Modeling and solving a distribution network design problem with multiple operational constraints. Application to a case-study in the automotive industry.

Mouna Kchaou-Boujelben

ÉCOLE CENTRALE DES ARTS ET MANUFACTURES
« ÉCOLE CENTRALE PARIS »

THÈSE

présentée par

Mouna KCHAOU BOUJELBEN

pour l'obtention du

GRADE DE DOCTEUR

Spécialité : Génie Industriel
Laboratoire d'accueil : Laboratoire Génie Industriel

SUJET : Modeling and solving a distribution network design problem with multiple operational constraints. Application to a case-study in the automotive industry.

Soutenue le : 2 Décembre 2013
Devant un jury composé de :

Pr Pierre Dejax, Ecole des Mines de Nantes	Président
Pr Eric Pinson, Université catholique de l'ouest	Rapporteur
Pr Van-Dat Cung, Grenoble INP Génie Industriel	Rapporteur
Pr Yves Dallery, Ecole Centrale Paris	Examinateur
Mr Alain Nguyen, Renault SAS	Examinateur
Pr Michel Minoux, Université Paris 6	Directeur de thèse
Mme Céline Gicquel, Université Paris 11	Co-encadrante de thèse

2013 – ECAP0067

A mes chers parents……à mon cher mari……à mes petits anges Mohamed et Eya……je dédie ce travail ♡

Remerciements

Je tiens tout d'abord à remercier Pr Eric Pinson, Pr Van-Dat Cung , Pr Pierre Dejax et Mr Alain Nguyen d'avoir accepté de faire partie du jury de ma thèse.

Cette thèse n'aurait pas vu le jour sans une première rencontre avec Pr Yves Dallery qui a cru en moi et m'a donné la chance de rencontrer mon directeur de thèse Pr Michel Minoux et l'équipe du DOP chez Renault. Avec son expertise en supply chain, il nous a énormément aidé au début de la thèse à orienter le sujet et à répondre aussi bien aux besoins industriels qu'aux exigences académiques. Aujourd'hui, il me fait l'honneur d'accepter de faire partie du jury de ma thèse.

Je souhaite exprimer toute ma gratitude à Pr Michel Minoux pour avoir accepté de diriger ce travail, pour la qualité de l'encadrement et du soutien qu'il m'a accordés et pour les nombreux conseils et les riches discussions que nous avons eues ensemble avec Céline Gicquel.

Céline, je ne te remercierai jamais assez pour tout le soutien que tu m'as donné tout au long de cette thèse sur le plan scientifique aussi bien que sur le plan humain. Tu n'es pas uniquement une co-encadrante pour ma thèse, tu es une amie...

Je remercie également Aimé Frédéric Rosenzweig pour m'avoir accueillie dans le DOP chez Renault et pour avoir encouragé l'esprit d'innovation dans le département à travers le recrutement de plusieurs thésards et à travers divers projets. Je remercie également Sébastien Habault pour son aide, ses conseils et la bonne humeur qu'il a su instaurer dans l'équipe DVN, sans oublier Sylvain Coppéré qui, bien que tôt parti de l'équipe, m'a beaucoup aidé au début de la thèse et m'a encouragé à rencontrer les opérationnels de la logistique. Chez Renault, j'ai également côtoyé plusieurs autres personnes auprès desquelles j'ai beaucoup appris sur la logistique des voitures et des pièces et sur les systèmes d'information. Merci donc à Martin, Catherine, Valérie, Philippe, Olivier, Alain, Joelle, Laurent et toute l'équipe du DOP.

Et même si j'ai passé la plupart de mes journées au technocentre de Renault, j'ai pu apprécier mes courts passages au LGI. Je regrette d'ailleurs que je n'ai pas pu passer plus de temps là- bas tellement les rencontres et les discussions étaient passionnantes et les personnes accueillantes. Un grand merci donc à Anne, Sylvie, Delphine, Lina, Toufik, Oualid, Emna, Myriam, Evren, Asma et tout le LGI. C'était un plaisir pour moi de vous avoir connus.

Le meilleur est pour la fin dit-on. Je suis et je serai toujours reconnaissante à mes parents sans qui je n'en serai pas là aujourd'hui. Maman, papa, je vous remercie pour votre amour et votre soutien inconditionnel tout au long de cette carrière. Je vous aime. Une mention particulire pour mon petit frère, qui grâce à lui, j'ai eu l'idée de faire une thèse.

Enfin, tous les mots ne suffiront pas pour remercier la personne qui m'a le plus encouragé à faire cette thèse, qui a toujours cru en moi, qui a été présente à mes côtés dans les hauts et les bas, dans les joies et les tristesses, qui m'a appris à prendre du recul et à faire la part des choses, mon cher mari Majed...Merci...Cette thèse est aussi la tienne...

Je dédie également la thèse à mon adorable fils Mohamed qui a supporté mon absence et mes sauts d'humeur. Une mention spéciale à mon bébé Eya qui va assister à ma soutenance... Mes trésors, un jour vous lirez peut être la thèse de votre maman et vous en serez fiers...

Résumé

A cause de leur aspect stratégique et des divers challenges qu'ils représentent en termes de modélisation et de résolution, les problèmes de localisation et de conception de réseaux ont été largement étudiés par les spécialistes en recherche opérationnelle. Par ailleurs, bien que les études de cas dans ce domaine soient rares dans la littérature, plusieurs travaux récents ont intégré certains aspects opérationnels afin de rendre ces problèmes d'optimisation plus réalistes.

L'objet de notre projet de recherche est le développement d'un modèle de conception d'un réseau de distribution prenant en compte plusieurs aspects opérationnels inspirés d'une étude de cas dans le domaine de l'automobile. Bien que nos choix de modélisation soient motivés par cette étude de cas, ils restent applicables dans d'autres secteurs industriels. Le réseau de distribution considéré se compose de trois niveaux : les usines au premier niveau, les centres de distribution (CD) au deuxième niveau et les clients au dernier niveau. Nous supposons que le nombre et la localisation des usines ainsi que le nombre et la localisation des clients sont connus. Etant donné la demande des clients et une liste de CD potentiels, l'objectif est de déterminer la localisation des CD à ouvrir et d'y affecter les clients de manière à minimiser le coût total.

Nos contributions par rapport aux travaux existants concernent la modélisation et la résolution du problème ainsi que les tests numériques effectués. En termes de modélisation, nous considérons divers aspects opérationnels qui ont été pris en compte séparément dans la littérature mais jamais combinés dans un même modèle. Plus particulièrement, nous introduisons un "clustering" en prétraitement afin de modéliser les tournées de camions. Nous intégrons également des contraintes de volume minimum sur les axes de transport pour assurer l'utilisation de camions pleins, des contraintes de volume minimum et de capacité maximale sur les centres de distribution, des contraintes de distance de couverture maximale et des contraintes d'uni-affectation. Par ailleurs, nous étudions une extension multi-périodes du problème en utilisant un "clustering" dynamique pour modéliser des tournées de camions multi-périodes. En termes de résolution, comme le problème étudié est NP-difficile au sens fort, nous proposons différentes méthodes heuristiques performantes basées sur la relaxation linéaire. A travers les tests effectués, nous montrons que ces méthodes fournissent des solutions proches de l'optimale en moins de temps de calcul que l'application directe d'un solveur linéaire. Nous analysons également la structure des réseaux de distribution obtenus et nous comparons les résultats issus de plusieurs versions du modèle afin de montrer la valeur ajoutée du "clustering" ainsi que de l'approche multi-périodes.

Mots clés: supply chain, conception d'un réseau de distribution, localisation-routing, contraintes de volume minimum, relaxation linéaire, industrie automobile

Abstract

Facility location and network design theories have been widely studied by OR researchers during the last decades. This interest might be explained by the strategic importance of these problems for industrial companies as well as by the research challenges to be tackled to model and solve them. Although real–life case–studies reported in the academic literature are rather scarce, several recent works have focused on improving the practical relevance of facility location models by considering operational features.

The purpose of our research project is to develop a distribution network design model taking into account many realistic features arising from a case–study in the field of car distribution. Our modeling choices were motivated by our practical application but can be relevant in other industrial contexts. The overall network structure consists of three levels: plants in the first level, distribution centres (DCs) in the second one and customers in the third one. We assume that the number and location of the plants as well as the number and location of the customers are fixed. Given the demand of customers and a list of potential DCs, our main concern is to locate DCs and to assign customers to them in such a way as to minimize the total distribution costs.

Our contributions relate to the modeling of a real–life problem, the development of efficient solution methods and the analysis of the obtained numerical results. In terms of problem modeling, we integrate various operational features that were considered separately in the literature but have never been combined in a same model. Namely, we introduce a clustering–based approach to model vehicle routing, minimum volume constraints to ensure full truckload transport, minimum and maximum throughput constraints on DCs, maximum covering distance constraints and single sourcing restrictions. Furthermore, we study a multi–period extension of the problem using an original dynamic clustering to model multi–period vehicle routing. In terms of solution method, as the problem we study is NP–Hard in the strong sense, we propose efficient heuristic procedures based on various types of linear relaxation. Through our numerical experiments, we show that the implemented heuristics offer near–optimal solutions with less computational effort than applying an exact MIP solver. We also analyze the structure of the obtained networks and compare the results of several versions of the model, highlighting the value of integrating a pre–processing clustering step and of using a multi–period approach.

Keywords: supply chain, network design, location-routing, minimum volume constraints, linear relaxation, automotive industry

Contents

Chapter 1

Introduction

The objective of logistics is to deliver products to customers at the best service level and the lowest possible cost, in other words to be the "*better, faster, cheaper, closer*" [Christopher, 2005]. Logistics deals with the management of activities related to procurement, inventory, transport and distribution. Supply chain could be considered as an extension of logistics integrating flows from the n^{th} tier supplier to the final customer. Two main concepts are closely linked to supply chain: supply chain management (SCM) and supply chain planning (SCP). Several definitions of SCM were suggested over the extensive literature devoted to this research field. For instance, [Christopher, 2005] defines it as "*the management of upstream and downstream relationships with suppliers and customers to deliver superior customer value at less cost to the supply chain as a whole*".

SCP consists in anticipating the future requirements in order to balance supply and demand and to deliver customer orders at the lowest possible cost. It aims at answering some questions that could be asked by a supply chain manager when thinking of the planning of its supply chain: where to manufacture this new product? Where to locate my warehouses? when and how to deliver the demand of this customer? To help industrial professionals optimizing their supply chain, researchers and software providers have been proposing SCP tools, also referred to as advanced planning systems (APS). Most of these software are structured into modules, each one covering a specific planning task. These tasks are themselves classified according to three planning levels [Anthony, 1965]: long–term (strategic), mid–term (tactical) and short–term (operational). Fig. 1.1 shows a possible illustration of the supply chain planning matrix, based on two dimensions: the planning horizon and the supply chain process.

The focus of the present work is on distribution network design which is a problem occurring at the strategic/tactical planning level as it involves mid to long–term decisions (location of facilities and assignments of customers to facilities). In this introductory chapter, we first present the industrial context of the study, namely the automotive supply chain with a focus on car distribution. Then, based on the main features highlighted from the industrial context, we describe the content

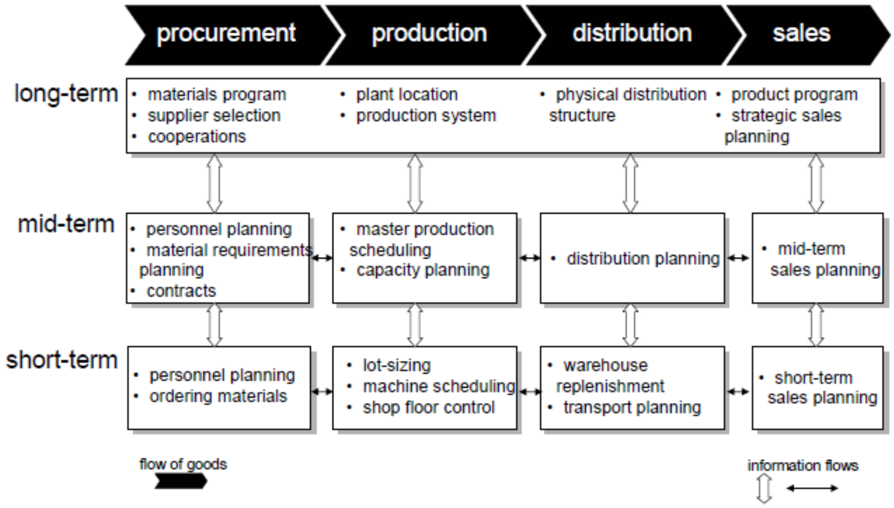


Fig. 1.1 Supply chain planning matrix. Source [Fleischmann et al., 2008]

of this research as well as the major scientific contributions as compared to the existing literature. We finally detail the outline of the present dissertation composed of five chapters.

1.1 Industrial context

1.1.1 The automotive supply chain

1.1.1.1 General overview

Modern cars are complex technological products involving a large number of mechanical and electronic sub–components. Accordingly, the automotive industry uses a large variety of production units (forge, foundries, mechanics, assembly, etc) but the car manufacturer outsources many of these activities to its suppliers. The resulting supply chain network is thus very complex due to the introduction of many levels of suppliers (1^{st} tier, 2^{nd} tier, 3^{rd} tier and even more) in addition to assembly plants, logistical compounds and customers. The members of this network should work as partners and efficiently coordinate flows in order to ensure high–quality products and good customer service while remaining economically competitive.

From the car manufacturer point of view, the supply chain is composed of three levels: inbound supply chain, manufacturing and outbound supply chain (see Fig. 1.2).

Inbound supply chain consists of flows and operations for parts from 1^{st} tier supplier plants to assembly plants. The collection of components from suppliers can be done through direct flows or using milk–run deliveries, i.e. consolidation of parts from several suppliers in one round trip (see Fig. 1.3). Once collected, parts are delivered to inbound distribution centres then sent to assembly plants where cars are manufactured. Each plant is designed to operate at a particular production rate measured using the number of vehicles produced per hour. For instance, Tangier

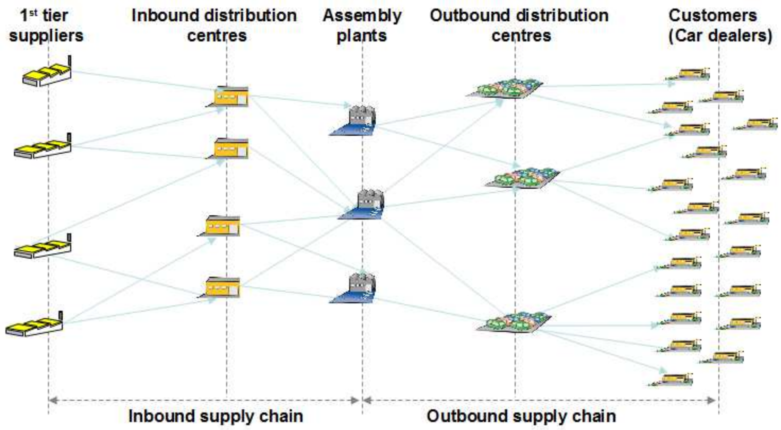


Fig. 1.2 Automotive supply chain from the car manufacturer point of view

Renault plant which produces Dacia Lodgy and Dacia Dokker has a current rate of 30 vehicles per hour and 170000 cars per year.

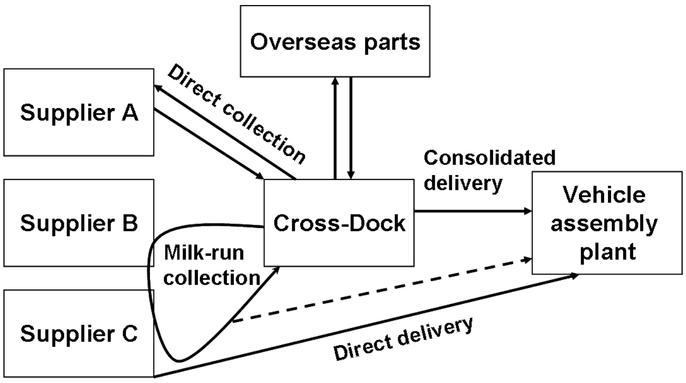


Fig. 1.3 Inbound supply chain process Source [Miemczyk and Holweg, 2004]. Cross–docking is a well–known practice in logistics, consisting in unloading materials from ingoing trucks and loading them in outgoing trucks with little or no storage in between.

Outbound supply chain consists of flows and operations for finished cars from assembly plants to car dealers which are the major end customers of the whole automotive supply chain. Other customers could include car rental companies and other public or private companies. All finished products wait on factory compounds (see Fig. 1.4 for an example of factory compound) to be loaded in trucks (see Fig. 1.5), trains (see Fig. 1.6), barges (see Fig. 1.7) or vessels (see Fig. 1.8) then routed to outbound distribution centres. They are then transported by truck to the final customers.

1.1.1.2 The example of Renault

Renault is a leading European car manufacturer producing more than 40 car types. The company is present all over the word through manufacturing sites, logistical compounds and commercial

Fig. 1.4 Factory compound of Renault Flins (France)

Fig. 1.5 Auto transport by truck

subsidiaries, Fig. 1.9 shows the assembly plants of Renault in the world (or those of the partner Nissan used by Renault).

The initial strategy of the company was to specialize plants by product but in order to conquer new markets, it was necessary to produce cars near the location where they will be commercialized. For instance, the Brazilian factory manufactures many car types which are already produced in European plants (Logan, Sandero, Duster, Megane2, Master). The company uses a mixed build–to–order (BTO) / build–to–stock (BTS) strategy, depending on the market and its specificities (refer to appendix A for a further discussion about BTO, BTS and the history of the automotive industry). In the sequel, we focus on the European continent as our practical application is related to this region (see Fig. 1.10). In European countries, inventories related to build–to–stock products

Fig. 1.6 Auto transport by train

Fig. 1.7 Auto transport by barge

are mainly managed by car dealers either on their own locations or on storage compounds they rent to have more space.

Within the supply chain management department of Renault there is a team devoted to outbound logistics. The objective of this team is to deliver cars on time, at the least cost and without any damage. Its short-term (operational) missions consist in the tracking of the distribution process and the control of costs. Its mid and long-term missions comprise among others developing and improving the distribution schemes, defining outbound logistics strategies and working on optimization/simulation tools. Cost reduction was identified as a major driver in the strategic plan of

Fig. 1.8 Auto transport by vessel

Fig. 1.9 Assembly plants of Renault in the world (or those of the partner Nissan used by Renault)

the company "Renault 2016: Drive the change". As inbound and outbound logistics costs account for 6% of the total delivery cost[1] of a finished car and as the only turnover of outbound logistics is estimated to hundreds of millions of euros, improving the supply chain performance through the optimization of the supply chain planning (SCP) seems to be a necessity for the company. In this context, comes the present work which focuses on the design of distribution networks taking into account the main specific features of car distribution.

1.1.2 Specific features of car distribution

The objective of outbound automotive supply chain is to deliver finished cars from assembly plants to car dealers. We show in Fig. 1.11 an example of outbound supply chain process for

[1]Total delivery cost of a product is the cost of manufacturing and delivering it.

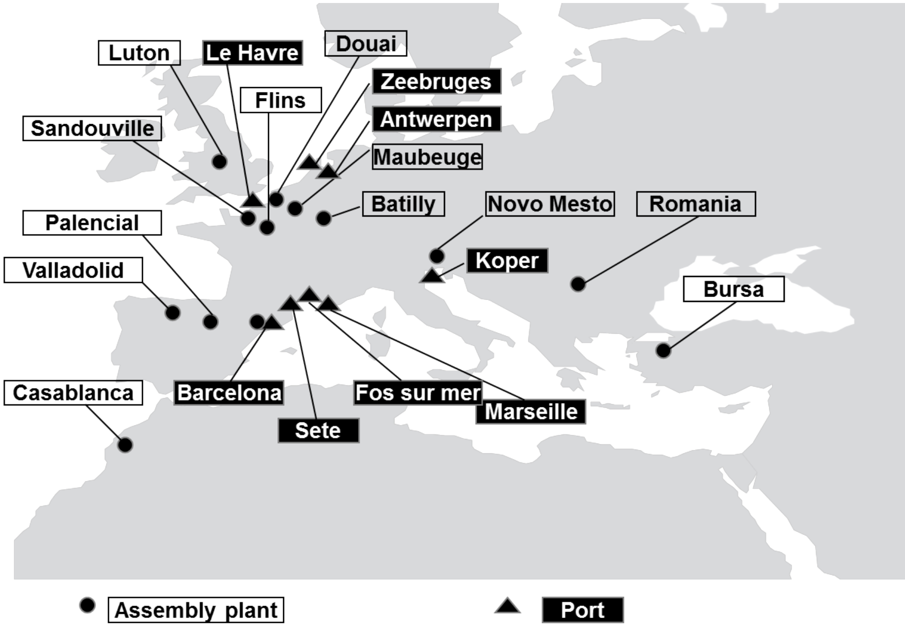


Fig. 1.10 Assembly plants supplying the European countries and main maritime ports

build–to–order cars. When build–to–stock inventories are managed by car dealers (as it is the case for Renault in European countries), the related distribution flows can also be managed through the process illustrated in Fig. 1.11 using the addresses of storage compounds instead of those of car dealers. In the sequel of the study, we will thus focus on build–to–order products.

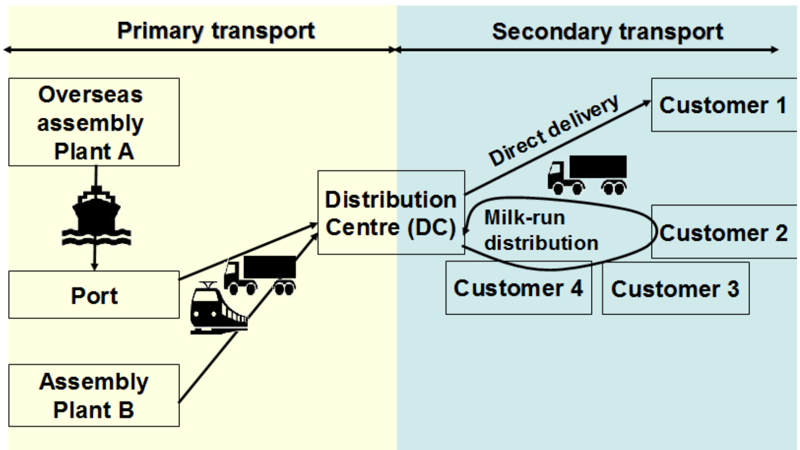


Fig. 1.11 Outbound supply chain process for build–to–order cars

The whole distribution process is mainly split into two sub–processes: primary transport from plants to distribution centres (DCs) and secondary transport from distribution centres to car dealers. One of the main advantages of using distribution centres is the consolidation of flows in order to make the best possible use of transport capacities. The main volume routes are from plants to distribution centres as these flows correspond to the aggregation of many customer demands transiting through intermediate DCs. This is why high–capacity modes of transport such as vessels and trains could be used, especially when manufacturing sites are scattered over several countries. In addition to their non–polluting reputation, these modes are known to be inexpensive and reliable for long distance, high volume flows. Once arrived at distribution centres, build–to–order cars are not stored but only held for a short transit time (typically a few days) before being sent to car dealers. In this second step (secondary transport), only trucks are used to deliver cars to car dealers as the corresponding transport routes are short and usually in urban areas.

The management of distribution centres as well as the preparation and handling activities are usually outsourced. In that case, distribution centres are owned by third party–logistics so that, from the car manufacturer point of view no heavy installation costs are incurred when using new distribution facilities. Only a unit transit cost has to be paid to the supplier each time a car goes through a DC. The determination of this unit cost is done during the establishment of the contract between the car maker and the logistics supplier through a commercial negotiation process. It depends on the total throughput that has to be handled in the given DC and only applies if this throughput is between a minimum volume and a maximum capacity.

Another distinctive feature of car distribution is the particularity of the transported products. In fact, cars are expensive, fragile and bulky products transported by dedicated trucks with limited capacities (see Fig. 1.5). Typically, a truck can carry up to 8 Renault Clio or 10 Renault Twingo. This implies a difficult backload[2] management with great impact on costs. In order to minimize empty kilometres, backloads have to be arranged by carriers either by contacting other competitors or by considering other flows for the same car manufacturer. Moreover, dealing with voluminous products results in the fact that load efficiency is a key parameter in car distribution. Thus, making the best possible use of transport capacities and in particular ensuring full truckload transport is one of the priorities of automotive outbound logistics. To show the impact of load efficiency on cost, we illustrate in Fig. 1.12 an example of the unit transport cost as a function of the transported volume when the frequency of shipping trucks is one week. In this illustration, we do not include the storage cost of the cars waiting to be loaded on trucks. This cost is indeed negligible as compared to the transport cost.

The unit transport cost is computed by dividing the cost of a truck (here we set it to 1000 Euro) by the load of the truck, i.e. the weekly volume on the transport link. Fig. 1.12 shows that the unit cost sharply increases when the weekly transported volume is less than the maximum capacity of a truck (10 in the example). For instance, if the truck transports only one car, the unit cost raises up to 1000 Euro, which is a significant amount as compared to the sale price of a car. When the weekly volume is equal to the maximum capacity of one truck (10), the unit cost reaches its minimum value (100 Euro). When the volume exceeds the maximum capacity of one truck, the unit transport cost is computed by dividing the cost of shipping n trucks (if n trucks are needed)

[2]Backloads are loads transported on the return journey of a delivery truck in order to reduce empty kilometres.

by the weekly transported volume, which leads to the dashed line in Fig. 1.12. In that case, we notice that the cost slightly increases but as it remains very close to 100, we consider it as equal to 100, which leads to the solid line in the same figure. This is indeed the assumption usually made by practitioners when designing a car distribution network.

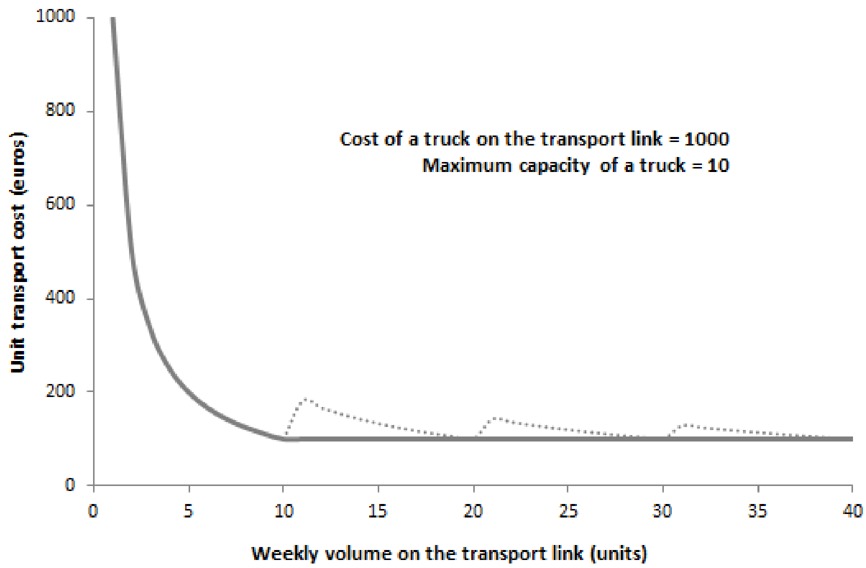


Fig. 1.12 Unit cost as a function of the weekly transported volume on a given transport link.

It is thus of high importance to consolidate enough volume on each opened transport link in order to ensure full truckloads within the allowed waiting time. For primary transport from plants to distribution centres, this involves reducing the number of transport links starting at a given plant. However, for secondary transport from distribution centres to car dealers, this is not always possible. The demand of some car dealers could be indeed below the threshold corresponding to reaching full truckload within the maximum waiting time allowed at a distribution centre. This is why it is necessary to group deliveries: a given truck starting from a distribution centre may have to visit two or three customers before coming back to the distribution centre. This is a general practice in car distribution and in many other industries.

1.2 Research focus and main contributions

1.2.1 Research focus

In the present work, we focus on modeling and solving a multi–product distribution network design problem taking into account the operational features discussed in §1.1.2. By introducing these operational features, we intend to develop a detailed model in terms of costs and constraints. Thus, we limit the scope of the distribution network to three levels: plants, distribution centres and customers. In case the transport from plants to distribution centres involves maritime shipping and several transshipments, we consider the port of arrival as the sourcing point instead of the originating plant. We use a mid–term planning horizon (typically one year) and we assume that the number and location of the plants as well as the number and location of the customers are fixed. Given the demand of each customer for each product, our main concern is to locate the DCs and to assign customers to them in such a way as to minimize the total distribution costs. In a first MIP (Mixed Integer Programming) model presented in chapter 3, the input demand is considered as static over the whole planning horizon whereas in chapter 5, we study, through a multi–period network design model, the robustness of location and assignment decisions when demand varies over time–periods according to some seasonality pattern.

One of the major issues addressed in the DNDMVD model presented in chapter 3 is truck loading. As we previously mentioned, due to the great impact of load efficiency on car transport costs (see Fig. 1.12), it is of primary importance to make the best use of transport capacities. Thus, deliveries from distribution centres to customers have to be grouped in order to optimize truck loading while ensuring acceptable lead times. This is introduced in our model through a two–step location–routing approach. The first step is a pre–processing clustering procedure whose main objective is to construct groups (clusters) of close customers meeting a minimum required volume (see chapter 2). Based on the results of this clustering, a distribution network design problem is formulated and solved in the second step of the approach. Using pre–processing clustering is a way to model routing in a supply chain network design problem while keeping a manageable size for the optimization problem. This allows dealing with real–life instances involving a large number of customers. The DNDMVD model also involves minimum volume constraints conditioning the opening of transport links. By carefully setting the minimum volume values, we will ensure that full truckload transport will be possible at the operational level and will thus contribute to the reduction of transport costs. Finally, we incorporate in our DNDMVD model several additional constraints, namely the total throughput of an opened distribution centre should be between a minimum volume and a maximum capacity, all transport demands having the same source and same destination are routed through the same distribution centre and the route distance between a distribution centre and any cluster of customers it serves should be less than a given limit called the maximum covering distance.

The main difference between the DNDMVD model developed in chapter 3 and the extension introduced in chapter 5 is demand variation. In fact, in the latter case, we assume that demand varies from period to period according to some seasonality pattern (the number of periods should be determined according to the context of the industrial application). The implemented model could also be used to express demand uncertainty through discrete scenarios with fixed probabilities of occurrence. The approach we propose is indeed very similar to the one employed in two–stage stochastic location problems where location decisions are made at the first stage and assignment decisions occur after random parameters become known.

1.2.2 Main contributions

The distribution network design problem we consider is inspired by a case–study in the automotive industry. Thus, one of the main contributions of our work is to provide some practical results in the field of facility location applied to supply chain management. Namely, as mentioned in [Melo et al., 2009a], there is a lack of real–life applications reported in the related literature. This could mainly be explained by the difficulty of data collection and the reluctance of managers to accept using quantitative models in the process of strategic planning.

The model we propose in chapter 3 shares common features with the classical problems of facility location and supply chain network design. It is however significantly different due to the introduction of various extra constraints needed to model real–life requirements (constraints are detailed in §3.2). Although many of these constraints were considered separately in the literature, we could not find any work *combining them in a same model* (see the detailed literature review in §3.1). Moreover, it was pointed out in [Melo et al., 2009a] that literature about multi–period facility location problems for supply chain is rather scarce as more than 80% of the papers surveyed by the authors deal with single–period problems. Thus, another contribution of the present work is *the introduction of a multi–period distribution network design problem* in chapter 5. More precisely, we propose *an original way of modeling multi–period location–routing through a dynamic clustering procedure* (see §5.2.1.2 for a detailed description of this procedure). There are in fact only a few papers simultaneously addressing routing and multi–period aspects in facility location and supply chain network design problems (some references are cited in §2.1.2) whereas static location–routing problems are more extensively studied. The choices that we made in our model were motivated by our case–study in the field of car distribution (see §1.1.2 for more details) but we would like to point out that they can be relevant in other contexts of application. In fact, routing, optimization of truck loading, maximum covering distances or multi–period demand are relevant features to take into account when modeling a generic distribution network design problem.

Furthermore, we develop in §3.3.4 a complexity analysis that shows that the studied problem is NP–complete in the strong sense. We thus propose efficient heuristic procedures with the aim of reducing computation times when dealing with large–size instances. We implement various types of linear relaxations in order to determine location and assignment variables, both sets of variables being required to be binary. The implemented heuristics are applied to the model based on static demand as well as to its multi–period extension. In the literature, we survey only a few works dealing with linear–relaxation based heuristics to solve facility location problems subject to minimum volume constraints (see §3.1). Moreover, most of these works study single–period models and mainly focus on rounding location variables either because assignment variables are not problematic in the solution procedure or because they are continuous.

Finally, in order to investigate computation times and to analyze the structure of the obtained networks, we carry out several numerical experiments based on industrial data from our practical application in the field of car distribution. Various tests show that obtaining optimal solutions using a state–of–the–art MIP solver sometimes requires extensive computation times and may run out of memory due to the difficulty of the problem. In these cases, it is interesting to apply the proposed heuristics, which provide good quality solutions within short computation times. Moreover, the use of a highly constrained model either in the case of static demand or in the case of seasonal demand leads to several trade–offs between conflicting constraints. This is why we propose a thorough qualitative analysis: in chapter 3 (§3.2.6), we discuss the influence of the various constraints on the specialization and the number of the opened DCs as well as on the existence of a feasible solution. In chapter 5, we analyze the main trade–offs resulting from demand variation in a highly constrained supply chain network design problem and show how to overcome the resulting ”infeasibilities”. We also study the outputs of the multi–period problem as compared to its single–period version (§5.4.1.2) and highlight the advantages of using static assignment decisions (§5.4.1.3).

1.3 Dissertation outline

We now present the general structure of the dissertation, which is summarized in Fig. 1.13.

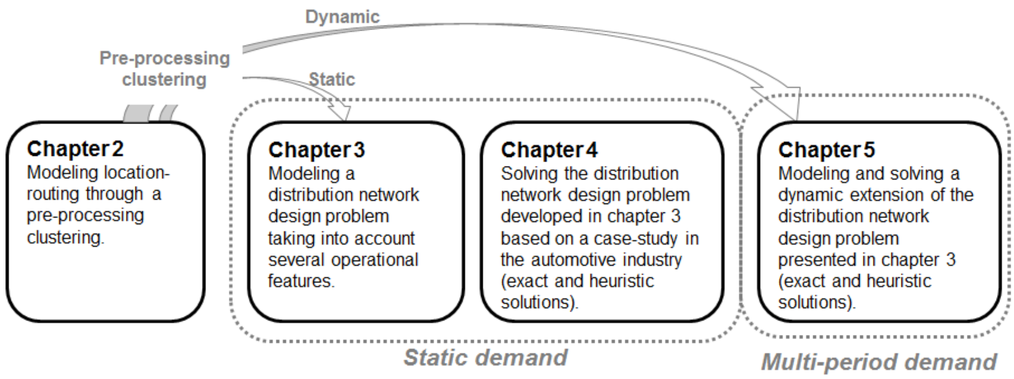


Fig. 1.13 General structure of the dissertation.

- Chapter 2: This chapter deals with a location–routing approach based on a pre–processing clustering. We first provide a general overview of location–routing problems (LRP) and propose to classify them according to their way of integrating routing in the main strategic problem of facility location. Then, we detail the three steps of the sequential location–routing approach we propose. For the clustering step, we present two methods: an exact one and a heuristic one which can be used when the exact method does not apply due to prohibitive computation times. Each of these methods is illustrated through numerical tests based on our practical application in the field of car distribution.
- Chapter 3: This chapter presents the main considerations taken into account when modeling

the distribution network design problem under study (DNDMVD). A literature review is proposed on facility location and supply chain network design problems with minimum volume and distance constraints as these two types of constraints are among the major characteristics of the implemented model. We then discuss into detail the features that motivated our modeling choices in terms of product representation, transport routes and various constraints. We also provide a qualitative analysis of the main trade–offs to be achieved in the problem. Finally, we present the mathematical formulation of the problem using a Mixed Integer Program (MIP).

- Chapter 4: This chapter is devoted to solving the DNDMVD problem described in chapter 3. We first present the case–study and the test instances we use. Then, we analyze the solution obtained with a state–of–the–art MIP solver on the reference dataset and study the impact of varying the main parameters of the problem on computation times. We show that in some cases, computation times can significantly increase and the program may even run out of memory. Thus, we propose MIP–based heuristic procedures and demonstrate through extensive numerical experiments their performance as compared to a commercial solver applied to the original MIP.

- Chapter 5: This chapter considers an extension of the DNDMVD model addressed in chapter 3 where demand varies over time–periods according to some seasonality pattern. In the first section of the chapter, we provide a literature review on dynamic facility location and supply chain network design problems. In the second section, we introduce a new aggregation approach and a dynamic pre–processing clustering in order to evaluate delivery routes and costs in each time–period. We present in the third section the mathematical formulation of the multi–period problem before providing in the fourth section a numerical study using a state–of–the–art commercial solver and based on a case–study in the field of car distribution. Finally, we introduce a MIP–based heuristic procedure to solve the multi–period problem and prove its efficiency as compared to the straight application of a state–of–the–art solver.

