
Multi-period Distribution Network Problems with Minimum Commitment Contracts

A dissertation submitted to the faculty of HEC Liège
in fulfilment of the requirement for the degree
of PhD in Economics and Management
by

Christian CLAVIJO LÓPEZ

Supervisors:

Yves Crama - Université de Liège

Thierry Pironet - Université de Liège

Committee member:

Frédéric Semet - École Centrale de Lille

Jury members:

An Caris - Hasselt University

Sabine Limbourg - Université de Liège

Célia Paquay - Université de Liège

Summary

Companies outsourcing road freight transportation services approach their procurement plan based on a combination of long-term annual contracts with dedicated carriers and spot-markets for immediate loads. This thesis proposes an integrated planning framework considering tactical and operational decisions for logistics services procurement. It allows individual shippers to formulate mid-term contractual agreements with several transporters that could reduce some of the risks encountered in shipper-carrier business relationships, such as unavailability of transportation resources and high-market prices.

Contractual agreements are based on minimum commitment (MC) contracts, where the shipper commits to a minimum business volume with certain regularity while carriers offer low and stable rates regardless of market variations. The management problem includes selecting among several contract proposals and deciding the distribution planning for every period along the time horizon.

Real-life applications may arise in the context of commercial companies or logistics firms with presence in widespread areas and subcontracting to multiple transportation and logistics service providers. The models are meaningful in dynamic environments, where the contracts represent both an opportunity and risk due to high demand and price variability. In the present work, the tactical planning framework with minimum commitment contracts is studied in the field of e-commerce logistics distribution from the perspective of an e-fulfilment company.

The mathematical problem is formulated with mixed-integer linear programs for several variants of MC contracts, in their deterministic and stochastic form. Several algorithmic methods are compared to prove efficiency in finding optimal or applicable solutions to the problem. Numerical tests were performed in random instances with well-defined parameter settings. Results are presented with performance profiles evaluating algorithms in terms of computational time and solution

quality. The analysis considers the economical advantages of approaching capacity planning from a tactical point of view.

Acknowledgements

The present thesis has been the product of a collaborative work with my supervisors, professors Yves Crama and Thierry Pironet at HEC Liège. I'm grateful for the opportunity that they gave me to embark myself upon this scientific journey and their influence played along the way that allowed me to accomplish this milestone in my career and my life. Your guidance and support during this years have been very important for my professional and individual growth. I particularly enjoyed our discussions and meetings where it was possible to express and share ideas and opinions on different aspects of my work. I appreciate your constant presence and responsiveness throughout this process and the flexibility given to conduct my research upon my personal interests.

I also want to say thanks to Professor Frédéric Semet, for being part of the thesis committee. His thoughtful insights and observations provided the best directions to improve and consolidate my advances. Also, professor An Caris for his acceptance as part of the jury. I thank the faculty members in the QuantOM research department who were part of my doctoral formation and thesis dissertation. My colleagues, who serve as an example and encouragement to finish my PhD. Not forgetting the people and friends that I had the chance to find in Belgium, and have made my stay more fun and interesting.

I dedicate this thesis to my beloved ones, my family who even though at far distance were always close to me and supportive. My parents who trust in me and my capabilities to follow and chase my aspirations. My sister, who never fails to restore my confidence and always had words of affection and support for me. My brother, who is also a great friend and with whom I always share joyful moments, specially in the last years. Special mention to my grandparents who will always rest in my heart.

Contents

Summary	i
Acknowledgements	iii
List of Abbreviations	ix
1 Introduction	1
1.1 Characterization of Logistics Service Providers	5
1.2 Road Freight Transportation	8
1.2.1 Long-Term Agreements	8
1.2.2 Short-Term Agreements	9
1.3 Issues Relative to Contracts in Transportation	10
1.4 Risk-sharing Contracts	12
1.5 Contextual Applications	13
1.6 Thesis Contents	14
2 Literature Review	17
2.1 Supply Contracts with Risk-sharing Provisions	18
2.1.1 Representation of Risk-sharing Provisions	21
2.1.2 Supply Contracts in the Literature	22
2.2 Supplier Selection in Procurement Planning	25
2.3 Network Optimization and Logistics Outsourcing Problems	27
2.4 Multi-Period Facility Location Problems	29
2.5 E-commerce Logistics Distribution	33
3 Distribution Network Problems with Minimum Purchase Commitment Contracts	35
3.1 Introduction	35
3.2 Two-Echelon Distribution System	37
3.3 Minimum Commitment Contracts	40
3.3.1 Duration (H)	41

3.3.2	Service and Penalty Costs (F, P)	42
3.3.3	Minimum Commitment (M)	43
3.4	The Minimum Purchase Commitment (MPC) Contract	44
3.5	Single-period Distribution Network Problem MPC($H = 1, P = 0, M = M_e$)	46
3.6	Multi-period Distribution Network Problem MPC($H > 1, P = 0, M = M_e^t$)	48
3.7	Multi-period Distribution Network Problem with Penalty Costs - MPC($H > 1, P > 0, M = M_e^t$)	51
3.8	Multi-period Distribution Network Problem with Penalty Costs - MPC($H > 1, P > 0, M = M_e^t$) and Off-Contract Services - ($H = 1, P > 0, M = 0$)	54
3.9	Multi-period Distribution Network Problem with Total Minimum Purchase Commitment MPC($H > 1, P = 0, M = M_e^H$)	56
3.10	Multi-period Distribution Network Problem with Total Minimum Purchase Commitment and Penalty Costs - MPC ($H > 1, P > 0, M = M_e^H$)	58
3.11	Conclusions	60

4 Multi-Period Distribution Network Problem with Minimum Purchase Commitment (MPC) Contracts 61

4.1	Introduction	61
4.2	Mathematical formulation	63
4.2.1	Notations	63
4.2.2	Mixed-integer programming formulation	65
4.3	Complexity	67
4.3.1	NP-hardness	67
4.3.2	Decomposability	68
4.4	Combinatorial Benders Algorithm - (CBA)	68
4.4.1	Benders with integer sub-problems	68
4.4.2	Decomposition	70
4.4.3	Algorithmic procedure	73
4.5	Relax-and-repair heuristic	75
4.5.1	Phase 1 - Relaxation - (H1)	76
4.5.2	Phase 2 - Repairing - (H2)	77
4.6	Experimental design and performance measures	79
4.6.1	Instances	79
4.6.2	Algorithms	81
4.6.3	Performance measures	82

4.7	Computational results	83
4.7.1	Small size instances	83
4.7.2	Medium size instances	88
4.7.3	Large size instances	91
4.7.4	Variations of contract duration	95
4.8	Conclusions	97
5	Distribution Network Problems with Stochastic Demand and Minimum Capacity Commitment Contracts	99
5.1	Introduction	99
5.2	The Minimum Capacity Commitment Contract - MCC	101
5.3	Stochastic Representation of the Demand	103
5.4	Two-stage Stochastic Model with Fixed Recourse	104
5.5	Stochastic Models with Capacity Request under Demand Certainty	109
5.5.1	First stage	109
5.5.2	Second stage	110
5.5.3	Single-period Distribution Network Problem: MCC ($H = 1, P = 0, M = M_e$)	111
5.5.4	Multi-period Distribution Network Problem: MCC ($H > 1, P = 0, M = M_e^t$)	114
5.5.5	Multi-period Distribution Network Problem with Penalty Costs: MCC ($H > 1, P > 0, M = M_e^t$)	115
5.6	Stochastic Models with Capacity Request under Demand Uncertainty	118
5.6.1	First stage	119
5.6.2	Second stage	120
5.6.3	Single-period Distribution Network Problem with Capacity Reservation: MCC($H = 1, P > 0, M = 0, R$)	121
5.6.4	Single-period Distribution Network Problem with Capacity Reservation: MCC($H = 1, P > 0, M = M_e, R$)	124
5.6.5	Multi-period Distribution Network Problem with Capacity Reservation: MCC($H > 1, P > 0, M = M_e, R$)	126
5.7	Conclusions	127
6	Multi-Period Distribution Network Problem with Stochastic Demand and Minimum Capacity Commitment (MCC) Contracts	129
6.1	Introduction	129
6.2	Stochastic Problem	131

6.2.1	First Stage	131
6.2.2	Second Stage	132
6.3	Mathematical Modelling for the Multi-stage Stochastic <i>mDNP_MCC</i> Model	132
6.4	Algorithmic Development	136
6.4.1	Rolling Horizon Procedure - (RHP)	138
6.4.2	Contract Length Relaxation - (CLR)	140
6.4.3	Off-contract Capacity Variables ($w_{i,l}^{f,s}$) Relaxation - (WR)	142
6.4.4	Contract Length and $w_{i,l}^{f,s}$ Variables Relaxation - (CLWR)	143
6.5	Experimental Design	144
6.5.1	Algorithms	145
6.5.2	Performance measures	147
6.5.3	Instance Parameters	148
6.6	Computational Results	150
6.6.1	Lower bounds	150
6.6.2	Algorithmic Performance	151
6.6.3	Algorithmic Results	156
6.6.4	Value of the stochastic solution	157
6.7	Managerial Insights	159
6.7.1	Value of mid-term MCC contracts	159
6.7.2	Structure of solutions	161
6.8	Conclusions	164
7	Conclusions, Limitations and Perspectives	167
7.1	Conclusions	167
7.2	Limitations	169
7.3	Perspectives	170
	Bibliography	173
	List of Tables	183
	List of Figures	185

List of Abbreviations

MC	Minimum Commitment
MPC	Minimum Purchase Commitment
MCC	Minimum Capacity Commitment
DNPMC	Distribution Network Problems with Minimum Commitment Contracts
MDPC	Multi-period Distribution Network Problem with Minimum Purchase Commitments
<i>mDNP</i>	Multi-period Distribution Network Problem
<i>mDNP_MCC</i>	Multi-period Distribution Network Problem with Minimum Capacity Commitments
MPFLP	Multi-period Facility Location Problem
3PL	Third Party Logistics
FTL	Full Truckload
CD	Cross Docking
TSP	Transportation Service Procurement
LT	Long-Term
T&D	Transportation and Distribution
LB	Lower Bound
CBA	Combinatorial Benders Algorithm
CLR	Contract Length Relaxation
CLWR	Contract Length and W-variables Relaxation
WR	W-variables Relaxation
MP	Master Problem
SP	Sub Problem

Chapter 1

Introduction

The sector of freight transportation and logistics services is of vital importance for the global economy. Several means of transport are responsible for the continuous flow of materials and products traded globally. In spite of strict environmental policies and regulations controlling the effects of fuel consumption and gas emissions, the freight transportation industry maintains a steady growth over the past years at an average annual rate of 1.5% in Europe, moving nearly 9.2 billion tonnes-km per day (within the EU zone) [Eurostat, 2020]. At the same time innovations in information technology and web-based tools lead logistics companies to take advantage of digitalization in order to improve their way of doing business, enhance operational performance, and provide higher quality of service. For instance, carriers make use of on-line platforms, increasing in that way visibility and allowing to display resources available in specific lanes, thereby, reaching a larger base of potential customers interested in transportation services.

In recent years, the intensified use of e-commerce channels is regarded as a determining factor in the economic growth of the freight transportation sector, with businesses and customers buying products on-line from a vast selection of sellers worldwide. Likewise, numerous retailers and manufacturers rely increasingly on e-markets as an alternative option of selling products at local and international markets. As some of the entry barriers to the use of e-channels such as high fees or long lead times are lowered, more individuals and businesses find e-commerce opportunities attractive to acquire supplies, goods and services at acceptable costs. Meanwhile, the demand for long-distance transportation and last-mile distribution keeps expanding. Leading players in the e-commerce arena (e.g. Amazon, Alibaba) continue to invest in physical assets, adding warehousing capacity and deploying

fulfilment services around the globe and at the same time, working together with logistics service providers (LSPs) in order to satisfy end-customer requests in terms of fast shipping and low-cost delivery.

The effect of e-commerce motivates commercial companies to seek out for specialists in product distribution and order fulfilment. Outsourcing logistics services is a common practice. Nowadays, most of the companies outsource domestic (74%) and international (62%) transportation as well as warehousing services (63%) [Infosys, 2020]. Some of the benefits mentioned in the academic literature are economical savings, efficient use of capital and resources, increased focus on main competences and capabilities as well as quick introduction of products and service innovations into the market. In addition to the management and operations of basic functions (transportation, storage), logistics service providers can offer a wide range of complementary solutions, including information-related and value-added services [Lai, 2004]. However, the extensive offer of outsourcing alternatives for freight transportation makes it challenging to select suitable logistics partners and to manage coordination among them, in some cases by different transport modes. Therefore, LSPs often make decisions involving the movements and handling of products on behalf of the shipper. A distinction of LSPs according to their core capabilities is presented in Section 1.1.

Companies outsourcing road freight transportation services approach their procurement plan based on a combination of long-term annual contracts (at least 1 year) with dedicated carriers and through spot marketplaces for immediate services. In both cases, the carrier selection criterion is primarily based on cost and taking into account expected variations and uncertain outcomes of transportation requirements in the future is not often analysed in contractual arrangements. A description of common practices and theoretical frameworks in road freight transportation procurement is described in Section 1.2. In practice, uncertain events can increase considerably logistics costs. Furthermore, undefined or unclear specifications in long-term contracts with respect to the shipper's capacity requirements can cause disruptions in the supply of transportation services. This is a major concern if we consider that a delayed or unfulfilled order can result in the loss of clients and undermine the shipper's reliability. Some of the problems encountered in the literature of transportation management regarding long-term business relationships between shippers and carriers are introduced in Section 1.3

An opportunity arises to consider transportation procurement methods from a tactical perspective (i.e., within a foreseeable horizon of several weeks but no longer than one year) taking into account deterministic and probabilistic scenarios of the

shippers' requirements in the near future. This thesis proposes a tactical planning approach for freight transportation management allowing to take preventive contractual decisions between a shipper and several carriers, for instance, when demand is likely to change drastically and alternative sources may be costly or insufficient. While planning methods aimed at the medium-to-long term are regularly studied and applied in production and manufacturing, less attention is given to procurement of logistics and transportation services. In that regard, this work attempts to provide a different perspective on the subject separate from standard practices. As observed by Huang & Xu [2013], transportation service procurement (TSP) mechanisms are worth investigating as they could affect the trading volume and efficiency of the freight market.

Given the mid-range planning horizon, our framework conceives contractual agreements between a shipper and carriers of shorter duration than commonly used annual contracts. The shortened length provides greater flexibility in adjusting transportation requirements to contracted resources. The shipper can dismiss more frequently certain contracts that are not being profitable, while at the same time diversify its choice of dedicated carriers. In addition, the use of risk-sharing contracts with minimum commitment conditions into the business agreement can be beneficial for both parties by clarifying norms about capacity (vehicle) availability, fleet utilization and shipping rates. Such conditions and terms may deal in part with the shipper's risks of unavailability of transportation resources and/or high market prices. On the other hand, carriers may enforce a minimum level of capacity utilization (e.g., volume, truck fleet) which guarantees a steady return on their assets independent of the market variability. A brief characterization of risk-sharing mechanisms is presented in Section 1.4 and further detailed in Section 2.1.

In the present work, the tactical planning framework with minimum commitment contracts is applied in the context of e-commerce logistics distribution. The shipper is in practice an e-commerce third-party logistics (3PL) company subcontracting with LSPs to support its distribution and delivery system. The selection of carriers is based on contract prices on their operational lanes subject to satisfying minimum commitment conditions. The shipper presents two main characteristics: it owns and manages central warehouses (so called e-fulfilment centres) where multiple items from multiple sellers are stored. In addition, it experiences large volumes of parcels bought daily, that need to be delivered to buyers as soon as possible. Therefore, full truck capacity provided by external carriers is best suited to ship parcels towards customer areas. Before arrival, parcels are forwarded through consolidation centres (i.e., cross-docking facilities where entering goods get sorted and placed into outbounds vehicles) close to customer areas and delivered by authorized postal

companies or last-mile delivery services to their destination points. The distribution system so described represents a two-echelon supply chain network.

The distribution network design problem is similar to a multi-period facility location problem. The mixed-integer linear mathematical programs developed aim at minimizing the total costs along a finite time horizon with time-varying demand. The mathematical formulation integrates managerial decisions at two hierarchical levels. At the tactical level, the shipper determines which carriers are able to provide services, therefore what resources (truck capacity, facilities) can be used to ship and redirect parcels upon arrival. At the operational level, the distribution plan dictates which of the available carriers' resources are functional in a certain period. Both, contract and distribution plans are interdependent in the decision-making process. A last purpose of this thesis is to derive algorithmic strategies to tackle the complex structure of the problems based on various decomposition procedures. The structure of the proposed mathematical models entails difficulties for solving large instances of the problem. The solution strategies are built on non-exact and exact methods applied previously for other optimization problems. The uncertainty of the customers' demand is also considered through several stochastic programming formulations.

The contributions of the thesis are summarized in Figure 1.1

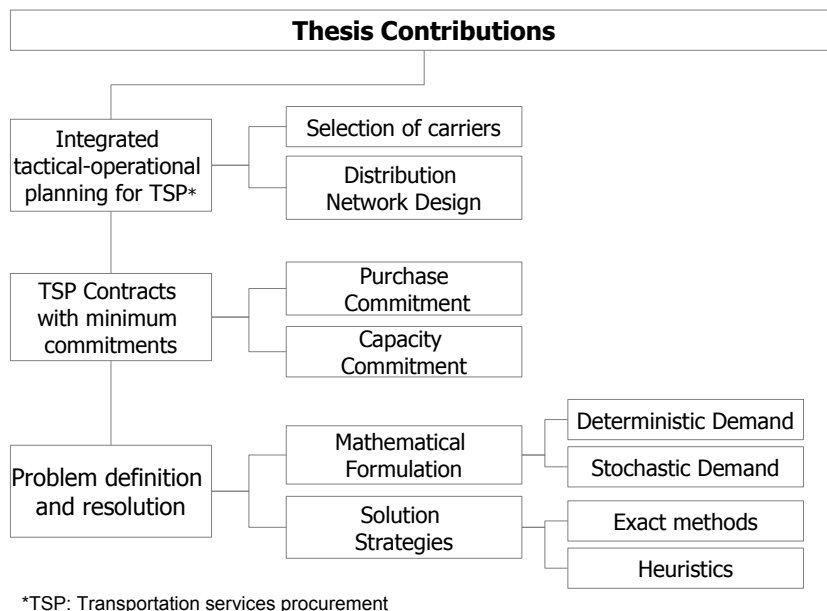


Figure 1.1: Thesis contributions

1.1 Characterization of Logistics Service Providers

The range of LSPs is well diversified as some operate and manage physical assets (transportation fleet, distribution centres, warehousing, equipment) while others act as intermediaries, partially or totally subcontracting transportation requests to the former type. The differentiation of logistics firms can be expressed according to their service capabilities and core competences which lead them to compete in different market segments [Lai, 2004]. Hence, distinct actors in the logistics service supply chain coexist and are interdependent with each other [Cui & Hertz, 2011]. A characterization of LSPs is valuable to the thesis subject since it allows us to define the nature of actors interested in approaching outsourcing of logistics services with the purpose of finding cost-optimal solutions to distribution logistics in e-commerce. Additionally, it gives a broader view of the applicability of the methods and models studied in this thesis.

Logistics service providers can be classified in three categories: Carriers, Logistics Service Intermediaries (LSI), and Third-Party Logistics companies (3PL) [Stefansson & Russell, 2008].

Carriers or transporters are agents that own and operate transportation assets [Mesa-Arango & Ukkusuri, 2014]. In the context of road freight logistics, carriers are for example trucking companies that conduct and operate resources (e.g. vehicles, terminals, drivers) within a limited geographical coverage, from narrow to wide areas [Cui & Hertz, 2011]. Their strategic capabilities and core competences stem from the exploitation of resources to carry out operations in their dedicated networks [Lai, 2004]. Carriers can vary in terms of size, service offerings and target customers. More commonly, they offer standard transportation services for regular use, although some may specialize in certain type of solutions investing in specific-purpose vehicles. Carriers often have preference for setting up long-term (LT) contracts with shippers by establishing fixed prices beforehand. LT contracts account for more than 90% percent of business of asset-based carriers [Scott, 2018]. In that sense, their fleet is mainly to cover their registered clients' needs. Some carriers even create strong alliances committing a large proportion of their resources to dedicated clients. Regardless of LT agreements, carriers seek alternative ways to cash in using to the maximum their resource capacity. For example by tendering to one-time jobs (or so-called loads) placed by unknown shippers in on-line auctions or placing offers for potential jobs in free slot times and convenient routes, according to their respective schedules, sometimes through logistics intermediaries who can connect easily with numerous shippers in need of transportation services.

Third-party logistics companies (3PLs) supply transportation services using either their own physical assets or those from others [Lindsey et al., 2013]. A distinction is made from those that rely completely on external resources (non-asset based). In spite of owning several transportation assets, some 3PLs subcontract with carriers when additional capacity is required [Mesa-Arango & Ukkusuri, 2014]. Regardless of the differences in ownership, 3PLs have in common to coordinate and synchronize different actors in the clients' supply chain. An important feature of 3PLs is the provision of value-added services beyond logistics functions, such as taking care of administrative paperwork, freight safety protocols, customs clearance among other activities. For doing so, 3PL companies invest in warehousing, and support integration of IT services with partners for smoother transition of physical flows and information. Furthermore, 3PLs may invest capital in customized assets that address specific transportation requirements of dedicated clients. The relationships 3PL-client can be considered as partnerships which are strongly developed with time and have a long-term view. Their expertise, knowledge and specialized skills are the primary intangibles available to shippers [Lindsey & Mahmassani, 2017]. Asset-based 3PLs have in common with carriers the use of spot markets as a mean to promote leftover capacity, either directly or through intermediaries (LSI). Therefore, 3PL providers can either supply capacity in the long-term or in spot markets [Lindsey et al., 2013]. Some 3PL companies defined themselves as 4PL organizations or supply chain integrators that are responsible to manage several third-party service providers of the same user [Buyukozkan et al., 2009]. Hence, 4PL, are examples of companies that manage logistics operations for their clients but rely also on external providers for physical flows.

In contrast, *logistics service intermediaries* (LSIs) are non-asset based companies specialized in linking carriers and shippers. Therefore, LSIs do not own any freight assets [Lindsey & Mahmassani, 2017], but handle a vast network of carriers, 3PLs and shippers. Their main goal is the consolidation of orders [Cui & Hertz, 2011], i.e., the allocation of loads or shipments to available capacity from multiple sources. There are several types of intermediaries in the market: brokers, forwarders or consolidators. [Coyle et al., 2010]. Brokers specialize in spot markets while forwarders concentrate on consolidation of shipments. Handling large number of clients reduces LSIs capabilities to manage physical flows in specific supply chains [Cui & Hertz, 2011]. The type of business relationship with clients can be considered as a low-level partnership, in which coordination and planning are key activities but do not necessarily involve integration and/or capital investment [Knemeyer et al., 2003]. Therefore, their main assets and capabilities come from making use of decision support systems and IT technologies as well as a comprehensive knowledge of procurement tactics in the field. The ability in processing, managing and transmit-

ting real-time data is key for their strategic competitiveness.

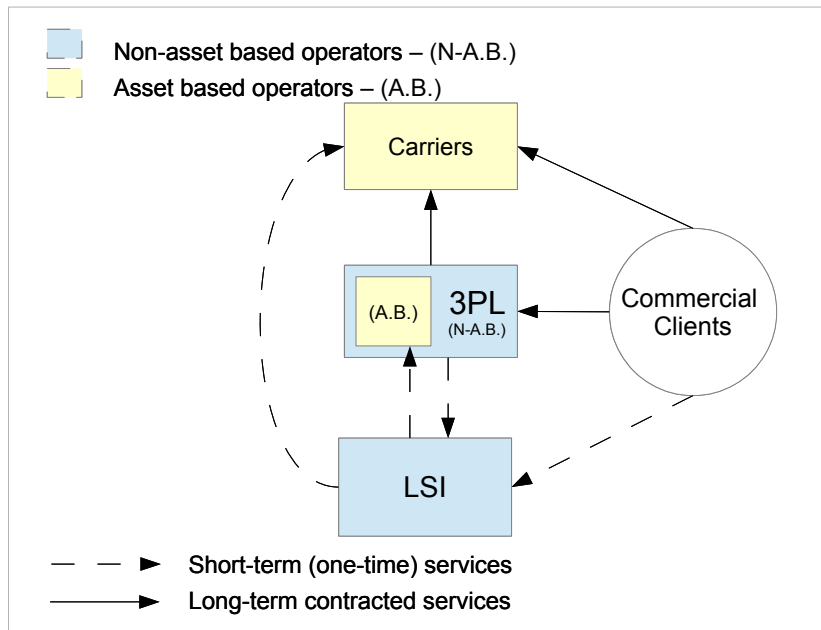


Figure 1.2: Relationships between logistics service providers

In Figure 1.2 are represented the different types of logistics service providers described above. Differentiation is made primarily between asset and non-asset based actors where 3PLs can range from having none to several physical assets. The interconnected links follows the direction from the requesting entity to the service supplier. Links are also marked differently to show how business relationships are usually established among players, either by long-term or short-term (one-time) agreements.

In addition, collaborative alliances in each market segment are increasing [Crujssen et al., 2007]. Figure 1.2 intends to simplify the nature of actors and business relationships although in reality the range of logistics firms can be more complex and further differences are presented in each market segment, for example regarding the main logistics functions provided.

1.2 Road Freight Transportation

Road transportation remains the most widely used mean of freight transportation worldwide. It's estimated that around 50% of goods are transported by road in the European Union (EU) zone with more than 500 thousands enterprises performing transportation and storage activities [Eurostat, 2020]. The range of trucking services is separated in full truckload (FTL) and less-than-truckload (LTL) according to capacity requirements of the shipper either for complete or partial vehicle use, respectively. Full truckload (FTL) services (i.e., full utilization) is more adequate for high-volume long-haul shipments, while LTL is intended for low-volume within short as well as long distances. A FTL shipment is also characterized by following a direct origin-destination path, while in LTL the shipment might be part of several routes and stops before destination. The trucking market is predominantly used for full truckload services, representing more than 78% of the total market [Caplice & Sheffi, 2003]. In this thesis, FTL transportation is considered within applications in the context of e-commerce distribution. For the concern of the study, a review of long-term and short-term (or one-time) contract practices between carrier and shippers is described next.

1.2.1 Long-Term Agreements

The most common formal commercial relationship between shippers and common carriers is based on contractual agreements, that usually hold for one to two years and sometimes longer [Caplice & Sheffi, 2003]. In that sense, the process of selecting logistics partners can be categorised at the strategic level and as one of the key success factors for a firm [Rekik & Mellouli, 2012]. Several mechanisms are used to find trading partners and formalize annual agreements, such as catalogues, negotiations and auctions as referred in Lafkihi et al. [2019]. Most regularly these mechanisms are available through on-line platforms (e-marketplaces) or transport management systems, which integrate strategic procurement tools between shippers and carriers. In catalogues, services and prices are relatively fixed although subject to negotiations [Lafkihi et al., 2019].

Among the most studied methods in the logistics business literature are auction mechanisms. Generally, auctions are conducted by shippers (hence, in reverse-form) for a determined lane (i.e., inter-cities). Interested carriers or intermediary agents bid on either one or multiple loads of the same shipper. Thereafter, the negotiation process with the winner determines conditions such as the contract du-

ration. Nonetheless, auctions studied in theory can take different forms, regarding the number of lanes offered, carrier allocation rules (one or several), bidding strategy or number of rounds. Forward auctions managed by carriers are also subject to analysis.

An extended area of research focuses on combinatorial auctions for freight transportation procurement services, particularly for full truckload (FTL) transportation. Studies make use of optimization models/OR techniques to determine auction winners for bundles of trips. The decision is based on intrinsic synergies in FTL transportation when combined loads could potentially reduce the cost of empty trips for a given carrier (economies of scope). Some of these theoretical approaches were implemented and applied successfully in practical contexts [Caplice & Sheffi, 2003]. Other combinatorial auction studies involve non-economic attributes into the decision process such as carriers reputation [Rekik & Mellouli, 2012] or service quality levels [Buer & Pankratz, 2010].

Considering long-term procurement effects, shippers are not only driven by the price of shipments, although it continues to be the main factor of carrier selection (in 60% of the cases) [Mesa-Arango & Ukkusuri, 2014]. Similarly, for 3PL companies the criteria to select carriers is inclined towards cost benefits [Patterson et al., 2010]. However, qualitative attributes are also taken into account like transport mode (rail, road), fleet size, capacity flexibility and experience. Other factors of concern in the selection process are performance attributes of carriers, such as service reliability, damage/security risks and satisfaction rates with previous partners. On the contrary, shipping times across destinations rarely imply a significant advantage in selecting a service provider, since transit times are rather standardized and commonly known by shippers [Patterson et al., 2010].

1.2.2 Short-Term Agreements

Short-term strategies are based on one-time transactions rather than contracts. Companies turn to spot markets where the excess of capacity coincides with load requests (i.e., shipping job). Hence, transactions take place at prices that are settled and governed by the market itself.

Although FTL transportation prices correlate with seasonality effects (e.g., a monthly index is used in forecast analysis) and principal characteristics (travel time, vehicle type), rates of services fluctuate constantly according to market dynamics (i.e., demand and supply of transportation services). Therefore, shippers experience price

uncertainty in the face of spot markets. In spite of its dynamic nature, the spot price of a truckload is typically much higher than the corresponding contract price [Tsai et al., 2011]. However little is known about the determinants of those prices [Scott, 2015].

Although the most commonly used alternative of outsourced transportation for commercial companies are for-hire carriers (i.e., third-party providers) with long-term contracts (in the US and Europe [Caplice, 2007]), the spot market is often utilized. Spot markets are predominantly used by logistics service intermediaries [Scott, 2018]. In particular brokerage services (or 3PL brokers) grant access to electronic marketplaces which are privately managed. The presence of intermediate agents is estimated of main importance to clear the market effectively. Brokers can negotiate better deals based on their market knowledge responding in that way to urgent or last-minute loads.

Capacity in the spot market can be found by contacting carriers, by offering loads on platforms, or by holding on-line auctions. Similar to annual procurement auctions, spot markets let shippers run one-load or multi-load auctions in search of bidders for urgent or on-demand requests. The short response time is likely to reduce the number of participants.

Logistics service providers of different nature (asset/non-asset based) approach spot loads offerings differently based upon individual interests: while intermediaries (LSIs) are capable to bid more frequently at relatively high prices, trucking firms bid less at more affordable prices [Scott, 2018]. The former focus on multiple transactions with better margins, while the latter react to few loads with acceptable margins. On the other hand, the shipper's allocation rule is merely based upon cost reduction, unlike longer term agreements where additional attributes and qualities are considered. One-time loads are often performed within 3 days after the auction determines a winner and the business relationship ends once the job is done [Scott, 2018].

1.3 Issues Relative to Contracts in Transportation

Long-term contracts in transportation are subject to potential breaches regarding conditions agreed upon initially. Two particular detrimental situations for the shipper's interests might occur.

On one hand, some requests for shipments may be refused by the hired carrier.

Even though, contractual prices for trucking services are binding in legal terms, the responsibility to satisfy shipping requirements is usually not strictly imposed [Caplice, 2007]. As a consequence, carriers may turn down shipments if trucks are unavailable or in presence of market opportunities to haul at higher prices in the spot market [Scott et al., 2017]. This could happen due to unplanned or unexpected requirements caused by surging needs of load movements. The declination rate could elevate up to 26% of requests [Caplice, 2007], which leads shippers to find capacity elsewhere, possibly at unfavourable off-contract prices. In some cases, urgent services are not guaranteed to be fulfilled on time [Lindsey & Mahmassani, 2017].

Sourcing capacity from spot market-places can be very expensive. The variation of trucking rates for the same direct route can range from 75% to 200% of the “official” rate [Brusset, 2010]. The variability of costs-per-mile can vary substantially regardless of the distance travelled (more than twice the average costs according to the observed data in [Lindsey et al., 2013]). In addition, contracting with logistics intermediaries allowing access to trade platforms represents extra costs.

On the other hand, the set of prices pre-agreed during initial negotiations may not be respected by carriers when circumstances allow opportunities to allocate resources to higher bidders, which implies a lack of compliance from carriers to shippers and a need to reinforce contractual provisions. For instance, in the context of agricultural production, prices are tentatively agreed before a harvest season, but carriers (or intermediaries) adjust at their convenience when the demand of transportation rises close to the peak period [Brusset, 2010]. In such circumstances, the shipper bears little control over its logistics costs and risks late deliveries at customers locations.

From a carrier’s perspective, long-term contracts do not specify the exact capacity to hold available along the contract time. This capacity depends upon the shipper’s requests, which can vary from one period to the next. Nonetheless, some shippers guarantee relative volumes as opposed to absolute volumes to the carrier (e.g., a percentage of truckload trips in a corridor) in exchange of securing capacity to some extent. In that way, avoiding refusals for long-haul service requests [Lindsey & Mahmassani, 2017].

1.4 Risk-sharing Contracts

In the literature of operations management, several contract mechanisms are proposed and studied to deal with supply chain uncertainties in buyer-supplier relationships. The element of risk is always present under unknown outcome of demand or purchasing prices. For that matter, special trade conditions are established for total order quantities and final unit costs.

Different risk-sharing mechanisms like fixed commitment, quantity flexibility and option contracts are extensively studied and applied in the literature for various purposes. Such risk-sharing contracts differ in specifications about reservation and execution terms. In the reservation phase, an initial agreement is made for a first order at some price. In the execution phase, transactions are completed up to or down to the final required quantities and pre-agreed costs.

The aforementioned mechanisms deal with potential risks in various way. For instance, in fixed commitment contracts the initial reservation quantities and negotiation prices are fixed and the total costs owed to the supplier. Hence, while the cost is certain, the risks of shortage and excess inventory are assumed by the buyer [Martínez-deAlbéniz & Simchi-Levi, 2005].

Flexibility contracts allow to modify the reservation order by increasing or decreasing that quantity by some percentage, in this way, letting the buyer adjust its initial requirement. Option contracts specify a tentative initial order and partial payment, that could be then denied or accepted totally or to some extent. The buyer could lose the initial payment, while the supplier risks to produce with low return and hold unused inventory.

In the same way, several contract forms can be translated for procurement of freight transportation services. Yet, unlike physical items, transportation capacity supplied by common carriers is neither storable or salvageable. Also, the risk of shortage of transportation capacity is always present for shippers and usually tackled through spot markets. On the other hand, the risk of unused capacity is often run by carriers. In that sense, the shipper could benefit from cost savings if the risk of empty or idle vehicles is addressed or shared partially with logistics partners.

In the literature on transportation and logistics procurement, some alternative risk-sharing mechanisms were formulated for the shippers' benefits, like the use of derivative contracts, (denominated truckload options [Tsai et al., 2011]), following the example of stock options in finance markets [Tibben-Lembke & Rogers, 2006].

Thus, the shipper earns the right to activate logistics resources at a given price but is not obliged to make use of the option.

Separately, some authors mention the use and application of flexible terms in practice [Lindsey & Mahmassani, 2017]. Little attention is paid to fixed commitment contracts with carriers, possibly due to the long-term nature of contracts.

1.5 Contextual Applications

In sum, this thesis addresses transportation procurement planning integrating tactical and operational decisions. The carrier selection and the distribution network design problems are handled in parallel. The combined approach in conjunction with mid-term minimum commitment contracts attempt to resolve some of the problems observed in shipper-carrier relationships in road freight transportation services.

In that way, the planning framework extends the range of alternatives established in standard practices when companies decide to subcontract or outsource to third-party LSPs, which generally consists in long-term strategic selection of partners and/or one-time (load-by-load) services. At the same time, the shipper is involved into actively deciding on the most suitable routes and pathways to move or collect freight across the network of facilities, starting from origin points (e.g., central hubs, distribution centres) to destinations. In that way, selecting the group of operators that are integrated into the logistics network to cover widespread areas.

Real-life applications may arise in the context of large commercial companies with multi-national presence, outsourcing to a wide variety of transportation and logistics companies, including services like warehousing, trucking (LTL, FTL), freight-forwarding (consolidation, cross-docking), parcel delivery and collection (returns). Nonetheless, 3PLs or LSIs connecting directly with multiple operators (see Figure 1.2) could foresee cost benefits in application of multi-period agreements which provide for instance minimum number of shipments.

Operational decisions (facility/route selection, vehicle routing, demand allocation) taken into account for the flow and movement of shipments, determines in great extent, the network flow optimization problem. In the present work, the models address situations which feature multilevel distribution networks, with multiple origin-destination paths. The models are meaningful in dynamic environments, where the contracts represent both an opportunity and a risk due to high demand and price variability. The former is typically encountered in e-commerce, and the latter in

the market of full-truckload services, but could arise in different environments as well, such as, more generally, parcel delivery. In such environments, the interaction between tactical and operational decisions is more relevant.

In line with these elements, we look into the e-commerce logistics distribution from a perspective of an e-fulfilment company in need of planning accordingly the required capacity to dispatch products in forecoming periods. The increasing difficulty in solving larger instances of the problem leads to develop algorithmic strategies based on decomposition of the structure of the mathematical models.

1.6 Thesis Contents

The outline of the thesis is summarized in Figure 1.3 and described as follows:

In Chapter 2, the literature review is proposed separately for the main aspects of the problem studied. Namely, revision of risk-sharing contracts with minimum commitments, procurement planning problems with supplier selection, network design optimization problems with outsourcing alternatives and multi-period facility location problems.

In Chapter 3, the description of the problem in the context of e-commerce fulfilment is introduced. The network design problem is formulated as a *distribution network problem with minimum commitments* (DNPMC). Minimum commitment contracts are characterised with parametric notation according to their duration, service, penalty costs, and commitment conditions. Generalities of the *minimum purchase commitment* (MPC) contract are clearly defined and several formulations for the DNPMC with different types of MPC contracts are derived. Insights on the various mathematical models are discussed and analysed.

In Chapter 4, the multi-period distribution network problem (MDPC) is studied for MPC contracts lasting for multiple periods, fixed service costs (no penalties) and compulsory per period commitments. The mathematical representation, introduced in chapter 3, is further analysed and solution methods based on decomposition strategies are developed and implemented for several instances of the problem. Managerial insights and algorithmic performance results are provided following extensive computational experiments.

In Chapter 5, the stochastic programming formulation of the distribution network optimization problem

(DNPMC) is presented. The demand is modelled with a discrete probabilistic function and the minimum commitment contract in terms of capacity is introduced, *minimum capacity commitment (MCC)*. Several two-stage and multi-stage stochastic programs with fixed recourse are then derived according to variants of the MCC contract and capacity reservation rules.

In Chapter 6, the stochastic multi-period distribution network problem with MCC contracts (*mDNP*) is studied, in particular for MCC contracts lasting for multiple periods, penalties, compulsory per period commitments and ex-ante capacity reservation. Several heuristic strategies are proposed and tested over generated instances of the problem. Managerial insights and algorithmic performance results are provided for several heuristics procedures.

Finally, in chapter 7 perspectives and conclusions of the thesis are presented.

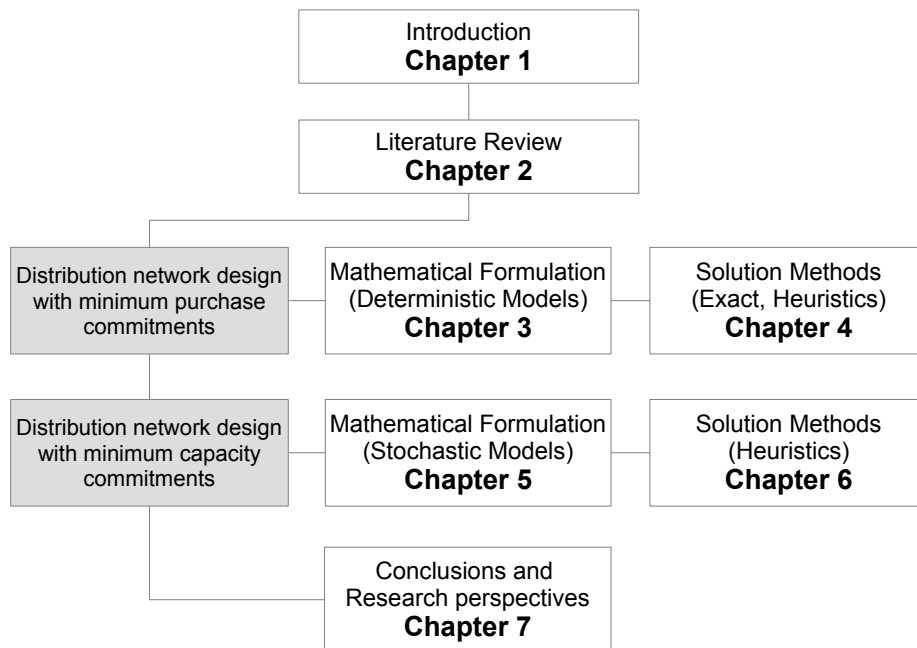


Figure 1.3: Thesis outline

Chapter 2

Literature Review

The literature review is separated in several subtopics and classes of problems approached in this thesis.

Designing and planning distribution network flows over a long-term dynamic horizon presents itself as one of the core subtopics. A part of it is concerned with the allocation of transportation corridors to eligible transporters working under risk-preventive minimum-commitment contracts. The decision-making procedure results in a finite selection of suppliers based on economical benefits in the long-term. Each of these topics will be handled in a separate section.

Section 2.1 discusses bilateral agreements that include risk-sharing provisions in buyer-supplier relationships. The supply chain literature describes under distinct names (e.g., capacity reservation, quantity flexibility, minimum quantity, options) several type of contracts sharing common characteristics. The literature on supply contracts is filtered through its ability to be translated in terms of freight transportation service procurement and is categorized according to common parameters.

Section 2.2 presents a review of supplier selection methods in procurement planning. Most of the OR/supply chain management literature that investigates supplier selection is applied in the context of manufacturing and commercial activities. In addition to classical procurement costs (fixed set-ups, inventory, purchasing), some applications take into account transportation costs and transport mode selection into the decision process.

Section 2.3 focuses on network design problems that include some form of out-

sourcing options in the decision making, meaning an external agent imposes an unconventional (typically, nonlinear) cost function for services provided. An important sub-theme in this area, addresses vehicle routing problems, which are generally grounded in the operational level of managerial decisions for day-to-day planning.

Section 2.4 centers on multi-period facility location problems which take place at a more strategic level. Particular interest is devoted to solution approaches and their applicability in the context of e-commerce distribution logistics. A classification of previous studies is proposed according to characteristics of each problem.

Section 2.5 discusses the particularities of e-commerce logistics distribution.

2.1 Supply Contracts with Risk-sharing Provisions

The literature on supply contracts provides managerial insights for tackling business risks in buyer-supplier relationships, in the face of uncertain future market behaviour. Several applications model formal agreements in the presence of uncertain demand and spot market costs. Setting up formal contracts in advance enables the buyer to negotiate fixed prices with suppliers. By contrast, sourcing 'on demand' from spot markets would entail risks of low market availability and high prices.

Reserving capacity through contracts requires to follow certain steps. For instance, the buyer needs to signal approximately the future demand levels (without knowing them exactly) in order for the supplier to reserve adequate resources, to allocate capacity and to prepare requests ahead in time. This estimation may not correctly reflect the subsequent behaviour of the market. As a consequence, falling short can require to consider sourcing alternatives on spot markets, which often entail higher-than-contract costs. On the other hand, overestimating demand requirements may result in excess inventory which, depending on the contractual agreement, may be either charged to one of the parties or split among them. Both detrimental situations arise, for example, in capital-intensive and high-tech industries with long lead times ([Spinler & Huchzermeier, 2006], [Jin & Wu, 2007]). In such industries, manufacturers are required to establish flexible mechanisms with suppliers, mitigating in that way overproduction costs or lost sales. Likewise, suppliers demand profitable conditions to commit production capacity regardless of the realized demand. These considerations give rise to various risk-sharing mechanisms.

Two distinct perspectives can be found in the discussion of supply chain contracts [Lariviere, 1999]. The first one deals with the design of contracts to achieve sup-

ply chain coordination between two partner firms in a decentralized supply chain. Setting up contract parameters in an appropriate way can provide a sufficient mechanism to achieve optimal global profitability, and to overcome inefficiency factors such as information asymmetry (undisclosed data), risk-aversion or power imbalance between players which motivate actions in their individual interest. In this setting, contract parameters are usually treated as endogenous within the problem formulation. Examples are found in Tsay [1999], Cachon & Lariviere [2005] and Corbett et al. [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 given. Examples are found, for example, in Martínez-deAlbéniz & Simchi-Levi [2005], Lian & Deshmukh [2009], Akbalik, Hadj-Alouane, Sauer, & Ghribi [2017]. This thesis falls in this type of framework: it approaches procurement planning of transportation services (from a single company perspective) under given terms of contracts.

Variants of contractual risk-sharing mechanisms are described in the literature under different names: *capacity reservation contracts* ([Jin & Wu, 2007], [Akbalik et al., 2017], [Li et al., 2020]), *total minimum order quantity* ([Bassok & Anupindi, 1997], [Chen & Krass, 2001]), *quantity flexibility contracts* ([Tsay, 1999], [Chen, Hum, & Sun, 2001], [Lian & Deshmukh, 2009], [Heydari et al., 2020], [Li et al., 2020]), *option contracts* ([Martínez-deAlbéniz & Simchi-Levi, 2005], [Spinler & Huchzermeier, 2006], [Nossohi & Nookabadi, 2016]). As a common feature, supply contracts with risk-sharing mechanisms specify general conditions at two moments in time, before and after the selling event, respectively. The procurement plan accordingly follows a two-stage process (see Figure 2.1).

