



Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Production, Manufacturing, Transportation and Logistics

Multi-period distribution networks with purchase commitment contracts

Christian Clavijo López^a, Yves Crama^a, Thierry Pironet^{a,*}, Frédéric Semet^b^a QuantOM, HEC Liège - Management School of the University of Liège, Belgium^b Univ. Lille, CNRS, Inria, Centrale Lille, UMR 9189 CRISTAL, F-59000 Lille, France

ARTICLE INFO

Article history:

Received 25 May 2022

Accepted 6 July 2023

Available online 16 July 2023

Keywords:

Transportation

Supply chain management

Contracts

Network design

Matheuristics

ABSTRACT

Retailers which deliver products directly to their customer locations often rely on Logistics Service Intermediaries (LSI) for order management, warehousing, transportation and distribution services. Usually, the LSI acts as a shipper and subcontracts the transportation to carriers for long-haul and last-mile delivery services. All agents interact and are connected through cross-docking facilities. As the demand from customers may vary significantly over time, the shipper's requirements for transportation evolve accordingly at the tactical level. This creates opportunities for the shipper to take advantage of medium-term contracts with the carriers at prices lower than those offered by the spot market. The study focuses on the tactical design, through dynamic contracts, of a suitable network of cross-docking facilities and related transportation capacities (belonging to different carriers) to reduce the shipper's operational costs. In this article, we propose an MILP formulation for the multi-period planning problem with minimum purchase commitment contracts faced by the shipper. We propose exact and heuristic decomposition methods for the model, respectively, based on combinatorial Benders cuts and on relax-and-repair approaches. The performance of these algorithms is experimentally compared to that of commercial solvers (branch-and-cut and classical Benders). The numerical results show that our methods perform comparatively well for the solution of large size instances and brings economic benefits to the shipper.

© 2023 Elsevier B.V. All rights reserved.

1. Introduction

Logistic firms differ according to the service capabilities and core competences that they deploy in their business networks (Lai, 2004). While some invest in a large number of physical assets, others may restrict their activities and partially outsource their operations. As observed for example by Cui & Hertz (2011), the terminology to designate these different actors is not completely standardized. We rely here on the framework adopted by Stefansson & Russell (2008), Cui & Hertz (2011), who distinguish three main groups of actors, namely, logistics service providers, logistics service intermediaries (LSIs), and sub-contract carriers. LSIs may carry out most administrative activities on behalf of their clients, but leave the physical activities to contracted third parties (Stefansson & Russell, 2008). So, LSIs are often non-asset-based service providers: their business is to coordinate and connect different actors and activities (Cui & Hertz, 2011). In contrast, asset-based companies may offer various services relying on their own physical resources and capabilities, like warehousing, trans-

portation, or freight consolidation, and may provide value-added services supported by IT technologies.

This study takes the perspective of an LSI, the so-called *shipper*, who is responsible for delivering to end-customers the products they purchased, for example, on e-commerce websites. The shipper subcontracts some operations to *carriers* who own and operate physical assets. Two main decisions must be taken by the shipper, one of a purely logistical nature and one of a contractual nature.

The logistical decision mostly bears on the carriers' resources which should be activated at every period to serve the customers at minimum cost: type and number of facilities (hubs, warehouses, cross-docking facilities, delivery points), connecting lanes (number, capacity), and customer service (single or multiple sourcing). It is essentially a classical distribution network design problem.

The contractual decision relates to the preselection of carriers to be included in the distribution network of the shipper ahead of time to secure access to resources at an attractive price when needed. This component of the decision problem is at the core of our work. Generally speaking, the shipper might choose short-

* Corresponding author.

E-mail addresses: cjclavijo@uliege.be (C. Clavijo López), yves.crama@uliege.be (Y. Crama), thierry.pironet@uliege.be (T. Pironet), frederic.semet@centralelille.fr (F. Semet).

term, mid-term, or long-term contractual agreements with some carriers. In the first, short-term case, the costs of contracted services vary along with market prices. This occurs when the shipper sources transportation capacities from the spot market. This policy handles shipments on a one-time, load-by-load basis, and it allows the transportation capacity to be flexibly adjusted to real demand requirements at each current period. As mentioned in [Lindsey & Mahmassani \(2017\)](#), however, “the spot market itself can be highly dynamic, and subject to considerable uncertainty in availability and/or pricing, making it difficult for shippers to utilize”.

At the other end of the spectrum, long-term contractual agreements guarantee more stable prices for the shipper. Thus, the shipper is interested in signing such contracts that could reduce its outbound logistics costs in the long run while ensuring the availability of transportation capacity, at the risk of buying overcapacity in a fixed network. In return, the carriers expect some level of regularity in demand for their services in order to make the contracts profitable: they agree to conclude discounted contracts when the shipper offers sufficient business volume, sufficient regularity, and service compatibility (see [Kuyzu, Akyol, Ergun, & Savelsbergh, 2015](#)).

In markets subject to seasonal variations and to significant uncertainty, both policies mentioned above may prove costly. Indeed, the short-term policy tends to incur high spot prices when the demand is large, and the long-term policy suffers from the reservation of excess capacity when the demand is low. Such situations are typically encountered in e-commerce, but also in classical supply chains (see [Section 2](#)). With this in mind, this article focuses on mid-term contracts whereby the shipper tactically manages a portfolio of carrier contracts that are dynamically revised to balance the reserved capacity over a mid-term horizon. The setting is inspired by real-world situations encountered in a collaboration with fast-delivery companies, as described in [Gendron & Semet \(2009\)](#). In this study, the shipper operates a two-echelon distribution system where the intermediate facilities are owned and operated by carriers which engage in mid-term contracts with the shipper. The system is highly adaptive, in the sense that the facilities can be used or not, depending on fluctuations of the demand. While [Gendron & Semet \(2009\)](#) focus on the mathematical properties of the single-period optimization model, our main interest here is in integrating the contractual options in a multi-period framework.

A main objective of this paper is therefore to propose a mathematical optimization approach of the planning problem faced by the shipper in this context, namely, the tactical design, through dynamic mid-term contracts, of a suitable network of cross-docking facilities and related transportation capacities belonging to different carriers. The optimization approach aims at the reduction of operational costs, assuming that the parameters of the contracts (price, duration) are known. When used in a simulation framework, it also allows for a comparison of the impact of different contractual conditions (e.g., contracts with different durations).

Accordingly, as a first contribution, the paper introduces an optimization model integrating multi-period purchase commitment contracts within the distribution network design. To the best of our knowledge, this problem has not been previously studied, and even less modeled or solved, in the literature. The optimization problem is proved to be NP-hard, and as expected, solving large-size instances is time-consuming. As a second contribution, the structure of the problem is used to decompose it into a collection of single-period network design problems. This algorithmic contribution is implemented in two ways: we propose, first, an exact method based on Benders’ canonical cut decomposition and second, an ad-hoc relax-and-repair heuristic. Finally, computational experiments are carried out to assess the efficiency of the algorithms, the applicability of the problem formulations, and the economic benefits derived from the flexibility of the contracts.

The article is organized as follows. [Section 2](#) reviews the literature on multi-period network design and capacity reservation contracts as applied in operations management. [Section 3](#) specifies the decision problem, which is then formulated in [Section 4](#) as a mixed-integer linear program. [Section 5](#) establishes the complexity of the problem. [Section 6](#) presents an exact optimization procedure based on combinatorial Benders decomposition, and [Section 7](#) proposes a relax-and-repair heuristic. [Section 8](#) describes the data sets used in the computational tests. The results of the tests are discussed in algorithmic terms in [Section 9](#) and in economic terms in [Section 10](#). Finally, [Section 11](#) draws conclusions and proposes future research paths.

2. Literature review

The literature review consists of two parts. The first one deals with the multi-period design of distribution networks. The second part covers contracts in supply chain management.

Network design is a broad research area. As mentioned in [Klose & Drexl \(2005\)](#), this class of problems features a least 9 core dimensions: topography, objective function, capacity and demand satisfaction constraints (single or multiple sourcing), number of stages or echelons, single or multiple products, the elasticity of demand with respect to location, static or dynamic design over a planning horizon, deterministic or uncertain data, inclusion or exclusion of routing decisions. Additional features could also be mentioned. This leads to a vast literature related to network design. Therefore, we aim to position our problem statement in this stream briefly.

The “multi-period” nature of our problem refers to a sequence of decisions made over a planning horizon discretized into a finite number of periods. In contrast with this assumption, many network design models aim at making a “static” strategic decision entailing large investments and engaging the firm for a long period of time. However, even when discussing one of the most basic network design models, namely, the facility location problem, the issue arises of dealing with time-dependent demand (see [Drezner, 1995](#) and [Owen & Daskin, 1998](#)). This leads to tactical problem statements, expressing that the firm wants to determine a “dynamic” sequence of distribution networks in which facilities can be opened or closed over time, at a cost, or with a limited number of changes from one period to the next, or should retain their status for a fixed, minimum number of periods; see, e.g., [Wesolowsky & Truscott \(1976\)](#), [Van Roy & Erlenkotter \(1982\)](#), [Drezner & Wesolowsky \(1991\)](#), [Klose & Drexl \(2005\)](#), [Melo, Nickel, & da Gama \(2005\)](#), [Dias, Captivo, & Clímacos \(2008\)](#), [Jena, Cordeau, & Gendron \(2015\)](#). As mentioned in [Klose & Drexl \(2005\)](#), this second, tactical approach leads to an increase in data requirements compared to an aggregated model, so the ability to solve the models reduces accordingly. Also, a main added difficulty with multi-period problems arises from the connections between periods. It is well-known in production planning that inventories or setup costs link successive periods with each other. In the case of multi-period network design problems, the status of the facilities (open, operating or closed) similarly creates connections between periods.

In single-echelon networks, it is usually assumed that the facilities are owned by the decision maker, who is therefore entitled to define the opening/closing conditions and related costs. When dealing with multi-echelon networks, it may be the case that the intermediate levels are owned by subcontractors who can be activated over time. This leads to the consideration of temporary contracts to use these facilities. A detailed introduction to such two-echelon network design problems can be found in [Ben Mohamed, Klibi, & Vanderbeck \(2020\)](#). These authors propose a classification based on the following notations for the multi-period setting:

- 1 vs. T for a mono-period or multi-period problem,
- O if opening new locations is allowed,
- C if closing previously opened locations is allowed, C implies O,
- Re if reopening closed locations is allowed, Re implies O and C.

Our problem falls in the class T/O/C/Re, like Pimentel, Mateus, & Almeida (2013) (one-echelon), Cortinhal, Lopes, & Melo (2015), or Ben Mohamed et al. (2020). Yet, in these papers, the contracts with the suppliers involve opening and operating costs for the facilities, but no minimal or maximal duration extending over several periods, nor any purchase commitment.

The latter comment leads to the second part of our literature review, which deals with contracts used in supply chain management. Two distinct perspectives can be found in the discussion of supply chain contracts (see Lariviere, 1999 for a global overview).

A first, large stream of literature focuses on the *design* of contracts aiming at the coordination of firms in a decentralized supply chain. Setting up contract parameters appropriately can provide a sufficient mechanism to achieve optimal global profitability and overcome inefficiency factors such as information asymmetry, risk aversion, or power imbalance between players that motivate actions in their individual interest. In this setting, contract parameters are usually considered as *endogenous* in the problem formulation. Illustrations can be found in Tsay & Lovejoy (1999), Cachon & Lariviere (2001), Corbett, Zhou, & Tang (2004).

The second stream of research focuses on determining optimal procurement plans for a single firm with one or multiple suppliers, assuming that the general terms and conditions of risk-sharing contracts are *exogenously* fixed. Examples are found in de Albéniz & Simchi-Levi (2005), Lian & Deshmukh (2009), Akbalik, Hadj-Alouane, Sauer, & Ghribi (2017). The present work falls in this second category: it investigates procurement planning of transportation services (from a single firm perspective) under given contract parameters, namely, fixed duration and minimum business volume commitment.

The supply chain literature describes different variants of such contractual mechanisms and related parameters, e.g., *capacity reservation contracts* (Akbalik et al., 2017; Jin & Wu, 2007; Li, Luo, Wang, & Zhou, 2020), *total minimum order quantity* (Bassok & Anupindi, 1997; Chen & Krass, 2001), *quantity flexibility contracts* (Chen, Hum, & Sun, 2001; Heydari, Govindan, Ebrahimi, & Taleizadeh, 2020; Li et al., 2020; Lian & Deshmukh, 2009; Tsay & Lovejoy, 1999), *option contracts* (de Albéniz & Simchi-Levi, 2005; Nosoohi & Nookabadi, 2016; Spinler & Huchzermeier, 2006).

In particular, the subject of contractual carrier selection has been investigated in several forms, from simple to more advanced ones. For instance, the basic Carrier Selection Problem (CSP) was addressed by several authors in the context of combinatorial auctions, usually viewed from the strategic perspective of yearly contracts. Caplice & Sheffi (2003) develop optimization models to determine how shippers should procure transportation services after receiving bids from carriers. Song & Regan (2005) describe a unit auction where shippers predefine sets of routes for bidding purposes. In maritime transportation, Lim, Wang, & Xu (2006) further incorporate a minimum quantity commitment requirement in the classical transportation model. In their model, the freight owner decides how to allocate shipments among multiple carriers while respecting the constraint that each selected carrier must handle a minimum volume of cargo. In Brusset (2009, 2010), several types of multi-period contracts are compared to formalize long-term relationships between one carrier and one supplier. In addition to the simple price-only contract without commitments, contracts with minimum purchase commitments (per period or over several periods) can be selected based on the economic return for both actors. Patel & Swartz (2019) consider a supply chain design problem motivated by applications in the chemical industry, in which the

transportation links are subject to contracts with a fixed duration. This model shares some similarities with ours, with the significant difference that the contracts do not entail any minimum purchase commitment.

