timing for capacity expansions and contractions as well as the flow of multiple products in the network. Uncertainty in future demand and the budget available for replacing capacity levels is considered. Small-sized problems are solved by a Lagrangian-based heuristic.

Furthermore, stochastic programs with second-stage capacity expansion decisions (but without the option of planning for capacity contraction) are developed by Correia et al. [18] for a hub location problem, and Heckmann [29] for a facility location problem with disruptive events. To alleviate the capacity reduction caused by such events, suitable expansion plans are determined. Contrary to our work, decisions on opening and operating facilities are made once at the beginning of the planning horizon.

The contributions of the present work are summarized as follows: (1) We propose two alternative modeling strategies for a new multi-period facility location and capacity scalability problem with uncertain demand; (2) For each strategy, we formulate the *Deterministic Equivalent Program* and develop valid inequalities that significantly improve the polyhedral description of the feasible region; (3) We evaluate the potential benefit obtained from solving the stochastic programs instead of their deterministic counterparts by means of two measures, namely the expected value of perfect information and the value of the stochastic solution; (4) We assess the validity of the proposed two-stage stochastic mixed-integer linear programming (MILP) models by reporting and analyzing the results of an extensive computational study. In particular, we provide insights on the impact of considering capacity scalability decisions as second-stage (tactical) decisions as opposed to the typical approach of treating them as first-stage (strategic) decisions.

The remainder of this paper is organized as follows. In Section 2, we formally describe our problem and present two different stochastic programs along with the corresponding deterministic equivalent forms. Several sets of valid inequalities are developed in Section 3 to enhance the original formulations and therefore make them more amenable to standard off-the-shelf optimization software. Computational results are reported in Section 4 and the proposed formulations are compared using various metrics. Finally, Section 5 includes a summary of our findings and gives directions for future research.

2 Problem statement and stochastic formulations

The stochastic multi-period facility location and capacity scalability problem addressed in this paper builds upon a deterministic variant studied by Correia and Melo [15]. For convenience, we restate the assumptions that are common to the deterministic and stochastic settings, and introduce additional information pertaining to the stochastic problem. Subsequently, two stochastic formulations are proposed that differ in the framework under which capacity scalability decisions can be made.

A planning horizon comprising a finite number of time periods is considered along with a set of customers with demands for a single commodity in each period. Customers are differentiated according to their service requirements. Specifically, some customers require their demands to be met on time, whereas the remaining customers accept delayed deliveries as long as lateness does not exceed a pre-specified threshold. In this case, an order can be split over multiple time periods for the same customer. Typically, customers belonging to the first segment are willing to pay higher prices in order to have their demands satisfied in the same period in which they occur. To customers in the second segment, discount prices are often offered in exchange for longer response times. In our modeling framework, this scheme is translated into a tardiness penalty cost that is incurred per unit of backorder and per period. Customer service-differentiation offers greater flexibility to a firm in designing and managing its facility network.

A set of potential locations where facilities can be established is assumed to be available.

When a facility is opened, its initial capacity level must also be specified. The latter is expressed by the number of modular units that are installed, all having the same size. The adoption of a base modular unit is common in different application areas, such as in the public sector (Delmelle et al. [20]) and in telecommunications (Gourdin and Klopfenstein [26]). Over the planning horizon, the capacity of an operating facility can be adjusted to increase responsiveness to changes in the level of customer demand. Two forms of capacity scalability are considered: capacity expansion and capacity contraction. This means that one or several modular units can be added to or removed from a facility. Naturally, these options are mutually exclusive for the same facility at the same time period. A limit on the total number of modular units that a facility can hold at any time is pre-specified. Due to the sizeable investment associated with opening facilities, locations cannot be temporarily closed and reopened. Therefore, once a facility is established, it must remain open until the end of the planning horizon.

The objective is to find a schedule for the deployment of facilities and the expansion and contraction of their capacities in order to satisfy customer demands at minimal total cost. Different time scales are considered for decision-making. The capital-intensive decisions involving facility location and capacity scalability decisions can be made at the beginning of selected time periods, hereafter termed *design periods*. All other decisions concerning the processing of the commodity at the open facilities and the allocation of customer demands can be made at any time period. This setting is also adopted by other authors, e.g. Mohamed et al. [42].

In the problem addressed in the companion paper [15], all relevant data (i.e., costs, customer demands, and other parameters) are assumed to be known in advance. In the present work, this assumption is relaxed through considering that future customer demands are uncertain. Uncertainty is described by a finite random vector whose joint probability distribution is assumed to be foreknown, e.g. using historical data. The finitely many possible realizations of this random vector are called *scenarios*. Specifically, each scenario describes all customer demands over all periods of the planning horizon.

Before detailing the two stochastic programming models that we propose, we first introduce the notation that is common to both formulations.

2.1 Notation

Sets:

T Set of time periods along the planning horizon.

T_L Subset of *design* time periods at which location and capacity scalability decisions can be made ($T_L \subset T$).

I Set of potential facility sites.

J^0 Set of customers requiring timely deliveries (i.e., their demands must be satisfied in the same time periods in which they occur).

J^1 Set of customers that tolerate delayed demand fulfillment.

J Set of all customers ($J = J^0 \cup J^1$; $J^0 \cap J^1 = \emptyset$).

Let $\ell = 1$, resp. $\ell_{max} = \max\{\ell \in T_L\}$, be the first, resp. last, design time period in which decisions can be made on opening facilities in potential locations and adjusting their capacities through adding or removing modular units. The size of a facility at time period $\ell \in T_L$ is the outcome of the design and scalability decisions taken until that period. For $\ell < \ell_{max}$, this size remains unchanged over all intermediate periods between two consecutive design periods, ℓ and ℓ' ($\ell < \ell'$). For $\ell = \ell_{max}$, the capacity of the facility is maintained until the last period of the planning horizon. Let $\phi(\ell)$ denote the last time period between two consecutive design periods ℓ and ℓ'. It follows that

$$\phi(\ell) = \begin{cases} \max\{t \in T : t < \ell'\} & \text{if } \ell < \ell_{max}, \\ |T| & \text{if } \ell = \ell_{max}. \end{cases} \tag{1}$$

Deterministic parameters:

Q Capacity of a single modular unit.

n_i Maximum number of modular units that can be available at location i at any time period ($i \in I$).

ρ_j Maximum allowed delay (expressed by the total number of time periods) to satisfy the demand of customer j ($j \in J^1$).

FO_{ik}^{ℓ} Fixed cost of opening facility i with k modular units at design time period ℓ ($i \in I$; $k = 1, \ldots, n_i$; $\ell \in T_L$).

FE_{ik}^{ℓ} Fixed cost of expanding the capacity of facility i with k modular units at design time period ℓ ($i \in I$; $k = 1, \ldots, n_i - 1$; $\ell \in T_L \setminus \{1\}$).

FR_{ik}^{ℓ} Fixed cost of removing k modular units from facility i at design time period ℓ ($i \in I$; $k = 1, \ldots, n_i - 1$; $\ell \in T_L \setminus \{1\}$).

M_{ik}^t Fixed cost of operating facility i with k modular units at time period t $(i \in I;\ k = 1, \ldots, n_i;\ t \in T)$.

o_{ik}^t Unit processing cost charged by facility i operating with k modular units at time period t $(i \in I;\ k = 1, \ldots, n_i;\ t \in T)$.

c_{ij}^t Unit distribution cost from facility i to customer j at time period t $(i \in I;\ j \in J;\ t \in T)$.

$p_j^{tt'}$ Unit tardiness penalty cost for satisfying demand of customer j in period t' that was originally placed in period t $(j \in J^1;\ t \in T;\ t' = t, t+1, \ldots, \min\{t + \rho_j,\ |T|\})$; in particular, $p_j^{tt'} = 0$ for $t' = t$.

The customer-specified time lag for demand satisfaction ρ_j $(j \in J^1)$ imposes that an order placed by customer j for time period t $(t \in T)$ must be filled over periods $t, t+1, \ldots, t + \rho_j$. In case $t + \rho_j > |T|$, then the last delivery must occur at period $|T|$. This condition ensures that demand cannot be carried over to periods beyond the planning horizon. If $\rho_j = 0$ for every customer $j \in J^1$ then our problem reduces to the classical setting in multi-period facility location.

We note that some cost parameters capture economies of scale, namely in all fixed costs $(FO_{ik}^\ell,\ FE_{ik}^\ell,\ FR_{ik}^\ell,\ M_{ik}^t)$ and in the variable processing costs (o_{ik}^t).

Stochastic parameters:

$d_j^t(\xi)$ Random variable representing the demand of customer j at time period t $(j \in J; t \in T)$.

Ξ Finite discrete random vector with $\Xi = \big((d_j^t)_{j \in J;\, t \in T}\big)$ and $\Xi \subset \mathbb{R}_+^{|J| \times |T|}$.

It is assumed that the joint probability distribution of the random vector Ξ is foreknown.

2.2 A stochastic formulation

Given the problem assumptions, a natural approach to the decision-making process is to develop a two-stage stochastic programming model, where the facility location and capacity scalability decisions are made in the first stage, and the remaining processing and distribution decisions are deferred to the second stage. In other words, the first-stage (or *ex ante*) decisions are associated with the definition of a schedule for location, capacity acquisition, and capacity adjustment to be implemented over the entire planning horizon, before a realization of the demand becomes known. This results from the strategic nature assumed for these decisions as they have a long-term impact. The following binary variables represent the first-stage decisions.

$$z_{ik}^{\ell} = 1 \text{ if a new facility is established with } k \text{ modular units at potential location } i \text{ at time period } \ell, \text{ 0 otherwise } (i \in I;\ k = 1, \ldots, n_i;\ \ell \in T_L). \quad (2)$$

$$e_{ik}^{\ell} = 1 \text{ if the capacity of a facility in location } i \text{ is expanded with } k \text{ modular units at time period } \ell, \text{ 0 otherwise } (i \in I;\ k = 1, \ldots, n_i - 1;\ \ell \in T_L \setminus \{1\}). \quad (3)$$

$$r_{ik}^{\ell} = 1 \text{ if the capacity of a facility in location } i \text{ is reduced by removing } k \text{ modular units at time period } \ell, \text{ 0 otherwise } (i \in I;\ k = 1, \ldots, n_i - 1;\ \ell \in T_L \setminus \{1\}). \quad (4)$$

$$v_{ik}^{\ell} = 1 \text{ if a facility is operated at location } i \text{ with } k \text{ modular units at time period } \ell, \text{ 0 otherwise } (i \in I;\ k = 1, \ldots, n_i;\ \ell \in T_L). \quad (5)$$

After customer demand is observed, operational decisions are made by specifying the quantities of product to be processed at operating facilities and distributed to customers over the planning horizon. Therefore, second-stage (or *ex post*) decisions are represented by the following continuous variables:

$$x_{ij}^{t}(\xi) : \text{Amount of product distributed from facility } i \text{ to customer } j \text{ at time period } t\ (i \in I;\ j \in J^0;\ t \in T). \quad (6)$$

$$y_{ij}^{tt'}(\xi) : \text{Amount of product distributed from facility } i \text{ to customer } j \text{ at time period } t' \text{ to (partially) satisfy demand of period } t\ (i \in I;\ j \in J^1;\ t \in T;\ t' = t, t+1, \ldots, \min\{t + \rho_j,\ |T|\}). \quad (7)$$

$$w_{ik}^{t}(\xi) : \text{Total quantity of product handled by facility } i \text{ operating with } k \text{ modular units at time period } t\ (i \in I;\ k = 1, \ldots, n_i;\ t \in T). \quad (8)$$