In the first stage, the buyer reserves some of the supplier's capacity, based on its expectations regarding future demand and sales. As a risk-protection mechanism against lower-than-expected demand, the supplier may impose a minimum order commitment (and hence, a minimum guaranteed payment) to the buyer for instance to purchase a certain number of items, in order to cover at least the set-up costs that it incurs. This commitment reduces the risk of business failure for the supplier and stimulates it to invest in capacity. Thus, the supplier assign enough resources to cover *at least* the expected demand of the buyer (i.e, the reservation level). The supplier may also accept to designate more capacity than the reserved so as to earn the benefits in case the buyer experiences higher-than-expected demand. In that way, it concedes an additional safety stock. Consequently, the buyer enjoys a certain degree of flexibility in ordering up to the allocated capacity level at same contract

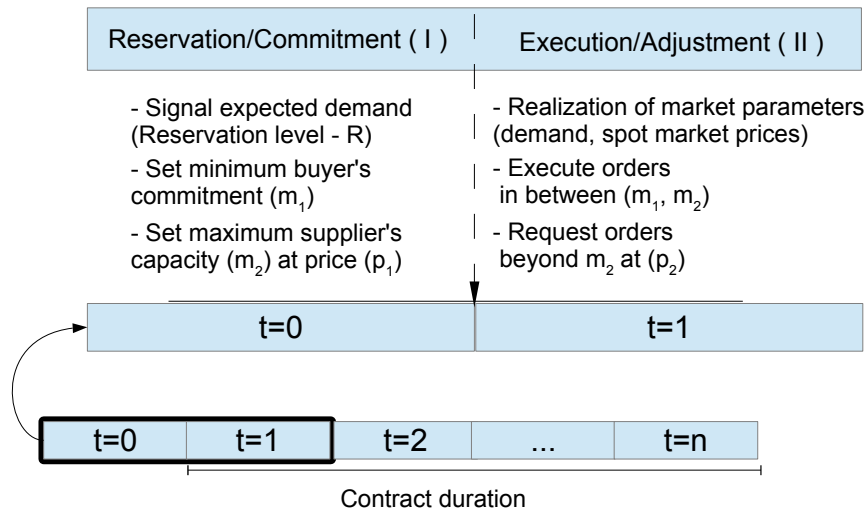


Figure 2.1: Two-stage procurement process

costs.

In the second stage, the demand is observed during the selling period and the buyer submits its final order to the supplier. The final order can vary with respect to the initial reservation in two ways: it can fall below the reservation level, R (but is limited downwards by the minimum order commitment, m_1), or it can rise above the reservation level (while limited upwards by the supplier's capacity allocated to the buyer, m_2). An additional pre-agreed convention may allow the buyer to order with the same supplier at outside-commitments costs above the prepared capacity.

2.1.1 Representation of Risk-sharing Provisions

A general class of two-stage risk-sharing mechanisms (involving several common types) can be described with the following notations: The amount R denotes the reservation level, that is, the estimated capacity utilization level to partially or completely fulfil the buyer's future demand requirements. The forecast value of R is complemented by the determination of two distinct parameters, setting on one side the minimum order commitment (m_1) with input value $\mu \in [0, 1]$, (where $m_1 = R(1 - \mu)$) and on the other side the maximum supplier's capacity allocated to the buyer (m_2) at contract prices (p_1), with input value $\omega \in [0, \infty)$, (where $m_2 = R(1 + \omega)$).

A premium price for orders beyond m_2 is denoted with p_2 . In addition, a fixed initial payment independent of parameters m_1 and m_2 , is defined as F_0 .

Figure 2.2 shows a general scheme for a contract with parameters $(m_1, m_2, p_1, p_2, F_0)$ settled in the reservation/commitment stage, previous to the selling season. In the execution/adjustment stage the range of executable orders are given by V_1 for orders (Q) at the agreed contract price p_1 and V_2 for orders above m_2 , charged at p_2 .

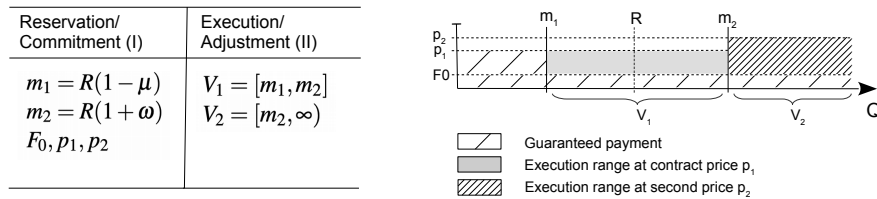


Figure 2.2: Configuration of contract parameters and risk-sharing provisions

Following the characterization of a given contract with risk-sharing provisions, two main features implicitly given are its *guaranteed payment* in the reservation (first) stage and its level of *flexibility* at the execution (second) stage for orders.

- The guaranteed payment can be implied by the minimum order commitment (m_1). Indeed, suppliers may secure a minimum payment linked to m_1 and the corresponding contract price p_1 . Thus, the parameter m_1 is strictly greater than 0, by setting $\mu < 1$ (e.g., in *quantity flexibility* contracts). In some other cases, the buyer is not obliged to execute orders at any level (i.e., setting $\mu = 1$, thus $m_1 = 0$) and can completely dismiss the available capacity m_2 , however, an initial (guar-

anteed) payment F_0 is stipulated. The value of F_0 can be relative to the maximum allocated capacity m_2 (e.g., in *options* contracts). In a more restricted setting, both fixed payment (F_0) and minimum order commitment (m_1) may be part of the agreement. In Figure 2.2, the areas with long-spaced diagonal lines represent the guaranteed payment for the supplier.

- Flexibility is characterized by μ and ω . When $\mu > 0$ (hence, $m_1 < R$), the buyer may choose to order less than the reservation level R , as long as the ordered quantity remains larger than m_1 . On the other end, when $\omega = \infty$ the buyer can order as much as needed at same contract prices, whereas when $\omega < \infty$, quantities requested by the buyer beyond m_2 may be charged at a higher price, or may not be available at all. Optionally, when $\omega = 0$ the supplier does not accept to allocate capacity beyond the reservation level. In Figure 2.2 the grey area represents the range of execution for orders at contract prices p_1 (V_1), and the area with short-spaced diagonal lines represent the range of execution for orders at a premium prices p_2 (V_2).

The two-column table and graph in Figure 2.2 are used along this thesis to describe and differentiate the type of contracts with risk-sharing provisions applied in the context of transportation procurement with multiple suppliers.

- When considering multi-period planning frameworks, a third defining characteristic of a procurement contract is its *duration*. The contract provisions can be specified in some contexts for a single period, with an eventual possibility to renovate a contract under new first-stage conditions (m_1, m_2, p_1, p_2, F_0). In other contexts, some contract parameters like the minimum order commitment (m_1), or the maximum available capacity at contract price (m_2) remain constant for multiple consecutive periods (while the contract is in effect), even though execution orders (in the second-stage) can be placed independently for each period.

2.1.2 Supply Contracts in the Literature

According to the characterization made in the previous section, Table 2.1 displays a classification of the literature on risk-sharing contracts

Authors	Name of contract	Contract duration		Commitment level			Flexibility (m_1 level)			Flexibility (m_2 level)			Flexibility (above m_2 level)
		Single	Multiple	F_0	m_1	m_2	$\mu = 0$	$\mu \in \{0, 1\}$	$\mu = 1$	$\omega = 0$	$\omega \in \{0, 1\}$	$\omega = \infty$	
Tsay [1999]	Quantity flexibility	x			x			x			x		x
Li et al. [2020]	Capacity reservation	x				x			x				x
Chen & Krass [2001]	Minimum total order quantity		x		x		x				x		x
Jin & Wu [2007]	Capacity reservation	x		x	x	x	x						
Brusset [2009]	Quantity flexibility		x	x	x			x			x		
Nosoohi & Nookabadi [2016]	Options	x				x		x			x		
Spinler & Huchzermeier [2006]	Options	x				x	x		x				x
Heydari et al. [2020]	Quantity flexibility	x			x			x			x		
Bassok & Anupindi [1997]	Total minimum commitment		x		x		x		x				x
Akbalik et al. [2017]	Capacity reservation		x	x					x				
Lian & Deshmukh [2009]	Quantity flexibility		x	x			x		x				x
Martínez-deAlbéniz & Simchi-Levi [2005]	Commitments with flexibilities		x	x	x			x			x		

Table 2.1: Classification of risk-sharing contracts

Only a few publications deal with risk-sharing contracts for the supply of transportation and logistics services. Chen et al. [2001] 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. During the planning period, any space used above the commitment level is charged at a premium price. Along the multi-period horizon, the *base* capacity commitment can be adjusted according to changing demand requirements.

A. Lim et al. [2006] incorporate minimum quantity commitment contracts into the classical transportation model. The freight owner decides how much to allocate shipments among multiple carriers, but a distinctive feature of the model is that a minimum volume has to be achieved with each selected carrier. The contracts provide the flexibility to assign as much load as wanted to any carrier, as long as the minimum quantity is respected.

Brusset [2009] considers three type of multi-period contracts to choose from, in order to establish a formal long-term relationship between a carrier and a supplier. Apart from a simple price-only contract without commitments, contracts with minimum purchase commitments (per period or over several periods) can be mutually selected based on numeric conditions of contract parameters. The presence of random demand and transportation prices in spot markets is modelled by stochastic processes.

In this thesis, the main focus is on optimizing transportation procurement costs for a single company in a multi-period and multi-supplier setting. We consider *minimum commitments* provisions expressed in two different metrics. The first metric expresses the commitments in monetary units (*purchase commitments*), as it is appropriately used in the presence of multiple products. (In transportation services, different corridors can be considered as different products with distinct prices). This contract holds similarities with the one of similar name described in Brusset [2009] which is also inspired from Cachon & Lariviere [2005] in a manufacturer/supplier context. A fixed fee is paid in advance to have got available some transportation capacity but does not entail the shipper to make use of it. The second metric is in terms of load or cargo units (*capacity commitments*) in order to account more explicitly for the reservation of transportation or logistical capacity. This type of quantity commitments is in the same vein as those described in A. Lim et al. [2006].

The separation in two type of metrics has two purposes : first to show an extensive (but not exhaustive) diversity of contracts with risk-sharing provisions which can result more or less advantageous in practical situations for modelling contractual

relationships in transportation. Second, to show how the diversity affects the mathematical representation (MILP) of the distribution network optimization problem under different contractual and contextual assumptions, in particular under stochastic/deterministic demand.

2.2 Supplier Selection in Procurement Planning

In this section, we revise the literature on multi-period procurement planning problems taking into account supplier selection. The revision is motivated by its relation with the tactical planning approach considered in this thesis, with the common objective of minimizing procurement costs. Differently we focus on procurement of services (transportation) rather than raw material or end products but similarly involving supplier (carrier) selection within a multi-period planning framework.

Supplier selection problems have been thoroughly studied in the context of production and manufacturing. Basically, the problem consists in buying one or multiple items offered by several candidate suppliers, of which a subset must be selected. Hence, finding optimal replenishment policies. Each supplier presents individual pricing schemes, which often includes volume and quantity discounts.

Following that description, several authors studied the single-item dynamic lot-sizing problem with selection of suppliers (SDSP-LSP). The buying firm may consider logistical and operational costs (inventory, handling, transportation, set-ups, production time) as well as demand variability features (backlogging, shortage) in the formulation of its cost minimization problem, leading to variants of similar problems.

Rosenblatt et al. [1998] examine a simplified case in which the demand is constant in each period and the firm must decide the frequency and amount of quantities to buy from each selected supplier. Zhao & Klabjan [2012] study the polytope associated with an MIP formulation of SDSP-LSP; they completely describe the polytope in the uncapacitated case, and they define valid inequalities in the capacitated case. Tempelmeier [2002] incorporates different discount structures under all-units and incremental policies. He develops a heuristic approach to be easily implemented in practical settings. Bai & Xu [2011] extend the diversity of cost proposals to three types of cost-structures including all-units, incremental, and multiple fixed costs. Dynamic programming algorithms are developed to solve certain cases in polynomial time. Ghaniabadi & Mazinani [2017] extend the dynamic LSP with multiple

suppliers to account for backlogging. A recursive formulation based on dynamic programming is proposed for this problem. Burke et al. [2008] propose a single-period LSP problem with three stages. First, the total order quantity is announced, next the suppliers submit their pricing schemes and capacity limits, and finally the buyer allocates quantity orders to the suppliers. Several discount schemes are considered such as linear, incremental and all-units. It was found that certain discount forms raise more computational difficulty than others, in which case heuristic-based algorithms are better suited. In a multi-product setting Xia & Wu [2007] combine qualitative criteria and supplier's discounts based on total business volume.

Some studies consider more detailed agreements with suppliers in the form of commitments contracts, that is, minimum required order quantities when contracting with a supplier. The delivery capacity may also be limited to a certain maximum value. For example, Chauhan & Proth [2003] incorporate minimum and maximum intervals of provisions for each supplier. Therefore, in case an order is received, no less than a minimum quantity is dispatched. The restriction of minimum order quantity serves to compensate for any fixed costs incurred in the supplier's production process, for example in batch-production manufacturing with fixed set-up costs. In the same logic Awasthi et al. [2009] considers supplier restrictions on minimum order quantities in addition to capacity limits, in the face of uncertain demand. Similarly, Zhang & Zhang [2011] assume an environment with min and max order levels and stochastic demand, in which a fixed cost is incurred whenever a supplier is selected.

Capacity reservation contracts have also been investigated in conjunction with supplier selection. Typically, a reservation (unitary) price is established beforehand. On top of that, an execution price is charged for the final order. Hazra & Mahadevan [2009] present the case of a buyer firm in need of capacity reservation from multiple sources in the face of demand uncertainty. A selection of suppliers together with contract reservation prices are obtained taking into account several parameters such as maximum capacity limit and minimum execution level. Park & Kim [2014] consider a replenishment model for a single-product from one or multiple suppliers under capacity reservation contracts. Akbalik et al. [2017] model the LSP problem with capacity reservation contracts and multiple suppliers, for a single-item that could be ordered per batches instead of units. An order quantity below a threshold is paid at the reservation price, while additional quantities are paid at a higher spot-market price.

A specific perspective on procurement planning involving shipping costs is found in a few applications. Mansini et al. [2012] address the problem from the stand-

point of a company willing to minimize its total procurement cost, including both the purchasing cost of multiple products and the associated transportation costs. The suppliers offer discounts depending on the quantity purchased. Moreover, the transportation costs are based on truckload shipping rates which depend on the total load volume. A paper by Jaruphongsa et al. [2005] includes transportation costs and shipment mode selection in a single-item dynamic LSP model. Multiple suppliers can be seen as transporters with distinct cost functions as well as lead times. The authors propose dynamic programming algorithms which allow to solve efficiently (in polynomial time) some special cases involving two transportation modes.

2.3 Network Optimization and Logistics Outsourcing Problems

In the domain of transportation and logistics, some applications in network design problems involve the selection of shipping companies. In a formulation of the freight allocation problem to multiple shipping alternatives, Qin et al. [2012] describe a procurement planning problem from a perspective of a large retail company buying products overseas and requiring transportation across several lanes. Shipping companies propose different cost structures with discounts according to the total volume transported. The allocation of freight has to comply with restrictions on minimum and maximum capacity allocation to every contract carrier in order to guarantee business profitability. A heuristic algorithm is proposed that relies on tabu search to solve instances of the problem.

As already mentioned in Section 2.1, A. Lim et al. [2006] consider minimum quantity contracts in a transportation problem. An international commercial company needs to ship customer's cargo overseas through carriers, which demand a minimum volume as a legal obligation. This enriched transportation problem is no longer solvable in polynomial time. The MIP formulation is enhanced with strong inequalities and its resolution is approached by branch-&-cut and rounding heuristics. In a complementary research [A. Lim et al., 2007], the minimum volume constraint is generalized to a number of containers. Linear and Lagrangian relaxation procedures are applied to the modified formulation. The work in our thesis generalizes this previous application in several ways. First, it considers facility selection variables for the distribution network problem. Furthermore, the time dimension is included to account for demand variability. Also, minimum quantity contracts have a determined duration that extends over several periods of a tactical medium-term planning horizon.

In a different context, the main interest is focused on vehicle routing problems (VRP) involving outsourcing alternatives (e.g., outside carriers). Beside the usual difficulty of routing a set of vehicles, an additional challenge consists in allocating ship-to locations to transport modes, either with private or public resources, with the objective to minimize transportation costs. Problems of this nature contemplate at least two transportation modes, primarily with the use of an internal fleet of vehicles at lower operational costs than the external source. The resulting integrated transportation planning problems deal with realistic managerial issues pertaining to limited private capacity and demand variability. In a seminal paper, Ball et al. [1983] consider a commercial firm that can service multiple customers with either private or public carriers. The problem aims at finding an optimal fleet size subject to limits on tour durations. More recently, similar problems are gathered under the name of VRP with private fleet and common carrier (VRPPC), whereby a shipper owning a limited fleet is interested in visiting multiple locations by subcontracting part of the service to external carriers. The outside carrier is expected to conduct LTL shipments from a central depot to the designated locations. Several variants of the VRPPC have been studied. For instance, Diaby & Ramesh [1995] study a distribution problem with one private vehicle and with outside carrier service. Zapfel & Bogl [2008] study an extension of the multi-period VRPTW with outsourcing options. In their work, a company should assign internal and external resources (drivers and trucks) to their routes. Krajewska & Kopfer [2009] contemplate two types of subcontracting options. Stenger, Vigo, et al. [2013] extend the model to include multiple central depots. Wang et al. [2014] diversify the type of subcontractors. Apart from common carriers in the market, vehicles may be sourced from allies in horizontal or vertical partnerships. Gham et al. [2017] cover vehicle rental options. Tours are designed for private and rented vehicles while some locations are left to outside carriers. To the best of our knowledge, multi-period or dynamic models have not yet received interest in VRPPC applications.

Subcontracting transportation services entails cost-related agreements with available carriers. In classical approaches, the cost of trucking services is based on constant or linear functions of distance and/or volume ([Chu, 2005], [Bolduc et al., 2008]). More recent studies integrate volume discounts (i.e., non-linear costs). In Stenger, Schneider, & Goeke [2013], subcontractors offer volume discounts when they deliver larger amounts of packages, and a minimum demand volume is allocated to the private fleet. Similarly, Dabia et al. [2019] propose a cost structure with quantity discounts. In Gham et al. [2017] multiple external (LTL) carriers offer volume discounts following independent cost functions.

The solution methodology for the VRPPC was initially based on heuristic methods

([Chu, 2005], [Bolduc et al., 2007]) followed by metaheuristics: tabu search algorithms ([Côté & Potvin, 2009], [Krajewska & Kopfer, 2009], [Euchi, 2017]), adaptive local neighbourhood search [Ziebuhr & Kopfer, 2014], iterated local search ([Castaneda et al., 2020]), variable neighbourhood search ([Stenger, Schneider, & Goeke, 2013], [Gham et al., 2017]), perturbation procedures ([Bolduc et al., 2008]). Some authors have attempted the use of exact methods ([Dabia et al., 2019]) or the inclusion of dual-based valid inequalities ([Diaby & Ramesh, 1995]).

In the present thesis, we consider a class of network design problems related to multi-period facility location models. Distinctively, the shipper relies entirely on external carriers for FTL shipping services departing from central warehouses. Eligible carriers propose different contract prices across a set of destinations. Linear and non-linear cost schemes (based on full truckloads rates) are allowed with the subcontractors. In addition, a minimum commitment payment (e.g., equivalent to a minimum number of shipments) is warranted to subcontractors in exchange for lower costs. The solution strategy developed in this work is based on decomposition algorithms for both exact and heuristics methods. The next section (2.4) provides a short literature review of multi-period location problems.

2.4 Multi-Period Facility Location Problems

In this section, the literature review focuses on multi-period facility location problems. Problems of this type are characterized by a finite planning horizon T , a set of candidate facilities I and directed arcs J , represented in a graph with nodes and arcs, $G = (N, A)$. For every $t \in T$, a subset of candidate nodes and arcs are functional to perform logistics activities. The modelling of a distribution network design problem comprises decisions about location of facilities and activation of distribution links (arcs) at each period of time. Conditions of parameters like demand or costs change dynamically from one period to another. Therefore, decisions regarding the network configuration are made per period in order to adjust adequately to the variations of the cost and demand patterns. In fact, expected changes in business and market conditions lead to redesign the logistics outlook (Cortinhal et al. [2015]), adapting in that way to operational requirements in the near future. Unlike single-period or "static" location problems, multi-period network design schemes capture the variations of key parameters.

Decisions about optimal location of logistics facilities can be made from scratch (*network design* problems) or from a current state of operational locations (*network*

redesign problems) Cortinhal et al. [2015]. Generally, problems of this nature involve long-term strategic decisions taking into account future market expectations. Clearly decisions at the strategic level have an impact on operational performance and on logistics costs. For instance, the allocation of customer demand to serving facilities is part of operational planning, but is influenced by network design decisions [Contreras & Fernandez, 2012]. Indeed, transferring supplies or commercial products to selling points entails transportation costs which are directly linked to the distance between supply and delivery points. Also, closeness to customer areas can be prioritized to improve responsiveness.

Multi-period facility location problems (MPFLP) allow a firm to expand or to reduce its overall distribution capacity. In one way, scaling up through acquisitions or construction of new buildings permits to increase the overall capacity and to reach a larger customer base, as part of the commercial and logistics strategy. On the other way, downsizing may entail closing or renting out unprofitable properties in certain areas where conditions are not favourable. The degree of flexibility in opening, modifying and closing can vary from one formulation of the problem to the next, depending on a particular context and/or assumptions.

The process of restructuring or redesigning a supply network is usually assumed to be subject to long lead times and considerable set-up costs. On one hand, capital investment is required to put in place new strategic assets (i.e., infrastructure, equipment), which are expected to remain active for a relatively long time before they return economic benefits. Therefore, within the formulation of network design problems, authors may impose minimum conditions on the operability of facilities once they are built or modified. Lead times are also considered in theoretical approaches. In other words, decisions are taken several periods ahead of time before new locations are integrated and operational in the supply network.

A feature commonly found in the literature is the position of the decision-maker as the owner and operator of facilities and distribution resources. It implies that the concerned company runs in-house logistics and transportation, as opposed to contracting with third-party service providers. It allows the model to handle transportation (allocation) costs in a simple form, e.g., as being linearly dependent on the quantity to be handled. On the other hand, making use of external transportation sources might lead to more complex, non-linear price schemes, or may require compliance with further restrictions, like limited possibilities for the allocation of customers based on the operability of sourcing companies in certain areas, or may require to select among several transportation modes with different costs.

Tables 2.2 and 2.3 classify the literature on MPFLP's according to different features

of the problems, including the degree of flexibility in network design decisions (i.e., capacity expansion and reduction). Different contexts lead to determine different assumptions in possible structure modifications. In a first type of situations (denoted as "1st Time Configuration" in the tables), new facilities can be opened and existing facilities can be closed at any time of the planning horizon ("Open NF" and "Close EF" are allowed); but once opened or closed, a facility must remain in the same state until the end of the horizon. Additional features may allow to expand ("Expand EF") or to reduce ("Reduce EF") the capacity of existing facilities, but here again only once in the horizon. In a second type of situations (denoted as "2nd Time Configuration"), two or more modifications of the same facility may be allowed over the horizon. This enables to reopen, re-close or change the capacity of facilities. A distinction is made between deterministic and stochastic mathematical formulation.

In the problems formulated in Chapters 3-6 of this thesis, it is assumed that candidate locations and distribution links are managed and operated by external carriers, but are accessible via fixed-duration leasing contracts with the carriers. A minimum commitment is established beforehand in order to reserve capacity as needed. Therefore, transportation resources (distribution links) and facilities are eligible through signed contracts. The underlying distribution network can be modified, expanded or reduced from time to time; operational decisions in terms of facility locations and allocation of customers are planned accordingly. Transportation capacity over available links is discretized per levels or modules, according to what is required. Similar to Correia & Melo [2017], capacity decisions may be subject to economies of scale.

In this context, contracts make it possible to "open" (i.e., to gain access to) subsets of facilities and links, corresponding to the regions of activity of certain carriers. "Closing" of facilities implicitly takes place at the predefined expiration period of a contract, although immediate renewal is always possible. In fact, it is possible to renew an ongoing contract at any time further along the planning horizon. A similar assumption is included in Dias et al. [2008] where locations remain open for at least a predetermined number of periods. The possibility to lease transportation and distribution resources from service providers allows firms to handle network design decisions at a tactical, mid-term level, rather than the more traditional strategic, long-term level. Thereby, large investment costs are replaced by periodic payments in relation with the capacity used. The resulting increased flexibility is especially beneficial in highly dynamic market conditions, like those outlined in Chapter 1.

Authors/Year	Network Configuration (Echelons)	Dynamic Network Design Flexibility								Solution Methodology
		1st Time Configuration				2nd Time Configuration				
		Open NF	Expand NF	Reduce EF	Close EF	Re-open EF	Expand EF	Reduce EF	Close NF	
Erlenkotter & Roy [1982]	1	x			x					Dual-based heuristic
Melachrinoudis & Min [2000]	2	x			x					MIP
Hormozi & Khumawala [1996]	1	x			x	x			x	MIP, dynamic programming
Canel et al. [2001]	3	x			x	x			x	B&B, dynamic programming
Dias et al. [2008]	2	x			x	x			x	Primal-Dual Heuristic
Albareda-Sambola et al. [2009]	1	x			x	x			x	Lagrangian Relaxation
Cortinhal et al. [2015]	3	x	x		x		x			MILP with additional inequalities
Melo et al. [2005]	1	x	x	x	x		x			MILP
Jena et al. [2015]	1	x	x	x	x	x	x	x	x	MILP
S.-K. Lim & Kim [1999]	1	x	x	x	x		x	x		Lagrangian relaxation-based heuristic, B&C
Correia & Melo [2017]	1	x	x	x	x	x	x	x	x	MILP with additional inequalities

Table 2.2: Multi-period facility location problems - deterministic models

Authors/Year	Network Configuration (Echelons)	Dynamic Network Design Flexibility								Solution Methodology
		1st Time Configuration				2nd Time Configuration				
		Open NF	Expand NF	Reduce EF	Close EF	Re-open EF	Expand EF	Reduce EF	Close NF	
Georgiadis et al. [2011]	3	x								B&B
Albareda-Sambola et al. [2013]	1	x			x					Fix-and-relax-coordination (matheuristic)
Mohamed et al. [2020]	1	x			x					MILP
Aghezzaf [2005]	2	x	x		x					Robust optimization, Lagrangian Relaxation
Zhuge et al. [2016]	1	x	x	x	x					MILP
Pimentel et al. [2013]	1	x			x	x			x	Lagrangian heuristic procedure

Table 2.3: Multi-period facility location problems - stochastic models

2.5 E-commerce Logistics Distribution

As more companies turn to e-commerce incorporating on-line sales and home delivery services, the need for cost-efficient and short-time transportation is increasing. However, the design of the underlying distribution process becomes more complex, even though storage facilities for e-fulfilment share common features with traditional warehouses [Agatz et al., 2008]. In fact, the presence of a traditional distribution structure adds multiple alternatives for the design of the delivery process in e-fulfillment. Moreover, the participation of companies into e-commerce channels entails opportunities to reach markets dispersed over larger geographical areas, which leads to longer transportation distances between customers and storage sites [Hesse, 2002].

The implementation of e-business strategies entails several modifications of the traditional distribution structure and network flows. For instance, an increasing need for distribution capabilities with closeness to high-density areas and incorporation of reverse flows supporting return policies. A general pattern is the rising establishment of transshipment centres, including large warehousing facilities and distribution centres [Hesse, 2002]. Market trends lead to transform old logistics platforms into high-throughput centres. The facility is more adequately adapted to support materials flow and consolidate outgoing products efficiently rather than purposely for storing functions [Hesse, 2002]. For instance, highly automated distribution centres help Ikea to achieve fast and effective transshipment [Yu et al., 2016]. The distribution center is composed of multiple docks for inbound and outgoing trucks, favouring cross-docking (CD) functions. Nonetheless, warehousing with large inventory capacity is still essential in a market inclined to fulfil the demand [Hesse, 2002].

The location of e-fulfilment centres in strategic points is also critical for the operational performance. Preferable locations are those close to large bases of customers in agglomerated areas. More likely, far from congested urban areas and with easy access to highways and transportation corridors, in addition to relative proximity to freight hubs (e.g., airports, seaports, logistics parks). For example, Amazon positions its transportation hubs for distribution purposes in districts with high population to reduce costs [Yu et al., 2016]. Other economical factors such as tax exemptions may induce the location decision. Distribution costs are taken as a priority but overall the main goal is to minimize total costs.

Different types of LSPs propose services to support multi-channel distribution for their clients. Some logistics integrators invest in their own distribution centres

(asset-based 3PL), which can provide a complete offer of solutions for clients selling through on-line platforms. Large e-retailers also decide to open and operate their own distribution centres and delivery fleet for e-fulfilment. They may also rely on the suppliers' facilities to keep storage of product and reduce inventory costs. Focusing on this strategy (known as *drop-shipping*), the e-retailers rely entirely on the suppliers for the delivery process and focus on sales and advertising [Agatz et al., 2008]. Express delivery companies and couriers also expanded their operation and located DCs facilities for effective transshipment of parcels. Therefore, multiple alternatives of logistics companies can be consider in the distribution service design for B2C e-commerce companies.

Chapter 3

Distribution Network Problems with Minimum Purchase Commitment Contracts

3.1 Introduction

In this thesis, we consider a tactical planning framework for procurement of freight transportation services. In addition to selecting third-party logistics companies with their own resources (asset-based carriers or 3PLs), the freight-owner (shipper) guides the movement of shipments that lead to minimize transportation costs. This approach differs from procurement strategies where companies leave complete responsibility to third-party operators about the way freight is directed from origin to destination points (for instance, selecting the transport mode and route). In our framework, the shipper considers multiple alternatives of asset-based companies which have access to private resources and networks of facilities. Finally, contracting with that combination which allows the shipper to minimize overall logistics costs.

The application of multi-period planning gains relevance in contexts with fluctuating demand when transportation requirements vary considerably from one period to another. The problem of serving multiple customers with fluctuating demand in widespread areas arises, typically, in e-commerce or in similar environments. This chapter discusses in more detail the context of the problem and the configuration of

the proposed distribution network in Section 3.2.

A fundamental component of the tactical planning framework, aiming at the cost-minimization objective, is to formalize mid-term contracts with carriers settling fixed costs per service, on condition that carriers receive benefits in terms of minimum business transactions (e.g., advanced payment) and specificity about future utilization of resources. Therefore, the contractual relationships are characterized by minimum commitment (MC) contracts, which offer risk-preventive mechanisms against market variabilities for both commercial agents. The planning framework is not oriented itself toward finding optimal minimum commitments, but rather toward selecting the most appropriate MC contracts from a pool of pre-designed agreements with carriers. Terms and conditions of minimum commitment (MC) contracts are described and expressed in parametric notation in Section 3.3. The parametric notation gives out numerous configurations, which helps to differentiate among several types of MC contracts, that could be put in practice in the context of transportation procurement. The extent of possibilities fits into the description of well-studied risk-sharing agreements in the literature.

The focal point of this chapter is about minimum purchase commitment (MPC) contracts, discussed in Section 3.4. From Sections 3.5 to 3.10, we present mathematical MILP formulations for variants of the problem of outsourcing and planning transportation and distribution (T&D) with MPC contracts. The purpose is to show the diversity of contractual mechanisms and their effects on the mathematical representation of a classical distribution network optimization model (such as the two-echelon location-allocation model).

The network optimization problem can be modelled as a multi-period facility location problem (MPFLP). Unlike traditional MPFLPs based on long-term strategic decisions for in-house operations, the tactical planning approach is based on temporary accessibility of transportation and distribution (T&D) capabilities provided by outside companies (carriers). Therefore, the eligibility of facilities is conditional to the existence of contracts and the optimal distribution flows vary according to the demand variations in customer areas. This class of network optimization problems is denominated *distribution network problems with minimum commitment contracts* (DNPMC).

The presented (MILP) models also demonstrate the mathematical structure similarities between models and the easiness to adapt for different settings of MPC contracts. Solving the mathematical models lead to assess the economical viability with respect to a non-contract short-term planning strategy, such as relying on spot markets. The chapter in itself does not address algorithmic questions related to the

solution of the models. These questions will be the topic of Chapter 4. Section 3.10 summarizes the contents by presenting findings and conclusions.

3.2 Two-Echelon Distribution System

A 3PL fulfilment (also called integrator) company operates dedicated distribution warehouses (DW) for the storage of products on behalf of multiple on-line sellers. Purchase orders are placed by end-customers via e-commerce marketplaces. They must be assembled into parcels and delivered within a few days. The warehouses are located strategically close to logistics hubs such as seaports, airports and urban logistics areas and are adequately equipped to hold inventory of a wide variety of products from multiple sellers worldwide. Every warehouse is replenished with equal range of products from regular sellers on a constant basis. Therefore, an order can be served from any of these large warehousing facilities. The high-volume of daily orders leaving the 3PL warehouses (outbounds flow) amounts to several truckloads (FTL) per day to be shipped in direction of customer areas.

The distribution system is divided in two parts or echelons (see Figure 3.1) in which first, parcels are directed towards cross-docking (CD) facilities (set I of size $|I|$) at regional zones, before being forwarded to differentiated customers areas (set K of size $|K|$). We assume that transportation in the first echelon and cross-docking services are provided by outside carriers (set E of size $|E|$), however, the decisions about the number and direction of FTL shipments are made by the shipper.

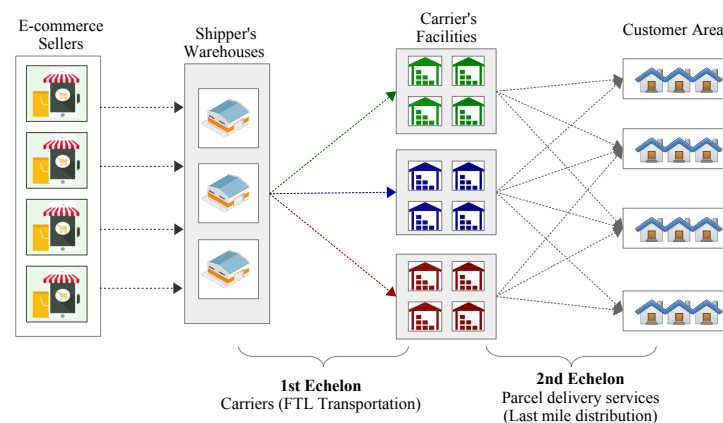


Figure 3.1: Two-echelon distribution network layout

In addition, the e-commerce integrator defines fast shipment delivery rules to guarantee express delivery times. First, it allocates each one of the externally operated CD facility $i \in I$ to the closest distribution warehouse, in order to reduce long-haul transportation distances. So, every CD facility i only receives parcels from one warehouse, nonetheless, a distribution warehouse can ship parcels towards multiple cross-docking points (as observed in Figure 3.1). This basic rule reduces the type of network design decisions in the two-echelon system (See Table 3.1). It considers only flow capacity decisions between the sets of facilities in the first echelon and disregards location of warehouses and allocation of CD points.

Network Design Decisions	1st-echelon	2nd-echelon
Outflow-Facility Location (Selection)	x	✓
Inflow-Facility Allocation	x	✓
Flow Capacity Selection	✓	✓

Table 3.1: Network design decisions

The combined services of long-haul transportation plus cross-docking is carried out by the same operator in the candidate set E . Each carrier $e \in E$ operates a subset I_e of CD facilities, where $\cup_{e \in E} I_e = I$. The number of shipments (between DW \rightarrow CD lanes) and the truck fleet size define the level of transportation capacity (set L of size $|L|$, as depicted in Figure 3.2) that every carrier has available. Note that the capacity Q_l for level $l \in L$ determines also the space required at the CD facilities for effective transshipment. Therefore, the cost of making use of a facility i is defined by the capacity level l , $F_{i,l}$. This cost function (per facility) is staircase and marginal costs might increase or decrease, even for an homogeneous set of vehicles.





Level	Fleet	Capacity
1		Q_1
2		Q_2
3		Q_3
4		Q_4

Figure 3.2: Capacity levels

As an additional feature, capacity levels could be differentiated per carrier $e \in E$, assuming for instance different truck fleets or limited resources for each company. In that case, the set $L_e \subseteq L$ would include only the capacity levels available for carrier e . In the models presented in this and following Chapters 4, 5 and 6, it is assumed that the capacity provided by any carrier is unlimited and enough to cover the total demand potentially allocated to CD facilities (i.e., set $L_e = L \forall e \in E$).

After parcels are sorted out and consolidated per destination in CD facilities, parcel delivery companies charge a unit cost per volume/weight of parcels. Thus, the shipping cost from outgoing facility $i \in I$ to customer area $k \in K$ is denoted $U_{i,k}$. A second fast shipment delivery rule forbids the delivery of parcels from those CD facilities which are far to customer areas. Therefore, a CD facility i can procure service over a finite subset of customer areas K_i . Likewise, a customer area k is designated to a subset of CD facilities I_k , from where it can be delivered.

In sum, the demands of customer area $k \in K$ (D_k^t) have to be completely satisfied at minimal cost. Several decisions have to be made by the shipper in every period $t \in T$. At the tactical level, it has to select carriers from the pool of eligible candidates E to be included into its transportation network. More specifically, the periods where contracts with selected carriers are valid, taking into account that a signed contract gives access to a network of facilities for a temporary lapse of time. Contracts terms and conditions (H, M, P) are further detailed in Section 3.3. The binary variables α_e^t , define whether to start the business agreement with carrier e at period t .

At a more operational level, the shipper determines network flow decisions. On one hand, the binary variable $v_{i,l}^t$ defines whether or not to make use of cross-docking facility i with entering transportation capacity at level l in period t (in which case the predefined cost $F_{i,l}$ is charged). The total payment due to carrier e is counted with auxiliary cost variables c_e^t and Λ_e^t , where t is the corresponding period. The auxiliary cost c_e^t helps to set a minimum payment owed to carrier e , given an initial commitment (in monetary terms) and regardless from the services effectively provided to the shipper, while Λ_e^t defines the costs to pay above the commitment level. On the other hand, the demand of customer area k is allocated to be delivered from one or several CD facilities in I_k , with continuous variable $q_{i,k}^t$.

The mathematical notation for the distribution network problem with minimum commitment contracts (DNPMC) is summarized in Table 3.2.

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 facilities operated by carrier e	i
I_k	Network of facilities that can serve customer area k	i
K_i	Customer areas that can be served from facility i	k
E_k	Carriers that can serve customer area k	e
L	Capacity levels	l

Input parameters

Symbol	Description
H_e	Duration of MC contracts with carrier e (number of periods)
M_e^t	Commitment payment (in monetary units) with carrier e in period t
P_e	Penalty rate charged to services provided above the commitment level
Q_l	Transportation and distribution capacity at level l (in transportation units)
$F_{i,l}$	Service costs for shipping to facility i at capacity level l with contracts (€)
D_k^t	Demand of customer area k in period t and scenario s (in transportation units)
$U_{i,k}$	Unit transportation costs for delivery from facility i to area k (€)

Decision variables

Symbol	Description
α_e^t	1 if a contract with carrier e starts at period t , 0 otherwise.
$v_{i,l}^t$	1 if facility i is operated at capacity level l , at period t , 0 otherwise.
$q_{i,k}^t$	Demand from customer area k allocated to facility i in period t (in transportation units)
c_e^t	Total payment charged by carrier e in period t
Λ_e^t	Costs charged above commitment by carrier e at period t

Table 3.2: Mathematical notations for DNPMC

3.3 Minimum Commitment Contracts

The class of business agreements that formalize the relationships between the 3PL company and outside carriers are *minimum commitment* contracts. Such contracts take into account risk-sharing conditions that reduce the likelihood of negative situations in the business relationship. In general terms, a minimum degree of business transactions is secured for the carrier in exchange of some transportation and logistical capacity at fixed prices.

Two families of minimum-commitment contracts are considered in this thesis. In this chapter and the next one, we focus on commitments made explicitly in monetary

terms. This means a fixed payment is guaranteed to carriers. In Chapters 5 and 6, we focus on commitments in terms of transportation capacity, which specify more explicitly future requirements of the shipper in relation to carriers' resources. Both type of contracts have in common the definition of the commitment in homogeneous units for an ensemble of logistics services. However, while the first stipulates explicitly financial benefits, the second clarifies operational utilization. The two types may be in their own way useful in real applications according to the shipper-carrier interests.

In our framework, the goal is not explicitly about contract design or contract configuration aiming at highlighting the most appropriate type of contract with respective parameters values in benefit of the shipper. Differently, it is about to optimize the transportation and distribution planning under given, predefined contractual provisions. Regardless of the contract type, additional common terms and conditions are subject to the accord and satisfaction of both parts, and are assumed to be previously agreed upon. They include, in particular, the terms and conditions regarding contract duration, services costs, penalty rates (if applicable), commitment payment policy and specifications. In this section, we describe each of these parameters that constitute an overall MC agreement.

3.3.1 Duration (H)

The duration of the MC contract is stated beforehand, as well as restrictions about renewals and termination. We assume that the planning horizon T is divided in discrete periods of equal length (months, quarters, etc.), i.e., $T = \{1, \dots, N\}$. In the tactical planning framework for transportation and logistics procurement, which occupies this thesis, MC agreements extend over more than one period, but are shorter than long-term contracts: typically, they would extend for a few weeks, and for at most one selling season. The contract duration could coincide with an H number of consecutive periods during which the shipper carries out a review of its logistics services requirements. If the contract duration differs per carrier, then it is more specifically denoted by H_e for carrier e .

The (3PL) e-commerce integrator negotiates the conditions of mid-term MC contracts with carriers. The shorter the contract, the greater flexibility the shipper has to adjust transportation and distribution resources to forthcoming demand requirements. However, the minimum commitment level (M) can be negotiated in relation with the contract duration. Carriers might impose high returns for short-term agreements and soften obligations for longer deals. In any case, parameters H and M

have to be previously agreed in accordance with each candidate carrier e , so as to obtain one eligible contract (H_e, M_e) .

In the mathematical formulation, each parameter H_e enters the model as a fixed value for each carrier $e \in E$. If a contract is to be renewed, this can only happen after expiration of the current one, and the renewed contract extends again over H_e periods. Cancellation or interruption are not allowed. In the particular case where $H_e = 1$ (e.g., monthly planning review and one month service contract), the tactical planning problem converts into a collection of independent single-period distribution network problems (sDNPMC). Otherwise, if $H_e > 1$, contracts last for more than one period (at least for one carrier) and a multi-period (mDNPMC) formulation is required.

3.3.2 Service and Penalty Costs (F, P)

Transportation and parcel consolidation services are combined together and priced at fixed rates by the carriers accepting to work under MC contracts. This means full truck-load (FTL) prices include together charges for transportation plus handling of parcels at cross-docking points. The prices agreed contractually (denoted by the generic parameter F) remain constant independently of market variations. Carriers' services prices $F_{i,l}$ are specified per shipping lane/facility ($i \in I$) and truck capacity used ($l \in L$).

In addition, MC contracts may or may not stipulate higher prices for services provided by a carrier above the commitment level (M), assuming more capacity is available. Comprehensibly, the carrier may charge an additional rate for services requested beyond the initial commitment in order to cover unanticipated costs, such as vehicle repositioning to the origin warehouse, or subcontracting services externally.

In our formulations, a parameter P defines the penalty or additional rate charged to the shipper. If $P = 0$, then no penalty is imposed and the totality of services are paid at contractual prices. Otherwise, if P is greater than 0, then the penalty is applied to the costs incurred above the commitment level M . The value of P can be seen as an adjustment factor as if services charged above commitment were off-contract and requested through the spot market. Its value can also be specified as P_e for carrier e .

3.3.3 Minimum Commitment (M)

As part of an MC contract, specifications related to the minimum commitment level must be made clear. The minimum commitment is denoted with the generic parameter M . Once a contract is signed, the shipper is obliged to comply with this terms in spite of market variations. In turn, carriers must respond to transportation and cross-docking requirements. For simplicity, we assume any carrier has more than enough available resources to respond to requests of the shipper.

Two distinct types of commitments are considered in this thesis. The *minimum purchase commitment* (MPC) expresses M in monetary units, whereas the *minimum capacity commitment* (MCC) expresses M in load/capacity units. Regardless of its nature, the commitment works as a mechanism to gain access to transportation and distribution resources at contract prices (F).

Moreover, in both cases, the commitment level M can relate either to the whole duration of the contract (*total commitment* - M^H) or to each individual period (*per period commitment* - M^t). In addition, it can be specified for the complete network of carrier e , (M_e) or for each specific facility/shipping lane (M_i). Therefore, several types of MC contracts can be derived (see Table 3.3).

MC Parameters	Option 1	Option 2
Type	Purchase Commitment (MPC)	Capacity Commitment (MCC)
Duration (H)	Single period ($H = 1$)	Multi-period ($H > 1$)
Penalty Costs (P)	No penalty ($P = 0$)	Penalty ($P > 0$)
MC Conditions (M): Policy	Per period (t)	Over the contract duration (H)
MC Conditions (M): Specification	Full carrier services (e)	Specified per service (i)

Table 3.3: Characterization of MC Contracts

The nature of an MC contract is differentiated (in the present work) by the acronym MPC or MCC followed by the list of parameters (H, P, M) with valid attribute values. These attributes have to be established according to the characteristics in Table 3.3: (1) The duration of contracts, (2) whether a penalty is imposed for costs above commitment or not, (3) the commitment policy (periodicity) and the specification per service. For instance, a single-period MC contract based on per period commitments, covering all carrier services, and with penalty costs is denoted - as $MPC(H = 1, P > 0, M = M_e^t)$.

For a given configuration of the parametric setting, one type of contract is obtained

to model the business relationships between the shipper and candidate carriers. Each particular contract is implemented into the mathematical modelling of DNPMC. In this chapter we focus on MPC contracts specifying commitments for full carrier services M_e . (The study and implementation of MCC contracts together with considerations of single-service commitments (M_i) are part of the contents of Chapter 5).

3.4 The Minimum Purchase Commitment (MPC) Contract

The *minimum purchase commitment* contract is a simple agreement between two parties that specifies a minimum monetary transaction in exchange of products or services during some period of time (e.g., one month, one semester). The buyer is committed to pay a certain amount of money equivalent to a bundle of products or services, in principle unspecified. Unless the supplier sells a single product/service with unique cost, the bundle of products can take several forms depending on their individual costs.

The guaranteed payment might be in relation with the expected amount of products/services to purchase or relative to a budget allocation limit. In any case, the MPC agreed has to be paid, independent of the number of products/services finally required. This initial agreement motivates the supplier to allocate resources so as to fulfil the buyer's requirements during the specified period. The MPC works as a cover or entry fee which guarantees at least a minimum return on the supplier's assets. On the other hand, the buyer can profit from advantageous price schemes, such as fixed contract prices and quantity-based discounts.