Chapter 2

Modeling location routing

The main objective of our work is to optimize the design of a distribution network featuring three levels: plants, distribution centres and customers. Final decisions include the location of distribution centres (DCs) and the assignments of customers to the selected DCs. Our work is thus related to strategic supply chain planning as location of distribution centres are mid to long–term decisions. In this context, one of the major difficulties is to define the perimeter of the study and its limitations, i.e. to identify which operational and tactical features should be taken into account and which are less important. When the problem is derived from a real–life case–study, this difficulty is more noticeable because decision–makers usually want to model every operational detail in the same optimization tool. However, this would lead to the formulation of mathematical models which are likely to be very challenging to solve, especially for the large size industrial instances we would like to tackle. We thus have to wisely select the features that could influence the main target of the study, i.e. DC locations and customer assignments to DCs.

An important feature that we propose to consider in our problem is vehicle routing, i.e. defining delivery routes from DCs to customers. In real–life operations, the demand of some customers could be insufficient to ensure full truckload transport within the maximum allowed waiting time at distribution centres. In this case, we may resort to grouping deliveries to avoid less–than–truckload transport: a given truck starting from a distribution centre may have to visit several customers before coming back to the distribution centre. Through delivery grouping, the demand of many customers is consolidated and thus truck loading can be optimized and unit transport cost reduced. The unit transport cost is indeed evaluated as the total cost on a given transport route divided by the load of the truck. If the truck is not fully loaded, then the unit cost per transported product could significantly increase. This observation is even more important for supply chains where products are voluminous like car distribution. In the introduction of the present work (§1.1.2), we discussed into details the importance of load efficiency in car distribution and its impact on delivery costs (see Fig. 1.12).

We have then to answer the following question: which customers to serve together in the same route? This leads to a Vehicle Routing Problem (VRP aims at determining the best delivery routes to serve customers using multiple vehicles). This problem as well as the problem of facility location (FLP) are rather classical and well–studied topics in the literature (FLP consists in locating sites in a given space, according to certain criteria, in order to satisfy the demand of customers). In this chapter, we aim at integrating these two issues into a more complex problem called the location–routing problem (LRP). The resulting model not only aims at locating sites and assigning demand points to them but also at determining the delivery routes from facilities to customers in such a way as to minimize the whole distribution cost. We thus propose a location–routing approach based on a pre–processing clustering procedure. The main objective of clustering is to construct clusters of customers in such a way as to minimize distribution costs and to ensure full truckload transport within acceptable waiting times.

By introducing a clustering approach, we intend to develop a detailed representation of costs and constraints in order to make the right location and assignment decisions that allow at the operational level to construct routes meeting all the transport constraints. Furthermore, using a clustering–based sequential approach instead of explicitly modeling routes through binary variables enables us to reach a good trade–off between a detailed representation of routing in the model and its computational tractability. Accordingly, we can deal with large size multi–product industrial instances either in the first model based on a static customer demand (chapter 3) or in the second model using a multi–period customer demand (chapter 5).

The present chapter is organized as follows: in the first section, we propose a literature review on location–routing problems (LRP) including a classification of LRP based on their approach of integrating routing in the master location problem. In the second section, we present the clustering–based location–routing approach that we intend to use before detailing in the third and fourth sections exact and heuristic clustering methods.

2.1 Literature review on location-routing problems (LRP)

2.1.1 General overview of location–routing

Integrated models combine decisions at different planning levels (strategic, tactical and operational) and aim at studying the trade–off between interdependent logistical components. In the case of a strategic facility location problem, [Shen, 2007] and [Daskin et al., 2003] point out that neglecting some operational aspects can lead to sub–optimality and thus to more expensive solutions. [Salhi and Rand, 1989] was among the first works to study the effect of ignoring routes when locating depots. However, few works addressed the effective economic impact of jointly modeling operational and strategic decisions. Exceptions could be found in [Shen and Qi, 2007] and [Javid and Azad, 2010].

In the present work, we focus on location–routing problems (LRP). The reader is referred to [Min et al., 1998] for an early comprehensive literature review on this topic. [Min et al., 1998] provide a two–way classification: first in terms of general problem features (single vs two–stage levels, deterministic vs stochastic demand, number of facilities, single vs multi–period, etc.) then in terms of solution methods (exact vs heuristic one). They point out that about 3/4 of the reviewed works concern deterministic models. Moreover, most of them deal with static problems and only one reference for dynamic LRP is mentioned in the paper ([Laporte and Dejax, 1989]). A more recent survey of location–routing problems can be found in [Nagy and Salhi, 2008]. The authors indicate that LRP "aims at solving a facility location problem (the master problem), but in order to achieve this, we simultaneously need to solve a vehicle routing problem (the subproblem)". They propose several classification criteria but their main focus is on exact/heuristic solution methods and dynamic/stochastic problems. Furthermore, they note that application–oriented papers account for a fifth of the LRP literature. They mention case–studies in consumer goods distribution, health, military, communications but not in the automotive industry.

In the context of our work, we will also consider the use of a set–partitioning formulation (see §2.3 below). Set–partitioning is a classical way of formulating vehicle routing problems (VRP). It consists in two steps: a first step generating all subsets (potential routes) of a given customer set then a second step to select the subsets optimizing the objective function while meeting the routing constraints. [Laporte, 1992] points out that a set–partitioning approach leads to two difficulties: the large number of binary variables in most real–life applications and the difficulty of computing route costs as this often involves solving traveling salesman problems (TSP). Among others, [Taillard, 1999] and [Baldacci et al., 2008] use set–partitioning formulations to solve VRP. In the location–routing literature, we could not find many works dealing with a set–partitioning formulation. [Berger, 1997] is the only reference we could find using this approach. The author formulates an integrated location–routing problem where a delivery vehicle may not be required to return to the distribution center after the final delivery is made. Although the problem does not involve other constraints, the author reports a difficulty in solving it and implements a column generation procedure as complete enumeration of all possible columns of the problem is prohibitive.

2.1.2 Classification of location–routing problems

In this subsection, we propose a classification of LRP works based on their way to integrate routing in the main strategic problem of facility location or network design. From the review of [Nagy and Salhi, 2008] and the literature relative to LRP, we can mainly identify three modeling approaches:

1. Using continuous approximation: A literature review on continuous approximation models in freight distribution is provided in [Langevin and Mbaraga, 1996]. The authors present continuous approximation as "relying on concise summaries of data and analytic models" whereas mathematical programming "relies on detailed data and numerical methods". In

[Daganzo, 1984], continuous approximation in the context of traveling salesman problems (TSP) is presented as "a descriptive effort attempting to give length formulas under different conditions" as opposed to "prescriptive research efforts attempting to derive algorithms for the construction of optimal or near-optimal tours". For instance, continuous approximation uses continuous functions to describe demand point distribution rather than the exact location of each demand point. We can cite among others the formula developed by [Beardwood et al., 1959] in order to estimate the length L of a traveling salesman visiting N points uniformly and independently scattered in a region of area A: $L = k\sqrt{AN}$, where k is a constant and N has to be large enough. In the field of location–routing, we can refer to [Shen and Qi, 2007] who use continuous approximation to estimate the optimal routing costs. The major disadvantage in this case is the need to resort to strong assumptions such as uniform demand distribution in order to simplify the resulting complex expressions.

2. Explicitly modeling the choice of routes as decisions in the optimization problem (see e.g. [Wu et al., 2002, Yu et al., 2001, Zarandi et al., 2013]): this usually results in the formulation of large size mixed integer programs, often leading to computational difficulties. This might explain why the literature on multi–product and multi–period LRP is scarce. We only could identify three papers addressing multi–product location–routing: [Yi and Ozdamar, 2007, Sajjadi, 2008, Afshar and Haghani, 2012] and five papers dealing with multi–period location–routing: [Yi and Ozdamar, 2007, Afshar and Haghani, 2012, Albareda-Sambola et al., 2012, Laporte and Dejax, 1989, Prodhon, 2011]). The scarcity of multi–product and multi–period models could be explained by the fact that LRP is already a complex problem and thus introducing several products or several time–periods leads to prohibitive computation times when using state–of–the–art commercial MIP solvers. Table. 2.1 presents the largest computational instances presented by the above–mentioned multi–product and multi–period LRPs. To tackle large size problems, four of the six papers illustrated in the table resort to heuristic methods. Notice that the location–routing problem is known to be NP–hard as it is the combination of two NP–hard problems (facility location and vehicle routing).

3. Using a pre–processing clustering method (clustering first and location/assignment second), as proposed in [Barreto et al., 2007]. The approach presented in this work consists of four steps: in the first step, customers are grouped into clusters according to a maximum volume constraint (the delivery vehicle capacity). The authors propose several heuristic approaches based on hierarchical[1] and non hierarchical[2] clustering. In the second step, a Traveling

[1]Hierarchical clustering is a heuristic approach aiming at grouping points into clusters beginning with clusters consisting of only one point and converging to the formation of one single cluster. The use of a capacity limit avoids the consecutive joining of clusters, acting as a stopping criterion of the procedure.

[2]Non hierarchical clustering is a heuristic approach aiming at grouping points into p clusters where p is known a priori (reaching p clusters is the stopping criterion in that case).

Salesman Problem (TSP) is solved for each cluster without considering distribution centres. In the third step, routes are improved using a local search. Finally, clusters are assigned to depots by solving a capacitated facility location problem. Other works integrate clustering and location decisions in a same mathematical model (see e.g. [Miranda et al., 2009] for a "Hub and Spoke" cost structure considering transport cost between depot and cluster hub, as well as transport cost between cluster hub and its customers). A clustering approach can also be used to build heuristic solutions to the location–routing problem (see e.g. [Escobar et al., 2013, Mehrjerdi and Nadizadeh, 2013]
[Ambrosinoa et al., 2009]).

Although using clustering as a pre–processing step is not a common way to handle routing in a supply chain network design problem, we chose to apply it in our study in order to have a good approximation of the routing costs while keeping a manageable size for the optimal location problem. This allows dealing with real–life instances involving many customers, many products (chapter 3) and even multiple periods (chapter 5). As compared to [Barreto et al., 2007], our approach shows the following differences/contributions:

- Introducing new constraints in the clustering procedure: minimum volume per cluster and maximum number of customers per cluster.
- Proposing an exact clustering method based on an overall objective function (sum of route lengths) and using a set–partitioning formulation in addition to a hierarchical heuristic clustering.
- Traveling Salesman Problems (TSP) are solved including distribution centres to calculate the assignment costs used later in the distribution network design problem. This is in contrast with [Barreto et al., 2007] who do not take into account the connection with centres.
- Numerical experiments are proposed for real life data whereas [Barreto et al., 2007] use only theoretical instances from the literature.
- The performance of the heuristic clustering approach is compared with the performance of the exact one based on a same objective function (sum of route lengths).

2.1.3 Contributions of our study

In the present work, we implement a location–routing approach based on a pre–processing clustering (clustering first and location/assignment second). To achieve this, we propose two clustering algorithms: an exact algorithm based on a set–partitioning formulation (see §2.3) for small–to–medium size instances and a heuristic one (see §2.4) for large size instances.

One of the main contributions of our study is to propose a location–routing approach in a multi–product and multi–period distribution network. As mentioned in the previous literature

Paper	Multi–period	Multi–product	Largest instance	Computational effort using a commercial solver	Heuristic implemented
[Yi and Ozdamar, 2007]	X	X	5 potential locations, 60 nodes (supply and demand), 8 time–periods, 4 product types (2 kinds of wounded people and 2 commodities), total of 7495 integer variables	The largest instance is solved within 140seconds	
[Sajjadi, 2008]		X	5 plants, 30 depots, 350 customers, 40 product types	No information	X
[Afshar and Haghani, 2012]	X	X	32 nodes (supply and demand), 96 time–periods, 2 product types	The largest instance could be solved within more than 2 days	
[Albareda-Sambola et al., 2012]	X		10 facilities, 70 customers, 12 time–periods	The largest instance can not be solved within 3 hours of CPU	X
[Laporte and Dejax, 1989]	X		10 facilities, 70 customers, 12 time–periods	No information	X
[Prodhon, 2011]	X		10 depots, 200 customers, 5 time–periods	Only small instances with few customers (≤20) can be solved exactly by commercial solvers	X

Table. 2.1 A look at the literature on multi–period and multi–product LRP

review (see Table. 2.1), there are few papers in these two fields as they involve computational difficulties due to the size of the resulting MIP models. On the contrary, using an approach based on a pre–processing clustering allows dealing with large size problems. This is why we chose to apply it in our study. For instance, in chapter 3, we present a multi–product distribution network design model subject to several operational constraints and we analyze the related numerical results in chapter 4 based on a real–life case–study in the automotive industry. In chapter 5, we propose a multi–period extension of the first model and an original way of modeling the impact of demand variation on delivery routes through a dynamic clustering procedure.

2.2 The location–routing approach based on a pre–processing clustering

Because of its considerable impact on delivery costs, load efficiency is one of the key drivers in transport. This is true in many sectors but even more important when the transported products are voluminous like cars (refer to the discussion presented in chapter 1 concerning the impact of truck loading on costs). Thus, in car distribution as in other fields, milk–run deliveries are used: a given truck starting from a distribution centre may have to visit several customers before coming back to the distribution centre. Deliveries are indeed grouped because some customers have such low demand that full truckloads cannot be consolidated to serve them within the maximum waiting time allowed at distribution centres. In this context, the clustering–based approach consists in running a pre–processing clustering procedure to construct clusters of customers that will be assigned to the same delivery routes for all products. As the main objective of the procedure is to ensure that each cluster is served using full truckloads within a pre–determined waiting time, clusters have to meet some minimum volume constraints.

We chose to use this sequential approach (i.e. clustering first then location and assignment decisions) rather than continuous approximation or explicit route modeling for the following reasons:

- This method allows constructing delivery routes ensuring full truckload transport.
- It leads to a detailed representation of individual customers instead of the standard approach using towns or districts as a way of demand aggregation.
- It results in a good approximation of the routing costs without increasing the size of the problem and thus it enables us to deal with large size industrial instances even in a multi–period context (chapter 5)

The clustering–based location–routing approach presented below consists of three main steps. In the first step, we have to group customers into clusters according to some given routing constraints (a minimum volume, a maximum number of customers, and a maximum distance between customers). In the second step, we compute the length of delivery routes from distribution centres (DCs) to customer groups by solving Traveling Salesman Problems (TSP). Finally, clusters are assigned to distribution centres by solving a facility location problem. Fig. 2.1 summarizes the clustering–based procedure and its main inputs/outputs.

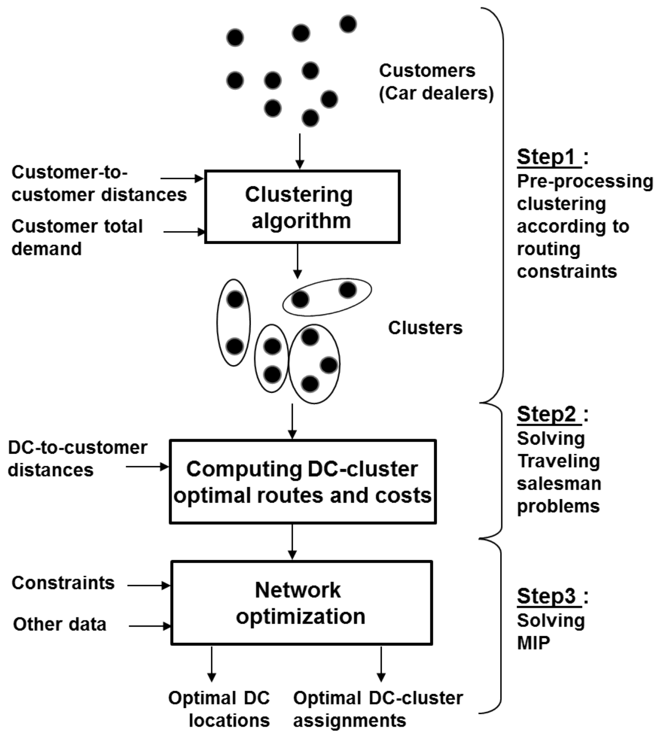


Fig. 2.1 The clustering–based approach for location–routing

2.2.1 First step: Clustering–based pre–processing

As discussed in the introduction of this section, the idea of the pre–processing clustering procedure is to construct clusters of customers that will be allocated to the same delivery routes. Customers of a same cluster have to be as close as possible from each other in order to minimize distribution costs. Moreover, clusters have to be constructed in such a way as each of them can be served using full truckloads within acceptable waiting times. Therefore, minimum volume constraints have to be considered in the clustering algorithm and have to be evaluated so as to ensure full truckloads within a pre–determined waiting time. Its value thus depends on the truck capacity and on the targeted waiting time. For instance, if we aim at delivering customers using full truckloads within five working days, we have to ensure that the total average demand of each cluster within five days is greater than a full truckload. One of the differences of our clustering procedure as compared to the one implemented in [Barreto et al., 2007] consists in the integration of minimum volume constraints that allow optimizing truck loading. We also take into account two additional constraints: the maximum number of customers per cluster and the maximum distance between customers of a same cluster. The maximum number of customers per cluster is an operational constraint which facilitates the work of the driver in particular when loading/unloading merchandises. Its value depends on the delivered products. In the automotive industry, if cars are concerned, a maximum of three customers per delivery route is usually considered.

In view of the minimum volume constraints, one important issue to overcome consists in clarifying the customer demand to be used as input to the clustering procedure, i.e. in defining the proper level of demand aggregation to be used in the clustering procedure. Three points have thus to be discussed:

- Space aggregation: using a pre–processing step means that we try to group customers before knowing their assignments to distribution centres. We thus consider the total demand of each customer without taking into account the fact that it may be split between several DCs. We construct clusters that will be used for deliveries from all DCs.
- Time aggregation: the combination of two planning levels (strategic for the network optimization problem and operational for the clustering/routing problem) leads to a major difficulty: how to evaluate customer demand in different time–horizons? Let us consider the above–mentioned example where the demand of each cluster within five working days has to be greater than the content of a full truckload. The question is how to compute the daily demand of each customer if the input of a strategic facility location problem is only the total demand for the whole planning horizon. The solution that we propose is to compute an average daily demand through a division of the total demand by the number of working days in the planning horizon.
- Product aggregation: due to the multi–product feature of the study, we have to answer the following question: should we consider as input of the clustering procedure the total demand for all products or the demand for each product separately? If the products were delivered via separate routes, we could use as input the demand per product and obtain a different clustering of customers for each product. However, our case–study involves mixed delivery routes, i.e. a truck can contain products of different types. Thus, we use the total demand quantity for all products as input to the clustering algorithm.

Another question arises from the multi–product feature of the study: how to evaluate the truck capacity that will be used when computing the minimum volume per cluster? The solution that we propose is to employ a weighted average capacity over all the delivered products with demand quantities as weights.

2.2.2 Second step: Route and cost computation

The outcome of the first step of the procedure is a list of clusters. For each cluster, we then have to compute the length of the optimal route that should be used to serve it from each potential distribution centre. This consists in determining the shortest route starting at the DC, visiting all the customers of the cluster then coming back to the DC. We have thus to solve a Traveling Salesman Problem (TSP) for each DC–cluster pair. If the number of customers is not very high (typically, for car distribution, the maximum number of customers per cluster does not exceed three), this can be done by carrying out a complete enumeration of the possible solutions.

Transport costs are then evaluated using the obtained route distances and a kilometric cost formula (see §3.2.3). Here, it is worth mentioning that our routing approach differs from the one proposed by [Barreto et al., 2007] in that it includes distribution centres in the evaluation of TSP routes and uses real–life distances and transport costs instead of theoretical euclidean formula.

2.2.3 Third step: Network optimization

Network optimization is the last step in our location–routing procedure. It uses as input data the results of the first step (i.e. a set of clusters), the transport costs computed in the second step as well as some other numerical parameters provided by the user (such as the plant–DC distances and the DC capacities). Other data like load factors[3], customer–to–customer distances and DC–to–customer distances are already employed in the first and second steps of the procedure. Using the list of clusters as final customers, a Mixed Integer Program (MIP) is formulated (see chapter 3, §3.3) in order to determine the best locations for DCs and the assignments of clusters to them with the aim of minimizing total distribution costs.

2.3 Exact clustering using set–partitioning

One of the main contributions of our clustering–based approach as compared to the one used in [Barreto et al., 2007] is the implementation of an exact optimization method based on a set–partitioning formulation. In this section, we intend to describe this approach which first generates a set of potential clusters meeting selection constraints, then, selects the clusters leading to the minimum objective function by solving to optimality an integer program. We also analyze the results of numerical experiments based on real–life data from the automotive industry. These results will serve to test the performance of the heuristic approach, presented in the next section.

2.3.1 Generation of potential clusters

The first step of the exact set–partitioning approach is the generation of all potential clusters. A cluster is a group of customers that will be served together from distribution centres. It could be modeled using a vector of 0–1 numbers, 1 if the corresponding customer belongs to the cluster and 0 otherwise (see Fig. 2.2 for an example). Potential clusters can be generated using simple procedures applying the selection constraints fixed by the user.

Selection constraints are applied when generating potential clusters in order to remove those not meeting the routing conditions fixed by the decision maker. In §2.2.1, we defined two major routing conditions, namely a minimum volume per cluster and a maximum number of customers

[3]The load factor of a product type is the number of units of this type that can be loaded on a truck. It can differ according to the size of the product.

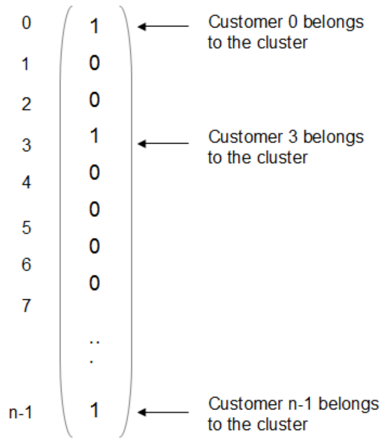


Fig. 2.2 Example of binary representation of a cluster in a distribution region of n customers

per cluster. However, if we study the clustering of customers over a large distribution zone, the following inconsistency is likely to happen: two customers situated very far from each other could be considered in a same potential cluster. This is inconsistent because the overall location–routing approach aims at minimizing distribution costs. We have thus to add another selection constraint related to distance between customers. Adding this constraint will also lead to reducing the list of potential clusters which will be input to the set–partitioning integer program. A first alternative is to consider the distance between each pair of customers belonging to the same cluster. Another alternative is to introduce intra–cluster routes: we define the intra–cluster route for a given cluster as the optimal route visiting all the customers of the cluster, starting and ending at a same customer. Estimating these routes consists in solving traveling salesman problems (TSP), which could be done using complete enumeration if the number of customers per cluster is not very high. Algorithm 1 illustrates the application of selection constraints in the procedure generating potential clusters of three customers.

Algorithm 1. Example of an algorithm generating potential clusters of three customers

Data:

V: 0–1 vector of size n representing the current cluster, n is the total number of customers in the distribution zone;

$Test$: Boolean created to check whether the current cluster should be added to the set of potential clusters;

$OptRoute(V)$: Intra–cluster distance, i.e. the length of the optimal route visiting all the customers of the cluster V, starting and ending at a same customer. $OptRoute(V)$ has to be less than a given limit *LimitRoute* defined by the user;

$MinVolCluster$: Minimum volume required per cluster;

Algorithm:

```
Initialize vector V to 0;
for i=0 to n-3 do
    V(i)=1;
    for k=i+1 to n-2 do
        V(k)=1;
        for m=k+1 to n-1 do
            V(m)=1;
            Test=1;
            if OptRoute(V) > LimitRoute then
                Test=0;
            end
            if Test=1 then
                Demand(V)=Demand(i)+Demand(k)+Demand(m);
                if Demand(V) < MinVolCluster then
                    Test=0;
                end
                if Test=1 then
                    Add the current cluster (vector V) to the set of potential clusters;
                end
            end
            V(m)=0;
        end
        V(k)=0;
    end
    V(i)=0;
end
```

2.3.2 Problem formulation

In order to optimize the clustering of customers in a given distribution zone, we formulate a set–partitioning problem based on the output of the generation procedure described in the previous paragraph.

2.3.2.1 Parameters

N	Set of customers belonging to the distribution zone ($n = 1..N$)
K	Set of potential clusters as given by the generation procedure detailed in §2.3.1 ($k = 1..K$)
$\mathbf{c}_k$	Cost of selecting potential cluster k in the solution
$\mathbf{a}_{nk}$	Parameter equal to 1 if customer n belongs to potential cluster k and 0 if not

To evaluate the cost of selecting a given cluster k in the set–partitioning solution, we consider the intra–cluster route earlier defined. Thus, c_k is set to the length of the optimal route visiting all the customers belonging to cluster k, starting and ending at a same customer.

2.3.2.2 Decision variables

$$x_k = \begin{cases} 1 & \text{if potential cluster number k is selected in the solution} \\ 0 & \text{otherwise} \end{cases}$$

2.3.2.3 Optimization problem

The set–partitioning clustering problem can now be defined as follows:

Minimize:

$$Cost = \sum_{k \in \mathbf{K}} c_k x_k \tag{2.1}$$

Subject to:

$$\sum_{k \in \mathbf{K}} a_{nk} x_k = 1 \;\; \forall n \in \mathbf{N} \tag{2.2}$$

Notice that through the objective function (2.1), we aim at minimizing the total length of intra–cluster routes. This means that we minimize the customer dispersion by encouraging small–size clusters, and if possible construct clusters with only one customer as the intra–cluster route for a 1–customer cluster is equal to 0. In other words, we aim at ensuring direct deliveries whenever possible as this is the cheapest and quickest way of transport if customer demand is high enough.

Constraints (2.2) stipulate that each customer n has to be assigned to exactly one cluster.

2.3.3 Tests

This subsection focuses on testing the computational performance of the set–partitioning based clustering approach and to analyze the results obtained. We use real–life data from the automotive industry, namely the distribution network of the car maker Renault in France. The overall network contains 448 car dealers scattered all over the country (see Fig. 2.3). We are also given the total demand of each customer per product for the whole planning horizon (one year) and we compute the point–to–point distance matrix between customers using the Geographic Information System (GIS) Microsoft MapPoint.

Fig. 2.3 Car dealers of Renault in France

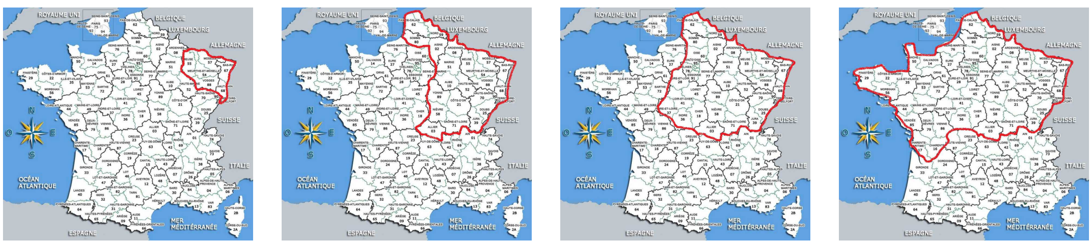
We chose to group customers according to the following constraints (reference dataset) based on real–life conditions in car distribution:

- Minimum demand volume per cluster: two full truckloads for each five working days
- Maximum number of customers per cluster: three
- Maximum distance between customers in a cluster composed of two customers: 80 kilometres
- Maximum intra–cluster route in a cluster composed of three customers: $80 * 3 = 240$ kilometres (intra–cluster route is the shortest route visiting all the customers of the cluster starting and ending at a same customer)

According to these constraints, we generate the potential clusters. Then we build the integer program of §2.3.2 with an objective function representing the total length of intra–cluster routes. We employ the C++ language to implement the model and the commercial solver ILOG Cplex version 12.5 to solve it. We carry out all the tests on a PC Intel Core(TM) i5–3210M (2.5 GHz) with 8 Gb of RAM, running under Windows 7. In the literature review, we mentioned that using a set–partitioning formulation leads to two difficulties: the large number of binary variables in most real–life cases and the difficulty of computing route costs as this often involves solving traveling salesman problems (TSP). In our case, we tried to decrease the number of potential clusters, and thus the number of binary variables, using the constraint of maximum intra–cluster routes. Moreover, we did not find any difficulty in using complete enumeration when solving TSPs as the maximum number of customers per cluster is limited to three.

Seven instances were used to test the performance of the set–partitioning based clustering approach. Table. 2.2 summarizes the details concerning these instances and Fig. 2.4 shows the corresponding geographical regions. From the computational results illustrated in Table. 2.3, we observe that:

- The computation time increases as the number of customers increases but remains acceptable even for the largest problem dealing with the whole country (i.e. 448 customers). Moreover, the maximum distance between customers influences the computation time (test 7 takes much more time than test 5). Increasing the maximum distance between customers of a same cluster means indeed that we authorize more potential clusters to be candidate in the set–partitioning problem. Thus, the number of binary variables in the integer program is greater, which makes the computation time increase.
- Increasing the maximum distance value does not seem to have an impact on the clustering results (test 7 and test 5 lead to the same number of clusters, same average cluster size and close objective values). On the contrary, we observe that decreasing customer demand leads to increasing the average cluster size and almost doubling the objective value (through comparison of tests 5 and 6). This can be explained by the fact that more demand points have to be grouped together to meet the minimum volume constraints.



(a) Region 1 (b) Region 2 (c) Region 3 (d) Region 4

Fig. 2.4 Geographical regions of the customer instances used to test the clustering approach

Fig. 2.5 illustrates the clustering results on a map for test instance 5 (whole country). The figure clearly shows that there are many 1–customer clusters and few 3–customer clusters, which

Num Test	Geographical region	Nb districts	Demand	Max distance (KM)
1	1	6 districts	RefData	80
2	2	23 districts	RefData	80
3	3	39 districts	RefData	80
4	4	56 districts	RefData	80
5	5	Whole country	RefData	80
6	5	Whole country	RefData/2	80
7	5	Whole country	RefData	150

Table. 2.2 Test instances used for clustering. RefData denotes the reference data given for customer demand

confirms the fact that the implemented set–partitioning clustering approach favors small–size clusters when they have enough demand and encourages direct deliveries whenever possible.

Num Test	Nb customers	Nb clusters	Average cluster size	Objective	CPU
1	33	24	1.37	525.23	$<$ 1s
2	115	78	1.47	1962.06	1s
3	196	135	1.45	2814.39	5s
4	286	198	1.44	4544.12	14s
5	448	314	1.43	7051.39	55s
6	448	218	2.05	13037.9	58s
7	448	314	1.43	6951.8	6.1min

Table. 2.3 Computational results of the set–partitioning based clustering approach

As a conclusion, we can point out the good performance of the exact clustering approach based on set–partitioning not only in terms of grouping results but also in terms of computation time. Thus, we will adopt this approach when testing the overall network optimization procedure in chapter 4 (static version) and chapter 5 (multi–period extension). However, tests showed that computation times could increase when the number of binary variables increases. This is likely

Fig. 2.5 Optimal clustering of Renault car dealers in France. Single points represent one–customer clusters, lines represent 2–customer clusters and triangles are 3–customer clusters.

to happen if we consider a distribution region with more customers than 448, if we increase the maximum allowed distance between customers of a same cluster or if we authorize more than three customers per cluster (as in that case solving TSP will be more difficult and also the number of potential clusters will significantly increase). Thus, we propose in the next section a heuristic clustering approach in order to obtain good quality groups within short computation times for large size instances.

2.4 Heuristic clustering

When the distribution network contains a large number of customers, it might be more convenient for the decision maker to use a heuristic pre–processing clustering in order to group customers within a short computation time. The purpose of this section is to provide a heuristic algorithm for clustering and to compare its results to the results of the exact clustering based on set–partitioning.

2.4.1 Algorithm

The proposed algorithm is a heuristic approach which consists in grouping the close customers into clusters. The clustering produced by the algorithm has to meet, as far as possible, three constraints: a maximum distance between the customers of a same cluster, a maximum number of customers per cluster and a minimum demand per cluster. Input distances between customers are calculated using a Geographic Information System (GIS) to form a point-to-point distance matrix.

Fig. 2.6 and Algorithm. 2 illustrate the first phase of the implemented algorithm. At the beginning of the procedure, the cluster list is initialized to single–element clusters (i.e. each customer corresponds to a cluster). Then, based on the distance matrix, the closest clusters are chosen in each iteration. If grouping them is possible with respect to the various clustering constraints, they are grouped together. If the resulting cluster already reaches the minimum required volume, it is removed from the clustering procedure. Otherwise, distances towards the other clusters are updated. We chose the following distance definition between two clusters p and q: $Distance(p, q) = Min\{Distance(i, j), i \in C_p, j \in C_q\}$. C_p is the set of customers of the cluster p. This proximity measure is called the "single linkage" measure; see [Barreto et al., 2007] for a discussion of other possible measures.

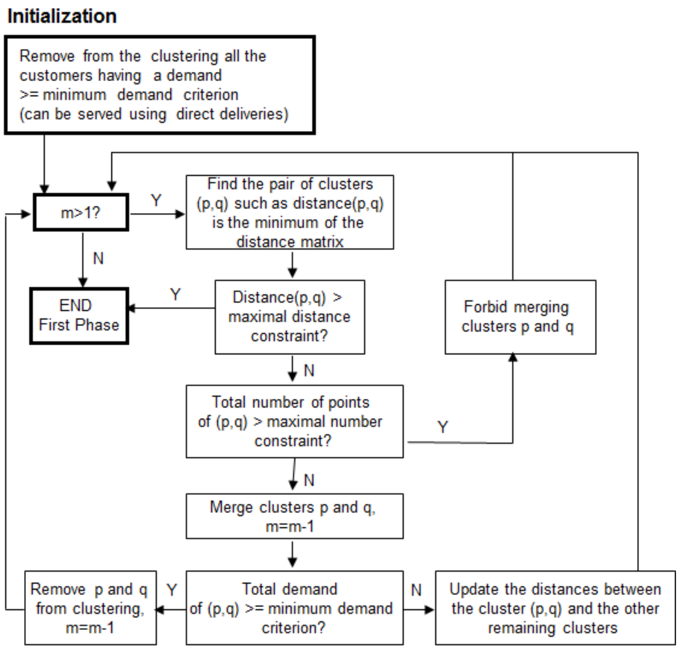


Fig. 2.6 The first phase of the heuristic clustering algorithm (m is the number of clusters in each step of the clustering algorithm)

Algorithm 2. The pseudo–code of the first phase of the heuristic clustering

Data:

n: Total number of customers;

m: Number of clusters remaining in the clustering procedure;

V_{min}: Minimum volume per cluster;

demand(p): Total demand of cluster p;

maxDistance: Maximum distance between two customers of a same cluster;

maxNumber: Maximum number of customers per cluster;

nbCustomers(p): Number of customers in cluster p;

Algorithm:

```
Initialize the n customers to n individual clusters;
for p=1 to n do
    if demand(p) ≥ V_min then
        Remove cluster p from the clustering procedure;
        m=m-1;
    end
end
while m > 1 do
    Find the pair of clusters (p,q) such as distance(p,q) is the minimum of the distance matrix;
    if distance(p,q) ≤ maxDistance then
        if nbCustomers(p)+nbCustomers(q) ≤ maxNumber then
            Merge clusters p and q;
            m=m-1;
            if demand(p)+demand(q) ≥ V_min then
                Remove p and q from the clustering procedure;
                m=m-1;
            else
                In the distance matrix, update the distances between the cluster (p,q) and the
                other remaining clusters;
            end
        else
            Forbid merging p and q;
        end
    else
        Break;
    end
end
```

This first phase does not assure that all the resulting clusters meet the minimum volume constraint. Thus, a second repairing phase has to be applied in order to consider the clusters discarded in the first phase due to the various constraints. In the second phase, only constraints of minimum volume and maximum number of customers are imposed. We check that each cluster resulting from the first phase has a total demand greater than the minimum demand required. If it is not the case for some cluster q then we attach it to the nearest cluster having fewer customers than the maximum allowed number of customers.

2.4.2 Tests

Our objective is to compare the heuristic clustering procedure with the exact one in terms of solution quality and computation time. To this end, we employ test instances 1–6 defined in Table. 2.2. We use the same parameter values for the maximum number of customers per cluster (i.e. three) and for the minimum volume per cluster (i.e. two full truckloads/five working days) but we introduce a different value for the maximum distance between customers of a same cluster, 50 kilometres instead of 80 kilometres (this gives better results with the heuristic). Results for the six test instances are summarized in Table. 2.4 where the fifth column represents the evaluation a posteriori of the objective value (sum of intra–cluster routes) on the solution given by the heuristic method. Fig. 2.7 illustrates the heuristic clustering output on a map for test instance 5.

Num Test	Nb customers	Nb clusters	Average cluster size	Objective	Quality (%)	CPU
1	33	23	1.43	529.50	0.81	$<$ 1s
2	78	75	1.53	2196.53	11.95	$<$ 1s
3	196	133	1.47	3129.71	11.20	$<$ 1s
4	286	192	1.49	5007.50	10.20	$<$ 1s
5	448	306	1.46	7670.28	8.78	$<$ 1s
6	448	197	2.27	13896.40	6.58	$<$ 1s

Table. 2.4 Computational results of the heuristic clustering approach. Quality is measured as the relative difference between the objective value of the heuristic solution and the objective value of the exact one.