Thus, the implicit representation of the two-stage stochastic program is

$$\text{Min} \sum_{\ell \in T_L} \sum_{i \in I} \sum_{k=1}^{n_i} FO_{ik}^{\ell}\, z_{ik}^{\ell} + \sum_{\ell \in T_L \setminus \{1\}} \sum_{i \in I} \sum_{k=1}^{n_i - 1} FE_{ik}^{\ell}\, e_{ik}^{\ell} + \sum_{\ell \in T_L \setminus \{1\}} \sum_{i \in I} \sum_{k=1}^{n_i - 1} FR_{ik}^{\ell}\, r_{ik}^{\ell} + \sum_{\ell \in T_L} \sum_{i \in I} \sum_{k=1}^{n_i} v_{ik}^{\ell} \sum_{t=\ell}^{\phi(\ell)} M_{ik}^{t} + \mathbb{E}_{\Xi}\left[Q(z, e, r, v, \xi)\right] \quad (9)$$

subject to

$$\sum_{\ell \in T_L} \sum_{k=1}^{n_i} z_{ik}^{\ell} \leq 1 \qquad i \in I \tag{10}$$

$$\sum_{k=1}^{n_i} v_{ik}^{\ell} = \sum_{\ell' \in T_L:\, \ell' \leq \ell} \sum_{k=1}^{n_i} z_{ik}^{\ell'} \qquad i \in I,\ \ell \in T_L \tag{11}$$

$$\sum_{k=1}^{n_i-1} e_{ik}^{\ell} + \sum_{k=1}^{n_i-1} r_{ik}^{\ell} \leq \sum_{\ell' \in T_L:\, \ell' < \ell} \sum_{k=1}^{n_i} z_{ik}^{\ell'} \qquad i \in I,\ \ell \in T_L \setminus \{1\} \tag{12}$$

$$\begin{aligned} \sum_{k=1}^{n_i} k\, v_{ik}^{\ell} = & \sum_{\ell' \in T_L:\, \ell' \leq \ell} \sum_{k=1}^{n_i} k\, z_{ik}^{\ell'} + \\ & \sum_{\ell' \in T_L \setminus \{1\}:\, \ell' \leq \ell} \sum_{k=1}^{n_i-1} k \left(e_{ik}^{\ell'} - r_{ik}^{\ell'} \right) \qquad i \in I,\ \ell \in T_L \end{aligned} \tag{13}$$

$$z_{ik}^{\ell} \in \{0,1\} \qquad i \in I,\ k = 1, \ldots, n_i,\ \ell \in T_L \tag{14}$$

$$v_{ik}^{\ell} \in \{0,1\} \qquad i \in I,\ k = 1, \ldots, n_i,\ \ell \in T_L \tag{15}$$

$$e_{ik}^{\ell},\ r_{ik}^{\ell} \in \{0,1\} \qquad i \in I,\ k = 1, \ldots, n_i - 1,\ \ell \in T_L \setminus \{1\} \tag{16}$$

with $\mathbb{E}_{\Xi}\left[Q(z,e,r,v,\xi)\right]$ denoting the recourse function, which represents the expected value of the second-stage problem. The optimal value of this problem, $Q(z,e,r,v,\xi)$, is given by

$$\begin{aligned} Q(z,e,r,v,\xi) = \text{Min} & \sum_{t \in T} \sum_{i \in I} \sum_{k=1}^{n_i} o_{ik}^{t}\, w_{ik}^{t}(\xi) + \sum_{t \in T} \sum_{i \in I} \sum_{j \in J^0} c_{ij}^{t}\, x_{ij}^{t}(\xi) + \\ & \sum_{t \in T} \sum_{i \in I} \sum_{j \in J^1} \sum_{t'=t}^{\min\{t+\rho_j,\, |T|\}} \left(c_{ij}^{t'} + p_j^{tt'} \right) y_{ij}^{tt'}(\xi) \end{aligned} \tag{17}$$

subject to

$$\sum_{i \in I} x_{ij}^{t}(\xi) = d_j^{t}(\xi) \qquad j \in J^0,\ t \in T \tag{18}$$

$$\sum_{i \in I} \sum_{t'=t}^{\min\{t+\rho_j,\, |T|\}} y_{ij}^{tt'}(\xi) = d_j^{t}(\xi) \qquad j \in J^1,\ t \in T \tag{19}$$

$$w_{ik}^{t}(\xi) \leq k\, Q\, v_{ik}^{\ell} \qquad i \in I,\ k = 1, \ldots, n_i,\ t \in T,\ \ell = \max\{\ell' \in T_L :\ \ell' \leq t\} \tag{20}$$

$$\sum_{k=1}^{n_i} w_{ik}^t(\xi) = \sum_{j \in J^0} x_{ij}^t(\xi) + \sum_{j \in J^1} \sum_{t'=\max\{1,\, t-\rho_j\}}^{t} y_{ij}^{t't}(\xi) \qquad i \in I,\ t \in T \tag{21}$$

$$x_{ij}^t(\xi) \geq 0 \qquad i \in I,\ j \in J^0,\ t \in T \tag{22}$$

$$y_{ij}^{tt'}(\xi) \geq 0 \qquad i \in I,\ j \in J^1,\ t \in T,\ t' = t, \ldots, \min\{t + \rho_j, |T|\} \tag{23}$$

$$w_{ik}^t(\xi) \geq 0 \qquad i \in I,\ k = 1, \ldots, n_i,\ t \in T \tag{24}$$

In the first-stage problem (9)–(16), the objective is to minimize the sum of fixed costs and expected processing and demand allocation costs. The fixed costs account for facility location, capacity expansion, capacity contraction, and facility operating costs. By including the expected total cost in (9), a neutral approach to risk is assumed, which is a common criterion to address risk in decision-making problems, often leading to computationally tractable models (Birge and Louveaux [13]). Constraints (10) ensure that at most one facility can be opened at a potential location over the planning horizon. Equalities (11) impose that facilities can only be operated provided they have previously been established. Constraints (12) guarantee that an open facility can be subjected to at most one type of capacity adjustment (either expansion or contraction) at each design period. The number of available modular units at open facilities is defined by equalities (13) for each design period. This number is the outcome of the sizing choice made when the facility was opened, and the number of modular units that were added or removed afterward. Finally, constraints (14)–(16) set the binary conditions on the first-stage (strategic) variables.

Every particular realization of customer demand ξ of Ξ yields a second-stage LP-model (17)–(24). Observe that in this model, the design variables z_{ik}^ℓ, v_{ik}^ℓ, e_{ik}^ℓ, and r_{ik}^ℓ take on fixed values. The objective function (17) minimizes the sum of processing costs at operating facilities, distribution costs to customers, and tardiness penalty costs for orders delivered with delay to customer segment J^1. Demand satisfaction is imposed by constraints (18) and (19). Capacity constraints are enforced by inequalities (20). Equalities (21) state that the outgoing flow from a facility at a given time period is split into deliveries to priority customers and to customers that receive late shipments. Non-negativity conditions on the second-stage (operational) variables are set by constraints (22)–(24). Observe that we are facing a stochastic problem with fixed recourse, since the coefficients of the second-stage variables (i.e., the unit cost parameters in (17) and the capacity of a modular unit in (20)) are all known in advance.

Recall that we assume that the random variable ξ has a finite support, i.e., ξ is defined by a finite probability distribution. Accordingly, let $s \in S$ be the index of the possible realizations of ξ (called *scenarios*), where S is a finite set, and let π_s represent the associated (positive) probabilities such that $\sum_{s\in S} \pi_s = 1$. In this case, $\mathbb{E}_{\boldsymbol{\Xi}}\left[Q(z, e, r, v, \xi)\right] = \sum_{s\in S} \pi_s \, Q(z, e, r, v, \xi_s)$. It follows that we can rewrite formulation (9)–(16) as a MILP model by defining scenario-indexed demands and recourse variables. The former are given by d^t_{js} ($t \in T$; $j \in J$; $s \in S$), while the latter replace the second-stage variables (6)–(8) and are represented by:

x^t_{ijs} : Amount of product distributed from facility i to customer j at time period t under scenario s $(i \in I;\ j \in J^0;\ t \in T;\ s \in S)$. $\quad (25)$

$y^{tt'}_{ijs}$: Amount of product distributed from facility i to customer j at time period t' to (partially) satisfy demand of period t under scenario s $(i \in I;\ j \in J^1;\ t \in T;\ t' = t, t+1, \ldots, \min\{t + \rho_j,\ |T|\};\ s \in S)$. $\quad (26)$

w^t_{iks} : Total quantity of product handled by facility i operating with k modular units at time period t under scenario s $(i \in I;\ k = 1, \ldots, n_i;\ t \in T;\ s \in S)$. $\quad (27)$

Using the above redefinition of the recourse variables, we obtain the so-called *deterministic equivalent* MILP model, or extensive form, hereafter denoted $(\mathbf{P}_{DE-1})$.

$$
\begin{aligned}
\text{Min} \ & \sum_{\ell\in T_L}\sum_{i\in I}\sum_{k=1}^{n_i} FO^\ell_{ik}\, z^\ell_{ik} + \sum_{\ell\in T_L\setminus\{1\}}\sum_{i\in I}\sum_{k=1}^{n_i-1} FE^\ell_{ik}\, e^\ell_{ik} + \\
& \sum_{\ell\in T_L\setminus\{1\}}\sum_{i\in I}\sum_{k=1}^{n_i-1} FR^\ell_{ik}\, r^\ell_{ik} + \sum_{\ell\in T_L}\sum_{i\in I}\sum_{k=1}^{n_i} v^\ell_{ik} \sum_{t=\ell}^{\phi(\ell)} M^t_{ik} + \\
& \sum_{s\in S}\pi_s \left[\sum_{t\in T}\sum_{i\in I}\sum_{k=1}^{n_i} o^t_{ik}\, w^t_{iks} + \sum_{t\in T}\sum_{i\in I}\sum_{j\in J^0} c^t_{ij}\, x^t_{ijs} + \right. \\
& \left. \sum_{t\in T}\sum_{i\in I}\sum_{j\in J^1}\sum_{t'=t}^{\min\{t+\rho_j,\,|T|\}} \left(c^{t'}_{ij} + p^{tt'}_j\right) y^{tt'}_{ijs} \right]
\end{aligned}
\tag{28}
$$

subject to

$$(10) - (16)$$

$$\sum_{i \in I} x_{ijs}^{t} \; = \; d_{js}^{t} \qquad j \in J^0, \; t \in T, \; s \in S \tag{29}$$

$$\sum_{i \in I} \sum_{t'=t}^{\min\{t+\rho_j, \, |T|\}} y_{ijs}^{tt'} \; = \; d_{js}^{t} \qquad j \in J^1, \; t \in T, \; s \in S \tag{30}$$

$$w_{iks}^{t} \; \leq \; k \, Q \, v_{ik}^{\ell} \qquad i \in I, \; k = 1, \ldots, n_i, \; t \in T, \; \ell = \max\{\ell' \in T_L : \ell' \leq t\}, \; s \in S \tag{31}$$

$$\sum_{k=1}^{n_i} w_{iks}^{t} \; = \; \sum_{j \in J^0} x_{ijs}^{t} + \sum_{j \in J^1} \sum_{t'=\max\{1, \, t-\rho_j\}}^{t} y_{ijs}^{t't} \qquad i \in I, \; t \in T, \; s \in S \tag{32}$$

$$x_{ijs}^{t} \geq 0 \qquad i \in I, \; j \in J^0, \; t \in T, \; s \in S \tag{33}$$

$$y_{ijs}^{tt'} \geq 0 \qquad i \in I, \; j \in J^1, \; t \in T, \; t' = t, \ldots, \min\{t + \rho_j, |T|\}, \; s \in S \tag{34}$$

$$w_{iks}^{t} \geq 0 \qquad i \in I, \; k = 1, \ldots, n_i, \; t \in T, \; s \in S \tag{35}$$

We note that the extensive form of the stochastic program becomes quite large even when the random variable ξ has a moderate number of possible realizations. Since all the strategic decisions made in the first stage do not depend on the realization of the random demand process, and so the location and capacity scalability schedule determined for the entire planning horizon is the same for all scenarios, formulation $(\mathbf{P}_{DE-1})$ satisfies the non-anticipativity principle. For this reason, $(\mathbf{P}_{DE-1})$ will be called the *scenario-independent* location and capacity scalability model.

2.3 A stochastic formulation for an alternative strategy

Discrepancies between the capacity of a firm and the demands of its customers result in inefficiency, either in underutilised resources or dissatisfied customers. The ability of a firm to adjust the capacity of its resources in response to varying customer demand is one of the key factors in decreasing these discrepancies and achieving competitiveness. Therefore, an alternative strategy to the settings considered in the previous section is to assume that capacity adjustment decisions have a tactical nature, and thus can be deferred to the second-stage problem.

In this strategy, the first-stage problem consists of defining a schedule for locating facilities and setting their initial capacities for the entire planning horizon. Accordingly, the strategic facility location decisions are the *here-and-now* decisions. The second-stage problem addresses the *wait-and-see* (tactical and operational) decisions by prescribing a scheme for adjusting the capacities of operating facilities (through expansion or contraction), and specifying the product flows from facilities to customers over all time periods. In contrast to the approach presented in Section 2.2, the capacity adjustment measures depend now on the realization of the uncertain customer demand. As mentioned earlier, the resulting two-stage stochastic program is suitable for those cases in which sizing decisions can be reverted in the medium-term. This arises, for example, in the context of leasing or renting space and equipment. Furthermore, subcontracting and adjusting labor (e.g., using temporary workers or furlough) are also measures that can be implemented in medium-range capacity planning.