In the context of transportation and logistics services, the MPC contract is applied to a business relationship between one shipper and one carrier. In the context of e-commerce logistics distribution, the 3PL integrator acts as the shipper which negotiates MPC agreements with carriers for logistics services including long-haul (FTL) transportation and cross-docking functions. As long as the MPC is satisfiable enough, the FTL carrier charges contract prices with favourable marginal costs compared to the spot market.

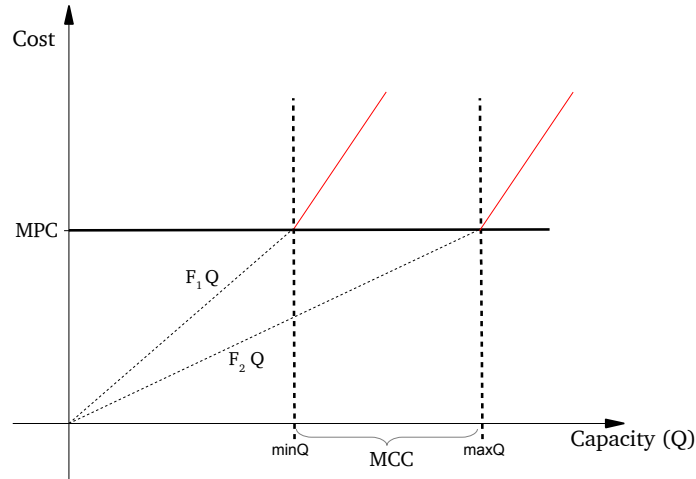


Figure 3.3: The minimum purchase commitment contract

The costs function associated with an MPC contract is represented in Figure 3.3 for two services with respective marginal contractual costs ($F_1 > F_2$). The minimum capacity utilization required to fulfil the MPC condition varies according to the amount ordered for each service, showing the capacity levels ($\min Q$, $\max Q$) at which the MPC is achieved if only one service is demanded. The marginal costs may become higher after surpassing the MPC limit to account for the penalty rate (or spot market costs) above the commitment level MPC. Note that if the supplier only offers one type of service with a fixed marginal cost, the MPC could be translated in capacity units. However, there might be discrepancies since logistical capacity is often valued according to resource sizes (e.g, truckload or facility capacity).

In this chapter, we propose different mixed-integer programming (MILP) formulations for distribution network problems DNPMC under various settings of MPC contract parameters, in terms of duration (H), penalty rates (P), and minimum commitment terms (M). This difference is further illustrated based on the discussion made in Section 2.1.1 for a general class of two-phase supply contracts with risk-sharing provisions. The illustrative graph is included in some of the models to observe how the MPC contracts fits within the general configuration scheme. We shall return to the case of minimum capacity commitments in Chapters 5 and 6.

3.5 Single-period Distribution Network Problem

MPC($H = 1, P = 0, M = M_e$)

In its simplest setting, the MPC contract takes the form MPC($H = 1, P = 0, M = M_e$): it lasts for a unique period of time (e.g., one month), the shipper can request as much as services as it needs provided that the minimum payment (M_e) is made. No additional costs or higher rates apply regardless of the amount of transportation and distribution services demanded. Note once again that, from the managerial perspective, the MPC value M_e may heavily impact the decisions made about the selection of contracts.

Figure 3.4 shows the configuration of the MPC contract without penalty rates above the commitment level M_e . No minimum requests (m_1) are enforced to the shipper, and the shipper profits from unlimited capacity allocation (m_2) at contractual costs ($F_{i,l}$). The guaranteed payment for the carrier is implied by the minimum purchase payment M_e . In the execution phase, the range of capacity requests at contract prices (V_1) is maximum.

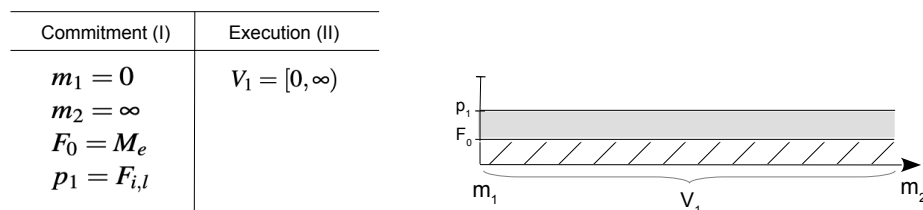


Figure 3.4: Configuration of an MPC contract without penalty rates

Considering contracts lasting for one determined period, the duration H can be viewed as being equal to the length of the planning horizon T . It means that, for a lengthy planning horizon (e.g., several months, one year), information about demand of customer areas is aggregated without taking into account internal fluctuations in that interval of time. After solving the single-period problem a new selection of contracts (carriers) is chosen. Therefore, deciding to get involved with certain carriers bears no consequences afterwards.

A slightly different perspective on the assumptions is that we consider a multi-period planning framework for the same horizon T and each period of the planning horizon can be handled independently without considering future conditions of the

market or availabilities of transportation. This short-sighted perspective restricts the mid-term tactical view of the problem, but we will see in Chapter 4 that the single-period model can be algorithmically useful even when dealing with more complex assumptions.

A mathematical formulation of the single-period DNPMC model is displayed here-under:

Objective Function

$$\min \sum_{e \in E} c_e + \sum_{k \in K} \sum_{i \in I_k} U_{i,k} q_{i,k} \quad (3.1)$$

Constraints

$$\sum_{i \in I_k} q_{i,k} = D_k \quad \forall k \in K \quad (3.2)$$

$$\sum_{k \in K_i} q_{i,k} \leq \sum_{l \in L} Q_l v_{i,l} \quad \forall i \in I \quad (3.3)$$

$$q_{i,k} \leq D_k \sum_{l \in L} v_{i,l} \quad \forall i \in I, k \in K \quad (3.4)$$

$$\sum_{l \in L} v_{i,l} \leq \alpha_e \quad \forall i \in I_e, e \in E \quad (3.5)$$

$$c_e \geq M_e \alpha_e \quad \forall e \in E \quad (3.6)$$

$$c_e \geq \sum_{l \in L} \sum_{i \in I_e} F_{i,l} v_{i,l} \quad \forall e \in E \quad (3.7)$$

$$v_{i,l} \in \{0, 1\} \quad \forall i \in I, l \in L \quad (3.8)$$

$$\alpha_e \in \{0, 1\} \quad \forall e \in E \quad (3.9)$$

$$q_{i,k} \geq 0 \quad \forall i \in I, k \in K \quad (3.10)$$

The objective function is shown on the top left side. The shipper's goal is to minimize the costs of transportation and distribution (T&D) services based on the two-echelon distribution network system. The first-stage costs, charged by FTL carriers, are expressed by auxiliary variables c_e for each potential carrier ($e \in E$), while the last-stage operations are performed by parcel delivery companies which charge a fixed cost per unit ($U_{i,k}$) on each origin-destination route.

The group of constraints relates to transportation capacity location and demand allocation. The demand D_k of every customer area has to be completely fulfilled (Eq. 3.2). Trucking capacity in load units (e.g. tons) is set by levels Q_l . Each

level corresponds to a capacity limit equivalent to a truck fleet. The truck capacity entering a cross-docking point (input flow) bounds the allocation to customer areas (output flow) (Eq. (3.3)). The output flow from a cross-docking facility i to a destination k is limited by the availability of the facility and by the demand of the destination (Eq. (3.4)). (When the $v_{i,l}$ variables take binary values, constraints in Eq. (3.4) are implied by Eqs.(3.2)-(3.3), but they strengthen the linear relaxation of the formulation.)

The second group of constraints (after the first dotted line) is linked to MPC contracts conditions and permissions. The subset of cross-docking facilities I_e and their potential inbound flows (according to the set of capacity levels L) are eligible only if a contract with carrier e is signed (Eq. (3.5)) in which case, the commitment payment M_e is charged (Eq. (3.6)). Then, a third group of constraints (after the second dotted line) define auxiliary costs functions for first-stage transportation costs with FTL carriers. The total cost of resource utilization is computed in Eq. (3.7). While a contract is ongoing, the auxiliary variable (c_e) for contract costs with carrier e is either lower-bounded by the commitment payment M_e or the total costs of services effectively provided. Finally, at the bottom is defined the nature of binary (Eqs. (3.8), (3.9)) and continuous (Eq. (3.10)) decision variables. Note that in this model, it is mandatory to open at least some contracts in order to have access to carriers' services.

The single-period formulation serves as a base on which more general models are build upon.

3.6 Multi-period Distribution Network Problem

MPC($H > 1, P = 0, M = M_e^t$)

Assume now that the shipper's procurement planning of transportation and distribution (T&D) services is reviewed in each period and more frequently than mid-term contracts (e.g., weekly planning and monthly contracts). Following that policy, the planning horizon is divided in multiple periods $t \in T$ wherein decisions are made. The MPC contracts now take the form MPC($H > 1, P = 0, M = M_e^t$). They extend over several planning periods for all $e \in E$, but not longer than the full planning horizon length: $1 < H_e \leq N$. The committed amount M_e^t must be at least paid in each period t where a contract with carrier e is in force. No extra rates or penalty costs apply, and resource utilization above the commitment is paid at contractual rates. The representation in Figure 3.4 also applies in the multi-period setting. The only

variation regards the guaranteed payment F_0 which can vary per period according to M_e^t . Nonetheless, the execution range for capacity requests remains unlimited.

The reviewing per period of T&D requirements allows the shipper to open or renew MPC contracts in any period $t \in T$ while others are still ongoing. Every new contract adds up to the list of available T&D resources. In that way, it is possible to raise capacity in periods where T&D requirements surge. Otherwise, when demand is decreasing, some contracts can be left to expire without renewal.

When $H_e > 1$, decisions made in any period $t \in T$ to open new contracts affect the future state of available capacity and as a consequence, they also condition future decisions made at the tactical level (selection of new contracts) and operational level (location and allocation plan). The mathematical model requires to specify decision variables in each period. In addition, some constraints handle the validity of contracts across several periods.

The multi-period DNPMC model is presented below and described as follows. The objective function (Eq. 3.11) minimizes total distribution costs along T . Transportation and distribution costs in the first and second echelon vary from one period to another according to demand fluctuations and cost-efficient allocation of resources.

Location-allocation constraints (bottom-left equations) are similar to those of the single-period model in Section 3.5 except for the added superscript t . The demand D_k^t of every customer area has to be satisfied at all times (Eq. (3.12)), the flow balance at cross-docking points is controlled with constraints in Eq. (3.13), while the output is limited by constraints in Eq. (3.14).

The resources (transportation lanes and facilities) of carrier e can only be used at period t if an MPC contract with e has been previously opened and is still in force at period t . This is expressed by constraints (3.15): note that each binary variable α_e^t takes value 1 only for the starting period of a contract. On the other hand, the value of the expression $\sum_{n=0}^{H_e-1} \alpha_e^{t-n}$ determines whether a contract with duration H_e is still valid or not in period t . When this is the case, the commitment payment M_e^t is simultaneously enforced in period t (Eq. 3.16). Eq. (3.17) expresses the assumption that a new contract cannot be signed with a carrier until the previous one has expired. (This modeling assumption could be simply removed in situations where it does not apply.) The auxiliary cost function (Eq. 3.18) captures resource utilization costs at contract prices $F_{i,l}$ in the first stage of the distribution network. Binary and continuous variables are defined by constraints (3.19)-(3.21).

Objective Function

$$\min \sum_{t \in T} \left[\sum_{e \in E} c_e^t + \sum_{k \in K} \sum_{i \in I_k} U_{i,k} q_{i,k}^t \right] \quad (3.11)$$

Constraints

$$\sum_{i \in I_k} q_{i,k}^t = D_k^t \quad \forall k \in K, t \in T \quad (3.12)$$

$$\sum_{k \in K_i} q_{i,k}^t \leq \sum_{l \in L} Q_l v_{i,l}^t \quad \forall i \in I, t \in T \quad (3.13)$$

$$q_{i,k}^t \leq D_k^t \sum_{l \in L} v_{i,l}^t \quad \forall i \in I, k \in K, t \in T \quad (3.14)$$

$$\sum_{l \in L} v_{i,l}^t \leq \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall i \in I_e, e \in E, t \in T \quad (3.15)$$

$$c_e^t \geq M_e^t \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall e \in E, t \in T \quad (3.16)$$

$$\sum_{n=0}^{H_e-1} \alpha_e^{t-n} \leq 1 \quad \forall e \in E, t \in T \quad (3.17)$$

$$c_e^t \geq \sum_{l \in L} \sum_{i \in I_e} F_{i,l} v_{i,l}^t \quad \forall e \in E, t \in T \quad (3.18)$$

$$v_{i,l}^t \in \{0, 1\} \quad \forall i \in I, l \in L, t \in T \quad (3.19)$$

$$\alpha_e^t \in \{0, 1\} \quad \forall e \in E, t \in T \quad (3.20)$$

$$q_{i,k}^t \geq 0 \quad \forall i \in I, k \in K, t \in T \quad (3.21)$$

Observe that in this model, constraints (3.15)-(3.17) are the only ones linking several planning periods together, and that the linkage is created by the contract variables α_e^t : this reflects the fact that planning decisions are independently made in successive periods, except for the contracting decisions which extend their effects over the duration of the contracts. In particular, we do not consider the possibility to carry inventory from one period to the next at the cross-docking facilities, nor to delay, nor to advance the deliveries to the end customers. It should also be noted that the beginning and the end of the planning horizon may require some adaptation of the expression of the constraints, in order to account for the existence of contracts taking their effects before period 1 or terminating after period N . For the sake of simplicity and of clarity, we do not explicitly spell out these boundary conditions here, nor in the following sections.

This mathematical model is much more difficult to solve than the single-period model in Section 3.5, since it does not only determines which candidate carriers to choose, but also when to start business relationships that extend over multiple periods. Several optimization algorithms for this model will be developed and evaluated in Chapter 4. In the remaining sections of the current chapter, we discuss a few variants of the model which may prove of interest in different applied settings, but which are not further explored in the thesis. The consideration of off-contract prices in Section 3.8, however, will prove of interest when we discuss stochastic models in Chapters 5 and 6.

3.7 Multi-period Distribution Network Problem with Penalty Costs - MPC($H > 1, P > 0, M = M_e^t$)

Considering further expansions of the multi-period DNPMC model, let us assume now that the carriers charge penalty costs for resource utilization above commitment (i.e., $P > 0$). The contract type takes the form MPC($H > 1, P > 0, M = M_e^t$). The penalty applies to the cumulative costs surpassing the MPC threshold. It can be specified per carrier (P_e) as part of the individual pre-agreed conditions, as carriers define themselves standard conditions for all their customers.

In comparison with the previous type, Figure 3.5 shows the configuration of the MPC contract with penalty rates above the commitment level M_e . Similarly, no minimum requests (m_1) are enforced to the shipper, but the carrier allocates limited capacity (m_2) at contractual costs ($F_{i,l}$). Observe that m_2 depends on the specific demanded services, due to the fact that services differ in price. As illustrated in Figure 3.3 the limits of m_2 are between the capacity obtained for the most expensive service ($minQ$) and the least expensive service ($maxQ$) assuming linear marginal cost. Above the undefined value m_2 , costs are penalized by factor P_e . The guaranteed payment for the carrier is implied by the minimum purchase payment M_e . In the execution phase, the range of capacity requests at contract prices (V_1) is limited by m_2 , and with penalized costs (V_2) is unlimited from there on. We can see that the flexibility in requesting T&D capacity is maximum as in the previous models, but comes at expense of two parameters, minimum payment M_e and penalty rate P_e .

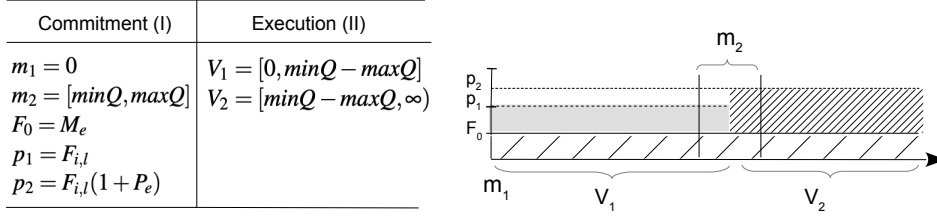


Figure 3.5: Configuration of an MPC contract with penalty rates

Naturally, high penalties prevent the shipper from deviating from using resources covered by the minimum commitment, while low penalties may not affect very much the solution obtained with the model in Section 3.6. An optimal decision must balance out between assuming additional charges and finding alternative ways to ship parcels in the interest of minimizing costs.

As a consequence, the mathematical model entails several modifications. The objective function (3.11) has a similar expression as in the previous model. However, the service costs c_e^t charged by each contracted carrier e in period t consist now of two components, as expressed by Eq. (3.22). First, the minimum payment M_e^t is charged. Next, additional costs beyond M_e^t are computed through the introduction of non-negative continuous auxiliary variables Λ_e^t : in view of constraints in Eqs. (3.23) and (3.24), Λ_e^t captures the amount due to carrier e in period t , calculated at the contract rates $F_{i,t}$, beyond the commitment payment. Because of constraint in Eq. (3.15), Λ_e^t is always set to 0 when there is no ongoing contract with carrier e at period t . Otherwise, utilization costs for requested services apply and the commitment payment M_e^t is subtracted to determine the value of Λ_e^t . It is again equal to zero if utilization costs do not surpass the MPC with carrier e . Finally, the penalty P_e is applied to Λ_e^t , obtaining in that way the second part of the total costs c_e^t in Eq. (3.22). Note that (3.22) could be equivalently written as an equality constraint, since no other restriction applies to c_e^t .

The remaining constraints are the same as in the model of Section 3.5, and include the location-allocation constraints regarding demand satisfiability (Eq. (3.12)), the flow balance constraints at cross-docking points (Eq. (3.13)), the upper bounds on the outflow of parcels (Eq. (3.14)), the constraints expressing that the shipper access to T&D services is conditioned by the existence of a valid contract (Eq. (3.15)), and the restriction (Eq. (3.17)) on contract renewal.

Objective Function

$$\min \sum_{t \in T} \left[\sum_{e \in E} c_e^t + \sum_{k \in K} \sum_{i \in I_k} U_{i,k} q_{i,k}^t \right] \quad (3.11)$$

Constraints

$$\sum_{i \in I_k} q_{i,k}^t = D_k^t \quad \forall k \in K, t \in T \quad (3.12)$$

$$\sum_{k \in K_i} q_{i,k}^t \leq \sum_{l \in L} Q_l v_{i,l}^t \quad \forall i \in I, t \in T \quad (3.13)$$

$$q_{i,k}^t \leq D_k^t \sum_{l \in L} v_{i,l}^t \quad \forall i \in I, k \in K, t \in T \quad (3.14)$$

$$\sum_{l \in L} v_{i,l}^t \leq \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall i \in I_e, e \in E, t \in T \quad (3.15)$$

$$c_e^t \geq M_e^t \sum_{n=0}^{H_e-1} \alpha_e^{t-n} + \Lambda_e^t (1 + P_e) \quad \forall e \in E, t \in T \quad (3.22)$$

$$\sum_{n=0}^{H_e-1} \alpha_e^{t-n} \leq 1 \quad \forall e \in E, t \in T \quad (3.17)$$

$$\Lambda_e^t \geq \sum_{l \in L} \sum_{i \in I_e} F_{i,l} v_{i,l}^t - M_e^t \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall e \in E, t \in T \quad (3.23)$$

$$v_{i,l}^t \in \{0, 1\} \quad \forall i \in I, l \in L, t \in T \quad (3.19)$$

$$\alpha_e^t \in \{0, 1\} \quad \forall e \in E, t \in T \quad (3.20)$$

$$q_{i,k}^t \geq 0 \quad \forall i \in I, k \in K, t \in T \quad (3.21)$$

$$\Lambda_e^t \geq 0 \quad \forall e \in E, t \in T \quad (3.24)$$

Note that the same penalty rate is applied indistinctly for the whole group of logistics services in I_e . In Chapter 5, we look at some models where the penalty rate or the cost associated to services above the commitment level can be differentiated per service $i \in I_e$. As a particular case, if no penalties are imposed by the carriers (i.e., $P_e = 0$ for all $e \in E$), then the model is equivalent to the one in Section 3.6. Otherwise, when $P_e > 0$ for some e , an optimal solution is likely to provide a higher cost value, except in the possible situation where minimum purchase commitments are not exceeded in any period.

3.8 Multi-period Distribution Network Problem with Penalty Costs - MPC($H > 1, P > 0, M = M_e^t$) and Off-Contract Services - ($H = 1, P > 0, M = 0$)

In previous models, an important assumption states that the shipper can only gain access to the resources of a candidate carrier through MPC contracts (constraints in Eq. (3.15)). As a result, the shipper must have at least some ongoing contracts at every period. However, common carriers may also offer logistics services on-demand without requiring a minimum cover payment, in which case T&D service costs resemble those at the spot market. An off-contract service is represented with the following attributes: ($H = 1, P > 0, M = 0$).

From the economic point of view, the shipper can foresee cost savings by avoiding to get involved in additional MPC contracts in periods of low demand where parcels can be distributed either by available contract resources or, if necessary, by spot services. The values of P_e are key determinants for the selection of the procurement plan. To understand this, assume that the penalty parameter P_e stands for the relative difference between spot market rates and contract rates. For instance, for a given service, the spot market price might be F_s while the same service under contract with carrier e is charged at F_e . The relative difference is computed as $P_e = (F_s - F_e)/F_e$. Hence, the multiplier $(1 + P_e)$ in constraint (3.22) is equal to F_s/F_e , and it can be interpreted as an adjustment factor from contract rate to spot market prices. Low values of P_e can lead to resort exclusively to single-period agreements. In turn, high values entice the shipper to open MPC contracts and to use low-cost contractual capacity as much as possible.

This motivates the consideration of models where, in any period, the shipper can procure off-contract services from some of the candidate carriers: when compared with the model formulated in the previous section (3.7), it means that penalty costs apply above the commitment level, but that this commitment level can be viewed as being zero (with all services charged at the spot price) when no contract is in force.

The multi-period DNPMC model can now be formulated as follows (although spot prices are typically dynamic with time, we consider for simplicity that the parameter P_e is constant over the planning horizon).

Objective Function

$$\min \sum_{t \in T} \left[\sum_{e \in E} c_e^t + \sum_{k \in K} \sum_{i \in I_k} U_{i,k} q_{i,k}^t \right] \quad (3.11)$$

Constraints

$$\sum_{i \in I_k} q_{i,k}^t = D_k^t \quad \forall k \in K, t \in T \quad (3.12)$$

$$\sum_{k \in K_i} q_{i,k}^t \leq \sum_{l \in L} Q_l v_{i,l}^t \quad \forall i \in I, t \in T \quad (3.13)$$

$$q_{i,k}^t \leq D_k^t \sum_{l \in L} v_{i,l}^t \quad \forall i \in I, k \in K, t \in T \quad (3.14)$$

$$\dots \dots \dots$$

$$c_e^t \geq M_e^t \sum_{n=0}^{H_e-1} \alpha_e^{t-n} + \Lambda_e^t (1 + P_e) \quad \forall e \in E, t \in T \quad (3.22)$$

$$\sum_{n=0}^{H_e-1} \alpha_e^{t-n} \leq 1 \quad \forall e \in E, t \in T \quad (3.17)$$

$$\dots \dots \dots$$

$$\Lambda_e^t \geq \sum_{l \in L} \sum_{i \in I_e} F_{i,l} v_{i,l}^t - M_e^t \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall e \in E, t \in T \quad (3.23)$$

$$\dots \dots \dots$$

$$v_{i,l}^t \in \{0, 1\} \quad \forall i \in I, l \in L, t \in T \quad (3.19)$$

$$\alpha_e^t \in \{0, 1\} \quad \forall e \in E, t \in T \quad (3.20)$$

$$q_{i,k}^t \geq 0 \quad \forall i \in I, k \in K, t \in T \quad (3.21)$$

$$\Lambda_e^t \geq 0 \quad \forall e \in E, t \in T \quad (3.24)$$

The mathematical model is almost identical to the model in Section 3.7. The only difference is the removal of constraints (3.15) which prevented the shipper from using carrier resources without a contract. The interpretation of the auxiliary cost function Λ_e^t (Eq. 3.23) changes slightly, as it can now be positive even in the absence of a contract with carrier e in period t (i.e., when $\sum_{n=0}^{H_e-1} \alpha_e^{t-n} = 0$). It expresses costs above commitment if there is an ongoing MPC contract at time t , or one-time service costs if there is no MPC contract. In the former case, the guaranteed payment M_e^t is deducted from the resource utilization costs and any excess value is then penalized with parameter P_e by Eq. 3.22. In the latter case, only resource utilization costs are charged and subject to the adjusted rate $(1 + P_e)$ in constraint (3.22). Note again that parameter P_e is applied without difference in the services provided. This limitation of the model may not reflect the conditions in transportation markets were spot

prices are more linked to the demand in particular lanes, rather than decided by carriers.

Formally, this model increases the space of feasible solutions of the model in Section 3.7 and hence, can yield solutions with lower optimal cost. This is obviously due to the additional possibility to resort to one-time services in order to meet distribution requirements.

In previous models, the MPC contract was looked at from the perspective of per period minimum payments, arguably with restricted adaptability to transportation requirements but more attractive from the carriers point of view. In the two following models, we open the discussion to consider now a single MPC commitment prolonged over the contract duration rather than multiple MPCs, one for every period. This type of agreement is motivated from applications in the literature (some examples can be found in Brusset [2010] and Bassok & Anupindi [1997]). Naturally, this type of contracts can only be handled in a multi-period planning framework. Then, $H_e > 1$. In a similar fashion of previous models, we derive two possibilities for the costs of services once the commitments M_e is achieved. Namely, without and with penalty rates.

3.9 Multi-period Distribution Network Problem with Total Minimum Purchase Commitment **MPC($H > 1, P = 0, M = M_e^H$)**

Now consider the case in which the minimum purchase commitment (MPC) is defined for the complete duration of the contract (H_e) instead of individual periods. Disregarding penalty costs above the commitment level, contracts are of the form MPC($H > 1, P = 0, M = M_e^H$).

In comparison with the previous models, the shipper is not required to comply with a minimum payment at every period covered by the contract. Instead, the MPC must be attained over the contract duration H_e : the shipper may register some periods with small transactions, while others are large enough to surpass the required global payment M_e^H by the expiration date.

As in Sections (3.5)-(3.7), we assume that an MPC contract is necessary to include a carrier's resources in the distribution network. Moreover, for notational simplicity, we assume that M_e^H does not depend on the period when the contract is signed (it

could be replaced by $M_e^{H,t}$).

The DNPMC is now formulated as follows.

Objective Function

$$\min \sum_{t \in T} \left[\sum_{e \in E} c_e^t + \sum_{k \in K} \sum_{i \in I_k} U_{i,k} q_{i,k}^t \right] \quad (3.11)$$

Constraints

$$\sum_{i \in I_k} q_{i,k}^t = D_k^t \quad \forall k \in K, t \in T \quad (3.12)$$

$$\sum_{k \in K_i} q_{i,k}^t \leq \sum_{l \in L} Q_l v_{i,l}^t \quad \forall i \in I, t \in T \quad (3.13)$$

$$q_{i,k}^t \leq D_k^t \sum_{l \in L} v_{i,l}^t \quad \forall i \in I, k \in K, t \in T \quad (3.14)$$

$$\sum_{l \in L} v_{i,l}^t \leq \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall i \in I_e, e \in E, t \in T \quad (3.15)$$

$$\sum_{n=0}^{H_e-1} c_e^{t+n} \geq M_e^H \alpha_e^t \quad \forall e \in E, t \in T \quad (3.25)$$

$$\sum_{n=0}^{H_e-1} \alpha_e^{t-n} \leq 1 \quad \forall e \in E, t \in T \quad (3.17)$$

$$c_e^t \geq \sum_{l \in L} \sum_{i \in I_e} F_{i,l} v_{i,l}^t \quad \forall e \in E, t \in T \quad (3.18)$$

$$v_{i,l}^t \in \{0, 1\} \quad \forall i \in I, l \in L, t \in T \quad (3.19)$$

$$\alpha_e^t \in \{0, 1\} \quad \forall e \in E, t \in T \quad (3.20)$$

$$q_{i,k}^t \geq 0 \quad \forall i \in I, k \in K, t \in T \quad (3.21)$$

The mathematical model resembles very closely the one in Section 3.6. The only difference consists in replacing constraint (3.16) by constraint (3.25). In case a contract is ongoing with a given carrier at period t , the shipper pays the carrier at least for resource utilization costs (as expressed by constraint (3.18)). When the MPC (M_e^H) is higher than the total resource utilization costs over the whole contract duration, the values of c_e^t are adjusted to match the commitment payment M_e^H (constraint (3.25)); recall that the binary variable α_e^t only indicates when the contract is

opened).

When compared with an equivalent minimum purchase commitment applied in every period, that is, where $M_e^t = M_e^H / H_e$ for all t , the total minimum purchase commitment clearly allows for greater flexibility and hence, for lower costs. This translates into the property that, with this uniform setting of M_e^t , any feasible solution of the model in Section 3.6 is also feasible for the current model, but not vice-versa.

3.10 Multi-period Distribution Network Problem with Total Minimum Purchase Commitment and Penalty Costs - MPC ($H > 1, P > 0, M = M_e^H$)

Finally, in this section, the previous model is extended to include penalty costs. So, the mid-term contract is structured with parameters MPC($H > 1, P > 0, M = M_e^H$). Similar to the previous model in Section 3.9, the minimum payment M_e^H can be settled by adding up costs of services requested during the contract duration H_e . The penalty P_e is applied to the cost of services requested above the commitment level. In this formulation, an auxiliary variable μ_e^t is introduced to represent the cost of services above commitment for a contract *starting* in period t with carrier e . Note the difference with auxiliary variable λ_e^t used in previous models, which expresses the costs above the commitment level for period t .

Likewise, a new auxiliary variable B_e^t is added, which determines the total costs paid to carrier e for a contract starting in period t . It is equal to the pre-agreed MPC value M_e^H plus the costs $\mu_e^t(1 + P_e)$ charged above commitment (constraint in Eq. (3.27).

This results in the following non-linear DNPMC model wherein the modified objective function (3.26) includes a component for the total contract cost incurred with carrier e , plus the cost of delivery to the end customers.

Objective Function

$$\min \sum_{t \in T} \left[\sum_{e \in E} B_e^t + \sum_{k \in K} \sum_{i \in I^k} U_{i,k} q_{i,k}^t \right] \quad (3.26)$$

Constraints

$$\sum_{i \in I_k} q_{i,k}^t = D_k^t \quad \forall k \in K, t \in T \quad (3.12)$$

$$\sum_{k \in K_i} q_{i,k}^t \leq \sum_{l \in L} Q_l v_{i,l}^t \quad \forall i \in I, t \in T \quad (3.13)$$

$$q_{i,k}^t \leq D_k^t \sum_{l \in L} v_{i,l}^t \quad \forall i \in I, k \in K, t \in T \quad (3.14)$$

$$\sum_{l \in L} v_{i,l}^t \leq \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall i \in I_e, e \in E, t \in T \quad (3.15)$$

$$B_e^t = M_e^H \alpha_e^t + \mu_e^t (1 + P_e) \quad \forall e \in E, t \in T \quad (3.27)$$

$$\sum_{n=0}^{H_e-1} \alpha_e^{t-n} \leq 1 \quad \forall e \in E, t \in T \quad (3.17)$$

$$c_e^t \geq \sum_{l \in L} \sum_{i \in I_e} F_{i,l} v_{i,l}^t \quad \forall e \in E, t \in T \quad (3.18)$$

$$\mu_e^t \geq \left[\sum_{n=0}^{H_e-1} c_e^{t+n} - M_e^H \right] \alpha_e^t \quad \forall e \in E, t \in T \quad (3.28)$$

$$v_{i,l}^t \in \{0, 1\} \quad \forall i \in I, l \in L, t \in T \quad (3.19)$$

$$\mu_e^t \in \{0, 1\} \quad \forall e \in E, t \in T \quad (3.20)$$

$$q_{i,k}^t \geq 0 \quad \forall i \in I, k \in K, t \in T \quad (3.21)$$

$$\mu_e^t \geq 0 \quad \forall e \in E, t \in T \quad (3.29)$$

$$B_e^t \geq 0 \quad \forall e \in E, t \in T \quad (3.30)$$

The costs charged by a contracted carrier in each period amount at least to the resource utilization costs (3.18). Moreover, Eq. (3.28) computes the costs above commitment per contract starting in period t . In case a contract is open, total resource utilization costs over all contract periods are compared with the minimum commitment payment (M_e^H). If the sum ($\sum_{n=0}^{H_e-1} c_e^{t+n}$) is larger than M_e^H , then additional charges are adjusted with the penalty rate. Otherwise, only the commitment payment is charged in ((3.27)).

Following its definition, μ_e^t must be zero if no contract with carrier e starts in period

t . Thus, the difference $(\sum_{n=0}^{H_e-1} c_e^{t+n} - M_e^H)$ must be calculated only in case a contract starts in period t , which justifies the multiplication by variable α_e^t in Eq. (3.28). As a consequence, the resulting formulation is non-linear as it involves the product of two decision variables. It can be linearized by standard methods. Alternatively, constraint in Eq. (3.28). could also be replaced by

$$\mu_e^t \geq \left[\sum_{n=0}^{H_e-1} c_e^{t+n} - M_e^H \right] - \kappa(1 - \alpha_e^t) \quad \forall e \in E, t \in T,$$

where κ is a large enough constant (larger than $(\sum_{n=0}^{H_e-1} c_e^{t+n} - M_e^H)$ for all feasible solutions).

Note that this formulation could be used to model problem in Section 3.9 when the penalties P_e are equal to zero for all $e \in E$. However, this still results in a nonlinear model and does not immediately boil down to the model of Section 3.9. Note also that the current formulation does not allow to model off-contract requests (adjusted with penalties P_e) by simply removing constraints in Eq. (3.15), as was the case in Section 3.7. It is possible to claim a more complex structure, distinctive from the previous ones. Notably the non-linearity or inclusion of a Big-M constraint may complicate its computational resolution or inefficacy by standard optimization solvers. However, as it was suggested, solving alternative models with different conditions of contracts, like in Section 3.9 could serve as an starting point to obtain valid lower bounds and feasible solutions, respectively.

3.11 Conclusions

This chapter presented a variety of models and mixed-integer linear programming formulations for transportation capacity location problems with minimum purchase commitment contracts. The formulations are built around a common core of constraints, but vary according to the terms of the contracts and to the features of the cost function: presence or absence of penalties beyond the minimum purchase commitment level, possibility or not to hire “spot” capacity without MPC contract, total or per period purchase commitment.

In the next chapter, we will turn to the development of algorithms for the solution of the model presented in Section 3.5, and to an evaluation of the performance of these algorithms. We will return to more complex models involving off-contract use of transportation and logistics facilities in Chapters 5 and 6, when we discuss stochastic extensions of the problem.

Chapter 4

Multi-Period Distribution Network Problem with Minimum Purchase Commitment (MPC) Contracts

4.1 Introduction

In this chapter, we consider the distribution network problem arising in e-commerce logistics from the perspective of an e-fulfilment agent, introduced in Section 3.2. The relationships with logistics partners is made on the basis of an MPC contract with specific attributes ($H > 0, M = M_e^t, P = 0$). This mathematical model was introduced in Section 3.6 and is described in detail in Section 4.2 with added notes about its complexity in Section 4.3. The present chapter is focused on solution strategies for the multi period problem based on decomposition methods. First, the combinatorial benders algorithm is presented in Section 4.4 while a two-phase heuristic is presented in Section 4.5. The experimental design taking into account multiple instances of the problem is described in Section 4.6, and computational results are detailed in Section 4.7. Finally, conclusions are provided in Section 4.8. The description of the problem is described as follows.

A logistics service intermediary known as the *shipper* acts on behalf of online sellers for the distribution of products sold worldwide on the Internet. The shipper

manages the inventory from warehouses where the sellers send their products on a continuous basis. Then, this 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, cross-docking operations at intermediate facilities and on *parcel delivery services* for last-mile transportation to customers areas. See Figure 4.1 for a representation of the distribution network.

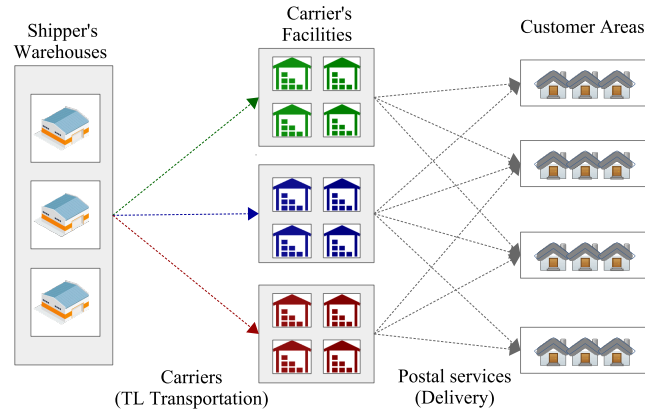


Figure 4.1: Distribution network layout

The focus of our work is on the optimization of 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 carrier's cross-docking facilities. We assume that the range of services offered by the carriers is limited to long-haul road transportation from the warehouses to their own facilities and to cross-docking operations (freight consolidation, vehicle un/loading...). Then, last-mile distribution of parcels from the cross-docking facilities to customer areas is performed by parcel delivery services.

The total cost incurred by the shipper depends on several factors. 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 actual total cost is charged by the carrier to the shipper. The minimum purchase commitment may vary in each period according

to related expected demands issued from forecasts. When a contract with carrier e expires (after H_e periods), it can be immediately renewed; but it cannot be modified nor extended before its expiration date.

In order to fit with real-life marginal decreasing prices, the amount charged by each carrier for transportation on a specific lane, from the warehouses to one of its cross-docking facilities, is a function of the total volume of freight. More precisely, we model it as a staircase function accounting for *capacity levels*, that is, the number of full truckload (FTL) shipments needed, with a decreasing marginal cost for each additional truck. We also assume that the number of FTL shipments to a facility determines the operational costs incurred for handling and cross-docking functions at this facility. Transportation of parcels from each cross-docking facility to each related customer area is performed by a specialized operator which charge a fixed rate (per distance and weight) for this last-mile parcel delivery service.

When trying to minimize its total distribution costs, the shipper tends to favor shipping lanes/facilities which minimize transportation and handling costs. However, e-commerce patterns and seasonality effects induce fluctuations of demand which are likely to affect the optimal selection of carriers and cross-docking points in each period, from a short term perspective. On the other hand, each contract ties the shipper to a carrier for successive periods of time, and this implies that the contract portfolio must be optimized from a mid-term perspective. These interactions between tactical and operational decision levels considerably increase the complexity of the resulting planning problem, that we call hereafter the *multi-period distribution network design problem with purchase-commitment contracts* (MDPC).

4.2 Mathematical formulation

In this section, we propose a mixed-integer linear programming (MILP) formulation of . We start with some definitions and notations (see also Table 4.1).

4.2.1 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

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 o.w.
$q_{i,k}^t$	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 (€)

Table 4.1: Mathematical notations for the MDPC model

$I = \cup_{e \in E} I_e$ forms the complete set of available cross-docking facilities. In order to be allowed 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 units, say, tonnes). 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 $k \in K$ can only be served on time but if needed in split deliveries from a subset of facilities, say $I_k \subseteq I$, which are considered to be close enough to get an offer from a parcel delivery operator. So 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 $k \in K_i$.

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 .

4.2.2 Mixed-integer programming formulation

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

In this formulation of , the objective function (4.1) 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 (4.2) enforce the demand satisfaction for all customer areas in each period. Constraints (4.3) ensure that the total demand allocated to a facility for the last mile-delivery does not exceed its selected long-haul capacity level. Constraints (4.4) establish that a customer area k can only be served from facilities belonging to carriers with ongoing contracts. (These constraints are actually redundant with (4.2)-(4.3), but they reinforce the linear re-

laxation of the formulation.) Constraints (4.5) and (4.6) 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 (that is, any contract signed in one of the periods from $t - H_e + 1$ to t). Constraints (4.7) express that a facility can be used only if there is a valid contract with its owner. Constraints (4.8) specify that two contracts cannot simultaneously be in effect with a same carrier. Note that in view of constraints (4.7) and (4.8), at most one capacity level can be selected for each facility. The formulation is strengthened by the covering constraints (4.9) which guarantee, that each customer area is served by at least one carrier. Finally, constraints (4.10)-(4.12) specify the range of the variables. We denote by Z_{opt} the optimal value of MDPC.

Objective function

$$\min Z = \sum_{t \in T} \left[\sum_{e \in E} c_e^t + \sum_{k \in K} \sum_{i \in I_k} U_{i,k} q_{i,k}^t \right] \quad (4.1)$$

Constraints

$$\sum_{i \in I_k} q_{i,k}^t = D_k^t \quad \forall k \in K, t \in T \quad (4.2)$$

$$\sum_{k \in K_i} q_{i,k}^t \leq \sum_{l \in L} Q_l v_{i,l}^t \quad \forall i \in I, t \in T \quad (4.3)$$

$$q_{i,k}^t \leq D_k^t \sum_{l \in L} v_{i,l}^t \quad \forall i \in I_k, k \in K, t \in T \quad (4.4)$$

$$c_e^t \geq \sum_{l \in L} \sum_{i \in I_e} F_{i,l} v_{i,l}^t \quad \forall e \in E, t \in T \quad (4.5)$$

$$c_e^t \geq M_e^t \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall e \in E, t \in T \quad (4.6)$$

$$\sum_{l \in L} v_{i,l}^t \leq \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall i \in I_e, e \in E, t \in T \quad (4.7)$$

$$\sum_{n=0}^{H_e-1} \alpha_e^{t-n} \leq 1 \quad \forall e \in E, t \in T \quad (4.8)$$

$$\sum_{e \in E_k} \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \geq 1 \quad \forall k \in K, t \in T \quad (4.9)$$

$$\alpha_e^t \in \{0, 1\} \quad \forall e \in E, t \in T \quad (4.10)$$

$$v_{i,l}^t \in \{0, 1\} \quad \forall i \in I, l \in L, t \in T \quad (4.11)$$

$$q_{i,k}^t \geq 0 \quad \forall k \in K, i \in I, t \in T \quad (4.12)$$

4.3 Complexity

4.3.1 NP-hardness

For certain values of the parameters, the MDPC model reduces to the well-known simple facility location problem (SPL; Ghiani et al. [2004]). Specifically, consider the special case of MDPC where

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

Under these conditions, the shipper has no choice but to contract with the unique carrier ($\alpha_e^1 = 1$), and MDPC boils down 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 $q_{i,k}$ so as to minimize the total cost of operating the facilities and of transporting the parcels: this is exactly the definition of the simple facility location problem. Since the SPL problem is NP-hard, so is MDPC.

The previous argument does not really 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 (4.1)-(4.12). An alternative argument, therefore, might consider instead the following special case of MDPC:

- there is a single period: $N = 1$;
- 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 $i \in I$;
- the contract duration is equal to one period: $H_e = 1$;
- there is only one operating level ($L = \{1\}$) and Q_1 is large enough to satisfy the total demand;
- 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 SPL, with the variables α_e^1 , $e \in E$, indicating the contracts to be signed in order to be allowed to operate the associated facilities.

4.3.2 Decomposability

From a practical point of view, the difficulty of MDPC mostly 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, and this 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 above-mentioned interrelationships, 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 (Benders [1962]; Rahmaniani et al. [2017]) can be useful for tackling the complex structure of the problem. An exact method based on this idea is presented in Section (4.4).

Similarly, when we relax the constraints (4.6)-(4.9), which bind consecutive periods over the duration of each contract, then the distribution sub-problem can be independently tackled for each period. Section 4.5 describes a heuristic algorithm based on this idea.

4.4 Combinatorial Benders Algorithm - (CBA)

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

4.4.1 Benders with integer sub-problems

The classical Benders procedure for mixed-integer linear programming requires to fix (iteratively) the value of all integer variables and to solve 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 Rahmani et al. [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.

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

$$\sum_{j \in C} y_j + \sum_{j \in D} (1 - y_j) \geq 1 \quad (4.13)$$

for some appropriate subsets C, D of indices.

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 nogood 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.

Authors in Fakhri et al. [2017] proposed a Benders decomposition method with integer sub-problems. Their work builds on ideas initially developed in Cordeau et al. [2000] for closing the duality gaps of ILP problems. Here, the Benders procedure is embedded in a branch-and-bound framework, and Benders cuts are inserted into the formulation at every node, in addition to branching constraints.

Related ideas were proposed to handle applications arising in a variety of fields. We only mention here a couple of examples. In Botton et al. [2013], the authors investigate a survivable network design problem for which they propose to relax the integer variables occurring in the Benders sub-problem. They characterize some cases where this procedure is exact. For the general case, reinforced logic-based cuts in the spirit of Hooker & Ottosson [2003] must in principle be generated, but

this turns out to be unnecessary in the computational experiments.

In Gendron et al. [2016], authors also rely on combinatorial cuts for an integer programming formulation of wireless LAN network design problem involving two subsets of binary variables, say (x, y) . In their approach, the master problem (MP) includes all the variables of the original problem, but the integrality constraints are relaxed on the x -variables. The Benders sub-problem is obtained by fixing the y -variables to their optimal value in (MP), and by reinstating the x -variables as integer variables. Canonical cuts of the form (4.13) are used to exclude previously visited solutions.

For our MDPC model, we develop a similar decomposition strategy, making use of combinatorial cuts that we add to the master problem in the course of the branch-and-cut optimization process. 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 algorithmic approach, which differs from the iterative classical procedure, is known as a *single-tree* Benders procedure [Rahmaniani et al., 2017].

4.4.2 Decomposition

The mathematical formulation (4.1)-(4.12) 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), which is obtained by relaxing the integrality constraints on the binary variables $v_{i,l}^t$, and into a sub-problem (SP), which is obtained by fixing the binary variables α_e^t to valid values $\hat{\alpha}_e^t$ (see Figure 4.2) derived from the optimal solution of the MP.