The comparison of Fig. 2.5 with Fig. 2.7 confirms that the results given by the heuristic approach are close to those of the set–partitioning based method when using the reference dataset. Besides, the comparison of Table. 2.4 with Table. 2.3 shows that using the heuristic approach leads to decreasing computation times as compared to the exact approach. However, it can be seen that the average cluster size (respectively the number of clusters) resulting from the heuristic approach is greater or equal (respectively less or equal) than the average size resulting from the exact method, which could lead to higher distribution costs. The value of the heuristic solution as

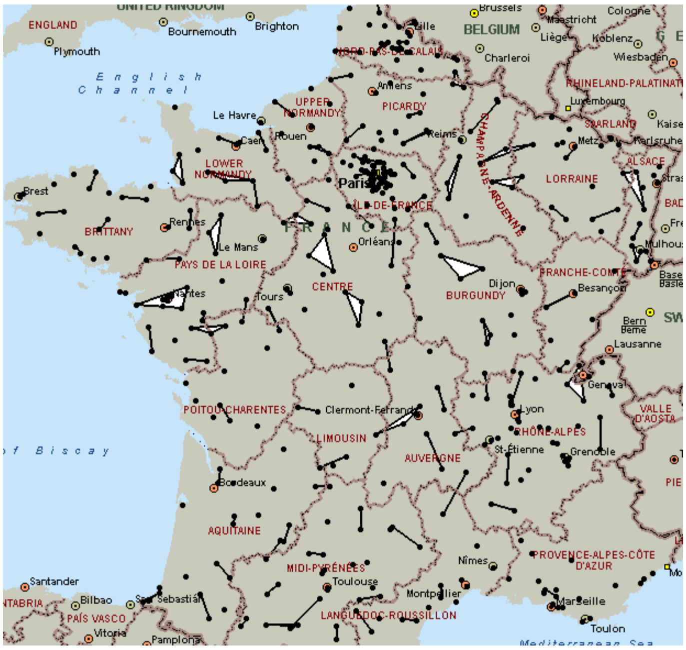


Fig. 2.7 Heuristic clustering of Renault car dealers in France. Single points represent one–customer clusters, lines represent 2–customer clusters and triangles are 3–customer clusters.

compared to the exact one on the basis of the same objective function shows indeed an increase lying between 0.81% and 11.95%. To confirm these values, we carry out other tests on 20 test instances (we generate these instances based on the reference dataset and in each of them we multiply the demand of each customer by a pseudo–random coefficient in [0.5,1.5]). Table. 2.5 shows the quality of the heuristic solutions obtained on these instances.

The heuristic solution quality thus varies between 6.42% and 15.70% with an average of 9.6%. This is an acceptable result but can be improved in a further research in order to obtain better grouping results when dealing with large size problems where the exact approach cannot be applied due to prohibitive computation times.

2.5 Conclusion

We presented in this chapter a sequential location–routing approach based on a pre–processing clustering, which consists of three steps:

- Step 1: Forming clusters of close customers subject to minimum volume requirements, maximum distances restrictions and maximum number of customer constraints.
- Step 2: Computing the length of optimal TSP routes from distribution centres to the resulting clusters.
- Step 3: Optimizing the distribution network based on the outcome of the first two steps.

Test instance	Quality of the heuristic clustering (%)
C.1	10.31
C.2	8.60
C.3	7.89
C.4	7.51
C.5	10.15
C.6	9.84
C.7	10.72
C.8	11.85
C.9	6.54
C.10	9.72
C.11	15.70
C.12	6.42
C.13	10.99
C.14	12.36
C.15	9.17
C.16	8.28
C.17	10.45
C.18	7.11
C.19	9.18
C.20	9.29

Table. 2.5 Evaluation of the quality of heuristic clustering as compared to optimal clustering on instances varying customer demand.

To carry out the first clustering step of the location–routing approach, we proposed an optimal procedure using a set–partitioning formulation that provided good quality solutions within short computation times when applied to our case–study in the field of car distribution. However, when varying some of the problem parameters, the problem size may significantly increase and thus lead to computational difficulties. To deal with these cases, we proposed a simple heuristic method and compared its performance to the performance of the exact method on the basis of a same measure (sum of intra–cluster routes). The average gap that we obtained was less than 10%, which is good enough at this level of the study but requires to be improved if we have to employ the heuristic to deal with large size instances.

The sequential clustering–based location–routing procedure that we proposed shows two main advantages:

- As compared to a standard aggregate approach: A standard aggregate approach consists in solving a distribution network design problem using towns or districts as aggregate customers then computing delivery routes in a second step. Our approach provides a more detailed representation of routing through a clustering–based procedure (routing first, location second). Thus it is worth comparing it to the standard method in terms of costs and operational constraint satisfaction. The relative experiments discussed in appendix B prove that using the standard aggregate approach (called AG1 in appendix B) results in wrong location and assignment decisions, which leads to impossible route construction at the operational level. In fact, in all the tests, more than half of the opened distribution centres are concerned by infeasible routing problem, i.e. customers assigned to them cannot be grouped into clusters meeting all the operational constraints (namely minimum volume, maximum number of customers and maximum route distance from the distribution centre). In the context of our application, the main contribution of the clustering–based approach is thus to solve a combined location–routing problem meeting all the operational constraints. In terms of costs, we cannot compare our detailed method with the aggregate one due to the violation of operational constraints in the latter case.

- As compared to a fully–integrated (explicit) location–routing approach: Using a sequential clustering–based method is a way to take into account multiple stop deliveries while keeping a manageable size for the optimization problem. Thus, we can handle large size industrial instances with many customers.

Furthermore, the optimal clustering procedure presented in this chapter can be employed either in a static environment or in a dynamic one. In a static environment (chapter 3), we use the clustering method on the whole distribution region in order to obtain clusters of customers that will represent the final customers of the optimization problem. In a dynamic environment (chapter 5), we iteratively apply the clustering on limited distribution regions (districts) in each time–period.

Chapter 3

Modeling a multi–product distribution network design problem with minimum volume and distance constraints (DNDMVD)

Major industrial companies put considerable efforts on improving the planning of their supply chains in order to reduce costs, decrease lead times and meet the due dates promised to customers. The problems encountered by firms are quite diverse and depend on the specificities of the supply chain in their sector as well as on the related planning level. Supply chain planning tasks are classified according to three planning levels [Anthony, 1965]: long–term (strategic), mid–term (tactical) and short–term (operational).

One key question at the strategic planning level is designing the supply chain network and locating the major facilities of the company, namely plants and distribution centres (DCs). This leads to classical and well–studied problems in the literature, which are facility location and network design. In the context of our study, we focus on a discrete location problem, i.e. facility locations have to be chosen among a list of eligible sites (see [ReVelle et al., 2008] for a survey on discrete location science). In this case, the objective is to locate new facilities and to determine the related product flows on a given network where the potential links between nodes are already constructed and the locations of demand points are known. A fixed installation cost is usually incurred for the location of each facility. In a network design problem, the links of the network have to be constructed for a fixed construction cost whereas the nodes (facilities) are already located. Some works like [Melkote and Daskin, 2001] proposed integrated models for both facility location and network design.

The attraction of researchers to these expanding areas could be explained by the numerous applications that can be found in various industries and service fields like health–care systems [Rahman and Smith, 2000], blood banks [Shen et al., 2003], solid waste management [Barros et al., 1998], telecommunication network [Carello et al., 2004], automotive industry [Nozick and Turnquist, 1998], etc. The expression of the objective function of the problem thus depends on the application context but is particularly related to the manager preferences and to the key factors impacting the strategic plan of the company. In this context, some works focus on customer coverage (see e.g. [Eiselt and Marianov, 2009]), others include the cost of carbon emissions in the objective function (see e.g.
[Elhedhli and Merrick, 2012]) or even combine two or several objectives (see e.g.
[Xifeng et al., 2013]) in order to let the decision maker balance his preferences and priorities. This kind of multi–objective tools is out of the scope of the present work which concentrates on a single objective minimizing the total distribution costs. However, some secondary objectives such as lead times are implicitly taken into account through the problem constraints and parameters.

The present chapter focuses on modeling a multi–product distribution network design problem based on a three–level structure consisting of plants in the first level, distribution centres (DCs) in the second one and final customers in the third one. We assume that the number and location of the plants as well as the number and location of the customers are fixed. Given a deterministic and static demand of customers for each product and a list of potential DCs, our main concern is to locate DCs and to assign customers to them in such a way as to minimize the total distribution costs. Our work shares common features with the classical problems of facility location and supply chain network design. It is however significantly different due to various extra constraints that makes it closer to reality. First of all, as already explained in chapter 2, we introduce a location–routing approach to model transport routes from DCs to customers. Other realistic features are also taken into account in our study through various constraints: maximum covering distances, maximum capacities on DCs, minimum volumes on transport links, minimum volumes on DCs and single sourcing assignments. As compared to classical works in the literature of facility location and network design, our model does not involve fixed costs neither for siting facilities nor for constructing transport links. However, minimum volume constraints can be viewed as an indirect way of representing fixed costs by imposing a minimum quantity (in other words a pre–determined cost) each time a facility or a link is used. In the present chapter, we provide more detail on these constraints and the motivation of their introduction in the studied model. Apart from minimum volume constraints, all the other restrictions are classical considerations in the literature of facility location and supply chain network design. They were however considered separately and we could not find any work *combining them in a same model*, as this is the case here.

Using this highly constrained model in our study makes it as close as possible to real–life requirements but this leads to several trade–offs to be achieved in the problem as well as to computational difficulties in the solution stage. In this chapter, we propose a qualitative analysis regarding the trade–offs, which will be supported in chapter 4 by numerical results based on a real–life case–study in the automotive industry. Chapter 4 involves also the implementation and tests of heuristic solution procedures in order to obtain good quality solutions within short computation times.

The present chapter is organized as follows: the first section presents a literature review on facility location and supply chain network design problems with minimum volume and distance constraints. In section 2, we look at the main modeling considerations of the problem and analyze the trade–off between different constraints. Section 3 is devoted to the detailed mathematical formulation of the problem as well as the analysis of its complexity. Finally, some conclusions are provided in section 4.

3.1 Literature review on facility location and supply chain network design problems with minimum volume and distance constraints

3.1.1 Literature considering minimum volume constraints

Economies of scale can be defined as the reduction of cost per unit due to an increasing quantity, it "can be present in nearly every function of a business, including manufacturing, purchasing, research and development, marketing, service network, sales force utilization, and distribution" [Porter, 1980]. For instance, dealing with much volume in transport leads to decreasing the unit cost by using full truckloads. Less than-truckload shipments proved indeed to be very expensive and ineffective. Thus, it is crucial to consider economies of scale in the supply chain network design problem so as to ensure that full truckload shipments will be possible at the operational planning level of the supply chain. In the present work, we focus on the use of minimum volume constraints to model flow consolidation. The literature devoted to facility location and supply chain network design using minimum volume constraints is relatively scarce. This is probably due to two reasons: the complexity of the resulting problem and the difficulty of evaluating the lower bound values in practice. Table. 3.1 summarizes the main characteristics of the 17 papers that we found in this field.

From a modeling point of view, we can first notice that only four works consider minimum volume constraints for transport quantity on each link of the network. Most of the papers deal with this kind of constraint for the throughput of facilities and none of them propose a model using minimum volume constraints for both facility throughput and transport quantities. Multi-

Paper		Min volume constraints		Complicating features		Solution procedure and numerical results					
	Facility location decisions	Throughput of facilities	Transport quantity	Multi-product	Single sourcing	Number of integer variables in the largest instances considered*	Number of continuous variables in the largest instances considered*	State-of-the-art MIP solver	Linear relaxation based heuristic	Other heuristic	Specialized exact method
[Melo et al., 2009b]	X	X		X		560	107057	X	X		
[Thanh et al., 2010]	X	X		X		2185	30870	X	X		
[Barros et al., 1998]	X	X		X		NM	NM		X		
[Seedig, 2010]			X			50	50				X
[Zhu et al., 2011]			X			119643	119643	X			
[Krumke and Thielen, 2011]			X			NM	NM			X	X
[Meyerson, 2001]	X	X			X	NM	NM			X	
[Svitkina, 2010]	X	X			X	NM	NM			X	
[Geoffrion and Graves, 1974]	X	X		X	X	5490	1295910				X
[Lim et al., 2006]			X			10890	0		X	X	X
[Melo et al., 2005]	X	X		X		270	732810	X			
[Sabri and Beamon, 2000]	X	X		X	X	27	214	X			
[Karger and Minkoff, 2000]	X	X			X	NM	NM			X	
[Guha et al., 2000]	X	X			X	NM	NM			X	
[Alumur et al., 2012]	X	X		X		1200	58000	X			
[Ndiaye and Alfares, 2008]	X	X			X	370	1810	X			
[Correia et al., 2013]	X	X		X		1728	83705	X			X
Present work	**X**	**X**	**X**	**X**	**X**	**65364**	**816**	**X**	**X**		

Table. 3.1 Literature review on facility location and network problems featuring minimum volume constraints (NM=Not Mentioned). **Largest instances are considered in view of the number of integer variables*

product or single sourcing are restrictions that could make the problem more difficult to solve as the former increases the variable number and the latter adds 0–1 constraints. Only two of the listed works ([Sabri and Beamon, 2000] and [Geoffrion and Graves, 1974]) simultaneously consider the multi-product and single sourcing features.

From a solution method perspective, the computational results provided in [Sabri and Beamon, 2000] concern only small instances (27 binary variables) and use a commercial solver. Other papers implement heuristic solution procedures, mainly constant–factor approximation[1] [Krumke and Thielen, 2011, Meyerson, 2001, Svitkina, 2010]

[1]Constant–factor approximation algorithms are polynomial–time approximation algorithms with an approximation ratio bounded by a constant. For instance, a given algorithm is called c–approximation

[Karger and Minkoff, 2000, Guha et al., 2000] and linear relaxation[2] based heuristics [Melo et al., 2009b, Thanh et al., 2010, Barros et al., 1998, Lim et al., 2006]. In the latter papers, the heuristic strategy is based on rounding fractional solutions to 0 or 1, depending on their values. In [Lim et al., 2006], the idea consists in rounding to 1 the carrier selection variable having the greatest fractional value. In the other works, the rounding of a location variable depends on the comparison of its fractional value to an upper bound close to 1 and a lower bound close to 0. None of the four papers focuses on rounding assignment variables either because they are not problematic in the solution procedure or because they are continuous variables.

3.1.2 Literature considering maximum distance constraints

In supply chain, one of the main objectives of facility location is covering as much of customer demand as possible. A customer is said "covered" if there exists an opened facility situated within a pre–specified distance of it. In order to formulate this kind of distance constraints, we can either use covering objectives or impose covering distance constraints with any type of objective, usually cost or distance minimization. Covering objectives are based on "equity", i.e. fairness between customers in accessing the new facilities. This is mainly expressed using maximum distance models such as set–covering location problems (SCLP), maximum–covering location problems (MCLP) and p–center problems (refer to [Farahani et al., 2012] for a recent literature review on covering problems in facility location). In set–covering location problems (see e.g. [Eiselt and Marianov, 2009]), the objective is to minimize the number of sites to locate in order to satisfy the demand of customers within a maximal distance called the "covering distance", typical applications concern the location of emergency facilities like hospitals or fire stations. In maximum–covering location problems first introduced by [Church and ReVelle, 1974], we aim at locating a predetermined number of facilities in such a way as to maximize the demand satisfied within the "covering distance". The purpose of p–center problems (see e.g. [Krumke, 1995]) is to minimize the maximum distance between any customer and its closest facility, given a predetermined number of facilities to be located.

In spite of the abundance of literature dealing with covering objectives, there are only few papers considering covering distance constraints. We can however mention some works in this field like [Moon and Chaudhry, 1984] where authors motivate the use of distance constraints between facilities and demand points with the necessity of ensuring a desired level of customer service. In [Saez-Aguado and Trandafir, 2012], authors develop a p–median problem minimizing the sum of traveled distances while the total demand covered at a distance greater than a given coverage dis-

algorithm (for a given constant c) if it can be proven that it can find a solution which is at most c times worse than the optimal solution.

[2] Linear relaxation consists in relaxing the constraints imposing the integrality of variables in order to obtain a linear program. The optimal solution of the resulting linear program represents a lower bound to the optimal solution of the original integer (or mixed–integer) program.

tance should not exceed a predetermined value. Authors of [Albareda-Sambolaa et al., 2011] propose a problem called the capacity and distance constrained plant location problem where the total distance traveled by each vehicle to serve its assigned customers is limited. In [Moon and Papayanopoulos, 1995], the problem of locating the smallest number of facilities to serve customer–nodes on a tree network is addressed with the condition that no node can be served by a facility beyond a given distance. None of these works discusses the interaction between distance constraints and other constraints imposed by real–life applications, they rather focus on implementing efficient solution methods to improve computation times.

3.1.3 Main contributions of our study

In the present chapter, we describe all the modeling considerations related to the distribution network design problem we propose to study, then we will discuss in chapter 4 exact and heuristic solution methods for this problem. The two chapters feature four main contributions as compared to the works mentioned in Table. 3.1 and those discussed in the previous paragraph. First, to fit real–life requirements, we consider maximum covering distance and single sourcing restrictions simultaneously with minimum volume constraints in connection with a multi–product distribution network design problem. Second, in our study, we assign minimum volume constraints to facility throughputs as well as to transport flows, which makes the model more realistic but increases the difficulty of solving it. Third, we propose in chapter 4 MIP–based heuristic procedures using various types of linear relaxations to determine location and assignment variables (both sets of variables being required to be binary). Finally, we carry out numerical experiments using a case–study from the automotive industry with large size instances leading to a great number of binary variables and minimum volume constraints (see chapter 4).

3.2 Problem modeling

3.2.1 Problem description

The problem under study involves an outbound supply chain whose purpose is to deliver finished products from plants to final customers. The planning horizon that we consider is not very long (one year in our case–study). The network features a three–level structure: plants at the first level, distribution centres at the second level and customers at the last one (see Fig. 3.1). We assume that the number and location of the plants as well as the number and location of the customers are fixed. Given a deterministic and static demand of customers for each product and a list of potential DCs, our main concern is to locate DCs and to assign customers to them in such a way as to minimize the total distribution costs. The model we propose takes into account several operational features that have to be integrated together in order to remain as close to reality as possible. First, multi–stop routes from DCs to customers are considered through a location–routing procedure. Customers are grouped according to a pre–processing clustering procedure based on a set–partitioning formulation (defined in §2.3). The overall procedure was detailed in chapter 2 and illustrated in Fig. 2.1. It leads to the construction of a set of clusters that is input to the network optimization model.

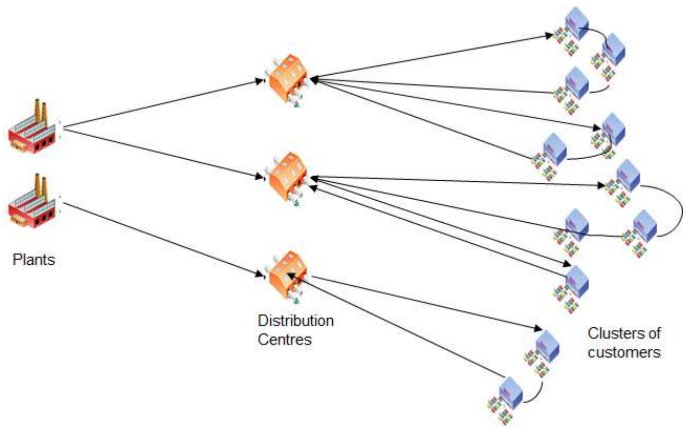


Fig. 3.1 The distribution network under study

The whole transport process is mainly split into two sub–processes: primary transport from plants to distribution centres and secondary transport from distribution centres to clusters. All finished products have to wait on factory compounds to be loaded in trucks (or trains). Then they are routed to distribution centres where they are held for a short transit time (typically a few days) before being sent by truck to (clusters of) customers. We assume that DCs do not manage inventories but are only used as intermediate facilities between plants and (clusters of) customers. As in most of the literature, the use of these DCs in our model is constrained by a limited physical capacity. No fixed costs are incurred for opening DCs but minimum throughputs are required (see §3.2.5 for more explanations on these constraints). A minimum volume is also required to open a primary or a secondary transport link, i.e. the flow on each of these links is either zero or above a given limit. This is indeed a way to maximize the use of transport capacities (see §3.2.5 for more details). At the same time, a DC cannot be situated farther than a given distance from a cluster that it serves, this is called the maximum covering distance constraint. Finally, for a given cluster, all the products manufactured in a same plant should go through a single DC. This implies three consequences in our model: products are aggregated per plant, average truck capacities are used and assignment decisions are binary and tri–indexed (plant, DC, cluster) with single sourcing constraints. In the following paragraphs, we provide more detail on the problem we are considering and on the features motivating our modeling approach.

3.2.2 Product representation

In the problem under study, we deal with a multi-product situation: a given plant can manufacture different product types and a given product type can be manufactured in several sites. It is therefore possible that a given customer be supplied with the same product type from different plants. Nevertheless, we assume that the demand of each customer for the different product types have already been assigned to plants. This is indeed a strategic (and even political) choice usually made by top–level management as it involves international exchanges and decisions on how to share the product portfolio between plants (and thus between countries). At our decision level, we can thus consider that demands are expressed as quantities of a given product type to be distributed from a given source (plant) to a given destination (customer) and that each plant capacity is sufficient to produce the total demand assigned to it. Consequently, in the studied problem, we have to focus on building a supply chain network capable of satisfying already known transport demands.

Moreover, we assume that all transport demands having the same source (plant) and same destination (cluster of customers) are routed through the same DC. This single sourcing constraint is considered by many companies in their distribution network in order to facilitate day–to–day operations. In this way, each plant has only one interlocutor when sending products to a given customer. Similarly, a customer knows that the products originating from a given plant are always distributed through the same DC. This assumption implies three consequences in our model:

- Products are aggregated per plant: When all the products originating from a given plant to a given customer are routed through a unique DC, it is meaningful to use an aggregate representation of the product types based on their sourcing plants. In fact, distribution decisions in this case depend only on the sourcing and the destination of the flow but not on the considered product type. Moreover, deliveries from plants to DCs are made through loads of mixed product types in order to optimize the use of transport capacities, which supports the decision of aggregating the products in each plant.
- Average truck capacities are used: Aggregation of products per plant and the use of mixed product loads on primary transport links lead to considering average capacities for trucks from plants to DCs. We propose to use an average truck capacity specific to each plant and to compute it as the weighted average load factor[3] over the whole demand of the various product types manufactured in the given plant with the demand of each product type as weight. As far as secondary transport links are considered, we assume for the same reason that trucks of mixed product types are shipped from DCs to clusters of customers. However, on the contrary of primary transport flows, secondary transport ones could concern products

[3]The load factor of a product type is the number of units of this type that can be loaded on a truck. It can differ according to the physical size of the product.

from different plants. Thus, we propose to use a same truck capacity for all secondary transport links, computed as the weighted average load factor over the whole demand of all product types.

- Assignment decisions are binary and tri–indexed (DC, cluster, plant) with single sourcing constraints: This means that to each plant/cluster pair, we assign a unique DC delivering the related flows but a given cluster could be assigned to different DCs, each for one or several products. Our study keeps thus its multi–product character even if we aggregate products per plant.

3.2.3 Transport costs

In order to evaluate transport costs in many of the network problems, theoretical distances like the euclidean one are usually used. This approach could lead to a good estimate of the objective function but is not very accurate. Thus, in our work, we propose to use a geographic information system to calculate real route distances between points. We then apply a classical formula to compute the cost C of a truck for a given distance x, $C = ax + bh + c$:

- a: variable cost per kilometre, it involves the cost of diesel oil in addition to maintenance and repairs per kilometre
- b: variable cost per hour, it consists of the hourly wage of the driver and the hourly cost of using the truck
- h: duration of the trip
- c: fixed cost typically related to the time needed for loading and unloading merchandises

If the average speed s of the truck is known, then it is easy to obtain a kilometric formula: $C = (a + \frac{b}{s})x + c$. Once the cost of a truck on a given transport link is computed using this formula, the average transport cost per unit is obtained by dividing the cost of a truck by the average truck capacity (minimum volume constraints on transport links ensure full truckload transport).

3.2.4 Transport routes

According to the distribution network illustrated in Fig. 3.1, our problem involves two kinds of transport routes: primary routes from plants to distribution centres and secondary routes from distribution centres to clusters of customers. On each primary transport link, we have to consider only one means of transport which could be for instance truck or train. In fact, taking into account successively more than one means of transport (i.e. using combined transport) would lead to a more complex network with several transshipment points, which is out of the scope of our study. For secondary transport, trucks are usually used because routes concern short deliveries in urban areas. In this context, we consider a maximum covering distance constraint, i.e. the length of the route between a distribution centre and the clusters that it serves must not exceed a given distance. In fact, these deliveries are usually made by drivers that have to come back to the distribution centre at the end of the working day. The traveled distance per delivery trip thus should not be greater than a given limit, allowing the driver to comply with the legal daily driving time (around 8 hours in Europe).

Primary transport routes are direct from plants to distribution centres, thus, it is sufficient to determine the corresponding distances in order to compute transport costs. This is not the case for secondary transport, which are concerned by multi–stop routes. Thus, we propose to evaluate secondary transport routes using the clustering–based approach detailed in chapter 2. Clusters have to meet two major constraints: a minimum volume and a maximum number of customers. The minimum volume is computed in such a way as to ensure full truckload transport within a given maximum waiting time expressed in days. As explained in §2.2.1, the average daily demand is obtained through a division of the total demand (for the whole planning horizon and for all products) by the number of working days in the whole planning horizon.

3.2.5 Minimum volume constraints

We assume in our study (like in many industries and in particular the automotive one) that the management of the distribution centres and the related activities are outsourced to third-party logistics. This is why there are no fixed opening costs to be incurred by the company before using a DC. There is only a transit cost to be paid to the supplier each time a product goes through a DC. This transit cost, resulting from a commercial negotiation between the car maker and the logistics supplier, only applies if the total throughput of the DC is between a minimum volume and a maximum capacity. Hence, we introduce in our model minimum volume and maximum capacity constraints conditioning the use of each DC.

A minimum volume is also required when opening a primary or a secondary transport link, which means that the flow on each of these links is either zero or above a given limit. As explained in the introduction of the present work, if a truck is not fully loaded, the unit cost per transported product could significantly increase. Thus, products to be distributed to the same destination have to wait on the plant or the distribution centre until a full truckload is consolidated. If the total flow rate on a given transport link is low, the waiting time could be considerable and results in excessive lead times. In this case, the best solution is to consolidate many flows on a same transport link. This is expressed in our DNDMVD model using minimum volume constraints that condition the use of primary and secondary transport links. The lower bound assigned to each link has to be greater than the volume ensuring on average a full truckload within the maximum waiting time allowed at the sourcing point. Similarly to the approach used in the clustering step, we compute the average daily flow on each transport link as the total flow divided by the number of working days. The result is multiplied by the maximum allowed waiting time then compared to the minimum required volume (a full truckload for instance).

It should also be remembered that we use a location–routing approach based on a pre–processing clustering that imposes a minimum volume per cluster. When using this sequential method, we have to ensure consistency between the clustering step and the optimization step. We thus have to make sure that the minimum volume required for each cluster is greater than the minimum volume required for secondary transport.

3.2.6 Main trade–off analysis

Before detailing the mathematical formulation of the optimization problem under study, we first provide a qualitative analysis of the main trade–offs to be achieved in the problem. These trade–offs not only influence the specialization and the number of the opened DCs but also the existence of a feasible solution.

Specialization of the opened DCs

In the multi-product context of our problem, we have to determine which products should be distributed to which clusters via the opened DCs. The following "simplistic" decisions would be possible:

- Specialization of each opened DC on a subset of products: A simple way of ensuring that the minimum volume constraints on primary transport links are satisfied would be to decide to distribute all the volume of each plant via a single DC. In the resulting distribution network, each DC would be in charge of distributing a subset of products to the clusters of the whole distribution region. However, this would mean that each cluster receives products from each opened DC. The total demand of each cluster would thus be split on many (DC-cluster) secondary transport links so that the corresponding solution is likely to violate the minimum volume constraints on secondary transport links.

- Specialization of each opened DC on a subset of clusters: To ensure that minimum volume constraints on secondary transport links are satisfied, we could decide to assign each cluster to a single DC. A given cluster would thus receive all the products from a single DC and a given DC would be in charge of distributing all the products of a subset of clusters. Nevertheless, this would mean that each plant has to serve each opened DC and it might be difficult or even impossible to satisfy the minimum volume constraints on primary transport links in the corresponding solution.

In the proposed optimization model, we thus seek to find the best trade-off regarding the degree of specialization of the opened DCs.

Number of opened DCs

In classical facility location problems, one of the main purposes of optimization is to decide about the number of DCs to open so as to achieve the best possible trade-off between fixed-charge opening costs and transport costs. In our problem, in addition to minimizing transport costs, we have also to cope with maximum covering distance and maximum capacity constraints. This could lead to opening many DCs in order to be close to customers and not to exceed the capacity of each DC. As previously explained, no fixed-charge opening costs are included in the proposed model but opening a DC at each possible location may not be feasible. Namely, this decision would result in splitting the global volume to be distributed from the plants in many DCs so that both the constraints of minimum throughput per DC and the constraints of minimum volume on each (plant-DC) primary transport link are likely to be violated in the corresponding solution.

Fig. 3.2 shows the impact of the problem constraints on the number of DCs to be opened in the optimal solution of the problem. One of the purposes of the proposed optimization model is thus to build a network so as to meet the maximum covering distance constraint and to minimize the total transport costs while ensuring that minimum volume and maximum capacity constraints are satisfied.

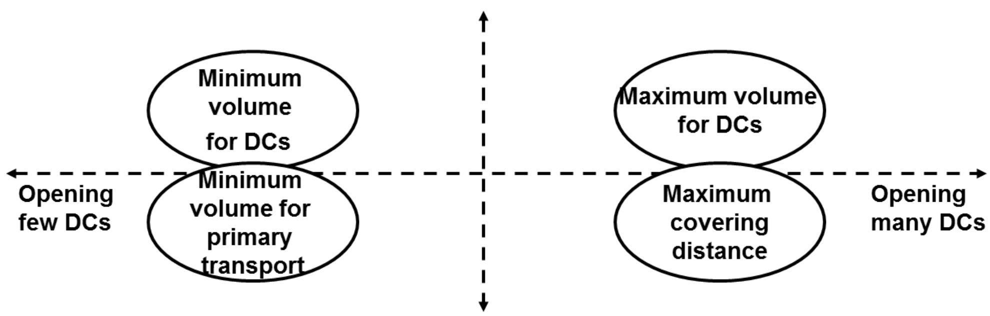


Fig. 3.2 Impact of the DNDMVD problem constraints on the number of DCs to be opened

Feasibility issue and using penalties

In the previous discussion, we explained why there is a conflict between minimum volume constraints on primary transport links and maximum covering distance constraints. Experiments with typical data and parameters of our case–study show that this conflict often leads to infeasible situations due to the presence of some plants that produce low quantities. The volume of these plants cannot be split between a large number of DCs (that should be opened due to the maximum covering distance constraint) as this violates the strict minimum volume constraints on primary transport links. Therefore, in the model, we propose to relax these constraints. We consider the possibility of falling under the minimum required volume on a given primary transport link but in that case this incurs an additional cost (penalty) in the objective function. This penalty can be viewed as the cost of increasing waiting times to consolidate enough volume or as the additional cost of shipping trucks not fully loaded within the targeted waiting times (see §4.1.4.1 for a further discussion about the choice of the penalty coefficients). It should be noted that minimum volume constraints on secondary transport links do not lead to the same feasibility issues. This is why we keep modeling them using strict constraints.

3.3 Problem formulation

3.3.1 Model parameters

I	Set of plant indices ($i = 1..I$)
J	Set of DC indices ($j = 1..J$)
Q	Set of cluster indices ($q = 1..Q$)
D_{qi}	Total demand of cluster q for the products manufactured in plant i during the whole planning period.
$minVol_j$	Minimum volume of cars that has to go through DC j if it is selected in the solution.
$maxVol_j$	Maximum volume of cars that can go through DC j if it is selected in the solution.
PTC_{ij}	Cost of a truck going from plant i to DC j (Primary transport cost).
STC_{jq}	Cost of a truck starting its route at DC j and visiting all the customers of cluster q before going back to j (Secondary transport cost).
TC_j	Unit transit cost for a car going through DC j.
W_i	Average truck capacity for the cars manufactured in plant i.
W	Average truck capacity for the whole volume of cars.
M	Big value.
NWD	Number of working days in the planning period.
$T_{max}(i)$	Maximum waiting time allowed at plant i before shipping is made to distribution centres.
T	Maximum waiting time allowed at a distribution centre before shipping is made to customers.
CD	Maximum covering distance parameter (i.e. the maximum length of a route starting at a DC, visiting the customers of a given cluster then coming back to the DC).
PI_{ij}	Low volume penalty amount for primary transport from plant i to DC j.
$R(j,q)$	Length of the optimal route starting at DC j and visiting the customers of cluster q before coming back to j.
$totProd_i$	Total volume of cars produced by plant i.

$V_{min}(i)$	Minimum volume of cars that has to go through any opened primary transport link starting at plant i.
V_{min}	Minimum volume of cars that has to go through any opened secondary transport link.

$V_{min}(i)$ has to be equal at least to the minimum volume ensuring on average a full truckload within $T_{max}(i)$ from plant i to a given DC hence $V_{min}(i) \geq \frac{W_i}{T_{max}(i)} NWD$.

V_{min} has to be equal at least to the minimum volume ensuring on average a full truckload within T from a DC to any cluster of customers hence $V_{min} \geq \frac{W}{T} NWD$.

It is worth pointing out that, in order to evaluate the average truck capacity on each transport link, we use the weighted average of the load factors with the demand rate of each car type as weights. The load factor of a car type is the number of cars of this type that can be loaded on a truck. It can differ according to the car size. On primary transport links starting at plant i, the flows concern only the products manufactured in plant i. This is why we propose to use a specific average capacity W_i. On secondary transport links, we do not know in advance what kind of product will be transported. As the trucks used for secondary transport are loaded with a mix of various car types, we propose to use a common weighted average capacity W over all the car types.

3.3.2 Decision variables

- Location variables:

$$y_j = \begin{cases} 1 & \text{if DC } j \text{ is selected} \\ 0 & \text{otherwise} \end{cases}$$

- Assignment variables:

$$x_{jqi} = \begin{cases} 1 & \text{if cluster } q \text{ is assigned to DC } j \text{ for the products of plant } i \\ 0 & \text{otherwise} \end{cases}$$

- Variables stating if the secondary transport links are selected:

$$a_{jq} = \begin{cases} 1 & \text{if cluster } q \text{ is assigned to DC } j \text{ for at least one product} \\ 0 & \text{otherwise} \end{cases}$$

x_{jqi} and a_{jq} are defined only if $R(j,q) <= CD$ (Maximum covering distance constraint).

- Variables stating if the primary transport links are selected:

$$z_{ij} = \begin{cases} 1 & \text{if the route from plant } i \text{ to DC } j \text{ is selected} \\ 0 & \text{otherwise} \end{cases}$$

- v'_{ij},v''_{ij}: Continuous variables used to write the minimum volume constraints for primary transport links.

  v'_{ij} is a positive variable that has to be greater than $V_{min}(i)$.

  v''_{ij} is a positive variable used to compute the amount of violation of the minimum volume constraint on a given primary transport link [ij]. It has to be less than $V_{min}(i)$ and will be minimized, null if possible, as it is penalized in the objective function.

3.3.3 MIP formulation of the optimization problem

The SCNDP with minimum volume constraints can now be defined as follows:

Minimize:

$$\begin{aligned} Total\ cost = & \sum_{i\in\mathbf{I}}\sum_{j\in\mathbf{J}} \frac{PTC_{ij}}{W_i} \sum_{\substack{q\in\mathbf{Q} \\ R(j,q)\le CD}} x_{jqi}D_{qi} && (Primary\ transport\ cost) \\ & + \sum_{j\in\mathbf{J}} \sum_{\substack{q\in\mathbf{Q} \\ R(j,q)\le CD}} \frac{STC_{jq}}{W} \sum_{i\in\mathbf{I}} x_{jqi}D_{qi} && (Secondary\ transport\ cost) \\ & + \sum_{j\in\mathbf{J}} TC_j \sum_{i\in\mathbf{I}} \sum_{\substack{q\in\mathbf{Q} \\ R(j,q)\le CD}} x_{jqi}D_{qi} && (Transit\ cost) \\ & + \sum_{i\in\mathbf{I}}\sum_{j\in\mathbf{J}} PI_{ij}v''_{ij} && (Low\ volume\ penalties\ for\ primary\ transport) \end{aligned}$$

Subject to:

$$\sum_{\substack{j\in\mathbf{J}\\ R(j,q)\leq CD}} x_{jqi} = 1 \qquad \forall i\in\mathbf{I}, q\in\mathbf{Q}; D_{qi}\geq 0 \tag{3.1}$$

$$\sum_{\substack{q\in\mathbf{Q}\\ R(j,q)\leq CD}} x_{jqi}D_{qi} = v'_{ij} - v''_{ij} \qquad \forall i\in\mathbf{I}, j\in\mathbf{J} \tag{3.2}$$

$$v'_{ij} \geq v_{min}(i)z_{ij} \qquad \forall i\in\mathbf{I}, j\in\mathbf{J} \tag{3.3}$$

$$v'_{ij} \leq Mz_{ij} \qquad \forall i\in\mathbf{I}, j\in\mathbf{J} \tag{3.4}$$

$$v''_{ij} \leq v_{min}(i)z_{ij} \qquad \forall i\in\mathbf{I}, j\in\mathbf{J} \tag{3.5}$$

$$\sum_{i\in\mathbf{I}} \sum_{\substack{q\in\mathbf{Q}\\ R(j,q)\leq CD}} x_{jqi}D_{qi} \geq minVol_j y_j \qquad \forall j\in\mathbf{J} \tag{3.6}$$

$$\sum_{i\in\mathbf{I}} \sum_{\substack{q\in\mathbf{Q}\\ R(j,q)\leq CD}} x_{jqi}D_{qi} \leq maxVol_j y_j \qquad \forall j\in\mathbf{J} \tag{3.7}$$

$$\sum_{i\in\mathbf{I}} x_{jqi}D_{qi} \geq V_{min}a_{jq} \qquad \forall j\in\mathbf{J}, q\in\mathbf{Q}; R(j,q)\leq CD \tag{3.8}$$

$$\sum_{i\in\mathbf{I}} x_{jqi} \leq Ia_{jq} \qquad \forall j\in\mathbf{J}, q\in\mathbf{Q}; R(j,q)\leq CD \tag{3.9}$$

$$z_{ij} \leq y_j \qquad \forall i\in\mathbf{I}, j\in\mathbf{J} \tag{3.10}$$

$$y_j,\ x_{jqi},\ a_{jq},\ z_{ij} \in \{0,1\} \qquad \forall i\in\mathbf{I}, j\in\mathbf{J}, q\in\mathbf{Q}; R(j,q)\leq CD \tag{3.11}$$

$$v'_{ij},\ v''_{ij} \geq 0 \qquad \forall i\in\mathbf{I}, j\in\mathbf{J} \tag{3.12}$$

With:
$M = min\,(maxVol_j, totProd_i)$.