To formulate a two-stage stochastic program for this alternative strategy, the binary location variables (2) and the continuous flow variables (6), (7), and (8) are used. Moreover, the binary variables associated with the sizing decisions, i.e., (3), (4), and (5), are replaced by $\widetilde{e}^{\ell}_{ik}(\xi)$, $\widetilde{r}^{\ell}_{ik}(\xi)$, and $\widetilde{v}^{\ell}_{ik}(\xi)$, respectively. Each one of these new variables depends on the random vector ξ ruling future customer demands. The total expected cost to be minimized is given by $\sum_{\ell\in T_L}\sum_{i\in I}\sum_{k=1}^{n_i} FO^{\ell}_{ik}\, z^{\ell}_{ik} + \mathbb{E}_{\Xi}[Q'(z,\xi)]$ under constraints (10) and (14). The optimal value $Q'(z,\xi)$ of the second-stage problem is determined by

$$
\begin{aligned}
Q'(z,\xi) = \text{Min} \; & \sum_{\ell\in T_L\setminus\{1\}}\sum_{i\in I}\sum_{k=1}^{n_i-1} FE^{\ell}_{ik}\,\widetilde{e}^{\ell}_{ik}(\xi) + \sum_{\ell\in T_L\setminus\{1\}}\sum_{i\in I}\sum_{k=1}^{n_i-1} FR^{\ell}_{ik}\,\widetilde{r}^{\ell}_{ik}(\xi) + \\
& \sum_{\ell\in T_L}\sum_{i\in I}\sum_{k=1}^{n_i}\widetilde{v}^{\ell}_{ik}(\xi)\sum_{t=\ell}^{\phi(\ell)} M^t_{ik} + \sum_{t\in T}\sum_{i\in I}\sum_{k=1}^{n_i} o^t_{ik}\, w^t_{ik}(\xi) + \\
& \sum_{t\in T}\sum_{i\in I}\sum_{j\in J^0} c^t_{ij}\, x^t_{ij}(\xi) + \sum_{t\in T}\sum_{i\in I}\sum_{j\in J^1}\sum_{t'=t}^{\min\{t+\rho_j,\,|T|\}}\left(c^{t'}_{ij} + p^{tt'}_j\right) y^{tt'}_{ij}(\xi)
\end{aligned}
\tag{36}
$$

subject to

$$(18), (19), (21) - (24)$$

$$\sum_{k=1}^{n_i}\widetilde{v}^{\ell}_{ik}(\xi) = \sum_{\ell'\in T_L:\,\ell'\leq\ell}\;\sum_{k=1}^{n_i} z^{\ell'}_{ik} \qquad i\in I,\ \ell\in T_L \tag{37}$$

$$\sum_{k=1}^{n_i-1}\widetilde{e}^{\ell}_{ik}(\xi) + \sum_{k=1}^{n_i-1}\widetilde{r}^{\ell}_{ik}(\xi) \leq \sum_{\ell'\in T_L:\,\ell'<\ell}\;\sum_{k=1}^{n_i} z^{\ell'}_{ik} \qquad i\in I,\ \ell\in T_L\setminus\{1\} \tag{38}$$

$$\sum_{k=1}^{n_i} k\, \widetilde{v}_{ik}^{\ell}(\xi) \;=\; \sum_{\ell' \in T_L:\, \ell' \leq \ell} \; \sum_{k=1}^{n_i} k\, z_{ik}^{\ell'} + \sum_{\ell' \in T_L \setminus \{1\}:\, \ell' \leq \ell} \; \sum_{k=1}^{n_i-1} k\, \left(\widetilde{e}_{ik}^{\ell'}(\xi) \,-\, \widetilde{r}_{ik}^{\ell'}(\xi)\right) \qquad i \in I,\; \ell \in T_L \tag{39}$$

$$w_{ik}^{t}(\xi) \;\leq\; k\, Q\, \widetilde{v}_{ik}^{\ell}(\xi) \qquad i \in I,\; k = 1, \ldots, n_i,\; t \in T,\; \ell = \max\{\ell' \in T_L :\; \ell' \leq t\} \tag{40}$$

$$\widetilde{v}_{ik}^{\ell}(\xi) \in \{0, 1\} \qquad i \in I,\; k = 1, \ldots, n_i,\; \ell \in T_L \tag{41}$$

$$\widetilde{e}_{ik}^{\ell}(\xi),\; \widetilde{r}_{ik}^{\ell}(\xi) \in \{0, 1\} \qquad i \in I,\; k = 1, \ldots, n_i - 1,\; \ell \in T_L \setminus \{1\} \tag{42}$$

Due to capacity adjustment decisions being now recourse decisions, the new constraints (37)–(39) are the counterpart of constraints (11)–(13). Likewise, the capacity constraints (40) replace inequalities (20).

By imposing the same conditions on the random demand vector ξ as in Section 2.2, we can rewrite the recourse function as a function of all scenarios, i.e., $\mathbb{E}_{\Xi}[Q'(z, \xi)] = \sum_{s \in S} \pi_s\, Q'(z, \xi_s)$. Accordingly, the set of scenario-dependent flow variables (25)–(27) is extended with the new scenario-indexed capacity scalability variables, $\widetilde{v}_{iks}^{\ell}$, $\widetilde{e}_{iks}^{\ell}$, and $\widetilde{r}_{iks}^{\ell}$. The definition of these variables and the resulting deterministic equivalent MILP formulation, hereafter denoted $(\mathbf{P}_{DE-2})$, are detailed in Appendix A. Since all capacity adjustment decisions depend on the realizations of the random parameters, $(\mathbf{P}_{DE-2})$ will be called the *scenario-dependent* capacity scalability model. $(\mathbf{P}_{DE-2})$ defines a large-scale MILP model with a significant number of binary variables and constraints, even for a small-sized scenario set.

3 Additional inequalities

In an attempt to improve the polyhedral description of the set of feasible solutions of models $(\mathbf{P}_{DE-1})$ and $(\mathbf{P}_{DE-2})$, we develop in this section several sets of additional inequalities. The proposed enhancements may contribute to more problem instances and with larger sizes being solved to (near-)optimality with off-the-shelf, general-purpose optimization software within reasonable computing time.

3.1 Inequalities for model ($\mathbf{P}_{DE-1}$)

Recall from (28)–(35) that in the extensive form of the deterministic equivalent model ($\mathbf{P}_{DE-1}$), all facility location and capacity-related decisions are made in the first stage, before uncertainty about future demand is disclosed. Next, we successively show how the information conveyed by the various scenarios can be used in this case to derive a lower bound on the total number of facilities that must be open at a particular design period.

Let us assume, without loss of generality, that the length of the planning horizon is a multiple of the number of design periods, and let us define $\sigma = |T|/|T_L|$. Hence, parameter σ is integer-valued and gives the total number of periods between two consecutive design periods. Since all demand requirements of the preferred customer segment J^0 must be serviced without delay in every scenario, they are used to set minimum capacity requirements between two consecutive design periods. The latter are given by

$$D^\ell \;=\; \max\left\{\sum_{j\in J^0} d^t_{js} : \; s \in S;\; t = \ell,\ldots,\phi(\ell)\right\},$$

for every $\ell \in T_L$ and $\phi(\ell)$ defined in (1). Hence, the following inequalities must hold for every design period:

$$Q\sum_{i\in I}\sum_{k=1}^{n_i} k\, v^\ell_{ik} \;\geq\; D^\ell, \qquad \ell \in T_L.$$

Dividing the above inequalities by the size of a modular unit, Q, and taking into account that the left-hand side must be integer-valued, it follows that

$$\sum_{i\in I}\sum_{k=1}^{n_i} k\, v^\ell_{ik} \;\geq\; \left\lceil \frac{D^\ell}{Q} \right\rceil, \qquad \ell \in T_L. \tag{43}$$

Furthermore, for a particular design period ℓ ($\ell \in T_L$), let $\widetilde{D}^\ell$ denote the largest minimum demand quantity over all scenarios that has to be covered from period ℓ through period $\phi(\ell)$.

$$\widetilde{D}^\ell = \begin{cases} \max\limits_{s\in S}\left\{\sum\limits_{j\in J^0}\sum\limits_{t=\ell}^{\phi(\ell)} d^t_{js} + \sum\limits_{j\in J^1}\sum\limits_{t=\ell}^{\phi(\ell)-\rho_j} d^t_{js}\right\} & \text{if } \ell < \ell_{max}, \\[2ex] \max\limits_{s\in S}\left\{\sum\limits_{j\in J}\sum\limits_{t=\ell}^{|T|} d^t_{js}\right\} & \text{if } \ell = \ell_{max}. \end{cases}$$

The above expression includes all orders from customer segment J^0 as well as the minimum quantity of demand that must be delivered to customers J^1 accepting late shipments. Since demand cannot be lost, in the last time interval (comprising all time periods from ℓ_{max} through $|T|$), the orders of both customer segments must be completely satisfied. Accordingly, the following inequalities impose that facilities open at time period ℓ (and therefore at all subsequent periods up to period $\phi(\ell)$) must have sufficient capacity to serve the demand $\widetilde{D}^\ell$:

$$\sigma\, Q \sum_{i\in I}\sum_{k=1}^{n_i} k\, v_{ik}^\ell \;\; \geq \;\; \widetilde{D}^\ell, \qquad \ell \in T_L.$$

Dividing these inequalities, first by Q and then by the time lag σ, we obtain

$$\sum_{i\in I}\sum_{k=1}^{n_i} k\, v_{ik}^\ell \;\; \geq \;\; \left\lceil \frac{\left\lceil \frac{\widetilde{D}^\ell}{Q} \right\rceil}{\sigma} \right\rceil, \qquad \ell \in T_L. \tag{44}$$

Combining inequalities (43) and (44), and defining

$$M^\ell \;\; = \;\; \max\left\{ \left\lceil \frac{D^\ell}{Q} \right\rceil, \left\lceil \frac{\left\lceil \frac{\widetilde{D}^\ell}{Q} \right\rceil}{\sigma} \right\rceil \right\}$$

yields

$$\sum_{i\in I}\sum_{k=1}^{n_i} k\, v_{ik}^\ell \;\; \geq \;\; M^\ell, \qquad \ell \in T_L.$$

By applying the Chvátal-Gomory rounding procedure a finite number of times p ($p \geq 1$) to the above inequalities (Wolsey [50]), we obtain

$$\sum_{i\in I}\sum_{k=1}^{n_i} \left\lceil \frac{k}{p} \right\rceil \, v_{ik}^\ell \;\; \geq \;\; \left\lceil \frac{M^\ell}{p} \right\rceil, \qquad \ell \in T_L,\; p = 1, \ldots, \max_{i\in I}\{n_i\}, \tag{45}$$

The reasoning followed to derive (45) is itself a proof for the following result:

Proposition 1. *The set of inequalities* (45) *is valid for formulation* ($\mathbf{P}_{DE-1}$).

The enhanced formulation is called hereafter ($\mathbf{P}^+_{DE-1}$) for the scenario-independent location and capacity scalability model.

3.2 Inequalities for model ($\mathbf{P}_{DE-2}$)

For the extensive form of the deterministic model given by (52)–(58) (see Appendix A), we develop additional inequalities that serve a purpose similar to (45). In this case, the variables $\widetilde{v}_{iks}^{\ell}$ (cf. (51)) representing the number of modular units available in a location at time period ℓ ($\ell \in T_L$) are scenario-indexed. Hence, the minimum quantity of demand that must be satisfied in each period between two consecutive design periods is scenario-dependent and defined by

$$\Delta_s^{\ell} \;=\; \max\left\{\sum_{j\in J^0} d_{js}^t : t = \ell,\ldots,\phi(\ell)\right\}, \qquad \ell \in T_L,\ s \in S.$$

Given a scenario s and a design period ℓ, it follows that the minimum number of open facilities must satisfy the following conditions:

$$\sum_{i\in I}\sum_{k=1}^{n_i} k\,\widetilde{v}_{iks}^{\ell} \;\geq\; \left\lceil \frac{\Delta_s^{\ell}}{Q} \right\rceil, \qquad \ell \in T_L,\ s \in S. \tag{46}$$

The above inequalities are the counterpart of (43) for the scenario-dependent capacity scalability model.

Additionally, let $\widetilde{\Delta}_s^{\ell}$ denote the minimum total demand requirements to be satisfied in scenario s for all periods from ℓ through $\phi(\ell)$, with $\ell \in T_L$.

$$\widetilde{\Delta}_s^{\ell} \;=\; \begin{cases} \displaystyle\sum_{j\in J^0}\sum_{t=\ell}^{\phi(\ell)} d_{js}^t + \sum_{j\in J^1}\sum_{t=\ell}^{\phi(\ell)-\rho_j} d_{js}^t & \text{if } \ell < \ell_{max}, \\[2ex] \displaystyle\sum_{j\in J}\sum_{t=\ell}^{|T|} d_{js}^t & \text{if } \ell = \ell_{max}. \end{cases}$$

Therefore, the following inequalities must also hold

$$\sum_{i\in I}\sum_{k=1}^{n_i} k\,\widetilde{v}_{iks}^{\ell} \;\geq\; \left\lceil \frac{\left\lceil \frac{\widetilde{\Delta}_s^{\ell}}{Q} \right\rceil}{\sigma} \right\rceil, \qquad \ell \in T_L,\ s \in S. \tag{47}$$

Using inequalities (46) and (47) yields the following relations

$$\sum_{i\in I}\sum_{k=1}^{n_i} k\,\widetilde{v}_{iks}^{\ell} \;\geq\; \widetilde{M}_s^{\ell}, \qquad \ell \in T_L,\ s \in S,$$

where the lower bound $\widetilde{M}_s^\ell$ is defined as the maximum of the two parameters on the right-hand sides of inequalities (46) and (47). Finally, we apply once again the Chvátal-Gomory rounding method to the above inequalities and obtain the following result:

Proposition 2. *The set of inequalities*

$$\sum_{i \in I} \sum_{k=1}^{n_i} \left\lceil \frac{k}{p} \right\rceil \widetilde{v}_{iks}^{\ell} \;\; \geq \;\; \left\lceil \frac{\widetilde{M}_s^\ell}{p} \right\rceil, \qquad \ell \in T_L, \; s \in S, \; p = 1, \ldots, \max_{i \in I}\{n_i\} \tag{48}$$

is valid for formulation $(\mathbf{P}_{DE-2})$.