Every time the MP generates a new array of values $\hat{\alpha}_e^t$ (that is, a new tentative plan of contracts), an additional combinatorial cut is introduced into the MP formulation. The procedure stops when no new array of values $\hat{\alpha}_e^t$ is 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.

$\hat{\alpha}_e^t$	1	2							N
1	1	0	...	0					
2		1	0	...	0				
3	1	0							
E									

Contract duration (H_e)

Figure 4.2: Plan of contracts (valid values of $\hat{\alpha}_e^t$)

Master problem (MP)

The initial MP formulation is identical to the formulation (4.1)-(4.12), except that the integrality constraints (4.11) are relaxed between 0 and 1 and replaced by

$$v_{i,l}^t \in [0, 1] \quad \forall i \in I, l \in L, t \in T. \quad (4.14)$$

The relaxed variables $v_{i,l}^t$ are taken into account 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 will additionally be 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. We denote by $Z_R(\hat{\alpha})$ the optimal value of the master problem when α is fixed to $\hat{\alpha}$ (the subscript R reminds us that the $v_{i,l}^t$ variables are relaxed). If $Z_{inc} \leq Z_R(\hat{\alpha})$, then the usual principles of B&C guarantee that $\hat{\alpha}$ 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 by a combinatorial cut (see below). Otherwise, a sub-problem $SP(\hat{\alpha})$ must be solved to evaluate the quality of the contract plan modeled by $\hat{\alpha}$, as explained next.

Sub-problem (SP)

For each feasible assignment $\hat{\alpha}$, the associated sub-problem $\text{SP}(\hat{\alpha})$ is generated by setting α_e^t to the value $\hat{\alpha}_e^t$ in the formulation of MDPC, for all $e \in E$, $t \in T$. So, $\text{SP}(\hat{\alpha})$ is an MILP model with objective function (4.1), with binary variables $v_{i,l}^t$ and with continuous variables $q_{i,k}^t$, c_e^t . It includes constraints (4.2)-(4.5) and (4.11)-(4.12), as well as constraints (4.6) and (4.7) which respectively boil down to the following simple inequalities:

$$c_e^t \geq \begin{cases} M_e^t & \text{if } \sum_{n=0}^{H_e-1} \hat{\alpha}_e^{t-n} = 1 \\ 0 & \text{otherwise} \end{cases} \quad \forall e \in E, t \in T, \quad (4.15)$$

$$\sum_{l \in L} v_{i,l}^t \leq \begin{cases} 1 & \text{if } \sum_{n=0}^{H_e-1} \hat{\alpha}_e^{t-n} = 1 \\ 0 & \text{otherwise} \end{cases} \quad \forall i \in I_e, e \in E, t \in T. \quad (4.16)$$

Constraints (4.15) and (4.16) 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 (4.8)-(4.10) can be discarded from the sub-problem formulation. Note, however, that the presence of the covering constraints (4.9) in the master problem ensures that all customers can be served from the facility network available with the plan of contracts represented by $\hat{\alpha}$. In other words, the sub-problem $\text{SP}(\hat{\alpha})$ is always feasible.

As mentioned in Section 4.3.2, a crucial feature of the sub-problem $\text{SP}(\hat{\alpha})$ is that it is separable per period: indeed, for each $t \in T$, the optimal values (v^{t*}, q^{t*}, c^{t*}) only depend on the contracts $(\hat{\alpha}^t)$ that are active in period t . If we denote by $Z_{\text{SP}}(\hat{\alpha})$ the optimal value of $\text{SP}(\hat{\alpha})$, and by $Z_{\text{SP}}(\hat{\alpha}^t)$ the optimal value of the sub-problem $\text{SP}(\hat{\alpha}^t)$ arising in period t , then

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

This property facilitates the solving of $\text{SP}(\hat{\alpha})$, and actually motivates the whole decomposition scheme.

Combinatorial cuts

The optimal solution of $\text{SP}(\hat{\alpha})$ 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:

$$\sum_{t \in T} \left(\sum_{e \in E: \hat{\alpha}_e^t = 0} \alpha_e^t + \sum_{e \in E: \hat{\alpha}_e^t = 1} (1 - \alpha_e^t) \right) \geq 1. \quad (4.17)$$

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 (4.17) is strengthened by disregarding the values $\hat{\alpha}_e^t = 0$ associated with a contract in effect, but not taking place at time t . The values of these variables are already implied from constraints (4.8). Therefore, the expression $\sum_{n=0}^{H_e-1} \alpha_e^{t-n} = 0$ validates for 0-values when no contract is in effect at time t with carrier e . The strengthened cut (4.18) is, in all cases, included into the formulation.

$$\sum_{t \in T} \left(\sum_{e \in E: \sum_{n=0}^{H_e-1} \alpha_e^{t-n} = 0} \alpha_e^t + \sum_{e \in E: \hat{\alpha}_e^t = 1} (1 - \alpha_e^t) \right) \geq 1. \quad (4.18)$$

Note that even though $SP(\hat{\alpha})$ is time-separable, the combinatorial cut (4.18) cannot be split by period.

4.4.3 Algorithmic procedure

The overall combinatorial Benders algorithm (CBA) is sketched in Figure 4.3. Commercial solvers enable the user to interrupt the MP branch-and-cut process at various points to launch predefined routines for different purposes. In our algorithm, this practical tool is used to conveniently launch the $SP(\hat{\alpha})$ model whenever a feasible integer solution $\hat{\alpha}$ meet certain conditions, as described by Algorithm 1. The solution method for the SP is summarized below by Algorithm 2.

The addition of the combinatorial cut to the MP formulation rules out candidate $\hat{\alpha}$ from the feasible space, regardless if it is optimal for MDPC or not. The branch-and-cut procedure continues searching for the best feasible solution until no feasible candidates remain. 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 by CPLEX

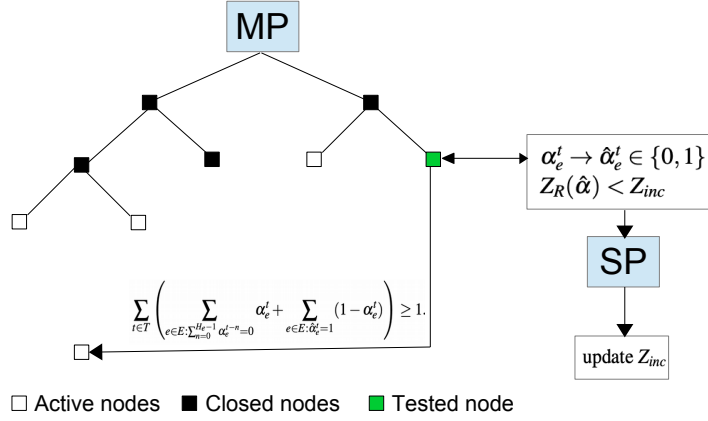


Figure 4.3: Single-tree Combinatorial Benders algorithm

is associated to the MP formulation. Therefore, it is never larger than the optimal value of the MP, which is in itself a LB for the MDPC model.

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 $\alpha \leftarrow \hat{\alpha}$
 - 2 **if** $Z_R(\hat{\alpha}) < Z_{inc}$ **then**
 - 3 **solve** sub-problem $SP(\hat{\alpha})$
 - 4 **retrieve** value $Z_{SP}(\hat{\alpha})$ {see Algorithm 2}
 - 5 **if** $Z_{SP}(\hat{\alpha}) < Z_{inc}$ **then**
 - 6 $Z_{inc} \leftarrow Z_{SP}(\hat{\alpha})$
 - 7 **add** combinatorial cut (4.18) to MP
-

The sub-problem SP to be solved by Algorithm 1 is an MILP, and its solution can take a significant amount of time. Moreover, it must be solved repeatedly. In order to facilitate its solution, as suggested in Section 4.4.2, we decompose it per period, and we sequentially solve the sub-problems $SP(\hat{\alpha}^t)$ for each $t \in T$. 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 $\bar{\alpha}$, aiming at delivering the required demand in each period, can involve

the same carriers in period t , even if these carriers differ in some other periods facing another demand. 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 $\hat{\alpha}^t$, for each period t , together with its optimal value in a storage set Ω^t . When the same assignment $\hat{\alpha}^t$ is found again in subsequent iterations, its value is retrieved from the storage set, and do not lead 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. $\hat{\alpha}^t$: Part of the contract plan assigned for period t .
 Ω^t : for each $t \in T$, 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      $\lfloor$  retrieve the optimal value  $Z_{SP}(\hat{\alpha}^t)$ 
5   else
6      $\lfloor$  solve  $SP(\hat{\alpha}^t)$ 
7      $\lfloor$  add  $\hat{\alpha}^t$  and  $Z_{SP}(\hat{\alpha}^t)$  to  $\Omega^t$ 
8    $\lfloor$   $Z_{SP}(\hat{\alpha}) \leftarrow Z_{SP}(\hat{\alpha}) + Z_{SP}(\hat{\alpha}^t)$ 
9 return  $Z_{SP}(\hat{\alpha})$ 

```

4.5 Relax-and-repair heuristic

In this section, we present a heuristic optimization algorithm based on a decomposition of MDPC along the time dimension. The heuristic proceeds in two phases, which respectively consist in solving a relaxation of the problem, then in “repairing” the obtained solution to make it feasible.

4.5.1 Phase 1 - Relaxation - (H1)

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 $e \in E$. An alternative 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 (4.1)-(4.12) by a single binary variable β_e^t , and add the constraints

$$\sum_{n=0}^{H_e-1} \alpha_e^{t-n} = \beta_e^t \quad \forall e \in E, t \in T \quad (4.19)$$

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 (4.19). 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 $H1$, and its optimal value as Z_{H1} . Clearly, Z_{H1} is a valid lower bound for MDPC: $Z_{H1} \leq Z_{opt}$.

The formulation $H1$ can be decomposed and solved independently for each period $t \in T$. Let $H1(t)$ define the single-period problem at period t and let $Z_{H1}(t)$ be its optimal value. Then, $Z_{H1} = \sum_{t \in T} Z_{H1}(t)$.

In order to avoid confusion, we stress that each sub-problem $H1(t)$ 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 4.4.2, 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 $H1(t)$, then, by definition, $Z_{SP}(\hat{\beta}^t) = Z_{H1}(t)$.

Algorithm 3: Relax-and-repair heuristic: Phase 1

Input: An instance of MDPC.

Output: Z_{H1} : Lower bound on the optimal value of MDPC.

$\hat{\beta}$: A feasible contract plan for the relaxation of MDPC.

```

1  $Z_{H1} \leftarrow 0$ 
2 for each  $t \in T$  do
3   solve  $H1(t)$ , obtain  $\hat{\beta}^t, Z_{H1}(t)$ 
4    $Z_{H1} \leftarrow Z_{H1} + Z_{H1}(t)$ 
5 return  $Z_{H1}, \hat{\beta} = (\hat{\beta}^t : t \in T)$ 

```

4.5.2 Phase 2 - Repairing - (H2)

Consider now an optimal solution $\hat{s} = (\hat{\beta}, \hat{v}, \hat{q}, \hat{c})$ of problem $H1$. With a slight abuse of terminology, we say that this solution is feasible for MDPC if the succession of single-period contracts defined by $(\hat{\beta}_e^t, t \in T)$ defines a collection of contracts with initial durations H_e not necessarily equal to a single period, 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 (4.19).) Note that this is very easy to check. When it happens to be 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 (4.9), that is, $Z_{H1} = Z_{opt}$, 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_{H1} 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 H2 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*. The incremental cost of each option ($\Delta CostC$, $\Delta CostR$) is estimated, and the least expensive one is implemented.

We next give a more formal description of the resulting algorithm with descriptive comments of in brackets.

The *Completion* option considers the number of periods needed for the unfinished contract to get completed (i.e, $H_e - r$), starting from time t . Next, the local variable *CurrentCostC* sums up the cost values $Z_{SP}(\hat{\beta}^t) + \dots + Z_{SP}(\hat{\beta}^{t+H_e-r-1})$ issued from the last iteration or Phase 1. Then, the algorithm sets temporarily the values $\hat{\beta}^t, \dots, \hat{\beta}^{t+H_e-r-1}$ to one (1) and computes again the costs values $Z_{SP}(\hat{\beta}^t), \dots, Z_{SP}(\hat{\beta}^{t+H_e-r-1})$ by solving the sub-problems with the modified values. The difference is saved with parameter $\Delta CostC$ and the values are reset back to zero (0).

The *Removal* option considers the number of periods before the interruption of the unfinished contract in time t , in order to get removed (i.e., r). Next, the local parameter *CurrentCostR* sums up the cost values $Z_{SP}(\hat{\beta}^{t-r}), \dots, Z_{SP}(\hat{\beta}^{t-1})$. Then, the algorithm sets temporarily the values $\hat{\beta}^{t-r}, \dots, \hat{\beta}^{t-1}$ to zero (0) and computes again the costs values $Z_{SP}(\hat{\beta}^{t-r}), \dots, Z_{SP}(\hat{\beta}^{t-1})$ by solving the sub-problem with the modified values. The difference is saved with parameter $\Delta CostR$. Finally the smallest

Algorithm 4: Relax-and-repair heuristic: Phase 2

Input: $\hat{\beta}$: A feasible contract plan for the relaxation of MDPC (H1)

Output: $\hat{\alpha}_{H2}$: A feasible contract plan. Z_{H2} : Value of the heuristic solution.

