



## Network design of a multi-period collaborative distribution system

Xin Tang, Fabien Lehuédé, Olivier Péton

### ► To cite this version:

Xin Tang, Fabien Lehuédé, Olivier Péton. Network design of a multi-period collaborative distribution system. International journal of machine learning and cybernetics, 2019, 10 (2), pp.279-290. 10.1007/s13042-017-0713-5 . hal-01591618

**HAL Id: hal-01591618**

**<https://hal.science/hal-01591618>**

Submitted on 14 Apr 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Network design of a multi-period collaborative distribution system

Xin TANG, Fabien LEHUÉDÉ, Olivier PÉTON

---

---

## Highlights

- Combination of full truckload and less-than truckload transportation cost.
- Direct shipments from suppliers to customers are possible.
- Additional constraints that improve the applicability of our model.
- Numerical experiments on a real-life case study.

# Network design of a multi-period collaborative distribution system

Xin TANG, Fabien LEHUÉDÉ, Olivier PÉTON

---

## Abstract

This paper deals with the design of a two-tiered distribution network in the context of the collaborative distribution of goods. It is based on a case study concerning the distribution of horticultural products in France. The logistics network includes a geographical cluster of collaborating suppliers, one consolidation facility near the suppliers, a set of Regional Distribution Centers (RDC) to be selected and thousands of customers spread over the whole country. Shipments from suppliers to RDCs have a full truckload rate structure. Shipments from RDCs to customers have a less-than-truckload rate structure. We propose a temporal and spatial aggregation of the data, which leads to a mixed integer linear programming formulation for optimizing the location of RDCs and the choice of distribution routes in this network. One strong feature of this model is that all logistics operations are outsourced to logistics service providers. Thus, all RDCs are facilities that are already running and there is no formal fixed cost associated with their selection. Decision makers have to find a trade-off solution that minimizes the logistics costs while keeping a reasonable complexity of the resulting distribution network.

*Keywords:* Facility location; network design; collaboration; distribution; mixed integer linear programming.

---

## 1. Introduction

This paper proposes a mixed integer linear programming formulation for the optimal location of intermediate facilities in a collaborative distribution network. It is motivated by a real-life case study concerning the distribution of horticultural products in Western France. This case study involves a cluster of four competing companies in the same geographical area, who decide to collaborate on the delivery of their goods to a large set of customers spread over the whole country. Since this delivery is paid by suppliers, Logistics Service Providers (LSP) and customers do not seek cost minimization in this delivery supply chain. The optimization relies on the efforts of suppliers only.

We consider horizontal collaboration in a geographical cluster of suppliers who are direct competitors on the same market. Horizontal collaboration is defined as a business agreement between two or more unrelated or competing organizations, that cooperate to share their private information or resource such as joint distribution centers between two

retailers (Simatupang & Sridharan, 2002). The partners aim at increasing productivity, e.g. by optimizing vehicle capacity utilization, reducing empty mileage and cutting costs of non-core/supporting activities to increase the competitiveness of their logistics networks (Cruijssen et al., 2007b).

In this study, the suppliers want to design a common distribution network, composed of one Consolidation and Distribution Center (CDC) located near the suppliers, a set of intermediate facilities called Regional Distribution Centers (RDC) and thousands of customers spread over the whole country. This network leads to a two-fold cost reduction opportunity. First, gathering together the goods of each collaborator can provide full truckload (FTL) shipments, which are much cheaper than individual less-than-truckload (LTL) shipments. Second, the cost structure of LTL shipments is generally concave, so that the consolidation of several small shipments brings some economies of scale.

The goal of this paper is to propose a mathematical model and a solution method for optimizing the location of RDCs. The contributions of this paper are twofold. First, we apply a fixed-charge multi-period facility location model. This model includes special characteristics, e.g. two transportation segments with distinct rate structure and the possibility of direct shipments from suppliers to customers. Second, it is an example of the applicability of operations research to a real-life problem. We propose and measure the impact of two additional constraints that improve the applicability of the mathematical model. The model is tested on a case study with one complete year of shipments and was used to help the collaborating suppliers make decisions about which facility should be selected.

The rest of this paper is organized as follows. In Section 2, we present several literature references that position our work in the field of supply chain network design. In Section 3, we detail the main settings of the optimization problem considered. In Section 4, we introduce the formulation of the MILP model. In Section 5, computational experiments on a real-life case study are presented. In Section 4.3, two classes of additional constraints are proposed. Finally, we draw conclusions in Section 6.