Only a few publications deal with risk-sharing contracts for logistic services other than transportation. For example, in Chen et al. (2001), the authors consider flexible commitment contracts proposed by a company which subcontracts warehousing space to third-party service providers. Each client specifies a base (reservation) commitment for storage space at the warehouse based on its expected demand. Any space used above the commitment level is charged at a premium price during the planning period. The base capacity commitment can be adjusted according to periodic demand requirements along the multi-period horizon.

In this paper, we consider a multi-period network design problem where multiple carriers operate their own subnetworks of facilities as well as their transportation resources. We focus on contracts with *minimum purchase commitments*, a type of risk-sharing contract which is appropriate in the context of multi-product procurement planning. In this model, the shipper commits to pay a deductible fee in order to benefit from discounted prices for a “reserved” business volume. An advantage of such contracts is that they express the commitment in homogeneous monetary units, and hence, they can be applied to the purchase of multiple product types. For instance, they are common in the electronics industry, see, e.g., Bassok & Anupindi (1997). In transportation, carriers offering loads on different corridors can be considered as selling different products with distinct prices. On the other hand, the reservation of transportation (truckloads) or service (handling) capacity is not explicitly modeled in the contract. Alternative models taking capacity commitments into account are discussed in Clavijo López (2021).

3. Description of the problem

A logistics service intermediary known as the *shipper* acts on behalf of online sellers for the distribution of products in many customer areas. The shipper manages the inventory in warehouses which are replenished by the sellers on a continuous basis. Moreover, the shipper is in charge of delivering parcels along time according to *customer orders*. To do so, it relies on specialized *carriers* for long-haul transportation from the warehouses to intermediate facilities and for cross-docking operations at the facilities, and it relies on *parcel delivery services* for last-mile transportation to customers areas. An example of such a distribution network is shown in Fig. 1.

Our work focuses on optimizing the shipper's decisions regarding its multi-period contractual relations with the carriers (contract portfolio). The shipper's planning process encompasses the selection of a sequence of contracts with the carriers, together with the associated transportation plans, over multiple sub-periods of a discrete time horizon. A contract with a particular carrier allows the shipper to move its parcels to the cross-docking facilities operated by the carrier. We assume that the range of services offered by the carriers is limited to long-haul road transportation from the warehouses to their facilities and to cross-docking operations (freight consolidation, vehicle loading, and unloading). Parcel delivery services provide last-mile distribution of parcels from cross-docking facilities to customer areas.

The two-echelon network displayed in Fig. 1 is similar to those discussed in papers which focus on the role of intermediate facilities for freight-consolidation and delivery to the final customers, e.g., Lin & Wu (2001), Gendron & Semet (2009), Guastaroba, Speranza, & Vigo (2016), Hanbazazah, Abril, Shaikh, & Erkok (2018), Hanbazazah, Abril, Erkok, & Shaikh (2019), Gu et al. (2022). Similarly to these authors, we assume that demand is

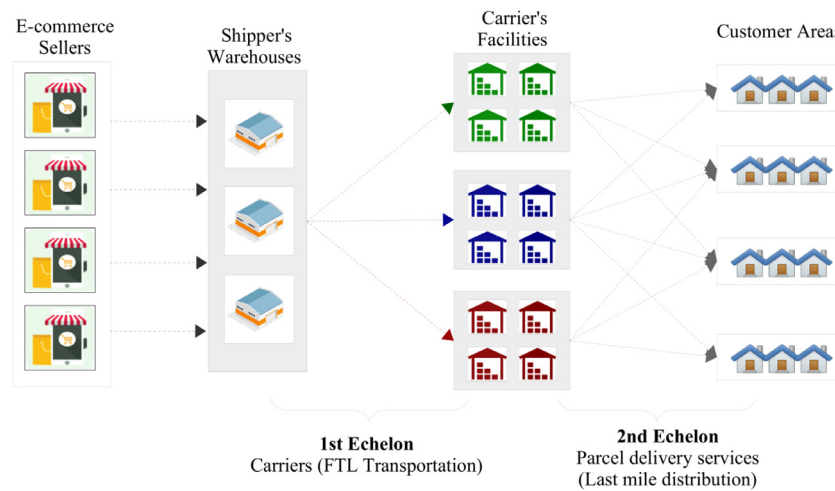


Fig. 1. Two-echelon distribution network layout.

known with full certainty, as is the case over relatively short planning horizons when delivery due dates must be strictly respected (“time-definite requirements”). The case of stochastic demand is handled in Clavijo López (2021), but is not further discussed here.

The total cost incurred by the shipper can be decomposed as follows. First, each contract with a given carrier, say carrier e , extends over a predefined duration, say H_e periods (weeks, months), and stipulates a minimum purchase commitment M_e^t , that is a minimum due fee for each period t covered by the contract, independently of the utilization of services. If the services requested at period t amount to a higher monetary value than M_e^t , then the carrier charges the actual total cost to the shipper. The minimum purchase commitment may vary in each period based on the expected demand. When a contract with carrier e expires (after H_e periods), it can be immediately renewed (i.e., access to facilities can be reopened, complying with a T/O/C/Re model, see Section 2); but it cannot be cancelled nor extended before its expiration date.

The amount charged by each carrier for transportation on a specific lane, from a warehouse to one of its cross-docking facilities, is a function of the total volume of freight. More precisely, we model it as a function of the required capacity level, that is, of the required number of full truckload shipments. In our computational experiments, we use a staircase function in order to translate a decreasing marginal price for each additional truck. We also assume that the number of shipments to a facility determines the operational costs incurred for handling and cross-docking activities at this facility. A last-mile delivery company performs the transportation of parcels from each cross-docking facility to each related customer area, and charges a fixed unit rate (per distance and weight) for this service.

In order to minimize its total distribution costs, the shipper tends to favor shipping lanes (from warehouses to intermediate facilities) which optimize transportation and handling costs. However, variable demand and seasonality effects induce fluctuations of transportation needs which, from a short-term perspective, affect the optimal selection of carriers and cross-docking points in each period. On the other hand, each contract binds the shipper to a carrier for successive periods, which means that the portfolio of contracts must be optimized from a mid-term perspective. These interactions between the tactical and operational decision levels significantly increase the difficulty of solving the resulting multi-period distribution network problem with purchase commitment contracts (MDPC).

4. Mathematical formulation

In this section, we propose a mixed-integer linear programming (MILP) formulation of the MDPC problem. We start with definitions and notations (see also Table 1).

4.1. Definitions and notations

The shipper's decision problem involves selecting carriers, from a set E of candidates, over a multi-period planning horizon $T = \{1, 2, \dots, N\}$. Every carrier $e \in E$ operates its own set of geographically dispersed facilities, I_e , each of which is equipped for cross-docking operations. From the point of view of the shipper, the collection $I = \cup_{e \in E} I_e$ forms the complete set of available cross-docking facilities. To use the facilities in I_e , the shipper must have previously signed a contract with carrier e . The conditions applying to a contract with e include its duration (H_e periods) and the minimum payment commitment (M_e^t) for every period $t \in T$. If a contract with a carrier e is in effect during period t , then the total payment due to e for period t , denoted c_e^t , is at least equal to M_e^t or if it exceeds M_e^t , the costs (handling and transportation) incurred during this period t .

For the long-haul part of the distribution network, the transportation and cross-docking capacity which can be requested by the shipper at any carrier' facility is discretized at increasing levels Q_1, Q_2, \dots, Q_l . The capacity Q_l at level l may be typically equal to the total load capacity of l vehicles expressed in weight unit. Transportation and cross-docking services are then charged depending on the capacity level requested at a given facility: the long-haul cost charged to the shipper is $F_{i,l}$ when facility i is used at level l .

For the parcel delivery part of the distribution network, we assume that the customers are located in a set K of distinct areas, where each area is small enough to be identified as a single point. The aggregated deterministic demand for parcels is D_k^t (weight units) for period t in area k . Each customer area $k \in K$ must be served on time from a subset of facilities, say $I_k \subseteq I$, considered close enough to receive an offer from a last-mile delivery operator. Conversely, vehicles departing from any facility $i \in I$ can only deliver parcels to a subset of customer areas, say, $K_i \subseteq K$. The unit cost of parcel delivery depends on the distance travelled: $U_{i,k}$ denotes the cost of transporting one weight unit from facility $i \in I$ to customer area $k \in K_i$.

Note that split deliveries from several facilities in I_k to a same area k are allowed. On the other hand, we assume that there is

Table 1
Mathematical notations for the MDPC model.

Datasets and indices		
Set	Description	Index
T	Planning horizon	t, n
E	Candidate carriers	e
I	Cross-docking facilities	i
K	Customer areas	k
I_e	Network of cross-docking facilities operated by carrier e	i
I_k	Network of cross-docking facilities that can serve customer area k	i
K_i	Customer areas that can be served from cross-docking facility i	k
E_k	Carriers that can serve customer area k	e
L	Capacity levels of cross-docking facilities	l
Input parameters		
Symbol	Description	
N	Number of periods in the planning horizon ($N = T $)	
H_e	Duration of contracts with carrier e (number of periods)	
M_e^t	Minimum Purchase Commitment with carrier e at period t (€)	
Q_l	Available capacity at level l (weight units)	
$F_{i,l}$	Cost for operating facility i at capacity level l (€)	
D_k^t	Demand of customer area k in period t (weight units)	
$U_{i,k}$	Unit transportation costs for delivery from facility i to area k (€ per unit)	
Decision variables		
Symbol	Description	
α_e^t	1 if a contract with carrier e takes effect at period t , 0 otherwise	
$v_{i,l}^t$	1 if cross-docking facility i is used at capacity level l in period t , 0 otherwise	
$q_{i,k}^t$	Demand from customer area k allocated to facility i in period t (weight units)	
c_e^t	Total fee due to carrier e in period t (€)	

enough transportation capacity in the system to satisfy demand in each period. This can be ensured, for example, by including a special carrier m representing the spot market, with enough capacity and no contractual requirement: say, $H_m = 1$, $M_m^t = 0$ for all t , large enough capacity and operating costs.

The MILP model involves four families of decision variables for every period $t \in T$ in the planning horizon. First, α_e^t is a 0–1 variable which takes value 1 if a contract with carrier $e \in E$ goes into effect in period t (and remains valid throughout periods $t, t + 1, \dots, t + H_e - 1$). Next, $v_{i,l}^t$ is a 0–1 variable which takes value 1 if facility $i \in I$ is operated at level $l \in L$ in period t . The continuous variable $q_{i,k}^t$ represents the quantity of parcels (number of weight units) shipped from facility $i \in I$ to customer area $k \in K_i$ in period t . Finally, as already introduced above, an auxiliary variable c_e^t stands for the total cost charged to the shipper by carrier e in period t .

4.2. Mixed-integer programming formulation

The multi-period distribution network design problem with purchase-commitment contracts (MDPC) can now be formulated as follows.

The objective function (1) has two components. The first one computes the fees due to all carriers over the planning horizon. The second one accounts for the parcel delivery costs. Constraints (2) enforce demand satisfaction for all customer areas in each period. Constraints (3) ensure that the total demand allocated to a facility for the last-mile delivery does not exceed its selected long-haul capacity level. Constraints (4) establish that a customer area k can only be served from facilities belonging to carriers with ongoing contracts. These constraints are actually redundant with (2)–(3), but they strengthen the linear relaxation of the formulation. Constraints (5) and (6) determine the payment due to carrier e in period t , taking into account the effective costs of services provided and the minimum purchase commitments pre-agreed by ongoing contracts, i.e., any contract signed in one of the periods from $t - H_e + 1$ to t . Constraints (7) express that a facility can be used only if there is a valid contract with its owner. Constraints

(8) specify that two contracts cannot be in effect simultaneously with the same carrier. Note that in view of constraints (7) and (8), at most one capacity level can be selected for each facility. The formulation is strengthened by the covering constraints (9) which guarantee that each customer area is served by at least one carrier. Finally, constraints (10)–(12) specify the range of the variables. We denote by Z_{opt} the optimal value of MDPC.

Objective function

$$\min Z = \sum_{t \in T} \left[\sum_{e \in E} c_e^t + \sum_{k \in K} \sum_{i \in I_k} U_{i,k} q_{i,k}^t \right] \quad (1)$$

Constraints

$$\sum_{i \in I_k} q_{i,k}^t = D_k^t \quad \forall k \in K, t \in T \quad (2)$$

$$\sum_{k \in K_i} q_{i,k}^t \leq \sum_{l \in L} Q_l v_{i,l}^t \quad \forall i \in I, t \in T \quad (3)$$

$$q_{i,k}^t \leq D_k^t \sum_{l \in L} v_{i,l}^t \quad \forall i \in I_k, k \in K, t \in T \quad (4)$$

$$c_e^t \geq \sum_{l \in L} \sum_{i \in I_e} F_{i,l} v_{i,l}^t \quad \forall e \in E, t \in T \quad (5)$$

$$c_e^t \geq M_e^t \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall e \in E, t \in T \quad (6)$$

$$\sum_{l \in L} v_{i,l}^t \leq \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall i \in I_e, e \in E, t \in T \quad (7)$$

$$\sum_{n=0}^{H_e-1} \alpha_e^{t-n} \leq 1 \quad \forall e \in E, t \in T \quad (8)$$

$$\sum_{e \in E_k} \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \geq 1 \quad \forall k \in K, t \in T \quad (9)$$

$$\alpha_e^t \in \{0, 1\} \quad \forall e \in E, t \in T \quad (10)$$

$$v_{i,l}^t \in \{0, 1\} \quad \forall i \in I, l \in L, t \in T \quad (11)$$

$$q_{i,k}^t \geq 0 \quad \forall k \in K, i \in I, t \in T \quad (12)$$

5. Complexity

5.1. NP-hardness

For specific values of the parameters, the MDPC model reduces to the well-known simple facility location problem (SFL), (see, e.g., Ghiani, Laporte, & Musmanno, 2004). Specifically, consider the special case of MDPC where

- (a) there is a single period: $N = 1$;
- (b) the contract duration is equal to one period: $H_e = 1$;
- (c) there is only one operating level ($L = \{1\}$) and Q_1 is large enough to satisfy the total demand.
- (d) a single carrier can reach all the customer areas from all its facilities: $E = \{e\}$, $I = I_e$, and $K_i = K$ for all $i \in I$;
- (e) there is no purchase commitment: $M_e^1 = 0$;

Under these conditions, the shipper has no choice but to contract with the unique carrier (i.e., to set $\alpha_e^1 = 1$), and MDPC reduces to the problem of selecting the facilities $i \in I$ (i.e., the values of the variables $v_{i,1}^1$) and the delivery quantities $q_{i,k}$ to minimize the total cost of operating the facilities and of transporting the parcels: which is precisely the SFL problem. Since SFL is NP-hard, so is MDPC.