The steps presented above serve as proof for this proposition. If for a particular design period ℓ and scenario s there exist several inequalities (48), all having the same right-hand side for different values of p, then we only need to consider the inequality with the strongest left-hand side. A similar remark also applies to (45). The enhanced formulation for the scenario-dependent capacity scalability case is called hereafter $(\mathbf{P}_{DE-2}^{+})$.

4 Computational results

In this section, we present the results of our numerical tests. The computational experiments were guided by five objectives, namely: (i) to evaluate the usefulness of using a standard off-the-shelf MILP solver to identify optimal or near-optimal solutions to the original deterministic equivalent models within reasonable computing time; (ii) to analyze the effectiveness of the proposed additional inequalities; (iii) to compare the two strategies for deciding on the timing and sizing of capacity adjustments; (iv) to discuss relevant insights into the characteristics of the (near-)optimal solutions identified; and (v) to assess the benefits of using a stochastic programming approach for the problem under study.

4.1 Characteristics of test instances

Since benchmark data sets are not available for the problem at hand, we have generated a set of test instances. The size of each instance is mainly dictated by the length of the planning horizon and the total number of customers according to the choices specified in Table 1. We have considered planning horizons with 12 and 24 time periods. In both cases, there are three design periods for making location and capacity scalability decisions. Specifically, for $|T| = 12$

(resp. $|T| = 24$), these opportunities occur at the beginning of periods 1, 5, and 9 (resp. 1, 9, and 17). Three different sizes for the customer set are assumed and three different partitions of J are considered. For example, a set of 60 customers can be partitioned into $|J^0| = 15$ and $|J^1| = 45$, $|J^0| = |J^1| = 30$, and $|J^0| = 45$ and $|J^1| = 15$. In all cases, customers in segment J^1 specify the same maximum number of periods for late deliveries, i.e., $\rho_j = \rho$ for all $j \in J^1$. Instances with $\rho = 1$, $\rho = 2$, and $\rho = 3$ were generated.

Parameter	Value	Parameter	Value
$\lvert T\rvert$	12, 24	$\lvert T_L\rvert$	3
$\lvert J\rvert$	20, 40, 60	$\lvert J^0\rvert$	$\beta^J\,\lvert J\rvert$ with $\beta^J \in \{0.25, 0.5, 0.75\}$
$\lvert I\rvert$	$0.25\,\lvert J\rvert$	ρ	1, 2, 3
$\lvert S\rvert$	3	π_s	$1/3$ $(s \in S)$

Table 1: Parameter values.

In the deterministic variant of our problem [15], three different patterns of demand fluctuation over the planning horizon were considered. In the present work, we follow this setting by assuming three different demand scenarios, an approach that is also followed in other works, e.g. Marín et al. [39]. In addition, equal probabilities are assigned to the scenarios. In all scenarios, the demand of each customer for the first time period is selected at random from the interval $[20, 500]$ according to a continuous uniform distribution. In scenario 1, and at each subsequent period, demands exhibit fluctuations ranging from -20% to $+20\%$ compared to the previous period. Therefore, this scenario represents an irregular demand pattern. In scenarios 2 and 3, demand variations go through three phases, showing a trapezoidal shape. Specifically, scenario 2 starts with a demand growth phase, followed by a maturity phase with small fluctuations (i.e., $\pm 1\%$), and ends with a decline phase. Scenario 3 has the opposite demand pattern; the associated inverted trapezoidal structure depicts demand contraction, recession, and gradual recovery. Table 11 in Appendix B details the three scenarios. These scenarios differ significantly from each other with respect to demand variations, making them a suitable choice in a setting with uncertain demand.

The capacity of a modular unit is defined as a function of the demand scenarios, namely $Q = \frac{1}{n_i\,|I|}\,\mathcal{U}[3,\,4]\,\frac{\sum_{j\in J}\sum_{t\in T}\max_{s\in S} d_{js}^t}{|T|}$, where $\mathcal{U}[3,\,4]$ denotes a continuous uniform distribution in $[3,\,4]$ and $n_i = 4$ for every $i \in I$. Appendix B provides further details on the procedure for randomly generating all cost parameters. Five instances were randomly generated for each combination of parameters shown in Table 1, giving rise to a total of 270 different instances.

4.2 Numerical results

Formulations ($\mathbf{P}_{DE-1}$) and ($\mathbf{P}_{DE-2}$) along with their enhancements were coded in C++ using IBM ILOG Concert Technology and run with IBM ILOG CPLEX 12.6. All experiments were performed on a workstation with a multi-core Intel Xeon E5-2650V3 processor (2.3 GHz, 10 cores), 32 GB RAM, and running the Ubuntu operating system (64-bit). Due to the strategic nature of the problem that we study, fast solution times are not of paramount importance. Therefore, a limit of 10 hours of CPU time was set for each solver run. Furthermore, CPLEX was used with default settings under a deterministic parallel mode.

Regarding goal (i) listed at the beginning of Section 4, we evaluate the effectiveness of using CPLEX to identify (near-)optimal solutions to the test instances for the proposed original formulations, ($\mathbf{P}_{DE-1}$) and ($\mathbf{P}_{DE-2}$). To do so, we compare the optimality gaps and the computing times reported by the optimization solver. Table 2 summarizes the relevant information for the different sizes of the customer set, while Table 3 provides information from a different perspective, namely for varying values for the maximum delivery delay tolerated by customer segment J^1. Column 1 gives the parameter selected ($|J|$ in Table 2 and ρ in Table 3). Columns 2–5 report the number of instances solved to optimality (# opt sol.) and the number of instances not solved to proven optimality within the specified time limit (# non-opt sol.) for all models (i.e., the original formulations and their enhancements). For those instances not solved optimally, the minimum, average, and maximum optimality gaps achieved by CPLEX after 10 h of computing time are given for each formulation in columns 7–10 (MIP gap (%) $=$ $(z^{UB} - z^{LB})/z^{UB} \times 100\%$, with z^{UB} denoting the objective value of the best feasible solution available and z^{LB} representing the objective value of the best lower bound identified during the branch-and-cut procedure). Columns 11–14 display the minimum, average, and maximum computing times, in seconds (MIP CPU (sec.)). The last row refers to all 270 instances. The best (average) values with respect to the evaluation criteria are given in boldface.

Table 2 shows that the capability of CPLEX to identify an optimal solution within the given time limit is affected by the formulation that is used. Regarding formulations ($\mathbf{P}_{DE-1}$) and ($\mathbf{P}_{DE-2}$), optimality is achieved in all test instances with 20 customers, but gradually decreases as the total number of customers grows. Not surprisingly, the even larger formulation ($\mathbf{P}_{DE-2}$) poses an additional challenge to CPLEX. In particular, for this formulation, optimal solutions are only available for less than half of the instances with 60 customers (44% or 40/90), whereas 74% (67/90) of the instances with the same number of customers could be solved to optimality with

$\lvert J \rvert$	# opt sol./# non-opt sol.					MIP gap (%)*				MIP CPU (sec.)			
	$(\mathbf{P}_{DE-1})$	$(\mathbf{P}^{+}_{DE-1})$	$(\mathbf{P}_{DE-2})$	$(\mathbf{P}^{+}_{DE-2})$		$(\mathbf{P}_{DE-1})$	$(\mathbf{P}^{+}_{DE-1})$	$(\mathbf{P}_{DE-2})$	$(\mathbf{P}^{+}_{DE-2})$	$(\mathbf{P}_{DE-1})$	$(\mathbf{P}^{+}_{DE-1})$	$(\mathbf{P}_{DE-2})$	$(\mathbf{P}^{+}_{DE-2})$
20	90/0	90/0	90/0	90/0	min	-	-	-	-	5.7	2.5	7.1	2.3
					avg	-	-	-	-	27.7	13.9	68.0	32.0
					max	-	-	-	-	201.1	46.2	477.6	140.3
40	87/3	90/0	71/19	84/6	min	0.89	-	1.15	1.59	22.9	16.5	105.5	46.7
					avg	3.26	-	4.04	4.43	3462.4	1198.5	11333.7	6615.7
					max	5.32	-	7.12	6.69	**	30074.9	**	**
60	67/23	75/15	40/50	51/39	min	1.47	0.27	0.62	0.32	70.0	61.2	111.4	71.4
					avg	2.43	2.28	2.29	2.04	13072.2	10699.9	24198.9	20359.2
					max	3.41	4.26	4.81	5.31	**	**	**	**
All	244/26	**255/15**	201/69	**225/45**	avg	2.53	**2.28**	2.77	**2.36**	5520.8	**3970.8**	11866.9	**9002.3**

Table 2: Optimality gaps and CPU times for all formulations according to the total number of customers; *MIP gap for instances not solved to optimality within 10 h; **time limit (10 h) reached.

ρ	# opt sol./# non-opt sol.					MIP gap (%)*				MIP CPU (sec.)			
	$(\mathbf{P}_{DE-1})$	$(\mathbf{P}^{+}_{DE-1})$	$(\mathbf{P}_{DE-2})$	$(\mathbf{P}^{+}_{DE-2})$		$(\mathbf{P}_{DE-1})$	$(\mathbf{P}^{+}_{DE-1})$	$(\mathbf{P}_{DE-2})$	$(\mathbf{P}^{+}_{DE-2})$	$(\mathbf{P}_{DE-1})$	$(\mathbf{P}^{+}_{DE-1})$	$(\mathbf{P}_{DE-2})$	$(\mathbf{P}^{+}_{DE-2})$
1	78/12	83/7	61/29	70/20	min	0.89	0.78	0.62	0.42	10.0	6.8	12.1	4.0
					avg	2.50	2.37	3.24	2.71	7314.4	5276.9	14269.4	11086.2
					max	3.57	4.26	7.12	6.69	**	**	**	**
2	84/6	86/4	68/22	78/12	min	2.19	0.27	0.69	0.32	9.6	2.5	9.1	2.3
					avg	2.49	1.84	2.44	2.00	4578.7	2818.0	11102.5	7423.6
					max	2.80	2.63	6.99	3.88	**	**	**	**
3	82/8	86/4	72/18	77/13	min	1.47	2.30	0.62	0.52	5.7	3.0	7.1	3.0
					avg	2.59	2.55	2.45	2.13	4669.1	3817.4	10228.7	8497.1
					max	5.32	2.89	6.20	5.12	**	**	**	**
All	244/26	**255/15**	201/69	**225/45**	avg	2.53	**2.28**	2.77	**2.36**	5520.8	**3970.8**	11866.9	**9002.3**

Table 3: Optimality gaps and CPU times for all formulations under different values for the maximum delivery delay, ρ; *MIP gap for instances not solved to optimality within 10 h; **time limit (10 h) reached.

model $(\mathbf{P}_{DE-1})$. These findings are also supported by the growth observed in the computing times. Interestingly, for those instances that could not be solved to proven optimality, the associated gaps are, on average, relatively small (cf. columns 7 and 9). The challenges posed by formulation $(\mathbf{P}_{DE-2})$ result in slightly larger gaps compared to formulation $(\mathbf{P}_{DE-1})$.

Table 2 also reveals the positive effect of the proposed valid inequalities (45) and (48), an aspect that was mentioned at the beginning of Section 4 as goal (ii) to be investigated. Specifically, the additional constraints prove to be very useful in identifying more optimal solutions in shorter computing times. Adding inequalities (45) to formulation $(\mathbf{P}_{DE-1})$ results in obtaining optimal solutions to 94% (255/270) of the instances. Moreover, CPLEX does not achieve optimality with formulation $(\mathbf{P}^{+}_{DE-1})$ for only 17% (15/90) of the instances with 60 customers. Regarding the enhanced model $(\mathbf{P}^{+}_{DE-2})$, similar findings can be observed despite the greater size of this formulation, with 83.3% (225/270) instances being solved to optimality.