The objective function consists in the total distribution cost, i.e. primary transport cost, secondary transport cost, transit cost and penalties for violating the minimum volume constraints on primary transport links. Each of the transport and transit cost components are computed as the unit cost per item multiplied by the number of transported items on each link/DC. Constraints (3.1) state that the demand of cluster q for the products of plant i is satisfied and is routed through a single DC (as x are binary variables). Constraints (3.2) stipulate that the total volume going from plant i to DC j is expressed as a difference between the continuous variables v'_{ij} and v''_{ij}. Constraints (3.3)-(3.5) mean that:

- If the link $[ij]$ is selected ($z_{ij} = 1$) then $v'_{ij} \geq V_{min}(i)$, $v'_{ij} \leq M$ and $v''_{ij} \leq V_{min}(i)$.
- If the link $[ij]$ is not selected ($z_{ij} = 0$) then $v'_{ij} = 0$ and $v''_{ij} = 0$.

This enables us to compute the violation of minimum volume constraints and to penalize it in the objective function. Constraints (3.6) state that if DC j is selected ($y_j = 1$), then the flows going through j have to be greater than the corresponding minimum volume limit. Constraints (3.7) stipulate that:

- If DC j is selected ($y_j = 1$) then the flows going through j must not exceed the corresponding maximum capacity.
- If DC j is not selected ($y_j = 0$) then there are no flows transiting through it (all the x_{jqi} have to be set equal to 0).

Constraints (3.8) ensure that if the link between j and q is selected ($a_{jq} = 1$) then the corresponding total volume has to be greater than the minimum volume V_{min} . Constraints (3.9) stipulate that if the link between j and q is not selected ($a_{jq} = 0$) then all of the variables x_{jqi} have to be set equal to 0. Constraints (3.10) stipulate that if DC j is not opened ($y_j = 0$) then all of the variables z_{ij} have to be set equal to 0. Constraints (3.11) and (3.12) are the integrality and non negativity constraints.

3.3.4 Complexity analysis

In this section, we aim at proving that the DNDMVD problem described in §3.3 is NP–Hard in the strong sense. The proof is completed by reducing the 3–partition problem (refer to Definition. 1), which is a NP–complete problem in the strong sense [Garey and Johnson, 1979] to the DNDMVD problem. In other words, we induce a special case of the DNDMVD problem into a 3–partition problem.

Definition 1. ***3–partition problem [Garey and Johnson, 1979]*** *Given the following instance: a finite set A of $3m$ elements, a bound $B \in \mathbb{Z}^+$ and a size $s(a) \in \mathbb{Z}^+ \ \forall a \in A$, such that each $s(a)$ satisfies $\frac{B}{4} < s(a) < \frac{B}{2}$ and such that $\sum_{a \in A} s(a) = mB$.*
Question is: can A be partitioned into m disjoint sets S_1, S_2,....., S_m such that, for $1 \leq i \leq m$, $\sum_{a \in S_i} s(a) = B$ (Notice that the above constraints on the item sizes imply that every such S_i must contain exactly three elements from A).

To this aim, we consider the DNDMVD problem on a special network as in Fig. 3.3, where:

- There is only one plant (and thus one product)
- There are m distribution centres having the same capacity $maxVol_j = B \forall j$ and a minimum throughput value $minVol_j \in [0, B] \, \forall j$, $B \in \mathbb{Z}^+$.
- There are $3m$ clusters having each a demand D_q such that $\frac{B}{4} < D_q < \frac{B}{2} \, \forall q$ and $\sum_{q \in Q} D_q = mB$
- All costs (transit and transport) are set to 0: $PTC_j = 0 \forall j$, $STC_{jq} = 0 \forall j \forall q$ and $TC_j = 0 \forall j$
- Minimum volumes on primary and secondary transport links are set to 0
- Maximum covering distance value CD is big enough such that all secondary transport links $[jq]$ are defined

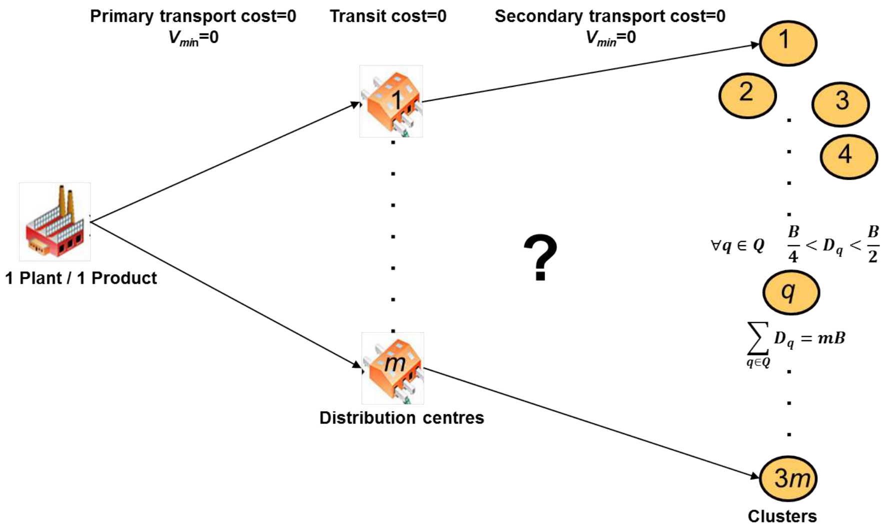


Fig. 3.3 Special distribution network corresponding to a 3–partition problem

As costs are set to zero and as all DCs have to be opened in order to satisfy customer demand, the specific instance of the DNDMVD model has a feasible solution (also optimal in that case) ***if and only if*** the 3–partition problem has a feasible solution. Therefore, the feasibility version of the DNDMVD problem is at least as hard as the 3–partition problem, and consequently NP–Hard in the strong sense.

3.4 Conclusion

In this chapter, we developed an original model for the design of distribution networks taking into account many realistic features that have never been combined together. Namely, in order to optimize truck loading, we considered minimum volume constraints on transport links. We also introduced minimum throughput and maximum capacity constraints for distribution centres, maximum covering distance constraints and single sourcing restrictions. After providing a detailed discussion about the modeling issues, we analyzed the different trade–offs that can occur in the optimization problem, in particular those related to the network structure (number and specialization of opened sites) and those impacting feasibility. Feasibility issues resulted from the conflict between minimum volume constraints on primary transport links and maximum covering distance constraints. This is why we chose to use penalties to model the first type of constraints.

In summary, economies of scale were taken into account using two methods:

- Implicit method through minimum volume constraints (for DC–cluster secondary transport and for the throughput of DCs).
- Explicit method through simple cost functions involving penalties (for plant–DC primary transport as considering strict minimum volume constraints in that case led to frequent

"infeasibilities").

Using more accurate cost functions was indeed impossible in view of the available data. This observation can be generalized to most real–life case–studies as it is very difficult to evaluate the cost as a function of the volume, especially when logistics operations depend on many third–party logistics and on varying market prices. Using minimum volume constraints makes the model easy to use by practitioners as this requires less data and mathematical expertise than if accurate cost functions were needed.

Based on the proposed mixed integer linear programming formulation, it is now possible to provide an exact solution using a state–of–the–art commercial solver. However, as the problem is NP–Hard, we expect to face computational difficulties when solving large size instances. This situation is likely to happen as we propose in the next chapter to carry out numerical experiments based on real–life data from the automotive industry. Thus, it may be necessary to implement heuristic methods to reduce running times.

Chapter 4

Application of the DNDMVD model to a case–study in the automotive industry: exact and heuristic solution methods

In the previous chapters, we discussed into details the modeling considerations related to the distribution network design problem with minimum volume and distance constraints (DNDMVD) as well as the pre–processing clustering procedure used in connection with a proper evaluation of the routing costs. In the present chapter, we propose exact and heuristic methods to solve the resulting MIP using real–life data from the automotive industry. The case–study we propose deals indeed with the vehicle distribution network of a car maker.

A recent literature review on facility location and supply chain management ([Melo et al., 2009a]) points out that a considerable number of works in this field propose heuristic solution procedures (about 50% of the surveyed papers). This is due to the increased number of discrete variables when dealing with "realistically sized problems". The authors mention that Lagrangian relaxation, linear programming based heuristics and metaheuristics are among the most popular techniques. In the specific field of facility location with minimum volume constraints, we showed in the literature review of chapter 3 that two main heuristic approaches are used, namely constant–factor approximation[1] and linear relaxation. However, the heuristics based on linear

[1] Constant–factor approximation algorithms are polynomial–time approximation algorithms with an approximation ratio bounded by a constant. For instance, a given algorithm is called c–approximation algorithm (for a given constant c) if it can be proven that it can find a solution which is at most c times worse than the optimal solution.

relaxation focused only on rounding fractional location variables because assignment variables were either not problematic in the solution procedure or continuous variables. Furthermore, according to [Melo et al., 2009a], there is a lack of applications in the literature of facility location applied to supply chain management. Three main reasons could explain this situation: 1. collecting data is difficult or even data are not available 2. when data are available, preparation and aggregation tasks are rather time–consuming 3. managers are not used to employ quantitative models for strategic decision support.

This chapter involves thus two main contributions of the present work to the existing literature. First, we provide detailed numerical experiments based on real–life data from a practical application. Second, we develop efficient MIP–based heuristic procedures using various types of linear relaxations to determine location and assignment variables (both sets of variables being required to be binary). The remainder of the chapter is organized as follows. Section 4.1 is devoted to the description and analysis of the case–study, the test instances and the exact solution given by a state–of–the–art MIP solver. In section 4.2, we investigate the heuristic solution procedures and the related computational experiments. Section 4.3 suggests some conclusions.

4.1 Case–study

4.1.1 Case–study description

The case–study that motivated this research deals with the car distribution network of the car maker Renault in France. The network is illustrated in Fig. 4.1 and consists of 16 assembly plants, 51 potential distribution centres and 448 car dealers. When an assembly plant is situated very far from the country and involves maritime transport, we consider the corresponding arrival port as a sourcing plant. In fact, combined transport with several transshipment points is out of the scope of this work.

Our study is based on the three–level distribution network structure illustrated in Fig. 3.1 with primary transport by truck from plants to distribution centres and secondary transport also by truck from distribution centres to car dealers. Given a deterministic and static demand of customers for each product, we aim at selecting the DCs to open among the list of potential DCs and to assign customers to them in such a way as to minimize the total distribution costs. As explained in §1.1.2, making the best use of transport capacities and in particular the loading of trucks is one of the priorities of automotive outbound logistics. This is why using the clustering approach defined in chapter 2 and introducing minimum volume constraints on transport links is relevant for this case–study. Moreover, as the distribution centres and the related activities are outsourced to third-party logistics, there are no fixed opening costs to be incurred by the car maker but there is a transit cost to be paid to the supplier each time a car goes through a DC. This transit cost only applies if the total throughput of the DC is between a minimum volume

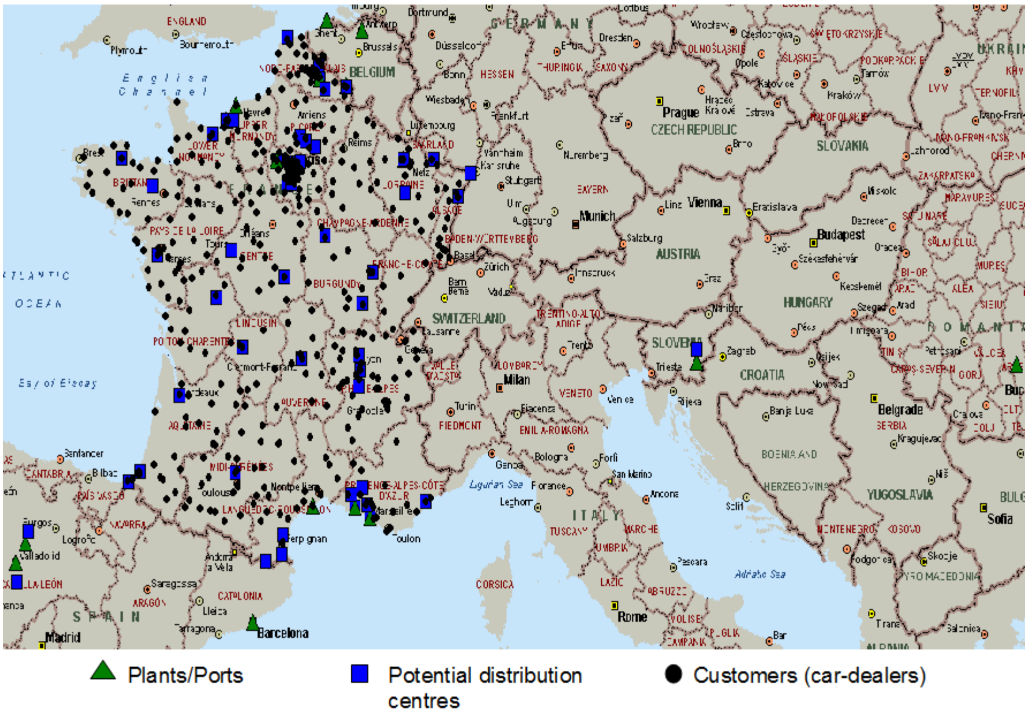


Fig. 4.1 Distribution network of Renault in France: plants/ports, potential distribution centres and car dealers

and a maximum capacity. Other constraints are considered, namely maximum covering distances and routing all the products related to a given plant/cluster flow through a unique DC. All these operational features are taken into account by the DNDMVD model presented in the previous chapter.

We assume that backloads[2] are already optimized through arrangements between the company and its carriers, thus a primary transport distance includes the length of the direct route from the plant to the DC together with an additional distance accounting for the empty truck trip necessary to get to a given destination (specified by the decision maker). Secondary transport distances are computed by solving Traveling Salesman Problems (TSP) on the customer groups (clusters) resulting from the pre–processing clustering procedure.

4.1.2 Implementation

We employed the C++ language to implement the program and the commercial solver ILOG Cplex version 12.5 to solve the resulting mixed integer programs. We carried out all the tests on a PC Intel Core(TM) i5–3210M (2.5 GHz) with 8 Gb of RAM, running under Windows 7. Input data were collected and stored in ”csv” files thanks to three different sources:

[2]Backloads are loads transported on the return journey of a delivery truck in order to minimize empty kilometers.

- Expertise of logistics managers
- Company databases
- Geographic Information Systems (GIS): We employed a free batch ([BatchGeocodeur, 2007]) to obtain the geospatial coordinates of the different sites and Microsoft MapPoint to evaluate the route distances

Fig. 4.2 summarizes the inputs and outputs of the implemented tool.

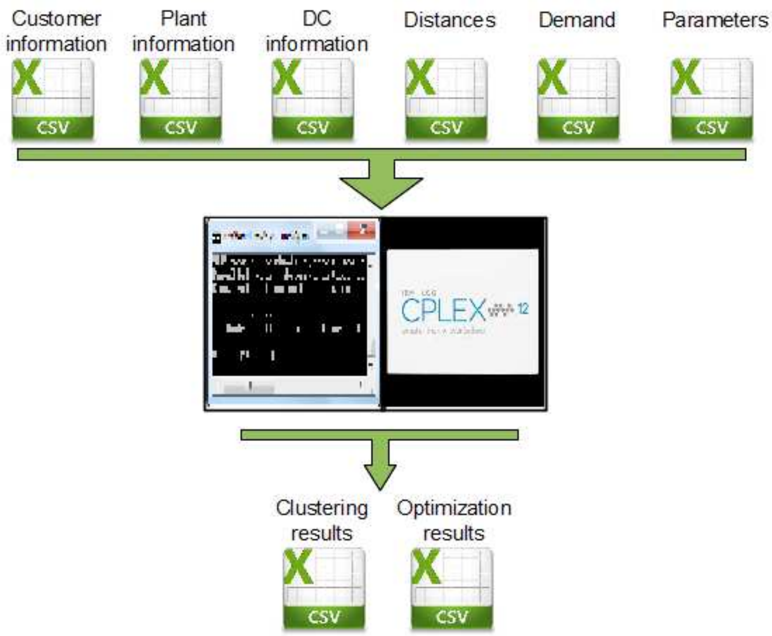


Fig. 4.2 Optimization tool: inputs and outputs

4.1.3 Test instances

Based on the preferences of the decision maker, we defined a reference dataset representing the typical values of input data and parameters for the distribution network in France. In order to test the performance of exact and heuristic solutions, we constructed also other instances by varying the main parameters of the problem in the reference dataset (using random perturbations in some cases).

4.1.3.1 Reference dataset

In the reference dataset, we fix the value of the maximum covering distance at 460 kilometres for all the secondary transport routes. The minimum and maximum volumes per DC are set according to some hypothesis on the flexibility of each DC. The minimum volume per primary transport link is fixed at 1 full truckload within a maximal waiting time of 5 days. The minimum volume per secondary transport link is also fixed at 1 full truckload within a maximal waiting time of 5 days. Other data like customer–to–customer distances, plant–to–DC distances and DC–to–customer distances are computed using geographic information systems.

As far as the clustering procedure is considered, we apply the optimal set–partitioning approach discussed into detail in chapter 2 (§2.3). We use the same clustering parameters employed in the tests of §2.3.3, namely a minimum volume per cluster set to 2 full truckloads per 5 working days, a maximum number of customers per cluster set to 3, a maximum distance between two customers in a cluster composed of two customers set to 80 kilometres and a maximum intra–cluster route in a cluster composed of three customers set to 240 kilometres (intra–cluster route is the shortest route visiting all the customers of the cluster starting and ending at a same customer). It should also be remembered that a difference of 1 truckload between the minimum quantity per cluster and the minimum quantity per secondary transport link gives more flexibility to the optimization algorithm with respect to the possibility of assigning a cluster to several DCs, each for one or several products. The clustering procedure found the optimal solution within 55 seconds and the map of results was already presented in chapter 2 (see Fig. 2.5). Applying an optimal clustering for this case–study is possible because the maximum number of customers per cluster is low, which leads to a set of potential clusters of manageable size, all the more so as we apply a maximum distance constraint when selecting potential clusters. However, if the problem involves more binary variables (if the total number of customer is high or if the maximum number of customers is greater than three) then the use of the heuristic clustering approach addressed in §2.4 may become the only feasible alternative.

4.1.3.2 Test instances A: varying the maximum covering distance value

We create 13 instances varying the value of the maximum covering distance from 460 kilometres to 700 kilometres and fixing the other parameters of the reference dataset. This leads to increasing the number of binary variables as well as the number of minimum volume constraints (see Table. 4.1).

4.1.3.3 Test instances B: varying the minimum and maximum volume parameters

We create 11 instances varying the minimum volume parameters for primary transport, secondary transport and DC throughputs as well as the maximum capacities for DCs (Table. 4.2). In two instances (B.6 and B.8) we also changed the maximum covering distance value because tight minimum volume constraints led to an infeasible problem.

4.1.3.4 Test instances C: varying customer demand

We generate 20 instances based on the reference dataset, each of which involves a pseudo–random coefficient in [0.5,1.5] per customer. We multiply the demand of each customer by the

Test instance	Maximum covering distance (KM)	Number of binary variables	Number of minimum volume constraints
A.1	460	37383	3015
A.2	480	39831	3159
A.3	500	42398	3310
A.4	520	44608	3440
A.5	540	46954	3578
A.6	560	49368	3720
A.7	580	51595	3851
A.8	600	54230	4006
A.9	620	56967	4167
A.10	640	60095	4351
A.11	660	62934	4518
A.12	680	65569	4673
A.13	700	68289	4833

Table. 4.1 Test instances A: varying the maximum covering distance value

Test instance	Min volume for primary transport (nb truckload/5 days)	Min volume for secondary transport (nb truckload/5 days)	Min throughput per DC	Max throughput per DC	Max covering distance (KM)
B.1	1	1	RefData	RefData	460
B.2	2	1	RefData	RefData	460
B.3	1	2	RefData	RefData	460
B.4	2	2	RefData	RefData	460
B.5	1	1	1.5*RefData	RefData	460
B.6	1	1	1.5*RefData	RefData	620
B.7	1	1	2*RefData	RefData	460
B.8	1	1	2*RefData	RefData	660
B.9	1	1	RefData	0.9*RefData	460
B.10	1	1	RefData	0.8*RefData	460
B.11	1	1	RefData	2*RefData	460

Table. 4.2 Test instances B: varying the minimum and maximum volume parameters. RefData denotes the data of the reference dataset.

corresponding pseudo–random coefficient.

4.1.3.5 Test instances D: varying the list of potential DCs

We generate 20 instances based on the reference dataset, each of which contains 46 potential DCs instead of 51. In fact, in each instance, we randomly select 5 DCs to remove from the list of potential DCs.

4.1.4 Exact solution using a commercial MIP solver

In the present section, we study the solution given by the commercial solver Cplex for the MIP proposed in §3.3. Different tests showed that proving optimality, in some cases, is very time–consuming. Consequently, we limited the final optimality gap of Cplex to 0.2% in order to reduce the computational effort (the optimality gap is defined as the relative difference between the best feasible integer solution obtained and the best lower bound found by the Branch & Bound procedure).

First, we study the setting of the penalty coefficients for minimum volume constraints on primary transport links. Then, we analyze the configuration of the distribution network obtained for the reference dataset. Finally, we study computation times through different test instances, particularly the impact of varying the main parameters of the problem on computation times.

4.1.4.1 Setting the penalty coefficients for minimum volume constraints on primary transport links

As already explained in chapter 3 (§3.2.6), due to the "infeasibilities" resulting from the conflict between minimum volume constraints on primary transport links and maximum covering distance constraints, we handle the minimum volume constraints on primary transport links using penalties. We penalize each unit below the targeted minimum volume by a given amount in the objective function. We suggested two possibilities to evaluate these penalties:

- The value of shipping trucks not fully loaded within the targeted waiting times
- The value of increasing waiting times in order to consolidate enough volume and to ensure full truckload transport

Nevertheless, the second alternative requires to evaluate the impact of lead time increase on the willing of customers to buy the products, which is very difficult to achieve either in our case–study or in other contexts of application. Thus, we chose the first alternative, which consists in keeping shipping trucks with the same frequency even if they are not fully loaded. To analyze this situation, we use the example of a primary transport link where the cost of a truck is set to 1000 and its capacity is fixed at 10 cars. The unit cost as a function of the weekly volume transported on the link was illustrated in Fig. 1.12 presented in chapter 1. Based on this first curve (consider the solid line), it is possible to represent the total transport cost as a function of the total transported volume (see Fig.4.3).

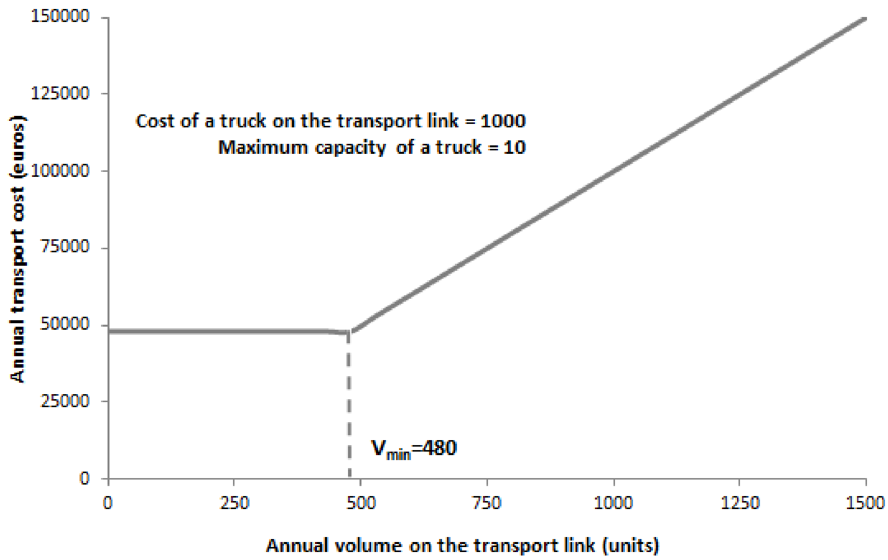


Fig. 4.3 Example of the annual transport cost as a function of the annual transported volume on a given link where the cost of a truck is set to 1000 and its capacity to 10.

Thus, if the annual volume on the transport link is greater than 480 units, all the products can be shipped using full truckloads, the unit transport cost in that case is 100 and no penalties are paid. On the contrary, if the volume on the link is less than 480, then regardless to the quantity, the total cost is equal to 48000, which corresponds to shipping a truck every week even if it is not fully loaded. In other words, an additional cost of 100 is paid for each unit below the minimum volume of 480. For instance, if the transported volume is 300, the amount of minimum volume violation is 180 units and the total cost is calculated as: $(100 * 300) + (100 * 180) = 48000$. The generalization of this situation is illustrated in Fig. 4.4.

Thus, if the volume on a primary transport link $[ij]$ is greater than the minimum required volume $V_{min}(i)$, all the products are shipped using full truckloads. A unit transport cost equal to $\frac{PTC_{ij}}{W_i}$ is paid, where PTC_{ij} is the cost of a truck going from plant i to DC j and W_i is the average truck capacity for the trips starting at plant i. On the contrary, if the volume on the link is less than $V_{min}(i)$, then regardless to the quantity, a total cost equal to $PTC_{ij} \cdot \frac{NWD}{T_{max}(i)}$ is paid, which corresponds to shipping trucks with a periodicity set to $T_{max}(i)$ even if they are not fully loaded.

4.1.4.2 Analysis of the solution obtained for the reference dataset

In this section, we run the DNDMVD problem on the reference dataset in order to examine the structure of the obtained network with regard to the main trade–off analysis provided in §3.2.6. Three features are thus discussed: the number of opened DCs, the penalties on primary transport

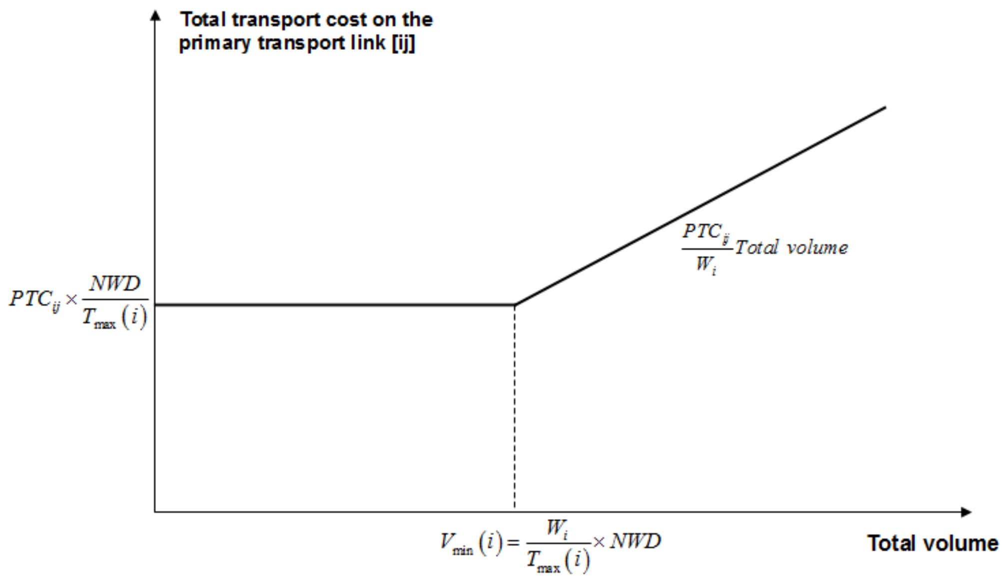


Fig. 4.4 Total transport cost as a function of the transported volume on a given primary transport link $[ij]$ when using unit transport cost as penalty.

links and the specialization of the opened DCs.

- Number of opened DCs: The network obtained when running the problem on the reference dataset is composed of 28 DCs: the geographic distribution of the opened DCs as well as the clusters of customers they serve are shown in Fig. 4.5.

  The optimal network configuration thus consists in opening a rather large proportion of the potential sites (28 among 51 $\sim$ 55%). This might be due to the tight maximum covering distance constraints and to the high value of the per car per kilometre cost on secondary transport links. Both parameters tend to reinforce the impact of secondary transport on the network structure. Basically, DCs are opened in such a way as to be as close as possible to customers in order to comply with the maximum covering distance constraints and to reduce the number of kilometres traveled on secondary transport links.

  Although a considerable number of DCs is opened, not all of the potential locations were used as this may violate minimum throughput constraints. The throughput of many opened DCs is indeed close to the minimum required quantity. This shows the impact of minimum throughput constraints on the network configuration.

- Penalties on primary transport links: as the number of opened DCs is relatively high (28), we can expect that minimum volume constraints are violated on several primary transport links. A further analysis shows however that violations of minimum volume constraints can be found only on 9.5% of the opened primary transport links. These violations are mainly related to the conflict between the minimum volume constraints and the maximum covering distance constraints. However, another reason could be that the amount of the penalty to

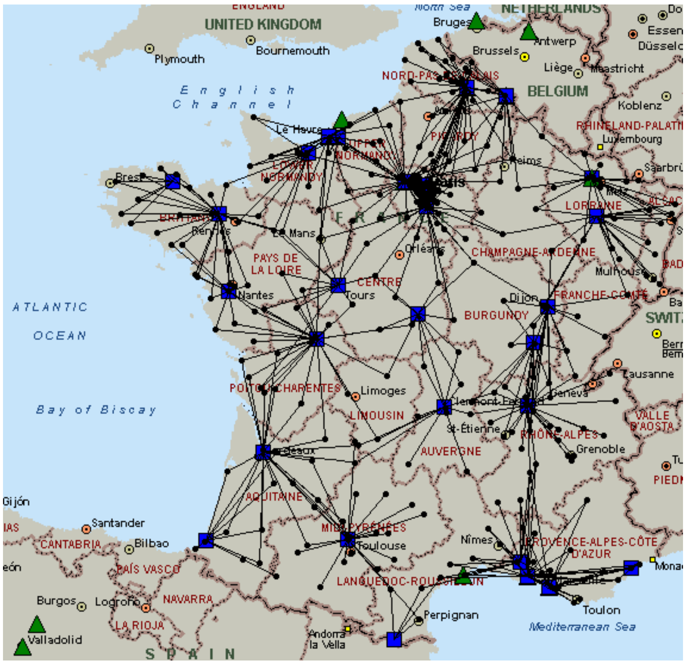


Fig. 4.5 Secondary transport in the optimal solution given by Cplex for the reference dataset. Plant locations are represented using triangles, opened DCs using squares, and cluster barycentres using dark circles.

be paid for violating the minimum volume constraints on some primary transport links is less than the amount saved in secondary transport when getting close to final customers.

- Specialization of the opened DCs: The geographic distribution of DCs illustrated in Fig. 4.5 shows that transport flows assigned to DCs are related to close customers and thus that DCs are likely to be specialized by region. Does that mean that we obtain a "specialization on a subset of customers" as explained in §3.2.6? This is not exactly the case but the resulting configuration is very close to that. In fact, the number of single assignments is important (34% of the clusters are assigned to only one DC). Other clusters are assigned to several DCs (between 2 and 5) either because the problem is highly constrained or because sometimes it is better to use a farther facility which costs less in terms of primary transport.

  Nevertheless, we noticed that the resulting configuration does not involve a specialization of DCs on a subset of products. The number of primary transport links per plant varies indeed between 16 and 28, which means that a given product is at least managed by 16 DCs. This shows again the great impact of secondary transport on the optimal decisions of the problem.

4.1.4.3 Analysis of computation times on different instances

In this section, we aim at studying the computation times of solving the DNDMVD problem with a commercial solver on various test instances. Even if the computation time needed by Cplex

MIP solver to solve the problem on the reference dataset is acceptable (42 minutes), varying the main parameters of the problem, namely the maximum covering distance and the minimum volume parameters, could lead either to increasing computation times or to running out of memory.

Impact of varying the maximum covering distance value Increasing the value of the maximum covering distance parameter leads to an increase in the number of possible assignments, which results in raising the number of the binary assignment variables and of the related constraints. For instance, according to Table. 4.1, the size of the problem goes from 37383 binary variables and 3015 minimum volume constraints (when the maximum covering distance is fixed at 460 kilometers) to 68289 binary variables and 4833 minimum volume constraints (when the maximum covering distance is set to 700 kilometers). Fig. 4.6 shows how the computation time increases when the maximum covering distance increases and the related detailed figures are summarized in Table. 4.3.

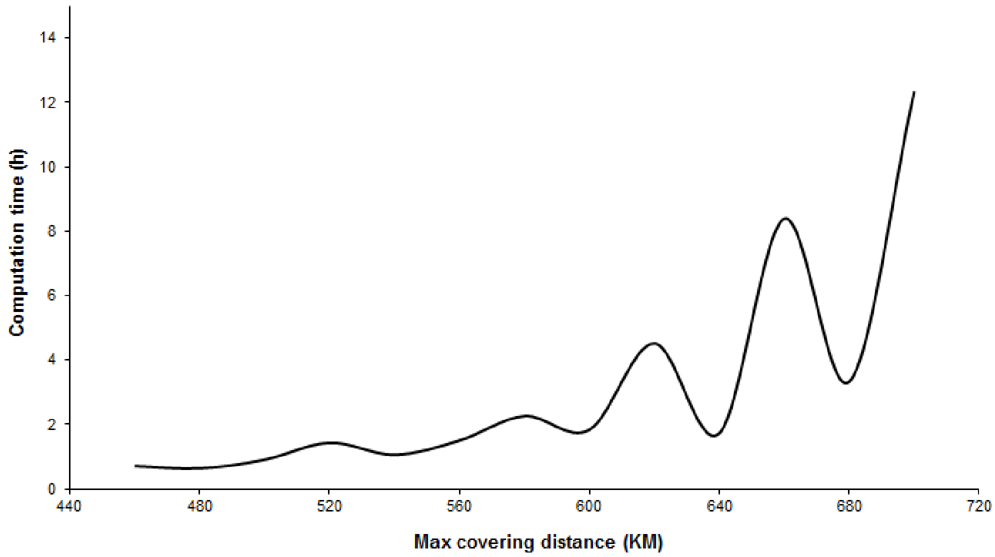


Fig. 4.6 Computation times of Cplex with test instances A varying the maximum covering distance value

We observe indeed that the computation time exceeds 12 hours when the distance is set to 700 kilometers.

Impact of varying the minimum and maximum volume parameters In Table. 4.2, we presented test instances B varying the minimum volume parameters on transport links and DCs as well as the maximum capacities of DCs. The related numerical results are summarized in Table. 4.4. Comparing the results of instances B.2 and B.3 with the result of instance B.1 shows that increasing the number of truckloads needed either on primary transport links or on secondary transport links makes the CPU increase. Simultaneously raising this number for primary and secondary transport even leads to running out of memory (instance B.4), which means that 8 Gb

Test instance	CPU of Cplex within 0.2% of exact optimality (hours)
A.1	0.7
A.2	0.6
A.3	0.9
A.4	1.4
A.5	1.1
A.6	1.5
A.7	2.3
A.8	1.8
A.9	4.5
A.10	1.7
A.11	8.4
A.12	3.3
A.13	12.3

Table. 4.3 Computation times of Cplex with test instances A varying the maximum covering distance value

of RAM was insufficient to cover the expansion of the branch and bound tree of Cplex in this case. The same behavior is also observed with instances B.6 and B.8 where we increased the minimum throughput per DC. Notice that "infeasibilities" were encountered with instances B.5 and B.7 due to the conflict between tight minimum volume constraints on DCs and maximum covering distance constraints. This is why we increased maximum distances for instances B.6 and B.8 but we faced all the same an "out of memory" status due to the difficulty of the problem. Varying the maximum capacities on DCs (instances B.9, B.10 and B.11) does not lead to increasing CPU but could result in an infeasible problem when capacities become insufficient to handle the customer demand subject to the other operational constraints of the problem (instance B.10).

Conclusion The analysis carried out in this section showed that the variation of the main parameters of the problem, in particular the maximum covering distance and the minimum volume parameters, can result in extensive computation times or in running out of memory. This is explained either by the difficulty to find the optimal solution due to tight constraints or by the fact that the problem is NP–Hard, thus it is difficult to manage instances with an increased number of integer variables and constraints. Moreover, this study has an interactive character: the decision maker wants to run many "what if" scenarios and requires to quickly view the impact of certain decisions. Therefore, implementing a heuristic method appears to be the best way to find a near–optimal solution in a reasonable runtime, so that the decision making process is facilitated without

Test instance	CPU of Cplex within 0.2% of exact optimality (hours)
B.1	0.7
B.2	3.0
B.3	1.8
B.4	Out of memory after 13 hours
B.5	Infeasible
B.6	Out of memory after 12 hours
B.7	Infeasible
B.8	Out of memory after 11 hours
B.9	0.4
B.10	Infeasible
B.11	0.6

Table. 4.4 Computation times of Cplex with test instances B varying the minimum and maximum volume parameters.

deteriorating the solution quality.