The previous argument does not tell us anything about the difficulty of handling the contracts, since under its assumptions, the unique variable α_e^1 is trivially set to 1 in all feasible solutions of model (1)–(12). An alternative argument, therefore, might consider instead the special case of MDPC where assumptions (a)–(c) above are satisfied, and

- (f) each carrier $e \in E$ owns a single facility, and each facility can reach all the customer areas: $E = I$, $I_e = \{e\}$, and $K_i = K$ for all $i \in I$;
- (g) there is no service cost for any of the facilities: $F_{i,1} = 0$ for all $i \in I$.

In this case, again, MDPC reduces to SFL, with the variables α_e^1 , $e \in E$, indicating the contracts to be signed in order to use the associated facilities.

5.2. Decomposability

From a practical point of view, the difficulty of MDPC mainly stems from the interrelationship among various subsets of decision variables. Indeed, the selection of carrier contracts in each period $t \in T$ (expressed by the values of the variables α_e^t) restricts the candidate facilities that can be used in period t (variables $v_{i,l}^t$) and hence, the quantities that can be shipped in the same period (variables $q_{i,k}^t$). Moreover, the contracts signed with carriers usually extend over several periods which creates a linkage between the decisions made in successive periods.

These observations suggest that appropriate solution methodologies may be developed by relaxing some of the interrelationships mentioned above in various ways. In particular, when feasible values are set for the decision variables α_e^t , finding the optimal values of $(v_{i,l}^t, q_{i,k}^t)$ becomes a time-separable sub-problem which is easier to handle, even though it remains theoretically hard. This suggests that Benders decomposition can be useful for tackling the

complex structure of the problem. An exact method based on this idea is presented in Section 6.

Similarly, when we relax the constraints (6)–(9), which bind consecutive periods for each contract, the distribution sub-problem can be independently tackled for each period. Section 7 describes a heuristic algorithm based on this idea.

6. Combinatorial Benders Algorithm - (CBA)

We start this section with a brief review of combinatorial Benders approaches.

6.1. Benders with integer sub-problems

The classical Benders procedure for mixed-integer linear programming requires fixing (iteratively) the value of all integer variables and solving the remaining sub-problem, which is by construction a linear programming problem with continuous variables. Then, duality theory allows the derivation of valid (feasibility or optimality) inequalities from the optimal solution of the sub-problem, and these inequalities can be added as cuts to the formulation of the original master problem (see Benders, 1962; Rahmaniiani, Crainic, Gendreau, & Rei, 2017).

In contrast, in some recent extensions of the classical approach, only a subset of the integer variables are fixed (in our case, the α_e^t variables), and the resulting sub-problem is still an MILP problem. When this is the case, classical Benders cuts cannot be used, due to the failure of duality relations, and different approaches need to be applied.

Hooker & Ottosson (2003) coined the term *combinatorial cuts* and applied Benders-type decomposition methods for sub-problems in binary variables by developing an abstract theory of “inference dual”. Similarly in Codato & Fischetti (2006), combinatorial cuts were implemented to model and solve MIP problems involving logical implications. The cuts are so-called *canonical cuts* (see Balas & Jeroslow, 1972) of the general form:

$$\sum_{j \in C} y_j + \sum_{j \in D} (1 - y_j) \geq 1 \quad (13)$$

for appropriate binary variables y_j and subsets of indices C, D .

Similar types of cuts appear as part of “nogood” learning techniques used in constraint programming together with mixed-integer programming. The main idea is to find combinations of variable assignments that cannot be part of an optimal solution, or no good combinations (see Sandholm & Shields, 2006). Such techniques can be embedded in a branch-and-cut framework in order to reduce the size of the search tree. Examples of such approaches with integer sub-problems can be found in Botton, Fortz, Gouveia, & Poss (2013), Gendron, Scutellà, Garroppo, Nencioni, & Tavanti (2016), Fakhri, Ghatee, Fragkogios, & Saharidis (2017).

For our MDPC model, we develop a decomposition strategy using combinatorial cuts that we add to the master problem in the branch-and-cut process (similar to Gendron et al., 2016). The master problem is solved only once. The branch-and-cut process stops at some nodes for evaluating integer solutions in the sub-problem. This approach, which differs from the iterative classical procedure, is known as a *single-tree* Benders procedure (Rahmaniiani et al., 2017).

6.2. Decomposition

The mathematical formulation (1)–(12) of MDPC contains three main sets of variables, respectively associated with the selection of contracts (binary variables α_e^t), the selection of facilities (binary variables $v_{i,l}^t$), and the allocation of demand to these facilities (continuous variables $q_{i,k}^t$). The optimal value of the cost variables c_e^t is

easily deduced from the values of the other variables. We decompose this formulation into a master problem (MP), obtained by relaxing the integrality requirements on the binary variables $v_{i,l}^t$, and into a sub-problem (SP), obtained by fixing the binary variables α_e^t to valid values $\hat{\alpha}_e^t$ derived from the optimal solution of the MP.

Whenever the MP generates new values for $\hat{\alpha}_e^t$ (that is, a new tentative contract plan), an additional combinatorial cut is introduced into the MP formulation. The procedure stops when no new values for $\hat{\alpha}_e^t$ are eligible to be optimal. Let us now describe this procedure in more detail. In the next sections, we denote by Z_{inc} the incumbent, or best-known value of a solution of MDPC at any time during the procedure.

6.2.1. Master problem (MP)

The initial MP formulation is identical to the formulation (1)–(12), except that the integrality constraints (11) are relaxed and replaced by: $v_{i,l}^t \in [0, 1]$ for all $i \in I, l \in L, t \in T$.

The relaxed variables $v_{i,l}^t$ are included in the MP in order to guide the search for good values of the contract variables α_e^t . The MP is solved by a standard branch-and-cut (B&C) procedure, and its formulation is additionally enriched by combinatorial cuts as explained later on.

In the course of solving the MP by branch-and-cut, (possibly many) feasible assignments of binary values $\hat{\alpha} = (\hat{\alpha}_e^t : e \in E, t \in T)$ are identified for the contract variables in intermediate nodes of the enumeration tree. We denote by $Z_R(\hat{\alpha})$ the optimal value of the master problem when α is fixed at $\hat{\alpha}$ (the subscript R reminds us that the $v_{i,l}^t$ variables are relaxed). If $Z_{inc} \leq Z_R(\hat{\alpha})$, then $\hat{\alpha}$ is dominated by the incumbent assignment for the contract variables, and hence $\hat{\alpha}$ can be further discarded from the search process, either by pruning or adding a combinatorial cut. Otherwise, a sub-problem $SP(\hat{\alpha})$ must be solved to evaluate the quality of the contract plan associated with $\hat{\alpha}$, as explained next.

6.2.2. Sub-problem (SP)

For each feasible assignment $\hat{\alpha}$, the associated sub-problem $SP(\hat{\alpha})$ is generated by setting α_e^t to the value $\hat{\alpha}_e^t$ in the formulation of MDPC, for all $e \in E, t \in T$. Thus, $SP(\hat{\alpha})$ is an MILP model with objective function (1), with binary variables $v_{i,l}^t$ and with continuous variables $q_{i,k}^t, c_e^t$. It includes constraints (2)–(5) and (11)–(12), as well as constraints (6) and (7) which can respectively be rewritten as :

$$c_e^t \geq \begin{cases} M_e^t & \text{if } \sum_{n=0}^{H_e-1} \hat{\alpha}_e^{t-n} = 1 \\ 0 & \text{otherwise} \end{cases} \quad \forall e \in E, t \in T, \quad (14)$$

$$\sum_{l \in L} v_{i,l}^t \leq \begin{cases} 1 & \text{if } \sum_{n=0}^{H_e-1} \hat{\alpha}_e^{t-n} = 1 \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in I_e, e \in E, t \in T. \quad (15)$$

Constraints (14) and (15) respectively guarantee the minimum payment due to carriers and the availability of facilities if and only if a contract is in effect at period t .

Constraints (8)–(10) do not appear in the sub-problem formulation. Note, however, that the presence of the covering constraints (9) in the master problem ensures that all customers can be served from the facility network available under the contract plan represented by $\hat{\alpha}$. In other words, the sub-problem $SP(\hat{\alpha})$ is always feasible.

As mentioned in Section 5.2, a key feature of the sub-problem $SP(\hat{\alpha})$ is that it is separable per period: indeed, for each $t \in T$, the optimal values of (v^t, q^t, c^t) only depend on the contracts $(\hat{\alpha}^t)$ that are active in period t . Thus, if we denote by $Z_{SP}(\hat{\alpha})$ the optimal value of $SP(\hat{\alpha})$, and by $Z_{SP}(\hat{\alpha}^t)$ the optimal value of the sub-problem $SP(\hat{\alpha}^t)$ arising in period t , then

$$Z_{SP}(\hat{\alpha}) = \sum_{t \in T} Z_{SP}(\hat{\alpha}^t).$$

This property facilitates the resolution of $SP(\hat{\alpha})$, and motivates the entire decomposition scheme.

6.2.3. Combinatorial cuts

The optimal solution of $SP(\hat{\alpha})$ provides the best assignment (v^*, q^*, c^*) associated with the contracts defined by $\hat{\alpha}$. In particular, it defines a feasible solution of MDPC, and its value $Z_{SP}(\hat{\alpha})$ is an upper bound on the optimal value of MDPC. The incumbent or best-known value, i.e., Z_{inc} , is compared with $Z_{SP}(\hat{\alpha})$ and is updated if needed, that is, if $Z_{SP}(\hat{\alpha}) < Z_{inc}$. Moreover, a combinatorial cut of the following type can be added to the MP formulation:

$$\sum_{t \in T} \left(\sum_{e \in E: \hat{\alpha}_e^t = 0} \alpha_e^t + \sum_{e \in E: \hat{\alpha}_e^t = 1} (1 - \alpha_e^t) \right) \geq 1. \quad (16)$$

This combinatorial cut removes the solution $\hat{\alpha}$ from the feasible space of MP. In other words, the cut expresses that the contract plan $\hat{\alpha}$ has already been handled, and therefore $\hat{\alpha}$ can be ruled out from the subsequent search process.

The combinatorial cut (16) can be further strengthened by restricting the first double sum to those pairs (t, e) such that $\sum_{n=0}^{H_e-1} \hat{\alpha}_e^{t-n} = 0$. Indeed, suppose that a new contract plan is to be considered by switching some value $\hat{\alpha}_e^t$ from 0 to 1 (i.e., by activating a new contract at time t) so as to satisfy (16), and suppose that t is the first time period for which such a switch takes place. Then, it must be the case that $\hat{\alpha}_e^{t-n} = 0$ for all $n = 0, \dots, H_e - 1$: otherwise, it would not be feasible to start a new contract at time t (see constraints (8)).

Thus, the strengthened cut (17) can be included in the formulation instead of (16):

$$\sum_{t \in T} \left(\sum_{e \in E: \sum_{n=0}^{H_e-1} \hat{\alpha}_e^{t-n} = 0} \alpha_e^t + \sum_{e \in E: \hat{\alpha}_e^t = 1} (1 - \alpha_e^t) \right) \geq 1. \quad (17)$$

Note that the cut (17) is not separable per period.

6.3. Algorithmic procedure

The cut insertion procedure of the Combinatorial Benders Algorithm (CBA) is sketched in Fig. 2. As mentioned above, the tree is generated when solving the master problem by branch-and-cut. Commercial solvers enable the user to interrupt the MP branch-and-cut process at various nodes to launch predefined routines. In our algorithm, this practical tool is used to solve the $SP(\hat{\alpha})$ model whenever the value $Z_R(\hat{\alpha})$ of a new feasible integer solution $\hat{\alpha}$ is lower than the incumbent value Z_{inc} , as described by Algorithm 1. The solution method for the SP is summarized below by Algorithm 2.

Algorithm 1: Single-tree Benders Subroutine (Feasible $\hat{\alpha}$).

Input: $\hat{\alpha}$: Contract plan,
 $Z_R(\hat{\alpha})$: Optimal value of MP when $\alpha = \hat{\alpha}$,
 Z_{inc} : Incumbent value of MDPC.