## 2. Literature review

### 2.1. Related network design problems

From a theoretical point of view, the design of collaborative distribution networks is an extension of the facility location problem, which has received considerable attention from academics and practitioners over the last decades (Melo et al., 2009; Laporte et al., 2015). In general, facility location problems deal with determining the optimal number, capacity, type, and geographic location of facilities while minimizing the network cost and satisfying customers demand. Facility location is often considered over a strategic planning horizon, generally at least several years.

The model proposed in this paper has several similarities with the well-known  $p$ -median problem, whose objective is to locate  $p$  facilities so that the weighted average distance between demand nodes and the nearest of the selected facilities is minimized (Daskin & Maass,

2015). One extension of the  $p$ -median problem to multi-layer logistics or transportation systems is the  $p$ -hub location problem (Campbell, 1994), which belongs to the broader class of hub location problems (Alumur & Kara, 2008; Contreras, 2015). In the  $p$ -hub location problem, the goal is to select exactly  $p$  candidate locations that are used as cross-docking facilities for the routing of transportation requests. More generally, in hub location problems, the number of hubs is not imposed but each candidate facility has a fixed set-up cost. Hub location problems have traditionally been applied to air transport, but other applications are also studied in the literature. For example, Campbell (2005) reviews the case of ground transport, which integrates LTL shipments and trips with multiple pickups or deliveries. Gelareh & Nickel (2011) propose an uncapacitated multiple allocation hub location problem tailored to urban transport and liner shipping network design.

Wu et al. (2015) propose a joint location-inventory model in a very similar network. A common characteristic with our work is that customers demand is considered over multiple periods, but location and RDC-customer allocation are static decisions. However, we adopt a more flexible distribution policy that allows direct shipments from the production zone to customers. Hence, each customer is delivered either from its allocated RDC or from PZ.

## 2.2. Building a collaborative network

Horizontal collaboration has raised much literature in recent years, but most papers focus on topics other than facility location. Among the main topics studied are the identification and selection of potential partners (Adenso-Díaz et al., 2014; Cruijssen et al., 2007b), building of trust within the coalition (Groothedde et al., 2005; Hingley et al., 2011; Leitner et al., 2011), collaborative information and communication technologies (Buijs & Wortmann, 2014; Cruijssen et al., 2007b; Mason et al., 2007), reduction of transportation costs and  $CO_2$  emissions (Leitner et al., 2011; Pan et al., 2013, 2014), optimization of load plans (Andersen et al., 2011; Crainic et al., 2014; Erera et al., 2013a,b) and cost/profit allocation (Audy et al., 2012; Dai & Chen, 2012; Vanovermeire & Sörensen, 2014).

As far as cost savings are concerned, various results have been published in the last fifteen years. Table 1 presents a sample of results presented in the supplier collaboration literature. As already mentioned, the cost savings are not the only goal of logistics collaboration, and the scope of collaboration and the method for calculating cost savings differ a lot from one study to another. Therefore, no direct conclusion should be drawn solely on cost savings. Cruijssen et al. (2007a) reports that cost savings are larger when order sizes are small compared to a standard truckload, which is our case.

## 2.3. Modeling transportation costs

One of the main incentives for horizontal collaboration is the LTL cost structure charged by the carriers. As detailed in Özkaya et al. (2010), this cost depends on both tangible and intangible factors. Examples of intangible factors are the degree of competition in the local markets, the desirability of the shipment, and the negotiating power of the shipper.

Reference	Case study / Application	Cost savings
Tyan et al. (2003)	Global 3PL (Taiwan)	6.7%
Z#apfel & Wasner (2002)	Parcel service provider (Austria)	on average 10%
Audy et al. (2011)	Furniture industry (Canada)	up to 12.9%
Pan et al. (2014)	Food industry (France)	13%
Frisk et al. (2010)	Forest industry (Sweden)	up to 14.2%
Leitner et al. (2011)	Automotive industry (Romania, Spain)	14%-15%
le Blanc et al. (2006)	Dry grocery (The Netherlands)	22% -26%

Table 1: Cost savings brought by shipper collaboration

Transportation costs are also influenced by inevitable empty truck repositioning (Ergun et al., 2007).