4.2 Heuristic solution procedures

In this section, we propose heuristic solution approaches for the MIP problem presented in chapter 3 (§3.3). In the previous section, we showed that a commercial solver such as Cplex can be directly used with the MIP to obtain exact optimal solutions. However, due to tight constraints or to a high number of binary variables when dealing with large size real–life instances, computation times can become prohibitively high and even running out of memory can occur. Therefore, we had to consider the development of specific and possibly more efficient heuristic solution methods. We first discuss into detail the various heuristic procedures that we implemented, then, we test their performances as compared to those of the solver Cplex applied to the original MIP.

4.2.1 Description of the heuristic procedures

4.2.1.1 Main idea

The main idea of the proposed heuristic methods is to exploit as much as possible the information provided by the optimal solution of the linear relaxation of the problem. An argument supporting this approach is the tightness of the lower bounds provided by the linear relaxation solution of the formulation defined in §3.3. Indeed, using the classical disaggregate formulation for facility location (see (4.1)), a strong linear relaxation is obtained leading to a lower bound very close to the exact optimal solution value (typical deviations are less than 1.3%).

$$x_{jqi} \leq y_j \qquad \forall j \in \mathbf{J}, q \in \mathbf{Q}, i \in \mathbf{I} \tag{4.1}$$

In the problem under study, we have two main types of binary variables: location variables and assignment variables. We propose a two–stage heuristic method: the first stage focuses on location decisions whereas the second one deals with assignment decisions. The location variables are fixed in the first stage and the corresponding values are used in the second stage. In the following paragraphs, we propose various heuristic methods for each stage of the solution procedure (see Table. 4.5).

First stage (S1): Location decisions	**Second stage (S2): Assignment decisions**
S1M1 *Rounding location variables*	S2M1 *Reintroduction of all the integrality constraints*
S1M2 *Partial linear relaxation of the original problem*	S2M2 *Gradual reintroduction of the integrality constraints*
	S2M3 *Fixing strategy cancelling inconsistent flows*
	S2M4 *Fixing strategy repairing inconsistent flows*

Table. 4.5 The different methods implemented in the two stages of the heuristic solution procedure

4.2.1.2 First stage (S1): Location decisions

In the first stage, we focus on deciding which DCs should be opened, i.e. we aim at determining the value of each location variable y in such a way as to be as close as possible to the optimal solution. According to the maximum covering distance and the minimum volume constraints for DCs, we can first derive some trivial decisions based on the following variable fixation rule:

For each DC j
Compute the maximum quantity Q_j that it could deliver according to the maximum covering distance constraint
If $Q_j \leq minVol_j$ **Then** impose not opening DC j ($y_j = 0$)

Then, two alternatives could be applied. The first alternative (S1M1) is a rounding strategy based on the optimal solution of the linear relaxation of the original MIP. The second one (S1M2) is to solve a partial linear relaxation of the original MIP.

Rounding location variables (S1M1) The first step in this method is to solve the linear relaxation of the original MIP, i.e. where all the integrality constraints are removed. Different tests showed that the optimal relaxed solution contains few fractional location variables y (typically less than 10%). Moreover, we observed that the location variables which are set to 1 in the optimal relaxed solution also take on the value 1 in the optimal solution of the original MIP. Thus, we fix all the variables having a value equal to 1 in the optimal relaxed solution. Then, we select the variable having the highest fractional value and set it to 0 or 1, depending on whether the value is below or above a given limit (L) and the linear relaxation is solved again subject to all variable fixations carried out so far. The algorithm continues iterating until there are no more fractional solutions for location variables (see Algorithm 3). This strategy appeared to drive the process to feasible solutions, yielding location decisions very similar to the ones provided in the optimal solution of the original MIP.

Algorithm 3. Rounding strategy in the location stage (S1M1)

```
Solve the linear relaxation (LR) of the original problem ;
Set to their current values all the location variables y having a solution equal to 1;
while there are fractional location solutions y do
    Find j such as y_j has the highest fractional value ;
    if y_j > L then
    |   set y_j to 1
    else
    |   set it to 0
    end
    Solve the LR of the model with the new constraint;
    if adding this rounding constraint leads to an infeasible problem then
    |   cancel it, add the opposite rounding constraint and solve the new problem
    end
end
Fix the location variables y at their final integer values
```

Solving a partial linear relaxation of the original MIP (S1M2) This method is based on solving a linear relaxation of the original MIP where the integrality constraints on variables x, z and a have been removed whereas they have been kept for y variables. Different tests showed that this is an efficient way to quickly identify the relevant main structure of the supply chain network, i.e. to determine the DCs to open.

4.2.1.3 Second stage (S2): Assignment decisions

In the second stage, we focus on deciding which DCs should serve which clusters, i.e. we aim at determining the value of each assignment variable x in such a way as to meet the minimum volume constraints for secondary transport. After fixing the location variables in the first stage of the heuristic approach, the resulting assignment problem is easier to solve than the original MIP. A rather large fraction (on average 44%) of the variables x are indeed set to 0 due to the fact that the corresponding DCs are not opened. In order to determine the values of the remaining assignment variables, four different methods were studied.

Reintroduction of all the integrality constraints (S2M1) The method S2M1 reintroduces the integrality constraints for all the free assignment variables x just after the location stage then solves the resulting MIP. This is much quicker than directly solving the original MIP as the number of binary variables is reduced after fixing the location variables.

Gradual reintroduction of the integrality constraints (S2M2) The method S2M2 consists in gradually including the relaxed integrality constraints. As previously mentioned, in the first stage, we fix the location variables y. Then, we add the integrality constraints to all the secondary transport related variables, namely, x and a. We solve the resulting MIP and we fix x and a at their values in the obtained solution. Finally, we add the integrality constraints to all the primary transport related variables (i.e. z) and we solve the resulting problem. Here, it is worth pointing out that the last step does not change any assignment decision as the x and a variables are fixed in the previous step. It only re–evaluates the objective function taking into account the penalty resulting from violating the minimum volume constraints on primary transport links.

Fixing strategies We now discuss two fixing strategies exploiting the information provided by the solution of the first stage of the heuristic. The idea is to try to fix as many assignment variables as possible before reintroducing the integrality constraints for the currently free variables and solving the resulting MIP with a commercial solver. The fixing strategy should enable us to decrease the number of binary variables before solving the final MIP and thus to reduce computation times in the second stage of the heuristic procedure. However, a key issue is to ensure that the fixing decisions thus made do not lead to infeasibility.

In order to implement the fixing strategies, we started at the optimal solution of the first location stage that we examined to identify the main issues. We noticed two problems: firstly, fractional values are obtained for only a small proportion of the assignment variables (of about 2%); secondly, there are many secondary transport links where the minimum volume constraint is violated (of about 50% of the secondary transport links). These figures clearly show that the violation of the minimum volume constraints for secondary transport links is the main issue in the optimal solution of the location stage. This is why we focus, in the fixing strategy, on modifying the related flows before reintroducing the integrality constraints and solving the final MIP. In this context, two possible approaches are studied: cancelling these flows by setting the corresponding variables to 0 (method S2M3) and repairing them in order to reach the required minimum value (method S2M4).

Fixing strategy by cancelling inconsistent flows (S2M3)

The fixing strategy based on cancelling flows is detailed in Algorithm 4. In a first step, we run a loop checking for each opened secondary transport link $[jq]$ violating the volume constraint if it is possible to set it to 0. Namely, we make sure that the throughput of DC j will be kept above its minimum volume $minVol_j$ if we move the $[jq]$ flow from DC j to another DC k. Furthermore, we check that there is another DC k able to deliver the demand of q without exceeding its maximum volume of throughput $maxVol_k$. If the two tests are successful, then we set the flow on $[jq]$ to 0. We repeat this procedure for all the secondary transport links then we solve the resulting linear problem. In case the problem becomes infeasible, we remove the lately added fixing constraints and stop the loop. Otherwise, we iterate the procedure until there are no further flows that can be fixed at 0 or the problem becomes infeasible.

The second step of the fixing strategy is to reintroduce the 0–1 constraints for all the free assignment variables and to solve the resulting MIP with a commercial solver.

Algorithm 4. Fixing strategy cancelling inconsistent flows (S2M3)

Data:

V_{jq}: Volume going through the secondary transport link $[jq]$ during the planning period;

$\forall j \in \mathbf{J}\ \forall q \in \mathbf{Q}$ *such as* $R(j,q) \leq CD \quad V_{jq} = \sum_{i \in \mathbf{I}} x_{jqi} D_{qi}$;

V_j: Volume going through the DC j during the planning period;

$\forall j \in \mathbf{J} \quad V_j = \sum_{\substack{q \in \mathbf{Q} \\ R(j,q) \leq CD}} \sum_{i \in \mathbf{I}} x_{jqi} D_{qi}$;

Algorithm:

Step 1

while *there are possible changes* **do**
- **for** *each opened transport link [jq] from DC j to cluster q* **do**
  - **if** $V_{jq} < V_{min}$ **then**
    - **if** $V_j - V_{jq} \geq minVol_j$ **then**
      - **if** $\exists$ *DC* k *within the covering distance* CD *of* q *such as* $(V_k + V_{jq}) \leq maxVol_k$ **then**
        - Add the constraint $V_{jq} = 0$ (i.e. $x_{jqi} = 0\ \forall i$)
      - **end**
    - **end**
  - **end**
- **end**
- Solve the resulting linear problem;
- **if** *infeasible problem* **then**
  - Cancel the lately added fixing constraints, break while
- **end**

end

Step 2

Reintroduce the integrality constraints for all the free variables then solve the resulting MIP;

Fixing strategy by repairing inconsistent flows (S2M4)

In a first step, we run a loop checking for each opened secondary transport flow $[jq]$ violating the volume constraint if there are DCs k, which are already serving the considered cluster q but able to give up these flows to DC j without falling under their minimum quantity of throughput. The objective here is to complete the $[jq]$ flow up to the required minimum volume without exceeding the maximum capacity of DC j. If the different tests are successful, then we set the $[kq]$ flows to 0. We repeat this procedure for all the secondary transport links then we solve the resulting linear problem. In case the problem becomes infeasible, we remove the lately added fixing constraints and stop the loop. Otherwise, we iterate the procedure until there are no further flows that can be fixed at 0 or the problem becomes infeasible. In the second step of the fixing algorithm, we reintroduce the 0–1 constraints for all the free assignment variables and solve the resulting MIP with a commercial solver.

Discussion about feasibility

In the implemented algorithms, we tried to develop a feasibility–care approach, which is illustrated through the tests that we made before fixing variables at 0 to prevent the algorithm from making inappropriate decisions, which will be hard to repair later.

As mentioned in the literature review, we found only four papers dealing with linear relaxation based heuristics for facility location or network problems with minimum volume constraints. The management of feasibility is among the main differences between our fixing algorithm and the solution methods implemented in these papers. In the present work, we try to keep feasibility while fixing assignment variables. In [Thanh et al., 2010], an infeasible solution is corrected after running the fixing algorithm using a feasibility pump. [Melo et al., 2009b] used a search in the neighborhood of the infeasible solution. In [Barros et al., 1998] and [Lim et al., 2006], the authors do not address the feasibility issues at the end of the heuristic procedure.

4.2.2 Experiments with the heuristic procedures

In this section, we analyze the computational performance of the two–stage heuristic approaches presented in §4.2.1 as compared to running a state-of-the-art commercial solver with the original MIP. We use the four test instances defined in §4.1.3 respectively varying the maximum covering distance value, the minimum and maximum volume parameters, the demand of customers and the list of potential DCs. We employ Cplex to solve the linear and MIP problems in the different steps of the heuristic approaches. We keep limiting the optimality gap of the solver to 0.2% off optimality as proving exact optimality, in some cases, is very time–consuming.

As previously explained, we implemented two heuristic methods in the location stage (S1M1 and S1M2) and four heuristic methods in the assignment stage (S2M1, S2M2, S2M3 and S2M4). The tests showed that the methods S1M1 and S1M2 have a similar computational behavior. Moreover, they usually lead to the same list of opened DCs, which is very similar to the list given by Cplex for the original MIP (1 or 2 different DCs at most). This is why, in the rest of this section, the numerical tests use only one of the two methods, which is S1M2.

S2M3 and S2M4 are both second stage heuristics based on a fixing strategy which aims at fixing as many assignment variables as possible before reintroducing the integrality constraints and solving the resulting MIP with the commercial solver Cplex. The objective of these two methods is to decrease the number of free binary variables before solving the final MIP and thus to reduce computation times in the second stage of the procedure. We thus propose to compare in Table. 4.6 the number of assignment variables x fixed by the two methods and measured as a percentage of the total number of variables x.

Test instance	Cumulative percentage (%) of variables x fixed in the location stage of the heuristic	Cumulative percentage (%) of variables x fixed after applying the fixing strategy of S2M3	Cumulative percentage (%) of variables x fixed after applying the fixing strategy of S2M4
A.1	46	64	73
A.2	44	63	73
A.3	45	63	73
A.4	43	62	71
A.5	42	42	72
A.6	42	61	70
A.7	43	62	73
A.8	42	62	72
A.9	42	62	72
A.10	42	62	74
A.11	41	60	74
A.12	41	60	76
A.13	40	59	74

Table. 4.6 Number of fixed assignment variables x with test instances A in the location stage of the heuristic procedure and after applying fixing strategy S2M3 or S2M4 (measured as a cumulative percentage of the total number of variables x)

Table. 4.6 shows that a significant percentage of assignment variables is already fixed in the location stage of the heuristic procedure. Then, if we apply the fixing strategy S2M4, we can fix until 76% of the total number of assignment variables. With S2M3, less variables are fixed and sometimes no variables are fixed (instance A.5), which simply results in applying heuristic S2M1. However, we cannot evaluate the performance of the two fixing strategies based only on these figures. We have to check computation times and solution qualities resulting from applying S2M3 and S2M4. In the following tests, we propose to compare the computational performance of the various heuristic procedures when applied to the test instances A. We measure the computation time in minutes and the solution quality as the relative difference between the objective value given by the considered heuristic solution and the objective value given by Cplex for the original MIP (within 0.2% of exact optimality). Table. 4.7 shows the computation time and solution quality for the four heuristic methods earlier defined and for Cplex applied to the original MIP when varying the maximum covering distance value. We also added the quality of the solver output for the original MIP within 10 minutes and 60 minutes respectively. In fact, this information is useful to analyze the performance of the two–stage heuristic approaches as compared to applying an exact solver.

On the one hand, it is seen that the best fixing strategy is the one cancelling the inconsistent flows (S2M3) as it yields solutions of better quality than those given by S2M4 within comparable computation times. The gradual reintroduction of the integrality constraints in the second stage (S2M2) leads to short running times but to higher deviations than S2M1 and S2M3 from the original MIP solutions (up to 1.75%). In fact, in the second step of S2M2, we ignore the minimum volume constraints for primary transport links as the corresponding variables could have fractional values. Consequently, some expensive decisions are made for the transport from plants to distribution centres. On the other hand, the computational results clearly show that S2M1 yields excellent quality solutions sometimes better compared to the MIP solver output for the original MIP (negative values in the second part of Table. 4.7). This could be explained by the fact that the MIP optimality gap was limited to 0.2% for the four heuristic methods and for Cplex with the original MIP. In return, the computation time of S2M1 could increase up to 40 minutes and in one of the test instances (A.7) the heuristic does not succeed to find any feasible solution due to an out of memory status. Therefore, to achieve a trade–off between time and value, we should apply the fixing strategy S2M3, which is indeed better than S2M1 in terms of running time and its solution is less expensive than the one provided by S2M2. According to the choice of the decision maker, it is possible to prioritize the solution quality (S2M1) or a trade–off between time and quality (S2M3).

In summary, using the two–stage heuristic approach appears to be significantly more competitive than running the solver on the whole MIP model of section 3.3. We can indeed notice according to the last two columns of Table. 4.7 that the solutions obtained by Cplex within 10 minutes CPU are far from the reference solution it provides within 0.2% of exact optimality. Within 60 minutes CPU, the solutions obtained by the solver in test instances A.11, A.12 and A.13 are worse than those found by S2M1 and S2M3 in less CPU. This conclusion was also validated through additional numerical experiments using test instances B, C and D defined in §4.1.3. The obtained results are summarized in Table. 4.8, Table. 4.9 and Table. 4.10.

CPU(mn)	Heuristic solution				Cplex applied to the original MIP
Test instance	S2M1	S2M2	S2M3	S2M4	Optimality gap set to 0.2%
A.1	1.2	0.3	0.3	0.2	42.3
A.2	1.4	0.4	0.3	0.2	38.0
A.3	1.4	0.6	0.3	0.3	54.0
A.4	5.5	0.6	0.4	0.4	85.7
A.5	6.8	0.7	6.4	0.3	63.3
A.6	9.2	1.2	0.5	0.4	89.9
A.7	OOM	0.7	0.4	0.4	135.4
A.8	2.5	0.6	0.4	0.4	109.8
A.9	16.9	1.3	0.4	0.3	270.9
A.10	10.4	0.7	0.6	0.4	103.4
A.11	21.6	1.4	1.4	0.4	503.1
A.12	39.8	1.0	1.4	0.5	199.6
A.13	26.3	1.8	1.3	0.8	740.7

Solution quality(%)	Heuristic solution				Cplex applied to the original MIP	
Test instance	S2M1	S2M2	S2M3	S2M4	Time limit set to 10mn	Time limit set to 60mn
A.1	-0.05	1.48	0.77	8.34	9.97	-0.01
A.2	-0.05	1.51	0.74	10.92	6.78	-0.06
A.3	0.09	1.54	0.86	11.29	6.34	-0.04
A.4	-0.06	1.65	0.76	11.13	10.62	0.05
A.5	0.03	1.70	0.03	11.51	14.80	0.02
A.6	0.03	1.63	0.98	10.18	18.35	1.13
A.7	-	1.57	1.00	14.01	18.96	6.69
A.8	-0.03	1.61	1.02	13.64	22.34	0.56
A.9	0.05	1.44	1.07	13.64	18.60	0.11
A.10	0.08	1.75	1.15	15.68	26.29	-0.01
A.11	0.03	1.68	0.92	16.26	27.53	10.69
A.12	0.05	1.63	0.96	17.52	28.75	14.05
A.13	0.02	1.51	0.94	19.26	15.67	9.58

Table. 4.7 CPU (mn) and solution quality (%) as a function of the maximum covering distance (test instances A). Solution quality is measured as the relative difference between the heuristic solution and the solution produced by Cplex applied to the original MIP within 0.2% of exact optimality. The heuristic uses method S1M2 in the 1^{st} stage and one of the methods S2Mi, $i = 1..4$ in the 2^{nd} stage. Optimality gap for any MIP used in the heuristics was limited to 0.2%. OOM denotes an out of memory status and "-" means that the solution quality cannot be evaluated as no reference solution was found.

4.3 Conclusion

Based on real–life data from our practical application in the field of car distribution, we focused in the present chapter on solving the complex distribution network design problem (DNDMVD)

CPU(mn)	Heuristic solution				Cplex applied to the original MIP
Test instance	S2M1	S2M2	S2M3	S2M4	Optimality gap set to 0.2%
B.1	1.1	0.3	0.3	0.2	42.4
B.2	5.4	0.5	1.0	0.4	182.0
B.3	2.7	1.0	0.4	0.2	109.9
B.4	69.1	1.0	3.0	0.2	OOM after 13 hours
B.5	INF	INF	INF	INF	INF
B.6	OOM	1.6	1.4	0.7	OOM after 12 hours
B.7	INF	INF	INF	INF	INF
B.8	7.2	1.4	7.1	884.7	OOM after 11 hours
B.9	2.8	0.5	0.3	0.3	26.8
B.10	INF	INF	INF	INF	INF
B.11	0.9	0.3	0.2	0.2	33.1

Solution quality(%)	Heuristic solution			
Test instance	S2M1	S2M2	S2M3	S2M4
B.1	-0.05	1.48	0.77	8.34
B.2	0.10	6.07	1.48	9.82
B.3	0.05	1.17	1.24	8.52
B.4	-	-	-	-
B.5	-	-	-	-
B.6	-	-	-	-
B.7	-	-	-	-
B.8	-	-	-	-
B.9	0.03	1.24	0.75	8.22
B.10	-	-	-	-
B.11	0.04	1.19	0.84	8.31

Table. 4.8 CPU (mn) and solution quality (%) when varying the minimum and maximum volume parameters (test instances B). INF denotes an infeasible problem. OOM denotes an out of memory status and ”-” means that the solution quality cannot be evaluated as no reference solution was found.

discussed in chapter 3. We first provided numerical results for the reference dataset defined by the logistics managers of the automotive company and studied the structure of the resulting distribution network. We then analyzed the impact of varying the main parameters of the problem on computation times. This analysis showed that computation times could significantly increase in some test cases and that the program may even run out of memory. To avoid these difficulties, we implemented several heuristic procedures. These procedures are based on various linear relaxations of the original MIP formulation of the problem. They were validated through extensive computational experiments where the produced solutions have been compared with those obtained using an efficient state–of–the–art MIP solver. The results of these experiments confirmed the good performance of the proposed heuristic approaches, both in terms of computation time and solution quality.

One of the limitations of the single–period DNDMVD model studied in chapters 3 and 4 consists in the assumption that the demand is static over the whole planning horizon. The demand of many industrial products and in particular cars displays however a clear seasonal pattern. We thus propose in the following chapter a multi–period extension of our DNDMVD model enabling us to take into account demand fluctuations from period to period.

CPU(mn)	Heuristic solution				Cplex applied to the original MIP
Test instance	S2M1	S2M2	S2M3	S2M4	Optimality gap set to 0.2%
C.1	1.0	0.8	0.2	0.2	45.4
C.2	0.7	0.6	0.2	0.2	32.2
C.3	1.8	0.6	0.2	0.2	27.2
C.4	1.0	0.6	0.3	0.3	26.0
C.5	1.1	0.4	0.3	0.3	78.3
C.6	0.8	0.3	0.2	0.2	11.8
C.7	0.9	0.2	0.2	0.2	19.6
C.8	1.5	0.2	0.2	0.2	19.8
C.9	1.0	0.8	0.3	0.3	30.6
C.10	1.0	0.9	0.3	0.2	32.6
C.11	0.7	0.8	0.3	0.3	19.3
C.12	1.1	0.4	0.2	0.2	25.8
C.13	0.8	0.6	0.2	0.2	20.0
C.14	1.1	0.4	0.3	0.2	37.6
C.15	1.5	0.5	0.4	0.3	19.9
C.16	0.8	0.7	0.2	0.2	36.5
C.17	1.1	0.2	0.2	0.2	34.8
C.18	0.7	0.4	0.2	0.2	6.2
C.19	0.8	0.6	0.3	0.3	26.2
C.20	1.1	0.9	1.8	0.2	23.9

Solution quality(%)	Heuristic solution			
Test instance	S2M1	S2M2	S2M3	S2M4
C.1	-0.02	1.00	0.82	7.74
C.2	0.14	1.21	0.90	7.41
C.3	-0.01	1.55	0.89	8.85
C.4	-0.03	1.14	0.75	7.87
C.5	0.00	1.70	0.90	7.56
C.6	-0.05	1.20	0.67	6.96
C.7	-0.07	1.44	0.87	8.51
C.8	-0.01	1.89	0.78	8.47
C.9	0.04	1.57	0.94	8.09
C.10	0.02	1.54	0.87	7.32
C.11	0.02	1.47	0.71	7.65
C.12	-0.06	1.51	0.92	8.11
C.13	0.04	1.33	0.83	8.22
C.14	0.01	1.49	0.91	8.53
C.15	0.07	1.45	0.50	8.00
C.16	0.07	1.47	0.88	8.43
C.17	-0.02	1.46	1.00	8.11
C.18	-0.06	1.41	0.77	8.17
C.19	0.03	1.42	0.46	8.48
C.20	-0.06	1.49	0.83	8.58

Table. 4.9 CPU (mn) and solution quality (%) when varying customer demand (test instances C).

CPU(mn)	Heuristic solution				Cplex applied to the original MIP
Test instance	S2M1	S2M2	S2M3	S2M4	Optimality gap set to 0.2%
D.1	1.1	0.8	0.6	0.8	43.0
D.2	3.5	0.4	0.5	0.4	20.3
D.3	0.8	0.4	0.4	0.4	23.5
D.4	1.0	0.6	0.3	0.2	47.9
D.5	1.4	0.3	0.2	0.2	16.2
D.6	0.7	0.5	0.3	0.3	73.5
D.7	0.9	0.3	0.2	0.2	20.8
D.8	0.7	0.5	0.2	0.2	18.6
D.9	1.0	0.4	0.2	0.2	29.3
D.10	0.9	0.4	0.2	0.2	16.1
D.11	1.3	0.4	0.2	0.2	19.1
D.12	1.1	0.4	0.4	0.4	34.9
D.13	0.8	0.5	0.2	0.3	23.0
D.14	1.0	0.3	0.3	0.2	45.9
D.15	0.9	0.5	0.2	0.2	20.9
D.16	0.8	0.6	0.2	0.2	25.7
D.17	0.7	0.3	0.1	0.2	11.8
D.18	0.9	0.5	0.3	0.3	16.2
D.19	0.9	0.5	0.3	0.4	45.9
D.20	2.6	0.6	0.2	0.2	23.5

Solution quality(%)	Heuristic solution			
Test instance	S2M1	S2M2	S2M3	S2M4
D.1	0.04	1.51	0.69	6.45
D.2	0.02	1.42	0.72	6.89
D.3	0.11	1.18	0.58	6.15
D.4	0.02	1.68	0.76	7.99
D.5	-0.08	0.81	0.48	7.18
D.6	0.14	1.34	0.67	7.46
D.7	0.01	0.86	0.60	6.67
D.8	0.11	1.09	0.63	7.22
D.9	-0.08	1.11	0.73	5.96
D.10	0.04	1.30	0.80	7.64
D.11	0.01	1.00	0.78	7.51
D.12	-0.02	0.89	0.50	5.95
D.13	-0.03	1.09	0.69	7.63
D.14	-0.06	1.43	0.75	8.57
D.15	-0.04	1.27	0.52	7.35
D.16	-0.11	1.08	0.55	7.96
D.17	-0.14	0.93	0.47	7.97
D.18	-0.05	1.26	0.59	7.88
D.19	-0.04	1.36	0.69	7.78
D.20	-0.01	1.35	0.67	7.42

Table. 4.10 CPU(mn) and solution quality(%) when varying the list of potential DCs (test instances D)

Chapter 5

A multi–period extension of the DNDMVD model

In chapter 3, we have proposed a mixed integer program allowing us to model a distribution network design problem capturing several operational features. Namely, the implemented model (DNDMVD) takes into account the location–routing aspect, maximum covering distance constraints, maximum capacity constraints on DCs, single sourcing restrictions and flow consolidation through a series of minimum volume constraints. Given a static demand of customers for the whole planning horizon, our main objective has been to determine the best location and assignment decisions while minimizing total distribution costs. However, one may wonder whether it makes sense to consider a static demand when there are fluctuations in sales. In fact, many industries are concerned by demand variation, in particular seasonality, i.e. fluctuations depending on seasons. For instance, Fig.5.1 illustrates the monthly fluctuations of passenger car registrations in France during years 2009–2012. The figure clearly shows that car registrations are seasonal with a considerable fall in summer (around the month of August).

Hence, it could be interesting to study the robustness of the problem decisions (mainly location decisions) when demand varies over time–periods according to some seasonality pattern. This is why we address in the present chapter a multi–period extension of the DNDMVD problem defined in chapter 3. We consider a small number of periods that should be determined according to the context of the industrial application. For instance, in the numerical experiments proposed in §5.4 we use a planning horizon of one year divided into four time–periods illustrating the seasonality in car distribution. Each period is characterized by a demand variation (seasonality) factor that will be applied to all customer demands (we assume that factors do not depend on products or on customers but only on periods).

Now, the question is how to deal with dynamic parameters and decisions in the distribution network design problem. On the one hand, we have to cope with the difficulty of modeling location–

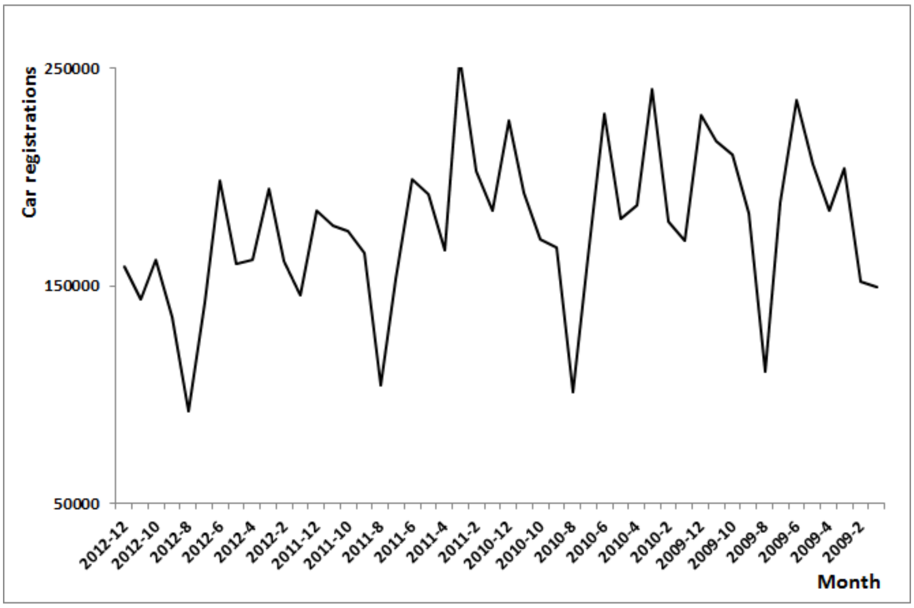


Fig. 5.1 Historical data of passenger car registrations in France. Source: INSEE

routing in a multi–period context. Hence, we introduce in §5.2.1.2 an iterative clustering procedure in order to estimate route and cost variation from period to period. On the other hand, we have to choose which decisions should be dynamic. As the planning horizon that we consider is not very long (one year in the case–study presented in §5.4) and distribution centres are related to fixed–term contracts, we cannot allow dynamic opening and closing of facilities. Those opened at the beginning of the planning horizon have to be operational over all the time–periods, which means that location decisions are static. However, this is not necessarily the case for assignment decisions that can be more easily adapted to demand variation. In our work, we propose to compare two situations : static and dynamic assignment decisions. A company can decide to keep the same distribution flows over the whole planning horizon in order to simplify day–to–day operations. However, this choice leads to additional costs that will be estimated in the numerical experiments detailed in §5.4.1.3.

As previously explained, the model we propose aims at taking into account the seasonality of demand according to time–periods but it could also be used to express demand uncertainty through discrete scenarios with fixed probabilities of occurrence. Our approach is indeed very similar to the one employed in two–stage stochastic location problems where location decisions are made in the first stage and assignment decisions occur after random parameters become known. Moreover, the introduction of time–periods or of scenarios makes the problem more difficult to solve due to an increased number of binary variables and constraints. Thus, to cope with the computational difficulties shown by numerical results, we also investigate a heuristic procedure to solve large instances of the multi–period distribution network design problem.

The present chapter is organized as follows. In the next section, we propose a literature review on dynamic facility location and dynamic supply chain network design problems. Section 5.2 is devoted to the explanation of the relevant modeling considerations of the multi–period problem and is followed by a detailed discussion on the proposed mathematical formulation in section 5.3. In section 5.4, we deal with a case–study in the automotive industry that helps us in the analysis of some features of the problem. We first discuss the main trade–offs resulting from demand variation. Then, we compare the solution of the multi–period problem with the solution of the single–period problem. Finally, we study the impact on location decisions and costs of using static assignments in the multi–period model as compared to using dynamic assignments. Numerical results show computational difficulties, thus we investigate in section 5.4.2 an efficient heuristic procedure based on linear–relaxation. Finally, some conclusions are provided in the last section.

5.1 Literature review on dynamic facility location and dynamic supply chain network design problems

In this section, we do not aim at reviewing all the works dealing with dynamic facility location and supply chain network design problems. Our intent is rather to propose a general overview of the main features of dynamic problems based on some review works. Then, we focus on a classification of the multi–period problems including minimum volume constraints with the aim of highlighting the main contributions of our work as compared to the existing literature.

5.1.1 An overview of dynamic facility location and dynamic supply chain network design

Dynamic systems are related to two main features: uncertainty (difficult prediction of input parameters) and time–dependency (parameters changing over time–periods). According to [Snyder, 2006], facility location problems under uncertainty could be modeled through different approaches, depending on the information that we have about the probability distribution of random parameters. If no probability information is known, then uncertain parameters are usually required to lie in some pre–specified interval. This leads to "robust" optimization where a standard goal is to optimize the performance of the system in the worst–case situation. "Stochastic optimization" deals with the cases where probability information is known and uses either continuous or discrete probability distributions. In the latter case, uncertain parameters are described through a set of discrete scenarios with a fixed probability of occurrence.

In our literature review, we mainly focused on time–dependent optimization problems that consider the way in which relevant parameters and decisions will change over time. In this context, we can talk about dynamic, multi–period or time–dependent facility location problems, as opposed to static facility location problems. As stated in [Owen and Daskin, 1998], "dynamic formulations transform snapshot models of one time decisions into extended horizon models which capture the temporal aspects of real–world problems". In real–world facility location problems, demand locations, demand quantities, costs, lead times, etc. are likely to vary from period to period over an extended time–horizon. To cope with these variations, sites could be opened/closed in different time–periods (see e.g. [Canel and Khumawalaz, 2001, Lee and Dong, 2009, Nickel et al., 2012, Melo et al., 2012]). However, closing facilities is costly and difficult at the operational level particularly if the planning horizon is not very long (only a few years). Thus, another alternative would be to open facilities for the whole planning horizon and to adapt assignment decisions to parameter variations. This is not a very common approach in the dynamic facility location literature but some references could be found in [Melo et al., 2009a]. Moreover, it is worth pointing out that this method is similar to the approach used in two–stage stochastic location problems where location decisions are made in the first stage and assignment decisions occur after random parameters become known (see [Kenyon and Morton, 2001] for more details).

In the context of our work, it was also interesting to look at multi–period location–routing problems (LRP). Some reference papers are those by [Yi and Ozdamar, 2007], [Afshar and Haghani, 2012], [Albareda-Sambola et al., 2012], [Laporte and Dejax, 1989] and [Prodhon, 2011]. However, the literature devoted to this topic is relatively scarce, mainly due to the large size of the mixed integer programs used in multi–period LRPs.

Even without taking into account the routing aspect, considering several periods could lead to computational difficulties when solving optimization problems. To explore the literature on solution methods and related numerical tests when using dynamic decisions, we focused on both multi–period and stochastic problems with discrete scenarios. The problem structure is indeed similar in the two cases. Most of the surveyed works study optimization problems which are already highly constrained in their static/deterministic versions, hence it is very difficult to consider several time–periods or discrete scenarios when running numerical tests. For instance, among the 22 papers that we surveyed in multi–period facility location and supply chain network design, only 6 of them present numerical experiments with a number of time–periods greater than or equal to 12. [Gebennini et al., 2009] (30 periods), [Lin et al., 2009] (52 periods), [Ghaderi and Jabalameli, 2013] (20 periods) and [Canel and Khumawalaz, 2001] (15 periods) use small instances in terms of plants, facilities and customers. In [Albareda-Sambolaa et al., 2009], numerical results are analyzed for tests with 12 periods leading to more than 180000 binary variables but the problem constraints are not particularly difficult to handle and most of the instances are successfully managed by the MIP solver Cplex within few hours. [Lee and Dong, 2009] study 50–period dynamic network design problem in a stochastic environment. The authors mention that the deterministic mean–value problem solution is manageable using Cplex and that the stochastic problem is solved only using a simulated annealing heuristic approach. However, no information is given about computation times with the exact and heuristic methods. Furthermore, few papers deal with facility location and network design problems using discrete scenarios with fixed probability of occurrence. We can, all the same, refer to [Crainic et al., 2009] for a stochastic network design problem and to [Nickel et al., 2012, Alonso-Ayuso et al., 2003] for combining multi–period and stochastic features in supply chain network and planning problems.

5.1.2 Multi–period facility location and supply chain network design problems with minimum volume constraints

Highly constrained facility location and supply chain network design problems are already difficult to solve in their static versions. This is what we showed when studying the distribution network design problem with minimum volume and distance constraints in chapters 3 and 4. Thus, it is not surprising that multi–period or stochastic facility location problems including minimum volume constraints are scarce in the literature. For instance, among all the works cited in the literature review of chapter 3 (see Table. 3.1), only 6 of them are dynamic. Table. 5.1 shows a classification of these works according to 6 different criteria:

- Dynamic decisions: specifying which decision variables depend on time–periods.
- Details on how to estimate parameter variation: explaining how to evaluate the parameters from period to period (demand, routes, costs, etc.).
- Case–study: telling if the paper involves a practical application or not.
- Discussion about feasibility issues resulting from parameter variation (when problems are highly constrained).
- Comparison of dynamic to static model outputs: this means comparing costs and decisions in the two cases and confirming or denying the benefits of dynamic modeling.
- Maximum number of periods considered in the numerical experiments presented in each paper.