The optimality gaps of the best solutions obtained after 10 h of CPU also decrease using inequalities (45) and (48). The only exception occurs for the set of instances with 40 customers, for which CPLEX reports an average optimality gap of 4.43% with formulation $(\mathbf{P}^{+}_{DE-2})$ against a gap of 4.04% without using the proposed inequalities. Observe that the latter value is obtained for 19 instances, whereas the former value results from six instances. If we consider the same set of six instances in both formulations, we realize that they exhibit an average optimality gap of 4.94% with model $(\mathbf{P}_{DE-1})$. Therefore, in fact, the valid inequalities (48) succeed in reducing this gap to 4.43%.

Table 3 shows that increasing the maximum lead time for customers accepting late shipments results in slightly less challenging problems for CPLEX. This fact is evidenced by the larger number of instances not solved to proven optimality for $\rho = 1$, as opposed to $\rho = 2$ or $\rho = 3$. We remark that increasing the value of ρ brings more flexibility to designing the facility network and adjusting the capacities of the operating facilities over the planning horizon. Interestingly, the quality of the best solutions returned by CPLEX upon reaching the time limit does not seem to be much affected by parameter ρ. The enhanced formulations succeed in solving more instances to optimality and in providing lower optimality gaps with a reduced computational burden.

As evidenced by the significantly shorter computing times of the tighter formulations $(\mathbf{P}^{+}_{DE-1})$ and $(\mathbf{P}^{+}_{DE-2})$ (cf. last column in Tables 2–3), the branch-and-cut tree generated by CPLEX is smaller, despite the greater number of constraints. However, additional computational effort may be required to solve the associated linear relaxations. This feature is present in solving the

LP-relaxation of model $(\mathbf{P}^{+}_{DE-2})$ but not for the LP-relaxation of model $(\mathbf{P}^{+}_{DE-1})$, as displayed in Tables 4 and 5. These tables report in the last four columns the minimum, average, and maximum solution times of the linear relaxation (LP CPU (sec.)) of all models, for different values of parameter $|J|$ (Table 4) and parameter ρ (Table 5). In addition, columns 3–6 show the minimum, average, and maximum integrality gaps (LP gap (%)). The latter give the relative percent deviation between the objective value of the best feasible solution available (z^{UB}) and the lower bound provided by the linear relaxation. The best average values over all 270 instances are highlighted in boldface in the last row of these tables.

$\|J\|$		LP gap (%)				LP CPU (sec.)			
		$(\mathbf{P}_{DE-1})$	$(\mathbf{P}^{+}_{DE-1})$	$(\mathbf{P}_{DE-2})$	$(\mathbf{P}^{+}_{DE-2})$	$(\mathbf{P}_{DE-1})$	$(\mathbf{P}^{+}_{DE-1})$	$(\mathbf{P}_{DE-2})$	$(\mathbf{P}^{+}_{DE-2})$
20	min	4.69	1.66	3.68	1.87	0.1	0.1	0.2	0.2
	avg	9.79	4.33	9.60	4.45	2.0	1.1	1.8	1.2
	max	16.18	6.69	16.91	8.33	14.3	13.3	7.2	7.0
40	min	1.50	0.87	1.80	1.47	2.7	0.8	3.1	2.7
	avg	3.75	2.65	4.16	3.19	17.2	10.2	16.3	19.0
	max	9.29	9.24	9.60	9.60	46.2	43.4	31.8	55.7
60	min	0.71	0.69	1.22	0.83	4.8	4.5	6.1	6.6
	avg	3.04	2.48	3.18	2.73	33.8	33.8	17.2	27.0
	max	6.38	6.30	6.76	5.73	124.5	98.5	62.5	100.0
All	avg	5.53	**3.15**	5.65	**3.46**	17.7	**15.0**	**11.8**	13.3

Table 4: Integrality gaps and solution times of the LP relaxation of all formulations according to the total number of customers.

Tables 4 and 5 indicate that the linear relaxations of the enhanced formulations provide substantially better lower bounds, and as a result lower integrality gaps, as opposed to the original models. Even though the number of available optimal solutions decreases with the size of J, and accordingly the quality of z^{UB} declines, the average integrality gap becomes smaller when inequalities (45) and (48) are added. In contrast, the magnitude of the maximum allowed tardiness for orders in customer segment J^1 does not seem to affect the average integrality gap in any of the four formulations, as shown in Table 5. It is also noticeable that the reduction of the average integrality gap is more pronounced for instances with 20 customers ($> 50\%$) than for instances with a larger number of customers.

Finally, we note that tight integrality gaps such as the ones we have obtained with the enhanced formulations can be very useful in evaluating the quality of a feasible solution produced, for example, by a tailored heuristic method. In particular, this feature is relevant for problems with large customer sets since in this case it becomes increasingly difficult for a state-of-the-art optimization solver to achieve optimality within acceptable time. From a computational point of view, the optimal solution to the linear relaxation is inexpensive ($\leq$ 100 seconds for the enhanced models, according to Tables 4–5).

ρ		LP gap (%)				LP CPU (sec.)			
		$(\mathbf{P}_{DE-1})$	$(\mathbf{P}^{+}_{DE-1})$	$(\mathbf{P}_{DE-2})$	$(\mathbf{P}^{+}_{DE-2})$	$(\mathbf{P}_{DE-1})$	$(\mathbf{P}^{+}_{DE-1})$	$(\mathbf{P}_{DE-2})$	$(\mathbf{P}^{+}_{DE-2})$
1	min	0.71	0.69	1.52	1.27	0.1	0.1	0.2	0.2
	avg	5.03	3.46	6.25	3.85	14.6	14.5	10.5	13.3
	max	14.76	9.24	16.91	9.60	63.8	56.8	33.5	48.4
2	min	0.71	0.69	1.22	0.83	0.1	0.2	0.2	0.2
	avg	5.37	2.84	5.37	3.18	17.7	14.0	11.4	14.9
	max	15.80	5.85	16.76	5.79	114.6	82.1	32.6	50.0
3	min	0.71	0.69	1.45	1.04	0.2	0.2	0.2	0.2
	avg	5.37	3.16	5.31	3.34	20.8	16.6	13.4	19.0
	max	16.18	6.86	16.53	7.43	124.5	98.5	62.5	100.0
All	avg	5.53	**3.15**	5.65	**3.46**	17.7	**15.0**	**11.8**	13.3

Table 5: Integrality gaps and solution times of the LP relaxation of all formulations under different values for the maximum delivery delay, ρ.

4.3 Deployment of facility location and capacity scalability strategies

In this section, we focus on goals (iii) and (iv) stated at the beginning of Section 4, and compare the characteristics of the two strategies proposed for locating facilities and adjusting their capacities over the planning horizon. For this purpose, Table 6 gives the minimum, average, and maximum relative deviation of the objective values of the best solutions reported by CPLEX for the enhanced formulations. Specifically, columns 2–4 (resp. 6–8) display the value of $(z^{+}_{DE-1} - z^{+}_{DE-2})/z^{+}_{DE-1} \times 100\%$ according to the size of the customer set (resp. the maximum delivery delay). We denote by z^{+}_{DE-1} and z^{+}_{DE-2} the best objective values identified by CPLEX to formulation $(\mathbf{P}^{+}_{DE-1})$ and $(\mathbf{P}^{+}_{DE-2})$, respectively, within the given time limit. Each row aggregates the results for 90 instances. Since the relative deviations presented in this table are all positive, they indicate that cost savings can be achieved with the scenario-dependent capacity scalability model (strategy 2) over the scenario-independent location and scalability model (strategy 1). This finding is not surprising because modeling capacity adjustment decisions as second-stage variables in formulation $(\mathbf{P}^{+}_{DE-2})$ allows a more agile response to temporal changes in demand. In other words, these decisions can be met taking into account the realization of demand. Therefore, in general, the resulting schedule for facility opening and capacity adjustment is cheaper compared to the case in which all design and capacity scalability decisions must be made in the first stage. In the latter case, such decisions do not depend on the particular realization of the random demand process.

$\|J\|$	% Cost reduction of $(\mathbf{P}^+_{DE-2})$ over $(\mathbf{P}^+_{DE-1})$			ρ	% Cost reduction of $(\mathbf{P}^+_{DE-2})$ over $(\mathbf{P}^+_{DE-1})$		
	Min	Avg	Max		Min	Avg	Max
20	0.00	3.32	5.90	1	0.04	2.84	5.23
40	0.33	2.64	6.19	2	0.00	3.07	6.19
60	0.04	3.01	5.80	3	0.00	3.07	6.19

Table 6: Comparison between the best objective values of formulations $(\mathbf{P}^+_{DE-1})$ and $(\mathbf{P}^+_{DE-2})$ for different parameters.

Over all 270 instances, strategy 2 allows for up to 6.19% reduction in total cost, with an average reduction of 3%. Even when an instance is not solved to optimality with formulation $(\mathbf{P}^+_{DE-2})$ and an optimal solution is available to formulation $(\mathbf{P}^+_{DE-1})$, strategy 2 always has lower total cost than strategy 1. This particular case occurs in 30 instances.

To further analyze the differences between the two strategies, Tables 7 and 8 provide additional information about the selected facilities for each combination of potential facilities ($|I|$), number of customers ($|J|$), and value of ρ (columns 1–3). Table 7 shows average results for strategy 1 by indicating the average values of the first-stage decisions in columns 4–6. The average demand of customer segment J^1 that is served with delay is displayed for each demand scenario in columns 7–9. Table 8 gives similar information but for strategy 2. Since adjustments in the number of modular units at open facilities are second-stage decisions, they are specified for each demand scenario separately.

Not surprisingly, the total number of selected facilities is, on average, similar in both strategies. This is explained by facility opening decisions being modeled as first-stage variables in both approaches. Furthermore, Table 8 clearly shows that the choice of facilities to have their capacities expanded or reduced is driven by the type of demand scenario. In contrast, when such choices are also included in the first-stage there is a decline in flexibility which results in greater total cost as seen previously. According to Table 8, when customer demand exhibits a trapezoidal pattern, the need to install additional modular units is higher compared to the other two demand scenarios. The removal of modular units from operating facilities is more noticeable when customer demand has an inverted trapezoidal structure. In the scenario with

$\|I\|$	$\|J\|$	ρ	Avg no. of open facilities	Avg no. of facilities with cap. adjustments		Scenario 1: irregular demand rate	Scenario 2: trapezoidal demand rate	Scenario 3: inverted trapezoidal demand rate
				Expansion	Contraction	Avg delayed demand (%)	Avg delayed demand (%)	Avg delayed demand (%)
5	20	1	2.00	0.90	0.40	0.14	4.29	0.08
		2	2.00	0.93	0.43	0.64	9.22	0.33
		3	2.00	0.93	0.43	1.05	10.21	0.60
10	40	1	3.50	1.00	0.43	0.04	3.55	0.01
		2	3.30	1.00	0.33	0.13	6.02	0.03
		3	3.30	1.00	0.30	0.16	6.21	0.03
15	60	1	5.07	1.33	0.60	0.03	2.70	0.01
		2	4.93	1.57	0.57	0.19	4.07	0.08
		3	4.83	1.37	0.33	0.18	4.25	0.07
All instances			3.44	1.11	0.43	0.29	5.61	0.14

Table 7: Average number of selected facilities in strategy 1 (formulation $(\mathbf{P}^{+}_{DE-1})$) and average demand served with delay to customer segment J^1.

$\|I\|$	$\|J\|$	ρ	Avg no. of open facilities	Scenario 1: irregular demand rate			Scenario 2: trapezoidal demand rate			Scenario 3: inverted trapezoidal demand rate		
				Avg no. of facilities with cap. adjustments		Avg delayed demand (%)	Avg no. of facilities with cap. adjustments		Avg delayed demand (%)	Avg no. of facilities with cap. adjustments		Avg delayed demand (%)
				Expansion	Contraction		Expansion	Contraction		Expansion	Contraction	
5	20	1	2.00	0.03	0.70	0.28	0.97	0.37	3.95	0.23	0.73	0.01
		2	2.00	0.10	0.43	0.78	1.00	0.23	8.02	0.23	0.40	0.16
		3	2.00	0.13	0.37	1.27	1.10	0.27	8.46	0.27	0.33	0.46
10	40	1	3.50	0.20	0.97	0.07	1.13	0.47	3.26	0.23	0.87	0.01
		2	3.30	0.27	0.87	0.27	1.13	0.20	5.87	0.40	0.77	0.03
		3	3.30	0.27	0.63	0.48	1.30	0.23	6.33	0.43	0.60	0.03
15	60	1	5.10	0.03	1.10	0.04	1.63	0.60	2.33	0.27	1.13	0.00
		2	4.93	0.27	0.77	0.31	1.83	0.33	4.06	0.40	0.73	0.09
		3	4.90	0.40	0.83	0.44	2.00	0.30	4.27	0.40	0.63	0.12
All instances			3.45	0.19	0.74	0.44	1.34	0.33	5.17	0.32	0.69	0.10

Table 8: Average number of selected facilities in strategy 2 (formulation $(\mathbf{P}^{+}_{DE-2})$) and average demand served with delay to customer segment J^1.

irregular demand pattern, both types of capacity adjustments occur, with the contraction of capacity having a slightly higher incidence.