- 1 **if** $Z_R(\hat{\alpha}) < Z_{inc}$ **then**
- 2 **solve** sub-problem $SP(\hat{\alpha})$ {see Algorithm-2}
- 3 **retrieve** value $Z_{SP}(\hat{\alpha})$
- 4 **if** $Z_{SP}(\hat{\alpha}) < Z_{inc}$ **then**
- 5 $Z_{inc} \leftarrow Z_{SP}(\hat{\alpha})$

6 **add** combinatorial cut (17) to MP

The addition of the combinatorial cut (17) to the MP formulation excludes candidate $\hat{\alpha}$ from the feasible space, whether it is optimal for MDPC or not. The branch-and-cut procedure continues

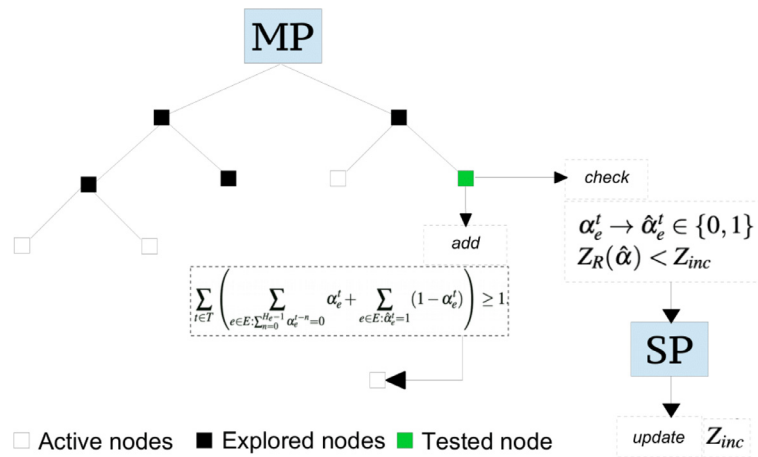


Fig. 2. Single-tree Combinatorial Benders Algorithm (cut-insertion process).

searching for better feasible solutions, and it may either generate new values for $\hat{\alpha}$, or eventually conclude that there is none left. When the solver stops, the optimality criterion is satisfied: the relaxed value $Z_R(\hat{\alpha})$ of any solution $\hat{\alpha}$ is above the incumbent value, which is therefore optimal ($Z_{inc} = Z_{opt}$). The lower bound obtained along the process is associated with the MP formulation. Therefore, it is never larger than the optimal value of the MP, which is itself a lower bound for the MDPC model.

The sub-problem SP to be solved by Algorithm 1 is a MILP that must be solved repeatedly and may take a considerable amount of time to solve. In order to facilitate its solution, as suggested in Section 6.2.2, we decompose it per period, and we solve each sub-problem $SP(\hat{\alpha}^t)$, for $t \in T$. This process can be further accelerated as follows. For the contract plan $\hat{\alpha}$ and a period t , consider the assignment $\hat{\alpha}^t$, that is, the restriction of $\hat{\alpha}$ to period t . It may very well happen (and in fact, it frequently happens in practice) that $\hat{\alpha}^t = \bar{\alpha}^t$ for another contract plan $\bar{\alpha}$ which was considered in a previous call to Algorithm 1: indeed, two contract plans $\hat{\alpha}$ and $\bar{\alpha}$ can involve the same carriers in period t , even if they differ in other periods. In such a case, obviously, $Z_{SP}(\hat{\alpha}^t) = Z_{SP}(\bar{\alpha}^t)$. In view of this observation, when solving $SP(\hat{\alpha})$, we find it advantageous to record each assignment $\hat{\alpha}^t$ together with the optimal value $Z_{SP}(\hat{\alpha}^t)$ in a set Ω^t , for each period t . When the same assignment $\hat{\alpha}^t$ is encountered again in subsequent iterations, its value is simply retrieved from Ω^t , and there is no need to solve the sub-problem again for that particular period t .

The resulting sub-problem procedure is described by the following pseudocode.

Algorithm 2: Sub-problem Procedure.

```

Input:  $\hat{\alpha}$ : Contract plan.
           $\hat{\alpha}^t$ : Restriction of the contract plan to period  $t$ .
           $\Omega^t, t \in T$ : storage sets of previously encountered
          contract plans and their optimal values.
Output:  $Z_{SP}(\hat{\alpha})$ : Optimal value of the sub-problem  $SP(\hat{\alpha})$ .
1  $Z_{SP}(\hat{\alpha}) \leftarrow 0$ 
2 for each  $t \in T$  do
3   if  $\hat{\alpha}^t \in \Omega^t$  then
4     retrieve the optimal value  $Z_{SP}(\hat{\alpha}^t)$ 
5   else
6     solve  $SP(\hat{\alpha}^t)$ 
7     add  $\hat{\alpha}^t$  and  $Z_{SP}(\hat{\alpha}^t)$  to  $\Omega^t$ 
8    $Z_{SP}(\hat{\alpha}) \leftarrow Z_{SP}(\hat{\alpha}) + Z_{SP}(\hat{\alpha}^t)$ 
9 return  $Z_{SP}(\hat{\alpha})$ 
    
```

7. Relax-and-repair heuristic

In this section, we present a heuristic optimization algorithm based on a decomposition of MDPC per period. The heuristic proceeds in two phases, which respectively consist in solving a "relaxation" of the problem for each period, then in "repairing" the obtained solutions $\hat{\alpha}^t$ to merge them over the horizon into a feasible solution $\hat{\alpha}$.

7.1. Phase 1 - Relaxation - (RL)

The idea of the first phase consists in relaxing the duration of the contracts to a single period, that is, in assuming that $H_e = 1$ for all $e \in E$. Another way of looking at this relaxation is as follows. Replace each occurrence of the expression $\sum_{n=0}^{H_e-1} \alpha_e^{t-n}$ in the formulation (1)-(12) by a single binary variable β_e^t , and add the constraints

$$\sum_{n=0}^{H_e-1} \alpha_e^{t-n} = \beta_e^t \quad \forall e \in E, t \in T. \tag{18}$$

This obviously yields an equivalent formulation of MDPC, with the interpretation that $\beta_e^t = 1$ if and only if a contract with carrier e is in effect in period t . Now, remove all constraints (18). In this way, we obtain a relaxation of MDPC where the variables β_e^t are not necessarily associated with contracts of duration H_e , but can be viewed as describing contracts of duration 1. Let us denote this relaxed problem as RL, and its optimal value as Z_{RL} . Clearly, Z_{RL} is a valid lower bound for MDPC: $Z_{RL} \leq Z_{opt}$.

The formulation RL can be decomposed and solved independently for each period $t \in T$. Let $RL(t)$ define the single-period problem at period t and let $Z_{RL}(t)$ be its optimal value. Then, $Z_{RL} = \sum_{t \in T} Z_{RL}(t)$.

In order to avoid confusion, we stress that each sub-problem $RL(t)$ involves the binary decision variables β_e^t in its formulation. Thus, it differs from the single-period sub-problem $SP(\hat{\alpha}_e^t)$ introduced in Section 6.2.2, which is associated with fixed values of the variables α_e^t . On the other hand, if $\hat{\beta}^t$ denotes the optimal value of β^t in the solution of $RL(t)$, then, by definition, $Z_{SP}(\hat{\beta}^t) = Z_{RL}(t)$.

7.2. Phase 2 - Repairing - (RP)

Consider now an optimal solution $\hat{s} = (\hat{\beta}, \hat{v}, \hat{q}, \hat{c})$ of problem RL. With a slight abuse of terminology, we say that this solution is feasible for MDPC if the succession of single-period contracts $(\hat{\beta}_e^t, t \in T)$ describes a collection of contracts with duration H_e for each $e \in E$ (meaning that there exists an assignment of values $\hat{\alpha}$ for the α -variables such that $(\hat{\alpha}, \hat{\beta})$ satisfies constraints (18)). Note

Algorithm 3: Relax-and-repair Heuristic: Phase 1.

Input: An instance of MDPC.
Output: Z_{RL} : Lower bound on the optimal value of MDPC.
 $\hat{\beta}$: A feasible contract plan for the relaxation of MDPC - (RL).

```

1  $Z_{RL} \leftarrow 0$ 
2 for each  $t \in T$  do
3   solve  $RL(t)$ , obtain  $\{\hat{\beta}^t, Z_{RL}(t)\}$ 
4    $Z_{RL} \leftarrow Z_{RL} + Z_{RL}(t)$ 
5 return  $Z_{RL}, \hat{\beta} = (\hat{\beta}^t : t \in T)$ 

```

that this is easy to check. When this is the case, then the solution \hat{s} (or more rigorously, $(\hat{\alpha}, \hat{v}, \hat{q}, \hat{c})$) is necessarily optimal for the original MDPC model as it satisfies constraints (8), i.e., $Z_{RL} = Z_{opt}$, and the second phase of the heuristic simply returns this solution.

On the other hand, if the conditions on the contract durations H_e are not respected, then the solution \hat{s} is infeasible for MDPC and Z_{RL} only provides a lower bound for Z_{opt} . The second phase then consists in “repairing” the infeasible solution to transform it into a feasible one. The repair heuristic RP iterates forward over periods $t \in T$. For each carrier e such that $\hat{\beta}_e^t = 0$, it examines whether the single-period contracts $(\hat{\beta}_e^{t-1}, \hat{\beta}_e^{t-2}, \dots)$ opened in previous periods define a contract of duration H_e (or a sequence of contracts of duration H_e). When this is not the case, the unfinished contract can be either *Completed* up to H_e periods, or *Removed*.

We next give a more formal description in Algorithm 4 of the phase 2 procedure (RP) with descriptive comments in brackets.

The *Completion* option considers the number of periods needed for the unfinished contract to get completed (namely, $H_e - r$ periods), starting from time t . The local variable *CurrentCostC* sums up the optimal cost values $Z_{SP}(\hat{\beta}^t), \dots, Z_{SP}(\hat{\beta}^{t+H_e-r-1})$ of the respective sub-problems $SP(\hat{\beta}^t), \dots, SP(\hat{\beta}^{t+H_e-r-1})$. Next, the algorithm temporarily sets the variables $\hat{\beta}^t, \dots, \hat{\beta}^{t+H_e-r-1}$ to one and recomputes the costs $Z_{SP}(\hat{\beta}^t), \dots, Z_{SP}(\hat{\beta}^{t+H_e-r-1})$ by solving the sub-problems with these modified values. The difference is saved in parameter $\Delta CostC$ and the values of $\hat{\beta}^t, \dots, \hat{\beta}^{t+H_e-r-1}$ are set back to zero.

The *Removal* option considers the elapsed duration of the unfinished contract running at period time $t - 1$ (namely, r) in order to discard it. The local variable *CurrentCostR* sums up the optimal cost values $Z_{SP}(\hat{\beta}^{t-r}), \dots, Z_{SP}(\hat{\beta}^{t-1})$ of the respective sub-problems $SP(\hat{\beta}^{t-r}), \dots, SP(\hat{\beta}^{t-1})$. Next, the algorithm temporarily sets the variables $\hat{\beta}^{t-r}, \dots, \hat{\beta}^{t-1}$ to zero and recomputes the costs $Z_{SP}(\hat{\beta}^{t-r}), \dots, Z_{SP}(\hat{\beta}^{t-1})$ by solving the sub-problem with the modified values. The difference is saved in parameter $\Delta CostR$, which is set to a large value in case of infeasibility induced by at least one uncovered customer area. The incremental cost of each option ($\Delta CostC, \Delta CostR$) is estimated, and the least expensive one is implemented.

It may be interesting to note that in a “real-world” setting, the heuristic could be implemented in a rolling-horizon fashion. More precisely, when using the relax-and-repair heuristic, we do not really need to solve the problem over the complete horizon of N periods in order to make decisions in any given period: indeed, the decisions in period t only depend on the contracts signed in the previous periods and on the demand for periods $t, t + 1, \dots, t + \max\{H_e : e \in E\} - 1$. Thus, the heuristic can be implemented with a forecast horizon of $\max\{H_e : e \in E\}$ periods.

8. Experimental design and performance measures

We have conducted a series of experiments comparing the proposed algorithms with a commercial solver on multiple sets of in-

Algorithm 4: Relax-and-repair Heuristic: Phase 2.

Input: $\hat{\beta}$: A feasible contract plan for the relaxation of MDPC (RL)
Output: $\hat{\alpha}_{RP}$: A feasible contract plan.
 Z_{RP} : Value of the heuristic solution.

```

1  $Z_{RP} \leftarrow Z_{RL}$ 
2 for each  $t \in T, t \neq 1$  do
3   for each  $e \in E$  do
4     if  $\hat{\beta}_e^t = 0$  and  $\hat{\beta}_e^{t-1} = 1$  then
5       /* {potentially incomplete contract} */
6       /* {determine  $s$ , the first period such that
7          $\hat{\beta}_e^s = \dots = \hat{\beta}_e^{t-1} = 1$ } */
8        $s \leftarrow t - 1$ 
9       while  $s > 1$  and  $\hat{\beta}_e^{s-1} = 1$  do
10         $s \leftarrow s - 1$ 
11         $r \leftarrow (t - s) \bmod H_e$ 
12        /* { $r$ : elapsed duration (mod  $H_e$ ) of the contract running
13          at period  $t - 1$ } */
14        if  $r \neq 0$  then
15          /* {the sequence  $\hat{\beta}_e^s, \dots, \hat{\beta}_e^t$  is not feasible} */
16          /* {estimate the extra cost of completing the
17            contract} */
18           $CurrentCostC \leftarrow \sum_{k=0}^{H_e-r-1} Z_{SP}(\hat{\beta}^{t+k})$ 
19          temporarily set  $\hat{\beta}_e^t, \dots, \hat{\beta}_e^{t+H_e-r-1} \leftarrow 1$ 
20          /* temporarily complete the contract */
21           $NewCostC \leftarrow \sum_{k=0}^{H_e-r-1} Z_{SP}(\hat{\beta}^{t+k})$ 
22           $\Delta CostC \leftarrow NewCostC - CurrentCostC$ 
23          reset the previous values of
24           $\hat{\beta}_e^t, \dots, \hat{\beta}_e^{t+H_e-r-1} \leftarrow 0$ 
25          /* {estimate the extra cost of removing the contract}
26          */
27           $CurrentCostR \leftarrow \sum_{k=1}^r Z_{SP}(\hat{\beta}^{t-k})$ 
28          temporarily set  $\hat{\beta}_e^{t-r}, \dots, \hat{\beta}_e^{t-1} \leftarrow 0$ 
29          /* temporarily remove the contract */
30          if  $SP(\hat{\beta}^{t-k})$  is feasible for  $k = 1$  to  $r$  then
31             $NewCostR \leftarrow \sum_{k=1}^r Z_{SP}(\hat{\beta}^{t-k})$ 
32          else
33             $NewCostR \leftarrow \infty$ 
34           $\Delta CostR \leftarrow NewCostR - CurrentCostR$ 
35          reset the previous values of  $\hat{\beta}_e^{t-r}, \dots, \hat{\beta}_e^{t-1} \leftarrow 1$ 
36          if  $\Delta CostC \leq \Delta CostR$  then
37            /* {the completion option is preferred} */
38             $\hat{\beta}_e^t, \dots, \hat{\beta}_e^{t+H_e-r-1} \leftarrow 1$ 
39             $Z_{RP} \leftarrow Z_{RP} + \Delta CostC$ 
40          else
41            /* {the removal option is preferred} */
42             $\hat{\beta}_e^{t-r}, \dots, \hat{\beta}_e^{t-1} \leftarrow 0$ 
43             $Z_{RP} \leftarrow Z_{RP} + \Delta CostR$ 
31 return  $Z_{RP}, \hat{\alpha}_{RP}$ 

```

Table 2
Description of instance classes.