One classic approach is to consider a cost structure with incremental discounts, leading to a concave piecewise linear function. This approximation is practical for optimization purposes but sometimes not realistic. The Modified All-Unit Discount (MAUD) cost is based on weight intervals for which decreasing prices are incurred (Mui Ann Chan et al., 2002). For example, shipments of less than 30kg are charged \$ 30/kg and shipments of more than 45 kg are charged \$ 20 / kg. This leads to a non-continuous piecewise affine cost curve with breakpoints. In order to avoid these breakpoints and keep the price curve monotonously non-decreasing, the whole interval [30,45] is charged as 45kg for \$ 20/kg. Examples of MAUD cost structures can be found in Hill & Galbreth (2008), Özkaya et al. (2010) or Wu et al. (2015).

The combination of FTL and LTL cost structures has been used in papers that optimize transportation in a predefined network. Croxton et al. (2003) propose a merge-in-transit distribution system with four modes: small package, less-than-truckload, truckload, and transport by air. Lapierre et al. (2004) present a compound cost function merging the best use of each category of transport: parcel, LTL or FTL depending on the shipment weight. Hill & Galbreth (2008) minimize transportation and inventory costs in a two-tiered distribution system: FTL shipments are used from a single factory to a single warehouse, and LTL shipments are used from the warehouse to customers. Krajewska & Kopfer (2009) combine four types of costs corresponding to vehicles from their own fleet, vehicles from subcontractors paid on a tour basis, vehicles from subcontractors paid on a daily basis and requests forwarded to independent carrier. The last two categories are roughly similar to FTL and LTL. Lindsey et al. (2013) study a more general network in which each shipment can be sent individually (LTL) or possibly consolidated on an FTL route with other shipments.

#### 2.4. Position of our problem

As stated above, the combination of FTL and LTL cost structures has been used in papers that optimize transportation in predefined networks. To the best of our knowledge, it has never been considered in the field of facility location.

In this paper, we do not consider any fixed set-up cost for selecting a candidate facility. Thus, we assume that  $p$  candidate locations must be selected, but the value of  $p$  is not known *a priori*. Our problem differs from  $p$ -hub location models most collaborative delivery models because the transshipment of goods at the RDCs is not mandatory. Direct shipments from the production zone to customers may happen (i) when orders from one customer are very large, enabling FTL shipment, and (ii) when the sum of customer orders delivered from an RDC exceeds a multiple of full truckload by only a few units, directly shipping these units avoids resorting to an extra FTL route.

Our mathematical model integrates time periods in the description of the demand but the facility location decision variables are static. RDCs are cross-docking facilities in which storage is forbidden. Time periods correspond to independent shipping campaigns. RDCs may not receive material at all periods. Since there is no set-up cost at RDCs, selecting an RDC that does not receive material at every period is not penalized.

### 3. Problem settings

#### 3.1. Facilities

We consider a distribution network composed of three types of facilities, represented in Figure 1.

- The Production Zone (PZ) includes a set of neighboring suppliers and one Consolidation and Distribution Center (CDC). The suppliers are the collaborating companies. They are the source of all material flows in the network. The CDC is a collaborative warehouse at the suppliers gate. Customer orders are consolidated at the CDC by so-called supplier routes. Since supplier routes are relatively short, they have almost no impact on facility location decisions. The set of all suppliers and the CDC are thus aggregated into a single artificial facility PZ. PZ is the unique origin of all material flows.
- The Regional Distribution Centers (RDCs) are the facilities to be located. They form an intermediate layer between the production zone and the customers. The delivery process generally follows a sequence of two successive transportation segments: an FTL route from PZ to an RDC and an LTL shipment from the RDC to a customer.
- A large set of customers who are the destinations of all product flows. Each customer is assigned to a single RDC during the whole time horizon.