Other information about modeling features and solution methods in these works were already illustrated in Table. 3.1.

As can be seen from Table. 5.1, no much importance is given by authors to qualitative analysis when developing dynamic problems. Only [Alumur et al., 2012] provide details on how to evaluate data variation from period to period using assumptions about volume increase and yearly inflation rates. Moreover, their work explains the advantages of a dynamic model as compared to a static one based on a case–study in the context of reverse logistics network design for washing machines and tumble dryers. It is also noticeable from the table that it is more common to use dynamic location decisions as well as dynamic flow decisions because extended planning horizons are considered.

Paper	Dynamic decisions		Qualitative study			Numerical tests	
	Locations	Assignments/Flows	Details on how to estimate route and cost variation	Discussion about "infeasibilities" resulting from parameter variation	Comparison of dynamic to static model outputs	Case–study	Max number of periods
[Melo et al., 2009b]	X	X					8
[Thanh et al., 2010]	X	X					5
[Melo et al., 2005]	X	X					10
[Alumur et al., 2012]	X	X	X		X	X	5
[Ndiaye and Alfares, 2008]		X				X	2
[Correia et al., 2013]	X	X					4
The present work		**X**	**X**	**X**	**X**	**X**	**4**

Table. 5.1 Literature review on multi–period facility location and network problems featuring minimum volume constraints.

5.1.3 Contributions of our study

As pointed out in [Melo et al., 2009a], literature about multi–period facility location problems for supply chain is rather scarce as more than 80% of the papers surveyed by the authors deal with single–period problems. In fact, the multi–period aspect increases the size of optimization problems and usually leads to computational difficulties, especially when the problem is already hard to solve in its static version. This is why there are few papers simultaneously addressing routing and multi–period features in facility location and supply chain network design problems whereas several works deal with static location–routing problems. One of the contributions of this chapter is thus to provide a model for multi–period supply chain network design taking into account the routing aspect as well as other operational features. Through a dynamic clustering procedure, we propose an original way of modeling the impact of demand variation on delivery routes. This approach enables us to reach a good trade–off between the representation of operational details in the model and its computational tractability.

Another contribution of this chapter is to provide a comprehensive qualitative study about the main features of dynamic modeling based on a real–life case–study in the automotive industry. Namely, we present in §5.2.1 a new aggregation approach as well as a clustering method to evaluate routes and costs per time–period. Numerical experiments related to the case–study are provided in §5.4 and used to draw some managerial conclusions. In §5.4.1.1, we discuss the main trade–offs resulting from demand variation in a highly constrained supply chain network design problem and show how to overcome the resulting "infeasibilities". In §5.4.1.2, we study the outputs of the multi–period extension of the DNDMVD problem as compared to its single–period version. In §5.4.1.3, we evaluate the difference between using static assignments and using dynamic assignments in terms of costs and location decisions.

Finally, an efficient heuristic procedure based on linear relaxations is proposed to solve the multi–period problem (see §5.4.2) whereas only few works used this kind of approach to tackle facility location problems with minimum volume constraints.

5.2 Problem modeling

In this section, we focus on modeling demand variation in the distribution network design problem introduced in the previous chapters. First, we explain the reasons for using a new aggregation approach and present the methods we studied. Then, we discuss the implementation of the multi–period problem.

5.2.1 Demand aggregation

5.2.1.1 Need for aggregation

One may wonder why it is necessary to consider an aggregation approach different from the clustering–based one used in the DNDMVD problem. In fact, considering multiple periods introduces in our problem two main difficulties that we propose to overcome by using a more aggregate representation of the problem.

- First difficulty: It concerns the pre–processing clustering step whose purpose is to construct groups of close customers meeting a required minimum volume. The objective of this step is to define delivery routes ensuring full truckload transport at the operational planning level. Now, the problem is that routes have to be adapted to demand changes. Basically, when demand decreases, we have to visit more customers to optimize truck loading within the allowed waiting time, hence routes become longer. This means that we have to define a different set of clusters for each considered period. Therefore, a given customer may belong to different clusters from period to period and thus be assigned to distinct DCs. However, as we aim at comparing a static assignment approach (i.e. for a given product, customers are assigned to the same DC whatever the period) to a dynamic assignment approach, using a detailed representation based on a different cluster set per period is not an appropriate modeling option.
- Second difficulty: The number of clusters in the single–period problem that we studied in chapters 3 and 4 is already large. Thus, if we introduce in our model a different cluster

set for each time–period, the problem size will significantly increase, which would result in prohibitive computation times.

Due to these difficulties, we propose to introduce a steady customer aggregation that can be adopted in all periods. We use country districts as final customers in our MIP formulation because the number of districts in a country is significantly less than the number of clusters. Using districts as final customers mainly impacts secondary transport starting from distribution centres. Fig. 5.2 illustrates the difference between the aggregation approach based on districts and the detailed approach explained in chapter 3 and based on clusters.

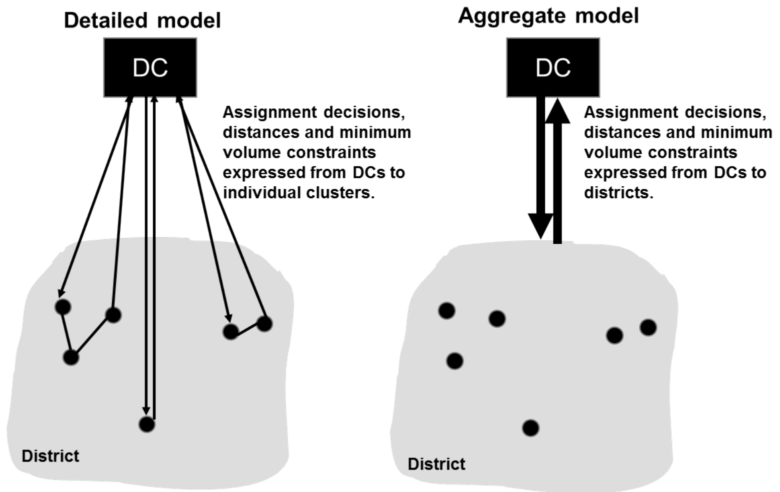


Fig. 5.2 Illustration of the aggregate distribution network design model as compared to the detailed model defined in chapter 3

5.2.1.2 Aggregation approach

In the aggregation approach, an assignment decision is made for a whole district and concerns all the customers belonging to the district. Similarly, distances are expressed for DC–district transport links instead of detailed routes from DCs to clusters. Minimum volume constraints on secondary transport links are also aggregated per district. The next question that we have to address concerns how to estimate secondary transport distances from DCs to districts in order to minimize the loss of information induced by aggregation as compared to the detailed method studied in chapter 3.

Two alternatives are possible A first alternative could be to adopt a rather standard approach (AG1) using distances to district barycentres whereas a second alternative (AG2) consists in applying a more sophisticated clustering approach per district to evaluate average distances from DCs to districts (see below for a detailed description of AG2). A numerical comparison of these two aggregation methods with the detailed method is presented in appendix B. It shows that the second alternative (AG2) results in a better approximation of the detailed problem than AG1. Namely:

- The solution given by AG2 presents a smaller number of violation of the detailed operational constraints (related to route construction).
- Thanks to a better estimation of secondary transport routes, AG2 provides a network structure (opened DCs) closer to the one obtained with the detailed model.

We therefore chose to apply aggregation approach AG2 in the sequel of our study.

Description of the chosen alternative (AG2) It consists in applying a pre–processing clustering for each district separately. This clustering only aims at estimating DC–district transport costs in each time–period and not at defining customer groups as was done in chapter 2. The clustering result within a given district is likely to be different according to the DC delivering products and according to the considered time–period (see appendix D for numerical details). We thus iteratively apply the clustering for each triple district/DC/period. The overall procedure is illustrated in Fig. 5.3. It involves two main steps:

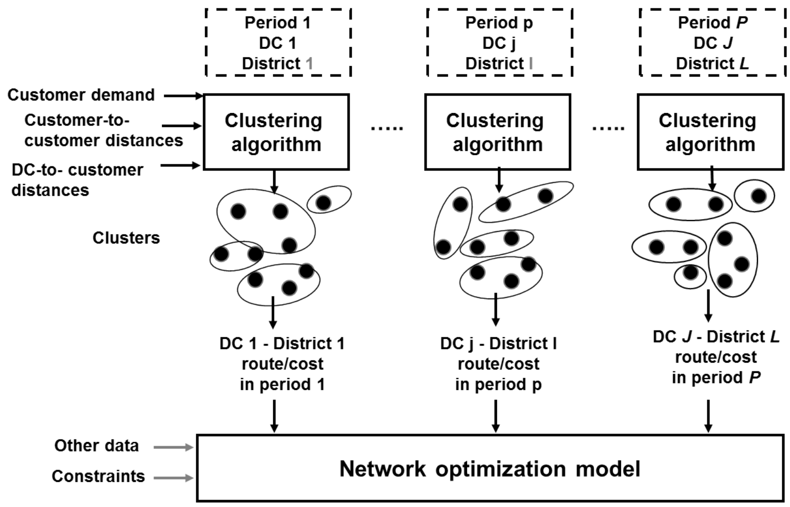


Fig. 5.3 Clustering approach in a multi–period network design problem

1. Applying optimal clustering for each district/DC/period triple

   The geographical region concerned by clustering is limited (only one district). We thus can use an exact clustering method based on set–partitioning (see chapter 2, §2.3). This procedure is applied to almost all district/DC/period triples except those where the district is situated very far from the DC. In fact, as the network optimization model contains a maximum covering distance constraint, a district cannot be served by a DC situated beyond a given distance from it. In this case, it is useless to apply clustering. We simply compute the distance from the DC to the nearest customer within the concerned district. If this distance is greater than the maximum covering distance, then we do not run the clustering. If not, we apply the two steps of clustering defined in §2.3:

- Generation of potential clusters: As the number of customers per district is not very high, we can easily generate all potential clusters. We consider the same selection constraints defined in 2.3.1, namely a minimum volume and a maximum number of customers per cluster but we do not include a distance constraint as customers are already close to each others in a same district.
- Set–partitioning integer program: We consider the integer program defined in §2.3.2.3 but we use a different way to estimate the costs assigned to the selection of each potential cluster (c_k). In fact, as we apply the clustering per DC, we can evaluate the cost of selecting cluster k as the total delivery cost from the DC to the cluster k. To achieve this, we first compute the shortest route starting at the DC, visiting all the customers of the cluster then coming back to the DC. This is a traveling salesman problem that can be solved using full enumeration if the number of customers per cluster is low. We then compute the cost of a truck $T_{optRoute}$ on this optimal route. Given the average truck capacity W and the total demand of the cluster k (D_k), we can compute the total delivery cost from the distribution centre to the cluster k as $\frac{T_{optRoute}}{W} D_k$.

  Notice that in practice, all the demand of a cluster is not necessarily delivered by a unique DC as a district could be assigned to different DCs according to products. The total delivery cost we obtain is therefore not fully accurate since it is based on the total demand of the cluster. Nevertheless, as we are in a pre–processing step, considering the total demand of customers is a good way to estimate secondary transport delivery routes and costs.

2. Calculating secondary transport costs

   Unit secondary transport costs from DCs to districts can be directly deduced from the output of the set–partitioning integer program. The total delivery cost of a given district l from a given DC j is computed as the cost of delivering all the clusters belonging to the district. Then, the average unit cost is obtained by dividing this total cost by the total demand of district l. This is equivalent to computing the weighted average cost over clusters with cluster volumes as weights (see Fig. 5.4 for an example of evaluating weighted average routes). As far as the minimum volume constraints are concerned, we exploit information that we have on routing inside each district. In the case–study presented in §5.4, we will consider the quantity ensuring on average full truckload transport from DCs to districts within the allowed waiting time multiplied by the number of clusters in the district.

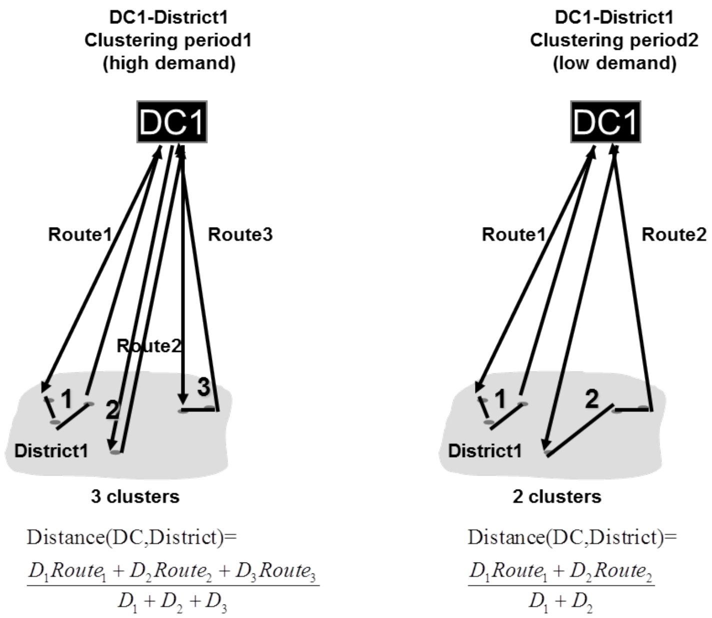


Fig. 5.4 Clustering per district/DC/period. Example of 2 time–periods for a given DC and a given district. D_i denotes the total demand of cluster i. $Route_i$ is the optimal route visiting the customers of cluster i

5.2.2 Multi–period approach

Before formulating the multi–period distribution network design problem in the next section, we first describe the principles of the multi–period approach we intend to use. The main difference between the multi–period and the single–period distribution network design problems consists in considering an input demand that varies over periods instead of a static one. This leads to three questions that we have to answer before discussing the mathematical formulation of the problem: how to evaluate the demand fluctuation? what is the impact of this fluctuation on the other parameters of the problem? what are the consequences on the location and assignment decisions? We assume that customers are aggregated per district as explained in the previous paragraph.

- How to represent the demand fluctuation?

  First, we have to determine the number of time–periods according to the related application, then, we apply a demand variation factor F_p for each period p. In our study, we consider only the impact of seasonality on demand and thus we do not differentiate demand variation according to customers or to products but we assume that F_p depends only on the time–period (season) p. Therefore, demand for district l, for the products of plant i, for period p is computed as: $D_{pli} = F_p \frac{D_{li}}{P}$ where D_{li} is the total demand of district l for the products of plant i for the whole planning horizon and P is the number of periods.

  Moreover, we do not use sophisticated demand forecast tools to estimate seasonality factors but we suppose that these factors are input to our study. We rather focus on studying the impact of demand fluctuation on the distribution network.

- What is the impact of demand fluctuation on the other parameters of the problem?
  Demand fluctuation mainly impacts secondary transport as delivery routes have to be changed in each time–period in order to ensure full truckload transport within the allowed maximum waiting time on DCs. This could lead to different clustering results and thus to different route computation according to time–periods. This may impact:
  - The evaluation of secondary transport costs per time–period
  - The expression of maximum covering distance constraints per time–period
  - The expression of minimum volume constraints for DC–district links per time–period (as it takes into account the number of clusters per district)
- What are the consequences on the location and assignment decisions?
  In a multi–period context, it is possible to use dynamic decision variables, i.e. to allow some modifications of the network configuration within the planning horizon, in order to decrease the overall cost. These modifications can involve opening a different set of DCs in each time–period or changing DC–district assignments from period to period.

  In our case, it is very difficult to allow dynamic DC openings and closings. In fact, the planning horizon we are considering is not very long (one year in the case–study presented in §5.4) and distribution centres are related to fixed–term contracts, hence it is not possible to change their locations from period to period. Accordingly, we assume that location decisions are static over time–periods.

  However, it is possible to use dynamic DC–district assignments, as it is easy to adapt these decisions at the operational level. We propose in the sequel to study two alternatives: static and dynamic assignments, in order to compare them. In fact, a company can either decide to change distribution flows from period to period or to keep the same flows over the whole planning horizon in order to simplify day–to–day operations. The latter choice leads however to additional costs that we propose to estimate in the numerical experiments presented in 5.4.1.3.

5.3 Formulation of the multi–period distribution network design problem

Using the results of aggregation method AG2 and the multi–period approach presented in the previous section, we can formulate the multi–period distribution network design problem. In this section, we present the variables and constraints used in the two versions that will be tested: dynamic versus static assignment decisions.

5.3.1 Model parameters

I	Set of plant indices ($i = 1..I$)
J	Set of DC indices ($j = 1..J$)
L	Set of district indices ($l = 1..L$)
P	Set of time–period indices ($p = 1..P$)
F_p	Demand variation factor corresponding to time–period p.
D_{li}	Demand of district l for the products manufactured in plant i during the whole planning period.
D_{pli}	Demand of district l for the products manufactured in plant i and time–period p, $D_{pli} = F_p \frac{D_{li}}{P}$.
$minVol_j$	Minimum volume of cars that has to go through DC j in each time–period if it is selected in the solution.
$maxVol_j$	Maximum volume of cars that can go through DC j in each time–period if it is selected in the solution.
PTC_{ij}	Cost of a truck going from plant i to DC j (Primary transport cost).
STC_{pjl}	Average cost of a truck starting its route at DC j and delivering products to district l then coming back to j (Secondary transport cost).
TC_j	Unit transit cost for a product going through DC j.
W_i	Average truck capacity for the products manufactured in plant i.
W	Average truck capacity for the whole volume of products.
M	Big value, set to $M = min\,(maxVol_j, totProd_{pi})$.
NWD	Number of working days in the whole planning horizon.
$T_{max}(i)$	Maximum waiting time allowed at plant i before shipping is made to distribution centres.
T	Maximum waiting time allowed at a distribution centre before shipping is made to customers.
CD	Maximum covering distance parameter (i.e. the maximum length of a route starting at a DC, delivering products to a given district then coming back to the DC).
$R(p, j, l)$	Average length of the route starting at DC j, delivering products to district l then coming back to j during time–period p.
$totProd_i$	Volume produced by plant i during the whole planning horizon.
$totProd_{pi}$	Volume produced by plant i during time–period p, $totProd_{pi} = F_p \frac{totProd_i}{P}$.
$V_{min}(i)$	Minimum volume that has to go through primary transport routes starting at plant i in each time–period. Has to be equal at least to the period volume ensuring on average a full truckload within $T_{max}(i)$,

	hence $V_{min}(i) \geq \frac{W_i}{T_{max}(i)} \frac{NWD}{P}$.
V_{min}	Minimum volume that has to go through secondary transport routes in each time–period. Has to be equal at least to the period volume ensuring on average a full truckload within T, hence $V_{min} \geq \frac{W}{T} \frac{NWD}{P}$.
PI_{ij}	Low volume penalty amount for primary transport from plant i to DC j.
$nbClusters(p, j, l)$	Number of clusters of district l during time–period p if delivered from DC j.
$V_{min}(p, j, l)$	Minimum volume that has to go through the secondary transport route delivering district l from DC j in time–period p. It is set to $nbClusters(p, j, l) * V_{min}$

Secondary transport cost STC_{pjl} is computed as the weighted average cost of a truck starting its route at DC j delivering products to clusters of district l then coming back to j. We consider cluster total demands as weights and we compute each DC–cluster route by solving a traveling salesman problem (TSP).

5.3.2 Decision variables

We first present variables not depending on the assumption of dynamic or static assignments, namely binary location variables y and continuous variables v' and v'' used to express the violation of minimum volume constraints on primary transport links.

- Location variables:

$$y_j = \begin{cases} 1 & \text{if DC } j \text{ is selected} \\ 0 & \text{otherwise} \end{cases}$$

- v'_{pij},v''_{pij}: Continuous variables used to write the minimum volume constraints on primary transport links in each time–period p.

  v'_{pij} is a positive variable that has to be greater than $V_{min}(i)$.

  v''_{pij} is a nonnegative variable used to compute the amount of violation of the minimum volume constraint on a given primary transport link [ij] in time–period p. It has to be less than $V_{min}(i)$ and will be minimized, null if possible, as it is penalized in the objective function.

5.3.2.1 Dynamic assignments

If we choose to use dynamic assignments then the following decision variables depend on time–periods:

- Assignment variables:

$$x_{pjli} = \begin{cases} 1 & \text{if district } l \text{ is assigned to DC } j \text{ for the products of plant } i \text{ in time–period } p \\ 0 & \text{otherwise} \end{cases}$$

- Variables stating if secondary transport links are selected:

$$a_{pjl} = \begin{cases} 1 & \text{if district } l \text{ is assigned to DC } j \text{ for at least one product in time–period } p \\ 0 & \text{otherwise} \end{cases}$$

x_{pjli} and a_{pjl} are defined only if $R(p, j, l) <= CD$ (maximum covering distance constraint).

- Variables stating if primary transport links are selected:

$$z_{pij} = \begin{cases} 1 & \text{if the route from plant } i \text{ to DC } j \text{ is selected in time–period } p \\ 0 & \text{otherwise} \end{cases}$$

5.3.2.2 Static assignments

If we choose to use static assignments then decision variables x, z and a do not depend on time–periods:

- Assignment variables:

$$x_{jli} = \begin{cases} 1 & \text{if district } l \text{ is assigned to DC } j \text{ for the products of plant } i \\ 0 & \text{otherwise} \end{cases}$$

- Variables stating if secondary transport links are selected:

$$a_{jl} = \begin{cases} 1 & \text{if district } l \text{ is assigned to DC } j \text{ for at least one product} \\ 0 & \text{otherwise} \end{cases}$$

x_{jli} and a_{jl} are defined only if $R(p, j, l) <= CD \, \forall p$.

- Variables stating if primary transport links are selected:

$$z_{ij} = \begin{cases} 1 & \text{if the route from plant } i \text{ to DC } j \text{ is selected} \\ 0 & \text{otherwise} \end{cases}$$

5.3.3 MIP formulation

In this subsection, we present a MIP formulation of the multi–period distribution network design problem with dynamic assignments. The formulation of the problem using static assignments can be simply obtained by replacing variables x_{pjli}, a_{pjl} and z_{pij} respectively by variables x_{jli}, a_{jl} and z_{ij}. Moreover, it is worth pointing out that in this case, minimum and maximum volume constraints have all the same to be met in each time–period.

Minimize:

$$\begin{aligned} Total\ cost = & \sum_{i\in\mathbf{I}}\sum_{j\in\mathbf{J}}\frac{PTC_{ij}}{W_i}\sum_{p\in\mathbf{P}}\sum_{\substack{l\in\mathbf{L}\\ R(p,j,l)\leq CD}} x_{pjli}D_{pli} && (Primary\ transport\ cost)\\ & + \sum_{p\in\mathbf{P}}\sum_{j\in\mathbf{J}}\sum_{\substack{l\in\mathbf{L}\\ R(p,j,l)\leq CD}}\frac{STC_{pjl}}{W}\sum_{i\in\mathbf{I}} x_{pjli}D_{pli} && (Secondary\ transport\ cost)\\ & + \sum_{j\in\mathbf{J}} TC_j \sum_{p\in\mathbf{P}}\sum_{\substack{l\in\mathbf{L}\\ R(p,j,l)\leq CD}}\sum_{i\in\mathbf{I}} x_{pjli}D_{pli} && (Transit\ cost)\\ & + \sum_{i\in\mathbf{I}}\sum_{j\in\mathbf{J}} PI_{ij}\sum_{p\in\mathbf{P}} v''_{pij} && (Low\ volume\ penalties\ for\ primary\ transport) \end{aligned}$$

The objective function consists in the total distribution cost (primary and secondary transport costs, transit cost and penalties for violating minimum volume constraints on primary transport links). Each of the transport and transit cost components are computed as the unit cost per product multiplied by the volume going through each link/DC per period.

Subject to:

$$\sum_{\substack{j\in\mathbf{J}\\R(p,j,l)\leq CD}} x_{pjli} = 1 \qquad \forall p\in\mathbf{P}, i\in\mathbf{I}, l\in\mathbf{L}; D_{pli}\geq 0 \tag{5.1}$$

$$\sum_{\substack{l\in\mathbf{L}\\R(p,j,l)\leq CD}} x_{pjli}D_{pli} = v'_{pij} - v''_{pij} \qquad \forall p\in\mathbf{P}, i\in\mathbf{I}, j\in\mathbf{J} \tag{5.2}$$

$$v'_{pij} \geq V_{min}(i)z_{pij} \qquad \forall p\in\mathbf{P}, i\in\mathbf{I}, j\in\mathbf{J} \tag{5.3}$$

$$v'_{pij} \leq Mz_{pij} \qquad \forall p\in\mathbf{P}, i\in\mathbf{I}, j\in\mathbf{J} \tag{5.4}$$

$$v''_{pij} \leq V_{min}(i)z_{pij} \qquad \forall p\in\mathbf{P}, i\in\mathbf{I}, j\in\mathbf{J} \tag{5.5}$$

$$\sum_{i\in\mathbf{I}}\sum_{\substack{l\in\mathbf{L}\\R(p,j,l)\leq CD}} x_{pjli}D_{pli} \geq minVol_j y_j \qquad \forall p\in\mathbf{P}, j\in\mathbf{J} \tag{5.6}$$

$$\sum_{i\in\mathbf{I}}\sum_{\substack{l\in\mathbf{L}\\R(p,j,l)\leq CD}} x_{pjli}D_{qi} \leq maxVol_j y_j \qquad \forall p\in\mathbf{P}, j\in\mathbf{J} \tag{5.7}$$

$$\sum_{i\in\mathbf{I}} x_{pjli}D_{pli} \geq V_{min}(p,j,l)a_{pjl} \qquad \forall p\in\mathbf{P}, j\in\mathbf{J}, l\in\mathbf{L}; R(p,j,l)\leq CD \tag{5.8}$$

$$\sum_{i\in\mathbf{I}} x_{pjli} \leq Ia_{pjl} \qquad \forall p\in\mathbf{P}, j\in\mathbf{J}, l\in\mathbf{L}; R(p,j,l)\leq CD \tag{5.9}$$

$$z_{pij} \leq y_j \qquad \forall p\in\mathbf{P}, i\in\mathbf{I}, j\in\mathbf{J} \tag{5.10}$$

$$y_j,\ x_{pjli},\ a_{pjl},\ z_{pij} \in \{0,1\} \qquad \forall p\in\mathbf{P}, i\in\mathbf{I}, j\in\mathbf{J}, l\in\mathbf{L}; R(p,j,l)\leq CD \tag{5.11}$$

$$v'_{pij},\ v''_{pij} \geq 0 \qquad \forall p\in\mathbf{P}, i\in\mathbf{I}, j\in\mathbf{J} \tag{5.12}$$

Constraints (5.1) state that the demand of district l for the products of plant i is satisfied in each time–period p and is routed through a single DC (as x are binary variables).

Constraints (5.2) stipulate that the total volume going from plant i to DC j in time–period p is expressed as a difference between the continuous variables v'_{pij} and v''_{pij}. Constraints (5.3)-(5.5) ensure that:

- If the link $[ij]$ is selected in time–period p ($z_{pij} = 1$) then $v'_{pij} \geq V_{min}(i)$, $v'_{pij} \leq M$ and $v''_{pij} \leq V_{min}(i)$.
- If the link $[ij]$ is not selected in time–period p ($z_{pij} = 0$) then $v'_{pij} = 0$ and $v''_{pij} = 0$.

This enables us to compute the violation of minimum volume constraints on primary transport links and to penalize it in the objective function.

Constraints (5.6) state that if DC j is selected ($y_j = 1$), then the flows going through j in each time–period p have to be greater than the corresponding minimum volume limit. Constraints (5.7) stipulate that:

- If DC j is selected ($y_j = 1$) then the flows going through j in each time–period p must not exceed the corresponding maximum capacity.

- If DC j is not selected ($y_j = 0$) then there are no flows transiting by it (all the x_{pjli} have to be set equal to 0).

Constraints (5.8) ensure that if the secondary transport link between DC j and district l is selected in time–period p ($a_{pjl} = 1$) then the corresponding total volume has to be greater than the minimum volume $V_{min}(p, j, l)$. Constraints (5.9) stipulate that if the link between j and l is not selected in time–period p ($a_{pjl} = 0$) then all of the variables x_{pjli} have to be set equal to 0. Constraints (5.10) state that if DC j is not opened ($y_j = 0$) then all of the variables z_{pij} have to be set equal to 0. Constraints (5.11) and (5.12) are the integrality and non negativity constraints.

5.4 Application to a case–study in the automotive industry

In this section, we propose to test the multi–period distribution network design problem based on aggregation approach AG2 (MIP previously defined in §5.3). We examine the results of the computational experiments carried out using the case–study of Renault car distribution in France, detailed in §4.1.1. The country is divided into 92 districts where the company distributes its cars and each of the 448 car dealers is assigned to a district. Concerning the planning horizon, we consider one year divided into 4 time–periods (quarters) expressing the seasonality of demand for cars. The demand variation factor F_p for each season p was estimated using historical data and led to the following figures:
$F_1 = 1.1$
$F_2 = 1.2$
$F_3 = 0.7$
$F_4 = 1$
Thus, demand for district l, for the products of plant i and for period p is computed as $D_{pli} = F_p \frac{D_{li}}{P}$ where $P = 4$ is the number of periods in the year and D_{li} is the total yearly demand of district l for the products of plant i.

In the sequel, we aim first at discussing the managerial insights related to the optimal solutions of the multi–period distribution network design problem using either static or dynamic assignment decisions. Then, as numerical experiments showed computational difficulties, we propose a heuristic solution approach that we validate through a comparison with the optimal approach.

We employed the C++ language to implement the program and the commercial solver ILOG Cplex version 12.5 to solve MIP models. We carried out all the tests on a PC Intel Core(TM) i5–3210M (2.5 GHz) with 8 Gb of RAM, running under Windows 7.

5.4.1 Analysis of the optimization output using a commercial solver

In the present section, we intend to analyze the optimal solution provided by the commercial solver Cplex on the reference dataset of our case–study. First, we examine the main trade–offs to be achieved in the problem due to demand fluctuation. Then, we compare the results given by the multi–period network design problem with those given by the single–period one, with a focus on network structure. Finally, we study the differences in terms of costs, location decisions and computation times between the two alternatives that can be used in the multi–period model, namely static assignments and dynamic assignments.

5.4.1.1 Main trade–off analysis

Due to several operational features taken into account, the network design problem under study is highly constrained. As previously discussed in chapter 3 §3.2.6 and illustrated in Fig. 3.2, there is a trade–off to be achieved between potentially conflicting constraints, namely minimum volume constraints on primary transport links/ minimum throughput on DCs on the one side and maximum covering distances/maximum capacities on DCs on the other side. The first set of constraints forces indeed the solution process to decrease the number of opened DCs whereas the second set drives it to reduce the DC–district distances and thus to open many DCs. Experiments of the single–period DNDMVD problem with typical data and parameters of our case–study showed that there are frequent "infeasibility" situations due to the conflict between minimum volume constraints on primary transport links and maximum covering distance constraints. This is why we chose to use penalties when modeling the first type of constraints, which leads to Fig. 5.5 illustrating the remaining trade–offs between constraints (considering only the strict ones).

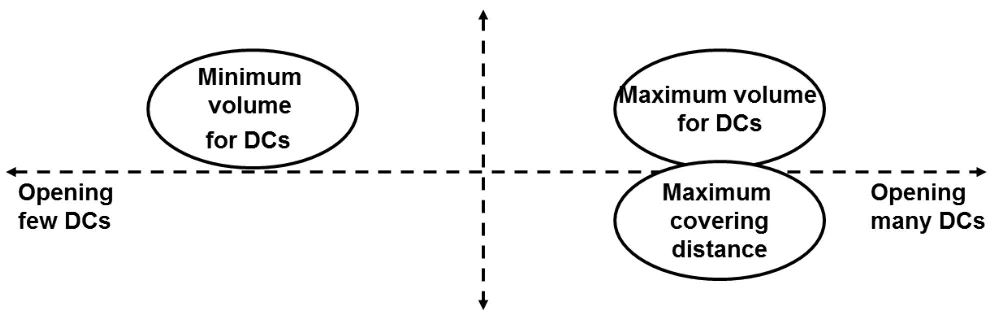


Fig. 5.5 Impact of the strict constraints of the problem on the number of DCs to be opened

The first tests of the multi–period extension of the DNDMVD problem showed that the remaining conflict between constraints can still lead to "infeasibilities". This is particularly due to low demand in some periods. For instance, in our case–study, as there is a fall of demand volumes in period 3, it is no more possible to ensure the minimum required volume per DC while keeping DCs close to customers. We thus identified two alternatives to overcome these feasibility issues:

1. First alternative: we relax minimum throughput constraints on DCs by modeling them using penalties. This strategy will be called feasibility strategy 1, it should favor "proximity to customers". Similarly to minimum volume constraints on primary transport links, we set the penalty of violating minimum throughput constraints to the unit transit cost which leads to Fig.5.6 illustrating the total transit cost per DC as a function of the total throughput. The figure shows that if DC j is opened, then, the company pays a fixed cost equal to $TC_j * minVol_j$ whatever the effective throughput volume. If the flow is greater than the minimum required volume $minVol_j$, then, an additional unit cost equal to TC_j is paid for each additional unit of product going through DC j.

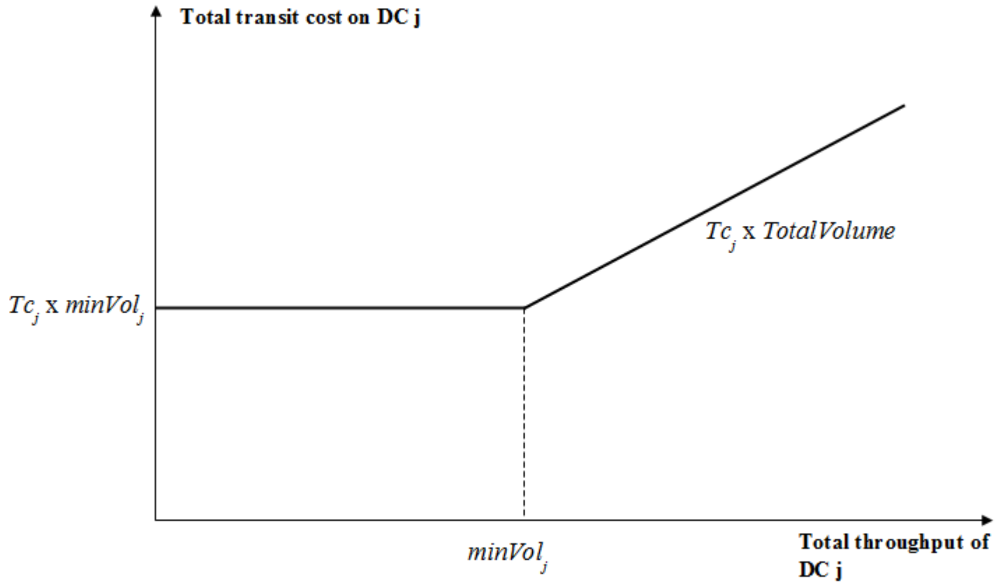


Fig. 5.6 Total transit cost as a function of throughput for a given DC j when using unit transit cost as penalty for violating minimum volume constraints on DC j.

Feasibility strategy 1 involves also some changes in the mathematical formulation of the problem defined in §5.3:

- Parameters: One additional parameter PI_j has to be added to the ones presented in 5.3.1. This is indeed the low volume penalty amount for the throughput of DC j.
- Decision variables: Two continuous variables r' and r'' have to be added in order to express the violation of minimum volume constraints on DCs.
- Constraints: Constraints 5.6 and 5.7 of the model described in §5.3.3 have to be re-

placed by the following constraints:

$$\sum_{\substack{l\in\mathbf{L}\\ R(p,j,l)\leq CD}} \sum_{i\in\mathbf{I}} x_{pjli}D_{pli} = r'_{pj} - r''_{pj} \qquad \forall p\in\mathbf{P}, j\in\mathbf{J} \tag{5.13}$$

$$r'_{pj} \geq minVol_j y_j \qquad \forall p\in\mathbf{P}, j\in\mathbf{J} \tag{5.14}$$

$$r'_{pj} \leq maxVol_j y_j \qquad \forall p\in\mathbf{P}, j\in\mathbf{J} \tag{5.15}$$

$$r''_{pj} \leq minVol_j y_j \qquad \forall p\in\mathbf{P}, j\in\mathbf{J} \tag{5.16}$$

Constraints (5.13) stipulate that the total volume going through DC j in time–period p is expressed as a difference between the continuous variables r'_{pj} and r''_{pj}. Constraints (5.14)-(5.16) state that:

– If DC j is selected ($y_j = 1$) then $r'_{pj} \geq minVol_j$, $r'_{pj} \leq maxVol_j$ and $r''_{pj} \leq minVol_j$.

– If DC j is not selected ($y_j = 0$) then $r'_{pj} = 0$ and $r''_{pj} = 0$.

This enables us to compute the violation of minimum volume constraints on DCs and to penalize it in the objective function.

- Objective function: The following term related to the penalties of violating minimum throughput constraints has to be added to the objective function used in §5.3.3: $\sum_{j\in\mathbf{J}} PI_j \sum_{p\in\mathbf{P}} r''_{pj}$

2. Second alternative: we relax the maximum covering distance and maximum capacity per DC constraints by increasing the values of the corresponding parameters. This strategy will be called feasibility strategy 2, it should lead to a "reduction in the DC number". We do not propose to relax them using penalties for two reasons. First, it is very difficult in our case–study to estimate penalties that make sense at the operational level. Second, if we remove the maximum covering distance constraints, then, the number of binary assignment variables will significantly increase and lead to considerable increase in computation times.