Instances	Small						Medium						Large					
Carriers $ E $	4	4	6	6	8	8	8	8	10	10	12	12	8	8	10	10	12	12
Periods $ T $	4	6	6	8	8	10	8	10	10	12	12	14	8	10	10	12	12	14
Customers $ K $	100						200						300					
Facilities per carrier $[i^-, i^+]$	[2,4]						[3,5]						[4,6]					
Contract duration H_e	2, 3, 4						2, 3, 4						2, 3, 4					

stances. This section first describes the set of instances that we generated, then the solution approaches considered, and finally the performance measures.

8.1. Instances

Data about the location of facilities and of customers, costs, demand and contract terms were randomly generated based on the following assumptions and parameter values.

Location and scope of sites

- The shipper operates two warehouses.
- The warehouses and customer areas are randomly uniformly located in a 2-dimensional (1000 × 1000) grid.
- The carriers' facilities are randomly located in a (700 × 700) centered area of the grid.
- For each carrier e , the size of I_e is randomly chosen from the uniform discrete distribution on an interval $[i^-, i^+]$: $U[i^-, i^+]$ (see Table 2).
- Customers can only be served from carriers' facilities that are within an Euclidean distance of 500 units. No carrier can supply all customers.

Service costs

- The long-haul cost and cross-docking services $F_{i,l}$ for operating facility i at capacity level l is of the form $F_{i,l} = C_i + T_{i,l}$, where C_i is the fixed operational cost of facility i , and $T_{i,l}$ is the cost charged by the carrier for (full truckload) shipment requiring l trucks from the shipper's warehouses to facility i and for operating it at level l .
- The fixed operational cost C_i is set randomly from a discrete uniform distribution $U[500, 1000]$.
- For a given operating level l , the variable cost $T_{i,l}$ for long-haul transportation is proportional to the Euclidean distance from facility i to the closest warehouse. As a function of the level l , it is modelled by a staircase function with decreasing marginal costs.
- The cost charged by the parcel delivery company for transporting one unit of good from facility i to customer area k , that is, $U_{i,k}$, is taken equal to the Euclidean distance from i to k .

Customers demand

- The planning horizon is subdivided into three demand seasons, starting with low, then high, and finally mid season. For each customer region $k \in K$ and each period $t \in T$, the demand quantity D_k^t (in weight units) is issued from a uniform distribution which depends on the season: $U[0.1, 0.4]$ in low season, $U[0.35, 0.65]$ in mid season, and $U[0.6, 0.9]$ in high season.

Contract terms and conditions

- In a given instance, all contracts have the same fixed duration $H_e \in \{2, 3, 4\}$, for all $e \in E$.
- Each carrier has enough available capacity to meet the demand of all customer areas that can be served from its facilities.
- The MPC M_e^t is equal to 10% of the total capacity reservation fee, that is, the minimum fee that would be charged by carrier e at period t if it were assigned all the demand ($\sum_{k:e \in E_k} D_k^t$) that it can possibly handle through its network of facilities

The main parameters that determine the size of the instances are the number of carriers $|E|$, the number of periods $N = |T|$, the number of customer areas $|K|$, and the range $[i^-, i^+]$ of the number of facilities per carrier. The instances are partitioned into three classes according to the value of these parameters, namely *small*, *medium* and *large* instances. In each class, 6 combinations of parameter values are considered as displayed in Table 2. Moreover, we are also interested in analyzing the effect of the contract duration on the difficulty to solve the MDPC model. Hence, three different values of H_e , namely, $H_e = 2, 3, 4$, are considered for each instance class. For each combination of parameters ($|E|, |T|, |K|, [i^-, i^+], H_e$), a set of five instances was randomly generated, for a grand total of $3 \times 6 \times 3 \times 5 = 270$ instances.

8.2. Algorithms

The computational study aims to evaluate the performance of different algorithms for solving the MDPC model. We compare the algorithms described in Sections 6 and 7 with two state-of-the-art generic algorithms, namely: CPLEX default implementations of branch-and-cut and of Benders decomposition. In contrast with our combinatorial Benders decomposition approach, CPLEX Benders decomposition classically separates the integer variables ($\alpha_k^t, v_{i,l}^t$), which are included in the master problem, from the continuous variables ($q_{i,k}^t$), which are handled in the LP sub-problem.

Thus, we consider four methods, respectively labeled as:

- (CP-B&C) - CPLEX Branch-and-Cut;
- (CP-Bend) - CPLEX Benders;
- (CBA) - Combinatorial Benders Decomposition Algorithm (Section 6);
- (RL-RP) - Relax-and-repair heuristic (Section 7).

The algorithms were coded in Java using CPLEX 12.8 Concert Technology. All tests were performed using four core processors Intel E5-2650 with 2.0 GHz and 16 GB of RAM (4GB/Core).

8.3. Performance measures

Each algorithm is run for a maximum of 3600 seconds on each of the 270 instances. The algorithms are compared based on two performance measures:

- The *efficiency* of algorithm A is assessed by its total running time, and by its running time until it obtains its best overall solution (within the time limit):

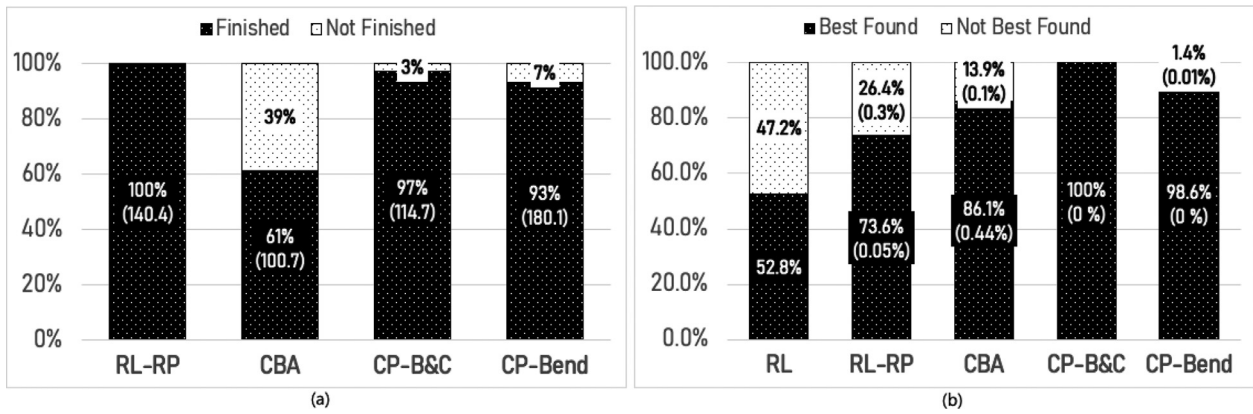


Fig. 3. Performance of algorithms on small instances.

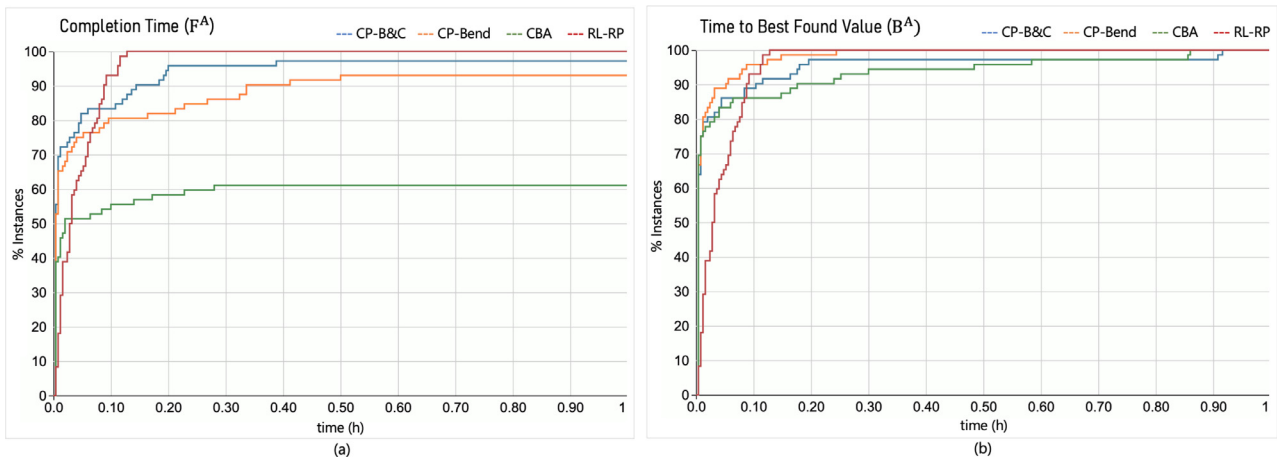


Fig. 4. Efficiency: performance profile of total running time F^A and of running time until best found solution B^A - small instances.

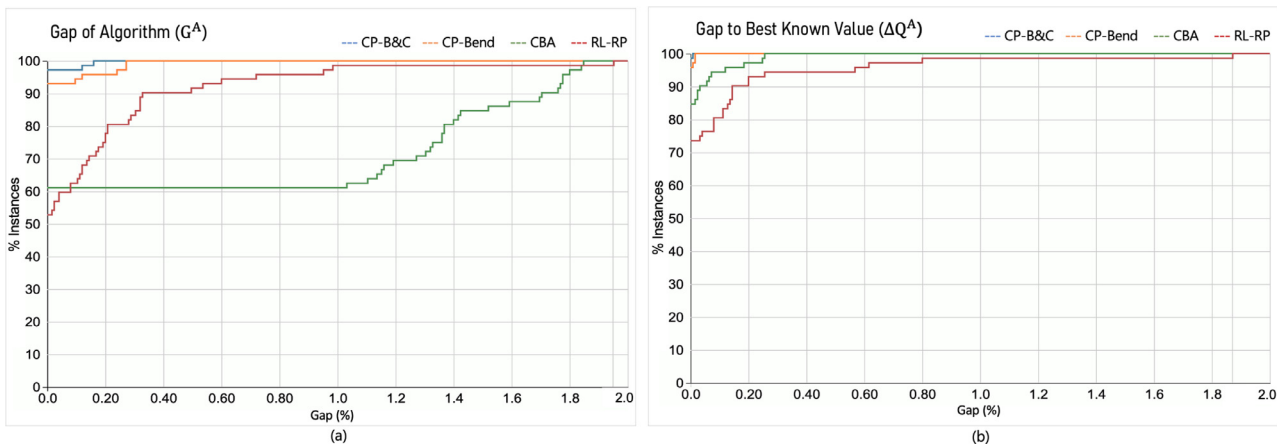


Fig. 5. Effectiveness: performance profile of optimality gap G^A and of gap to best known value ΔQ^A - small instances.

1. F^A : Running time of A until termination.
2. B^A : Running time of A until it obtains its best solution.

- The effectiveness of algorithm A is assessed by two distinct metrics, namely: the gap with respect to the best lower bound provided by the algorithm itself, and the gap with respect to the best solution found by any algorithm. Let Q^A be the best value obtained by algorithm A , L^A be the best lower bound obtained by A , and Q^{best} be the best solution value obtained by one of the algorithms. Then:

1. $G^A = \frac{Q^A - L^A}{Q^A} (\times 100\%) \rightarrow$ Relative optimality gap reported by A .
2. $\Delta Q^A = \frac{Q^A - Q^{best}}{Q^A} (\times 100\%) \rightarrow$ Relative gap of A with respect to the best known value.

There always holds $0 \leq G^A \leq 1$ and $0 \leq \Delta Q^A \leq 1$. If an exact algorithm A (CP-B&C, CP-Bend, CBA) terminates before the time limit, then the best found solution is optimal and $G^A = \Delta Q^A = 0$. The same holds for RL-RP when it finds a feasible solution in Phase 1.