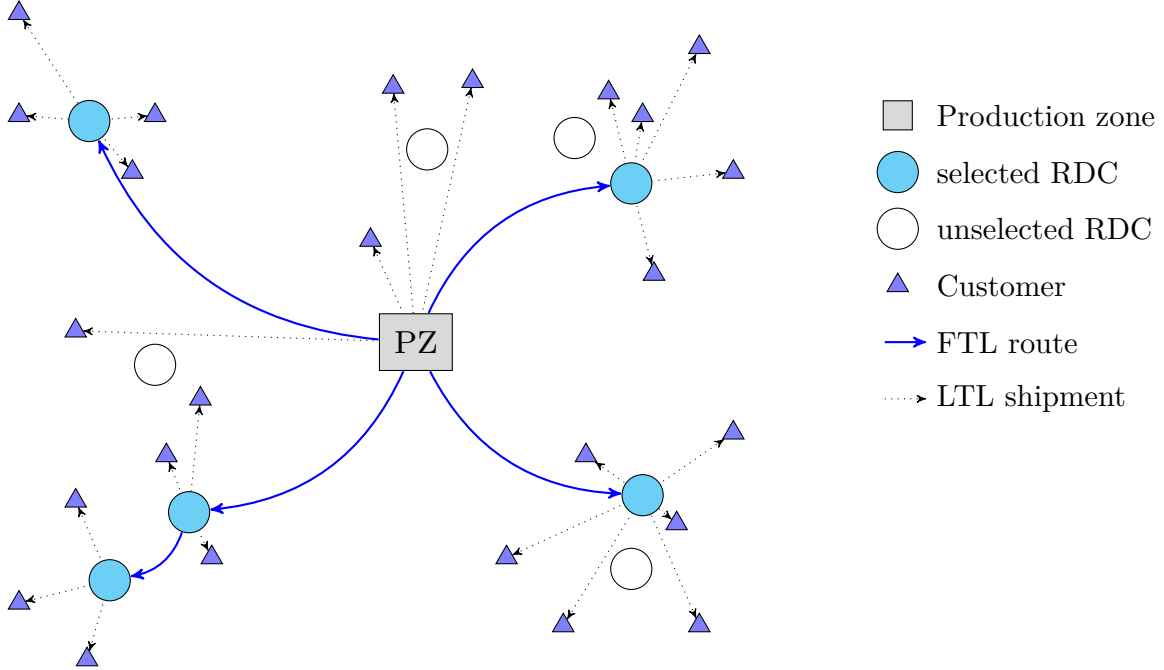


Figure 1: The collaborative distribution network

Our model assumes a homogeneous fleet of vehicles on each route. Since transportation is subcontracted, we assume that a large enough number of vehicles is available on every FTL or LTL route.

### 3.2. FTL routes

The use of FTL routes implicitly assumes that enough shipments have been consolidated, making it possible to fill a vehicle and travel directly from an origin to a destination. In theory, all FTL routes should be direct trips between PZ and an RDC. In practice, some carriers allow one stopover at an intermediate RDC, provided this creates no, or very small detour. This stopover incurs a fixed stopover cost.

The cost of an FTL route is independent of the number of units carried. It includes a fixed part representing tolls or stopover charges and a cost proportional to the distance traveled or hours worked (wages of drivers, fuel, use of truck). Thus, it is assumed constant and known a priori.

Since there can be only one or two RDCs on each FTL route, the set of all FTL routes can be easily enumerated, with a good cost estimate based on the previous experience of the companies.

We assume that customer orders can be split into several FTL routes, but partial orders must be consolidated at an RDC and then be shipped at once to the final customer. RDCs



are cross-docking facilities only. Therefore, only very short-term inventory is allowed at RDCs.

### 3.3. LTL shipments

LTL shipments concern the distribution of goods to final customers. These shipments are outsourced to regional LSPs who collect the goods at an RDC and define routes with multiples stops to deliver customers. Note that PZ plays the role of an RDC for customers located in the same region. Moreover, LTL shipments can be organized between PZ and any customer (including remote customers). Thus, PZ can be considered as a special RDC, which is always selected.

We empirically observed that the cost of LTL shipments increases with the number of units carried and incorporates discount rates. Figure 2 illustrates an example of our estimated LTL cost from a given RDC to a given department.

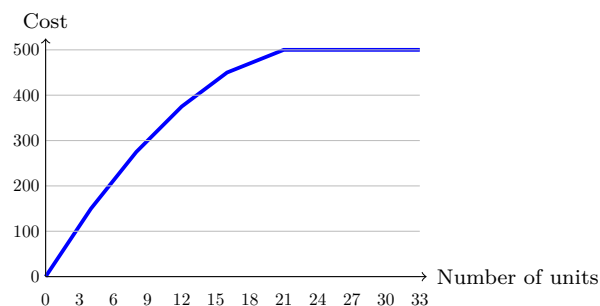


Figure 2: Cost of an LTL shipment

Unfortunately, since the shipments concern small discrete quantities, identifying break-points and capturing the parameters of the cost curve for each RDC is impossible in practice. Moreover, the available data concern only the shipments from PZ to the customers, and not from each candidate RDC location to the customers. Following the approach of le Blanc et al. (2006) and Özkaya et al. (2010), we estimated the costs of LTL shipments from each RDC by means of a regression model. The LTL cost is expressed as a function of the number of units shipped and the distance between the origin and the destination. Let  $c(q, j, d)$  represent the cost of shipping  $q$  units from one RDC  $j$  to one department  $d$ , and  $dist$  be the distance between the RDC and the department of the customer. The regression formula found is of the form  $c(q, j, d) = \alpha \times dist^\beta \times q^\gamma$ , where  $\alpha$ ,  $\beta$  and  $\gamma$  are constant coefficients, with an adjusted  $R^2$  of 0.94.