Conclusion 1: *A first contribution of considering a multi–period model is to realize that there are seasons of low demand volume where it is impossible to simultaneously meet all constraints. The decision–maker has thus to fix his preferences according to the company context and to prioritize either "proximity to customers" (feasibility strategy 1) or the "reduction in the DC number" (feasibility strategy 2). In the next section, we consider the two strategies when comparing the results of multi–period and single–period models.*

5.4.1.2 Multi–period vs. single–period models

We focus in this section on a computational comparison between a single–period model and a multi–period model based on the aggregation approach previously described. The main objective of these experiments is to gain a better understanding of the impact of demand seasonality on the network structure and to identify whether it is worth using a more complex multi–period model.

To carry out this comparison, we use static assignments, i.e. each district is assigned to the same DC in each time–period. This means that the two models tested here consider static decisions (locations and assignments) but the only difference is that the multi–period one takes into account seasonal demand and constraint satisfaction in each time–period. Feasibility issues are dealt with using one of the feasibility strategies mentioned in §5.4.1.1. With feasibility strategy 1 we keep the same values of the reference dataset for the maximum covering distance (460 kilometres) and for maximum capacities. With feasibility strategy 2, the maximum covering distance is set to 680 kilometres and maximum capacities on DCs are multiplied by four as compared to the reference dataset defined in §4.1.3.1. We run four tests using Cplex 12.5 with a final optimality gap of Cplex limited to 0.2% as proving exact optimality is very time–consuming.

Comparison of the results obtained with single–period and multi–period models could involve both the network structure (opened DCs) and the costs. However, as explained in appendix C, the fact that the considered models are highly constrained prevent us from carrying out an accurate cost analysis. We thus focus in what follows on the network structure (results are summarized in Table. 5.2). We observe that with both feasibility strategies, the network structure resulting from the multi–period model is different from the one resulting from the single–period model.

Instance	Feasibility strategy	Number of periods	Opened DCs
F1_1P	1	1	28 opened
F1_4P_S	1	4	30 opened (as compared to F1_1P: 2 added and 2 changed)
F2_1P	2	1	26 opened
F2_4P_S	2	4	21 opened (as compared to F2_1P: 5 removed and 1 changed)

Table. 5.2 Comparison of the network structure resulting from the multi–period model to the one resulting from the single–period model when using static assignments

When using feasibility strategy 1, we note that the single-period model (F1_1P) leads to opening 28 DCs whereas the multi-period model (F1_4P_S) leads to opening 30DCs. We carried out a detailed analysis to better understand this difference:

- We first tried to evaluate the cost of the network configuration resulting from the multi-period model (30DCs) in a single-period context: we found that this configuration remains feasible but leads to an increase of 0.19% in the total cost as compared to the configuration resulting from the single-period model (28DCs).

- We also tried to evaluate the cost of the network configuration resulting from the single-period model (28DCs) in a multi-period context and found that this configuration was not feasible due to the violation of the maximum capacity constraints for some DCs in the high demand periods 1 and 2. This can be explained by the fact that the single-period model uses an aggregate representation of these constraints (namely, the yearly volume should be less than a maximum yearly limit) whereas the multi–period model uses a more detailed representation (the quarterly volume should be less than the maximum yearly limit divided by 4).

With feasibility strategy 2, the difference in terms of network structure is much more noticeable as the multi–period version with static assignments has to meet strict minimum throughput constraints for all opened DCs in all time–periods while keeping the same network configuration and flows over the whole planning horizon. In order to meet minimum volume constraints in low–demand periods (namely period 3), it is thus necessary to reduce the number of opened DCs. This might explain why only 21 facilities are opened in this case. The single–period version does not capture demand variation and thus is not required to satisfy operational constraints in each time–period, which leads to opening more DCs (26).

Conclusion 2: *The second contribution of considering a multi–period model is to build a distribution network capturing demand seasonality. This is more noticeable with feasibility strategy 2 as strict minimum throughput constraints have to be met in each time–period. If the decision–maker prefers to use feasibility strategy 1, he can simply apply the single–period model as it leads to results close to those given by the multi–period model (favoring "proximity to customers"). This is why in the sequel of our tests (exact and heuristic), we only consider feasibility strategy 2 as in this case the necessity of considering a multi–period model is more apparent.*

5.4.1.3 Multi–period model: static vs. dynamic assignments

Using static assignment decisions might lead to a different network structure (opened DCs) and additional costs as compared to a situation where assignments can be dynamically adapted to demand seasonality. This is why we propose in this section to compare the results of a multi–period model featuring dynamic assignment decisions with those of a multi–period model featuring static assignment decisions (both based on feasibility strategy 2).

We first run two tests using the reference dataset and Cplex 12.5 with a final optimality gap limited to 0.2% (as proving exact optimality is very time–consuming). The obtained results are summarized in Table. 5.3.

We first note that using static assignment decisions leads to an increase in costs (column 4, line 2 in Table. 5.3). This is explained by the fact that the static assignment model (F2_4P_S) is more

Instance	Dynamic (D) or static (S) assignments	Number of opened DCs	Additional cost of using static assignments (%)	CPU (hours)
F2_4P_D	D	25	-	84.6
F2_4P_S	S	21	1.01	9.4

Table. 5.3 Comparison of the results given by the multi–period model using dynamic assignments with those given by the multi–period model using static assignments for the reference dataset (both based on feasibility strategy 2)

constrained than the dynamic assignment model (F2_4P_D). Nevertheless, this increase does not seem significant as it amounts only to 1%. Moreover, using static assignments forces the algorithm to reduce the number of opened DCs in order to meet minimum volume constraints in low–demand periods (hence opening only 21 DCs) whereas using dynamic assignments provides more flexibility to adapt flows to demand variation. This is why more DCs are opened in test instance F2_4P_D in order to get closer to customers and decrease delivery costs.

To confirm these results, we carry out further tests using instances C (varying annual demand per customer) and D (varying the list of potential DCs) defined in chapter 4, §4.1.3 (we use only the 5 first instances in each case). As heavy computation times are needed to carry out each test, especially for the version with dynamic assignments (more than 3 days as shown in Table. 5.3), we limit the optimality gap of Cplex to 0.5% and we use the solution of the problem featuring static assignments as initial solution of the problem featuring dynamic assignments. Computational results are summarized in Table. 5.4.

Table. 5.4 shows that using static assignments leads to opening fewer DCs than using dynamic assignments and that the related additional cost is less than 1%. These observations are in line with those made for the reference dataset (Table. 5.3). Furthermore, figures of Table. 5.3 and Table. 5.4 show considerable running times for multi–period versions, even when using static assignments (in instances F2_4P_S, C.2 and C.5, CPU exceeds 9 hours). Higher computation times are obviously shown by versions using dynamic assignments (exceeding 3 days in test instance F2_4P_D) as it involves a larger number of variables (assignment decisions for each period). These results motivate the development of a heuristic procedure in order to obtain good quality solutions within shorter computation times.

Conclusion 3: *The various experiments carried out in this section demonstrated that the additional cost to be incurred by the company when deciding to keep static assignment decisions in a multi–period context is not very significant. Using static assignments appears thus to be the best alternative at the operational level as it simplifies day–to–day operations (dealing with the same distribution flows over the whole planning horizon and managing fewer distribution centres) while ensuring a low increase in costs. In view of these observations, we propose in the next section*

Instance	Dynamic (D) or static (S) assignments	Number of opened DCs	Additional cost of using static assignments (%)	CPU (hours)
C.1	D	23	-	1.7
C.1	S	20	0.78	1.3
C.2	D	23	-	0.6
C.2	S	22	0.98	13.4
C.3	D	23	-	0.2
C.3	S	20	0.75	3.3
C.4	D	22	-	0.3
C.4	S	21	0.58	1.5
C.5	D	22	-	7.3
C.5	S	20	0.86	13.1
D.1	D	21	-	9.4
D.1	S	19	0.81	2.2
D.2	D	20	-	6.7
D.2	S	19	0.72	3.5
D.3	D	21	-	25.6
D.3	S	19	0.77	0.9
D.4	D	22	-	0.1
D.4	S	19	0.72	1.1
D.5	D	21	-	0.1
D.5	S	20	0.63	2.9

Table. 5.4 Comparison of the results given by the multi–period model using dynamic assignments with those given by the multi–period model using static assignments for various instances

to improve computation times for the multi–period problem featuring static assignments. We thus study a heuristic solution approach based on linear–relaxation.

5.4.2 A heuristic procedure for the multi–period extension of the DND-MVD problem

In order to reduce running times for the multi–period distribution network design problem featuring static assignments, we intend in this section to develop a heuristic solution approach. As the structure of the mixed integer program (MIP) of the multi–period problem using feasibility strategy 2 (defined in §5.3) is similar to the single–period DNDMVD MIP described in chapter 3 §3.3, we propose to adapt the heuristic approach S2M3 defined through Algorithm. 4. We first describe the heuristic method then we study its performance through a comparison with the straight application of Cplex to the original MIP.

5.4.2.1 Description of the heuristic method

As mentioned in chapter 4 §4.2.1, a first location stage has to be carried out in order to fix the values of variables y. We propose thus to solve a linear relaxation of the original MIP where the integrality constraints on variables x, z and a have been removed whereas they have been kept for y variables. In the second stage (assignment decisions), we focus on deciding which DCs should serve which districts, i.e. we aim at determining the value of each assignment variable x. We propose to apply a dynamic fixing strategy (MP–S2M3) to fix as many assignment variables as possible before reintroducing the integrality constraints for the currently free variables and solving the resulting MIP with a commercial solver. As discussed in chapter 4 §4.2.1.3, the fixing strategy should enable us to decrease the number of binary variables before solving the final MIP and thus to reduce computation times in the second stage of the heuristic procedure. The fixing strategy MP–S2M3 adapted to a multi–period model with static assignment decisions is described in Algorithm. 5.

Algorithm 5. Fixing strategy (MP–S2M3) adapted to the multi–period distribution network design model with static assignments (feasibility strategy 2)

Data:

V_{pjl}: Volume going through the secondary transport link $[jl]$ during the time-period p;

$\forall p \in \mathbf{P}\ \forall j \in \mathbf{J}\ \forall l \in \mathbf{L}$ *such as* $R(p,j,l) \leq CD \quad V_{pjl} = \sum_{i\in\mathbf{I}} x_{pjli} D_{pli}$;

V_{pj}: Volume going through the DC j during the time–period p;

$\forall p \in \mathbf{P} \quad \forall j \in \mathbf{J} \quad V_{pj} = \sum_{\substack{l\in\mathbf{L} \\ R(p,j,l)\leq CD}} \sum_{i\in\mathbf{I}} x_{pjli} D_{pli}$;

Algorithm:

Step 1

while *there are possible changes* **do**
- **for** *each opened transport link [jl] from DC j to district l* **do**
  - **if** $\exists$ *period p such as* $V_{pjl} < V_{min}(p,j,l)$ **then**
    - **if** $V_{p'j} - V_{p'jl} \geq minVol_j\ \forall p' \in \mathbf{P}$ **then**
      - **if** $\exists$ *DC k within the covering distance CD of l such as* $(V_{p'k} + V_{p'jl}) \leq maxVol_k\ \forall p' \in \mathbf{P}$ **then**
        - Add the constraint $V_{pjl} = 0\ \forall p \in \mathbf{P}$ (i.e. $x_{jli} = 0\ \forall i$)
      - **end**
    - **end**
  - **end**
- **end**
- Solve the resulting linear problem;
- **if** *infeasible problem* **then**
  - Cancel the lately added fixing constraints, break while
- **end**

end

Step 2

Reintroduce the integrality constraints for all the free variables then solve the resulting MIP;

The main difference that should be pointed out here as compared to the single–period version of the heuristic S2M3 is that checking the constraint satisfaction before setting variables x to 0 has to be made for all time–periods. This is explained by the fact that using static assignments forces the solution to meet operational constraints in all time–periods while keeping the same distribution flows for the whole planning horizon.

5.4.2.2 Numerical experiments

In order to study the performance of the heuristic MP–S2M3 above–described, we carry out 40 tests randomly varying either the annual demand per customer (instances C defined in 4.1.3.4) or the list of potential DCs (instances D defined in 4.1.3.5). For each test, we provide the computation time needed by a commercial solver to find an optimal solution to the original MIP as well as the computation time of the heuristic method and its quality.

All the linear and mixed integer programs (including those used in the heuristic) are solved with Cplex, limiting the MIP optimality gap to 0.2%. As for previous tests in the present work, we used a PC Intel Core(TM) i5–3210M (2.5 GHz) with 8 Gb of RAM, running under Windows 7. The obtained results are summarized in Table. 5.5 and Table. 5.6.

Instance	CPU(hours) of Cplex applied to the original MIP	CPU(hours) of the heuristic MP–S2M3	Quality of MP–S2M3 (%)
C.1	1.26	0.04	1.26
C.2	13.36	0.01	1.55
C.3	3.26	0.01	1.52
C.4	1.47	0.02	1.44
C.5	13.12	0.02	1.58
C.6	2.97	0.01	2.2
C.7	2.71	0.01	2.1
C.8	4.15	0.01	1.65
C.9	1.07	0.01	1.76
C.10	14.6	0.05	1.85
C.11	3.67	0.01	1.65
C.12	12.92	0.03	1.45
C.13	1.49	0.02	1.42
C.14	1.29	0.06	1.66
C.15	4.41	0.01	1.79
C.16	1.44	0.02	1.89
C.17	3.61	0.01	1.56
C.18	8.89	0.01	1.88
C.19	5.88	0.01	2.48
C.20	6.43	0.02	1.54

Table. 5.5 CPU and solution quality of the heuristic procedure MP–S2M3 applied to the multi–period problem with static assignments when varying customer demand (test instances C). Cplex optimality gap is fixed at 0.2%. Quality is measured as the relative difference between the solution provided by the heuristic MP–S2M3 and the solution given by Cplex applied to the original MIP.

The numerical results show that the heuristic approach significantly reduces computation times from several hours to few seconds while providing near–optimal solutions. The heuristic quality varies indeed between 0.82% and 2.48% with an average of 1.64%.

Instance	CPU(hours) of Cplex applied to the original MIP	CPU(hours) of the heuristic MP–S2M3	Quality of MP–S2M3
D.1	1.44	0.02	1.13
D.2	3.46	0.02	1.92
D.3	0.86	<0.01	1.49
D.4	1.13	0.02	1.84
D.5	2.85	<0.01	1.58
D.6	4.40	0.01	1.27
D.7	5.08	0.01	0.82
D.8	2.58	0.01	1.83
D.9	6.43	0.01	1.78
D.10	OOM	0.01	-
D.11	5.33	0.01	1.82
D.12	2.06	0.01	1.55
D.13	4.01	0.01	1.14
D.14	3.42	0.01	1.73
D.15	5.19	0.02	1.38
D.16	1.83	0.01	1.89
D.17	5.29	0.02	1.59
D.18	6.05	0.01	1.46
D.19	1.19	0.02	1.59
D.20	0.89	0.02	1.75

Table. 5.6 CPU and solution quality of the heuristic procedure MP–S2M3 applied to the multi–period problem with static assignments when varying the list of potential DCs (test instances D). Cplex optimality gap is fixed at 0.2%. OOM denotes running out of memory before achieving a gap equal to 0.2%. Quality is measured as the relative difference between the heuristic solution MP–S2M3 and the solution given by Cplex applied to the original MIP.

Conclusion

In this chapter, we studied a multi–period distribution network design problem involving several operational features. We analyzed the various modeling considerations and proposed an original way of introducing routing based on a dynamic clustering procedure. We assumed that location decisions are static over time–periods whereas assignments of districts to distribution centres can be either static or dynamic.

We carried out various tests based on our case–study in the field of car distribution, which lead to the following conclusions:

- There are seasons of low volume where it is impossible to simultaneously meet all constraints. The decision–maker has thus to relax some of the operational constraints and fix his preferences according to the company context: either "proximity to customers" (feasibility strategy 1) or the "reduction in the DC number" (feasibility strategy 2).

- Multi–period modeling allow building a distribution network which captures demand seasonality. This is more noticeable with feasibility strategy 2 as strict minimum throughput constraints have to be met in each time–period, which results in opening fewer distribution centres.
- The additional cost to be incurred by the company when deciding to keep static assignment decisions in a multi–period context is not very significant. Thus, using static assignments is the best alternative at the operational level as it allows dealing with the same distribution flows over the whole planning horizon and managing fewer distribution centres while ensuring a low increase in costs.

As far as computation times are considered, we studied them for the multi–period model featuring static assignment decisions. As obtaining exact solutions took several hours, we proposed a heuristic procedure (MP–S2M3) which was compared with the straight application of a state–of–the–art MIP solver. MP–S2M3 provided good quality solutions (about 1.6%) within short computation times (a few seconds). It involves the same contributions than the single–period version S2M3 presented in chapter 4 in addition to the fact that it is applied in a multi–period context.

Finally, the proposed model can be applied to a distribution network design problem with seasonal demand but can also be used in a context of uncertain demand with discrete scenarios having each a fixed probability of occurrence. Our approach is indeed very similar to the one employed in two–stage stochastic location problems where location decisions are made in the first stage and assignment decisions occur after random parameters become known. However, as it is highly constrained, our model does not always lead to a feasible solution even when using feasibility strategies 1 or 2. Both strategies reduce indeed the conflict between constraints but do not guarantee the feasibility of the problem whatever the value of demand and whatever the problem parameters. Infeasibility could indeed occur even if there is no conflict between constraints. For instance, when the maximum covering distance constraint is very low, we could find some district which is situated very far from all DCs and then its demand could never be met. This is why we do not achieve the complete recourse property needed in stochastic programming.

Chapter 6

Conclusion and future research

6.1 Conclusion

In the present work, we dealt with a multi–product distribution network design problem where the main objective was to determine the location of distribution centres and to assign customers to them while minimizing total costs.

- From an academic point of view, our first contribution relate to the combination of many operational features based on realistic assumptions in the field of car distribution. These features and in particular minimum volume, maximum capacity, maximum covering distance and single sourcing constraints can also be relevant in other contexts of application. Furthermore, the minimum volume constraints that we introduced concern both transport flows and distribution centre throughputs, which was not the case of the previous works taking into account this kind of constraints.

  Our second contribution is the implementation of a multi–period distribution network design problem including a location–routing approach in addition to the operational features above mentioned. In fact, due to the complexity of multi–period location–routing problems, there are few works dealing with this subject in the literature. In our study, we proposed a dynamic pre–processing clustering approach in order to estimate the length of transport routes from DCs to customers without increasing the complexity of the problem.

  Finally, our third contribution consists in studying exact and heuristic solutions through extensive numerical experiments based on real–life data from the automotive industry. The problem under study, either in its static or in its multi–period versions, proved to be difficult to solve when the number of integer variables and constraints increases or when tight constraints make it difficult to quickly find optimal solutions. We thus proposed several heuristics based on various linear relaxations of the original MIP formulation of the problem to determine location and assignment variables (both sets of variables being required to be binary).

- From an industrial point of view, our main contribution is to propose an optimization tool that can be used in the tactical/strategic planning of the supply chain. The main advantage of this tool is to be easy–to–use by logistics managers. On the contrary of complex commercial software that, trying to be as generic as possible, become very difficult to manipulate, our tool is simple to use and understand. Moreover, it can be employed as it is or slightly adapted to address many kinds of case–studies. For instance, in the context of our work, we studied the distribution network in France but other countries or groups of countries can be considered. We also compared a detailed approach based on a cluster representation of customers to the standard approach using districts (see appendix. B). Finally, the implemented tool can be used to analyze the impact of introducing direct flows (from plants to car dealers) on costs and on truck loading, to compare transport by train to transport by truck on primary transport links or to include additional linear costs like taxes related to carbon emissions.

  Furthermore, the various studies that we carried out showed that the problem is highly constrained. The interaction of numerous constraints, in particular when the corresponding parameter values are very tight, frequently led to "infeasibilities". Thus, the decision–maker should be less demanding as it is usually possible to relax some constraints using violation penalties instead of strict restrictions. This is the approach that we proposed to implement for minimum volume constraints on plant–DC transport links. In reality, some constraints are anyway violated (which explains delays in product deliveries for instance) but in a "perfect world" illustrated through the mathematical model, the decision–maker does not accept it.

  Logistics managers have also to be aware that the use of advanced planning system requires to efficiently manage data like site/customer locations, geographical coordinates, distances, costs, etc. In fact, one of the difficulties that we faced when analyzing our case–study is data collection. Some information were scattered over several databases and files which required a great manual effort to consolidate them. Therefore, the main axis of improvement in the context of industrial applications is the automation of data generation using appropriate databases. This can help motivating decision makers to use advanced planning systems as one of the heaviest parts in an optimization study is data collection.

6.2 Future research

We identified many outlooks that may be highly interesting for future research but that we could not study due to a lack of data or an absence of information in our industrial application. These are the features that we could not consider due to this context:

- Including inventory management in our model: From the beginning of the present study, we assumed that we only deal with build–to–order (BTO) products. Thus, no inventories are managed at distribution centres (DCs) but products are held for a few days of transit on DCs before being sent to their final destinations determined by car dealers. This situation does not mean that all products are build–to–order but that build–to–stock (BTS) inventories are individually managed by car dealers. Each car dealer can in fact make its own inventory control decisions and choose a storage compound that may be different from the distribution centre used by the car manufacturer. This decentralized policy makes it difficult to build a location–inventory model where inventories are stored on distribution centres and managed according to a same strategy.

- Including lead time minimization in the objective function: In addition to cost minimization which is a major performance driver for every company, customer satisfaction is also one of the key success factors. It is thus important to ensure quick deliveries and to reduce lead times. This can be achieved by introducing the minimization of time parameter in the objective function of our model either by using multi–objective programming or by assigning appropriate costs to the non–satisfaction of pre–determined lead times. The first alternative is difficult to implement in our case as managers are not used to employ optimization software and multi–objective tools are not easy to handle and require a great motivation and maturity from the user. The second alternative is also difficult to implement due to the lack of data. In fact, the penalties of delaying deliveries can be evaluated based on two main components: the storage cost of waiting products and the cost of loosing sales, i.e. the impact of increasing waiting times on the willingness of customers to buy cars. Nevertheless, it is very hard to estimate the second component and in our case–study, data was not available.

- Using accurate cost functions expressing economies of scale for transport and transit: In the present work, economies of scale were taken into account through two methods: implicit modeling based on minimum volume constraints (transit and secondary transport) and explicit modeling using simple cost functions (primary transport). Another alternative could be to use more accurate functions expressing cost depending on the handled quantities. For instance, it would have been interesting to consider a piecewise linear function for the cost of transit through distribution centres. Unfortunately, this was not possible in our case–study because data was not available. This is why introducing minimum volume constraints makes the model easy to use by practitioners as it requires less data and mathematical expertise than if accurate functions were needed.

Furthermore, we determined other challenging options that are worth being considered for future research.

- Introducing stochastic demand: As mentioned in chapter 5, the multi–period distribution network design model that we proposed can also be used in a context of uncertain demand with discrete scenarios having each a fixed probability of occurrence. Our approach is indeed very similar to the one employed in two–stage stochastic location problems where location decisions are made at the first stage and assignment decisions occur after random parameters become known. However, to achieve the complete recourse property needed in stochastic programming, it is necessary to change the model in order to ensure feasibility whatever the input data. We have thus to use penalties instead of strict constraints.

- Implementing an international model: One of the limitations of our model is to consider only three levels in the supply chain network, namely plants, distribution centres and final customers. This structure limits the geographic scope of the study as only one transport mode can be used to move products from plants to distribution centres. Using a combined transport involves indeed many transshipments and thus the introduction of many transit points in the distribution network. Therefore, one interesting research direction could be to consider an international model implicating several supply chain levels that may spread over many countries. In this case, the focus will not be on modeling detailed costs and constraints but on the logistical features related to international flows like custom duties, regulatory aspects, using different transport modes, etc. The consideration of minimum volume constraints makes also sense in an international context as the use of high capacity modes like trains and vessels requires the consolidation of flows.
- Optimizing backload management: Backloads are loads transported on the return journey of a delivery truck in order to avoid empty kilometres. This kind of transport is difficult to manage for cars as specific trucks are used in car transport. Backloads are arranged by carriers either by contacting competitors or by considering other flows for the same car manufacturer. In our model, we used a pre–determined distance related to empty kilometres but a specific tool was recently implemented to help logistics managers in minimizing empty kilometres given a fixed network structure and product flows. It would thus be possible to take advantage of this tool in order to analyze the impact of backload management on location and assignment decisions. One of the possibilities would be to study sequential iterations between our distribution network design tool and the backload optimization tool.
- Applying the model to case–studies in other fields than car distribution: this can present new challenges in the pre–processing clustering step if the number of customers significantly increases or the maximum cardinality of a cluster is much greater than three. In this case, applying a MIP solver to the clustering problem based on a set–partitioning formulation will lead to extensive computation times and using complete enumeration to compute optimal traveling salesman (TSP) tours in secondary transport will not be possible. Thus, one interesting research direction could be to implement efficient heuristics for these two problems. TSP is a widely studied problem in the literature, we can refer among others to

[Laporte, 1992] for a literature review on exact and heuristic methods to solve it.

Clustering of demand points into groups can be related to many real–life applications like transport, telecommunication networks, irrigation systems, etc. As this kind of problem becomes intractable for large–size instances, many authors proposed heuristic approaches to solve it. A first alternative could be to study the literature related to set–covering heuristics, [Umetani and Yagiura, 2007] review a number of approximate algorithms including linear relaxation based heuristics, Lagrangian heuristics, construction algorithms, meta–heuristics, etc.

Another close problem that may be interesting to investigate in our context is the cycle cover problem[1] (CCP). For instance, [Labbe et al., 1998] propose a heuristic for the edge CCP exploiting the similarity with the Chinese Postman Problem[2] (CPP). [Hochbaum and Olinick, 2001] consider k–cycle covering for edges, k is the maximum cardinality per cycle. They formulate the problem as a set–covering one and implement a heuristic method based on two steps: the first step consists in efficiently generating a set of candidate cycles but not necessarily the optimal one. The second step is to solve the set–covering integer program either by applying an exact branch and bound procedure or by rounding the solution of the linear–relaxation of the problem.

However, works on cycle covering do not deal with minimum volume or capacity constraints on each group of demand points. This is why it may be interesting to look at another close problem: the construction of SONET[3] (Synchronous Optical Network) networks in the field of telecommunications. In this context, we found some works in the literature that focus on heuristic approaches such as Tabu search [Laguna, 1994], greedy heuristic with performance guarantee [Brauner and Lemaire, 2002], linear–time approximation algorithm [Goldschmidt et al., 2003] and simulated annealing [Sutter et al., 1998] .

[1] A vertex (respectively edge) cycle cover of a graph G is defined as a set of cycles which are sub–graphs of G and contain all vertices (respectively edges) of G.

[2] The Chinese Postman Problem consists in determining the shortest tour such that each edge is traversed at least once.

[3] For SONET networks, the first problem to cope with is to partition the edges of the demand graph into sub–graphs called rings. Due to a limited capacity of treatment, rings are constrained to a maximum capacity.

Appendix A

Historical overview of the automotive industry

At the beginning of the 20^{th} century, cars were only designed for rich people but Henry Ford was determined to build a popular car, affordable by an average american worker. Hence, were created in 1913 the assembly line and the mass production concept, which consists in producing large quantities of standardized products. At the end of the second world war, Toyoda Kiichiro, president of Toyoda Motor Company, said "Catch up with America in three years. Otherwise, the automobile industry of Japan will not survive". The Japanese started then a continuous learning process from the american automotive industry, which led to the famous Toyota production system (TPS), developed by the mechanical engineer Taiichi Ohno. This system altered the "classical" logic of mass production by producing a great variety of cars in low volumes at a competitive cost. TPS was mainly based on just in time (JIT): "in a comprehensive industry such as automobile manufacturing, the best way to work would be to have all the parts for assembly at the side of the line just in time for their user" [Ohno, 1978]. Another important feature of the toyota production system was waste reduction to decrease costs. Ohno defined wastes in seven fields: overproduction, time on hand, transportation, processing, stock on hand, movement and making defective products. The japanese JIT approach is based on a "pull" philosophy, which means that products are not manufactured unless there are customers needing them. In other words, customer demand at the end of the chain pulls the products towards the market. This is also called the build–to–order (BTO) strategy and is opposed to the traditional "push" or build–to–stock (BTS) strategy, which consists in manufacturing products in anticipation of demand and thus in building inventories (see Fig. A.1).

The Toyota management way revolutionized the automotive industry and made other car makers rethink their manufacturing and supply chain strategies. Many of them have been indeed working on developing BTO strategies mainly to reduce inventory levels and to provide customized

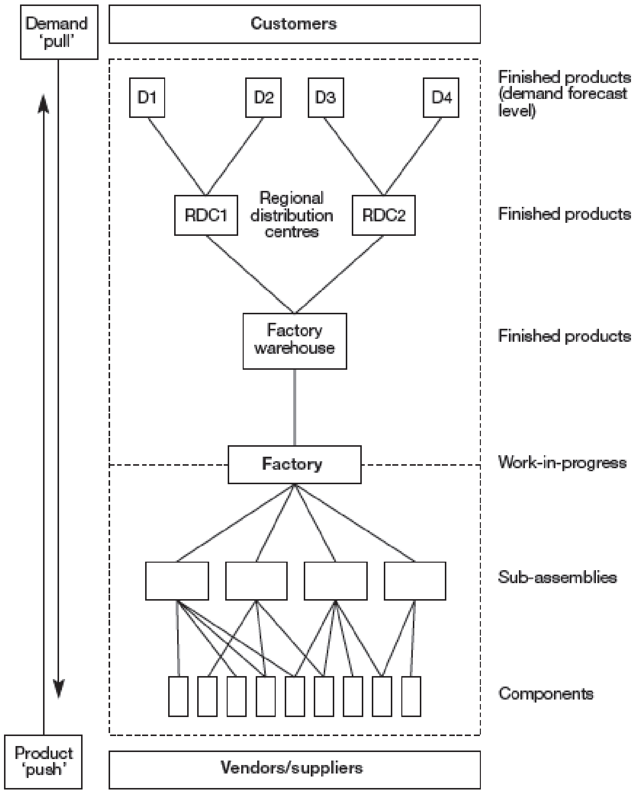


Fig. A.1 Push versus pull in the supply chain. Source [Christopher, 2005]

cars based on the specific requirements of each customer. The transition to a customer–driven production process was very complex for the auto companies that tried to implement it, not only because the automotive supply chain is very complex but also because the objectives they fixed were very optimistic. Most of them aimed, as a way to attract the public to customized products, at drastically reducing the order–to–delivery time while maintaining the reliability of the customer promised deadlines. The target time varied between 10 days (BMW) to 14 days (Renault–Nissan) and 21 days (Volvo). For Renault–Nissan, the objective of 14 days fixed in 1999 with the "Projet Nouvelle Distribution (PND)" was revised to 21 days in 2002 because supply chain operations could not cope with the lead time reduction [Saint-Seine, 2002]. [Klenau, 2005] presented a comparative study over 13 companies that implemented BTO strategies in the computer and automotive industry and concluded that the BTO transformation in the latter case was much more complex and led to disappointing financial results especially in terms of inventory reduction. This is why car makers are continuously trying to find the best balance between BTO and BTS proportions and to develop a more flexible supply chain network .

Appendix B

Comparison of two aggregate approaches based on districts with the detailed approach based on clusters

B.1 Introduction

In this appendix, we study two alternative methods using districts as an aggregate representation of customers. The first method (AG1) is a rather standard one employing distances from distribution centres to district barycentres whereas the second method (AG2) applies a more sophisticated clustering approach per district (AG2 is discussed into detail in chapter 5, §5.2.1.2). We compare them with the detailed approach studied in chapters 2 and 3 in order to show that:

- It is better to apply the detailed approach whenever possible.
- When we have to apply an aggregate approach (like in chapter 5), it is better to choose method AG2.

B.2 Review of the three tested versions

Here are the main features of the three approaches (which are also summarized in Table B.1):

1. Detailed approach based on clusters

   This approach is presented in chapters 2 and 3. Basically, it consists in applying a pre–processing clustering on the whole distribution region in order to group customers into clusters. Then, in the mathematical formulation of the problem, clusters are used as final customers. Detailed distances from DCs to clusters are used to compute secondary transport costs and minimum volume constraints are expressed for each DC–cluster link.

2. Aggregate approach AG1: using districts as final customers and calculating costs based on distances from DCs to district barycentres

   In most of the literature and consulting studies, a standard consists in using barycentre to represent a customer zone. In our study, we are given, for each customer, its demand and its distance to each DC. Then, we estimate the distance from a DC to a district barycentre as the weighted average of the distances from the DC to the customers belonging to the concerned district, using demands as weights. The secondary transport cost from a DC to a district is computed using the truck cost formula defined in §3.2.3 and the round–trip distance from the DC to the barycentre. Notice that in our case, we will use a demand variation factor per time–period that will be applied to all customer demands. Consequently, weights of customers do not change over periods and for a given district, the barycentre will be the same for the whole planning horizon. This is why in the aggregate model using AG1, DC–district delivery distances and costs are constant over time–periods and the impact of demand variation on the length of delivery routes is not taken into account. As far as the secondary transport minimum volume constraints are concerned, we propose to consider the quantity ensuring on average full truckload transport from DCs to districts within the allowed waiting time.

3. Aggregate approach AG2: using districts as final customers and calculating costs based on a clustering per district

   This approach is discussed into detail in chapter 5, §5.2.1.2. It consists in applying a pre–processing clustering per DC/district/period triple in order to better estimate secondary transport costs. In the mathematical formulation of the problem, districts are used as final customers but the unit cost from a given DC to a given district is estimated as the weighted average cost of serving the clusters of the district. As far as the secondary transport minimum volume constraints are concerned, we propose to consider the quantity ensuring on average full truckload transport within the allowed waiting time multiplied by the number of clusters in the district.

B.3 Numerical experiments

B.3.0.3 Main features of the study

We propose here to compare the three approaches described above. In the three cases, we consider a single–period model and we use the reference dataset defined in §4.1.3 as well as test

Model	Final customers	Secondary transport costs	Minimum volume per secondary transport route
Detailed	Clusters	Cost of optimal routes from DCs to clusters	Volume ensuring in average full truckload transport within the allowed waiting time
AG1	Districts	Cost of round–trip distances from DCs to district barycentres	Volume ensuring in average full truckload transport within the allowed waiting time
AG2	Districts	Weighted average costs of optimal routes from DCs to the clusters belonging to each district with cluster demands as weights	Number of clusters in the district for the period $*$ Volume ensuring in average full truckload transport within the allowed waiting time

Table. B.1 Comparison between the detailed model and the aggregate models based on AG1 and AG2

instances C (varying demand per customer presented also in §4.1.3). An important idea to point out here is that the cost of a solution to the aggregate problem (either using AG1 or using AG2) cannot immediately be compared with the cost of a solution of the detailed problem, since the first solution may not satisfy all constraints at the operational level. In fact, in the aggregate models, unit costs and constraints are expressed using transport links from DC to districts. Thus, it will not necessarily be possible at the operational planning level to construct transport routes from DCs to car dealers meeting the restrictions of maximum covering distance, minimum transport volume and maximum number of customers. For that reason, we use the aggregate model to set the location and assignment decisions but have to check "a posteriori" whether the transport to car dealers could really meet all the constraints. Therefore, after solving the aggregate network design problem, we apply "a posteriori" clustering procedure using a set–partitioning formulation for each opened DC with the aim to determine the secondary transport routes and their effective cost. We summarize in Table. B.2 the main parameters of customer clustering and network optimization for the three modeling options. Other details concerning the expression of secondary transport costs and minimum volume constraints were already illustrated in Table. B.1.

Table. B.2 shows that "a posteriori" clustering is also applied to the detailed model. This is due to the fact that the pre–processing clustering of the detailed model overestimates the secondary transport costs. In fact, the minimum volume per cluster in the pre–processing clustering is two full

Model	**Pre–processing clustering**				**Network optimization**				**"A posteriori" clustering**			
	Scope	Min volume per cluster (nb truck/week)	Max number of customers	Max route	Max covering distance (KM)	Min volume for primary transport (nb truck/week)	Penalty for violating primary transport min volume	Min volume for secondary transport (nb truck/week)	Scope	Min volume per cluster (nb truck/week)	Max number of customers	Max route (KM)
Detailed	Whole distribution region	2	3	-	460	1	Unit transport cost	1	Customers assigned to each DC	1	3	460
AG2	DC/District pair	1	3	-	460	1	Unit transport cost	Depends on the number of clusters in each district	Customers assigned to each DC	1	3	460
AG1	-	-	-	-	460	1	Unit transport cost	1	Customers assigned to each DC	1	3	460

Table. B.2 Main parameters of customer clustering and network optimization for the detailed model and the aggregate models AG1 and AG2

truckloads per week whereas the minimum required volume for transport is only one full truckload per week. As mentioned in chapter 3, a difference of 1 truckload between these two quantities gives more flexibility to the optimization algorithm with respect to the possibility of assigning a cluster to several DCs. Nevertheless, if some clusters are assigned to only one DC for all their products, then it is possible at the operational level to rebuild shorter delivery routes ensuring the minimum volume of transport which is one full truckload per week.