Regardless of the strategy chosen, the relative amount of late shipments to customer segment J^1 is always the largest in demand scenario 2. This is the outcome of the trade-off between the investment in additional capacity to deal with higher demand requirements (either through opening more facilities or expanding operating facilities) and incurring a tardiness penalty cost for not serving all the demand on time. In general, and independently of the scenario and the strategy followed for determining a schedule for facility opening and capacity adjustment, the percentage of late deliveries increases with the maximum allowed delay, as expected.

4.4 The relevance of the stochastic setting

This section focuses on goal (v) of the computational study, namely on the evaluation of the benefits of the stochastic programming approach. This is a relevant issue because of the additional complexity posed by the stochastic setting compared to the deterministic counterpart. Even though there are no 'robust measures' for assessing the relevance of using a stochastic model instead of a (simpler) deterministic model (Birge and Louveaux [13]), two metrics are often used. The first is the *expected value of perfect information* (EVPI), while the second is the *value of the stochastic solution* (VSS). Both metrics will be calculated to measure the benefit of working with the stochastic programs $(\mathbf{P}^+_{DE-1})$ and $(\mathbf{P}^+_{DE-2})$. As discussed by Birge and Louveaux [13], EVPI and VSS are distinct indicators that measure different types of uncertainty.

The EVPI represents the maximum amount that the decision maker would be willing to pay in return for complete information about the future. It is defined as the difference between the optimal value of the stochastic program (in our case, z^+_{DE-1} or z^+_{DE-2}) and the optimal value of the so-called 'wait-and-see' (WS) problem. For each demand scenario s, formulation $(\mathbf{P}^+_{DE-1})$ (or $(\mathbf{P}^+_{DE-2})$) is solved under the assumption that the scenario occurs with probability 1. In this way, we obtain the optimal decisions that should be made if it were possible to fully predict future demands. Let $z^+_{DE-1,s}$ (resp. $z^+_{DE-2,s}$) denote the optimal value of the deterministic problem associated with a particular scenario $s \in S$ and strategy 1 (resp. strategy 2). It follows that the optimal value of the WS problem is given by $\sum_{s\in S} \pi_s \, z^+_{DE-1,s}$ for strategy 1. A similar calculation is performed for strategy 2.

The VSS measures the difference between the optimal objective value of the so-called 'expected value problem' (EEV) and the optimal objective value of the stochastic program.

The former is obtained by creating a single scenario in which all random variables are replaced by their expected values, and then determining the optimal values of the first-stage variables for that single scenario. Next, the expected cost over all scenarios using those first-stage variable values is calculated. In the case of our problem, the average scenario would be constructed by taking the average demand of every customer in each time period, $d_j^t = \sum_{s \in S} \pi_s \, d_{js}^t$ ($j \in J$; $t \in T$). Unfortunately, the associated values of the first-stage variables may result in an infeasible stochastic problem, either for model $(\mathbf{P}_{DE-1}^{+})$ or model $(\mathbf{P}_{DE-2}^{+})$, or even both. To overcome this difficulty, a 'reference scenario' is often used (Birge and Louveaux [13]). Typically, the worst-case scenario, i.e. the scenario with the highest demand level for every customer and time period, is adopted. Due to the features of our problem, it is not possible to identify such a scenario for every instance. Therefore, we adapt the standard definition of VSS by considering a 'modified reference scenario' such that $d_j^t = \max_{s \in S} \left\{ d_{js}^t \right\}$ for every $j \in J$ and $t \in T$.

Tables 9 and 10 report the average relative values of EVPI (column 5) and modified VSS (column 7) for strategy 1 and strategy 2, respectively. The relative values are calculated by dividing each metric by the optimal objective value of the associated stochastic problem. Column 4 indicates the number of instances for which optimal solutions are available and for that reason were used for calculating these average relative gaps. In addition, columns 6 and 8 give the average total computing time required to solve the WS and EEV problems to optimality, respectively.

$\lvert I \rvert$	$\lvert J \rvert$	ρ	# opt sol. $(\mathbf{P}_{DE-1}^{+})$	Metric 1		Metric 2	
				Avg EVPI gap (%)	Avg total CPU (sec.)	Avg modified VSS gap (%)	Avg total CPU (sec.)
5	20	1	30	9.57	234.8	0.52	11.3
		2	30	8.95	28.2	0.96	86.6
		3	30	8.96	26.1	0.97	14.0
10	40	1	30	9.29	1386.7	2.25	230.8
		2	30	8.18	341.5	3.31	455.2
		3	30	8.11	453.1	3.42	391.5
15	60	1	23	10.16	6949.6	1.35	720.0
		2	26	9.41	4104.3	2.04	903.3
		3	26	9.20	8086.0	2.50	1193.6
All instances			255	9.06	2400.8	1.93	445.2

Table 9: EVPI and modified VSS gaps for the scenario-independent location and capacity scalability model.

The EVPI and modified VSS gaps reported in the tables indicate that it is appropriate to consider a stochastic approach rather than its deterministic counterpart. For the scenario-independent location and capacity scalability model (strategy 1), the relative EVPI gap ranges from 1.45% to 17.70%, with an average of 9.06%. Concerning the scenario-dependent capacity scalability model (strategy 2), smaller, but still significant, EVPI gaps are obtained (average: 5.87%) that vary between 0.45% and 14.28%. Interestingly, the average EVPI gap does not seem to be greatly affected neither by the size of the customer set nor by the maximum delivery delay imposed on J^1. This is an indication that capturing uncertainty is relevant under different conditions.

For the modified VSS gap, we observe that it gradually increases with growing value of ρ. Recall that this metric measures how far away the value of the optimal solution of a deterministic simplified version of the problem is from the optimal value of a stochastic problem. By increasing ρ, more opportunities arise to satisfy the demand of some customers with delay which, in turn, enable a lower investment in configuring the facility network (e.g., through reducing the need to establish more facilities). This feature becomes even more relevant in a stochastic setting, whereas the quality of the optimal solution to the deterministic EEV problem deteriorates. Finally, and as expected, the computational burden required to optimally solve the WS and EEV problems is considerably lower than solving the stochastic programming models.

$\|I\|$	$\|J\|$	ρ	# opt sol. ($\mathbf{P}^+_{DE-2}$)	Metric 1 Avg EVPI gap (%)	Metric 1 Avg total CPU (sec.)	Metric 2 Avg modified VSS gap (%)	Metric 2 Avg total CPU (sec.)
5	20	1	30	6.69	239.7	0.03	144.8
		2	30	5.77	235.4	0.48	92.6
		3	30	5.75	89.5	0.53	21.3
10	40	1	27	6.61	1595.7	2.55	486.9
		2	28	5.48	745.8	3.52	552.7
		3	29	5.27	1201.9	3.74	648.8
15	60	1	13	5.90	8035.8	1.66	1261.5
		2	20	5.74	7554.8	2.24	1201.0
		3	18	5.50	8532.7	2.48	1382.4
All instances			225	5.87	3136.8	1.86	643.6

Table 10: EVPI and modified VSS gaps for the scenario-dependent capacity scalability model.

5 Conclusions

We have studied a stochastic multi-period facility location and capacity scalability problem taking into account the sensitivity of each individual customer to delivery lead times. In addition to determining a schedule for opening facilities and setting their initial number of modules, this problem also allows for several ways to adjust capacity over the planning horizon. Specifically, capacities can be increased incrementally by installing additional modules or decreased by removing existing modular units. These two features, expansion and contraction, can help substantially improve the flexibility in responding to spatial and temporal changes in demand. For the case of uncertain future customer demand, we have proposed two stochastic models that differ in the framework in which capacity adjustments can be planned. The first two-stage stochastic model follows the natural approach of defining all decisions related to the design of the facility network as first-stage decisions. Demand allocation decisions are deferred to the second stage. The second model presents an alternative strategy which is particularly relevant when capacity adjustment decisions have a tactical nature. In this case, first-stage decisions include the selection of locations to open new facilities and setting their initial capacities. The phased capacity expansion and contraction of facilities along with demand allocation represent the recourse decisions. Under the assumption that demand uncertainty can be captured by a finite set of scenarios, we were able to develop the extensive form of the deterministic equivalent for the two stochastic programs. Additionally, valid inequalities were developed to enhance the resulting large-scale formulations.

Numerical experiments with randomly generated instances indicate that the proposed additional inequalities substantially increase the capability of an off-the-shelf MILP solver to identify optimal solutions in much shorter computing times. Moreover, improved solution quality is also achieved for instances for which the specified time limit is reached without guaranteed optimality. Deferring the decisions on capacity expansion and capacity contraction to the second-stage problem results in network configurations with noticeable lower total cost compared to the approach where such decisions belong to the first-stage problem. This cost advantage arises due to the possibility of adapting the network configuration to particular demand realizations. However, this approach leads to considerably larger models which in turn require higher computational burden. Based on our comparative analysis, a decision maker becomes aware of the opportunities that arise from different strategies. Finally, we have shown that when demand is uncertain it is relevant to adopt a stochastic framework despite its additional complexity.

In further work, the proposed models could be extended by considering the uncertainty present in other parameters (e.g. costs) along with the unknown demand. Naturally, from a computational viewpoint this would pose additional challenges to an off-the-shelf solver. Therefore, it would be relevant to develop an efficient and effective heuristic method. In this case, the enhancements proposed in Section 3 would also be helpful as they sharpen the linear relaxation. Hence, the LP-bounds would be useful to evaluate the quality of the solutions returned by a heuristic procedure. Finally, another line of research could be directed toward the development of a multi-stage stochastic framework to enable the revision of the planning decisions as more information regarding the uncertain data is revealed. Even though a multi-stage model yields a better characterization of the dynamic planning process, solution techniques remain computationally challenging (Bakker et al. [9], Nickel et al. [44]).

Appendix A: Deterministic equivalent formulation of the scenario-dependent capacity scalability problem

For the scenario-dependent capacity scalability problem introduced in Section 2.3, the extensive form of the deterministic equivalent model of the two-stage stochastic program uses the continuous flow variables (25)–(27) and the following (scenario-indexed) binary variables that define the capacity adjustment decisions over the planning horizon:

$\widetilde{e}_{iks}^{\ell} =$ 1 if the capacity of a facility in location i is expanded with k modular units at time period ℓ under scenario s, 0 otherwise $(i \in I;\ k = 1, \ldots, n_i - 1;\ \ell \in T_L \setminus \{1\};\ s \in S)$. (49)

$\widetilde{r}_{iks}^{\ell} =$ 1 if the capacity of a facility in location i is reduced by removing k modular units at time period ℓ under scenario s, 0 otherwise $(i \in I;\ k = 1, \ldots, n_i - 1;\ \ell \in T_L \setminus \{1\};\ s \in S)$. (50)

$\widetilde{v}_{iks}^{\ell} =$ 1 if a facility is operated at location i with k modular units at time period ℓ under scenario $s \in S$, 0 otherwise $(i \in I;\ k = 1, \ldots, n_i;\ \ell \in T_L;\ s \in S)$. (51)