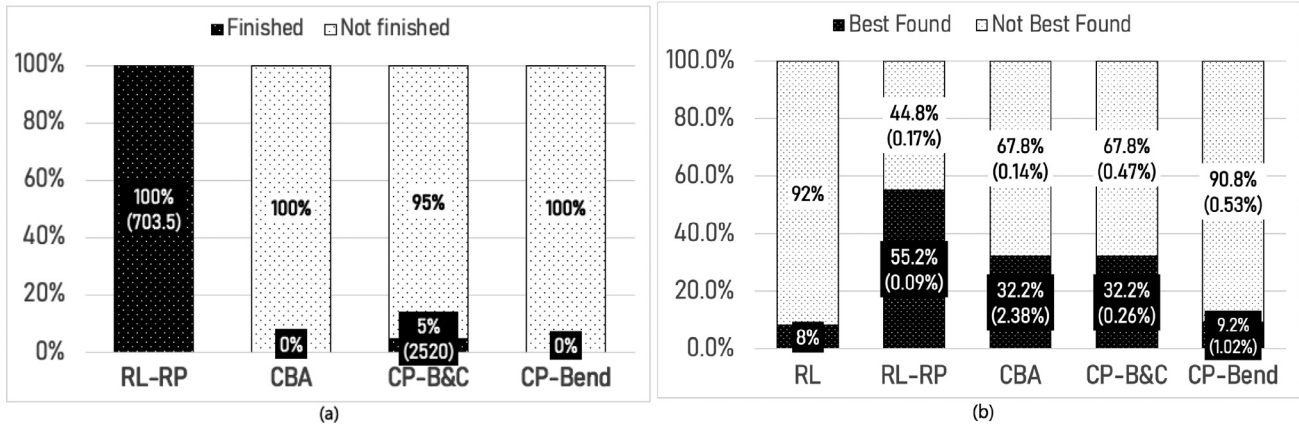


Fig. 6. Performance of algorithms on medium instances.

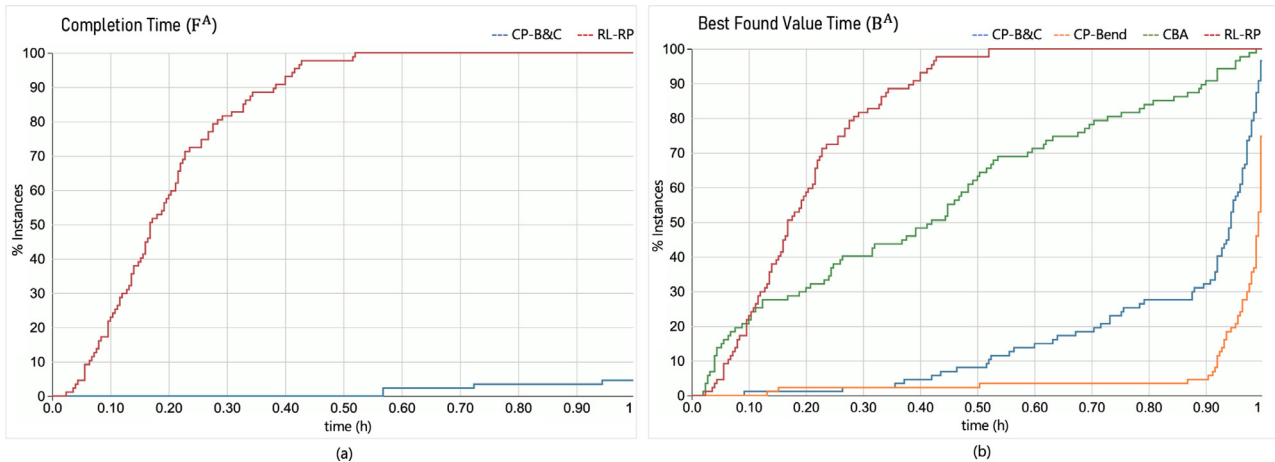


Fig. 7. Efficiency: performance profile of total running time F^A and of running time until best found solution B^A - medium instances.

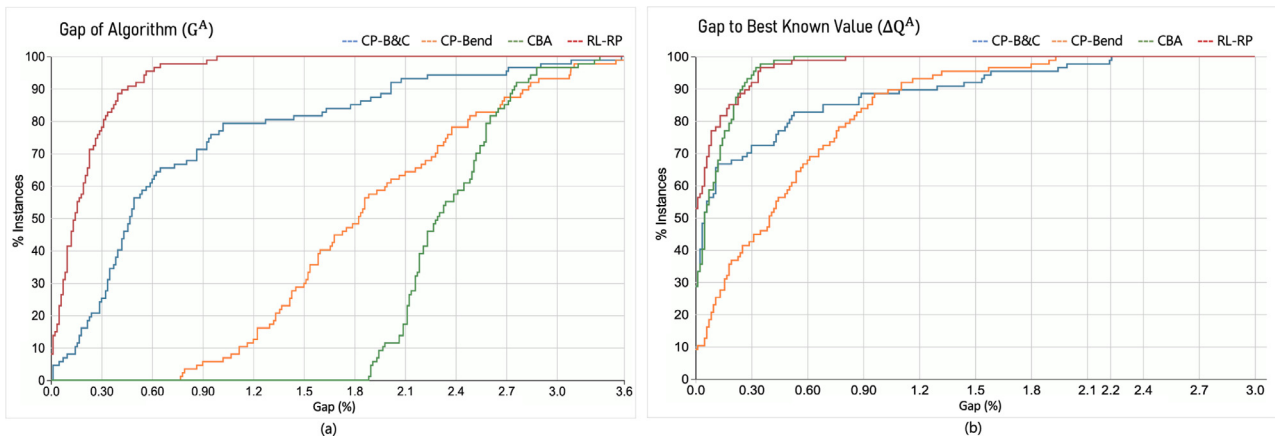


Fig. 8. Effectiveness: performance profile of optimality gap G^A and of gap to best known value ΔQ^A - medium instances.

9. Computational performance

We report in this section on the computational results obtained for instances of different sizes and with different contract durations.

9.1. Small size instances

A first overview of the results for the set of 90 small instances is displayed in Fig. 3. On the left side, the bar chart (a) presents results with respect to the *efficiency* criteria. For each algorithm, it

shows the percentage of instances for which the algorithm terminates within the time limit (3600 seconds). The average running time (in seconds) for these instances is shown in parentheses. On the right side, the bar chart (b) summarizes results in terms of the *effectiveness* criteria. It shows the percentage of instances for which each algorithm A obtains the best overall solution ($\Delta Q^A = 0$). The first bar in Fig. 3(b), marked “RL”, shows the percentage of instances for which the contract-duration relaxed model yields a feasible (and hence, optimal) solution in the first phase of the relax-and-repair heuristic (see Section 7). For each algorithm A , the average gap G^A over the set of instances for which A finds the best solution is shown in parentheses. On the other hand, when A does

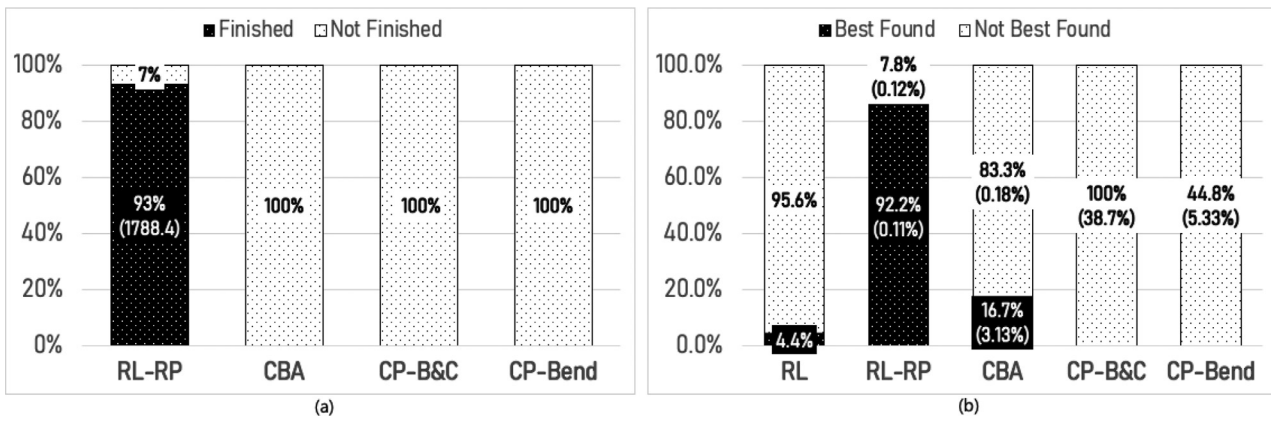


Fig. 9. Performance of algorithms on large instances.

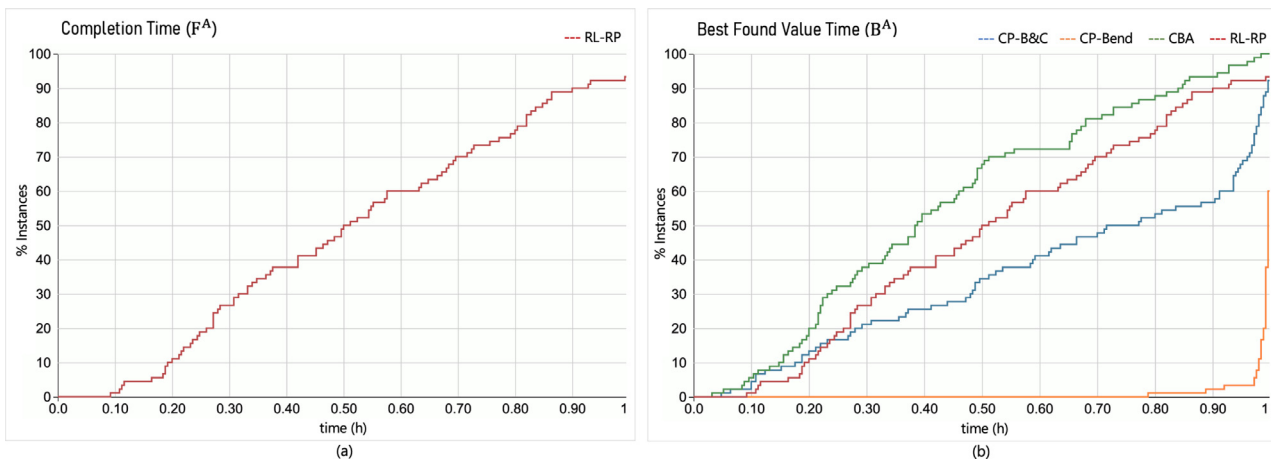


Fig. 10. Efficiency: performance profile of total running time F^A and of running time until best found solution B^A - large instances.

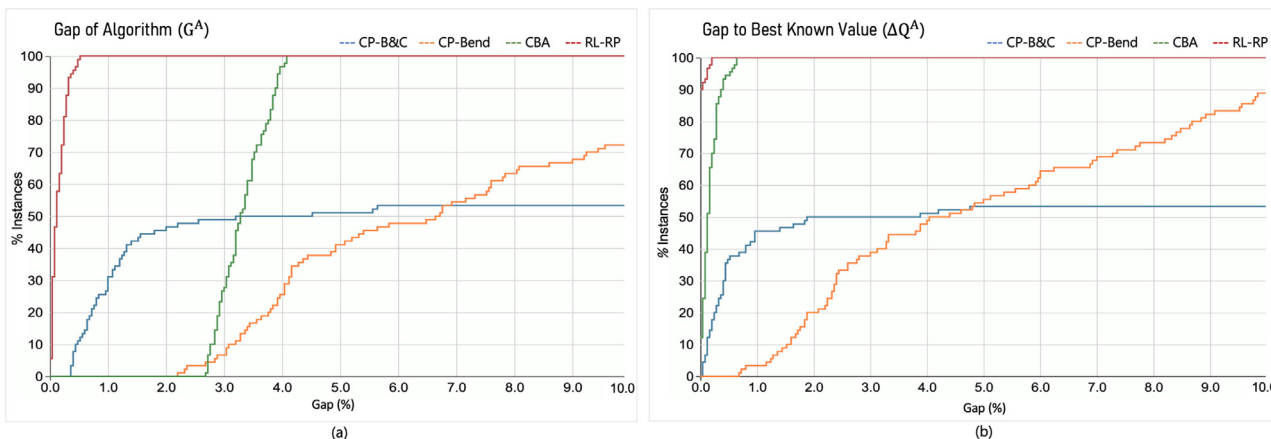


Fig. 11. Effectiveness: performance profile of optimality gap G^A and of gap to best known value ΔQ^A - large instances.

not find the best solution, the average relative difference ΔQ^A to the best-known value is displayed in parentheses.

Another overview of the results is provided in Figs. 4 and 5, which display the performance profile of each algorithm for the criteria F^A , B^A , G^A , and ΔQ^A . The performance profile can be viewed as the empirical distribution function of the performance criterion of interest (Dolan & Moré, 2002). More precisely, for an algorithm A , a criterion C^A , and a value x on the horizontal axis, the performance profile indicates the percentage of instances for which $C^A \leq x$. The profiles allow for easy visualization and comparison of the performance of different algorithms over a range of instances (in the present case, the collection of 90 small instances).

The relax-and-repair heuristic generates the optimal solution for approximately 53% of the instances in its Phase 1 - RL, by relaxing the contract-duration constraints. Moreover, this percentage reaches 73% after repairing infeasibilities in Phase 2 - RP (Figs. 3(b) and 5(b)). For the remaining instances, the heuristic provides suboptimal solutions with a relative gap (G^{RL-RP} or ΔQ^{RL-RP}) smaller than 1% for most of the small instances (Fig. 5).

CPLEX is able to solve almost all small instances (97% for B&C, 93% for CP-Bend) to optimality in less than one hour. In both cases, CPLEX is faster than the heuristic RL-RP for 80–85% of the instances (Fig. 4(a)).

Finally, the combinatorial Benders algorithm (CBA) terminates in 61% of the cases only (Figs. 3(a) and 4(a)), although it generally finds the optimal value within the time limit (for 86% of the instances, see Figs. 3(b) and 5(b)), in about the same time as CPLEX (Fig. 5(b)). Its main weakness lies in slightly weaker lower bounds which do not completely close the optimality gap, as shown by the value of the optimality gap: $G^{CBA} = 0.44\%$ on average when CBA finds the optimal solution (see also Fig. 5(a)). It is interesting to observe, however, that the performance profile of ΔQ^A lies lower for the heuristic RL-RP than for CBA, which means that CBA is generally able to find better feasible solutions.