```
1  $Z_{H2} \leftarrow Z_{H1}$ 
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  $\hat{\beta}_e^{t-1} = 1$  {/* potentially incomplete contract } then
5        $s \leftarrow t$  {/* s: the closest period to  $t$  such that
6          $\hat{\beta}_e^s = \dots = \hat{\beta}_e^{t-1} = 1$ }
7       while  $s > 1$  and  $\hat{\beta}_e^{s-1} = 1$  do
8          $s \leftarrow s - 1$ 
9        $r \leftarrow (t - s) \bmod H_e$  {/* r: the elapsed duration of the
10        contract running at period  $t - 1$ }
11       if  $r \neq 0$  {/* the sequence  $\hat{\beta}_e^s, \dots, \hat{\beta}_e^t$  is not feasible} then
12          $CurrentCostC \leftarrow \sum_{k=0}^{H_e-r-1} Z_{SP}(\hat{\beta}^{t+k})$  {/* estimate extra
13          cost of completing the contract running at period  $t - 1$ }
14         temporarily set  $\hat{\beta}_e^t, \dots, \hat{\beta}_e^{t+H_e-r-1} \leftarrow 1$ 
15          $NewCostC \leftarrow \sum_{k=0}^{H_e-r-1} Z_{SP}(\hat{\beta}^{t+k})$ 
16          $\Delta CostC \leftarrow NewCostC - CurrentCostC$ 
17         reset the previous values of  $\hat{\beta}_e^t, \dots, \hat{\beta}_e^{t+H_e-r-1} \leftarrow 0$ 
18          $CurrentCostR \leftarrow \sum_{k=1}^r Z_{SP}(\hat{\beta}^{t-k})$  {/* estimate extra cost
19          of removing the contract running at period  $t - 1$ }
20         temporarily set  $\hat{\beta}_e^{t-r}, \dots, \hat{\beta}_e^{t-1} \leftarrow 0$ 
21          $NewCostR \leftarrow \sum_{k=1}^r Z_{SP}(\hat{\beta}^{t-k})$ 
22          $\Delta CostR \leftarrow NewCostR - CurrentCostR$ 
23         reset the previous values of  $\hat{\beta}_e^{t-r}, \dots, \hat{\beta}_e^{t-1} \leftarrow 1$ 
24         if  $\Delta CostC \leq \Delta CostR$  then
25           /* { the completion option is preferred} */
26            $\hat{\beta}_e^t, \dots, \hat{\beta}_e^{t+H_e-r-1} \leftarrow 1$ 
27            $Z_{H2} \leftarrow Z_{H2} + \Delta CostC$ 
28         else
29           /* {the removal option is preferred} */
30            $\hat{\beta}_e^{t-r}, \dots, \hat{\beta}_e^{t-1} \leftarrow 0$ 
31            $Z_{H2} \leftarrow Z_{H2} + \Delta CostR$ 
32
33 26 return  $Z_{H2}, \hat{\alpha}_{H2}$  {/* contracts defined by  $\hat{\beta}$ }
```

difference between $\Delta CostC$ and $\Delta CostR$ determines whether to complete or remove the contract.

It may be interesting to note that in a "real-world" setting, the heuristic could be implemented in a rolling-horizon fashion. More precisely, 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 $t, t + 1, \dots, t + \max\{H_e : e \in E\}$. Thus, the heuristic can be implemented with a forecast horizon of $\max\{H_e : e \in E\}$ periods.

4.6 Experimental design and performance measures

We have conducted a series of experiments with different algorithms on multiple sets of instances. This section first describes the set of instances that we generated, then the algorithms, and finally the performance measures that we consider.

4.6.1 Instances

Data about the location of facilities and of customers, cost components, demand forecast and contract terms were generated randomly according to the following assumptions and parameter choices.

Location and scope of sites

- The shipper operates two warehouses.
- Cross-docking facilities and customer areas are randomly located in a two-dimensional grid (1000×1000).
- The carriers' facilities are randomly located in a centered area of the grid (700×700).
- For each carrier e , the size of I_e is randomly chosen from the discrete uniform distribution on an interval $[i^-, i^+]$: $U[i^-, i^+]$ (see Table 4.2).
- Customers can only be served from carriers' facilities that are within an Euclidean distance of 500 units. No carrier may 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 $U[500, 1000]$.
- The variable cost $T_{i,l}$ for long-haul transportation is modeled by a staircase function with decreasing marginal costs. The cost for operating at level one, i.e., $T_{i,1}$, is two times the Euclidean distance from facility i to the closest warehouse. When operating at level l , a total discount of $(l - 1)$ percent is applied to the cost $l \times T_{i,1}$.
- 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 middle 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 carriers $e \in E$.
- Each carrier has enough available capacity to fulfill the demand of all the 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 driving the size of 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 4.2. Moreover, we are also interested in analyzing the effect of the contract duration on the hardness of 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 randomized instances of the same size are generated, for a grand total of $3 \times 6 \times 3 \times 5 = 270$ instances.

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					

Table 4.2: Description of instance classes

4.6.2 Algorithms

The objective of the computational study is to evaluate the performance of different algorithms for solving instances of the MDPC model. The algorithms under consideration are those described in Sections 4.4 and 4.5, but as benchmarks, we also tested two additional 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 $(\alpha_e^t, v_{i,l}^t)$, 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 4.4);

- (H1 - H2) - Relax-and-repair heuristic (Section 4.5).

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).

4.6.3 Performance measures

Each algorithm is run for at most 3600 seconds on each of the 270 instances. If an algorithm terminates sooner, the computational time to completion and the optimal value are recorded (all times are expressed in seconds); otherwise, performance parameters like the best found solution value, the time to obtain it and the best lower bound are recorded.

More precisely, the algorithms are compared based on two performance attributes:

- The *efficiency* of each algorithm A is measured by its total running time, and by its running time until it obtains its best overall solution (within the time limit):
 1. $F^A \rightarrow$ Running time up to termination of algorithm A .
 2. $B^A \rightarrow$ Running time of algorithm A until it obtains its best solution.
- The *effectiveness* of algorithm A is measured 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 any of our algorithms. Then:
 1. $G^A = \frac{Q^A - L^A}{Q^A} (\times 100\%) \rightarrow$ Relative optimality gap obtained solely by algorithm A .
 2. $\Delta Q^A = \frac{Q^A - Q^{best}}{Q^A} (\times 100\%) \rightarrow$ Relative gap of algorithm 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 or if H1 found a feasible solution, then the best found solution is optimal and $G^A = \Delta Q^A = 0$.

4.7 Computational results

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

4.7.1 Small size instances

A first overview of the results for the set of 90 small instances is displayed in Figure 4.4. 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 Figure 4.4(b), marked “H1”, 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 H2 (see Section 4.5). 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 not find the best solution, the average relative difference ΔQ^A to the best-known value is displayed in parentheses.

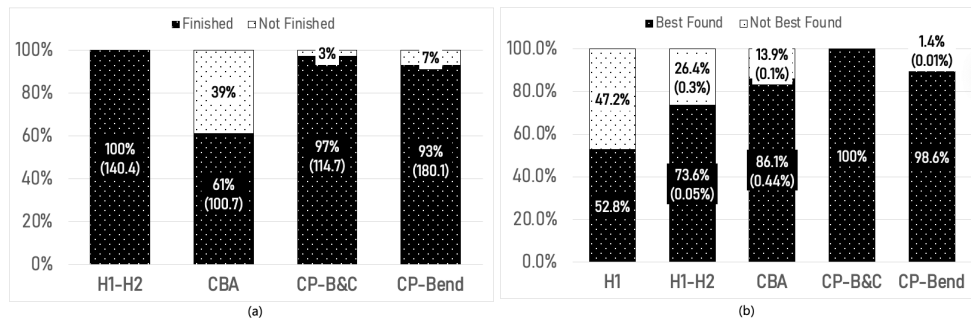


Figure 4.4: Performance of algorithms on small instances

Another look at the results is provided in Figure 4.5 and Figure 4.6, 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).

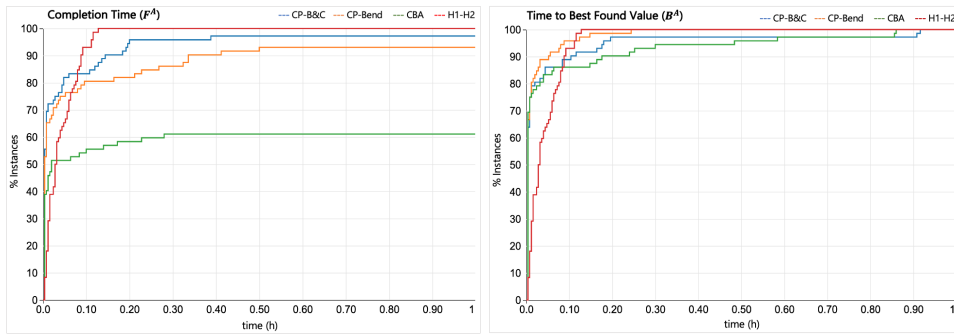


Figure 4.5: Efficiency: performance profile of total running time F^A and of running time until best found solution B^A - small instances

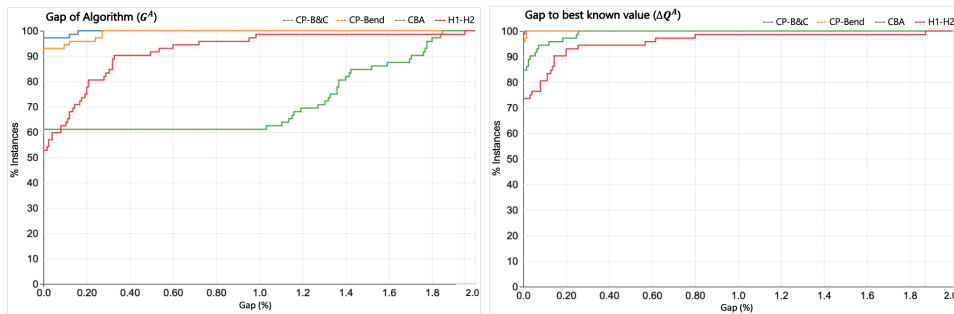


Figure 4.6: Effectiveness: performance profile of optimality gap G^A and of gap to best known value ΔQ^A - small instances

The relax-and-repair heuristic generates the optimal solution for approximately 53% of the instances in its Phase 1 H1, by relaxing the contract-duration constraints alone. Moreover, this percentage climbs to 73% after repairing infeasibilities in Phase 2 H2 (Figure 4.4(b) and Figure 4.6). For the remaining instances, the heuristic provides suboptimal solutions with a relative gap (G^{H2} or ΔQ^{H2}) smaller than 1% for most of the small instances (Figure 4.6).

CPLEX methods prove optimality and terminate in less than one hour for almost all small instances (97% for B&C, 93% for CPLEX Benders), but they are naturally slower than the heuristic H1-H2 (Figure 4.5). Finally, the combinatorial Benders algorithm (CBA) terminates in 61% of the cases only (Figure 4.4(a) and Figure 4.5), although it generally finds the optimal value within the time limit (for 86% of the instances, see Figure 4.4(b) and Figure 4.6), in about the same time as both CPLEX methods. Its main weakness resides in slightly weaker lower bounds which do not completely close the optimality gap, as evidenced by the value of the optimality gap: $G^{CBA} = 0.44\%$ on average when CBA finds the optimal solution (see also Figure 4.6). It is interesting to observe, however, that the performance profile of ΔQ^A lies lower for the heuristic H2 than for CBA, which means that CBA is generally able to find better feasible solutions.

Algorithmic performance per instance size

Figure 4.7 and Figure 4.8 present more detailed information regarding the algorithmic performance as a function of size parameters, as described in Table 4.2. For every sample of 5 instances in each class of small instances, the average value of F^A , B^A , G^A , and ΔQ^A is calculated and reported. In Figure 4.5, the white bar shows the total completion time (F^A), while the internal black bar is the time until the best solution is found (B^A).

In terms of efficiency, the heuristic method is fast in obtaining a feasible solution, and its running time grows slowly with respect to size increments (Figure 4.7(a)). The CPLEX methods (Figure 4.7(c,d)) are able to achieve the optimal value in relatively short time, but nonetheless, their completion time tends to increase rapidly with the size of the instances. Whereas instances with 6 carriers and 6 periods are solved in less than 8 (respectively, 10) seconds by CP-B&C (respectively, CP-Bend), the average running time increases to 202 (respectively, 496) seconds when the horizon is expanded to 8 periods, and to 1134 (respectively, 1338) seconds when 8 carriers and 10 periods are considered. This increasing trend is even more pronounced for the CBA algorithm (Figure 4.7(b)), although the running time of CBA until the best solution gets found is considerably shorter than its completion time, as we already observed.

In Figure 4.8, the white bar depicts the optimality gap derived independently by each method (G^A), whereas the internal black bar represents the relative difference to the best found value (ΔQ^A). Note again that CPLEX actually provides the optimal value for most of the instances. If we compare the two tailor-made algorithms, we see that

the gap G^A is smaller on average for the heuristic than for CBA: 0.04%, 0.34% and 0.53% vs. 1.48%, 1.55% and 1.25% for the three largest parameter values, respectively. In contrast with this, the relative difference ΔQ^A to the best found solution is smaller for CBA (0.00%, 0.02% and 0.10% for the largest parameter values) than for the heuristic (0.014%, 0.24% and 0.38%, respectively), meaning that CBA generally finds better solutions.

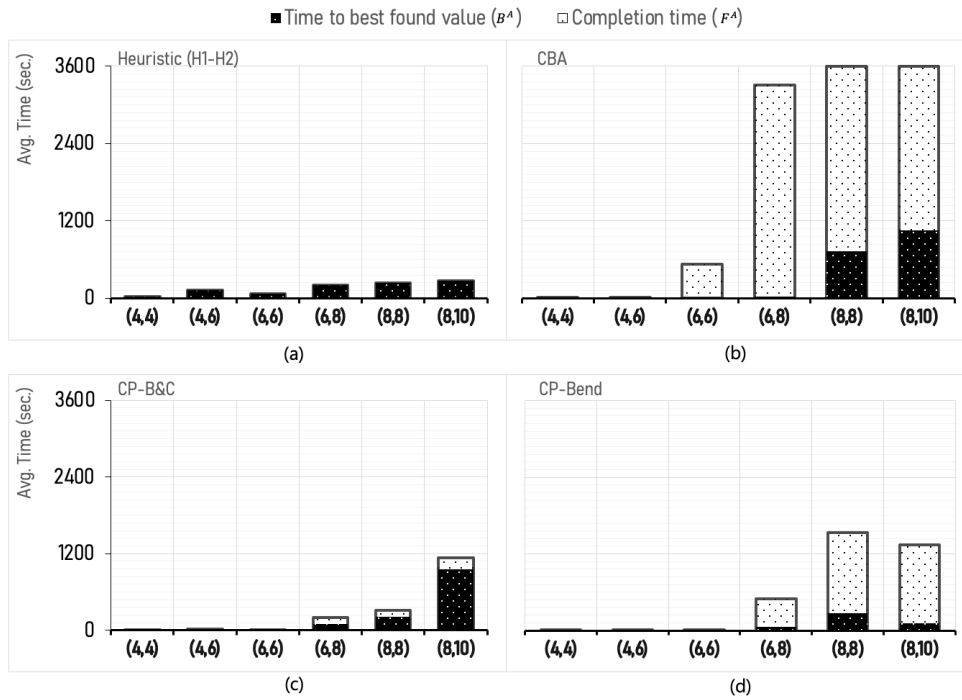


Figure 4.7: Efficiency (total computational time F^A and running time B^A until best solution) of algorithms as a function of parameter values for small instances

(E , T)	(4,4)		(4,6)		(6,6)		(6,8)		(8,8)		(8,10)	
	F^A	B^A	F^A	B^A	F^A	B^A	F^A	B^A	F^A	B^A	F^A	B^A
H1-H2	30	30	133	133	75	75	207	207	242	242	268	268
CBA	3	1	16	3	528	5	3306	22	3600	733	3600	1057
CP-B&C	1	1	15	12	8	6	202	95	308	212	1134	958
CP-Bend	1	1	11	7	10	7	496	62	1532	277	1338	111

Table 4.3: Numeric values for efficiency measures (F^A , B^A) in small instances

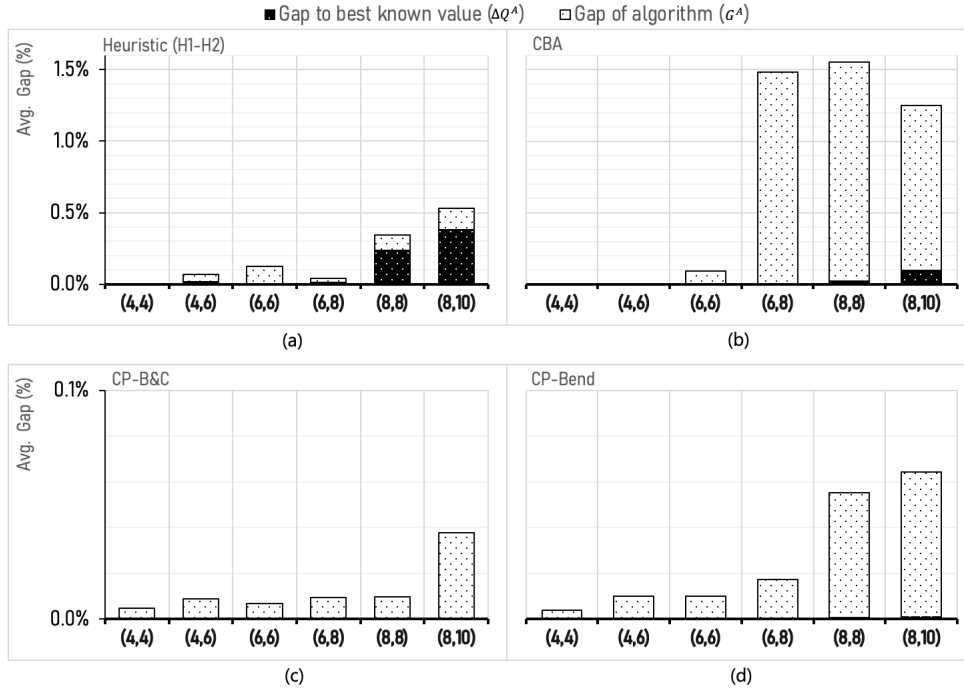


Figure 4.8: Effectiveness (optimality gap G^A and gap ΔQ^A to best known value) of algorithms as a function of parameter values for small instances

(E , T)	(4,4)		(4,6)		(6,6)		(6,8)		(8,8)		(8,10)	
Algorithm	G^A	ΔQ^A	G^A	ΔQ^A	G^A	ΔQ^A	G^A	ΔQ^A	G^A	ΔQ^A	G^A	ΔQ^A
H1-H2	0%	0%	0.07%	0.02%	0.12%	0%	0.04%	0.01%	0.34%	0.24%	0.53%	0.38%
CBA	0%	0%	0%	0%	0.1%	0%	1.48%	0%	1.55%	0.02%	1.25%	0.1%
CP-B&C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.04%	0%
CP-Bend	0%	0%	0%	0%	0%	0%	0.02%	0%	0.06%	0%	0.06%	0%

Table 4.4: Numeric values for effectiveness measures (G^A , ΔQ^A) in small instances

4.7.2 Medium size instances

Figures 4.9, 4.10 and 4.11 present the aggregated results for the set of medium size instances, with the same conventions as in the previous section.

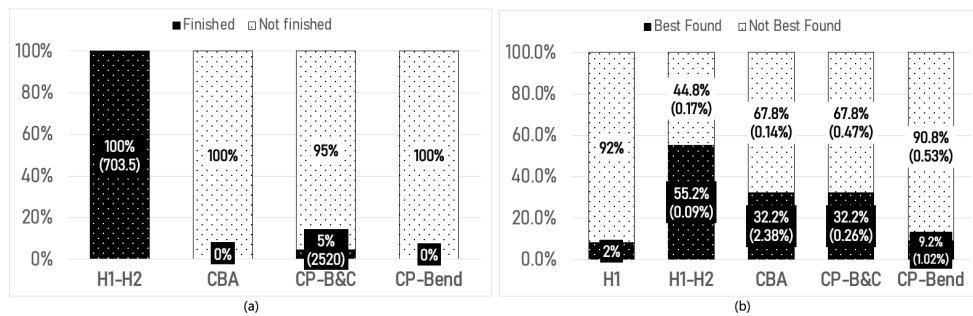


Figure 4.9: Performance of algorithms on medium instances

For most of the instances, none of the exact methods is able to terminate, and hence to prove optimality, within one hour. The heuristic method always terminates before the time limit and for 55.2% of the instances, it obtains a solution that is at least as good as the one given by the exact methods and even exact for 8% of the instances solved by H1. Among exact methods, the combinatorial algorithm CBA provides the best feasible solution faster than either of the CPLEX methods (performance profile B^{CBA} in Figure 4.10). 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% (performance profile G^{CBA} in Figure 4.11).

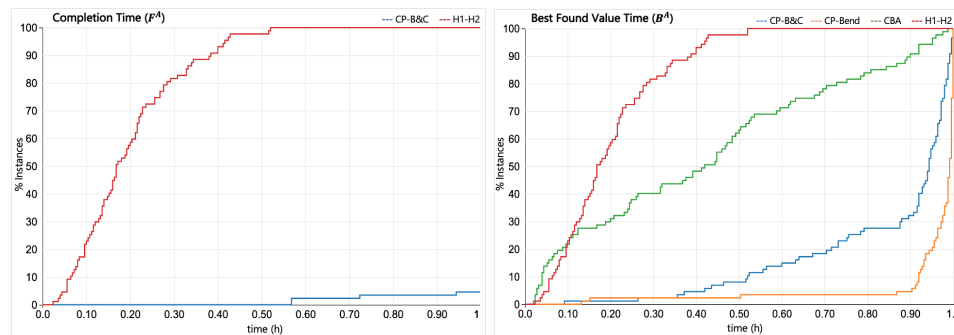


Figure 4.10: Efficiency: performance profile of total running time F^A and of running time until best found solution B^A - medium instances

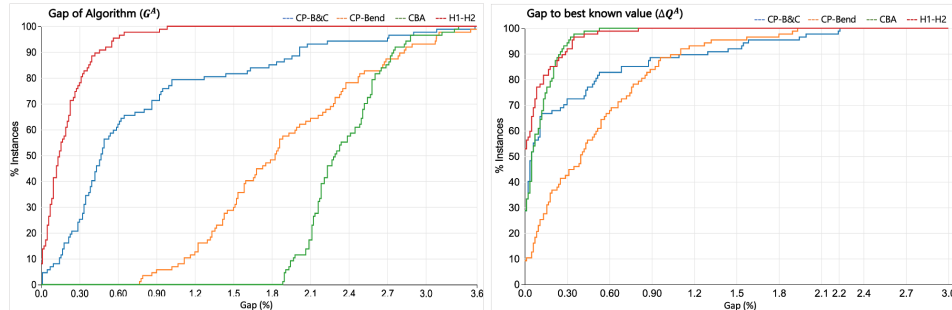


Figure 4.11: Effectiveness: performance profile of optimality gap G^A and of gap to best known value ΔQ^A - medium instances

Algorithmic performance per instance size

A more detailed presentation of the results is shown in Figures 4.12 and 4.13 for different size parameters and performance criteria. The heuristic method computes a feasible solution rather fast (less than 1300 seconds), and maintains a slow increase in running time as the instance parameters (E, T) grow larger (Figure 4.12(a)). The CBA algorithm identifies its best feasible solution early in the search process, almost independently of the parameters' size black bar in Figure 4.12(b)). Comparing to the time CPLEX's methods finds their best found solution (black bar in Figures 4.12(c,d)), it appears the CBA achieves its best found solution (within 1 hour) in less time. In fact, CPLEX's methods tend to improve continuously their incumbent solution, closing slowly the gap with their respective LBs.

In terms of solution quality, we observe that the heuristic method has a very small integrality gap, below 0.29% on average for all values of the size parameters. This means, in particular, that the lower bound computed in the first phase of the heuristic is always very tight.

On the other hand, the optimality gap G^A derived by the CBA method fluctuates around 2.5% for all medium size instances (Figure 4.13(b)). The small value of ΔQ^A for this algorithm (at most 0.15%) suggests, however, that the solution it returns is usually close to optimal.

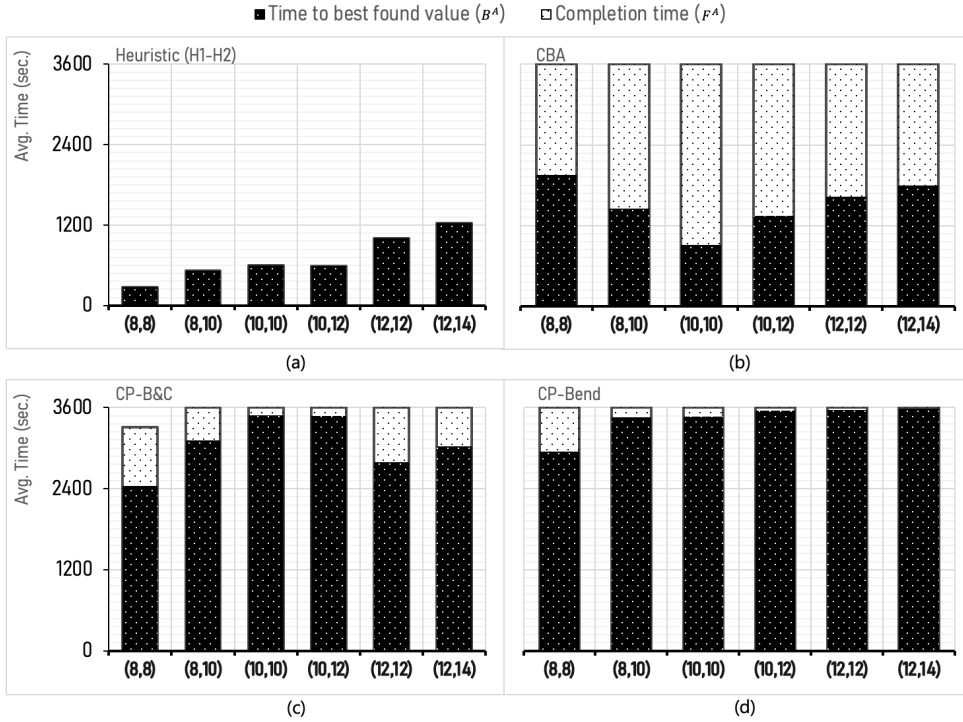


Figure 4.12: Efficiency (total computational time F^A and running time B^A until best solution) of algorithms as a function of parameter values for medium instances

(E , T)	(8,8)		(8,10)		(10,10)		(10,12)		(12,12)		(12,14)	
Algorithm	F^A	B^A	F^A	B^A	F^A	B^A	F^A	B^A	F^A	B^A	F^A	B^A
H1-H2	276	276	523	523	602	602	594	594	1007	1007	1231	1231
CBA	3600	1949	3600	1454	3600	915	3600	1340	3600	1625	3600	1794
CP-B&C	3313	2436	3600	3111	3600	3483	3600	3479	3600	2789	3600	3029
CP-Bend	3600	2949	3600	3458	3600	3469	3600	3552	3600	3567	3600	3582

Table 4.5: Numeric values for efficiency measures (F^A , B^A) in medium instances

In contrast with CBA, the optimality gap computed by CPLEX methods tends to increase along with the instance size. It even attains a larger value than CBA's gap for the largest instance size $(|E|, |T|) = (12, 14)$ (Figure 4.13(c,d)), a trend that will be confirmed in Section 4.7.3. CP-B&C, in particular, fails to find a good feasible solution, reason why the average relative gap ΔQ^A steadily increases with the instance size.

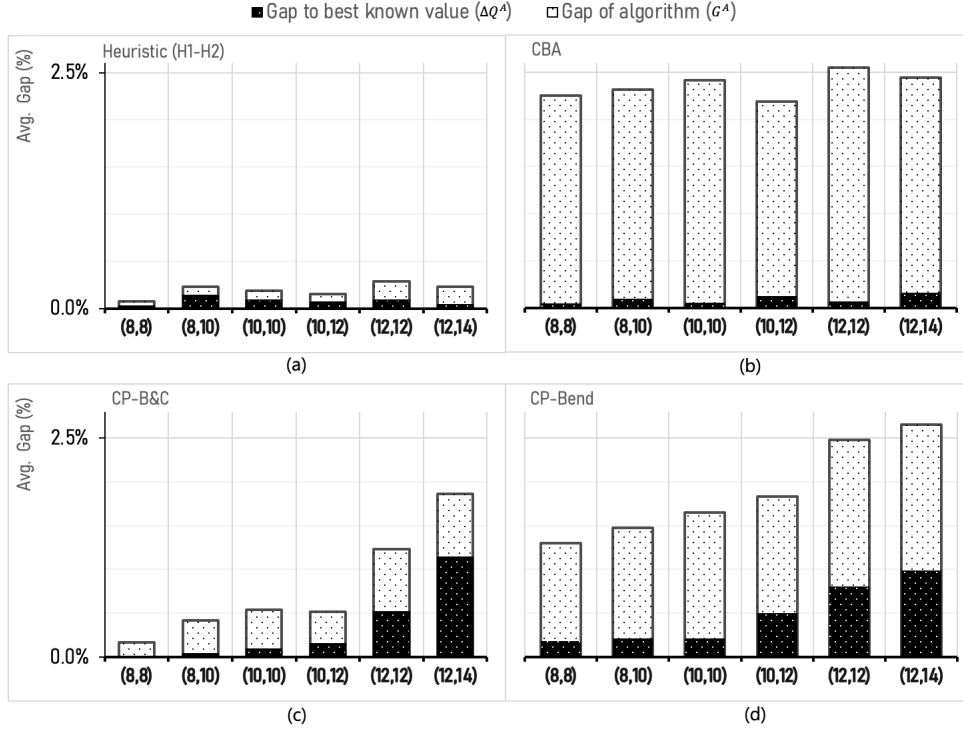


Figure 4.13: Effectiveness (optimality gap G^A and gap ΔQ^A to best known value) of algorithms as a function of parameter values for medium instances

(E , T)	(8,8)		(8,10)		(10,10)		(10,12)		(12,12)		(12,14)	
Algorithm	G^A	ΔQ^A	G^A	ΔQ^A	G^A	ΔQ^A	G^A	ΔQ^A	G^A	ΔQ^A	G^A	ΔQ^A
H1-H2	0.08%	0.03%	0.23%	0.14%	0.19%	0.09%	0.15%	0.07%	0.29%	0.09%	0.24%	0.04%
CBA	2.25%	0.05%	2.32%	0.10%	2.42%	0.05%	2.19%	0.13%	2.55%	0.07%	2.44%	0.16%
CP-B&C	0.17%	0.01%	0.42%	0.04%	0.54%	0.09%	0.52%	0.15%	1.23%	0.52%	1.86%	1.14%
CP-Bend	1.3%	0.19%	1.47%	0.21%	1.65%	0.21%	1.83%	0.5%	2.48%	0.8%	2.65%	0.99%

Table 4.6: Numeric values for effectiveness measures (G^A , ΔQ^A) in medium instances

4.7.3 Large size instances

The results of our tests on instances of large size are shown in Figures 4.14, 4.15 and 4.16. The trends observed for medium size instances are again present and accentuated here. Even though its completion time increases consistently, the heuristic method H2 is the only one for which instances (93%) terminates within one hour of running time, and it delivers the best solution for more than 90% of the instances.

The value of the heuristic solution is equal to the lower bound computed in the first phase for 4.4% of the large instances only, but the optimality gap G^{H2} is smaller than for any other method, and actually always smaller than 0.5%.

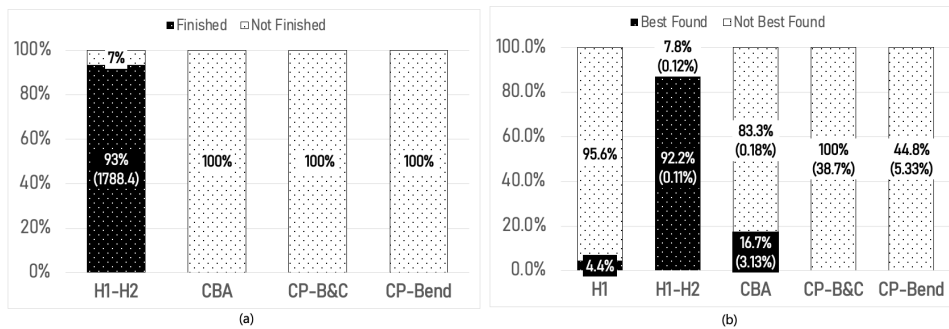


Figure 4.14: Performance of algorithms on large instances

Only our Combinatorial Benders Algorithm is sometimes able to match or to improve the solution found by H2 (for about 16% of the large instances). Moreover, interestingly, CBA usually finds it best solution faster than any other method, including the heuristic.

When applied to large size instances, CPLEX branch-and-cut frequently fails to detect a single feasible solution, as opposed to the classical Benders decomposition which is always able to find one. However, the lower bound provided by CPLEX B&C is tighter than by CPLEX-Benders.

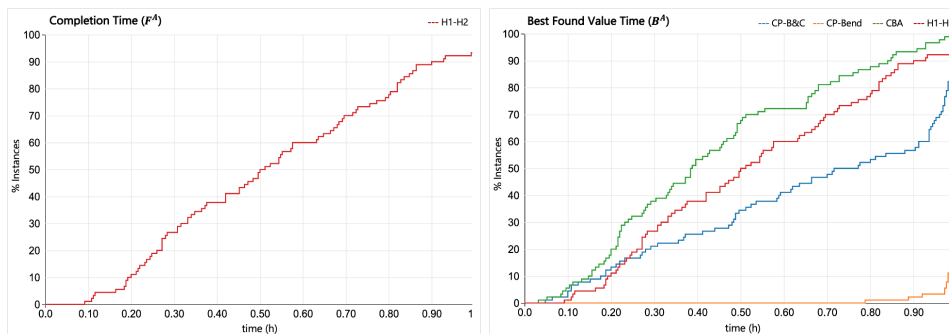


Figure 4.15: Efficiency: performance profile of total running time F^A and of running time until best found solution B^A - large instances

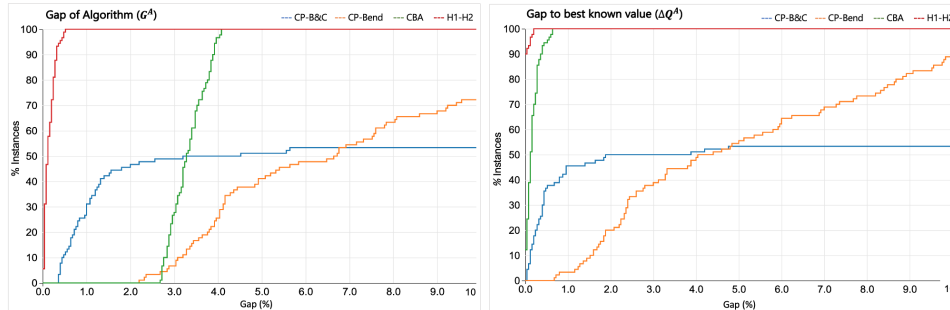


Figure 4.16: Effectiveness: performance profile of optimality gap G^A and of gap to best known value ΔQ^A - large instances

Algorithmic performance per instance size

Figure 4.17 and Figure 4.18 displays more detailed information. For these large instances, it is actually important to jointly examine the running time of the algorithms and the quality of the solutions that they produce.

The heuristic method takes on average less than 50 minutes (3000 seconds) to solve even the largest instances involving 12 carriers and 14 periods. This includes the time to solve the contract-duration relaxation and to perform the second-phase repair mechanism. The computational time is seemingly more affected by the number of carriers than by the number of periods. The cost of the solutions provided by the heuristic comes within 0.22% of the optimal cost in all cases (0.11% on average), and is always extremely close to the cost of the best solution found by any other method. These observations remain valid for all large instances.

As we already noted, the CBA algorithm reaches its best feasible solution faster than the heuristic (within 1900 seconds in all cases), especially for larger instance sizes. Its optimality gap G^{CBA} lies between 3% and 4%, but the relative gap ΔQ^{CBA} to the best found value is much smaller, around 0.15% on average, and smaller than 0.28% for all subclasses.

Regarding the CPLEX methods, CP-B&C takes a lot of time to obtain what turns out to be a poor feasible solution for the largest instance sizes. As the size of parameters increases, CP-B&C reaches a point where it is unable to find a good feasible solution, in spite of a decent lower bound obtained at the root node (Figure 4.18(c)).

In the course of its iterations, CP-Benders keeps improving its best incumbent solution until the time limit is reached. In general, however, it eventually returns (after

3600 seconds) a feasible solution which is far from the best known one, and the value of both gaps $G^{CP-Bend}$ and $\Delta Q^{CP-Bend}$ steadily increases with the instance size.

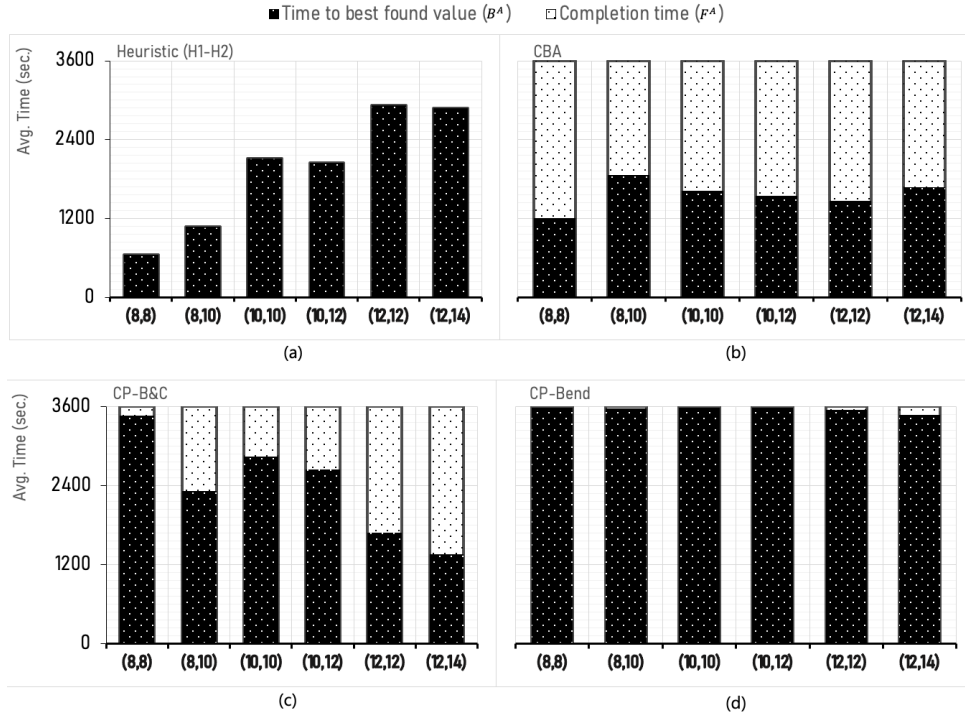


Figure 4.17: Efficiency (total computational time F^A and running time B^A until best solution) of algorithms as a function of parameter values for large instances

(E , T)	(8,8)		(8,10)		(10,10)		(10,12)		(12,12)		(12,14)	
Algorithm	F^A	B^A	F^A	B^A	F^A	B^A	F^A	B^A	F^A	B^A	F^A	B^A
H1-H2	646	646	1080	1080	2110	2110	2045	2045	2918	2918	2881	2881
CBA	3600	1204	3600	1867	3600	1623	3600	1544	3600	1466	3600	1676
CP-B&C	3600	3463	3600	2319	3600	2847	3600	2651	3600	1688	3600	1360
CP-Bend	3600	3581	3600	3577	3600	3578	3600	3583	3600	3554	3600	3473

Table 4.7: Numeric values for efficiency measures (F^A, B^A) in large instances

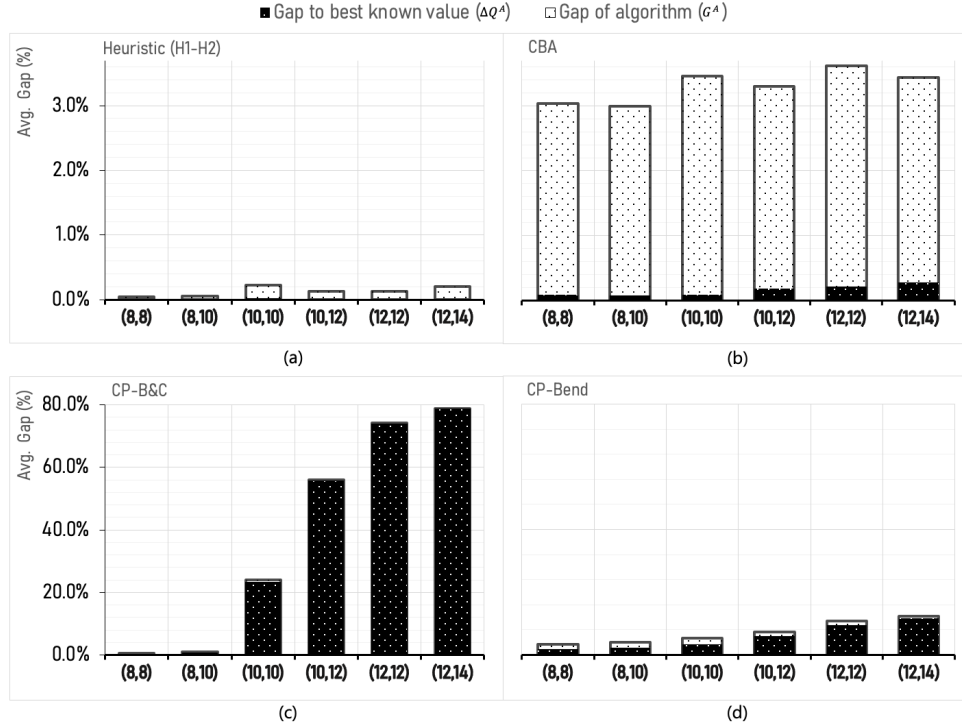


Figure 4.18: Effectiveness (optimality gap G^A and gap ΔQ^A to best known value) of algorithms as a function of parameter values for large instances

(E , T)	(8,8)		(8,10)		(10,10)		(10,12)		(12,12)		(12,14)	
Algorithm	G^A	ΔQ^A	G^A	ΔQ^A	G^A	ΔQ^A	G^A	ΔQ^A	G^A	ΔQ^A	G^A	ΔQ^A
H1-H2	0.04%	0%	0.06%	0.01%	0.22%	0.02%	0.13%	0%	0.13%	0.01%	0.2%	0.01%
CBA	3.04%	0.08%	2.99%	0.08%	3.46%	0.08%	3.30%	0.19%	3.62%	0.21%	3.43%	0.28%
CP-B&C	0.49%	0.13%	1.11%	0.65%	23.92%	23.25%	55.97%	55.63%	74.16%	73.93%	78.74%	78.6%
CP-Bend	3.37%	1.84%	3.93%	2.26%	5.32%	3.23%	7.41%	5.6%	10.75%	8.7%	12.30%	10.36%

Table 4.8: Numeric values for effectiveness measures (G^A , ΔQ^A) in large instances

4.7.4 Variations of contract duration

In this section, we examine the behavior of the computational time with respect to variations in the contract duration ($H_e = 2, 3, 4$) for the subclass of small instances, in which case it is possible to visualize the effect on the completion time (most of medium and large size instances cannot be solved within the allocated time limit and hence are not suited for this analysis). The contracts establish the only linkage among different periods (via constraints (4.7)-(4.9) of the MILP formulation) and

hence, at first sight, they represent a main complexifying component of the MDPC model. On the other hand, some of the algorithmic methods are based on a decomposition of the problem per time period, and it is unclear if the efficiency of these algorithms is considerably affected by the value of H_e .

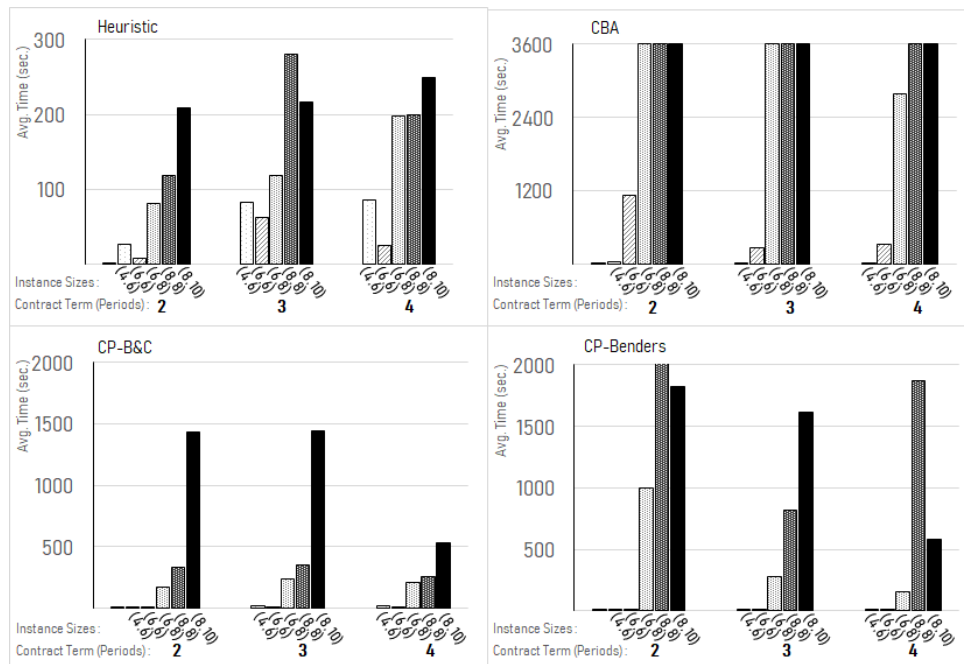


Figure 4.19: Variation of the computational time with respect to the contract duration

The heuristic algorithm, 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 . It could be argued that, when H_e gets larger, the first phase H1 is more likely to result in infeasible contract plans, so that the repair mechanism H2 has to 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 period, and hence the decision space is reduced when H_e increases. The overall result of these effects is hard to predict.

Experimentally, the multi-chart displayed in Figure 4.19(a) does not show any con-

sistent trend. The heuristic tends to run longer when H_e increases from 2 to 3 periods, for all instance sizes. But the differences are less noticeable when H_e is extended from 3 to 4 periods, and for some instance sizes (2 out of 5) the computational time even decreases.

The trend is also unclear for the combinatorial Benders algorithm (CBA), because it fails to terminate for most of the instances under consideration. For those instances that are solved to optimality, the running time decreases when H_e varies from 2 to 4 periods.

A similar observation applies for the performance of the CPLEX branch-and-cut method: there is little change in the computational time when H_e increases, except in the case of the largest instance size (8, 10) for which it is significantly reduced. On the other hand, with CPLEX Benders, the average computational effort shows a clearer trend to diminish when H_e increases, especially when extending the contract duration from 2 to 3 periods.

4.8 Conclusions

In this study we proposed a mathematical formulation to model the distribution network design problem under minimum-purchase commitment. The multi-period formulation is proven to be challenging for state-of-the-art algorithms when datasets increase in size, particularly if the decision-maker considers multiple carriers and separates the horizon in several periods. Aiming at producing satisfactory solutions for larger instances, we developed a solution methodology taking advantages of the decomposable structure of the mathematical model. Our developed algorithms were tested in several instances accounting for size variations and compared with CPLEX algorithms in two main performance criteria.

The exact solution methodology makes use of combinatorial cuts added while solving a relaxed version of the model by branch and cut. Compared with CPLEX methods, our exact methodology stays behind in terms of the resolution time for *small* instances. However, for larger instances it produces more satisfactory results overall, with better quality solutions in less computational time. Therefore, outperforming CPLEX in both, branch-and-cut and Benders procedures. It proves the usefulness of our algorithmic approach, which could be further developed and enriched with additional strategies, like our heuristic method for enhancing weak lower bounds. Also, it shows a difficulty faced by CPLEX in realizing the decomposable struc-

ture of a multi-period problem. A separate analysis can be done from the classical Benders procedure, where the decomposition in integer and continuous variables do not seem to deliver satisfactory results from *medium* size instances on. Instead, the combinatorial Benders decomposition with a mixed-integer sub-problem is able to obtain better results.

In general, the heuristic methodology from medium size to larger instances was proven not only more efficient by terminating in less time but also producing better quality solutions than any other exact method in one hour. This demonstrates the power of the relaxation which produces tight lower bounds to the objective function and the effectiveness of the simple mechanism of repairing an infeasible solution that violates contract-term constraints. An even better solution might be obtained by including local search algorithms. It was revealed that the contract-term parameter, which enlarge constraints linking several periods, produce little to no effect applying branch-and-cut, while it generates a positive effect applying decomposition methods and is more likely to increase the heuristics' solving time.

In the following Chapters 5 and 6, we examine a different type of capacity reservation mechanisms beyond the purchase commitment requirement expressed in capacity terms. The minimum capacity commitment is motivated from quantity flexibility contracts. These are based on reserving some resource capacity (e.g., a number of truckloads), while letting additional capacity to be charged at a higher cost. Such flexible options become particularly meaningful when the environment features significant uncertainty with regard to demand or cost parameters. Stochastic models that take uncertainty into account are also part of the analysis in next Chapters.

Chapter 5

Distribution Network Problems with Stochastic Demand and Minimum Capacity Commitment Contracts

5.1 Introduction

In Chapter 3, distribution network design problems arising in the context of e-commerce logistics were described and modelled in the presence of minimum purchase commitment (MPC) contracts and under the assumption of deterministic data. The use of MPC contracts grants access to transportation and distribution (T&D) services within a carrier's network, for a limited time. The shipper commits to paying at least a pre-agreed amount in exchange for transportation and distribution services at contract prices. However, the MPC contract does not specify mandatory use (take-it policy) of resources/capacity made available by the initial payment. Therefore, the shipper may decide to disregard partially or totally available capacity in periods where it does not require it, at his convenience.

In this chapter, we consider contract arrangements between the shipper and carriers which enforce a minimum commitment in terms of transportation and facility (throughput) *capacity* to be used in exchange for attractive contract costs, presumably lower and more stable than those offered by alternative sources like spot mar-

kets. This implies that a minimum amount of logistics services, expressed in load/cargo units, are requested along the contract duration and indirectly guarantee to the carrier a minimum payment for the services provided. By specifying a minimum utilization of capacity, the carrier has a better estimation of resources to reserve in order to fulfill the shipper's minimum requirements than in the case of minimum purchase commitment expressed in monetary units. This class of contracts are denominated *minimum capacity commitment (MCC)* contracts.

In addition, the present chapter addresses the uncertainty of customer demand. Indeed, with the recent surge of e-commerce and fast dynamic evolution of demand, it becomes increasingly difficult to predict with reasonable certainty the demand for goods and products purchased on-line. Although variations along the year might be attributed to seasonal effects and predicted to some extent, the uncertainty regarding individual decisions of consumers is more difficult to capture. In this work, we assume that the aggregated demand of customer areas in successive periods can be modelled as discrete random variables, which give rise to random scenarios over the course of the horizon.

In the deterministic case, the e-commerce fulfilment company is certain of the total T&D capacity required to satisfy customer orders in a certain period, regardless of the carriers used to that end. Once a network of contract carriers is established, it is possible to determine the least-cost distribution plan and to request T&D services accordingly. When dealing with stochastic demand, however, the shipper faces additional challenges in order to completely satisfy the demand while attempting to minimize its total costs. The total capacity required might differ depending on the observed scenario and the least-cost distribution plan varies accordingly.

Additional assumptions will have to be explicitly specified regarding the restrictions placed by carriers on the timing of the preparation and allocation of T&D resources. In some cases, the carriers may be able to fulfill the shipper's requests immediately (i.e, without lead time) after they are stated; in other cases, the carriers may need sufficient time in order to respond effectively to the shipper's requests. Therefore, the shipper may or may not be required to communicate to the carrier the capacity needed along specific lanes or at specific facilities early in the planning process. These additional assumptions can be incorporated into the business agreement, as part of the contractual conditions. Naturally, this affects the sequence of the decision-making process. Both cases are considered further in this chapter in the development of *two-stage stochastic programming* formulations of *distribution network design problems with minimum commitment contracts (DNPMC)*. They will give rise to different models where decisions will be separated according to whether

they are made before or after random demand is observed.

In our framework, the optimization of DNPMC models is not aimed at generating the optimal parameters for MCC contracts, but instead it helps to find the most suitable selection from a set of different alternatives that were previously conceived. This chapter is restricted to modelling and to formulating different situations which may be encountered in practical applications. The sequence of the contents is as follows. In Section 5.2, the family of contracts labelled as minimum commitment contracts (MCC) is further described and its terms are expressed in parametric notation. Similar to the MPC, three main parameters describe the nature of the contracts. Section 5.3 proposes a mathematical model for the uncertainty of the demand. In Section 5.4, the general form of two-stage stochastic problems is briefly reviewed. Several configurations of the MCC contract are considered to derive different mathematical formulations of the stochastic DNPMC according to their conditions and specifications. Some implications of working under a type of contract are discussed from a mathematical modelling perspective. Two families of problems are distinguished according to reservation/requests terms. In Section 5.5, we assume T&D requests can be made after the demand is realized, whereas in Section 5.6 T&D requests should be made before the real demand is known. In the final Section 5.7 we summarize the contents. Solution approaches for one of the models in Section 5.6 will be developed and tested in Chapter 6.

5.2 The Minimum Capacity Commitment Contract - MCC

The *minimum capacity commitment* contract (MCC) here described represents again a risk-sharing mechanism in business agreements between buyers and suppliers. In general terms, the buyer commits to purchasing at least a certain amount of products or services. Even if the supplier provides multiple products/services, the total minimum order quantity is stated in homogeneous units (e.g., number of items, volume, weight, etc.) therefore, minimum order quantities for each individual product/service are not clearly specified. As a consequence, the minimum payment is not quantifiable. This repeats in several periods throughout the contract duration.

In the context of freight transportation services, the minimum commitment (M) can be expressed in terms of capacity or load units, for instance transportation capacity (say, a number of full truck-loads) or storage capacity at facilities (say, total volume/weight handled). A carrier operating on multiple lanes and/or facilities is notified of the minimum freight volume that the shipper intends to move across its

complete network during the contract time. Considering the two-echelon distribution network (as described in Section 3.2), each shipping lane/cross-docking facility combination at the first distribution level represents a distinct service with its own cost function, whose value depends on the amount of freight handled. In exchange for the secured level of business, carriers agree to charge contract prices for services requested, up to the commitment level or more.

Several types of MCC contracts can be derived according to the classification displayed in Table 3.3. The main contract qualities are its (1) duration, (2) penalty costs, (3) commitment policy (periodicity) and commitment specification (differentiation per services). Terms and conditions are summed up by the generic parameters (H, P, M) , denoting the attributes of a contract.

As in Chapter 3, the contract duration can be for one planning period ($H_e = 1$), or extend over multiple planning periods ($H_e > 1$). Penalty rates may ($P > 0$) or may not ($P = 0$) apply to service costs charged above the commitment level. The commitment M can be imposed for each period ($M = M^t$) or for to the whole duration of the contract ($M = M^H$). Additionally, it can be specified per service (lane/facility) i ($M = M_i$) or for the totality of services provided by each carrier e ($M = M_e$).

A graphical representation of the cost functions associated with an MCC contract is shown in Figure 5.1, for two distinct services 1, 2 with different marginal contractual costs, namely $F_1 < F_2$. The commitment in capacity units (MCC) covers both services (i.e., $M = M_e$). The marginal cost above the commitment level MCC is higher, to reflect for instance the higher cost on the spot market modelled by a penalty rate P on costs F_1 and F_2 . Note that the final payment to be made by the shipper depends on the capacity allocated to each service, either 1, or 2. An absolute minimum (guaranteed) payment can be, however, computed on the basis of the cheaper service.

Let us express some of the common characteristics of the proposed MCC contracts, in relation with the two-phase procurement procedure for contracts with risk-sharing provisions (described in Section 2.1.1). In the first phase is negotiated the configuration of contract parameters $(m_1, m_2, p_1, p_2, F_0)$ explained next:

In the proposed MCC contracts, a common characteristic is to neglect an advanced payment ($F_0 = 0$), distinctive of MPC contracts. The minimum order (service or capacity requests in transportation) level, (denoted as m_1) is the commitment value M_e (or M_e^t specified per period). The price agreed for service requests (indicated as p_1) is related with input parameters $F_{i,l}$. Differently, the maximum capacity available (referred to as m_2) is up to the contract attribute P . Likewise, the penalty value

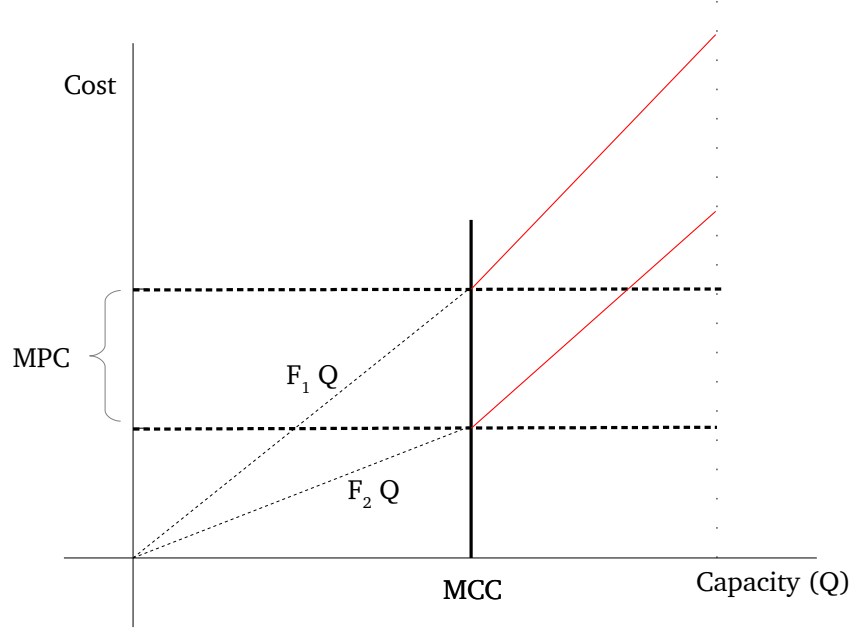


Figure 5.1: The minimum capacity commitment contract

P_e defines the price of service requests above level m_2 (p_2). In the execution phase, capacity requests take place in the execution interval V_1 : orders between m_1 and m_2 , and/or execution interval V_2 : orders above m_2 . Note that in the two-stage stochastic problem, the first stage relates with the phase of parameter definition and the second-stage with the execution phase. In the presenting models, we differentiate a certain type of MCC contract (with given attributes) through a common configuration graph as in Figure 2.2 (introduced in Section 2.1.1).

5.3 Stochastic Representation of the Demand

In the mathematical representation of the distribution network problems with deterministic demand (Chapters 3 and 4), we have used a parameter D_k^t representing the demand of customer area $k \in K$ in period $t \in T$. In the stochastic formulation, the demand of area k at time t is modelled as a discrete random variable d_k^t with values in the set $S_k^t = \{D_k^{t,1}, \dots, D_k^{t,|S_k^t|}\}$, taking value $D_k^{t,s}$ with probability $p_k^{t,s}$.

The demand of all areas in period t is represented by the random vector $d^t = (d_1^t, d_2^t, \dots, d_\kappa^t)$, where $\kappa = |K|$. A particular *demand scenario at time t* is therefore defined as a vector of possible demand realizations in period t , one realization for each area. In other words, a demand scenario at time t is a vector of the form $D^{t,s} = (D_1^{t,s_1}, \dots, D_\kappa^{t,s_\kappa})$ where $s = (s_1, \dots, s_\kappa)$ and $s_k \in \{1, \dots, |S_k^t|\}$ for $k = 1, \dots, \kappa$. We denote by S^t the set of feasible scenarios at period t . Its size is at most equal to $|S_1^t| \times \dots \times |S_\kappa^t|$.

For simplicity, we often refer to the vector s itself as the demand scenario at time t . So, we write $s \in S^t$ in the mathematical models to denote a feasible scenario s at period t .

The unconditional probability of realization of scenario s at time t is denoted as $p^{t,s}$: it is the joint probability of the area-specific demand realization $(D_1^{t,s_1}, \dots, D_\kappa^{t,s_\kappa})$. Hence, $\sum_{s \in S^t} p^{t,s} = 1$. Finally, a *complete demand scenario* is a vector $D^s = (D^{1,s^1}, \dots, D^{T,s^T})$ specifying the demand of all areas over all periods of the planning horizon. Again, we use the shorthand $s = (s^1, \dots, s^T)$ for the complete demand scenario D^s , and we denote by S the set of all complete demand scenarios.

5.4 Two-stage Stochastic Model with Fixed Recourse

The stochastic problems presented in this chapter belong to the general class of two-stage stochastic programs. Decisions taken in the first stage are in the presence of uncertainty of random events (i.e, without full information), while decisions made in the second stage (also called corrective actions or recourse decisions) are dependent on realizations of the random events.

The two-stage stochastic linear program with fixed recourse is formulated as follows (see Beale [1955]; Dantzig [1955]; our description is based on Birge & Louveau [2011]):

$$\begin{aligned} \min z &= c^T x + E_\xi [\min q(\omega)^T y(\omega)] \\ \text{subject to } Ax &= b \\ T(\omega)x + Wy(\omega) &= h(\omega) \\ x \in X, y(\omega) &\in Y \end{aligned}$$

Here, the first stage decision variables are represented by the vector x and the second stage variables by the vector $y(\omega)$, which depends on the random vector $\omega \in \Omega$. The constraints $x \in X, y(\omega) \in Y$ may express, for example, that the variables must take nonnegative, or integer values. The vectors b, c and the matrix A are associated with the variables x . The random vector ω potentially influences all components of $(q(\omega), T(\omega), h(\omega))$, which are gathered in the random vector ξ . E_ξ denote mathematical expectation with respect to ξ .

For a given vector x and a given realization of ω , the second-stage value function is defined as

$$Q(x, \xi(\omega)) = \min_y \{q(\omega)^T y \mid Wy = h(\omega) - T(\omega)x, y(\omega) \in Y\}.$$

It determines $y(\omega)$ as the solution of a linear program. Then, the two-stage stochastic linear program can be rewritten as the equivalent deterministic program

$$\begin{aligned} \min z &= c^T x + E_\xi [Q(x, \xi(\omega))] \\ \text{subject to } Ax &= b \\ x &\in X \end{aligned}$$

As we will see, in the distribution network design problem (DNPMC) with stochastic demand, the random components in ξ only consist of the right-hand side vector $h(\omega)$. This vector depends on the random demand scenarios D^s , for $s \in S$. The values of q and T are non-stochastic.

In DNPMC, decisions about contract opening/renewals are considered to be taken before realization of the demand: naturally, contracts are always formalized before their starting period. The corresponding decision variables, therefore, will be first-stage binary variables. Also, the cost vector c will be absent (i.e., no fee is explicitly charged for opening a new contract). Regarding the selection of transportation lanes and distribution facilities, two different types of assumptions can be formulated. In Section 5.5, it is assumed that this type of decisions are taken with full information about the demand: hence, they are scenario-dependent. In Section 5.6, those decisions are taken ex-ante or in anticipation of random events, implying that reservation of capacity is made in advance, without full information about the demand. Then, additional recourse actions for satisfying customer demand may be allowed in the final distribution plan for each given scenario. As a common feature in both type of models, the second-stage value function is a mixed integer-linear program. All models share a similar objective function aimed at minimizing the expected distribution costs along the horizon T .

In a single-period stochastic DNPMC, decisions are taken only once prior and once after realization of the demand. In a multi-period problem, on the other hand, first-stage decisions and corrective actions must be taken repeatedly at each time t , which gives rise to multi-stage stochastic problem. In this case, nonanticipativity of future demand scenarios is strictly enforced by not linking first-stage variables to *future* realizations of the demand. An important feature of the models, however, is that none of the decisions made at period t depend on *previous* realizations of the demand either. This is due to the fact that the demand must be fully satisfied at every period, and that neither inventory nor backlogged demand are carried over from one period to the next. In other words, the realization of a demand scenario in period S^t does not affect, in any form, the decisions regarding opening/renewal of contracts in the next periods ($t + 1, \dots, |T|$), nor the capacity reservation plans in those periods. As a result, non-anticipativity constraints do not need to be explicitly modelled; however, constraints expressing the duration of each contract will link consecutive periods.

This observation distinguishes our models from those described by classical scenario-trees for multi-stage stochastic problems. The decision-making process for DNPMC is more adequately represented as in Figure 5.2.

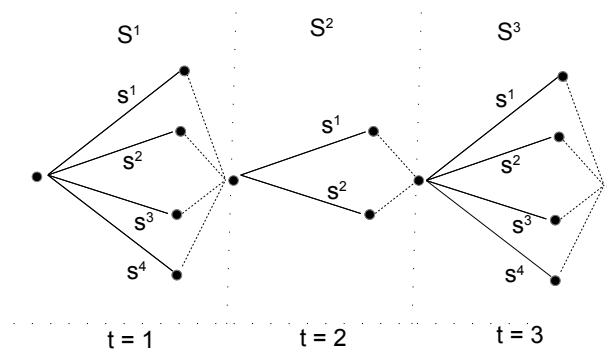


Figure 5.2: Scenario-tree for the stochastic DNPMC problem

The contract decisions in period t , rather than the outcomes of random events in periods $1, \dots, t$, influence the initial state of the system in period $t + 1$. In the following, we present general notation (datasets, parameters, decision variables) for the two subgroups of DNPMC problems, namely with capacity reservation (requests made before demand realizations) and without capacity reservation (requests made after demand realizations).

Commitment Contracts with Demand Uncertainty

Due to the stochastic character of the demand, business agreements may lay down conditions on *reservation* of capacity in the face of uncertainty, before the realization of random events.

In the deterministic case, studied in Chapter 3, reservation terms were not explicitly stated, as part of the MPC agreements. The certainty about the demand allows the shipper to make T&D requests at any point in time. In other words, once a selection of destination points and capacity levels (i.e., cross-docking facilities and truck fleet) is defined by the shipper (with variables $v_{i,l}^t$) the shipper is free to notify this information when the carrier requires, in line with its resources planning schedule for period t , and the fact of notifying in advance or not does not influence the logistics costs. Indeed, the values of $v_{i,l}^t$ determine enough capacity to fulfil the deterministic transportation demand.

Under stochastic demand the total expected costs are affected by whether or not this condition is imposed. If the carrier need the shipper to announce and specify services in advance (of time t), this decision is taken under uncertainty of future events. In that situation, it may occur that the specified capacity ($v_{i,l}^t$) is not sufficient to deliver all purchase orders. An alternative source such as spot markets, are commonly used for urgent one-time loads.

Alternatively, the shipper may be allowed to request T&D capacity after a scenario of demand is observed, in which case $v_{i,l}^{t,s}$ reflects the capacity associated with scenario s and the alternative source is not necessary. Reservation is part of the contractual agreements and denoted with attribute R . Note that R does not imply an input value, but can be considered as a an additional contractual feature of the business context along with the triad (H,P,M) , however, variables $v_{i,l}^t$ determine the reservation level (i.e., the reservation for a service $i \in I_e$ is given by the solution value of variable $v_{i,l}^t$)

Nonetheless, the reserved capacity can alternatively be viewed as an endogenously determined commitment (rather than exogenously given, as it has been the case for M). General mathematical notation for parameters and datasets presented in all stochastic models are summarized in Table 5.1. Decision variables for the first family of stochastic models in Section 5.5 are introduced in Table 5.2.

Datasets and indices

Set	Description	Index
T	Planning horizon	t
E	Candidate carriers	e
I	Cross-docking facilities	i
K	Customer areas	k
I_e	Network of facilities operated by carrier e	i
I_k	Network of facilities that can serve customer area k	i
K_i	Customer areas that can be served from facility i	k
E_k	Carriers that can serve customer area k	e
L_1	Capacity levels when using contracts with carriers in stage 1	l
L_2	Capacity levels when using contracts with carriers in stage 2	l
S^t	Demand scenarios in period t	s
S	Demand scenarios over the planning horizon	s

Input parameters

Symbol	Description
H_e	Duration of contract with carrier e (number of periods)
M_e	Minimum capacity commitment level agreed with carrier (in load units) e
Q_l	Standard transportation and distribution capacity at level l (in load units)
$F_{i,l}$	Service costs for shipping to facility i at capacity level l with contracts (€)
$\bar{F}_{i,l}$	Service costs for shipping to facility i at capacity level l with penalty costs (€)
$D_k^{t,s}$	Demand of customer area k in period t and scenario s (in load units)
$U_{i,k}$	Unit transportation costs for delivery from facility i to area k (€)
P_e	Penalty costs applied to costs charged above the commitment level
$p^{t,s}$	Probability of scenario s in period t

Table 5.1: Notations: Datasets and parameters for stochastic DNPMC problems

Decision variables

Single Period Model

Symbol	Description
α_e	1 if a contract with carrier e takes effect, 0 otherwise. (Stage 1)
$v_{i,l}^s$	Capacity purchased under contract related to facility i at level l (Stage 2)
$q_{i,k}^s$	Demand from customer area k allocated to facility i in scenario s (Stage 2)
c_e^s	Total distribution costs charged by carrier e in scenario s (Stage 2)

Multi Period Models

α_e^t	1 if a contract with carrier e takes effect at period t , 0 otherwise. (Stage 1)
$v_{i,l}^{t,s}$	Capacity purchased under contract related to facility i at level l and scenario s at period t (Stage 2)
$q_{i,k}^{t,s}$	Demand from customer area k allocated to facility i in scenario s at period t (Stage 2)
$c_e^{t,s}$	Total distribution costs charged by carrier e in scenario s at period t (Stage 2)
$\Lambda_e^{t,s}$	Total costs charged for capacity utilization above the commitment level M_e (Stage 2)

Table 5.2: Decision variables for models in Section 5.5

5.5 Stochastic Models with Capacity Request under Demand Certainty

In the mathematical formulations presented in this section is assumed that contracted carriers are able to provide transportation and distribution (T&D) services on-demand without preparation time. This means, carriers can arrange and allocate resources without the need for reservation. Therefore, the shipper can request services with certainty about the demand. After a *contract plan* is decided for period t , it is possible to determine the least-cost *distribution plan* comprising lanes/facilities selection and demand allocation decisions, one for every possible demand scenario $s \in S^t$ in period t . The decision-making follows the sequence illustrated in figure 5.3 for a planing horizon composed of 2 periods.

Decision-Making	t = 0	t = 1	t = 2
First Stage	- Selection of new contracts: Starting from period 1 with carrier $e \rightarrow \alpha_e^1$	- Selection of new contracts: Starting from period 2 with carrier $e \rightarrow \alpha_e^2$	
Second Stage		- Request for T&D services Specify service i at level l for period 1 given scenario realized $s^1 \rightarrow v_{i,l}^{1,s^1}$	- Request for T&D services Specify service i at level l for period 2 given scenario realized $s^2 \rightarrow v_{i,l}^{2,s^2}$
		- Demand allocation Last-mile delivery plan in period 1 for demand scenario $s^1 \rightarrow q_{i,k}^{1,s^1}$	- Demand allocation : Last-mile delivery plan in period 2 for demand scenario $s^2 \rightarrow q_{i,k}^{2,s^2}$
Demand Realizations	$s^1 = (s_1^1, \dots, s_k^1)$		$s^2 = (s_1^2, \dots, s_k^2)$

Figure 5.3: Multi-stage stochastic decision-making under MCC contracts without capacity reservation

5.5.1 First stage

Before information about the demands is known, the shipper may initiate business relationships with contract carriers from the set of candidates E , starting from the next period. The decision variable $\alpha_e^t \in \{0, 1\}$ determines whether or not to undertake an optional contract with carrier $e \in E$ from period t . Every new contract

expands the network of logistics facilities and resources taken into account in the *distribution planning* problem (Second-stage). For simplicity, the accessible capacity with a particular carrier is considered unlimited, however different prices may apply depending on the volume of services requested. Furthermore, the MCC contract forces the shipper to utilize resources to at least the capacity commitment level.

As shown in Figure 5.3, the multi-period approach requires to consider establishing new relationships at each time t , taking into account ongoing contracts and the demand probabilistic function. The stochastic demand may lead to change or modify the arrangement of contracts accordingly, with the purpose to minimize overall expected costs.

In principle, low-demand scenarios (requiring not much capacity) lead to avoid multiple active contracts at the same time, since the total capacity committed, might be underused or not used at all. On the other hand, high-demand scenarios encourage to activate several contracts, making sure to utilize fully purchased capacity at contract prices and profiting from cost-savings as a result of an extensive network of facilities/resources. Nonetheless, other factors such as individual service costs, penalties or parcel delivery costs may also affect the volume of contracts. In general, dealing with multiple probabilistic scenarios S as well as multiple alternatives in E increase the difficulty in determining an optimal selection of carriers.

5.5.2 Second stage

After the demand is realized in period t , it is up to the shipper to request transportation and distribution capacity in order to completely deliver customers' orders. The second-stage problem consists in determining the least-cost distribution plan for each possible scenario $s \in S^t$, taking into account T&D resources enabled by active contracts. Second-stage decision variables $(v_{i,l}^{t,s}, q_{i,k}^{t,s})$ account for selection of shipping lane plus transportation and customer demand allocation, respectively. In the multi-stage framework, both types of operational decisions at time t correspond to recourse actions, which are taken independently from the outcome of the demand in previous periods (i.e., s^1, \dots, s^{t-1}). This characteristic is explained by the fact that the demand is always completely satisfied. Therefore, the distribution planning problem can be treated independently in every period $t \in T$, conditioned only by the active contracts opted in the first-stage.

Figure 5.4 (based on Akbalik et al. [2017]) depicts an example of a multi-period procurement plan with carrier e following an MCC contract without need of reser-

vation. In this particular example, the levels of capacity increase linearly (e.g., homogeneous fleet of vehicles), and the minimum commitment M with carrier e throughout the contract duration is the same in all periods.

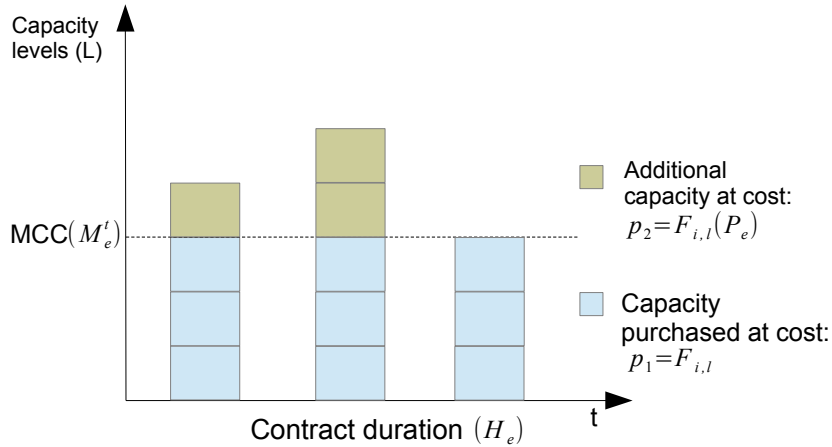


Figure 5.4: The MCC contract - capacity reservation

5.5.3 Single-period Distribution Network Problem: MCC ($H = 1, P = 0, M = M_e$)

In this mathematical representation the relationships between the shipper and carriers are established by contracts of the form MCC ($H = 1, P = 0, M = M_e$). The duration is agreed for a single period with all carriers (i.e., $H_e = 1 \forall e \in E$), penalty costs do not apply regardless of the utilization level, and the commitment level M_e bears on the complete carrier's network e . Taking into account these assumptions, the configuration of the MCC contract is given in Figure 5.5:

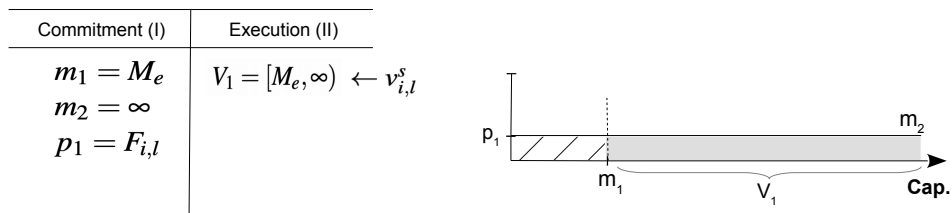


Figure 5.5: Configuration of an MCC contract without penalty rates

A signed contract enables unlimited capacity in transportation lanes/facilities ($m_2 = \infty$) at contract prices ($p_1 = F_{i,l}$ in the model). Hence, penalties are not considered. The minimum service requests pre-agreed is set in capacity units ($m_1 = M_e$). In the execution step, capacity requests (handled by $v_{i,l}^s$ variables) range over $V_1 = [M_e, \infty)$. The area with diagonal lines, indicates the minimum capacity level, while the shaded area shows the interval V_1 along which the shipper could execute orders.

The mathematical formulation for the DNPMC is presented below.

Objective Function

$$\min \sum_{s \in S} p^s \left[\sum_{e \in E} c_e^s + \sum_{k \in K} \sum_{i \in I_k} U_{i,k} q_{i,k}^s \right] \quad (5.1)$$

Constraints

$$\sum_{i \in I_k} q_{i,k}^s = D_k^s \quad \forall k \in K, s \in S \quad (5.2)$$

$$\sum_{k \in K_i} q_{i,k}^s \leq \sum_{l \in L} Q_l v_{i,l}^s \quad \forall i \in I, s \in S \quad (5.3)$$

$$q_{i,k}^s \leq D_k^s \sum_{l \in L} v_{i,l}^s \quad \forall i \in I_k, k \in K, s \in S \quad (5.4)$$

$$\sum_{l \in L} v_{i,l}^s \leq \alpha_e \quad \forall i \in I_e, e \in E, s \in S \quad (5.5)$$