Formulation $(\mathbf{P}_{DE-2})$ is given by

$$\text{Min} \sum_{\ell \in T_L} \sum_{i \in I} \sum_{k=1}^{n_i} FO_{ik}^{\ell}\, z_{ik}^{\ell} + \sum_{s \in S} \pi_s \Bigg[\sum_{\ell \in T_L \setminus \{1\}} \sum_{i \in I} \sum_{k=1}^{n_i-1} FE_{ik}^{\ell}\, \widetilde{e}_{iks}^{\ell} +$$
$$\sum_{\ell \in T_L \setminus \{1\}} \sum_{i \in I} \sum_{k=1}^{n_i-1} FR_{ik}^{\ell}\, \widetilde{r}_{iks}^{\ell} + \sum_{\ell \in T_L} \sum_{i \in I} \sum_{k=1}^{n_i} \widetilde{v}_{iks}^{\ell} \sum_{t=\ell}^{\phi(\ell)} M_{ik}^{t} +$$
$$\sum_{t \in T} \sum_{i \in I} \sum_{k=1}^{n_i} o_{ik}^{t}\, w_{iks}^{t} + \sum_{t \in T} \sum_{i \in I} \sum_{j \in J^0} c_{ij}^{t}\, x_{ijs}^{t} +$$
$$\sum_{t \in T} \sum_{i \in I} \sum_{j \in J^1} \sum_{t'=t}^{\min\{t+\rho_j,\, |T|\}} \left(c_{ij}^{t'} + p_j^{tt'} \right) y_{ijs}^{tt'} \Bigg] \tag{52}$$

subject to

$$(10), (14), (29) - (35)$$

$$\sum_{k=1}^{n_i} \widetilde{v}_{iks}^{\ell} = \sum_{\ell' \in T_L :\, \ell' \leq \ell} \sum_{k=1}^{n_i} z_{ik}^{\ell'} \qquad i \in I,\ \ell \in T_L,\ s \in S \tag{53}$$

$$\sum_{k=1}^{n_i-1} \widetilde{e}_{iks}^{\ell} + \sum_{k=1}^{n_i-1} \widetilde{r}_{iks}^{\ell} \leq \sum_{\ell' \in T_L :\, \ell' < \ell} \sum_{k=1}^{n_i} z_{ik}^{\ell'} \qquad i \in I,\ \ell \in T_L \setminus \{1\},\ s \in S \tag{54}$$

$$\sum_{k=1}^{n_i} k\, \widetilde{v}_{iks}^{\ell} = \sum_{\ell' \in T_L :\, \ell' \leq \ell} \sum_{k=1}^{n_i} k\, z_{ik}^{\ell'} + \sum_{\ell' \in T_L \setminus \{1\} :\, \ell' \leq \ell} \sum_{k=1}^{n_i-1} k \left(\widetilde{e}_{iks}^{\ell'} - \widetilde{r}_{iks}^{\ell'} \right) \qquad i \in I,\ \ell \in T_L,\ s \in S \tag{55}$$

$$w_{iks}^{t} \leq k\, Q\, \widetilde{v}_{iks}^{\ell} \qquad i \in I,\ k = 1, \ldots, n_i,\ t \in T,\ \ell = \max\{\ell' \in T_L : \ \ell' \leq t\},\ s \in S \tag{56}$$

$$\widetilde{v}_{iks}^{\ell} \in \{0, 1\} \qquad i \in I,\ k = 1, \ldots, n_i,\ \ell \in T_L,\ s \in S \tag{57}$$

$$\widetilde{e}_{iks}^{\ell},\ \widetilde{r}_{iks}^{\ell} \in \{0, 1\} \qquad i \in I,\ k = 1, \ldots, n_i - 1,\ \ell \in T_L \setminus \{1\},\ s \in S \tag{58}$$

Appendix B: Data generation

The data generation scheme is based on the method developed by Correia and Melo [15] for a deterministic variant of our problem. In what follows, we denote by $\mathcal{U}[a,\, b]$ the generation of random numbers over the range $[a,\, b]$ according to a continuous uniform distribution.

For every customer j ($j \in J$), time period t ($t \in T$), and scenario s ($s \in S$), demand d^t_{js} is randomly generated according to the procedure described in Table 11.

Scenario	Demand pattern	Customer demand	Periods	Coefficient
$1, 2, 3$		$d^1_{js} = \mathcal{U}[20,\, 500]$		
1	irregular	$d^t_{js} = \beta^t_j\, d^{t-1}_j$	$t = 2, \ldots, \|T\|$	$\beta^t_j \in \mathcal{U}[0.8,\, 1.2]$
2	trapezoid	$d^t_{js} = \beta^t_j\, d^{t-1}_j$	$t = 2, \ldots, \|T\|/3$	$\beta^t_j \in \mathcal{U}[1.0,\, 1.2]$
		$d^t_{js} = \beta^t_j\, d^{t-1}_j$	$t = 1 + \|T\|/3, \ldots, 2\|T\|/3$	$\beta^t_j \in \mathcal{U}[0.99,\, 1.01]$
		$d^t_{js} = \beta^t_j\, d^{t-1}_j$	$t = 1 + 2\|T\|/3, \ldots, \|T\|$	$\beta^t_j \in \mathcal{U}[0.8,\, 1.0]$
3	inverted trapezoid	$d^t_{js} = \beta^t_j\, d^{t-1}_j$	$t = 2, \ldots, \|T\|/3$	$\beta^t_j \in \mathcal{U}[0.8,\, 1.0]$
		$d^t_{js} = \beta^t_j\, d^{t-1}_j$	$t = 1 + \|T\|/3, \ldots, 2\|T\|/3$	$\beta^t_j \in \mathcal{U}[0.99,\, 1.01]$
		$d^t_{js} = \beta^t_j\, d^{t-1}_j$	$t = 1 + 2\|T\|/3, \ldots, \|T\|$	$\beta^t_j \in \mathcal{U}[1.0,\, 1.2]$

Table 11: Generation of demand scenarios.

The generation of the variable costs relies on two real random numbers, β_1 and β_2, in the range 1.01 and 1.03 (details are given next). For $i \in I$ and $j \in J$, the unit variable distribution costs are generated as follows:

$$
\begin{aligned}
c^1_{ij} &= \mathcal{U}[5,\, 10], \\
c^t_{ij} &= c^1_{ij} && \text{for } t = 2, \ldots, \sigma, \\
c^t_{ij} &= \beta_1\, c^{12}_{ij} && \text{for } t = 1 + \sigma, \ldots, 2\,\sigma, \\
c^t_{ij} &= \beta_2\, c^{24}_{ij} && \text{for } t = 1 + 2\,\sigma, \ldots, |T|.
\end{aligned}
$$

Recall that $\sigma = |T|/|T_L|$ gives the total number of periods between two consecutive design periods. Since $|T_L| = 3$, the planning horizon is divided into three sections. For $|T| = 12$ (resp. $|T| = 24$), each section spans 4 (resp. 8) time periods. The unit distribution costs are assumed to be constant in all time periods within a section. However, the costs increase between 1% and 3% from one section to the next. This type of pattern is also present in the unit variable processing costs. The latter reflect economies of scale by considering the number of modular units that are available at a particular location. For $i \in I$, we set

$$\begin{aligned}
o_{ik}^{1} &= 100 \,/\, \sqrt{Q} && \text{for } t = 1, \ldots, \sigma, \\
o_{ik}^{t} &= 0.9\, o_{ik-1}^{t} && \text{for } t = 2, \ldots, \sigma;\ k = 2, \ldots, n_i, \\
o_{ik}^{t} &= \beta_1\, o_{ik}^{\sigma} && \text{for } t = 1+\sigma, \ldots, 2\,\sigma;\ k = 2, \ldots, n_i, \\
o_{ik}^{t} &= \beta_2\, o_{ik}^{2\sigma} && \text{for } t = 1+2\,\sigma, \ldots, |T|.
\end{aligned}$$

Recall that at most four modular units may be available at a location, i.e. $n_i = 4$ $(i \in I)$.

Regarding the generation of the fixed costs, the following procedure was employed.

- At the first design period, $\ell = 1$, the fixed cost of opening facility i $(i \in I)$ with k modular units $(k = 1, \ldots, n_i)$ is defined by

  $$FO_{ik}^{1} = \alpha_i + \gamma_i \sqrt{k\,Q}, \qquad \alpha_i \in \mathcal{U}[500,\ 1000];\ \gamma_i \in \mathcal{U}[4000,\ 6000].$$

  In the remaining strategic periods $\ell \in T_L \setminus \{1\}$, we set $FO_{ik}^{\ell} = \mathcal{U}[1.01,\ 1.03]\, FO_{ik}^{\ell-1}$ for every $k = 1, \ldots, n_i$. Once again, the capacity size of a facility at the time it is established also affects the above costs.

- All remaining fixed costs at location i $(i \in I)$ represent a given percentage of the fixed opening costs:

  $$\begin{aligned}
  FE_{ik}^{\ell} &= 0.25\, FO_{ik}^{\ell} && k = 1, \ldots, n_i - 1;\ \ell \in T_L \setminus \{1\}, \\
  FR_{ik}^{\ell} &= 0.10\, FO_{ik}^{\ell} && k = 1, \ldots, n_i - 1;\ \ell \in T_L \setminus \{1\}, \\
  M_{ik}^{t} &= 0.05\, FO_{ik}^{\ell} && k = 1, \ldots, n_i;\ \ell \in T_L;\ t = \ell, \ldots, t';\ t' \in T : t' < \ell + 1.
  \end{aligned}$$

Finally, we adapt the scheme proposed in [15] to generate the unit tardiness penalty costs for late deliveries to customers $j \in J^1$ as follows:

$$p_j^{tt'} = 0.1\, \theta_j^t\, (t' - t)^2 \qquad \text{for } t \in T \text{ and } t' = t, \ldots, \min\{t + \rho_j,\ |T|\},$$

with

$$\theta_j^t = \frac{\sum_{i \in I} \sum_{k=1}^{n_i} M_{ik}^t}{TD\, |I| \sum_{i \in I} n_i} + \frac{\sum_{i \in I} c_{ij}^t}{|I|} + \frac{\sum_{i \in I} \sum_{k=1}^{n_i} o_{ik}^t}{|I| \sum_{i \in I} n_i},$$

and TD denoting the maximum quantity demanded across all scenarios. Therefore, $TD = \sum_{j\in J}\sum_{t\in T}\max_{s\in S} d^t_{js}$. Observe that the unit tardiness penalty cost is a function of the average fixed facility operating costs, and the average variable distribution and processing costs.

Acknowledgements

This work was partially supported by FCT - Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UID/MAT/00297/2019 (CMA/FCT/UNL). This support is gratefully acknowledged.

References

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

[2] S. Ahmed and R. Garcia. Dynamic capacity acquisition and assignment under uncertainty. *Annals of Operations Research*, 124:267–283, 2004.

[3] S. Ahmed, A.J. King, and G. Parija. A multi-stage stochastic integer programming approach for capacity expansion under uncertainty. *Journal of Global Optimization*, 26:3–24, 2003.

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

[5] A. Alonso-Ayuso, L.F. Escudero, A. Garín, M.T. Ortuño, and G. Pérez. An approach for strategic supply chain planning under uncertainty based on stochastic 0-1 programming. *Journal of Global Optimization*, 26:97–124, 2003.

[6] S.A. Alumur, S. Nickel, and F. Saldanha da Gama. Hub location under uncertainty. *Transportation Research Part B*, 46:529–543, 2012.

[7] E.M. Alvarez, M.C. van der Heijden, and W.H.M. Zijm. Service differentiation in spare parts supply through dedicated stocks. *Annals of Operations Research*, 231:283–303, 2015.

[8] A. Antunes and D. Peeters. On solving complex multi-period location models using simulated annealing. *European Journal of Operational Research*, 130:190–201, 2001.

[9] H. Bakker, F. Dunke, and S. Nickel. A structuring review on multi-stage optimization under uncertainty: Aligning concepts from theory and practice. *Omega*, 2019. doi: 10.1016/j.omega.2019.06.006. *In press*.

[10] O. Baron, J. Milner, and H. Naseraldin. Facility location: A robust optimization approach. *Production and Operations Management*, 20:772–785, 2011.

[11] P. Beraldi and M.E. Bruni. A probabilistic model applied to emergency service vehicle location. *European Journal of Operational Research*, 196:323–331, 2009.

[12] O. Berman and J. Wang. The probabilistic 1-maximal covering problem on a network with discrete demand weights. *Journal of the Operational Research Society*, 59:1398–1405, 2008.

[13] J.R. Birge and F.V. Louveaux. *Introduction to Stochastic Programming*. Springer Series in Operations Research and Financial Engineering. Springer, New York, Dordrecht, Heidelberg, London, second edition, 2011.

[14] I. Contreras, J.-F. Cordeau, and G. Laporte. Stochastic uncapacitated hub location. *European Journal of Operational Research*, 212:518–528, 2011.

[15] I. Correia and T. Melo. A multi-period facility location problem with modular capacity adjustments and flexible demand fulfillment. *Computers & Industrial Engineering*, 255:729–746, 2017.

[16] I. Correia and F. Saldanha da Gama. Facility location under uncertainty. In G. Laporte, S. Nickel, and F. Saldanha da Gama, editors, *Location Science*, chapter 8, pages 177–204. Springer, Heidelberg, 2015.

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

[18] I. Correia, S. Nickel, and F. Saldanha da Gama. A stochastic multi-period capacitated multiple allocation hub location problem: Formulation and inequalities. *Omega*, 74:122–134, 2018.

[19] M.J. Cortinhal, M.J. Lopes, and M.T. Melo. Dynamic design and re-design of multi-echelon, multi-product logistics networks with outsourcing opportunities: A computational study. *Computers & Industrial Engineering*, 90:118–131, 2015.

[20] E.M. Delmelle, J.-C. Thill, D. Peeters, and I. Thomas. A multi-period capacitated school location problem with modular equipment and closest assignment considerations. *Journal of Geographical Systems*, 16:263–286, 2014.

[21] J. Dias, M.E. Captivo, and J. Clímaco. Dynamic location problems with discrete expansion and reduction sizes of available capacities. *Investigação Operacional*, 27:107–130, 2007.

[22] A. Döyen, N. Aras, and G. Barbarosoğlu. A two-echelon stochastic facility location model for humanitarian relief logistics. *Operations Research Letters*, 6:1123–1145, 2012.

[23] S. Duran, T. Liu, D. Simchi-Levi, and J.L. Swann. Policies utilizing tactical inventory for service-differentiated customers. *Operations Research Letters*, 36:259–264, 2008.