"A posteriori" clustering is applied to the detailed and aggregate models using the same features:

- Considering a clustering procedure for each opened DC and all the customers that it serves
- Ensuring the minimum volume of transport per cluster (1 truck/week)
- Allowing at most three customers per cluster

- Constructing delivery routes shorter than the maximum covering distance of 460 kilometres

B.3.0.4 Numerical comparison

We propose to test the three modeling options: detailed model, aggregate model AG1 and aggregate model AG2 and to compare the resulting costs, location decisions and constraint satisfaction at the operational level (i.e. feasibility of the "a posteriori" clustering procedure). We use the parameters defined in Table. B.2 to run the numerical experiments for the reference dataset and test instance C varying demand per customer presented in §4.1.3. The main results obtained for aggregate approaches AG1 and AG2 as compared to the detailed approach are summarized in Table. B.3

"A posteriori" clustering could be infeasible for some DCs in the aggregate models, we thus display the number of observed "infeasibilites" in columns 'Inf1' and 'Inf2' ('Inf1' denotes the number of DCs where the "a posteriori clustering" is infeasible due to the maximum route constraint and 'Inf2' denotes the number of DCs where the "a posteriori clustering" is infeasible due to the minimum volume and maximum number per cluster constraints). In column 'Diff', we show the number of differences between the location decisions made in the detailed model and the ones made in the aggregate model.

Table. B.3 shows that the number of "infeasibilities" in the "a posteriori" clustering step (sum of columns Inf1 and Inf2) is always greater than 0 for the two aggregate approaches. This means that the location and assignment decisions resulting from aggregate models do not satisfy the routing constraints at the operational level whereas the detailed model provides a solution complying with all the constraints. However, figures summarized in the table show that aggregate approach AG2 leads to less "infeasibilities" than aggregate approach AG1. This means that, by using aggregate approach AG2, we obtain a better representation of the routing aspect in the network optimization problem, leading to location and assignment decisions closer to the "optimal" ones. This is confirmed by the figures of column 'Diff2' which show that aggregate approach AG2 is closer than aggregate approach AG1 to the detailed model in terms of location decisions. As far as costs are concerned, it is very difficult to carry out a realistic comparison between the aggregate models and the detailed model. In fact, AG1 and AG2 do not meet the secondary transport constraints at the operational level in all the test instances ("a posteriori" clustering is infeasible for many DCs), hence the resulting cost is underestimated. This is why we chose not to compare these approaches in terms of costs.

B.4 Conclusion

As a conclusion, our numerical results show that using one of the two aggregate approaches leads to the violation of the detailed operational constraints related to route construction, namely

Model	**Aggregate model using approach AG1**			**Aggregate model using approach AG2**		
Instance	Inf1	Inf2	Diff	Inf1	Inf2	Diff
R	15	2	4	8	0	2
C.1	15	2	4	7	0	3
C.2	16	3	0	7	0	0
C.3	14	2	2	7	0	0
C.4	14	2	7	7	0	1
C.5	15	2	3	9	0	1
C.6	15	2	3	8	1	0
C.7	15	2	2	6	0	2
C.8	13	3	2	6	0	2
C.9	14	3	0	5	1	0
C.10	16	2	3	7	0	1
C.11	16	2	4	8	0	4
C.12	14	2	2	5	0	0
C.13	15	2	2	8	0	0
C.14	15	2	1	7	0	3
C.15	13	3	2	4	0	2
C.16	15	2	2	5	0	0
C.17	15	3	2	7	1	2
C.18	16	2	4	8	0	2
C.19	15	3	2	7	0	2
C.20	14	2	3	6	0	1

Table. B.3 Main results obtained with the aggregate models using AG1 and AG2 on test instance C varying customer demand. 'R' is the reference dataset instance. 'Inf1' denotes the number of DCs where the "a posteriori clustering" is infeasible due to the maximum route constraint. 'Inf2' denotes the number of DCs where the "a posteriori clustering" is infeasible due to the minimum volume and maximum number per cluster constraints. 'Diff' is the number of differences in location decisions between the aggregate model and the detailed model.

the minimum volume, the maximum number of customers and the maximum distance per route. Thus, it is better to apply the detailed approach whenever possible. However, if we have to use an aggregate approach, we should choose aggregate approach AG2 as it leads to a better approximation of the detailed problem than approach AG1. First, the solution given by AG2 presents a smaller number of violation of the detailed operational constraints. Second, thanks to a better estimation of secondary transport costs, this approach provides a network structure (opened DCs) closer to the one obtained with the detailed model.

Appendix C

Why is it difficult to carry out a cost analysis?

In the literature, there are very few works studying the economic impact of jointly modeling operational and strategic decisions or of using dynamic models instead of static ones. It could be however motivating to show the economic interest of introducing complexity in the mathematical models. This is why we tried in our work to compare the various versions we studied in terms of costs, namely to compare the aggregate models with the detailed one and the multi–period model with the single–period one. Nevertheless, it was not possible to achieve accurate economic results due to the fact that our models are highly constrained.

C.1 Aggregate model vs. detailed model

The comparison between the aggregate models (2 versions proposed) and the detailed one was discussed in appendix. B. The numerical experiments showed that when using aggregate approaches, it was not possible to construct (a posteriori) secondary transport routes meeting all the operational constraints (i.e. minimum volume, maximum number of customers and maximum distance per route). This is due to the fact that aggregate approaches roughly model costs and constraints (minimum volume and maximum covering distance) for secondary transport, using districts instead of detailed clusters. Thus it was not possible to study the costs resulting from the aggregate models as compared to those resulting from the detailed model because estimating the operational constraint violations appeared to be a major difficulty in the first case. We compared however the models on the basis of the resulting network structures (opened DCs).

C.2 Multi–period model vs. static model

In chapter 5, we carried out some numerical experiments (see §5.4.1) to compare the multi–period model with the single–period one. Through the analysis of the network structure (opened DCs), our aim was to study the robustness of the decisions given by the static model when demand varies over time–periods. It would also have been interesting to compare the resulting costs but this does not seem relevant. Namely, the solution of the single–period model violates some of the detailed period by period constraints of the multi–period model (see Table. C.1). A fair comparison would thus require to include in the compared models a numerical estimation of the financial penalties to be incurred whenever minimum volume and maximum covering distance constraints are violated. However, as obtaining an accurate estimation of these penalties was very difficult in our case, we chose not to compare the models in terms of costs.

Strict constraints	Feasibility strategy 1		Feasibility strategy 2		Consequence
	Multi–period	Single–period	Multi–period	Single–period	
Minimum volume per DC	-	-	Per period	On average	The solution of the single–period problem violates the constraint in some periods
Maximum covering distance	Considering the clustering result per period	Considering the average clustering result	Considering the clustering result per period	Considering the average clustering result	
Minimum volume on DC–district links (1 full truck-load * number of clusters of the district)	Considering the number of clusters per period	Considering the average number of clusters	Considering the number of clusters per period	Considering the average number of clusters	

Table. C.1 Strict constraints of the multi–period model as compared to the single–period one.

Appendix D

Variation of the clustering results according to distribution centres and time–periods

D.1 Introduction

In this appendix, we aim at analyzing the results of the pre–processing clustering applied to each district/DC/period triple when using aggregation approach AG2 (used in chapter 5).

We consider the case–study of Renault car distribution in France, detailed in §4.1.1. The country is divided into 92 districts where the company distributes its cars and each of the 448 car dealers is assigned to a district. Concerning the planning horizon, we consider one year divided into 4 time–periods (quarters) expressing the seasonality of demand for cars. The demand variation factor F_p for each season p was estimated using historical data and led to the following figures:
$F_1 = 1.1$
$F_2 = 1.2$
$F_3 = 0.7$
$F_4 = 1$
Thus, demand for district l, for the products of plant i and for period p is computed as $D_{pli} = F_p \frac{D_{li}}{P}$ where $P = 4$ is the number of periods in the year and D_{li} is the total yearly demand of district l for the products of plant i.

We set the clustering constraints using the parameters defined in Table. B.2 for AG2. First, we study cluster variation according to the DC delivering products, then, we examine their variation from period to period and finally, we show how average secondary transport costs change over periods.

D.2 Numerical experiments

D.2.1 Variation of the clustering results according to DCs

We consider here the solution of the clustering for district 51, comprising six customers, when delivered from different DCs in time–period 3. Fig. D.1 shows the clusters obtained in this district when served by DCs Auxerre, Batilly and Chaligny. DC Auxerre situated at the south–west of the district leads to three clusters of one customer and one cluster of three customers. DC Batilly situated at the east of the district leads to two clusters of one customer and two clusters of two customers. DC Chaligny situated at the south–east of the district leads to one cluster of one customer, one cluster of two customers and one cluster of three customers. The obtained results show thus that clusters of a given district could vary according to the DC serving it. This is the reason why we chose to iteratively apply the clustering procedure for each DC in the pre–processing step of AG2.

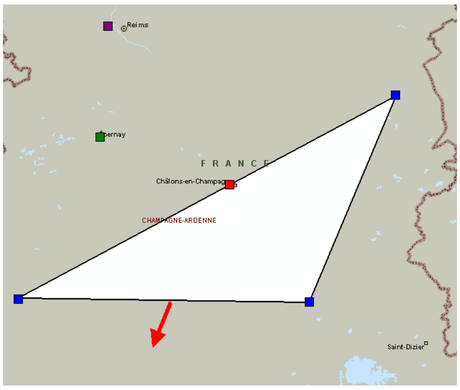
D.2.2 Variation of the clustering results according to time–periods

We study here the clustering output of district 1, comprising five customers, when delivered from DC Bourges in different time–periods (period 2 with $F_2 = 1.2$, period 3 with $F_3 = 0.7$ and period 4 with $F_4 = 1$). Fig. D.2 shows the resulting clusters according to time–periods. Period 2 having the highest volume leads to three clusters of one customer and one cluster of two customers. Period 4 with demand variation factor set to 1 leads to two clusters of two customers and one cluster of one customer. Period 3 having the lowest volume leads to one cluster of three customers and one cluster of two customers. This means that when demand gets lower, the average size of clusters gets higher in order to ensure full truckloads within maximum waiting time of 5 working days. The obtained results show thus that clusters of a given district could vary according to time–periods. This is why we chose to iterate applying the clustering procedure for each time–period.

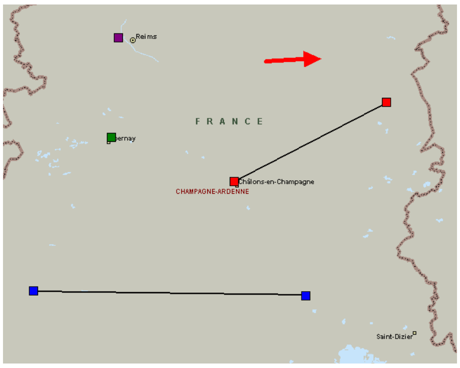
D.2.3 Variation of DC-district costs according to periods

A clustering procedure for each district/DC/period triple leads to different DC–district delivery costs according to periods. As shown in Fig. D.2, the average cluster size and thus the average delivery route length increases when demand decreases. Thus, we expect that secondary transport costs in period 3 ($F_3 = 0.7$) are greater than those in period 2 ($F_2 = 1.2$). To evaluate this rise, we illustrate in Fig. D.3 the average increase per DC and in Fig. D.4 the average increase per district.

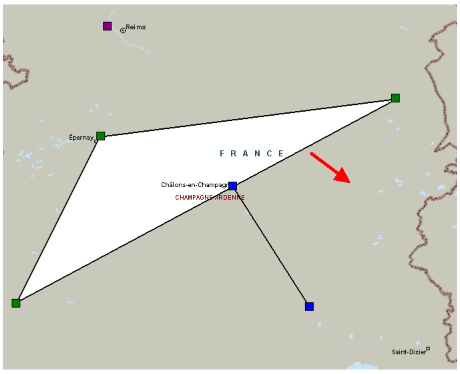
The average cost increase for a given DC (respectively a given district) is computed over all the DC–district transport links starting at this DC (respectively arriving to this district). Fig.



(a) Clusters of district 51 when delivered from DC Auxerre in period 3



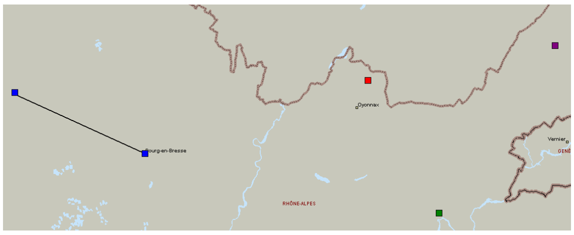
(b) Clusters of district 51 when delivered from DC Batilly in period 3



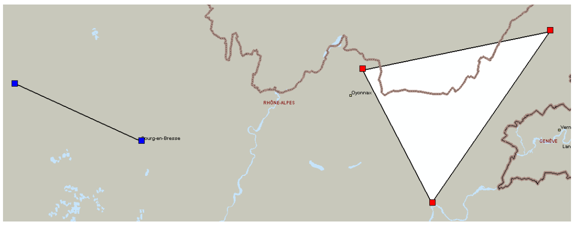
(c) Clusters of district 51 when delivered from DC Chaligny in period 3

Fig. D.1 Cluster variation in district 51 when delivered from DCs Auxerre, Batilly and Chaligny. The red arrow indicates the direction situating the DC.

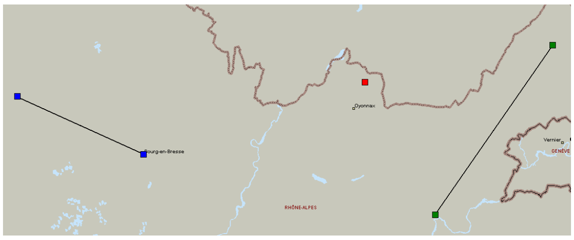
D.3 shows that the greater average cost increase per DC is about 5% for DC 6 (Bourges) but cost actually varies between 0% and 14%. Similarly, the greater average cost increase per district is



(a) Clusters of district 1 when delivered from DC Bourges in period 2 ($F_2 = 1.2$)



(b) Clusters of district 1 when delivered from DC Bourges in period 3 ($F_3 = 0.7$)



(c) Clusters of district 1 when delivered from DC Bourges in period 4 ($F_4 = 1$)

Fig. D.2 Cluster variation in district 1 when delivered from DC Bourges in different time–periods.

about 15% for the district indexed 74 but cost actually varies between 0% and 44%.

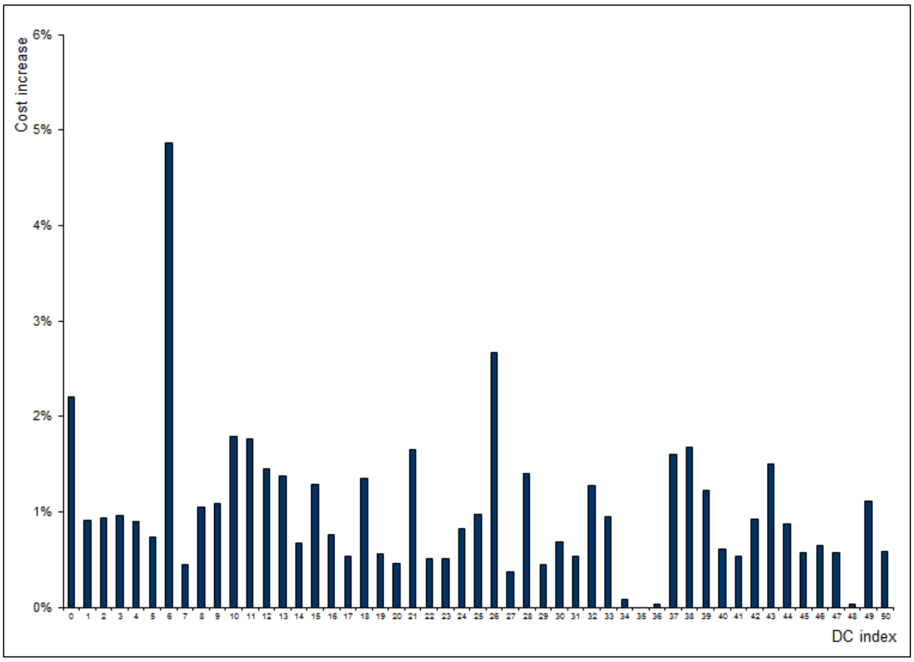


Fig. D.3 Average secondary transport cost increase per DC (from period 2, $F_2 = 1.2$, to period 3, $F_3 = 0.7$)

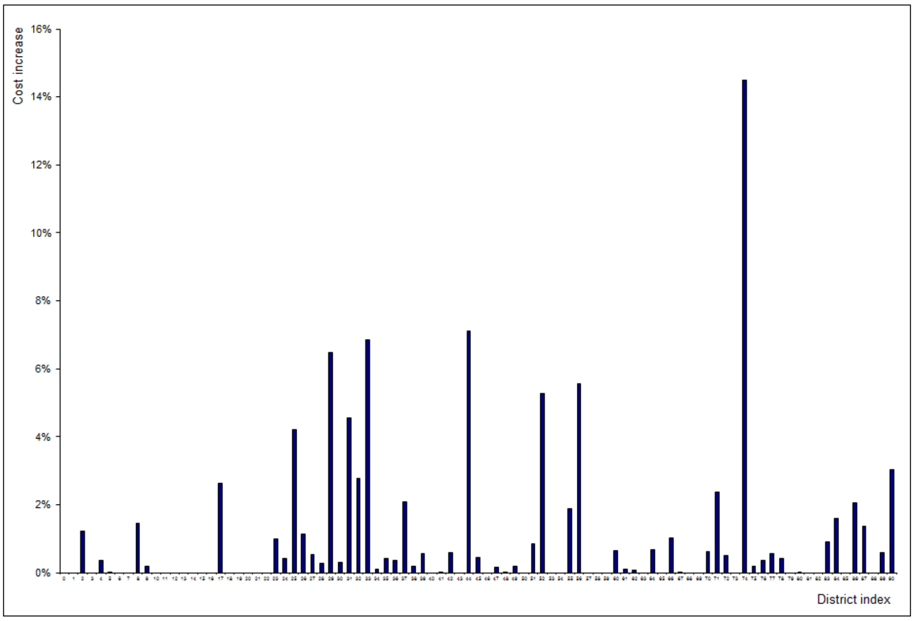


Fig. D.4 Average secondary transport cost increase per district (from period 2, $F_2 = 1.2$, to period 3, $F_3 = 0.7$)

List of Figures

List of Tables

List of scientific publications

Peer reviewed

Kchaou Boujelben, M., Gicquel, C. and Minoux, M. (2013) A distribution network design problem in the automotive industry: MIP formulation and heuristics. *In revision for publication in Computers and Operations Research.*

Conference proceedings

Kchaou Boujelben, M., Gicquel, C. and Minoux, M. (2012) A MILP model and heuristic approach for supply chain network design with minimum volume constraints. *International Conference on Industrial Engineering and Engineering Management, IEEM Hong Kong.*

Kchaou Boujelben, M., Gicquel, C. and Minoux, M. (2012) A supply chain network design problem under flow consolidation constraints. *9th International Conference of Modeling, Optimization and Simulation, MOSIM12, Bordeaux (France).*

Presentations

In English

Kchaou Boujelben, M., Gicquel, C. and Minoux, M. (2011) A supply chain network design problem under flow consolidation constraints. *Optimeo workshop, Orsay (France).*

In French

Kchaou Boujelben, M., Gicquel, C. and Minoux, M. (2013) Conception d'un réseau logistique avec contraintes de volume minimum : une heuristique basée sur la relaxation linéaire. *ROADEF Conference, Troyes (France).*

Kchaou Boujelben, M., Gicquel, C. and Minoux, M. (2012) Modélisation et résolution d'un problème de conception d'un réseau logistique avec contraintes de volume minimum. *STP Gdr Macs meeting, Clermont-Ferrand (France).*

Kchaou Boujelben, M., Gicquel, C. and Minoux, M. (2012) Un problème de conception de réseaux logistiques avec des contraintes de massification des flux. *ROADEF conference, Angers (France).*

Bibliography

[Afshar and Haghani, 2012] Afshar, A. and Haghani, A. (2012). Modeling integrated supply chain logistics in real–time large–scale disaster relief operations. *Socio–Economic Planning Sciences 46 pp. 327–338.*

[Albareda-Sambola et al., 2012] Albareda-Sambola, M., Fernandez, E., and Nickel, S. (2012). Multiperiod location–routing with decoupled time scales. *European Journal of Operational Research 217 pp. 248–258.*

[Albareda-Sambolaa et al., 2011] Albareda-Sambolaa, M., Fernandez, E., and Laporte, G. (2011). The capacity and distance constrained plant location problem. *Computers and Operations Research 36 pp. 597–611.*

[Albareda-Sambolaa et al., 2009] Albareda-Sambolaa, M., Fernandeza, E., Hinojosab, Y., and Puertoc, J. (2009). The multi–period incremental service facility location problem. *Computers and Operations Research 36 pp.1356–1375.*

[Alonso-Ayuso et al., 2003] Alonso-Ayuso, A., Escudero, L., Garin, A., Ortuno, M., and Perez, G. (2003). An approach for strategic supply chain planning under uncertainty based on stochastic 0–1 programming. *Journal of Global Optimization 26 pp.97-124.*

[Alumur et al., 2012] Alumur, S. A., Nickel, S., Saldanha-Da-Gama, F., and Verter, V. (2012). Multi–period reverse logistics network design. *European Journal of Operational Research 220 pp. 67-78.*

[Ambrosinoa et al., 2009] Ambrosinoa, D., Sciomachena, A., and Scutella, M. G. (2009). A heuristic based on multi–exchange techniques for a regional fleet assignment location–routing problem. *Computers and Operations Research 36 pp. 442–460.*

[Anthony, 1965] Anthony, R. (1965). Planning and control systems: A framework for analysis. *A framework for analysis, Harvard University, Boston.*

[Baldacci et al., 2008] Baldacci, R., Christofides, N., and Mingozzi, A. (2008). An exact algorithm for the vehicle routing problem based on the set partitioning formulation with additional cut. *Mathematical programming Ser.A 115 pp.351–385.*

[Barreto et al., 2007] Barreto, S., Ferreira, C., Paixa, J., and Santos, B. S. (2007). Using clustering analysis in a capacitated location-routing problem. *European Journal of Operational Research 179 pp. 968-977.*

[Barros et al., 1998] Barros, A., Dekker, R., and Scholten, V. (1998). A two-level network for recycling sand: a case study. *European Journal of Operational Research 110 pp. 199–214.*

[BatchGeocodeur, 2007] BatchGeocodeur, J. (2007). *http://www.batchgeocodeur.mapjmz.com/.*

[Beardwood et al., 1959] Beardwood, J., Halton, J. H., and Hammersley, J. M. (1959). The shortest path through many points. *Mathematical proceedings of the Cambridge Philosophical Society 55 No.4 pp.299–327.*

[Berger, 1997] Berger, R. T. (1997). Location–routing models for distribution system design. *Ph.D. Dissertation, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL.*

[Brauner and Lemaire, 2002] Brauner, N. and Lemaire, P. (2002). A set covering approach for sonet network design. *Les cahiers du laboratoire Leibnitz N62.*

[Canel and Khumawalaz, 2001] Canel, C. and Khumawalaz, B. M. (2001). International facilities location: A heuristic procedure for the dynamic uncapacitated problem. *International Journal of Production Research 39 N17 pp.3975–4000.*

[Carello et al., 2004] Carello, G., Croce, F. D., Ghirardi, M., and Tadei, R. (2004). Solving the hub location problem in telecommunication network design: A local search approach. *Networks 44 No.2 pp.94–105.*

[Christopher, 2005] Christopher, M. (2005). (book) logistics and supply chain management. creating value–adding networks. *Prentice Hall 3rd edition.*

[Church and ReVelle, 1974] Church, R. and ReVelle, C. (1974). The maximal covering location problem. *Papers of the Regional Science Association 32 pp. 101–118.*

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

[Crainic et al., 2009] Crainic, T. G., Fu, X., Gendreau, M., Rei, W., and Wallace, S. W. (2009). Progressive hedging–based meta–heuristics for stochastic network design. *CIRRELT–2009–03.*

[Daganzo, 1984] Daganzo, C. F. (1984). The length of tours in zones of different shapes. *Transportation Research Part B 18B No.2 pp.135–145.*

[Daskin et al., 2003] Daskin, M. S., Snyder, L. V., and Berger, R. T. (2003). Facility location in supply chain design. *(Book) Logistics Systems: Design and Optimization, Chapter 2.*

[Eiselt and Marianov, 2009] Eiselt, H. and Marianov, V. (2009). Gradual location set covering with service quality. *Socio–Economic Planning Sciences 43 pp. 121-130.*

[Elhedhli and Merrick, 2012] Elhedhli, S. and Merrick, R. (2012). Green supply chain network design to reduce carbon emissions. *Transportation Research Part D 17 pp. 370-379.*

[Escobar et al., 2013] Escobar, J. W., Linfati, R., and Toth, P. (2013). A two–phase hybrid heuristic algorithm for the capacitated location–routing problem. *Computers and Operations Research 40 pp. 70–79.*

[Farahani et al., 2012] Farahani, R. Z., Asgari, N., Heidari, N., Hosseininia, M., and Goh, M. (2012). Covering problems in facility location: a review. *Computers and Industrial Engineering 62 pp. 368-407.*

[Fleischmann et al., 2008] Fleischmann, B., Meyr, H., and Wagner, M. (2008). Advanced planning. *(Book) Supply Chain Management and Advanced Planning Concepts, Models, Software, and Case Studies, Springer–Verlag Berlin Heidelberg fourth edition.*

[Garey and Johnson, 1979] Garey, M. R. and Johnson, D. S. (1979). (book) computers and intractability: A guide to the theory of np–completeness. *Bell Telephone Laboratories, Incorporated.*

[Gebennini et al., 2009] Gebennini, E., Gamberini, R., and Manzini, R. (2009). An integrated production-distribution model for the dynamic location and allocation problem with safety stock optimization. *International Journal of Production Economics 122 pp.286-304.*

[Geoffrion and Graves, 1974] Geoffrion, A. and Graves, G. (1974). Multicommodity distribution system design by benders decomposition. *Management Science 20 No.5 pp. 822–844.*

[Ghaderi and Jabalameli, 2013] Ghaderi, A. and Jabalameli, M. S. (2013). Modeling the budget–constrained dynamic uncapacitated facility location-network design problem and solving it via two efficient heuristics: A case study of health care. *Article in press, Mathematical and computer modelling.*

[Goldschmidt et al., 2003] Goldschmidt, O., Hochbaum, D. S., Levin, A., and Olinick, E. V. (2003). The sonet edge-partition problem. *Networks 41N1 pp.13-23.*

[Guha et al., 2000] Guha, S., Meyerson, A., and Munagala, K. (2000). Hierarchical placement and network design problems. *FOCS00 Proceedings of the 41st annual symposium on foundations of computer science pp. 603.*

[Hochbaum and Olinick, 2001] Hochbaum, D. S. and Olinick, E. V. (2001). The bounded cycle cover problem. *INFORMS Journal on Computing 13 N2 pp.104–119.*

[Javid and Azad, 2010] Javid, A. A. and Azad, N. (2010). Incorporating location, routing and inventory decisions in supply chain network design. *Transportation Research Part E 46 pp. 582-597.*

[Karger and Minkoff, 2000] Karger, D. R. and Minkoff, M. (2000). Building steiner trees with incomplete global knowledge. *FOCS00 Proceedings of the 41st annual symposium on foundations of computer science pp. 613.*

[Kenyon and Morton, 2001] Kenyon, A. S. and Morton, D. P. (2001). A survey on stochastic location and routing problems. *CEJOR Physica–Verlag 9 pp.277–328.*

[Klenau, 2005] Klenau, S. (2005). The build-to-order transformation. an analysis of the financial impacts of build-to-order practices in the automobile and computer industries. *Working paper, Tennenbaum institute, Georgia institute of technology.*

[Krumke, 1995] Krumke, S. (1995). On a generalization of the p–center problem. *Information processing letters 56 pp. 67–71.*

[Krumke and Thielen, 2011] Krumke, S. O. and Thielen, C. (2011). Minimum cost flows with minimum quantities. *Information processing letters 111 pp. 533–537.*

[Labbe et al., 1998] Labbe, M., Laporte, G., and Soriano, P. (1998). Covering a graph with cycles. *Computers And Operations Research 25 N6 pp. 499–504.*

[Laguna, 1994] Laguna, M. (1994). Clustering for the design of sonet rings in interoffice telecommunications. *Management Science 40 N11 pp.1533–1541.*

[Langevin and Mbaraga, 1996] Langevin, A. and Mbaraga, P. (1996). Continuous approximation models in freight distribution: an overview. *Transportation Research Part E 30 No.3 pp.163–188.*

[Laporte, 1992] Laporte, G. (1992). The vehicle routing problem: An overview of exact and approximate algorithms. *European Journal of Operational Research 59 pp.345–358.*

[Laporte and Dejax, 1989] Laporte, G. and Dejax, P. J. (1989). Dynamic location-routing problems. *Journal of Operational Research society 40 N5 pp.471–482.*

[Lee and Dong, 2009] Lee, D.-H. and Dong, M. (2009). Dynamic network design for reverse logistics operations under uncertainty. *Transportation Research Part E 45 pp 61-71.*

[Lim et al., 2006] Lim, A., Fan, W., and Xu, Z. (2006). A transportation problem with minimum quantity commitments. *Transportation science 40 No. 1 pp. 117–129.*

[Lin et al., 2009] Lin, L., Mitsuo, G., and Xiaoguang, W. (2009). Integrated multistage logistics network design by using hybrid evolutionary algorithm. *Computers and Industrial Engineering 56 pp. 854-873.*

[Mehrjerdi and Nadizadeh, 2013] Mehrjerdi, Y. Z. and Nadizadeh, A. (2013). Using greedy glustering method to solve capacitated location–routing problem with fuzzy demands. *Accepted paper in European Journal of Operational Research 2013.*

[Melkote and Daskin, 2001] Melkote, S. and Daskin, M. S. (2001). An integrated model of facility location and transportation network design. *Transportation Research Part A 35 pp. 515–538.*

[Melo et al., 2005] Melo, M., Nickel, S., and da Gama, F. S. (2005). Dynamic multi-commodity capacitated facility location: a mathematical modeling framework for strategic supply chain planning. *Computers and Operations Research 33 pp. 181–208.*

[Melo et al., 2012] Melo, M., Nickel, S., and da Gama, F. S. (2012). A tabu search heuristic for redesigning a multi–echelon supply chain network over a planning horizon. *International Journal of Production Economics 136 pp.218-230.*

[Melo et al., 2009a] Melo, M., Nickel, S., and F.Saldanha-Da-Gama (2009a). Facility location and supply chain management: a review. *European Journal of Operational Research 196 pp.401–412.*

[Melo et al., 2009b] Melo, T., Nickel, S., and da Gama, F. S. (2009b). An lp-rounding heuristic to solve a multi-period facility relocation problem. *Berichte des Fraunhofer ITWM Nr. 168.*

[Meyerson, 2001] Meyerson, A. (2001). Profit-earning facility location. *STOC01 Crete Greece, July 6–8.*

[Miemczyk and Holweg, 2004] Miemczyk, J. and Holweg, M. (2004). Building cars to customer order - what does it mean for inbound logistics operations? *Journal of Business Logistics 25 N2 pp. 171–197.*

[Min et al., 1998] Min, H., Vaidyanathan, and Srivastava, R. (1998). Combined location–routing problems: A synthesis and future research directions. *European Journal of Operational Research 108 pp.1–15.*

[Miranda et al., 2009] Miranda, P. A., Garrido, R. A., and Ceroni, J. A. (2009). E–work based collaborative optimization approach for strategic logistic network design problem. *Computers and Industrial Engineering 57 pp. 3–13.*

[Moon and Chaudhry, 1984] Moon, I. D. and Chaudhry, S. S. (1984). An analysis of network location problems with distance constraints. *Management Science 30 N3 pp. 290–307.*

[Moon and Papayanopoulos, 1995] Moon, I. D. and Papayanopoulos, L. (1995). Facility location on a tree with maximum distance constraints. *Computers and Operations Research 22 N9 pp. 905–914.*

[Nagy and Salhi, 2008] Nagy, G. and Salhi, S. (2008). Location-routing: Issues, models and methods. *European Journal of Operational Research 117, pp. 649-672.*

[Ndiaye and Alfares, 2008] Ndiaye, M. and Alfares, H. (2008). Modeling healthcare facility location for moving population groups. *Computers and Operations Research 35 pp. 2154-2161.*

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

[Nozick and Turnquist, 1998] Nozick, L. K. and Turnquist, M. A. (1998). Integrating inventory impacts into a fixed-charge model for locating distribution centers. *Transportation Research Part E 34 No.3 pp.173–186.*

[Ohno, 1978] Ohno, T. (1978). (book) toyota production system. beyond large–scale production. *Diamond Inc. Tokyo Japan, english translation copyright 1988 by Productivity Inc.*

[Owen and Daskin, 1998] Owen, S. and Daskin, M. (1998). Strategic facility location: a review. *European Journal of Operational Research 111 N3 pp.423–447.*

[Porter, 1980] Porter, M. (1980). (Book) Competitive strategy. *Simon and Schuster, New York pp. 7.*

[Prodhon, 2011] Prodhon, C. (2011). A hybrid evolutionary algorithm for the periodic location–routing problem. *European Journal of Operational Research 210 N2 pp. 204–212.*

[Rahman and Smith, 2000] Rahman, S. and Smith, D. (2000). Use of location-allocation models in health service development planning in developing nations. *European Journal of Operational Research 123 N3 pp. 437–452.*

[ReVelle et al., 2008] ReVelle, C., Eiselt, H., and Daskin, M. (2008). A bibliography for some fundamental problem categories in discrete location science. *European Journal of Operational Research 184 pp.817–848.*

[Sabri and Beamon, 2000] Sabri, E. H. and Beamon, B. M. (2000). A multi-objective approach to simultaneous strategic and operational planning in supply chain design. *Omega 28 pp. 581–598.*

[Saez-Aguado and Trandafir, 2012] Saez-Aguado, J. and Trandafir, P. C. (2012). Some heuristic methods for solving p–median problems with a coverage constraint. *European Journal of Operational Research 220 pp. 320–327.*

[Saint-Seine, 2002] Saint-Seine, S. D. (2002). Renault quits build-to-order quest. *Automotive News Europe 7 N4.*

[Sajjadi, 2008] Sajjadi, S. R. (2008). Integrated supply chain: multi–products location routing problem integrated with inventory under stochastic demand. *Dissertation submitted to the department of industrial and manufacturing engineering and the faculty of the graduate school of Wichita State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.*

[Salhi and Rand, 1989] Salhi, S. and Rand, G. K. (1989). The effect of ignoring routes when locating depots. *European Journal of Operational Research 39 pp.150–156.*

[Seedig, 2010] Seedig, H. G. (2010). Network flow optimization with minimum quantities. *Proceedings of the international conference on operations research.*

[Shen and Qi, 2007] Shen, Z. and Qi, L. (2007). Incorporating inventory and routing costs in strategic location models. *European Journal of Operational Research 179 pp.372–389.*

[Shen et al., 2003] Shen, Z.-J. M., Coullard, C., and Daskin, M. S. (2003). A joint location-inventory model. *Transportation Science 37 No.1 pp.40–55.*

[Shen, 2007] Shen, Z. M. (2007). Integrated supply chain design models: A survey and future research directions. *Journal of Industrial and Management Optimization 3 N1 pp.1–27.*

[Snyder, 2006] Snyder, L. V. (2006). Facility location under uncertainty: A review. *IIE Transcations 38 N7 pp.537–554.*

[Sutter et al., 1998] Sutter, A., rancois Vanderbeck, and Wolsey, L. (1998). Optimal placement of add/drop multiplexers: heuristic and exact algorithms. *Operations Research 46 N5.*

[Svitkina, 2010] Svitkina, Z. (2010). Lower-bounded facility location. *ACM Transactions on algorithms (TALG) Volume 6 Issue 4.*

[Taillard, 1999] Taillard, E. (1999). A heuristic column generation method for the heterogeneous fleet vrp. *RAIRO 33 N1 pp.1–14.*

[Thanh et al., 2010] Thanh, P. N., Peton, O., and Bostel, N. (2010). A linear relaxation-based heuristic approach for logistics network design. *Computers and Industrial engineering 59 pp. 964–975.*

[Umetani and Yagiura, 2007] Umetani, S. and Yagiura, M. (2007). Relaxation heuristics for the set covering problem. *Journal of the Operations Research Society of Japan 50 N4 pp.350–375.*

[Wu et al., 2002] Wu, T.-H., , Low, C., and Bai, J.-W. (2002). Heuristic solutions to multi-depot location–routing problems. *Computers and Operations Research 29 pp. 1393–1415.*

[Xifeng et al., 2013] Xifeng, T., Ji, Z., and Peng, X. (2013). A multi–objective optimization model for sustainable logistics facility location. *Transportation Research Part D 22 pp. 45-48.*

[Yi and Ozdamar, 2007] Yi, W. and Ozdamar, L. (2007). A dynamic logistics coordination model for evacuation and support in disaster response activities. *European Journal of Operational Research 179 pp.1177–1193.*

[Yu et al., 2001] Yu, V. F., Lin, S.-W., Lee, W., and Ting, C.-J. (2001). A simulated annealing heuristic for the capacitated location–routing problem. *Computers and Operations Research 28 pp. 803-826.*

[Zarandi et al., 2013] Zarandi, M. H. F., Hemmati, A., Davari, S., and Turksen, I. B. (2013). Capacitated location-routing problem with time windows under uncertainty. *Knowledge–Based Systems 37 pp. 480–489.*

[Zhu et al., 2011] Zhu, X., Yuan, Q., Garcia-Diaz, A., and Dong, L. (2011). Minimal-cost network flow problems with variable lower bounds on arc flows. *Computers and Operations Research 38 pp. 1210–1218*.