$$\sum_{i \in I_e} \sum_{l \in L} Q_l v_{i,l}^s \geq M_e \alpha_e \quad \forall e \in E, s \in S \quad (5.6)$$

$$c_e^s = \sum_{l \in L} \sum_{i \in I_e} F_{i,l} v_{i,l}^s \quad \forall e \in E, s \in S \quad (5.7)$$

$$q_{i,k}^s \geq 0 \quad \forall k \in K, i \in I, s \in S \quad (5.8)$$

$$v_{i,l}^s \in \{0, 1\} \quad \forall i \in I, l \in L, s \in S \quad (5.9)$$

$$\alpha_e \in \{0, 1\} \quad \forall e \in E \quad (5.10)$$

The objective function (5.1) aims to minimize the expected transportation and distribution costs. It comprises first, cost of FTL carriers represented with auxiliary variables c_e^s for candidates in E and second, last-mile parcel delivery services charged at fixed unit costs $U_{i,k}$. According to the discrete probabilistic distribution of the demand, the value of the objective function can be computed. A single set of demand scenarios S is foreseen.

The first group of constraints regard transportation capacity and demand allocation decisions. The satisfaction of the demand is enforced for all customer areas in every demand scenario (Eq. (5.2)). The distribution output flow of facility i is limited according to the incoming transportation capacity, specified per levels Q_l (Eq. (5.3)). Each level corresponds for instance to a number of trucks. A customer area k can only be served from facilities operated by contracted carriers (Eq. (5.4)). This last set of constraints help to reinforce the linear relaxation formulation.

The second group of constraints (placed after the first dotted line), express conditions and permissions linked to MCC contracts. A valid contract allows to use a carrier's network, in which case only one transportation capacity level per facility i can be selected (Eq. (5.5)) and the minimum capacity commitment must be attained (Eq. (5.6)). In this formulation is compulsory to open MCC contracts to meet distribution requirements. The total costs for capacity utilization with carrier e in scenario s are computed with the auxiliary cost function in Eq. (5.7). Finally, constraints in Eqs. (5.8)-(5.10) specify the range and nature of decision variables. Note that parameters paired with second-stage variables (i.e., contract costs per service $(F_{i,l})$ and capacity levels (Q_l) are independent of random events (Fixed Recourse).

Note that setting the contract parameter $M_e = 0, \forall e \in E$, removing in that way restrictions on resource utilization renders unnecessary the use of contract selection variables. This implies, every carrier can be selected without implication in the total cost function. The stochastic model can be split per scenario $s \in S$ and multiple single-period deterministic models can be solved separately.

Considering only one scenario ($|S| = 1$) of the demand (e.g., taking the expected demand value for each area k), produces a similar version of the deterministic DNPMC model introduced in Chapter 3, for the single-period DNP problem with MPC contracts (Section 3.5). It differs just in terms of the type of commitment enforced by Eq. (5.6). Unlike the model in Section 3.5, the set of constraints in Eq. (5.7) is stated as an equality.

The above formulation serves as a basis which is slightly modified in the following models presented in this section.

5.5.4 Multi-period Distribution Network Problem: MCC ($H > 1, P = 0, M = M_e^t$)

The multi-period model becomes necessary if contracts last for several periods of the shipper's per period procurement review. The duration is agreed upon individually per carrier ($H_e > 1$ for candidate e in set E). Under this assumption, the shipper may consider contracts of the form MCC ($H > 1, P = 0, M = M_e^t$). In this framework, potential carriers provide unlimited capacity at contract costs on the condition that capacity requests amount to a minimum commitment level agreed for each contract period M_e^t . This implies that, penalty costs after reaching the minimum commitment do not come into effect ($P = 0$). The commitment is expressed in load/capacity units without specifying T&D services at a given location of the carrier's network. The complete formulation is presented as follows:

Objective Function

$$\min \sum_{t \in T} \sum_{s \in S^t} p^{t,s} \left[\sum_{e \in E} c_e^{t,s} + \sum_{k \in K} \sum_{i \in I_k} U_{i,k} q_{i,k}^{t,s} \right] \quad (5.11)$$

Constraints

$$\sum_{i \in I_k} q_{i,k}^{t,s} = D_k^{t,s} \quad \forall k \in K, t \in T, s \in S^t \quad (5.12)$$

$$\sum_{k \in K_i} q_{i,k}^{t,s} \leq \sum_{l \in L} Q_l v_{i,l}^{t,s} \quad \forall i \in I, t \in T, s \in S^t \quad (5.13)$$

$$q_{i,k}^{t,s} \leq D_k^{t,s} \sum_{l \in L} v_{i,l}^{t,s} \quad \forall i \in I_k, k \in K, t \in T, s \in S^t \quad (5.14)$$

$$\sum_{l \in L} v_{i,l}^{t,s} \leq \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall i \in I_e, e \in E, t \in T, s \in S^t \quad (5.15)$$

$$\sum_{i \in I_e} \sum_{l \in L} Q_l v_{i,l}^{t,s} \geq M_e \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall e \in E, t \in T, s \in S^t \quad (5.16)$$

$$\sum_{n=0}^{H_e-1} \alpha_e^{t-n} \leq 1 \quad \forall e \in E, t \in T \quad (5.17)$$

$$c_e^{t,s} = \sum_{l \in L_i \in I_e} F_{i,l} v_{i,l}^{t,s} \quad \forall e \in E, t \in T, s \in S^t \quad (5.18)$$

$$q_{i,k}^{t,s} \geq 0 \quad \forall k \in K, i \in I, t \in T, s \in S^t \quad (5.19)$$

$$v_{i,l}^{t,s} \in \{0, 1\} \quad \forall i \in I, l \in L, t \in T, s \in S^t \quad (5.20)$$

$$\alpha_e^t \in \{0, 1\} \quad \forall e \in E, t \in T \quad (5.21)$$

The mathematical formulation is similar to the one presented in the previous section. The time dimension (superscript t) is added to decision variables in both stages, defining a multi-stage stochastic problem (as in Figure 5.3). The stochastic parameter of the demand is also specified per period, capturing in that way time-dependent fluctuations. The set of scenarios is then specified individually per time t , S^t , in which the probability of scenario s at time t is $p^{t,s}$.

As a consequence, the expected cost minimization function (5.11) accounts for multiple periods, separating costs of first and second echelon of the distribution layout. A noticeable difference with respect to the single-period model regards the addition of Eq. (5.17) stipulating to validate a contract from its starting period up to its expiration, without possibility of cancellation, nor extension before it is finished. Therefore, the right side of Eqs. (5.5), (5.6) are modified in Eqs. (5.15), (5.16) to express validity of a contract in time t . These set of constraints, link decision variables α_e^t for consecutive periods, which makes the multi-period model not separable or not decomposable.

The model here described serves as a generalization of the multi-period deterministic model introduced in Section 3.6 and further described in Chapter 4, with same contract conditions, but different type of commitment. The solution methods developed in Chapter 4 (Combinatorial Benders, Relax-and-Repair Heuristic) can be adapted to the stochastic problem given similar contract conditions. Both procedures having in common to separate the overall problem in two phases. First, determining a solution for contract related variables (α_e^t) in the master problem/relaxation phase. Second, finding a solution for second-stage variables ($q_{i,k}^{t,s}, v_{i,l}^{t,s}$) in the corresponding sub-problem/repairing phase. The increased difficulty will come from solving the recourse problem for multiple scenarios of demand, taking into account capacity allocation variables $v_{i,l}^{t,s}$ are integer.

The L-shaped method is also commonly known to tackle two-stage stochastic models with decomposition structures. However its classical formulation is only valid for linear programs. The combinatorial approach may prove useful in dealing with integer variables in both first and second-stage of the stochastic problem.

5.5.5 Multi-period Distribution Network Problem with Penalty Costs: MCC ($H > 1, P > 0, M = M_e^t$)

In this version of the multi-period problem, we consider a variant of the MCC contract assuming now that carriers impose additional costs on the capacity used above

the commitment. The contract takes the form MCC ($H > 1, P > 0, M = M_e^l$). Under this policy, the shipper adds up T&D capacity to the distribution network at contract costs, up to a fixed limit. In addition, unlimited capacity is available at higher off-contract costs (adjusted by penalties). The penalty cost is specified per carrier (P_e) related to standard practices to each individual firm, and is charged on the capacity used above the minimum commitment level (M_e^l). Certainly, taking into account penalties, the MCC level M_e could represent with more fidelity the exact requirements that the shipper intends to use with a certain carrier, instead of a minimum (best-case scenario) which will just be sufficient to have access to unlimited services. This difference in costs presents a problem when the demand is stochastic, since random events could lead to one of the two sub-optimal situations, underutilised capacity or overpriced capacity.

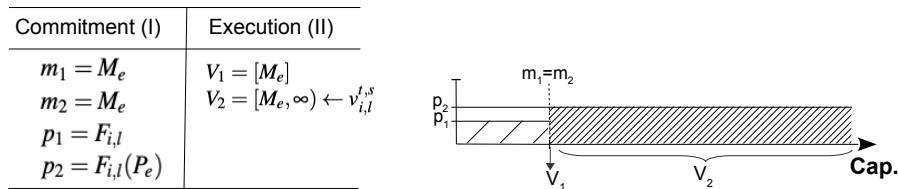


Figure 5.6: Configuration of and MCC contract with penalty rates)

The configuration of an MCC contract with penalty rates is depicted in Figure 5.6. The minimum level of service requests m_1 is the commitment M_e . Carriers provide limited capacity in transportation lanes/facilities at contract prices ($p_1 = F_{i,l}$). This limit is also according to M_e , ($m_2 = M_e$). Service costs above level m_2 are adjusted by the penalty rate P_e ($p_2 = F_{i,l}(P_e)$). Note that the penalty rate P_e should be stated in monetary units (per unit of capacity), instead of as a percentage to the difference with respect to the commitment level, as it was the case in Section 3.7, for the MPC contract. In the execution step, capacity requests at p_1 are enforced ($V_1 = [M_e]$, long-spaced lined area) while capacity requests at p_2 range in the interval $V_2 = [M_e, \infty)$ (short-spaced lined area). This decision is handled by variables $v_{i,l}^{f,s}$.

The mathematical model is presented as follows.

Objective Function

$$\min \sum_{t \in T} \sum_{s \in \mathcal{S}^t} p^{t,s} \left[\sum_{e \in E} c_e^{t,s} + \sum_{k \in K} \sum_{i \in I_k} U_{i,k} q_{i,k}^{t,s} \right] \quad (5.11)$$

Constraints

$$\sum_{i \in I_k} q_{i,k}^{t,s} = D_k^{t,s} \quad \forall k \in K, t \in T, s \in \mathcal{S}^t \quad (5.12)$$

$$\sum_{k \in K_i} q_{i,k}^{t,s} \leq \sum_{l \in L} Q_l v_{i,l}^{t,s} \quad \forall i \in I, t \in T, s \in \mathcal{S}^t \quad (5.13)$$

$$q_{i,k}^{t,s} \leq D_k^{t,s} \sum_{l \in L} v_{i,l}^{t,s} \quad \forall i \in I_k, k \in K, t \in T, s \in \mathcal{S}^t \quad (5.14)$$

$$\sum_{l \in L} v_{i,l}^{t,s} \leq \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall i \in I_e, e \in E, t \in T, s \in \mathcal{S}^t \quad (5.15)$$

$$\Lambda_e^{t,s} \geq \sum_{i \in I_e} \sum_{l \in L} Q_l v_{i,l}^{t,s} - M_e^t \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall e \in E, t \in T, s \in \mathcal{S}^t \quad (5.22)$$

$$\sum_{n=0}^{H_e-1} \alpha_e^{t-n} \leq 1 \quad \forall e \in E, t \in T \quad (5.17)$$

$$c_e^{t,s} \geq \sum_{l \in L} \sum_{i \in I_e} F_{i,l} v_{i,l}^{t,s} + \Lambda_e^{t,s} P_e \quad \forall e \in E, t \in T, s \in \mathcal{S}^t \quad (5.23)$$

$$q_{i,k}^{t,s} \geq 0 \quad \forall k \in K, i \in I, t \in T, s \in \mathcal{S}^t \quad (5.19)$$

$$v_{i,l}^{t,s} \in \{0, 1\} \quad \forall i \in I, l \in L, t \in T, s \in \mathcal{S}^t \quad (5.20)$$

$$\alpha_e^t \in \{0, 1\} \quad \forall e \in E, t \in T \quad (5.21)$$

$$\Lambda_e^{t,s} \geq 0 \quad \forall e \in E, t \in T, s \in \mathcal{S}^t \quad (5.24)$$

The mathematical formulation above differs from the previous model by including an auxiliary variable $\Lambda_e^{t,s}$, that computes the capacity used above the commitment level in Eq. (5.22). It is declared distinctly per carrier e , period t and scenario s . The non-negativity of variable $\Lambda_e^{t,s}$ is also enforced by Eq. (5.24), which implies that the commitment M_e^t is always satisfied. In addition, the auxiliary cost function in Eq. (5.23) regarding the total costs due to carrier e , adds the penalty rate to every additional unit of capacity.

The model here described is clearly a generalized model which can be easily converted to the particular case of non-penalized contracts ($P_e = 0$). Furthermore, removing constraints in Eq. (5.15) allows to request off-contract capacity with any

carrier. This capacity is penalized and charged at higher than contract prices.

Note also that the penalty P_e is assumed to be established individually per carrier and applied without distinction of the services required. The penalty could be related to each particular service i according to their individual costs. For so doing, the minimum commitments should be stated per service ($M_i^t : \forall i \in I_e, t \in T$) as well as the parameter ($P_i : \forall i \in I$). Finally, the variable $\Lambda_e^{t,s}$ can be modified to $\Lambda_i^{t,s}$ to account for the extra capacity utilized for service i period t and scenario s . Constraints (5.22), (5.23) and (5.24) are replaced by constraints (5.25), (5.26) and (5.27):

$$\Lambda_i^{t,s} \geq \sum_{l \in L} Q_l v_{i,l}^{t,s} - M_i^t \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall i \in I_e, e \in E, t \in T, s \in S^t \quad (5.25)$$

$$c_e^{t,s} \geq \sum_{l \in L} \sum_{i \in I_e} F_{i,l} v_{i,l}^{t,s} + \sum_{i \in I_e} \Lambda_i^{t,s} P_i \quad \forall e \in E, t \in T, s \in S^t \quad (5.26)$$

$$\Lambda_i^{t,s} \geq 0 \quad \forall i \in I, t \in T, s \in S^t \quad (5.27)$$

A similar conversion could be applied for the deterministic models in Section 3.7 and 3.8. Note that when the capacity commitment is stated per service M_i , it is possible to translate it in monetary terms according to the MCC level l (i.e., $F_{i,l}$). The specification of the commitment per service restricts the alternatives for taking decisions at the operational level given that some utilization of capacity is enforced for each service. The next family of stochastic models handles the possibility to maintain the shipper's flexibility to decide where to allocate the committed capacity, while carriers set higher-prices above the MCC level per service.

5.6 Stochastic Models with Capacity Request under Demand Uncertainty

In the following sections, we consider the case when contracted carriers ask shippers to request capacity with anticipation in order to grant lower rates. The shipper must then *reserve* the required services in the presence of uncertainty about the demand, knowing that the minimum capacity commitment (MCC) must be satisfied. Otherwise, late requests are also accepted at a higher off-contract rates with any candidate carrier.

A distinction is made between first-stage and second-stage decision variables. In general, for this class of problems the first-stage comprises decisions on contract openings and/or reservation of T&D services which take place before the demand is observed. Second-stage decisions are based on the outcome of the demand, and consists in requesting additional capacity upon necessity in order to fulfil completely customer orders, and allocating the demand to distribution facilities. In that way, the definite *distribution plan* for scenario $s \in S^t$ is determined. Note that the *distribution plan* is partially sketched in the first-stage and then it is completed in the second-stage. The sequence of decision-making is illustrated in Figure 5.7 for a two-period planning horizon and further described below.

Decision-Making	t = 0	t = 1	t = 2
First Stage	<ul style="list-style-type: none"> - Selection of new contracts: Starting from period 1 with carrier $e \rightarrow \alpha_e^1$ - Request for T&D services: Specify service i at level l for period 1 $\rightarrow v_{i,l}^1$ 	<ul style="list-style-type: none"> - Selection of new contracts: Starting from period 2 with carrier $e \rightarrow \alpha_e^2$ - Request for T&D services: Specify service i at level l for period 2 $\rightarrow v_{i,l}^2$ 	
Second Stage		<ul style="list-style-type: none"> - Request for additional T&D services: Specify service i at level l for period 1, given scenario realized $s^1 \rightarrow w_{i,l}^{1,s_1}$ - Demand allocation: Last-mile delivery plan in period 1 for demand scenario $s^1 \rightarrow q_{i,k}^{1,s_1}$ 	<ul style="list-style-type: none"> - Request for additional T&D services: Specify service i at level l for period 1, given scenario realized $s^2 \rightarrow w_{i,l}^{2,s_2}$ - Demand allocation: Last-mile delivery plan in period 2 for demand scenario $s^2 \rightarrow q_{i,k}^{2,s_2}$
Demand Realizations	$s^1 = (s_1^1, \dots, s_k^1)$		$s^2 = (s_1^2, \dots, s_k^2)$

Figure 5.7: Multi-stage stochastic decision-making under MCC contracts with capacity reservation

5.6.1 First stage

Besides formalizing new business relationships which grant access to T&D services in the carrier's network, the shipper must reserve capacity ahead in time with contracted carriers, before the demand is known. Once a contract becomes effective in time t it is allowed to request services in advance from that same period. Therefore, the possibility to reserve is attached to a formal agreement in force. Under the described situation, first stage decisions are represented with variables α_e^t for contract

openings and $v_{i,l}^f$ for capacity reservation, specified per service (i.e., shipping lane cross-docking point) and level. In accordance to contract stipulations, the total capacity used must reach at least the corresponding commitment level (M). However, the reserved capacity must be paid at pre-agreed contracts rates $F_{i,l}$.

5.6.2 Second stage

Second-stage decisions are represented with variables $w_{i,l}^{f,s}$ for the additional capacity and $q_{i,l}^{f,s}$ for the allocation of demand. Both variables characterize recourse actions linked to a determined scenario. The recourse function defines the *additional* expected costs taking into account the distribution of the demand. After the demand is realized, it is possible to decide whether or not to request additional capacity from the list of candidate carriers E in order to satisfy transportation requirements. Late requests are priced differently (presumably following spot market rates) and can be estimated by some penalty rate P which defines off-contract costs for each service and capacity level $\bar{F}_{i,l} = F_{i,l}(1 + P)$, where P can be seen as an adjustment factor associated with the proportional difference with respect to spot market costs. The penalty rate can be differentiated per carrier P_e . The factor P has a relevant effect on the shipper's decisions about how much to reserve in the first-stage. Assuming for instance that $P = 1$, the best decision is to wait until the demand is realized and to ask for the capacity that is needed. In that sense, the contract relationship is considered to have penalty $P > 0$. On the contrary, a sufficiently large value of P encourages the shipper to secure enough capacity in advance to reduce the likelihood of paying at expensive rates.

Figure 5.8 depicts an example of a procurement planning solution for an MCC contract with capacity reservation. The reservation level R is placed above the minimum commitment initially established (M_e), and can vary according to demand expectations. Capacity up to R is priced at contract costs, and above at off-contract costs.

The developed models hereafter differ in terms of two MCC contracts attributes. More exactly, the contract duration (H) and minimum commitment level (M). In Table 5.3 common notation is introduced for decision variables presented for this family of stochastic models.

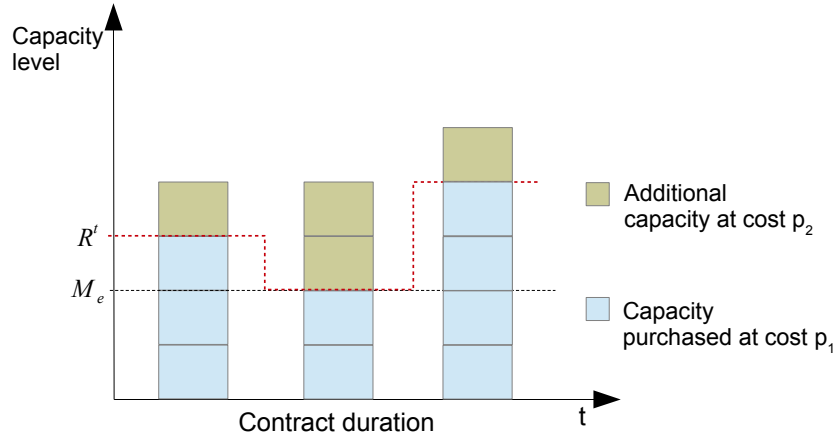


Figure 5.8: The MCC contract with capacity reservation

Decision variables

Single Period Model

Symbol	Description
α_e	1 if a contract with carrier e takes effect, 0 otherwise. (Stage 1)
$v_{i,l}$	Capacity purchased under contract related to facility i at level l (Stage 1)
$w_{i,l}^s$	Capacity purchased off-contract related to facility i at level l and scenario s (Stage 2)
$q_{i,k}^s$	Demand from customer area k allocated to facility i in scenario s (in transportation units) (Stage 2)
c_e^s	Total distribution costs charged by carrier e in scenario s

Multi Period Models

α_e^t	1 if a contract with carrier e takes effect at period t , 0 otherwise. (Stage 1)
$v_{i,l}^t$	Capacity purchased under contract related to facility i at level l at period t (Stage 1)
$w_{i,l}^{t,s}$	Capacity purchased off-contract related to facility i at level l and scenario s at period t (Stage 2)
$q_{i,k}^{t,s}$	Demand from customer area k allocated to facility i in scenario s at period t and (in transportation units) (Stage 2)
$c_e^{t,s}$	Total distribution costs charged by carrier e in scenario s at period t

Table 5.3: Decision variables for models in Section 5.6

5.6.3 Single-period Distribution Network Problem with Capacity Reservation: $MCC(H = 1, P > 0, M = 0, R)$

Consider a contract mechanism of the form $MCC(H = 1, P > 0, M = 0, R)$. The contractual agreement is established for a single period $H = 1$, in which the capacity commitment (M) is not stipulated. However, the carrier requires to be notified the capacity to use before period t , so as to reserve resources. The reservation level is

modelled as a decision to make in the formulation handled by variables $v_{i,l}$ upon which the R level depends, then $R(v_{i,l})$.

Figure 5.9 depicts the configuration of an MCC contracts with reservation requirements. The single-period agreement, does not stipulate a condition on the minimum capacity utilization ($m_1 = 0$), and the maximum capacity available is equal to the reservation level ($m_2 = R(v_{i,l})$), which grants contract prices ($p_1 = F_{i,l}$). Additional capacity is always possible in unlimited quantities at higher off-contract costs ($p_2 = \bar{F}_{i,l}$). In the execution step, it is decided the capacity to take at p_1 in interval $V_1 = [0, R(v_{i,l})]$, and at p_2 in the interval V_2 beyond the reservation level. This execution range is handled by variables $w_{i,l}^s$ in the model.

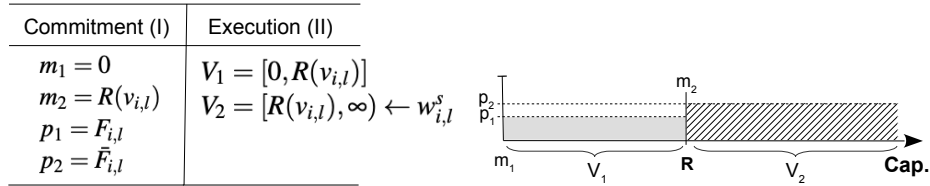


Figure 5.9: Configuration of an MCC contract with capacity reservation

By relaxing completely the minimum commitment ($M_e = 0$), providing in that way greater flexibility to the shipper, reduces the DNPMC formulation to a single-period planning problem. This follows assumptions about the acquired capacity not being available for consecutive periods and the demand being always satisfied. Therefore, the mathematical formulation for the planning horizon T can be separable in multiple two-stage stochastic problems, one for every $t \in T$. Moreover, the use of contract variables (α_e^t) becomes irrelevant since they do not incur any costs. Basically, a relationship with a carrier is activated whenever the shipper decides on using its resources either in the first or second stage. In this setting, the application of lower 'contract' rates is more related to the act of reservation rather than the formalization of an MCC contract.

The two-stage stochastic problem is formulated in the following mathematical form.

Objective Function:

$$\min \sum_{s \in S} p^s \left[\sum_{e \in E} c_e^s + \sum_{k \in K} \sum_{i \in I_k} U_{i,k} q_{i,k}^s \right] \quad (5.1)$$

Constraints

$$\sum_{i \in I_k} q_{i,k}^s = D_k^s \quad \forall k \in K, s \in S \quad (5.2)$$

$$q_{i,k}^s \leq D_k^s \left(\sum_{l \in L_1} v_{i,l} + \sum_{l \in L_2} w_{i,l}^s \right) \quad \forall i \in I, k \in K, s \in S \quad (5.28)$$

$$\sum_{k \in K_i} q_{i,k}^s \leq \sum_{l \in L_1} Q_l v_{i,l} + \sum_{l \in L_2} Q_l w_{i,l}^s \quad \forall i \in I, s \in S \quad (5.29)$$

$$\sum_{l \in L_1} v_{i,l} \leq 1 \quad \forall i \in I \quad (5.30)$$

$$\sum_{l \in L_2} w_{i,l}^s \leq 1 \quad \forall i \in I, s \in S \quad (5.31)$$

$$c_e^s = \sum_{l \in L_1} \sum_{i \in I_e} F_{i,l} v_{i,l} + \sum_{l \in L_2} \sum_{i \in I_e} \bar{F}_{i,l} w_{i,l}^s \quad \forall e \in E, s \in S \quad (5.32)$$

$$v_{i,l} \in \{0, 1\} \quad \forall i \in I, l \in L_1 \quad (5.33)$$

$$w_{i,l}^s \in \{0, 1\} \quad \forall i \in I, l \in L_2, s \in S \quad (5.34)$$

$$q_{i,k}^s \geq 0 \quad \forall k \in K, i \in I, s \in S \quad (5.8)$$

The model formulation suffers several modifications with respect to the structure shown in models in Sections 5.5.3-5.5.5. The objective function (5.1) is similar with a global purpose of minimizing total expected distribution costs. However, the total costs per carrier (c_e^s) follow the auxiliary costs function in Eq. (5.32) which comprises two terms for capacity requests in the first and second stage, respectively. The recourse function now takes into account additional capacity ($w_{i,l}^s$) linked to scenario s .

The first set of constraints (before the first dotted line) are related to network design decisions. The first three assure demand satisfaction (Eq.(5.2)), limit the outflow/inflow from facilities towards customer areas (Eq.(5.28)) and limit the total outflow from each cross-docking facility (Eq. (5.29)) to the total capacity entering facility i in both stages. Note that the sum between parenthesis in Eq.(5.28) can take either value zero, one, or two. In the latter case, the outflow is not completely restricted by the value of the demand, however by constraints (Eq.(5.2)) this is implied. Nonetheless, constraints in Eq. (5.28) aid to reinforce the linear relaxation of the problem.

In this model capacity utilization in the first stage (reserved before demand realization) is not restricted upon a valid contract. Naturally, at most one level per service $i \in I$ can be selected (Eq. (5.30)), also for capacity requests in the second stage (Eq. (5.31)). As mentioned constraint in Eq. (5.32) simply helps to differentiate total costs per carrier and per scenario, while holding the same objective function form. The last set of constraints in Eqs. (5.33), (5.34) and (5.8) specify the range nature of variables.

5.6.4 Single-period Distribution Network Problem with Capacity

Reservation: MCC($H = 1, P > 0, M = M_e, R$)

In this version of the stochastic DNPMC, a condition is imposed on the minimum capacity level requested by the shipper. The MCC contract with capacity reservation takes the form ($T = 1, P > 0, M = M_e, R$). The contract remains valid for a single-period. Besides to early capacity reservation, the shipper must comply with a minimum resource utilization level (M_e). Then, the reserved capacity must be at least the M_e level agreed upon so as to benefit from contract rates $F_{i,l}$. Alternatively, the shipper can always purchase capacity at off-contract costs ($\bar{F}_{i,l}$). The binary variable regarding contract opening is included to differentiate contract carriers from which capacity reservation is allowed.

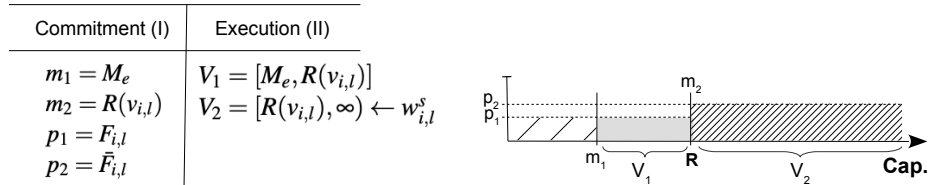


Figure 5.10: Configuration of an MCC contract with capacity reservation and minimum commitment

Figure 5.10 illustrates the configuration of the MCC contract type. In the first step, the contract parameter M_e defines the minimum capacity utilization ($m_1 = M_e$) for the shipper. The maximum capacity available at contract price ($p_1 = F_{i,l}$) depends upon a given reservation, which is obtained from a solution for variables $v_{i,l}$, ($m_2 = R(v_{i,l})$). Service costs outside reservation (p_2) are modelled with input parameters $\bar{F}_{i,l}$. Note that if no contract is taking place at time t with carrier e , $R = 0$ and any

service requests is charged at p_2 . For the second step, the execution of capacity requests at p_1 range in the interval $V_1 = [M_e, R(v_{i,l})]$. On the other end, capacity requests at p_2 are considered unlimited beyond level $R(v_{i,l})$, which is obtained from the solution of variables $w_{i,l}^s$.

In sum, for a single period the decision making follows. In the first-stage decisions are made about selecting MCC contracts with a subset of potential carriers and reserving therein the capacity level to use at determined shipping lanes/facilities. In the second stage, recourse decisions are taken about extra capacity needed from any carrier (with or without contract) and demand allocation. The MILP model takes the form:

$$\min \sum_{s \in S} p^s \left[\sum_{e \in E} c_e^s + \sum_{k \in K} \sum_{i \in I_k} U_{i,k} q_{i,k}^s \right] \quad (5.1)$$

Constraints

$$\sum_{i \in I_k} q_{i,k}^s = D_k^s \quad \forall k \in K, s \in S \quad (5.2)$$

$$q_{i,k}^s \leq D_k^s \left(\sum_{l \in L_1} v_{i,l} + \sum_{l \in L_2} w_{i,l}^s \right) \quad \forall i \in I, k \in K, s \in S \quad (5.27)$$

$$\sum_{k \in K_i} q_{i,k}^s \leq \sum_{l \in L_1} Q_l v_{i,l} + \sum_{l \in L_2} Q_l w_{i,l}^s \quad \forall i \in I, s \in S \quad (5.28)$$

$$\sum_{l \in L_2} w_{i,l}^s \leq 1 \quad \forall i \in I, s \in S \quad (5.31)$$

$$\sum_{i \in I_e} \sum_{l \in L_1} Q_l v_{i,l} \geq M_e \alpha_e \quad \forall e \in E \quad (5.35)$$

$$\sum_{l \in L_1} v_{i,l} \leq \alpha_e \quad \forall i \in I_e, e \in E \quad (5.36)$$

$$c_e^s = \sum_{l \in L_1} \sum_{i \in I_e} F_{i,l} v_{i,l} + \sum_{l \in L_2} \sum_{i \in I_e} \bar{F}_{i,l} w_{i,l}^s \quad \forall e \in E, s \in S \quad (5.32)$$

$$v_{i,l} \in \{0, 1\} \quad \forall i \in I, l \in L_1 \quad (5.33)$$

$$w_{i,l}^s \in \{0, 1\} \quad \forall i \in I, l \in L_2, s \in S \quad (5.34)$$

$$q_{i,k}^s \geq 0 \quad \forall k \in K, i \in I, s \in S \quad (5.8)$$

$$\alpha_e \in \{0, 1\} \quad \forall e \in E \quad (5.10)$$

The mathematical model builds upon the previous formulation in Section 5.6.3. It

includes decision variables α_e (Eq. (5.10)) and constraints (Eq. 5.35) enforcing the minimum capacity level (M_e). The right of reservation is now exclusive upon contract validation, by constrains (Eq. (5.36)). As a particular case, if the commitment level is set to zero ($M_e = 0 : \forall e \in E$) the model can adopt the form of the previous one in Section 5.6.3. From this observation, it is clear that the optimal value of function (Eq. (5.1)) is as high in this model as in the model of Section 5.6.3.

5.6.5 Multi-period Distribution Network Problem with Capacity Reservation: MCC($H > 1, P > 0, M = M_e, R$)

In this section, the multi-period stochastic DNPMC integrates contracts of the form MCC ($H > 1, P > 0, M = M_e, R$). The contracts lasts for several consecutive periods of the planning horizon T . While the contract is in effect with carrier e , the shipper should submit early requests of transportation and distribution services, equivalent to at least the predetermined capacity M_e^t . In that way, the shipper secures service costs and avoid the risks of overpriced services. On the other hand, the carrier secures resource utilization for several periods. However, in a fluctuating and dynamic market environment, the length of contracts may restrict the flexibility of the shipper in adjusting transportation capacity required to the expected needs, which may lead to overpay for the capacity available (M_e^t).

Considering stochastic demand, the multi-period stochastic DNPMC is described next. The formulation maintains the same structure as the single-period model in Section 5.6.4. It adds the time dimension to account for mid-term duration of contracts for several planning periods (H_e) and variability of the random distribution of the demand. The multi-period model includes constraints in Eq. (5.17) to represent the validity of a contract in period t with expression ($\sum_{n=0}^{H_e-1} \alpha_e^{t-n}$). Thereby, adding it to constraints in Eq.(5.41) to comply with minimum capacity utilization and to Eq. (5.42) to allow for reservation in the first-stage.

The formulation presented here above generalizes previous models for DNPMC with MCC contracts plus capacity reservation. It takes into account the possibility to separate decisions in multiple periods of the planning horizon in a multi-stage framework.

Objective Function:

$$\min \sum_{t \in T} \sum_{s \in S^t} p^{t,s} \left[\sum_{e \in E} c_e^{t,s} + \sum_{k \in K} \sum_{i \in I_k} U_{i,k} q_{i,k}^{t,s} \right] \quad (5.11)$$

Constraints

$$\sum_{i \in I_k} q_{i,k}^{t,s} = D_k^s \quad \forall k \in K, t \in T, s \in S^t \quad (5.12)$$

$$q_{i,k}^{t,s} \leq D_k^{t,s} \left(\sum_{l \in L_1} v_{i,l}^t + \sum_{l \in L_2} w_{i,l}^{t,s} \right) \quad \forall i \in I, k \in K, t \in T, s \in S^t \quad (5.38)$$

$$\sum_{k \in K_i} q_{i,k}^{t,s} \leq \sum_{l \in L_1} Q_l v_{i,l}^t + \sum_{l \in L_2} Q_l w_{i,l}^{t,s} \quad \forall i \in I, t \in T, s \in S^t \quad (5.39)$$

$$\sum_{l \in L_2} w_{i,l}^{t,s} \leq 1 \quad \forall i \in I, t \in T, s \in S^t \quad (5.40)$$

$$\sum_{i \in I_e} \sum_{l \in L_1} Q_l v_{i,l}^t \geq M_e^t \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall e \in E, t \in T \quad (5.41)$$

$$\sum_{l \in L_1} v_{i,l}^t \leq \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall i \in I_e, e \in E, t \in T \quad (5.42)$$

$$\sum_{n=0}^{H_e-1} \alpha_e^{t-n} \leq 1 \quad \forall e \in E, t \in T \quad (5.17)$$

$$c_e^{t,s} = \sum_{l \in L_1} \sum_{i \in I_e} F_{i,l} v_{i,l}^t + \sum_{l \in L_2} \sum_{i \in I_e} \bar{F}_{i,l} w_{i,l}^{t,s} \quad \forall e \in E, t \in T, s \in S^t \quad (5.42)$$

$$v_{i,l}^t \in \{0, 1\} \quad \forall i \in I, l \in L_1, t \in T \quad (5.43)$$

$$w_{i,l}^{t,s} \in \{0, 1\} \quad \forall i \in I, l \in L_2, t \in T, s \in S^t \quad (5.44)$$

$$q_{i,k}^{t,s} \geq 0 \quad \forall k \in K, i \in I, t \in T, s \in S^t \quad (5.19)$$

$$\alpha_e^t \in \{0, 1\} \quad \forall e \in E, t \in T \quad (5.21)$$

5.7 Conclusions

In this chapter, several variants of the stochastic DNPMC are introduced and their respective mixed-linear programming formulations described. The mathematical formulations include decisions made in the first-stage under uncertainty and in the second-stage. The models build upon simpler versions characterized by less-

stringent conditions on resource/capacity utilization. The models presented exhibit a common structure that is extended to more general versions, which integrate more stringent conditions of the MCC agreement. The stochastic formulation of the demand led to make further assumptions about the appropriate timing of decisions made in the presence of random events and the way recourse actions are deployed to satisfy relevant features of the problem.

In the next chapter, we focus on the algorithmic development and solution methodologies for one of the most complex and interesting models, namely, the model presented in Section 5.6.5, attempting to find efficiently close-to-optimal solutions, particularly when dealing with large instances of the problem.

Chapter 6

Multi-Period Distribution Network Problem with Stochastic Demand and Minimum Capacity Commitment (MCC) Contracts

6.1 Introduction

In this chapter we focus on the multi-period stochastic problem presented in Section 5.6.5 of chapter 5. The context of the problem is similar to the one presented in chapter 4. The decision-maker is represented by an intermediary logistics company - also known as the *shipper* - providing warehousing and home delivery services to e-commerce sellers. However, transportation services such as truckload (FTL) and last-mile parcel delivery are outsourced to specialized carriers and postal companies. The outbound distribution layout is separated in two legs as depicted in Figure 6.1. In the first leg, FTL carriers transport parcels to intermediate cross-docking facilities while in the second leg postal companies deliver parcels to customers. Cross-docking facilities are operated independently by carriers and multiple geographical areas should be served at every time period.

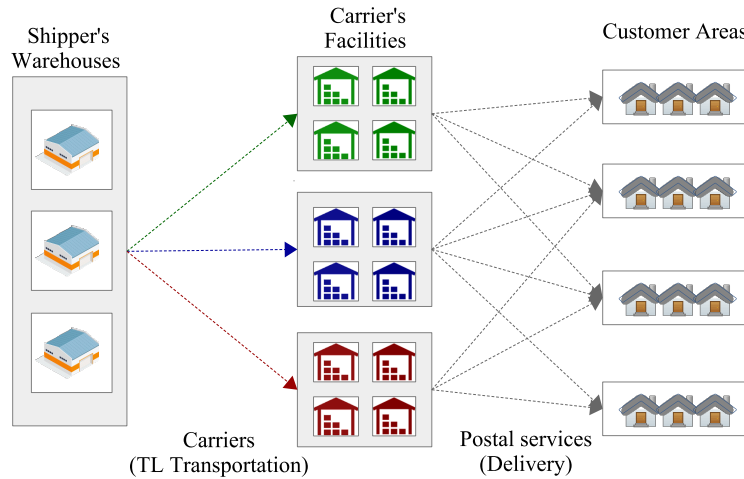


Figure 6.1: Distribution network layout

The planning horizon is divided in a discrete number of periods where the demand of products is dynamic and subject to uncertainty. Therefore, it fluctuates from one period to another following discrete probability distributions. Business agreements between the shipper and carriers are modelled by minimum capacity commitment (MCC) contracts which establish conditions regarding contract duration, minimum service levels and additional costs. The proposed MCC contract is intended to favour both parties by mitigating the risk of capacity shortage and high-prices for the shipper and underutilized resources for the carriers, due to random events. No formal agreements are subscribed with postal companies assuming unlimited available capacity at fixed rates.

The *shipper's* objective is to minimize the expected value of its transportation and distribution costs in the medium to long term, taking into account uncertainty of the demand and contracting options with candidate carriers. The stochastic problem is explained in detail in Section 6.2. The cost-minimization problem is modelled in the form of a stochastic MILP presented in Section 6.3. Algorithmic methods developed are described in Section 6.4. The experimental design and computational tests are presented in sections 6.5 and 6.6, respectively. Finally, conclusions are laid out in Section 6.8.

6.2 Stochastic Problem

The optimization problem of consideration in this chapter is denominated *the multi-period distribution network design problem with stochastic demand and minimum capacity commitment contracts*, *mDNP_MCC*, or in simple form it is referred as *mDNP*. The problem is formulated as a multi-stage stochastic MILP. At every period the decision-making sequence is followed in two stages. A more detailed description before and after the occurrence of random events is described hereunder.

6.2.1 First Stage

In the first stage the *shipper* considers decisions at the *tactical level*.

Transportation capacity required to haul parcels to regional cross-docking points is contracted to external carriers via *minimum capacity commitment contracts*, which could be signed at any planning period and remain valid up to the expiration date. The shipper needs to decide whether or not to open new contracts so as to expand its current distribution network.

As described in Section 5.2, MCC contracts declare conditions that vary in terms of time, costs and service requirements. The contract applied in this context is MCC ($H > 1, P > 0, M = M_e^t, R$) which stands for contracts lasting for several planning periods, penalty costs for additional capacity above the *reservation* level, and periodic commitment over the complete carrier's network. The term *off-contract* is used here to refer to capacity purchased with penalty rates, presumably in association with spot-market prices. As part of the risk-sharing conditions, transportation capacity has to be reserved with anticipation, that means in advance of demand realization and specified per shipping lane in the carrier's network. In that way, the shipper can profit from lower contractual costs. These conditions apply from the period the contract begins.

Alongside opening new contracts, the *shipper* needs to determine the capacity level (e.g., fleet of trucks, load space) to reserve in a particular lane, unknowing the resulting demand. Notice that the commitment M_e^t could be satisfied by capacity requests in both stages. However, under the assumption of higher off-contract costs in the second-stage, it results more economically beneficial to satisfy the MCC in the first stage, at contractual costs. Therefore, the minimum capacity commitment determines a minimum on the reservation level (i.e., $M_e^t \leq R_e^t$). The *reservation plan* only has effect for the immediate planning period independent of the realized demand

scenario, meaning that excess of reserved capacity can not be utilized later and shortage of capacity has to be handled from the alternative sourcing option. Hence, once the demand has been completely fulfilled for one period, a new reservation plan is made for the subsequent one.

6.2.2 Second Stage

In the second stage decisions are taken at the *operational level*.

At this stage the demand of the current period is revealed and known with certainty, we assume that it corresponds to one of the possible scenarios from the discrete probabilistic distribution. The *shipper* takes recourse actions based on the realization of random events. In response to the revealed demand, two main decisions are taken.

On one side, the additional capacity to request (from any candidate carrier) in case the reserved capacity with contracted carriers is insufficient to fulfil customers' orders. Therefore, the capacity available in shipping lanes can be expanded. The additional capacity is unlimited and we can assume that the maximum purchasable capacity ($\max. Q_l : l \in L_2$) respond to any scenario of demand. On the other side, the definite *distribution plan* is completed after determining the way parcels are carried from the shipper's warehouses through cross-docking facilities, and finally to customer areas. The flow of parcels is limited by the transportation capacity acquired in first-leg lanes. In the last leg, parcels are delivered by postal companies.

6.3 Mathematical Modelling for the Multi-stage Stochastic *mDNP_MCC* Model

The multi-stage stochastic MILP model presented in Section 5.6.5 is repeated here, for the sake of completeness, with additional comments regarding the interpretation of feasible solutions. The mathematical notation is summarized in Table 6.1.

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	Cross-docking facilities operated by carrier e	i
I_k	Cross-docking facilities that can serve customer area k	i
K_i	Customer areas that can be served from facility i	k
E_k	Carriers that can serve customer area k	e
L_1	Capacity levels using contracts with carriers in stage 1	l
L_2	Capacity levels using contracts with carriers in stage 2	l
S^t	Demand scenarios in period t	s

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	Capacity commitment with carrier e in period t (in load/cargo units)
Q_l	Service capacity at level l (in load/cargo units)
$F_{i,l}$	Service costs for shipping to facility i at capacity level l at reservation costs
$\bar{F}_{i,l}$	Service costs for shipping to facility i at capacity level l with penalty costs
$D_k^{t,s}$	Demand of customer area k in period t and scenario s (in cargo units)
$U_{i,k}$	Parcel delivery costs (per unit of load) from facility i to area k
$p^{t,s}$	Probability of scenario s in period t

First stage decision variables

Symbol	Description
α_e^t	1 if a contract with carrier e starts at period t , 0 otherwise.
$v_{i,l}^t$	1 if shipping lane to facility i is reserved at capacity level l (from a contracted carrier) in period t , 0 otherwise.

Second stage decision variables

Symbol	Description
$w_{i,l}^{t,s}$	1 if shipping lane to facility i is used without reservation at capacity level l in period t in scenario s , 0 otherwise.
$q_{i,k}^{t,s}$	Demand of customer area k distributed from facility i in period t and scenario s (in load/cargo units)
$c_e^{t,s}$	Total payment due to carrier e at period t and scenario s

Table 6.1: Mathematical notation for the *mDNP_MCC* model

Objective Function:

$$\min \sum_{t \in T} \sum_{s \in S^t} p^{t,s} \left[\sum_{e \in E} c_e^{t,s} + \sum_{k \in K} \sum_{i \in I_k} U_{i,k} q_{i,k}^{t,s} \right] \quad (6.1)$$

Constraints

$$\sum_{i \in I_k} q_{i,k}^{t,s} = D_k^s \quad \forall k \in K, t \in T, s \in S^t \quad (6.2)$$

$$q_{i,k}^{t,s} \leq D_k^{t,s} \left(\sum_{l \in L_1} v_{i,l} + \sum_{l \in L_2} w_{i,l}^{t,s} \right) \quad \forall i \in I, k \in K, t \in T, s \in S^t \quad (6.3)$$

$$\sum_{k \in K_i} q_{i,k}^{t,s} \leq \sum_{l \in L_1} Q_l v_{i,l} + \sum_{l \in L_2} Q_l w_{i,l}^{t,s} \quad \forall i \in I, t \in T, s \in S^t \quad (6.4)$$

$$\sum_{i \in I_e} \sum_{l \in L_1} Q_l v_{i,l}^t \geq M_e^t \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall e \in E, t \in T \quad (6.5)$$

$$\sum_{l \in L_1} v_{i,l}^t \leq \sum_{n=0}^{H_e-1} \alpha_e^{t-n} \quad \forall i \in I_e, e \in E, t \in T \quad (6.6)$$

$$\sum_{n=0}^{H_e-1} \alpha_e^{t-n} \leq 1 \quad \forall e \in E, t \in T \quad (6.7)$$

$$\sum_{l \in L_2} w_{i,l}^{t,s} \leq 1 \quad \forall i \in I, t \in T, s \in S^t \quad (6.8)$$

$$c_e^{t,s} = \sum_{l \in L_1} \sum_{i \in I_e} F_{i,l} v_{i,l}^t + \sum_{l \in L_2} \sum_{i \in I_e} \bar{F}_{i,l} w_{i,l}^{t,s} \quad \forall e \in E, t \in T, s \in S^t \quad (6.9)$$

$$\alpha_e^t \in \{0, 1\} \quad \forall e \in E, t \in T \quad (6.10)$$

$$v_{i,l}^t \in \{0, 1\} \quad \forall i \in I, l \in L_1, t \in T \quad (6.11)$$

$$w_{i,l}^{t,s} \in \{0, 1\} \quad \forall i \in I, l \in L_2, t \in T, s \in S^t \quad (6.12)$$

$$q_{i,k}^{t,s} \geq 0 \quad \forall k \in K, i \in I, t \in T, s \in S^t \quad (6.13)$$

The objective function (Eq. (6.1)) in the MILP minimizes the expected total cost of transportation services for distribution and delivery of parcels, during the whole planning horizon. The two components of the cost function account for expectation of total payment to carriers in separation to parcel delivery charges. The stochastic formulation takes into account probabilities of time-dependent scenarios of demand.

The set of constraints in Eq. (6.2) guarantee complete order fulfilment for all customer areas in every period, independent of the demand scenario. Constraints in

Eq. (6.3) enable delivery only from facilities used for cross-docking. Constraints in Eq. (6.4) ensure that the total transportation capacity entering a cross-docking facility covers the orders delivered from that facility. Constraints in Eq. (6.5) force the minimum capacity commitment with a contracted carrier. Constraints in Eq. (6.6) express that capacity reservation in the first stage is allowed if there is a valid contract, in which case at most one capacity level is selected per shipping lane. Constraints in Eq. (6.7) make sure that at most one contract is ongoing with any carrier at any time. Constraints in Eq. (6.8) specify at most one capacity level per lane in the second stage. Constraints in Eq. (6.9) determine the payment due to a carrier for services provided on and off contracts. Finally, constraints in Eqs. (6.10)-(6.13) specify the range and nature of variables.

A feasible solution $\hat{s} = (\hat{\alpha}, \hat{v}, \hat{q}, \hat{w}, \hat{c})$ of the problem can be broken down in two managerial perspectives. At the broader *tactical* perspective, a feasible solution for decision variables α_e^t determines the contracts schedule with selected carriers along T . A candidate solution ($\hat{\alpha} = \hat{\alpha}_e^t : e \in E, t \in T$), is therefore denominated the *contract plan*. Likewise, a solution ($\hat{v} = \hat{v}_{i,l}^t : i \in I, l \in L_1, t \in T$) determines the *capacity reservation* plan with each carrier under contract.

Any contract plan ($\hat{\alpha}$) that conforms with constraints (6.7), and which satisfies conditions about the duration of contracts (H_e), provides a feasible solution for the *tactical* problem. An open contract injects directly some level of capacity available with carriers, in their respective networks. Clearly, a feasible solution $\hat{\alpha}$ has an effect on decisions made at a more operational level. In addition, a feasible solution \hat{v} , defines the capacity reserved in certain lanes with contracted carriers $E^t \subseteq E$, where E^t is the subset of carriers with ongoing contracts in period t . The feasibility of a capacity reservation plan is linked with constraints (6.5) and (6.6) that force capacity reservation to reach at least the commitment level M_e with carriers in E^t .

After demand realization, decisions at the *operational* level are recourse actions, which are dependent on the scenario of demand. A feasible solution for the capacity requested off-contract ($\hat{w} = \hat{w}_{i,l}^{t,s} : i \in I, l \in L_2, t \in T, s \in S^t$) as a consequence of shortage, is decided upon scenario $s \in S^t$ and reservation plan \hat{v} . Similarly, the *last-mile distribution plan* to customer areas ($q_{i,k}^{t,s}$) is configured according to the capacity level available in the set of lanes I (given by \hat{v} and \hat{w}) at period t and under scenario s .

6.4 Algorithmic Development

From a computational point of view, an MIP solver can make use of built-in general purpose algorithms to obtain an optimal solution of the mathematical model for any instance of the problem. However as the size of the instances grows, this results in prohibitively long computing times, as we will see in the sections reporting on numerical experiments. Therefore, it becomes useful to develop *taylor-made* heuristic methods adapted for the *mDNP* problem, that lead to solve large instances in reasonable time. In this section, the algorithmic methods implemented to solve the *mDNP* problem are described. The general idea behind the proposed heuristics is to reduce the difficulty of the complete model by relaxing some of its constraints or assumptions. Two main strategies based on well-known methodologies are applied in this context.

A first approach is to apply a *rolling horizon procedure* (RHP). Rolling horizon decision-making is a common managerial practice for making decisions in a dynamic stochastic environment [Sethi & Sorger, 1991]. It is commonly applied to multi-period planning problems in practical contexts. It consists in reducing discrete-time problems to smaller sub-problems with shorter time horizons. RHP procedures are used for real-time applications where data is updated in continuous basis for future parameters, but can also be viewed as providing a meta-heuristic framework for the solution of multi-period problems. This is the point of view that we shall adopt.

A second approach is based on *relax-and-repair heuristics*. These kind of heuristics have been used widely in MILP models with applications in different areas like production planning, lot-sizing and scheduling problems. The general purpose is to decompose the monolithic model into easier to solve sub-models in order to obtain good feasible solutions [Nabil & van den Heuvel, 2019] in a decreased amount of time. The effectiveness highly depends on the way the problem is decomposed. Three heuristics methods will be developed using different relaxations of the model.

Let now introduce mathematical notation for the methods described in this section. In order to represent the decomposition structure of the model, an auxiliary decision variable $\beta_e^t \in \{0, 1\}$ is included in the model described in Section 6.3, in order to indicate the *status of contracts with a carrier e in period t* (similar to the transformation performed in Section 4.5). The variable validates or not a contract according to the following relationship :

$$\sum_{n=0}^{H_e-1} \alpha_e^{t-n} = \beta_e^t \quad \forall e \in E, t \in T \quad (6.14)$$

$$\beta_e^t \in \{0, 1\} \quad \forall e \in E, t \in T \quad (6.15)$$

If the binary variable β_e^t takes value 1, it means that an MCC contract with carrier e is in effect in period t , otherwise, it takes value 0. Replacing Eq. (6.7) in the original formulation by Eqs. (6.14, 6.15) and then replacing $\sum \alpha_e^{t-n}$ by β_e^t in Eqs. (6.5) and (6.6) yields an alternative equivalent model *mDNP*. The proposed methods make use of sub-models stemming from the alternative equivalent model, after relaxing or removing some constraints and/or fixing some decision variables to a given value. The equivalent *mDNP* formulation is also referred to as the *complete* model.

Before introducing the algorithmic methods, the distinction between sub-models is summarized as follows:

DNP(W^t): *Distribution network problem for the planning window W starting from period t.*

Model resulting from shortening the planning horizon T to W^t where t indicates the first planning period. Depending on the size of W^t , the resulting sub-model may account for a single or multiple consecutive periods.

sDNP(t): *Single-period distribution network problem for period t.*

Model resulting from applying a relaxation on the duration of contracts, as if $H_e = 1$ instead of their real duration. It converts the *mDNP* into multiple single-period problems, *sDNP(t)* for every $t \in T$. In this sub-model, variables α_e^t become unnecessary as their value can be deduced from the value of β_e^t . Therefore, equations (6.14) and (6.10) can be removed.

mRDNP: *Multi-period relaxed distribution network problem.*

Model resulting from relaxing integrality conditions for off-contract capacity variables (Eq. (6.12)) as if $w_{i,l}^{t,s}$ were continuous instead of binary. The *mRDNP* model is yet not decomposable per period.

sRDNP(t): *Single-period relaxed distribution network problem for period t.*

Model resulting from relaxing both contract lengths ($H_e = 1$) and off-contract capacity variables ($w_{i,l}^{t,s} \leq 0$). The *mDNP* converts into multiple single-period relaxed problems - *sRDNP(t)* for every $t \in T$. Here again, constraints related to α_e^t (6.14), (6.10) are not necessary.

The following sub-models come from fixing values of the first-stage variables ($\beta_e^t, v_{i,l}^t$):

$\hat{\beta}$ specifies a complete solution for the status of contracts on the planning horizon T . More specifically $\hat{\beta}^t$ is the solution values for period t . Likewise, \hat{v} specifies a feasible and complete solution for the reservation plans on T , in accordance with $\hat{\beta}$, and \hat{v}^t is concretely for period t .

$sDNP(t, \hat{\beta}^t)$: *Single-period distribution network problem for period t given contract plan $\hat{\beta}$.*

Model resulting from obtaining a feasible solution $\hat{\beta}$ for first-stage variables β_e^t . The fixation of contract variables according to $\hat{\beta}$ converts the $mDNP$ model into multiple single-period problems $sDNP(t, \hat{\beta}^t)$. A solution for variables (v) in the first stage and second stage (w, q, c) are yet to determine for each period $t \in T$.

$sDNP(t, s, \hat{\beta}^t, \hat{v}^t)$: *Single-period distribution network problem for period t and scenario s , given contract plan $\hat{\beta}$ and capacity reservation plan \hat{v} .*

Model resulting from obtaining a feasible solution $\hat{\beta}$ and \hat{v} , for first-stage variables β_e^t and $v_{i,l}^t$. The resulting sub-models are associated to a certain demand scenario $s \in S^t$ and period $t \in T$. Recourse variables (w, q, c) can be separately determined for each sub-model.

6.4.1 Rolling Horizon Procedure - (RHP)

The RHP strategy consists in reducing the planning horizon T with length n to a shorter planning window $W \subset T$ of fixed size r ($r < n$). For every period t , the time window W rolls over and accounts for periods in $\{t, t+1, \dots, t+r-1\} = W^t$ (indexed τ). This yields a sub-model with the shortened time horizon, aiming at minimizing the total expected costs.

After solving the sub-problem associated with time window W^t denoted as $DNP(W^t)$, the solution associated with variables in the first stage ($\hat{\beta}, \hat{\alpha}, \hat{v}$) and second stage ($\hat{q}, \hat{w}, \hat{c}$) for period t is put into effect, while the solution obtained for the next periods is discarded. However, the validity of contracts implied by decisions made in periods $1, \dots, t$ is enforced by fixing the corresponding variables to their appropriate values. if the solution value for variable α_e^t is 1, then the values for $\beta_e^{t+1}, \dots, \beta_e^{t+H_e-1}$ are set to 1 in the subsequent sub-problems.

More specifically, the sub-problem for period t can be denoted as $DNP(W^t, E^\tau)$ where $E^\tau \subseteq E$ is the subset of carriers with ongoing contracts in periods $\tau \in W^t$. Note that the sub-model $DNP(W^t, E^\tau)$ is not decomposable, since time-linking constraints (in Eq. (6.14)) still apply. Given an optimal solution of $DNP(W^t, E^\tau)$, we

denote as Z^t the cost incurred in period t under this optimal solution (that is, the value of the cost function (6.1) corresponding to period t).

The parameter r enters as an input value in the algorithm to configure the length of the planning window. This value is fixed for every sub-model $DNP(W^t, E^\tau)$. During the algorithmic process, the time window unfolds one period at a time and the subset E^τ is continuously updated for all τ in the time window W^t . The information about the remaining validity of an ongoing contract at time t needs to be passed for every new $DNP(W^t, E^\tau)$. Therefore, variables α_e^{t-n} for $n \in \{0, \dots, H_e - 1\}$ are fixed to their respective values given in preceding models. These are also required to account for constraints in Eq.(6.15). The structure of the algorithm is described in the following pseudocode (Algorithm 5).

Algorithm 5: Rolling Horizon Procedure

Input: r : Time window length.
 E^τ : Subset of carriers with contracts initiated before period 1 and still ongoing at period τ , for $\tau = 1, \dots, H_e - 1$.