9.2. Medium size instances

Figs. 6, 7 and 8 present the aggregated results for the set of medium size instances, using the same conventions as in the previous Section 9.1.

For most instances, none of the exact methods terminate within one hour, proving optimality. The heuristic RL-RP always stops before the time limit (in less than 1900 seconds - Figs. 6(a), 7(a)). It obtains an exact solution for 8% of the instances (Figs. 6(b), 8(a)). RL-RP provides solutions that are at least as good as the exact methods for 55.2% of the instances and mostly outperforms CPLEX methods (CP-B&C, CP-Bend) (see Fig. 8(b)). We observe that the heuristic method has a small integrality gap, below 1% (Fig. 8(a)). This means that the lower bound computed in the first phase is tight.

Among exact methods, the combinatorial algorithm CBA provides the best feasible solution faster than the two CPLEX methods (see performance profile B^{CBA} in Fig. 7(b)). CPLEX methods tend to continuously improve their incumbent solution, slowly closing the gap to their respective LBs. However, the lower bound computed by the CBA method alone remains relatively weak, which translates into values of G^{CBA} between 1.9% and 3.5% (see performance profile G^{CBA} in Fig. 8(a)).

9.3. Large size instances

The results of our tests on large instances are shown in Figs. 9, 10 and 11. The trends observed for medium size instances

are again present and accentuated here. Even though its completion time is steadily increasing, the heuristic RL-RP is the only one that terminates (for 93% of the instances) within one hour of running time (Fig. 9(a)). The remaining 7% take on average 16.7% more than 1 hour, and the last instance takes up to 29% above the time limit. Yet, RL-RP provides the best solution for more than 90% of the instances (Fig. 9(b)). The first phase RL provides an exact solution for 4.4% of the large instances. However, the optimality gap G^{RL-RP} is smaller than for any other method and in fact always smaller than 0.5% (Fig. 11(a)).

Only our Combinatorial Benders Algorithm can sometimes match or improve the solution found by RL-RP (for about 16% of the large instances). For the remaining instances, the relative gap ΔQ^{CBA} to the heuristic value is 0.18% on average. Furthermore, it is interesting to note that CBA generally finds its best solution faster than any other method, including the heuristic (Fig. 10(b)).

On large size instances, CP-B&C frequently (41%) fails to detect a single feasible solution (upper bound). As a consequence, $\Delta Q^{CP-B&C}$ and $G^{CP-B&C}$ are very large for these instances (Fig. 11). Conversely, CP-Bend (classical) decomposition is still able to find one. However, the lower bound provided by CP-B&C is systematically tighter than that of CP-Bend. On average, $L^{CP-B&C}$ is 1.26% and 2.48% closer to the optimal value than $L^{CP-Bend}$ and L^{CBA} respectively, while it is looser than L^{RL-RP} in 0.61%.

9.4. Variations of contract duration

Finally, we examine the behavior of the computational time of all algorithms with respect to variations in the contract duration ($H_e = 2, 3, 4$). The contracts establish the only link between the different periods (via constraints (7)-(9) of the MILP formulation) and thus, at first sight, they represent the main complexifying component of the MDPC model. On the other hand, some of the methods are based on a decomposition of the problem per time period, and it is not clear whether the efficiency of these algorithms is significantly affected by the value of H_e .

The heuristic RL-RP, in particular, is based in its first phase on the relaxation of the contract duration, and hence this phase is insensitive to the value of H_e . One could argue that as H_e becomes

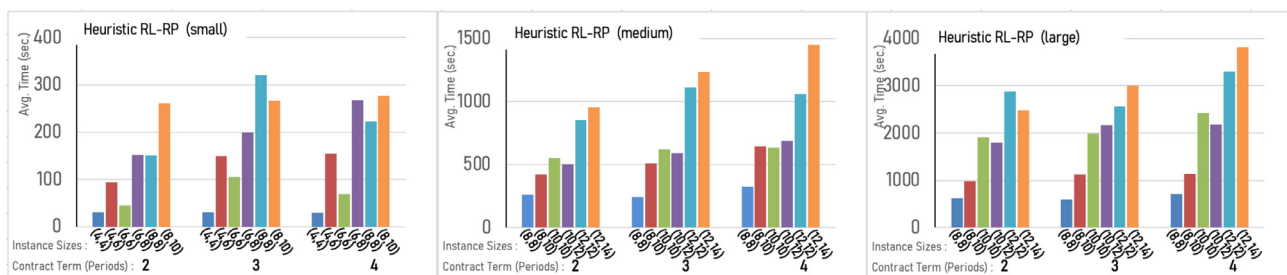


Fig. 12. Variation of the computing time of RL-RP with respect to H_e and $(|E|, |T|)$.

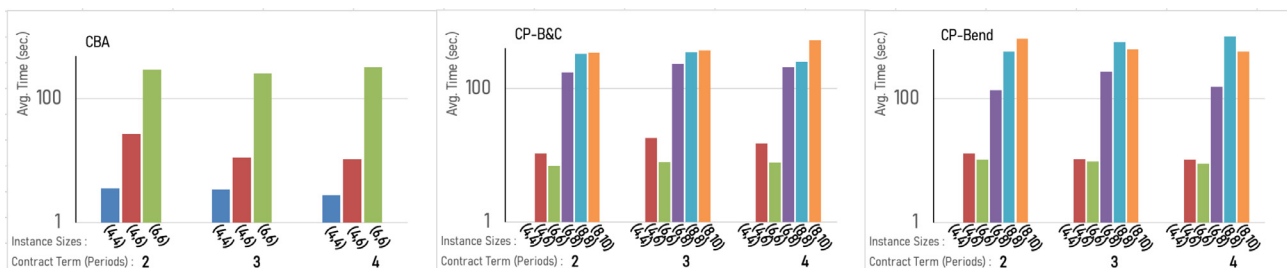


Fig. 13. Variation of the computing time of exact methods with respect to H_e and $(|E|, |T|)$.

Table 3
Variation of total cost and carriers' cost as a function of contract duration .

Cost increase	$H_e : 1 \rightarrow 2$			$H_e : 2 \rightarrow 3$		
	total: average	total: (min,max)	carriers: average	total: average	total: (min,max)	carriers: average
Small	0.02%	(0.00, 0.18)%	0.01%	0.04%	(-0.11, 0.50)%	-0.15%
Medium	0.08%	(0.00, 0.38)%	0.59%	0.04%	(-0.07, 0.22)%	-0.25%
Large	0.11%	(0.00, 0.39)%	0.31%	0.03%	(-0.15, 0.18)%	-0.11%
Cost increase	$H_e : 3 \rightarrow 4$			$H_e : 4 \rightarrow LT$		
	total: average	total: (min,max)	carriers: average	total: average	total: (min,max)	carriers: average
Small	0.05%	(-0.17, 0.46)%	-0.41%	0.57%	(0.00, 1.92)%	1.16%
Medium	-0.01%	(-0.17, 0.13)%	0.12%	0.21%	(-0.01, 1.71)%	-0.13%
Large	0.00%	(-0.20, 0.21)%	0.03%	0.94%	(0.00, 4.19)%	2.98%

larger, the first phase of RL is more likely to result in an infeasible contract plan, so the RP repair mechanism should be called more frequently in the second phase. But conversely, when a contract with carrier e is signed at period t ($\alpha_e^t = 1$), its consequences extend over more periods, and hence the decision space is reduced when H_e increases. The overall outcome of these effects is hard to predict a priori.

Experimentally, the multi-chart displayed in Fig. 12 shows that the running time of RL-RP tends to increase with the value of H_e , for all instance sizes. The variation is noticeable when H_e increases from 2 to 3 periods and from 3 to 4 periods (with a few exceptions). This behavior is particularly significant for large instances.

For the exact algorithms, most medium and large size instances cannot be solved within the allocated time limit, and are therefore not considered in this analysis. Fig. 13 displays the behavior of these algorithms for the set of small instances only. The trend here is not clear. For the considered instance, there is little change in the computational time when H_e increases. This may simply be a sign that the instances are too easy to draw meaningful conclusions.

10. Managerial insights

In this section, we discuss some economic features of the best solutions obtained in our experiments. As follows from the previous section, these solutions are close to optimal ones.

10.1. Impact of contract duration

Table 3 shows the variation of the total cost and of the cost of carriers when the contract duration H_e increases from 1 to 2, or 2 to 3, or 3 to 4, or 4 to LT. Here, LT refers to the situation where the shipper engages into “long term” agreements with the carriers. More precisely, this case is modelled by substituting the expression $\sum_{n=0}^{H_e-1} \alpha_e^{t-n}$ by α_e in constraints (6)–(9) of the MIP model (i.e., removing the time-dimension from contract decision variables: when $\alpha_e = 1$, a contract is in effect with carrier e over the complete horizon). The table also displays the minimum and maximum variation of the total cost over all 30 instances of a given size. Finally, the table shows the (average) variation of the cost of carriers for long-haul transportation and cross-docking services, i.e., of the sum $\sum_{t \in T} \sum_{e \in E} c_e^t$ that appears in the objective function (1).

Note that extending the contract length intuitively entails less flexibility to modify the set of active carriers, and may therefore lead to more costly solutions. However, this is not necessarily the case since dynamic variations of demand and of other parameters may occasionally cause a succession or non-renewal of longer contracts to be more advantageous than a sequence of shorter ones. Generally speaking, in our experiments, variations of H_e lead to

Table 4
Increase of total cost with respect to the case without purchase commitments.

	$H_e = 1$	$H_e = 2$	$H_e = 3$	$H_e = 4$
Small	0.01%	0.03%	0.07%	0.12%
Medium	0.13%	0.21%	0.24%	0.24%
Large	0.37%	0.48%	0.51%	0.51%

Table 5
MPC utilization rate (and share of MPC in the carriers cost).

	$H_e = 2$
Small	99.87% (33.82%)
Medium	97.10% (48.63%)
Large	95.62% (53.32%)

small differences in total cost (less than 1% on average, less than 4.2% in the worst case). These small differences are observed for all types of instances. The cost of carriers may show larger fluctuations (up to 14% increases for some instances – not shown in the table), but these variations are amortized in the total cost.

10.2. Impact of purchase commitment

We next assess the impact of minimum purchase commitments on the total cost. For this purpose, we compute the difference between the total cost of solutions with or without purchase commitment quantities. The second case is obtained by setting $M_e^t = 0$ for all $e \in E$ and $t \in T$, meaning that the shipper only pays for the resources that it actually uses. Table 4 displays the average cost differences computed over instances of a same class (small, medium, or large) for different values of H_e . (Columns associated with $H_e = 2, 3, 4$ can be directly deduced from column $H_e = 1$ and from Table 3, but we display them here for convenience.)

The table shows that the existence of minimum purchase commitments results, on average, in a very small increase of the total cost. This suggests that in the solutions selected by the optimization process, the shipper manages to take full advantage of the capacity that it reserves with each carrier through the MPC contracts, so that the payments M_e^t are not “wasted”. This is confirmed by the results displayed in Table 5. The figures in this table are obtained as follows. For the solution of an arbitrary instance, we let C be the collection of pairs (e, t) such that a contract with carrier e is in effect in period t . We record the proportion of pairs (e, t) in C such that the cost c_e^t paid to the carrier e is equal to $\sum_{i \in L} \sum_{i \in e} F_{i,l} v_{i,l}^t$ (meaning that the MPC is fully utilized by the shipper; see constraints (5)–(6) in the MIP model). Table 5 first displays this proportion, averaged over all instances of the same class. Next to this value, between parentheses, we also indicate the share of the carriers costs due to the minimum purchase commitment, that is, the

Table 6
% of carriers selected (% permanent) – % selected in low and high season.

% Carriers	$H_e = 1$			$H_e = 2$		
	Full horizon	Season		Full horizon	Season	
		Low	High		Low	High
Small	80.30% (50.6%)	52.3%	65.0%	78.61% (51.6%)	52.6%	64.8%
Medium	81.36% (32.2%)	43.8%	57.1%	77.97% (37.3%)	43.8%	58.5%
Large	85.46% (32.1%)	49.1%	57.9%	80.63% (39.0%)	48.8%	59.6%
% Carriers	$H_e = 3$			$H_e = 4$		
	Full horizon	Season		Full horizon	Season	
		Low	High		Low	High
Small	77.36% (54.6%)	51.3%	65.3%	74.17% (60.2%)	52.0%	66.0%
Medium	73.69% (41.0%)	43.6%	57.9%	72.89% (44.8%)	43.6%	58.5%
Large	78.85% (37.8%)	48.1%	60.4%	77.18% (44.3%)	47.9%	60.0%

quantity $100 \sum_{(e,t) \in C} \frac{M_e^t}{c_e^t}$ (averaged over 30 instances). We only display here the values obtained when $H_e = 2$ since the other cases are very similar in view of Table 3.

We see that when a contract is in effect, the purchase commitment is exceeded in more than 95% of the cases. Accordingly, the MPC only represents 33% to 53% of the carriers' cost. Even though this share is significant and tends to increase with the instance size, it remains much smaller than 100%.

Of course, these observations are very specific to the numerical instances that we have generated. Taken together, however, the results in this section indicate that the existence of mid-term contracts involving minimum purchase commitments does not heavily penalize the shipper. By carefully selecting its partners, the shipper can avoid major cost impacts in the long term, whereas the carriers benefit from the organizational advantages of the contracts (reduced uncertainty, improved communication and trust, etc.; see Brusset, 2009; Brusset, 2010).

10.3. Selection of suppliers

The existence and the duration of contracts are likely to affect the number of carriers selected over the planning horizon. Table 6 displays the average percentage of carriers activated by the shipper in the optimal (or best found) plans. It also gives the percentage of permanent carriers (i.e., those who remain active in all periods of the horizon) and the percentage of carriers selected in low (respectively, high) season.