[24] M. Fattahi, K. Govindan, and E. Keyvanshokooh. Responsive and resilient supply chain network design under operational and disruption risks with delivery lead-time sensitive customers. *Transportation Research Part E*, 101:176–200, 2017.

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

[26] E. Gourdin and O. Klopfenstein. Multi-period capacitated location with modular equipments. *Computers & Operations Research*, 35:661–682, 2008.

[27] K. Govindan, M. Fattahi, and E. Keyvanshokooh. Supply chain network design under uncertainty: A comprehensive review and future research directions. *European Journal of Operational Research*, 263:108–141, 2017.

[28] E. Grass, K. Fischer, and A. Rams. An accelerated L-shaped method for solving two-stage stochastic programs in disaster management. *Annals of Operations Research*, 2018. doi: https://doi.org/10.1007/s10479-018-2880-5.

[29] I. Heckmann. *Towards Supply Chain Risk Analytics*, chapter 12, pages 313–361. Springer Gabler, Wiesbaden, 2016.

[30] P. Hernández, A. Alonso-Ayuso, F. Bravo, L.F. Escudero, M. Guignard, V. Marianov, and A. Weintraub. A branch-and-cluster coordination scheme for selecting prison facility sites under uncertainty. *Computers & Operations Research*, 39:2232–2241, 2012.

[31] H.-C. Hung, E.P. Chew, L.H. Lee, and S. Liu. Dynamic inventory rationing for systems with multiple demand classes and general demand processes. *International Journal of Production Economics*, 139:351–358, 2012.

[32] S.D. Jena, J.-F. Cordeau, and B. Gendron. Solving a dynamic facility location problem with partial closing and reopening. *Computers & Operations Research*, 67:143–154, 2016.

[33] S. Khodaparasti, M.E. Bruni, P. Beraldi, H.R. Maleki, and S. Jahedi. A multi-period location-allocation model for nursing home network planning under uncertainty. *Operations Research for Health Care*, 18:4–15, 2018.

[34] D. Kim, K. Lee, and I. Moon. Stochastic facility location model for drones considering uncertain flight distance. *Annals of Operations Research*, 2018. doi: https://doi.org/10.1007/s10479-018-3114-6.

[35] Ö.B. Kınay, B.Y. Kara, F. Saldanha-da-Gama, and I. Correia. Modeling the shelter site location problem using chance constraints: A case study for Istanbul. *European Journal of Operational Research*, 270:132–145, 2018.

[36] G. Laporte, F.V. Louveaux, and L. van Hamme. Exact solution to a location problem with stochastic demands. *Transportation Science*, 28:95–103, 1994.

[37] Z. Li, Q. Lu, and M. Talebian. Online versus bricks-and-mortar retailing: a comparison of price, assortment and delivery time. *International Journal of Production Research*, 53:3823–3835, 2015.

[38] F.V. Louveaux. Discrete stochastic location models. *Annals of Operations Research*, 6:23–34, 1986.

[39] A. Marín, L.I. Martínez-Merino, A.M. Rodríguez-Chia, and F. Saldanha da Gama. Multi-period stochastic covering location problems: Modeling framework and solution approach. *European Journal of Operational Research*, 268:432–449, 2018.

[40] N. Marković, I.O. Ryzho, and P. Schonfeld. Evasive flow capture: A multi-period stochastic facility location problem with independent demand. *European Journal of Operational Research*, 257:687–703, 2017.

[41] M.T. Melo, S. Nickel, and F. Saldanha da Gama. Dynamic multi-commodity capacitated facility location: a mathematical modeling framework for strategic supply chain planning. *Computers & Operations Research*, 33:181–208, 2006.

[42] I.B. Mohamed, W. Klibi, and F. Vanderbeck. Designing a two-echelon distribution network under demand uncertainty. *European Journal of Operational Research*, 280:102–123, 2020.

[43] S. Nickel and F. Saldanha da Gama. Multi-period facility location. In G. Laporte, S. Nickel, and F. Saldanha da Gama, editors, *Location Science*, chapter 11, pages 289–310. Springer, Heidelberg, 2015.

[44] S. Nickel, F. Saldanha-da-Gama, and H.-P. Ziegler. A multi-stage stochastic supply network design problem with financial decisions and risk management. *Omega*, 40:511–524, 2012.

[45] S.H. Owen and M.S. Daskin. Strategic facility location: A review. *European Journal of Operational Research*, 111:423–447, 1998.

[46] L. Snyder. Facility location under uncertainty: A review. *IIE Transactions*, 38:537–554, 2006.

[47] V. Verter and M.C. Dincer. Facility location and capacity acquisition: An integrated approach. *Naval Research Logistics*, 42:1141–1160, 1995.

[48] Y. Wang, M.A. Cohen, and Y.-S. Zheng. Differentiating customer service on the basis of delivery lead-times. *IIE Transactions*, 34:979–489, 2002.

[49] W. Wilhelm, X. Han, and C. Lee. Computational comparison of two formulations for dynamic supply chain reconfiguration with capacity expansion and contraction. *Computers & Operations Research*, 40:2340–2356, 2013.

[50] L. A. Wolsey. *Integer programming*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, New York, 1998.

[51] D. Zhuge, S. Yu, L. Zhen, and W. Wang. Multi-period distribution center location and scale decision in supply chain network. *Computers & Industrial Engineering*, 101:216–226, 2016.

Veröffentlichte Berichte der Fakultät für Wirtschaftswissenschaften

Die PDF-Dateien der folgenden Berichte sind verfügbar unter:

Published reports of the Saarland Business School

The PDF files of the following reports are available under:

http://www.htwsaar.de/wiwi

1 I. Correia, T. Melo, F. Saldanha da Gama

Comparing classical performance measures for a multi-period, two-echelon supply chain network design problem with sizing decisions

Keywords: supply chain network design, facility location, capacity acquisition, profit maximization, cost minimization

(43 pages, 2012)

2 T. Melo

A note on challenges and opportunities for Operations Research in hospital logistics

Keywords: hospital logistics, Operations Research, application areas

(13 pages, 2012)

3 S. Hütter, A. Steinhaus

Forschung an Fachhochschulen – Treiber für Innovation im Mittelstand: Ergebnisse der Qbing-Trendumfrage 2013

Keywords: Innovation, Umfrage, Trendbarometer, Logistik-Konzepte, Logistik-Technologien, Mittelstand, KMU

(5 pages, 2012)

4 A. Steinhaus, S. Hütter

Leitfaden zur Implementierung von RFID in kleinen und mittelständischen Unternehmen

Keywords: RFID, KMU, schlanke Prozesse, Prozessoptimierung, Produktion, Forschungsgruppe Qbing

(49 pages, 2013)

5 S.A. Alumur, B.Y. Kara, M.T. Melo

Location and Logistics

Keywords: forward logistics network design, reverse logistics network design, models, applications

(26 pages, 2013)

6 S. Hütter, A. Steinhaus

Forschung an Fachhochschulen – Treiber für Innovation im Mittelstand: Ergebnisse der Qbing-Trendumfrage 2014

Keywords: Innovation, Umfrage, Trendbarometer, Logistik-Konzepte, Logistik-Technologien, Mittelstand, KMU

(6 pages, 2014)

7 M.J. Cortinhal, M.J. Lopes, M.T. Melo

Redesigning a three-echelon logistics network over multiple time periods with transportation mode selection and outsourcing opportunities

Keywords: logistics network design/re-design, multiple periods, transportation mode selection, product outsourcing, mixed-integer linear programming

(49 pages, 2014)

8 T. Bousonville, C. Ebert, J. Rath

A comparison of reward systems for truck drivers based on telematics data and driving behavior assessments

Keywords: telematics, driving behavior, incentives, award systems

(9 pages, 2015)

9 I. Correia, T. Melo

Multi-period capacitated facility location under delayed demand satisfaction

Keywords: location, multi-period, capacity choice, delivery lateness, MILP models

(35 pages, 2015)

10 C.L. Martins, M.T. Melo, M.V. Pato

Redesigning a food bank supply chain network, Part I: Background and mathematical formulation

Keywords: supply chain, sustainability, tri-objective problem, MILP model

(30 pages, 2016)

11 I. Correia, T. Melo

A computational comparison of formulations for a multi-period facility location problem with modular capacity adjustments and flexible demand fulfillment

Keywords: facility location, multi-period, capacity expansion and contraction, delivery lateness, mixed-integer linear models

(42 pages, 2016)

12 A. Bernhardt, T. Melo, T. Bousonville, H. Kopfer

Scheduling of driver activities with multiple soft time windows considering European regulations on rest periods and breaks

Keywords: road transportation, driver scheduling, rest periods, breaks, driving hours, Regulation (EC) No 561/2006, mixed integer linear programming models

(137 pages, 2016)

13 C.L. Martins, M.T. Melo, M.V. Pato

Redesigning a food bank supply chain network, Part II: Computational study

Keywords: food rescue and delivery, sustainability, supply chain network design, tri-objective problem, social impact, economic and environmental performance

(57 pages, 2017)

14 A. Bernhardt, T. Melo, T. Bousonville, H. Kopfer

Truck driver scheduling with combined planning of rest periods, breaks and vehicle refueling

Keywords: road transportation, refueling, fuel cost, driver scheduling, rest periods, breaks, driving hours, Regulation (EC) No 561/2006, mixed integer linear programming

(90 pages, 2017)

15 M.J. Cortinhal, M.J. Lopes, M.T. Melo

Impact of partial product outsourcing, transportation mode selection, and single-assignment requirements on the design of a multi-stage supply chain network

Keywords: supply chain network design, facility location, supplier selection, in-house production, product outsourcing, transportation mode selection, single-assignment, mixed-integer linear programming

(51 pages, 2018)

16 C. Sauvey, T. Melo, I. Correia

Two-phase heuristics for a multi-period capacitated facility location problem with service-differentiated customers

Keywords: facility location, multi-period, delivery lateness, constructive heuristics, local improvements

(41 pages, 2019)

17 I. Correia, T. Melo

Dynamic facility location problem with modular capacity adjustments under uncertainty

Keywords: facility location, dynamic capacity adjustment, delivery lateness, stochastic programming, valid inequalities

(42 pages, 2019)

htw saar

Hochschule für
Technik und Wirtschaft
des Saarlandes
University of
Applied Sciences

Hochschule für Technik und Wirtschaft des Saarlandes

Die Hochschule für Technik und Wirtschaft des Saarlandes (htw saar) wurde im Jahre 1971 als saarländische Fachhochschule gegründet. Insgesamt studieren rund 6000 Studentinnen und Studenten in 46 verschiedenen Studiengängen an der htw saar, aufgeteilt auf vier Fakultäten.

In den vergangenen zwanzig Jahren hat die Logistik immens an Bedeutung gewonnen. Die htw saar hat dieser Entwicklung frühzeitig Rechnung getragen und einschlägige Studienprogramme sowie signifikante Forschungs- und Technologietransferaktivitäten entwickelt. Die Veröffentlichung der Schriftenreihe Logistik soll die Ergebnisse aus Forschung und Projektpraxis der Öffentlichkeit zugänglich machen.

Weitere Informationen finden Sie unter **http://logistik.htwsaar.de**

ISCOM
INSTITUT FÜR SUPPLY CHAIN
UND OPERATIONS MANAGEMENT

Institut für Supply Chain und Operations Management

Das Institut für Supply Chain und Operations Management (ISCOM) der htw saar ist auf die Anwendung quantitativer Methoden in der Logistik und deren Implementierung in IT-Systemen spezialisiert. Neben öffentlich geförderten Forschungsprojekten zu innovativen Themen arbeitet ISCOM eng mit Projektpartnern aus der Wirtschaft zusammen, wodurch der Wissens- und Technologietransfer in die Praxis gewährleistet wird. Zu den Arbeitsgebieten zählen unter anderem Distributions- und Transportplanung, Supply Chain Design, Bestandsmanagement in Supply Chains, Materialflussanalyse und -gestaltung sowie Revenue Management.

Weitere Informationen finden Sie unter **http://iscom.htwsaar.de**

Qbing
Produktion | Logistik | Technologie

Forschungsgruppe Qbing

Qbing ist eine Forschungsgruppe an der Hochschule für Technik und Wirtschaft des Saarlandes, die spezialisiert ist auf interdisziplinäre Projekte in den Bereichen Produktion, Logistik und Technologie. Ein Team aus derzeit acht Ingenieuren und Logistikexperten arbeitet unter der wissenschaftlichen Leitung von Prof. Dr. Steffen Hütter sowohl in öffentlich geförderten Projekten als auch zusammen mit Industriepartnern an aktuellen Fragestellungen zur Optimierung von logistischen Prozessabläufen in Handel und Industrie unter Einbeziehung modernster Sensortechnologie und Telemetrie. Qbing hat auch und gerade auf dem Gebiet der angewandten Forschung Erfahrung in der Zusammenarbeit mit kleinen und mittelständischen Unternehmen.

Weitere Informationen finden Sie unter **http://www.qbing.de**

ISSN 2193-7761