```

1  $t \leftarrow 1$ ,  $Z_{RHP} \leftarrow 0$ ,  $W^t \leftarrow \{1, \dots, r\}$ 
2 while  $t \leq N$  do
3   for  $\tau \in W^t$ ,  $e \in E^\tau$  do
4     fix variable  $\beta_e^\tau = 1$ 
5   solve  $DNP(W^t, E^\tau)$  to optimality
6    $Z_{RHP} \leftarrow Z_{RHP} + Z^t$ 
7   retrieve the solution values for  $\alpha_e^t$ ,  $e \in E$  (stage 1)
8   for each  $e \in E$  do
9     if  $\alpha_e^t = 1$  then
10      for  $s = t + 1, \dots, t + H_e - 1$  {/*iterates until the expiration period} do
11         $E^s \leftarrow E^s \cup \{e\}$ 
12    $t \leftarrow t + 1$ 

```

The downside of the RHP comes from disregarding information beyond the time window (W^t). This may yield sub-optimal solutions in comparison with taking into account the complete planning horizon T . In the $mDNP$, this is equivalent to ignore effects of the demand from time $t + r$ in the sub-problem $DNP(W^t, E^\tau)$ at time t . As an advantage, shortening the number of periods allows to reduce the input data involved in each sub-model. Therefore, even if there are several sub-models to solve

in sequence (as many as the number of periods n) the total computational time to find an optimal solution is expected to be reduced overall.

In managerial terms, decisions at the tactical level about opting for new contracts (α_e^t) have mid-term effects since they have to remain open until completion. Operational decisions, such as, the reserved capacity $v_{i,l}^t$ and the recourse actions ($w_{i,l}^{t,s}$, $q_{i,k}^{t,s}$) are taken accordingly, but do not have mid-term effects. The solutions obtained for periods $\{t+1, \dots, t+r-1\}$ at time t are subject to modifications as the rolling horizon unfolds, but serve as a guide to look at potential actions into the future. The RHP could be considered as a *myopic* strategy with reduced visibility of future events and the window length r can be adapted according to the required time to find an optimal or close-enough solution, since large values of r may require a significant computing time.

6.4.2 Contract Length Relaxation - (CLR)

In this first relax-and-repair heuristic, MCC contracts are considered to last for a single-period ($H_e = 1, \forall e \in E$). In that way, the complete model is decomposable per period and handled as a sequence of independent two-stage stochastic problems ($sDNP(t), \forall t \in T$). The description of the algorithm is presented in pseudocode (Algorithm 6).

In the *first phase*, each single-period problem $sDNP(t)$ is solved iteratively to optimality. The optimal value $Z(sDNP(t))$ is aggregated $\sum_{t \in T} Z(sDNP(t))$ to generate a valid lower bound LB_1 for the $mDNP$. In each period a solution for variables β_e^t is obtained (i.e., $\hat{\beta}^t$). The complete contract plan $\hat{\beta}$ is still to be checked for feasibility and adjusted in case of rupture of contracts.

In the *second phase* the feasibility checking procedure iterates through every period starting from the second (in the first period is not possible to determine infeasibility of a contract plan solution). At each period $t \in T$, the algorithm goes through every carrier $e \in E$ and checks if a contract has been initiated but not continued as required. This takes place whenever an interrupted contract is detected (i.e., $\hat{\beta}_e^t = 0$ but $\hat{\beta}_e^{t-1} = 1$). The elapsed duration h of the contract is computed to verify its feasibility (h should be a multiple of H_e).

The repairing procedure consists in extending unfinished contracts for 1 period. When the algorithm detects an unfinished contract, it is enforced to continue for period t (set $\hat{\beta}_e^t = 1$). After checking for all carriers $e \in E$, if any extensions were

Algorithm 6: Contract Length Relaxation

```
/* Phase 1: Single-Period Problems */
1  $Z_{CLR} \leftarrow 0$ 
2 for each  $t \in T$  do
3   solve  $sDNP(t)$ 
4   retrieve  $Z(sDNP(t)), \hat{\beta}^t$ 
5    $Z_{CLR} \leftarrow Z_{CLR} + Z(sDNP(t))$ 
6  $LB^{CLR} \leftarrow Z_{CLR}$ 

/* Phase 2: Repairing Procedure */
Input:  $\hat{\beta}$ : Relaxed contract plan for the planning horizon
7 for each  $t \in T, t \neq 1$  do
8   for each  $e \in E$  do
9     if  $\hat{\beta}_e^t = 0$  and  $\hat{\beta}_e^{t-1} = 1$  then
10      compute  $h$  {/*  $h$ : number of consecutive periods with
11                 $\hat{\beta}_e^{t-h} = \dots = \hat{\beta}_e^{t-1} = 1$  */}
12      if  $h \bmod H_e \neq 0$  then
13         $\hat{\beta}_e^t \leftarrow 1$  {/* Set values to complete contract */}
14      if contract infeasibility then
15        solve  $sDNP(t, \hat{\beta}^t)$ 
16        update  $Z_{CLR}$ 
```

made, the sub-model $sDNP(t, \hat{\beta}^t)$ is solved taking into account the modified contract plan $\hat{\beta}^t$.

The CLR heuristic is a simpler version of the Relax-and-repair heuristic described in Chapter 4 for the MDPC problem. While the relaxation is made in the same terms, the repairing phase here, only chooses to complete contracts to comply with initial conditions for H_e , but does not consider the option to reverse this choice. The decision to disregard contract removals comes first from results in preliminary tests wherein the repairing procedure yielded better costs by completing instead of removing the unfinished contract. A second advantage is the acceleration of the repairing procedure given that only one sub-problem $sDNP(t, \hat{\beta}^t)$ is solved associated with infeasibilities found in period t . Since the focus of the solution methodology is in comparing algorithmic methods performance, the simple CLR heuristic appears to provide the best cost vs. time trade-off for our purpose.

6.4.3 Off-contract Capacity Variables ($w_{i,l}^{t,s}$) Relaxation - (WR)

In this section, we describe a second relax-and-repair heuristic based on a different relaxation of the complete formulation. The set of constraints in Eq. (6.12) is relaxed to the continuous form of variables $w_{i,l}^{t,s}$

$$w_{i,l}^{t,s} \geq 0 \quad \forall i \in I, l \in L_2, t \in T, s \in S_t \quad (6.16)$$

In that form, recourse variables regarding off-contract capacity $w_{i,l}^{t,s}$ are modelled as continuous. As a consequence, the problem consists in solving sub-model *mRDNP*. It remains a multi-stage stochastic problem with time-linking constraints, which is not decomposable per period. It is clear that the optimal value for the *mRDNP* is at most as high as the optimal value for the complete formulation. Hence, a lower bound can be derived if the optimal value is obtained (LB^{WR}). The computational effort taken to solve the *mRDNP* model is expected to be reduced compared to the complete *mDNP* model, while the loss in quality with respect to the optimal solution requires to be measured.

After resolution of the *mRDNP*, the solution values for variables $w_{i,l}^{t,s}$ need to be translated into integer values to conform with appropriate levels of capacity. For that matter, a repairing procedure is implemented. The WR algorithm is described in pseudocode (Algorithm 7).

Algorithm 7: Off-Contract Capacity Relaxation

```

/* Phase 1: Multi-Period Problem */
1 solve mRDNP
2 retrieve  $Z(\textit{mRDNP}), \hat{\beta}$ 
3  $LB^{WR} \leftarrow Z(\textit{mRDNP})$ 

/* Phase 2: Repairing Procedure */
Input:  $\hat{\beta}$  Contract plan for the planning horizon
4  $Z_{WR} \leftarrow 0$ 
5 for each  $t \in T$  do
6   solve  $sDNP(t, \hat{\beta}^t)$ 
7    $Z_{WR} \leftarrow Z_{WR} + Z(sDNP(t, \hat{\beta}^t))$ 

```

In order to repair the infeasibility in constraints (6.12), the solution obtained for the contract status variables ($\hat{\beta}$) is retrieved from the first phase and used as in-

put parameters for the repairing procedure. Integrality conditions (Eq. (6.12)), for variables $w_{i,l}^{t,s}$ are reinserted.

After variables β_e^t are fixed, the sub-model is now decomposable per period. The $sDNP(t, \hat{\beta}^t)$ addresses the cost-minimization problem for the sequence of two-stage distribution network models for every period t . Note that decision variables (v, w, q) , including first-stage variables $v_{i,l}^t$, are reoptimized by solving $sDNP(t, \hat{\beta}^t)$. The aggregation of the optimal values of $sDNP(t, \hat{\beta}^t)$ equals the total value obtained with the WR heuristic method (Z_{WR}).

6.4.4 Contract Length and $w_{i,l}^{t,s}$ Variables Relaxation - (CLWR)

In this section, we describe a third relax-and-repair heuristic method that mixes both relaxations previously presented. Hence, the contract is assumed valid for a single-period ($H_e = 1$ for every carrier), while $w_{i,l}^{t,s}$ variables are continuous. Under these assumptions, the single period contract allows the decomposition of the complete formulation per period. The single-period problems $sRDNP(t)$ for all $t \in T$, are solved in sequence. The algorithmic procedure finds optimal solutions for the $sRDNP(t)$, thus producing tentative values for the variables β, v, w and q .

Similar to the CLR method (*Heuristic 1*), the relaxed contract plan $\hat{\beta}$ needs to be checked to verify it conforms with the actual contracts durations H_e . Therefore, in the repairing procedure contracts are forced to remain in effect until their expiration time. In addition, variables w have to be transformed into their integer form. The algorithm is summarized in pseudocode (Algorithm 8).

The repairing procedure uses the feasibility checking step described previously in the CLR method (*Heuristic 1*). It scans through the relaxed contract plan obtained from the first phase. At every occurrence of an inappropriately interrupted contract, it extends its validity for the current period t and the contract status variable β_e^t is set to 1. If a contract infeasibility is detected, the $sDNP(t, \hat{\beta}_t)$ is solved taking into account the modified relaxed contract plan ($\hat{\beta}_t$) and enforcing integrality conditions for variables $w_{i,l}^{t,s}$. Solving to optimality the $sDNP(t, \hat{\beta}_t)$ model returns integer values for the first-stage variables $v_{i,l}^t$ and the second-stage variables $w_{i,l}^{t,s}$, as well as a feasible solution for variables $q_{i,k}^{t,s}$.

In comparison with the model described in Section 6.4.2, the first phase deals with the $sRDNP$ models which returns clearly a less tight lower bound (LB^{CLWR}) than the one obtained by solving the $sDNP$ models (LB^{CLR}). This comes with the benefit of

Algorithm 8: Contract Length and Off-Contract Capacity Relaxation

```
/* Phase 1: Single-Period Problems */
1  $Z_{CLWR} \leftarrow 0$ 
2 for each  $t \in T$  do
3   solve  $sRDNP(t)$ 
4   retrieve  $Z(sRDNP(t)), \hat{\beta}^t$ 
5    $Z_{CLWR} \leftarrow Z_{CLWR} + Z(sRDNP(t))$ 
6  $LB^{CLWR} \leftarrow Z_{CLWR}$ 