### 3.4. Supplier and Customer aggregation

Our case study concerns the distribution of horticultural products in France. In France, the territory is partitioned into 95 *departments*. Most carriers calculate their tariffs with a matrix of transportation costs between all pairs of departments. Since all suppliers and

the CDC are located in the same department, they are charged identical transportation costs. Thus, they can be aggregated in the mathematical model. To do so, we define the dummy facility PZ which represents the whole set of suppliers and the CDC.

When the number of customers is very large, a popular approach is to aggregate all customers from a given area to the centroid of this area. This generates smaller and thus more tractable problems, at the price of imprecision due to the aggregation error (Francis & Lowe, 2015). In our case, LTL shippers intrinsically assume aggregation, since shipment costs are given for each department. All customers in the same department have the same LTL shipment cost. Moreover, static single-sourcing is adopted by many firms in their strategic network design (Wu et al., 2015): each department is delivered from the same RDC during the whole time horizon considered.

### 3.5. Seasonality and time aggregation of shipments

Figure 3 illustrates the seasonality effect during one year of shipments.

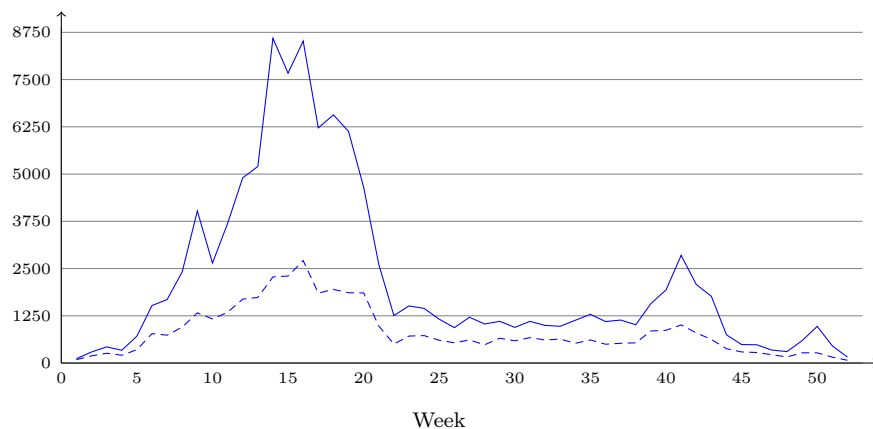


Figure 3: Seasonality of customer demands

The solid line represent the number of units delivered and the dashed line represents the number of deliveries per week. The period between weeks 10 and 21 represents 51.5% of the yearly activity. Given the seasonality, the potential for flow consolidation cannot be identical at every period. We define three main seasons: the *high season* (week 13-21) represents 9 weeks with more than 4 500 units shipped weekly, the *mid-season* represents 29 weeks with between 900 and 4 500 units shipped weekly, and the *low season* represents 14 weeks with less than 900 units shipped weekly.

During the high season, the customer orders are frequent and larger. This reinforces the opportunity to use FTL routes to RDCs. During the low season, customer orders are less frequent and smaller. The incentive for FTL routes is limited, but large relative savings are expected from flow consolidation and direct shipments for PZ.

A necessary condition for efficiently running collaborative distribution is the synchronization of shipments. Shipments are not performed daily: every week is split into a number of independent campaigns. During the high season, we considered three weekly deliveries. For example, if the delivery pattern to an RDC is “Mo Tu / We Th / Fr Sa Su”, shipments leave PZ on Tuesdays, Thursdays, Sundays. These days are called *shipping dates*. During the mid-season, there are 2 weekly shipping dates. During the low season, there is only 1 weekly shipping date. Voluntary postponement of shipments in order to improve consolidation at the next shipping date is not authorized. We assumed identical delivery patterns for all RDCs. In practice, in order to smooth the working load in the production zone, the RDCs have distinct shipping dates. This slight simplification of the reality has very limited impact on numerical results.

## 4. MILP formulation

In this section, we formulate the multi-period facility location problem that jointly optimizes the location of RDCs, the definition of FTL routes and the assignment of customers to RDCs.