We see that the selected carriers represent on average between 72% and 86% of the number of available candidates. As one might expect, this proportion tends to decrease with the duration of the contracts, and to increase slightly with the instance size. More interesting is the relatively large gap between, on the one hand, the percentage of carriers whose services are utilized at least once, and on the other hand, the percentage of permanent carriers. For medium or large instances, in particular, the proportion of permanent carriers is only between 30% and 45%. Moreover, the carriers used in low season are not necessarily the same as in high season: indeed, on average, 43% to 53% of the carriers are used in low season, whereas this number ranges from 57% to 65% in high season, still significantly less than the total number of carriers selected overall. All these observations point again to the fact that flexibility in the choice of carriers is effectively exploited in the optimization process.

11. Conclusions

In this article, we have proposed a mathematical formulation of a multi-period distribution network design problem with

minimum-purchase commitment contracts. The formulation is challenging for state-of-the-art algorithms when the size of instances increases, mainly when a long horizon is considered and when the shipper considers many carriers. To produce satisfactory solutions for large instances, we developed two algorithms, namely, CBA and RL-RP, which decompose the mathematical model to alleviate the difficulty created by the existence of contracts. These algorithms were tested on classes of instances of various sizes and compared against CPLEX branch-and-cut and Benders algorithms on two main performance criteria.

The CBA algorithm adds combinatorial cuts while solving a relaxed version of the model by branch-and-cut. Compared with CPLEX methods, this exact algorithm underperforms in terms of resolution time for small instances. However, for larger instances, CBA produces more satisfactory results overall, as it finds better solutions than both CPLEX procedures in less computational time. This confirms the usefulness of the CBA algorithmic approach, which could possibly be further enhanced with additional strategies to improve its weak lower bounds. It also underlines the difficulties faced by CPLEX methods to fully exploit the decomposable structure of the multi-period problem. This suggests that variants of our combinatorial Benders approach might tackle problems with similar structures more effectively than the classical methods implemented in CPLEX.

On the other hand, for medium to large-size instances, the ad-hoc heuristic RL-RP can produce better solutions than exact methods (usually, within 1% of optimality), in less time. This demonstrates the relevance of the decomposition, which produces tight lower bounds on the optimal value, and the effectiveness of the simple repairing mechanism for solutions that violate contract-term constraints. Even better solutions might be obtained by including local search in the repair phase.

From a managerial point of view, our study shows that despite its conceptual and computational complexity, the tactical MDPC problem can be efficiently solved to produce near-optimal contract portfolios and distribution plans. This provides managers with the opportunity to move away from the alternative of only relying on the spot market or on long-term binding contracts. The numerical results suggest that as long as the contracts are optimally selected, the existence of minimum purchase requirements does not necessarily increase the total cost in a significant way.

As mentioned in Section 2, the paper does not cover the design of contracts, in the sense that the parameters of the contracts (duration, minimum purchase commitment) are assumed to be fixed ex ante. However, our models and algorithms could be used to simulate the implications of different contract specifications, e.g., in the context of negotiations with the carriers. They can also be extended to examine different types of contracts with risk-sharing mechanisms beyond the purchase commitment expressed in mon-

etary terms that we have introduced in this paper. For example, quantity flexibility contracts can be used to reserve some resource capacity (e.g., a number of truckloads), while allowing additional capacity to be charged at a higher cost. Such flexible options become particularly meaningful when the environment presents significant uncertainty with regard to demand or cost parameters. Stochastic models arising in this context are investigated in Clavijo López (2021). Finally, it may be interesting to investigate the existence and the effect of short-term contracts in different settings, e.g., for the procurement of warehousing capacity. The decomposition approach used in the relax-and-repair heuristic would be a natural candidate for the solution of the associated planning problems.

Acknowledgments

We thank several anonymous reviewers for their comments and suggestions that helped us improve the article. This research was supported by the project PRESupply funded by the Walloon Region of Belgium (Convention #7566). Computational resources were provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under Grant No. 2.5020.11.

References

- Akbalik, A., Hadj-Alouane, A., Sauer, N., & Ghribi, H. (2017). NP-hard and polynomial cases for the single-item lot-sizing problem with batch ordering under capacity reservation contract. *European Journal of Operational Research*, 257(2), 483–493.
- Balas, E., & Jeroslow, R. (1972). Canonical cuts on the unit hypercube. *SIAM Journal on Applied Mathematics*, 23(1), 61–69.
- Bassok, Y., & Anupindi, R. (1997). Analysis of supply contracts with total minimum commitment. *IIE Transactions*, 29, 373–381.
- Ben Mohamed, I., Klibi, W., & Vanderbeck, F. (2020). Designing a two-echelon distribution network under demand uncertainty. *European Journal of Operational Research*, 280(1), 102–123.
- Benders, J. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4, 238–252.
- Botton, Q., Fortz, B., Gouveia, L., & Poss, M. (2013). Benders decomposition for the hop-constrained survivable network design problem. *INFORMS Journal on Computing*, 25(1), 13–26.
- Brusset, X. (2009). Choosing a transport contract over multiple periods. *International Journal of Logistics Systems and Management*, 5(2-3), 273–322.
- Brusset, X. (2010). *Modeling contractual relationships in transport*. Belgium: Louvain School of Management, Université Catholique de Louvain Ph.D. thesis. <http://hdl.handle.net/2078.1/32560>
- Cachon, G., & Lariviere, M. (2001). Contracting to assure supply: How to share demand forecasts in a supply chain. *Management Science*, 47(5), 629–646.
- Caplice, C., & Sheffi, Y. (2003). Optimization-based procurement of transportation services. *Journal of Business Logistics*, 24(2), 109–128.
- Chen, F., Hum, S., & Sun, J. (2001). Analysis of third-party warehousing contracts with commitments. *European Journal of Operational Research*, 131(3), 603–610.
- Chen, F., & Krass, D. (2001). Analysis of supply contracts with minimum total order quantity commitments and non-stationary demands. *European Journal of Operational Research*, 131(2), 309–323.
- Clavijo López, C. (2021). *Multi-period distribution network problems with minimum commitment contracts*. Belgium: HEC - Management School, University of Liège Ph.D. thesis. <http://hdl.handle.net/2268/263230>
- Codato, G., & Fischetti, M. (2006). Combinatorial Benders' cuts for mixed-integer linear programming. *Operations Research*, 54(4), 756–766.
- Corbett, C., Zhou, D., & Tang, C. (2004). Designing supply contracts: Contract type and information asymmetry. *Management Science*, 50(4), 425–559.
- Cortinhal, M., Lopes, M., & Melo, M. (2015). Dynamic design and re-design of multi-echelon, multi-product logistics networks with outsourcing opportunities: A computational study. *Computers & Industrial Engineering*, 90, 118–131.
- Cui, L., & Hertz, S. (2011). Networks and capabilities as characteristics of logistics firms. *Industrial Marketing Management*, 40(6), 1004–1011.
- de Albéniz, V. M., & Simchi-Levi, D. (2005). A portfolio approach to procurement contracts. *Production and Operations Management*, 14(1), 90–114.
- Dias, J., Captivo, M., & Clímacos, J. (2008). Dynamic multi-level capacitated and uncapacitated location problems: An approach using primal-dual heuristics. *Operational Research. An International Journal*, 7(3), 345–379.
- Dolan, E. D., & Moré, J. J. (2002). Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91(2), 201–213.
- Drezner, Z. (1995). Dynamic facility location: The progressive p-median problem. *Location Science*, 3(1), 1–7.
- Drezner, Z., & Wesolowsky, G. O. (1991). Facility location when demand is time dependent. *Naval Research Logistics*, 38, 763–777.
- Fakhri, A., Ghatee, M., Fragkogios, A., & Saharidis, G. (2017). Benders decomposition with integer subproblem. *Expert Systems with Applications*, 89, 20–30.
- Gendron, B., Scutellà, M., Garroppo, R., Nencioni, G., & Tavanti, L. (2016). A branch-and-Benders-cut method for nonlinear power design in green wireless local area networks. *European Journal of Operational Research*, 255(1), 151–162.
- Gendron, B., & Semet, F. (2009). Formulations and relaxations for a multi-echelon capacitated location–distribution problem. *Computers and Operations Research*, 36(5), 1335–1355.
- Ghiani, G., Laporte, G., & Musmanno, R. (2004). *Introduction to logistics systems planning and control*. Wiley.
- Gu, W., Archetti, C., Cattaruzza, D., Ogier, M., Semet, F., & Speranza, M. G. (2022). A sequential approach for a multi-commodity two-echelon distribution problem. *Computers & Industrial Engineering*, 163, 107793.
- Guastaroba, G., Speranza, M. G., & Vigo, D. (2016). Intermediate facilities in freight transportation planning: A survey. *Transportation Science*, 50(3), 763–789.
- Hanbazazah, A. S., Abril, L., Erkoc, M., & Shaikh, N. (2019). Freight consolidation with divisible shipments, delivery time windows, and piecewise transportation costs. *European Journal of Operational Research*, 276(1), 187–201.
- Hanbazazah, A. S., Abril, L. E., Shaikh, N. I., & Erkoc, M. (2018). A redesigned Benders decomposition approach for large-scale in-transit freight consolidation operations. *International Journal of Information Systems and Supply Chain Management*, 11(2), 1–15.
- Heydari, J., Govindan, K., Ebrahimi, H., & Taleizadeh, A. (2020). Coordination by quantity flexibility contract in a two echelon supply chain system: Effect of outsourcing decisions. *International Journal of Production Economics*, 225, 107586.
- Hooker, J., & Ottosson, G. (2003). Logic-based Benders decomposition. *Mathematical Programming, Series A*, 96(January), 33–60.
- Jena, S., Cordeau, J., & Gendron, B. (2015). Dynamic facility location with generalized modular capacities. *Transportation Science*, 49(3), 484–499.
- Jin, M., & Wu, S. D. (2007). Capacity reservation contracts for high-tech industry. *European Journal of Operational Research*, 176(3), 1659–1677.
- Klose, A., & Drexl, A. (2005). Facility location models for distribution system design. *European Journal of Operational Research*, 162(1), 4–29.
- Kuyzu, G., Akyol, Ç., Ergun, Ö., & Savelsbergh, M. (2015). Bid price optimization for truckload carriers in simultaneous transportation procurement auctions. *Transportation Research Part B*, 73, 34–58.
- Lai, K. (2004). Service capability and performance of logistics service providers. *Transportation Research Part E: Logistics and Transportation Review*, 40(5), 385–399.
- Lariviere, M. (1999). In S. Tayur, R. Ganeshan, & M. Magazine (Eds.), *Supply chain contracting and coordination with stochastic demand*. Boston, Massachusetts: Kluwer Academic Publisher.
- Li, J., Luo, X., Wang, Q., & Zhou, W. (2020). Supply chain coordination through capacity reservation contract and quantity flexibility contract. *Omega*, 99, 305–483.
- Lian, Z., & Deshmukh, A. (2009). Analysis of supply contracts with quantity flexibility. *European Journal of Operational Research*, 196(2), 526–533.
- Lim, A., Wang, F., & Xu, Z. (2006). A transportation problem with minimum quantity commitments. *Transportation Science*, 40(1), 117–129.
- Lin, C.-C., & Wu, Y.-C. (2001). The multiple frequency delivery operations in the time-definite delivery industry. *Journal of the Operational Research Society*, 52(11), 1215–1224.
- Lindsey, C., & Mahmassani, H. (2017). Sourcing truckload capacity in the transportation spot market: A framework for third party providers. *Transportation Research Part A*, 102, 261–273.
- Melo, M., Nickel, S., & da Gama, F. S. (2005). Dynamic multi-commodity capacitated facility location: A mathematical modelling framework for strategic supply chain planning. *Computers and Operations Research*, 33, 181–208.
- Nosohi, I., & Nookabadi, A.-S. (2016). Outsource planning through option contracts with demand and cost uncertainty. *European Journal of Operational Research*, 250, 131–142.
- Owen, S., & Daskin, M. (1998). Strategic facility location: A review. *European Journal of Operational Research*, 111, 423–447.
- Patel, S., & Swartz, C. (2019). Supply chain design with time-limited transportation contracts. *Computers and Chemical Engineering*, 131, 106579.
- Pimentel, B. S., Mateus, G. R., & Almeida, F. A. (2013). Stochastic capacity planning and dynamic network design. *International Journal of Production Economics*, 145(1), 139–149.
- Rahmaniani, R., Crainic, T. G., Gendreau, M., & Rei, W. (2017). The Benders decomposition algorithm: A literature review. *European Journal of Operational Research*, 259(3), 801–817.
- Sandholm, T., & Shields, R. (2006). Nogood learning for mixed integer programming. *Technical Report CMU-CS-06-155*. Carnegie Mellon University, Computer Science Department.
- Song, J., & Regan, A. (2005). Approximation algorithms for the bid construction problem in combinatorial auctions for the procurement of freight transportation contracts. *Transportation Research Part B Methodological*, 39(10), 914–933.
- Spinler, S., & Huchzermeier, A. (2006). The valuation of options on capacity cost and demand uncertainty. *European Journal of Operational Research*, 171(3), 915–934.
- Stefansson, G., & Russell, D. (2008). Supply chain interfaces: Defining attributes and attribute values for collaborative logistics management. *Journal of Business Logistics*, 29(1), 347–359.
- Tsay, A., & Lovejoy, W. (1999). Quantity flexibility contracts and supply chain performance. *Manufacturing & Service Operations Management*, 1(2), 89–111.
- Van Roy, T., & Erlenkotter, D. (1982). A dual-based procedure for dynamic facility location. *Management Science*, 28(10), 1091–1105.
- Wesolowsky, G., & Truscott, W. G. (1976). The multiperiod location-allocation problem with relocation of facilities. *Management Science*, 22(1), 57–65.