/* Phase 2: Repairing Procedure */
Input:  $\hat{\beta}$ : Relaxed contract plan for the planning horizon
7 for each  $t \in T, t \neq 1$  do
8   for each  $e \in E$  do
9     if  $\hat{\beta}_e^t = 0$  and  $\hat{\beta}_e^{t-1} = 1$  then
10      compute  $h$  {/*  $h$ : number of consecutive periods with
11                 $\hat{\beta}_e^{t-h} = \dots = \hat{\beta}_e^{t-1} = 1$  */}
12      if  $h \bmod H_e \neq 0$  then
13         $\hat{\beta}_e^t \leftarrow 1$  {/* Set values to complete contract */}
14      if contract infeasibility then
15        solve  $sDNP(t, \hat{\beta}^t)$ 
16        update  $Z_{CLWR}$ 
```

shorter computational time, which might affect the effectiveness or applicability of a certain heuristic. The computational experiments shed lights on this trade-off.

6.5 Experimental Design

The experimental study intends to evaluate and compare the computational performance of several methods including standard *off-the-shelf* practices and tailor-made heuristic approaches applied to the *mDNP* problem. Two performance criteria are taken into account to assess the efficiency and quality of alternative methods.

From one point of view, the **completion time** is registered to measure the algorithm's efficiency and capacity to finish within the testing time (3 hours). At the same time, the **relative gap** (with respect to the best known lower bound) is com-

puted as an indicator of the quality of the solution provided by a certain algorithm in the time limit. To account for data size effects in the performance metrics, computational tests were performed on different instance sizes of the *mDNP* problem. The classification by categories is defined in our experimental design with different values of input parameters (see Table 6.4 further).

In the following, we introduce the notation and characterization of algorithmic methods and performance metrics as presented later in the numerical results. Then, the experimental design applied in order to generate instances of the problem is described. Finally, results are presented in the form of performance profiles. In real-life contexts decision makers are interested in the applicability of the methods with restricted decision-making time, for instance in a weekly or daily periodic planning agenda. Analysis and comparison of results in performance metrics and insights from a managerial standpoint are laid down at the end of the section.

6.5.1 Algorithms

In this section, we summarize the solution strategies tested in the performance analysis. Algorithmic methods are classified into two groups so as to better represent computational results and to conduct further analysis separately.

The first group considers methods which have in common to apply, in some way, general-purpose algorithms for MIP problems. In that group, a first approach consists in solving the complete stochastic *mDNP* model by applying the B&C algorithm with default settings embedded in a commercial solver in our case CPLEX 12.8.

A second approach is to use the rolling horizon procedure RHP, iteratively per planning period using the same default B&C algorithm. Preliminary tests were made with different time window sizes ($r = 1, 2, \dots, H_\ell$), showing that the computational time was significantly affected to solve the sub-model related with time t ($DNP(W^t, E^\tau)$), but the best-found value Z^t did not improve or even resulted in an overall worse solution Z_{RHP} for longer time windows. This is associated with the fact, that $DNP(W^t, E^\tau)$ models with longer time windows take more time to be solved, therefore, their optimal value is not completely achieved within the testing time limits, affecting the quality of the best-found solution obtained. In view of the initial tests and for comparison analysis with algorithmic methods of different nature, we select a single-period window ($r = 1$), which runs the fastest and appears to provide as good quality solutions as the rest of the time windows r . From the man-

agerial standpoint, this implies that tactical decisions are based only on the demand information for the next selling period.

A third standard approach consists in solving the complete *mDNP* problem with deterministic (expected) values of the demand. This deterministic approach solves a simplified model with a single demand scenario in the second stage, which might increase the computational capacity to solve large instances. In any case, it is assumed that the underlying distribution of the demand is well represented by a discrete probabilistic function, following the description in Section 5.3 and a finite number of scenarios are possible. Therefore, the solution given by a deterministic approach is assessed ex post by computing its expected cost. Both, the RHP and deterministic approach can be considered as heuristics disregarding certain aspects of the model and providing upper bounds to the optimal value.

Standard methods are summarized in Table 6.2.

Name	Method
STO	Stochastic model solved by CPLEX branch & cut by default.
RHP	Rolling horizon procedure with single time window, $r = 1$
EVD	Deterministic (expected demand) model solved by CPLEX branch & cut by default.

Table 6.2: Standard methods

The second group involves **tailored** heuristics based on the decomposition structure of the *mDNP* problem. The group includes relax-and-repair heuristics described in Section 6.4, namely, contract length relaxation (CLR), off-contract capacity variables relaxation (WR) and double relaxation (CLWR). These methods have in common their separation in two phases and their decomposition of the sub-problem in the second phase. CPLEX is also used as the optimization software to deal with the sub-problems in each heuristic algorithm.

Relax-and-repair heuristic methods are summarized in Table 6.3.

Name	Method
CLR	Contract length relax-and-repair heuristic
WR	Integer variables ($w_{i,l}^{f,s}$) relax-and-repair heuristic
CLWR	Contract length and Integer variables ($w_{i,l}^{f,s}$) relax-and-repair heuristic

Table 6.3: Relax-and-repair heuristics

For notation we refer to set A as the group of algorithms mentioned above (i.e., $A = \text{STO, RHP, EVD, CLR, WR, CLWR}$) and to a given algorithm $a \in A$. An alternative separation of algorithms A can be made between exact and approximate methods. From the six algorithmic methods, only the default B&C solver (STO) guarantees to return the optimal value of $mDNP$. The remaining methods find valid and feasible solutions (hence, upper bounds on the optimal value) but are unable to guarantee optimality. However, the feasibility checking step mentioned in some relax-and-repair heuristics, applied to the relaxed solution may prove optimality in case it satisfies all model constraints. On the other hand, a valid lower bound can be retrieved either by the CPLEX's B&C algorithm or by relax-and-repair heuristics from their relaxation in the first-phase. The tightest (largest) lower bound is selected as a reference to compare upper bounds given by each of the methods.

6.5.2 Performance measures

The performance metrics are described and computed in the following terms:

- *Completion Time (F^a)* : The computational time is limited to 3 hours for every instance. As a result two possibilities may occur, either algorithm a manages to finish before the time limit or not. In the first case, the complete stochastic model (STO) method returns the optimal value while in the second case, it does not prove optimality. Regardless of the situation, the best-solutions found in the testing time can be compared in terms of quality. In addition, the completion rate, defined as the percentage of instances terminated by algorithm a sheds light on how fast an heuristic/exact method is able to find a solution.
- *Relative Gap Percentage (G^a)* : The gap is computed as the relative difference between the best value found by algorithm a - (Z^a), and the best known lower bound (LB_{best}). This difference is shown in percentage relative to best found value. The determination of the lower bound is computed by finding the tightest LB from either one of the methods that return a valid lower bound: $LB_{best} = \max\{LB^{STO}, LB^{CLR}, LB^{WR}, LB^{CLWR}\}$. For every instance the tightest lower bound is used in the calculations of the relative gap for algorithm a : $G^a = \frac{Z^a - LB_{best}}{Z^a} * 100$, where Z^a is the best solution value obtained by method a .

The LB given by CLWR (LB^{CLWR}) is not considered due to its double relaxation. Indeed, the optimal value for the $sDNP(t)$ in the CLR method is tighter than the one

for the $sRDNP(t)$ in the CLWR method. On top of that, the solution of the $mRDNP$ model with time-linking constraints in the WR method also offers a tighter bound than the complete solution given by $sRDNP(t) : \forall t \in T$. Therefore, for comparison purposes only the CLR and WR heuristics are considered.

6.5.3 Instance Parameters

The algorithms described in Section 6.5.1 were tested on a variety of random instances. Instances are classified depending on the value of input parameters, more specifically length of the planning horizon, number of candidate carriers, number of customer areas and number of logistics services (transportation lanes/cross-docking facilities) per carrier ($|T|, |E|, |K|, |I_e|$).

The classification with respective parameter values is detailed in Table 6.4.

Category	$ T $	$ E $	$ K $	$ I_e $
Small	8	6	50	3
Medium	10	8	100	3
Large	10	12	200	4

Table 6.4: Classification of datasets per size

Input values regarding the demand of customer areas in their respective scenarios $D_k^{t,s}$ were generated according to the following specifications:

- We assume the shipper owns two e-fulfilment warehouses for inventory of products sold by multiple clients.
- Customer areas are categorized in three different sizes. Thus, the expected value of demand of products changes from one to another customer area according to their size.
- The seasonality and variability of the demand is represented by two demand seasons (High, Low) fluctuating dynamically along the time horizon. Every period is predetermined to either one of the two possible demand seasons. The demand values in a high-demand seasons are twice as much larger than in low-demand seasons.

- At every period, there are two possible demand scenarios in $S^t = \{s^1, s^2\}$ with a given probability. Hypothetically, in the first scenario (s^1) customers' purchasing behaviour is based on forecast, which is more likely to occur. The second scenario (s^2) is generated from a potential situation where customer orders amount to larger-than-predicted values by a certain margin. The probability of scenario s^1 is equal to 0.85, while for scenario s^2 it is 0.15.
- The quantities demanded in scenario s^2 are twice as much those demanded in scenario s^1 . The differences in scenario s^2 are applied to all customer areas without exception. Scenario s^2 is considerably less likely to occur but likely enough to affect the expected distribution costs in the long-term. Meaning that random effects on the purchasing habits affect all customer areas at the same time. In other words, it is supposed that the difference in the *total* demand between s^1 and s^2 , is distributed evenly among all customers.

Parameters about contract conditions (H_e, P, M_e^t) and extra-capacity costs ($\bar{F}_{i,l}$) were configured according to the following criteria :

- The duration of contracts H_e is set to 3 periods for each carrier e .
- The minimum commitment level in load units (M_e^t) is established separately for each carrier e and period t , computed as a proportion of the maximum capacity that can be requested in a scenario s^1 of forecast demand. This proportion is set to 15% and remains fixed in all periods. Due to variability of the demand, the minimum commitment level varies accordingly.
- The penalty cost P for capacity after reservation is set to 50% of the rates paid under an MCC contract. In other terms, services off-contract are paid 1.5 times higher than contract costs. (i.e., $\bar{F}_{i,l} = 1.5 F_{i,l}$)
- Contractual services costs $F_{i,l}$ were generated following a similar procedure described in Section 4.6.1 based on the Euclidean distance of two randomly generated points in a square matrix of dimension [1000 x 1000]. The fixed costs for facility utilization is disregarded. The cost for operating at level one, i.e., $F_{i,1}$, is twice the Euclidean distance from facility i to the closest warehouse.
- Similarly, 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 .

6.6 Computational Results

6.6.1 Lower bounds

As it was mentioned, a valid lower bound (LB) can be obtained either by solving the stochastic model by CPLEXs B&C (STO algorithm) or by one of the relax-and-repair methods (CLR, WR) in their first 'relaxed' phase. In case CPLEX fails to prove optimality within the time limit, the tightest (max. value) lower bound is used to measure the relative gap and to compare the solution quality of algorithmic methods. As a general rule when the sub-model in the first relaxed phase fails to produce the optimal solution, is not taken into account to determine LB_{best} .

In order to provide insights on the tightness of relaxed models and its closeness with the optimal value, a comparison is made between the lower bounds obtained by each method in every instance. The bar chart in Figure 6.2 represents the percentage of instances for which a method gets the best lower bound.

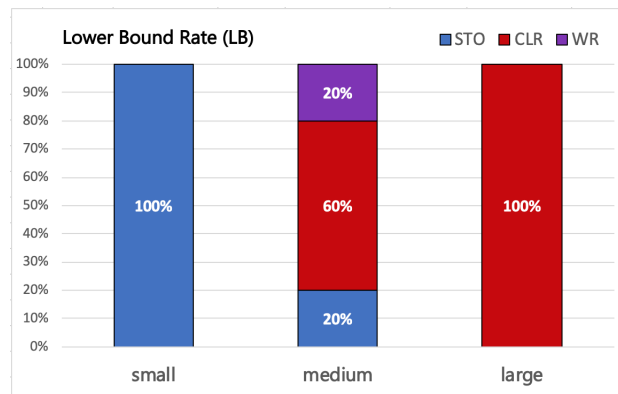


Figure 6.2: Comparison of Lower bounds among algorithmic methods

According to the results in Figure 6.2 the best lower bound varies according to the instance category. For the small size category, the STO method returns the best lower bound in every instance tested. In fact, for 20% of small instances it achieves the optimal value (Figure 6.3 bottom-right side), while for the remaining 80% the relative gap is lower than 0.24%. However, the outlook differs when instances get bigger.

For the medium size category, results are divided. Only 20% of times CPLEX gives

the best lower bound while having obtained the optimal solution in only $\approx 7\%$ of cases. The WR method achieves the best lower bound for a similar percentage (20%) out of 47% of times that the relaxed model *mRDNP* terminates. Thus, sometimes solving the *mRDNP* problem yields to tighter lower bound in the medium size category. The relaxed models *sDNP*(t) in the CLR method terminate always in the time limit and return the best lower bound for most of instances (60%). Thus, the single period models provide an even tighter value. For the large size category, the tendency leans clearly towards the CLR relaxation which gets the best lower bound in every instance, while the WR relaxation is unable to terminate on time, and the STO model obtains weaker values.

6.6.2 Algorithmic Performance

The focus of this chapter is to display computational results of algorithmic methods regarding performance metrics measuring the efficiency (F^a) and solution quality (G^a). Following the classification of algorithms in Section 6.5.1, results are presented first for standard methods (STO, EVD, RHP) followed by tailored heuristics (CLR, WR, CLWR). As a benchmark between both strategies, the RHP is included in all comparative graphs.

Graphs are presented in the form of performance profiles, which are cumulative distributions of performance metrics (F^a , G^a) based on the empirical results. More precisely, for an algorithm a , a criterion (say F^a), and a value x on the horizontal axis, the performance profile indicates the percentage of instances for which $F^a \leq x$. In terms of the computational time F^a , the maximum value (upper limit - U) of x is the testing time (3 hrs), while for the percentage gap G^a , the maximum value of x is 4%. The profiles allow for easy visualization and comparison of the performance of different algorithms over a range of instances (in our experimental design, the collection of 15 small/medium instances and 10 large instances).

Note that if an algorithm a generates performance values lower than U for all instances, is possible to notice the worst-case value x . Otherwise, is possible to check the percentage of instances for which the performance value is lower or higher than U . This opens our discussion of results for standard and tailored methods, separately.

Standard Methods

In Figure 6.3 results are presented for the computational time (F^a) of standard algorithms. The comparative graphs are shown in order from left to right, respectively, for small, medium and large instances. As instances get bigger, it is possible to observe the trend or evolution in performance profiles for each method.

For *small* instances, the optimal solution of the complete stochastic model *mDNP* by CPLEX (STO method) is obtained before 3 hrs. on roughly 20% of the instances. For the expected value deterministic method (EVD) this proportion of terminated instances is $\approx 73\%$ and for the rolling horizon procedure (RHP) it is 100%. In the worst-case the RHP method ends in ≈ 0.53 hrs.

For *medium* instances, the percentage of completed instances decreases to $\approx 7\%$ (STO) and $\approx 13\%$ (EVD). The RHP still terminates in all instances but the worst-case value increases to 2.3 hours, however about 93% of instances finish within 1 hr.

For *large* instances, the STO and EVD methods fail to complete even one instance within the time limit, while the RHP terminates in $\approx 90\%$ of cases.

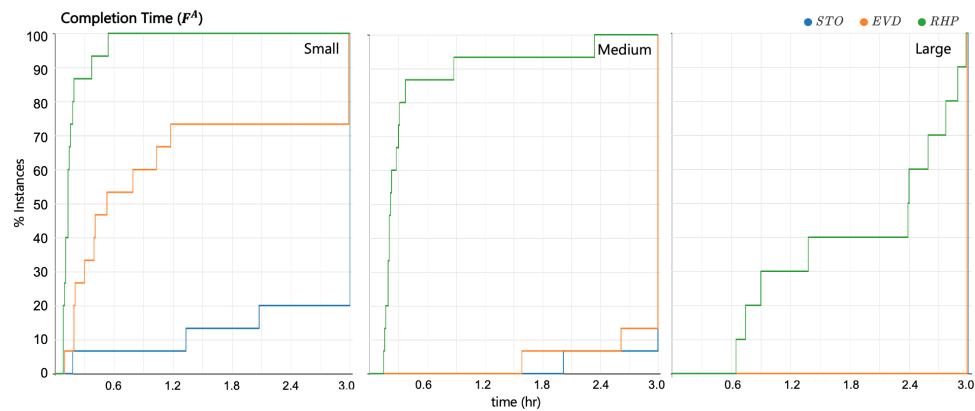


Figure 6.3: Computational Time for Standard Methods

In Figure 6.4 results are presented for the relative gap (G^a) in a similar way

For *small* instances, all standard methods achieve with certainty a gap lower than the upper limit (4%). The STO method has its performance profile above others, indicating a larger probability to provide lower gaps. The worst-case performance

value of G^a is approximately 0.24% for STO, 1.80% for EVD and 1.20% for RHP. For *medium* instances the performance profiles get lowered. The STO and RHP methods produce similar performance or percentage of instances for gaps less than $\approx 1.2\%$. For larger gaps, RHP outperforms STO, increasing the percentage of instances. The worst-case gap of G^a for RHP is about 2.27%, for EVD 3.86% and for STO is beyond limit in $\approx 7\%$ of cases.

For *large* instances, performance profiles for the STO and EVD methods are considerably decreased while the RHP maintains similar shape with respect to medium sizes. The gap provided by STO is lower than 4% in only $\approx 30\%$ of instances, and by EVD in only $\approx 50\%$. The RHP method returns a gap percentage of 1.22% in the worst-case.

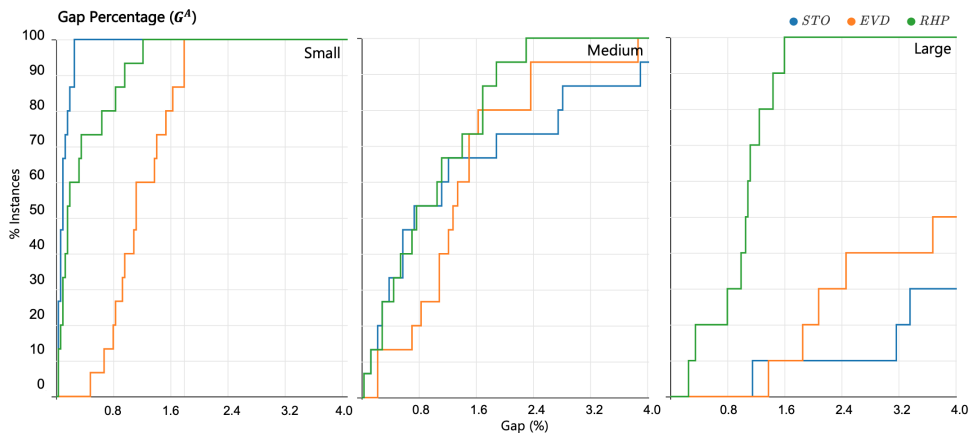


Figure 6.4: Gap Percentage for Standard Methods

Taking into account performance profiles for both metrics, it is noticeable that the RHP profile outperforms other methods in almost every graph. For large-size instances, the difference is significant in time and quality. Nonetheless, in small instances, the STO method is more likely to provide better quality solutions even though it is unlikely to finish before the time limit. The EVD method is outperformed by the RHP approach both in time and quality in every instance size, although compared with the STO method is expected to provide better quality solutions for large instances.

Tailored Methods

Results are now presented for the set of tailored heuristic methods, comprising relax-and-repair heuristics (CLR, WR, CLWR). The RHP is also included for comparison purposes with respect to standard methods.

In Figure 6.5 for *small* instances, it is noticed that all heuristic methods are able to finish rapidly within the time limit. Most of the performance profiles overlap each other indicating similar computational times. However the CLR and CLWR profiles are slightly faster than the RHP and WR ones. In the worst-case, the CLR and CLWR methods end in $\approx 0.57\text{h}$ and $\approx 0.18\text{h}$, respectively, while for the RHP and WR methods these values are $\approx 0.56\text{h}$ and $\approx 0.53\text{h}$. For *medium*-size instances, performance profiles for the CLR, CLWR and RHP methods are mildly lowered but still overlap each other, however the WR performance is considerably decreased and the method terminates for about 47% of instances. This indicates more sensibility to larger datasets. For large-size instances the profiles are distant from each other and a clear dominance is visible: the CLWR and CLR methods always complete within the time limit, with a worst-case time of 1.86h and 2.0h, respectively. In turn, the RHP fails to finish in $\approx 10\%$ of cases, while the WR method always used up time completely.

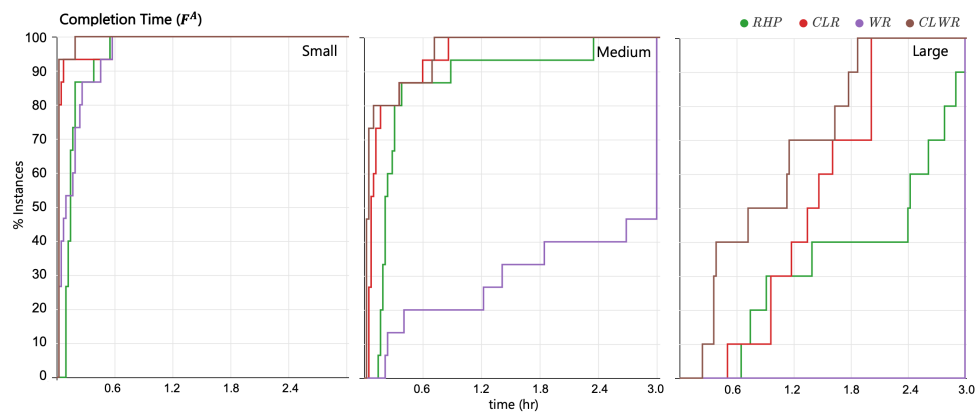


Figure 6.5: Computational Time for Heuristics Methods

From the perspective of the solution quality, Figure 6.6 shows that in small instances all tailor-made methods follow similar patterns, however, the WR performance is a little above the rest. In the worst-case the WR method obtains a gap of $\approx 0.67\%$, the

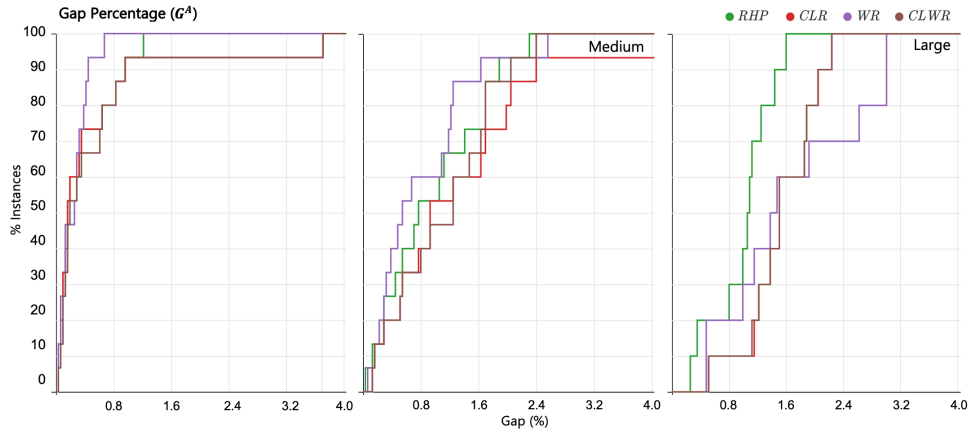


Figure 6.6: Gap Percentage for Heuristics Methods

RHP $\approx 1.20\%$, while the CLR and CLWR methods $\approx 3.69\%$ both. For medium size instances all performance profiles are lowered with small differences among them. Yet, the WR method (capacity relaxation) is slightly better overall, particularly with respect to the CLR, and CLWR methods. However, the worst-case gap is fairly similar for all methods. For the WR, CLWR and RHP it is about 2.54% , 2.38% and 2.27% respectively, while the CLR worst-case gap is beyond 4% . For large size instances, the RHP performance profile is shown to be above others. The CLR and CLWR performance profiles overlap directly each other while the WR gap performance is notoriously decreasing in quality. The worst-case gaps are then, 1.57% for the RHP, 2.04% for the CLR and CLWR methods, and 3.0% for the WR method.

Comparing both performance metrics, it is possible to make several observations. Methods including contract length relaxations (CLR, CLWR) show very similar performance in both metrics, however the CLWR method is faster in large instances.

This shows that the solution (contract plan) issued from the first phase is likely to be similar with either relaxed or non-relaxed w binary variables. However as it is expected solving the $mRDNP$ model accelerates the process while solving the second phase impacts less in the computational time. In view of these results, it is recommendable to opt for a double relaxation (CLWR), specially as instances increase in size. WR is highly sensitive to the instance size in terms of the completion time. Similar to complete models (STO, EVD), WR fails to terminate for any large instance within time limit. This indicates the difficulty raised by time-linking constraints for general-purpose algorithms in spite of the relaxation of some integer

variables.

6.6.3 Algorithmic Results

Running the complete model (STO) for large instances of the problem with general-purpose algorithms results in poor quality solutions, therefore some features of the problem have to be overlooked. However, neglecting the stochastic nature of the demand with the use of the 'expected-value' deterministic model doesn't provide an advantage at computational level. This shows the inherent difficulty in dealing with the multi-period aspect of the problem due to mid-term contracts, even with deterministic demands.

Alternatively, considering a shortened time horizon disregarding future demand behavior, as in the RHP, turns out to be more efficient computationally specially when instances increase in size. The RHP procedure does not only terminate faster but also yields better solutions while taking into account the stochastic modelling of the demand. Thus, removing the effect of future demand helps to reduce the complexity induced by a long planning horizon. In sum, the RHP outperforms other standard methods (EVD, STO) both in computational time and quality. Nonetheless, the RHP method alone does not provide a valid lower bound.

Relax-and-repair heuristics exploit the decomposition nature of the model after sorting out first-stage decision variables. They show to be effective in finding good solutions by relaxing some constraints of the model. The relaxation of off-contract capacity variables WR does not affect the computational time significantly. However, WR lags behind in providing competitive (short-gap) solutions as instances grow in size. The relaxation of contract lengths CLR does accelerate the computational process to a larger extent and is also able to return relatively good solutions for any instance size. By mixing both relaxations in CLWR, the computation of first stage variables is even faster, while holding good quality solutions overall. In sum the double-relaxed CLWR method outperforms the single CLR method by a small margin, consuming less time in obtaining solutions of similar quality. Nevertheless, the CLR method obtains tighter lower bounds for large instances. The same can be said of the WR method, however WR holds a small advantage over CLR in small-to-medium instances in terms of quality.

For the comparative analysis, it is possible to decide an appropriate approach depending on the instance size and performance criteria. For small to medium size instances it will be appropriate to solve the complete full horizon model with a

general purpose algorithm (STO) or relaxing off-contract capacity variables (WR) yields close-to-optimal solutions. For medium to large instances the best alternatives are to apply a rolling horizon procedure (RHP) which is likely to return best quality solutions or run a double-relaxed heuristic CLWR that proves to be faster than any other. The advantage with the CLWR heuristic model is its possibility to derive multiple solutions from its repairing phase by deciding differently upon unfinished contracts in the relaxed solution. CLWR could be coupled with local search algorithms that take advantage of its fast computation. In general, relax-and-repair algorithms could be enhanced with meta-heuristics that explore further the solution space with single modifications of the contract plan solution.

Taking into account stochasticity in the mathematical model or discretizing capacity per levels does not add up to the difficulty of the problem as much as contracts lasting for multiple durations. It proves time-linking constraints play an important role in the complexity of the problem. Decomposition strategies solving single-period problems are able to focus on the distribution network optimization at the expense of working on sub-optimal contract plans from relaxed formulations. Further development of algorithms could emphasize the efficient solution of single-period (two-stage) stochastic problems to handle even larger datasets.

6.6.4 Value of the stochastic solution

The value of the stochastic solution (VSS) is defined as the possible gain from solving the stochastic model as opposed to solving a deterministic approximation of it. It is an indicator that assesses the value of knowing and using probability distributions of uncertain parameters. In this study, we are interested in knowing the utility of using a discrete distribution of the demand in the network design problem (*DNPMC*). On the contrary, estimating and modelling the demand as a deterministic parameter could imply an economic loss in the long-term.

Following the categorization of instances (Table 6.4) and methods in our experimental design, the VSS is calculated for each size category with respect to standard procedures used commonly in practice (EVD, RHP), which make use of stochastic modelling of the demand. For every instance the relative difference between a stochastic approach and the deterministic model is computed and averaged over all instances of similar size. Let J be the set of instances of similar size indexed by j , then:
$$VSS = \frac{1}{|J|} \sum_{j \in J} \frac{Z_j^{DET} - Z_j^{STO}}{Z_j^{STO}}$$

The economic savings of modelling the demand as a stochastic quantity varies ac-

ording to the instance size due to the limited testing time. Results are shown in table 6.5 for every instance size.

Category	<i>STO</i>	<i>RHP</i>
Small	1.06%	0.82%
Medium	-0.06%	0.46%
Large	-1.88%	2.90%

Table 6.5: Relative value of the stochastic solution on average

For small size instances, the complete stochastic model returns economical savings in the time horizon ($|T| = 8$ periods). More exactly, the deterministic model entails a cost increase of 1.06% over the stochastic one. Taking into account that the EVD method gets the optimal value in $\approx 73\%$ of times, this margin is an accurate estimate of the potential benefits from solving the stochastic model for the long-term. In the same way, the benefits from implementing the rolling horizon approach (with $r = 1$) is estimated to yield 0.82% of economical savings on average, as compared to the EVD approach. The difference in margins (0.24%) represents the average benefit from taking into account the complete horizon as opposed to a single-period planning window.

The outlook looks different when instances increase in size. For medium and large instances opting for solving the complete stochastic model up to the testing time returns a negative VSS. In either size the solver does not prove optimality and running the deterministic model is more likely to result in a more profitable solution. For large-size instances the VSS goes down to -1.88%. Indeed, the higher the computational effort is spent in solving the stochastic model the worse is the sub-optimal solution value. In contrast, the rolling horizon approach is more likely to finish or find close-to-optimal solutions for the $sDNP(t, \beta^t)$ sub-models. In that framework, the stochastic approach is likely to improve on the deterministic solution. For large-size instances the average benefit is about 2.9%.

6.7 Managerial Insights

6.7.1 Value of mid-term MCC contracts

In addition to the value of the stochastic solution, it is possible to assess the value of including mid-term contracts with minimum commitments. For that purpose, we compare results in term of costs (best-found solution) with respect to two alternative strategies. First, we consider that the shipper adopts a *spot-market strategy* sourcing capacity for the immediate period once the demand is realized (i.e. at off-contract prices). This strategy provides full flexibility to request capacity to the extent that is required, but in practice involves the risk of low resource availability which often translates in high market prices.

This strategy removes the stochastic component of the problem and reduces to solving the single-period distribution network problem $sDNP(t, s, \hat{\beta}^t, \hat{v}^t)$ for each period $t \in T$ and scenario $s \in S^t$, where the values for the contract plan $\hat{\beta}^t$, and capacity reservation plan \hat{v}^t are set to zero. Subsequently, we calculate the weighted costs according to the probability of scenario s , for each period t . The total costs over T are compared with respect to the optimal or best-found expected value taking into account MCC contracts and the difference determined in percentage and averaged over all instances of the same size category.

As described in the experimental design, the parameter P that determines off-contract prices was given a value of 1.5, which is equivalent to say, FTL prices are 50% higher for immediate unreserved loads. Results are summarized in Table 6.6.

Category	Spot-market planning	Mixed long-term/spot planning	Single-period contracts
Small	7.87 %	2.91 %	- 0.50 %
Medium	8.46 %	1.92 %	- 0.83 %
Large	11.45 %	2.00 %	- 0.93 %

Table 6.6: Value of mid-term MCC contracts

In addition to have access to spot-market capacity as ultimate recourse, a second strategy consists in fixing long-term contracts with only a subset of carriers, that lasts for the complete horizon T . Hence, the shipper adopts a *mixed long-term/spot-market planning strategy*.

Indeed, a common feature of the optimal solutions after solving the $mDNP$ problem

with single-period MCC contracts, showcase a small group (i.e., 0, 1 or 2) of carriers permanently contracted along T , while others are signed temporarily for some periods (usually across high-demand periods). Based on this observation, we establish these subsets of carriers as permanent, from which capacity can be requested before demand realization (and respecting commitment M_e^t) throughout T . Hence, the contract duration for permanent carriers is $H_e = |T|$.

The above situation can be handled by solving the single-period distribution network problem for period t , $sDNP(t, \hat{\beta}^t)$, where the values $\hat{\beta}^t$ are set to 1 for permanent carriers and 0 for the rest. This mixed short-and-long term strategy reduces the flexibility in adjusting the distribution system according to the resulting demand, but in practice lowers the risk of low resource availability, given that an initial capacity is already in place throughout T . Similarly, we compute the difference in cost in percentages and averaged over all instances of the same size category (see results in Table 6.6).

Lastly, in order to measure the impact of multi-period MCC contracts as opposed to a single-period MCC contract, the negative difference in cost with respect to the latter optional plan is computed and reported in percentage terms. Representing a lower bound, this single-period solution is obtained from the relaxation phase during the CLR heuristic procedure.

From Table 6.6, it is possible to observe the economical benefits of opting for a procurement strategy with mid-term MCC contracts. In comparison with a strategy based only on requesting transportation capacity on-demand (according to the realized scenario), there is a substantial difference in the average gap. Optimal solutions obtain on average costlier values above 7.97%, 8.46% and 11.45% for small, medium and large instances, respectively. Incorporating now, some long-term deals with MCC agreements with few carriers allows to obtain considerable cost reductions, however, the optimal cost values (on average) remain larger by 2.91%, 1.92% and 2.00%, for small, medium and large instances, respectively. Thus, including mid-term agreements achieves an important decrease in long-term costs.

In general, larger values of H_e entails additional costs. By considering only long-term contracts (i.e., $H_e = |T|$), it is possible to see the maximum effect that the configuration of the parameter can have. On the other hand, if mid-term MCC contracts are pre-agreed for a single-period ideally (i.e., $H_e = 1$), then an improvement of less than 1% could be achieved.

6.7.2 Structure of solutions

After running several tests over various instances, some information can be drawn regarding the structure of the solutions obtained. As a side note, we consider only the best found solution, regardless of from which method it was obtained. We define some output parameters to characterize certain aspects of the decision variables.

Considering general information about carriers selection, we compute the number of permanent as well as occasional carriers and determine the percentage over the complete set size $|E|$ (averaging over instances of same size category). On top of that, the percentage of carriers selected for off-contract services with a distinction made for low demand (LD) and high demand (HD) periods. The maximum, minimum, and average values are reported in Table 6.7. These values hint an idea of the amount of operators that are finally selected from the complete pool of candidates and their frequency.

From Table 6.7, the percentage of permanent carriers is (on average) not higher than 15% and at most 25% (medium-size instances), while in some cases not one single carrier is selected for the whole horizon. The number of occasional carriers is slightly higher but not surpassing 30% on average (large-size instances), however, it can reach up to 66.7% of candidates. It is also observed that, for any instance, at least some occasional carriers are selected to work under MCC contracts. The percentage of carriers selected for off-contract services vary a lot depending on the realized scenario and demand period. In scenarios with larger-than-predicted demand (s^2), the selection of carriers is on average 67.2% (small), 60.4% (medium) and 77.9% (larges) for low demand periods, while it raises to 86.1% (small), 86.5% (medium) and 87.5 (larges), for high-demand periods. This shows the utility in having a second optional strategy as recourse, even if it implies higher service costs.

In addition, we define parameters highlighting information regarding resource utilization under both alternatives: MCC contracts and off-contract services.

The triggering effect of the MCC level is measured by computing the difference between the total capacity (trucks) requested or reserved *before* demand realizations and the pre-agreed capacity commitment M_e^t . For every instance, when an MCC contract is ongoing at period t , we measure the difference between the capacity effectively requested (*before* demand realization) and the pre-agreed capacity commitment M_e^t , which is the least value. Then, we calculate this difference in percentage terms to determine how far is the utilization from the commitment level.

Category	Small			Medium			Large		
Parameter	max	min	avg	max	min	avg	max	min	avg
Permanent Carriers (%)	16.7%	0%	10%	25%	0%	14.6%	0%	0%	0%
Occasional Carriers (%)	66.7%	16.7%	27.8%	25%	12.5%	18.8%	30%	30%	30%
Off-contract Carriers (LD, s^2) (%)	83.3%	50%	67.2%	95.8%	43.8%	60.4%	85%	75%	77.9%
Off-contract Carriers (HD, s^2) (%)	100%	66.6%	86.1%	100%	75%	86.5%	90%	80%	87.5%

Table 6.7: Output parameters : carrier selection data

Parameter	Small			Medium			Large		
	max	min	avg	max	min	avg	max	min	avg
Cap. utilization above MCC (LD)	30.9%	63.3%	48.4%	33.3%	21.2%	28.3%	14.9%	12.5%	13.6%
Cap. utilization above MCC (HD)	33.3%	87.5%	54.3%	68.4%	20.5%	32.4%	28.4%	12.8%	17.3%

Table 6.8: Output parameters : capacity utilization above commitment

Parameter	Small		Medium		Large	
	LD	HD	LD	HD	LD	HD
Cap. utilization with MCC contracts (scenario s^1)	75.3%	92.8%	74.2%	92.0%	77.6%	94.1%
Cap. utilization off-contract (scenario s^1)	24.7%	7.2%	25.8%	8.0%	22.4%	5.9%
Cap. utilization with MCC contracts (scenario s^2)	47.4%	50.1%	49.6%	46.7%	47.1%	48.9%
Cap. utilization off-contract (scenario s^2)	52.6%	49.9%	50.4%	53.3%	52.9%	51.1%

Table 6.9: Output parameters : capacity utilization between MCC contracts and off-contract services

Once again, a distinction is made between low (LD) and high (HD) demand periods. This measure is reported for maximum, minimum and average values in Table 6.8, taking into account instances of same size category.

From Table 6.8, it is possible to observe that the capacity utilization above commitments decreases as the instance size grows. Indeed, the commitment level is determined from the potential demand to be allocated to one carrier (approx. ratio customers $|K|$ /number of carriers $|E|$), which increases from one to the next category. Therefore, it becomes more stringent for the shipper to satisfy the pre-agreed level M'_e . However, there is always some slack of capacity effectively used even for large instances, which demonstrates the incorporation of the capacity initially committed into the operational solution. Also, it is notable that the difference is slightly higher in periods of larger demand. The above information exemplifies how the tactical planning framework could be used as an adaptive negotiation tool for the shipper so as to offer the largest possible minimum commitments without deteriorating shipper's costs.

Finally, information about the balance between the capacity requested *before* (i.e. by contracts) and *after* demand realization (i.e. off-contract) is given in percentages. A distinction is made in terms of both, the period seasonality and scenario. Results are summarized in table 6.9.

Table 6.9 features a similar behaviour for the indicated parameters regardless of the instance size. When the scenario of demands equals the predicted values (s^1), the proportion of capacity used from MCC contracts is considerably higher than from off-contract services. This difference is accentuated in high-demand periods, with values around 92.8%(small), 92.0% (medium) and 94.1%(large) on average. Differently, when the scenario results in larger demands than expected (s^2) the percentage is equally divided between both type of sources, independent from the seasonality.

6.8 Conclusions

The stochastic *mDNP* problem was introduced in this chapter. For its resolution several algorithmic methods were tested and compared in different settings of the problem. The computational results were presented in comparative performance profiles for two main performance criteria: completion time and best found value, testing separately their speed and quality.

In this study, the selection of carriers is based on the costs-savings criterion, involv-

ing decisions at the tactical and operational level. MCC contracts define clear liabilities for both parties including conditions on capacity reservation and utilization on specific shipping lanes/facilities of the distribution network. As an alternative procurement strategy the mathematical representation permits to request capacity off-contract which represents the alternative short-term planning strategy. The interested shipper may also consider multiple candidate carriers and/or contract alternatives with different parameter settings, like duration, service costs and penalty costs.

From the computational study, it was possible to observe possible economical advantages by modelling uncertainty of the demand as opposed to the expected value deterministic model. In a small size setting (with 8 carriers and 6 periods), benefits claimed were about 1% which for a large 3PL firm moving a large amount of freight may be significant in the long term. This margin is not clear as the size of instances increase due to computational limitations. However, with very limited time (like taking decisions daily), heuristic approaches are better suited to find good quality solutions.

In effect, the computational difficulty increases rapidly with the size of input parameters, such as number of carriers, periods and customers, and could be difficult to solve in practice in short planning periods. In particular, dealing with multiple demand scenarios per period can be computationally challenging. However, the stochastic model presents the particularity that decisions made at any period are independent from previous realizations of the demand (the random vector of scenarios S^t is linked to each period). Hence, taking into account several scenarios of demand to period t increase the size of the model linearly rather than exponentially.

In order to solve the multi-period model in reasonable time, certain aspects need to be neglected to allow to find better solutions of the problem. Notably, disregarding distant future information of the demand or assuming single-period contracts prove to be effective. Further modelling development could account for uncertainty in other unpredictable parameters like transportation costs in spot-market rates which are volatile and subject to random changes.

Chapter 7

Conclusions, Limitations and Perspectives

7.1 Conclusions

In this thesis, we studied multi-period distribution network problems with minimum commitment contracts (DNPMC). The purpose of the network optimization problem is to minimize total costs of transportation and distribution over a planning horizon. The tactical planning works as an alternative approach to common practices for freight transportation procurement, such as long-term deals and one-time services. It balances out the flexibility to adapt logistical capacity to fluctuating demands (due to seasonality effects) and stable services costs. Mid-term contracts with minimum commitment provisions are proposed as a win-win strategy with transporters, in order to tackle potential shipper's risks, such as limited vehicle capacity and high prices (in spot markets). In that framework, we consider a 3PL company in the business of e-commerce fulfilment, that subcontracts with outside partners.

We present the mathematical representation (MILP) of the DNPMC problem in Chapter 3 and prove that it can be adapted for several types of minimum commitment (MC) agreements with distinct attributes, regarding the duration, commitment type, and penalty rates. Moreover, a more general formulation was included that exemplifies the integration of several procurement options, for instance to denote a carrier working under two contractual settings. The diversification of portfolio

is beneficial due to its potential in reduce further logistics costs, however, multiple options increase the difficulty in finding good quality solutions.

In Chapter 4, the algorithms developed for MDPC take advantage of the decomposable structure of the problem. Their convenience was shown in comparison with state-of-the-art solvers, which are not well-suited to the characteristics of the models, particularly, its multi-periodicity and mixed-integer and linear features. This difficulty extends to the standard Benders algorithm, where the partition in complex (integer) and soft (linear) variables does not produces satisfactory results, in terms of computational timing and/or solution quality for MDPC.

Therefore, we propose an implementation of the combinatorial Benders algorithm (CBA) that separates only a subset of its integer variables, and left to the subproblem to deal with a decomposable MILP. Numeric results in exhaustive tests, show the convenience of the approach, especially, in solving large instances of the problem as compared to general-purpose B&C and Benders algorithms. Moreover, a heuristic method based on the relaxation of time-linking constraints provides even better results. On one hand, it significantly improves the lower bounds obtained from the commercial solvers' algorithms. On the other hand, it generates better quality solutions for large instances of the problem. This indicates that the difficulty in solving the multi-period DNPMC problem comes primarily from the interconnection of constraints linking periods related to contract duration rather than dealing with binary variables linked to the specification of capacity utilization.

In Chapter 5, the mathematical modelling was extended to the case of demand uncertainty. Considering the situation where capacity requests are placed as the demands realized (scenario-dependent), decisions in the first stage comprises only carrier selection. The stochastic modelling produces higher cost advantages in the situation where reservation is required. The shipper is not obliged to meet a minimum commitment level to gain access to resources. Instead, it could simply announce the capacity to use in advance. The consideration of reservation in place of commitments as a way to negotiate better prices, can be handled in a single-period two-stage stochastic model. The multi-period DNPMC is reasonable with an MCC contract of multi-period duration. Thereby, forcing the shipper to comply with minimum requests for consecutive periods in addition to reservation requisites.

The algorithmic development in Chapter 6 is based on the developed methods (in Chapter 4) for the deterministic DNPMC problem and explore further possibilities to tackle the MILP. We compare several solution approaches to solve the *mDNP_MCC*. Trying to solve the complete stochastic *mDNP_MCC* model results in poor quality solutions. Notably, for large instances up to 12 candidate carriers and 10 periods.

A better approach is to solve parts of the multi-period model independently like in the rolling horizon procedure or to relax some of the constraints and then correct the values of a given solution to make it feasible, like in relax-and-repair heuristics. Ultimately, four strategies were compared to show their performances in terms of quality and time. Among these, the RHP stands out for finding the best quality solution within the testing time limit, and the double relaxation CLWR strategy turns out to be the fastest. The CLR heuristics turns out tied in second place for both criteria (over large instances). The WR heuristic based on the relaxation of integer variables is not favourable for large instances. Once again, It shows the intrinsic difficulties related to time-linking constraints in the *mDNP* model.

7.2 Limitations

The implementations of the studied models may present several challenges. Two main aspects concern the collection of data and the difficulty in finding optimal (or close-to) solutions by implementing algorithmic procedures.

The most relevant data concerns precise information about customers' demands and service (FTL, cross-docking) costs at different periods in the future. The first one might be difficult to predict accurately, nonetheless, the stochastic models allow to represent the distribution of the demand with multiple scenarios. The underlying probability distribution could be approximated for instance by making use of sample average approximation (SAA) methods, that deal with continuous probability functions. More elaborated forecasting models could be useful to update and reduce error margins in the predictions for periods in the future. The price of truck-loads for a given lane is highly dynamic. In long-term contracts, prices given by carriers are often private and difficult to access. Moreover, prices associated to one lane vary for each carrier. Spot market prices are also hard to track, although some mathematical models may address forecasting methods, as in [Budak et al., 2017].

A second difficulty comes from the willingness of carriers (or LSPs), to agree upon a pre-fixed set of prices for specific services under the premise of acceptable business volume, in the short-to-mid term. In that regards, the parameters of duration (H) and respective commitments values (M) might result from a series of bids until a common agreement is accepted and satisfy both parties. In fact, the mathematical model can serve as an analytical tool for the shipper in order to determine acceptable conditions, such as, the largest MCC or MPC values that it can afford so as to generate cost savings in the long-term. The application of the model as an negotia-

tion tool with multiple partners might present itself some challenges in terms of the computational effort to retrieve good quality solutions in short time.

The possibility of solving large instances of the mathematical model (e.g., multiple carriers/contracts alternatives, several demand scenarios and/or for an exhaustive customer base), is directly linked with the computational capacity at hand. Based on the numeric results, making use of the commercial solver CPLEX to solve large instances of the multi-period DNPMC model does not yield good bounds. An advantage of decomposition methods such as the contract length relaxation also lies in the possibility to solve in the first phase several sub-problems, which can be allocated to different machines. Due to the tactical perspective of the studied problems (i.e., planning periods of one or several weeks), the time allowed to solve the tactical plan is large enough to obtain a good quality solution with standard computational resources. If the computational time is short, more efficient methods like the application of heuristics are proven to be a better alternative.

This is also the most suitable alternative, in case the decision-maker should solve the problem under different conditions to try for example different values of commitments/duration. Hence, the computation can be accelerated by the means of non-exact methods. A more detailed resolution strategy may combine the rolling horizon procedure with relax and repairing heuristics in order to obtain the best benefits in terms of efficiency and solution quality.

7.3 Perspectives

The underlying distribution network model with two-echelons was inspired from the domain of e-commerce, but could be applied more generally in other supply chain contexts for an agent dealing with outsourcing or subcontracting logistics services to outside providers. For instance, it could be extended in multi-period vehicle routing problems with private fleet and common carriers referred in the literature review. Applications could be extended, for example, to hub-location problems in contexts where finding appropriate hub-and-spokes facilities lead to the optimization of commercial flows in wide-areas, implying long-distance transportation.

The utility of the CBA algorithm might be tested in MILP models with distinct subset of integer variables, for which a subset complicates its resolution to a greater extent and/or, a possible fixation of integer variables (to certain values) allows to decompose or treat sub-models separately. Further developments, might be in

the direction of applying acceleration strategies or faster resolution methods to the single sub-problems $sDNP(t, \tilde{\beta}^t)$. In our case, algorithmic strategies for location-allocation models could be included into the single-tree Benders approach. A different direction goes to strengthen combinatorial cuts to restrict even further the solution space, by for instance, pruning the selection of a carrier based on dominance rules.

Heuristics developed in Chapters 4 and 6 for the deterministic and stochastic DNPMC models, could be enhanced by implementing different reparation rules in the second-phase. Simple rules like, totally removing or enforcing incomplete contracts could yield to poor quality solutions, while evaluating both possibilities, (like in Heuristic 4.5) may be time-consuming. It may be interesting to assess different rules for the eligibility, for instance, in consideration of the difference with respect to the pre-agreed duration. Moreover, a local search procedure can explore the space of possible corrected solutions (from the one relaxed), focusing on subgroups of incomplete contracts that are overlapped in time and therefore, the decision of excluding/including one of them affect the usefulness for the rest. In general, a local search procedure can be embedded in meta-heuristic framework.

The contract type and its corresponding parameter values (H_e, M_e, P_e) are considered to be stated or known a priori to the network optimization problem. That means the agreement process (without any obligation) takes place beforehand, which returns as an output eligible contracts. The utility of the models could be extended as a negotiation tool to analyse the implication of possible values of contract parameters. Another perspective could be to make use of the optimization model to determine optimal parameter values before negotiation (i.e., as endogenous variables). The stochastic DNPMC model in the study of capacity requests ahead in time provides a possibility in that direction. The study of combined types of commitment contracts may also be possible in an integrated model.

Interestingly, the applicability and utility of the algorithmic methods could be tested with real-data provided from the e-commerce industry. Notably, input values for realistic probability functions for demand scenarios and truckload estimation prices at spot markets. The dynamic nature of spot market prices could be also modelled as stochastic parameter.

Bibliography

- Agatz, N., Fleischmann, M., & van Nunen, J. (2008). E-fulfillment and multi-channel distribution – a review. *European Journal of Operational Research*, 187(2), 339–356.
- Aghezzaf, E. (2005). Capacity planning and warehouse location in supply chains with uncertain demands. *Journal of the Operational Research Society*, 56(4), 453–462.
- 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.
- Albareda-Sambola, M., Alonso-Ayuso, A., Escudero, L., Fernández, E., & Pizarro, C. (2013). Fix-and-relax-coordination for a multi-period location–allocation problem under uncertainty. *Computers and Operations Research*, 40(12), 2878–2892.
- Albareda-Sambola, M., Fernández, E., Hinojosa, Y., & Puerto, J. (2009). The multi-period incremental service facility location problem. *Computers and Operations Research*, 36(5), 1356–1375.
- Awasthi, A., Chauhan, S., Goyal, S., & Proth, J. (2009). Supplier selection problem for a single manufacturing unit under stochastic demand. *International Journal of Production Economics*, 117(1), 229–233.
- Bai, Q., & Xu, J. (2011). Optimal solutions for the economic lot-sizing problem with multiple suppliers and cost structures. *Journal of Applied Mathematics and Computing*, 37, 331–345.
- Balas, E., & Jeroslow, R. (1972). Canonical cuts on the unit hypercube. *SIAM Journal on Applied Mathematics*, 23(1), 61–69.

- Ball, M., B.L. Golden, Assad, A., & Bodin, L. (1983). Planning for truck fleet size in the presence of a common-carrier option. *Decision Sciences*, 14(1), 103-120.
- Bassok, Y., & Anupindi, R. (1997). Analysis of supply contracts with total minimum commitment. *IIE Transactions*, 29, 373-381.
- Beale, E. (1955). On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society: Series B (Methodological)*, 17(2), 173-184.
- Benders, J. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4, 238-252.
- Birge, J., & Louveau, F. (2011). *Stochastic programming*. Springer.
- Bolduc, M., Renaud, J., & Boctor, F. (2007). A heuristic for the routing and carrier selection problem. *European Journal of Operations Research*, 183(2), 926-932.
- Bolduc, M., Renaud, J., Boctor, F., & Laporte, G. (2008). A perturbation meta-heuristic for the vehicle routing problem with private fleet and common carriers. *Journal of The Operations Research Society*, 59(6), 776-787.
- 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 Logistics Systems and Management*, 5(2-3), 273-322.
- Brusset, X. (2010). *Modeling contractual relationships in transport* (Unpublished doctoral dissertation). Louvain School of Management.
- Budak, A., Ustundag, A., & Guloglu, B. (2017). A forecasting approach for truck-load spot market pricing. *Transportation Research Part A*, 97, 55-68.
- Buer, T., & Pankratz, G. (2010). Solving a biobjective winner determination problem in a transportation procurement auction. *Logistics Research*, 2, 65-78.
- Burke, G., Carrillo, J., & Vakharia, A. (2008). Heuristics for sourcing from multiple suppliers with alternative quantity discounts. *European Journal of Operational Research*, 186(1), 317-329.
- Buyukozkan, G., Feyzioglu, O., & Ersoy, M. S. (2009). Evaluation of 4pl operating models: A decision making approach based on 2-additive choquet integral. *International Journal Production Economics*, 121, 112-120.

- Cachon, G., & Lariviere, M. (2005). Contracting to assure supply: How to share demand forecasts in a supply chain. *Management Science*, 47(5), 629–646.
- Canel, C., Khumawala, B., Law, J., & Loh, A. (2001). An algorithm for the capacitated, multi-commodity multi-period facility location problem. *Computers and Operations Research*, 28(5), 411–427.
- Caplice, C. (2007). Electronic markets for truckload transportation. *Production and Operations Management*, 16(4), 423–436.
- Caplice, C., & Sheffi, Y. (2003). Optimization-based procurement for transportation services. *Journal of Business Logistics*, 24(2), 109–129.
- Castaneda, J., Toro, E., & Gallego, R. (2020). Iterated local search for the vehicle routing problem with a private fleet and a common carrier. *Engineering Optimization*, 52(10), 1796–1813.
- Chauhan, S., & Proth, J. (2003). The concave cost supply problem. *European Journal of Operational Research*, 148(2), 374–383.
- 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.
- Chu, C. (2005). A heuristic algorithm for the truckload and less-than-truckload problem. *European Journal of Operations Research*, 165(3), 657–667.
- Codato, G., & Fischetti, M. (2006). Combinatorial Benders' cuts for mixed-integer linear programming. *Operations Research*, 54(4), 756–766.
- Contreras, I., & Fernandez, E. (2012). General network design: A unified view of combined location and network design problems. *European Journal of Operational Research*, 219(3), 680–697.
- Corbett, C., Zhou, D., & Tang, C. (2004). Designing supply contracts: Contract type and information asymmetry. *Management Science*, 50(4), 425–559.
- Cordeau, J., Soumis, F., & Desrosiers, J. (2000). A Benders decomposition approach for the locomotive and car assignment problem. *Transportation Science*, 34(2), 133–149.

- Correia, I., & Melo, T. (2017). A multi-period facility location problem with modular capacity adjustments and flexible demand fulfillment. *Computers and Industrial Engineering*, *110*, 307–321.
- 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 and Industrial Engineering*, *90*, 118–131.
- Coyle, J., Novack, R., Gibson, B., & Bardi, E. (2010). *Transportation: a supply chain perspective* (7th ed.). Cengage Learning.
- Crujssen, F., Dullaert, W., & Fleuren, H. (2007). Horizontal cooperation in transport and logistics: A literature review. *Transportation Journal*, *46*(3), 22–39.
- Cui, L., & Hertz, S. (2011). Networks and capabilities as characteristics of logistics firms. *Industrial Marketing Management*, *40*(6), 1004–1011.
- Côté, J.-F., & Potvin, J.-Y. (2009). A tabu search heuristic for the vehicle routing problem with private fleet and common carrier. *European Journal of Operations Research*, *198*(2), 464–469.
- Dabia, S., Lai, D., & Vigo, D. (2019). An exact algorithm for a rich vehicle routing problem with private fleet and common carrier. *Transportation Science*, *53*(4), 917–1212.
- Dantzig, G. (1955). Linear programming under uncertainty. *Management Science*, *1*(3-4), 197–206.
- Diaby, M., & Ramesh, R. (1995). The distribution problem with carrier service: A dual based penalty approach. *INFORMS Journal of Computing*, *7*(1), 1–116.
- 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., & Moré, J. (2002). Benchmarking optimization software with performance profiles. *Mathematical Programming*, *91*, 201–213.
- Erlenkotter, D., & Roy, T. V. (1982). A dual-based procedure for dynamic facility location. *Management Science*, *28*(10), 1091–1105.
- Euchi, J. (2017). The vehicle routing problem with private fleet and multiple common carriers: Solution with hybrid metaheuristic algorithm. *Vehicular Communications*, *9*, 97–108.

- Eurostat. (2020). *Eu transport in figures*. Retrieved from https://ec.europa.eu/transport/facts-fundings/statistics/pocketbook-2020_en
- 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.
- Georgiadis, M., Tsiakis, P., Longinidis, P., & Sofioglu, M. (2011). Optimal design of supply chain networks under uncertain transient demand variations. *Omega*, 39(3), 254–272.
- Gham, C., Brabander, C., & Tuma, A. (2017). Vehicle routing with private fleet, multiple common carriers offering volume discounts, and rental options. *Transportation Research Part E: Logistics and Transportation Review*, 97, 192–216.
- Ghaniabadi, M., & Mazinani, A. (2017). Dynamic lot-sizing with multiple suppliers, backlogging and quantity discounts. *Computers and Industrial Engineering*, 110, 67–74.
- Ghiani, G., Laporte, G., & Musmanno, R. (2004). *Introduction to logistics systems planning and control*. Wiley.
- Hazra, J., & Mahadevan, B. (2009). A procurement model using capacity reservation. *European Journal of Operational Research*, 193(1), 303–316.
- Hesse, M. (2002). Shipping news: the implications of electronic commerce for logistics and freight transport. *Resources, Conservation & Recycling*, 36(3), 211–240.
- 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.
- Hormozi, A., & Khumawala, B. (1996). An improved algorithm for solving a multi-period facility location problem. *IIE Transactions*, 28(2), 105–114.

- Huang, G., & Xu, S. (2013). Transportation service procurement in periodic sealed double auctions with stochastic demand and supply. *Transportation Research Part B*, 56, 136–160.
- Infosys. (2020). *2021-third-party logistics study. the state of logistics outsourcing*. Retrieved from <https://www.3plstudy.com/ic3pl/ic3pl.ic3pl.home>
- Jaruphongsa, W., C*etinkaya, S., & Lee, C.-Y. (2005). A dynamic lot-sizing model with multi-mode replenishments: polynomial algorithms for special cases with dual and multiple modes. *IIE Transactions*, 37(5), 453–467.
- 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.
- Knemeyer, A., Corsi, T., & Murphy, P. (2003). Logistics outsourcing: relationships customer perspectives. *Journal of Business Logistics*, 24(1), 77–109.
- Krajewska, M., & Kopfer, H. (2009). Transportation planning in freight forwarding companies: Tabu search algorithm for the integrated operational transportation planning problem. *European Journal of Operations Research*, 197(2), 741-751.
- Lafkihi, M., Pan, S., & Ballot, E. (2019). Freight transportation procurement: Alitearatur review and future research opportunities in omnichannel e-commerce. *Transportation Research Part E*, 125, 348–365.
- 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.), *Quantitative models for supply chain management* (chap. Supply chain contracting and coordination with stochastic demand). Kluwer Academic Publisher, Boston, Massachusetts.
- 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), 117–29.
- Lim, A., Xu, Z., & Wang, F. (2007). The bidding selection and assignment problem with minimum quantity commitment. *Journal of the Operations Research Society*, *59*(5), 693–702.
- Lim, S.-K., & Kim, Y.-D. (1999). An integrated approach to dynamic plant location and capacity planning. *Journal of the Operational Research Society*, *50*(12), 1205–1216.
- Lindsey, C., Frei, A., Alibabai, H., Mahmassani, H., Park, Y., Klabjan, D., . . . Keating, T. (2013). *Modeling carrier truckload freight rates in spot markets*.
- 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.
- Mansini, R., Savelsbergh, M., & Tocchella, B. (2012). The supplier selection problem with quantity discounts and truckload shipping. *Omega*, *40*, 445–455.
- Martínez-deAlbéniz, V., & Simchi-Levi, D. (2005). A portfolio approach to procurement contracts. *Production and Operations Management*, *14*(1), 90–114.
- Melachrinoudis, E., & Min, H. (2000). The dynamic relocation and phase-out of a hybrid, two-echelon plant/warehousing facility: A multiple objective approach. *European Journal of Operations Research*, *123*(1), 1–15.
- 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.
- Mesa-Arango, R., & Ukkusuri, S. (2014). Attributes driving the selection of trucking services and the quantification of the shippers’s willingness to pay. *Transportation Research Part E*, *71*, 142–158.
- Mohamed, I., Klibi, W., & Vanderbeck, F. (2020). Designing a two-echelon distribution network under demand. *European Journal of Operational Research*, *280*(1), 102–123.
- Nabil, A., & van den Heuvel, W. (2019). Worst case analysis of relax and fix heuristics for lot sizing problems. *European Journal of Operational Research*, *279*(2), 449–458.

- Nosoohi, I., & Nookabadi, A.-S. (2016). Outsource planning through option contracts with demand and cost uncertainty. *European Journal of Operational Research*, 250, 131–142.
- Park, S., & Kim, J. (2014). A mathematical model for a capacity reservation contract. *Applied Mathematical Modelling*, 38(5-6), 1866–1880.
- Patterson, Z., Ewing, G., & Haider, M. (2010). How different is carrier choice for third-party logistics companies. *Transportation Research Part E*, 46, 764–774.
- Pimentel, B., Mateus, G., & Almeida, F. (2013). Stochastic capacity planning and dynamic network design. *International Journal of Production Economics*, 145(1), 139–149.
- Qin, H., Luo, M., Gao, X., & Lim, A. (2012). The freight allocation problem with all-units quantity-based discount: A heuristic algorithm. *Omega*, 40(4), 415–423.
- 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.
- Rekik, M., & Mellouli, S. (2012). Reputation-based winner determination problem for combinatorial transportation procurement auctions. *Journal of the Operational Research Society*, 63(10), 1400–1409.
- Rosenblatt, M., Herer, Y., & Hefter, I. (1998). Note. an acquisition policy for a single item multi-supplier system. *Management Science*, 44(11), S1-S144.
- Sandholm, T., & Shields, R. (2006). Nogood learning for mixed integer programming. *Workshop on Hybrid Methods and Branching Rules in Combinatorial Optimization, Montréal*(11), 138.
- Scott, A. (2015). The value of information sharing for truckload shippers. *Transportation Research Part E*, 83, 203-214.
- Scott, A. (2018). Carrier bidding behavior in truckload spot auctions. *Journal of Business Logistics*, 39(4), 267–281.
- Scott, A., Parker, C., & Craighead, C. (2017). Service refusals in supply chains : Drivers and deterrents of freight rejection. *Transportation Science*, 51(4), 1086-1101.
- Sethi, S., & Sorger, G. (1991). A theory of rolling horizon decision making. *Annals of Operations Research*, 29(1).

- 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.
- Stenger, A., Schneider, M., & Goeke, D. (2013). The prize-collecting vehicle routing problem with single and multiple depots and non-linear cost. *EURO Journal on Transportation and Logistics*, 2, 57-87.
- Stenger, A., Vigo, D., Enz, S., & Schwind, M. (2013). An adaptive variable neighborhood search algorithm for a vehicle routing problem arising in small package shipping. *Transportation Science*, 47(1), 64-80.
- Tempelmeier, H. (2002). A simple heuristic for dynamic order sizing and supplier selection with time-varying data. *Production and Operations Management*, 11(4), 499–515.
- Tibben-Lembke, R., & Rogers, D. (2006). Real options: applications to logistics and transportation. *International Journal of Physical Distribution & Logistics Management*, 36(4), 252–270.
- Tsai, M.-T., Saphores, J.-D., & Regan, A. (2011). Valuation of freight transportation contracts under uncertainty. *Transportation Research Part E*, 47, 920–932.
- Tsay, A. (1999). The quantity flexibility contract and supplier-customer incentives. *Management Science*, 45(10), 1339-1358.
- Wang, X., Kopfer, H., & Gendreau, M. (2014). Operational transportation planning of freight forwarding companies in horizontal coalitions. *European Journal of Operational Research*, 237(3), 1133-1141.
- Xia, W., & Wu, Z. (2007). Supplier selection with multiple criteria in volume discount environments. *Omega*, 35(5), 494-594.
- Yu, Y., Wang, X., Zhong, R., & Huang, G. (2016). E-commerce logistics in supply chain management: Practice perspective. *Procedia CIRP*, 52, 179-185.
- Zapfel, G., & Bogl, M. (2008). Multi-period vehicle routing and crew scheduling with outsourcing options. *International Journal of Production Economics*, 113(2), 980–996.

- Zhang, J., & Zhang, M. (2011). Supplier selection and purchase problem with fixed cost and constrained order quantities under stochastic demand. *International Journal of Production Economics*, 129(1), 1–7.
- Zhao, Y., & Klabjan, D. (2012). A polyhedral study of lot-sizing with supplier selection. *Discrete Optimization*, 9(2), 65-76.
- Zhuge, D., Yu, S., Zhen, L., & Wang, W. (2016). Multi-period distribution center location and scale decision in supply chain network. *Computers and Industrial Engineering*, 101, 216–226.
- Ziebuhr, M., & Kopfer, H. (2014). The integrated operational transportation planning problem with compulsory requests. In R. González-Ramírez, F. Schulte, S. Voß, & J. C. Díaz (Eds.), *Computational logistics iccl 2014*. Springer.

List of Tables

2.1	Classification of risk-sharing contracts	23
2.2	Multi-period facility location problems - deterministic models	32
2.3	Multi-period facility location problems - stochastic models	32
3.1	Network design decisions	38
3.2	Mathematical notations for DNPMC	40
3.3	Characterization of MC Contracts	43
4.1	Mathematical notations for the MDPC model	64
4.2	Description of instance classes	81
4.3	Numeric values for efficiency measures (F^A, B^A) in small instances	86
4.4	Numeric values for effectiveness measures ($G^A, \Delta Q^A$) in small instances	87
4.5	Numeric values for efficiency measures (F^A, B^A) in medium instances	90
4.6	Numeric values for effectiveness measures ($G^A, \Delta Q^A$) in medium instances	91
4.7	Numeric values for efficiency measures (F^A, B^A) in large instances	94
4.8	Numeric values for effectiveness measures ($G^A, \Delta Q^A$) in large instances	95
5.1	Notations: Datasets and parameters for stochastic DNPMC problems	108
5.2	Decision variables for models in Section 5.5	108
5.3	Decision variables for models in Section 5.6	121
6.1	Mathematical notation for the <i>mDNP_MCC</i> model	133
6.2	Standard methods	146
6.3	Relax-and-repair heuristics	146
6.4	Classification of datasets per size	148
6.5	Relative value of the stochastic solution on average	158
6.6	Value of mid-term MCC contracts	159
6.7	Output parameters : carrier selection data	162

6.8	Output parameters : capacity utilization above commitment	163
6.9	Output parameters : capacity utilization between MCC contracts and off-contract services	163

List of Figures

1.1	Thesis contributions	4
1.2	Relationships between logistics service providers	7
1.3	Thesis outline	15
2.1	Two-stage procurement process	20
2.2	Configuration of contract parameters and risk-sharing provisions	21
3.1	Two-echelon distribution network layout	37
3.2	Capacity levels	38
3.3	The minimum purchase commitment contract	45
3.4	Configuration of an MPC contract without penalty rates	46
3.5	Configuration of an MPC contract with penalty rates	52
4.1	Distribution network layout	62
4.2	Plan of contracts (valid values of $\hat{\alpha}_e^t$)	71
4.3	Single-tree Combinatorial Benders algorithm	74
4.4	Performance of algorithms on small instances	83
4.5	Efficiency: performance profile of total running time F^A and of running time until best found solution B^A - small instances	84
4.6	Effectiveness: performance profile of optimality gap G^A and of gap to best known value ΔQ^A - small instances	84
4.7	Efficiency (total computational time F^A and running time B^A until best solution) of algorithms as a function of parameter values for small instances	86
4.8	Effectiveness (optimality gap G^A and gap ΔQ^A to best known value) of algorithms as a function of parameter values for small instances	87
4.9	Performance of algorithms on medium instances	88
4.10	Efficiency: performance profile of total running time F^A and of running time until best found solution B^A - medium instances	88
4.11	Effectiveness: performance profile of optimality gap G^A and of gap to best known value ΔQ^A - medium instances	89

4.12	Efficiency (total computational time F^A and running time B^A until best solution) of algorithms as a function of parameter values for medium instances	90
4.13	Effectiveness (optimality gap G^A and gap ΔQ^A to best known value) of algorithms as a function of parameter values for medium instances	91
4.14	Performance of algorithms on large instances	92
4.15	Efficiency: performance profile of total running time F^A and of running time until best found solution B^A - large instances	92
4.16	Effectiveness: performance profile of optimality gap G^A and of gap to best known value ΔQ^A - large instances	93
4.17	Efficiency (total computational time F^A and running time B^A until best solution) of algorithms as a function of parameter values for large instances	94
4.18	Effectiveness (optimality gap G^A and gap ΔQ^A to best known value) of algorithms as a function of parameter values for large instances	95
4.19	Variation of the computational time with respect to the contract duration	96
5.1	The minimum capacity commitment contract	103
5.2	Scenario-tree for the stochastic DNPMC problem	106
5.3	Multi-stage stochastic decision-making under MCC contracts without capacity reservation	109
5.4	The MCC contract - capacity reservation	111
5.5	Configuration of an MCC contract without penalty rates	111
5.6	Configuration of and MCC contract with penalty rates)	116
5.7	Multi-stage stochastic decision-making under MCC contracts with capacity reservation	119
5.8	The MCC contract with capacity reservation	121
5.9	Configuration of an MCC contract with capacity reservation	122
5.10	Configuration of an MCC contract with capacity reservation and minimum commitment	124
6.1	Distribution network layout	130
6.2	Comparison of Lower bounds among algorithmic methods	150
6.3	Computational Time for Standard Methods	152
6.4	Gap Percentage for Standard Methods	153
6.5	Computational Time for Heuristics Methods	154
6.6	Gap Percentage for Heuristics Methods	155