### 4.1. A $p$ -median formulation

One strong characteristic of the optimization problem considered is that the objective function contains no fixed cost associated with the selection of candidate facilities. Each RDC is operated by a regional LSP in charge of cross-docking operations at the RDC and LTL distribution to customers. LSPs charge a variable cost for each logistic unit processed. This cost is already included in the LTL shipment cost, so that it does not need to appear in our mathematical formulation.

In practice, there are some initial fixed costs corresponding to the research and assessment of all potential LSPs as well as an administrative cost of establishing contracts with selected LSPs. However, these costs can be included in the general project management cost and are not specific to each RDC. Thus, we ignore them.

Since shipments represent a large part of the distribution costs, having many RDCs is likely to decrease the value of the objective function. On the other hand, having too many RDCs increases supply chain complexity because it involves managing many stakeholders. Complexity in the supply chain is not a desirable feature and has been argued to decrease the performance of operations and complicate decision-making (Perona & Miragliotta, 2004; Bode & Wagner, 2015). The optimal number of facilities achieves some trade-off between financial and organizational aspects but some of the hidden costs are roughly estimated. Thus, we modeled the facility location problem as an extension of the  $p$ -median problem. The number  $p$  of selected RDCs is imposed, but it is possible to build several scenarios corresponding to various relevant values of  $p$ . This practice corresponds to the approach of many decision-makers, who generally have a rough idea of the desirable number of facilities to be selected, but wish to compare a few scenarios.

<i>Sets</i>	
$J$	set of RDCs
$j_0 \in J$	Production Zone (PZ)
$J^* = J \setminus \{j_0\}$	
Set of candidate locations $I$	Set of customers
$D$	Set of Departments
$I_d \subset I$	Subset of customers located in department $d \in D$
$D_j \subset D$	Subset of departments that can be served from $j \in J$
$\Omega$	Set of FTL routes
$\Omega_j \subset \Omega$	Subset of FTL routes visiting $j \in J^*$
$J_\omega \subset J^*$	Subset of candidate RDCs visited by route $\omega \in \Omega$
$T$	Set of shipping dates
<i>Parameters</i>	
$p$	Number of candidate RDCs to be selected
$Q_\omega$	Capacity of the vehicle on the FTL route $\omega \in \Omega$
$c_\omega$	Cost of the vehicle on the FTL route $\omega \in \Omega$ , including stopover costs
$q_i^t$	Demand of customer $i$ on date $t$
$c_{ij}^t$	Cost of LTL distribution of each of the $q_i^t$ units from distribution center $j \in J$ to customer $i \in I$

Table 2: Notation used in the mathematical model

#### 4.2. MILP formulation

The mathematical model uses notations defined in Table 2.

Since direct distribution from PZ to customers is authorized, PZ is modeled as a special element of  $J$ , denoted  $j_0$ . Then, the set of RDCs that are candidate for selection is denoted by  $J^* = J \setminus \{j_0\}$ . For obvious logistics reasons, an RDC  $j \in J^*$  can deliver only a subset  $D_j \subset D$  of departments. The production zone PZ can deliver all departments, thus  $D_{j_0} = D$ . Each FTL route  $\omega \in \Omega$  is operated by a homogeneous fleet of vehicles of capacity  $Q_\omega$ . Note that cost  $c_{ij}^t$  includes handling costs at RDCs.

The model variables are defined in Table 3.

The location-allocation problem described can be modeled as follows:

$$\min z = \sum_{t \in T} \sum_{\omega \in \Omega} c_\omega n_\omega^t + \sum_{t \in T} \sum_{i \in I} \sum_{j \in J} c_{ij}^t f_{ij}^t \quad (1)$$

---

*Binary and integer variables*

$y_j$	Set at 1 if RDC $j \in J^*$ is selected, and 0 otherwise
$x_{dj}$	Set at 1 if department $d \in D$ is served by RDC $j \in J$ , and 0 otherwise
$n_\omega^t$	Number of FTL vehicles used on route $\omega \in \Omega$ on shipping date $t \in T$

---

*Continuous variables*

$u_{\omega j}^t$	Number of units on route $\omega \in \Omega$ unloaded at RDC $j \in J^*$ on shipping date $t \in T$
$f_{ij}^t$	Number of units shipped from RDC $j \in J$ to customer $i \in I$ , on shipping date $t \in T$

---

Table 3: Variables of the mathematical model

s.t.

$$\sum_{j \in J^*} y_j = p \quad (2)$$

$$\sum_{j \in J} x_{dj} = 1 \quad \forall d \in D \quad (3)$$

$$\sum_{j \in J} f_{ij}^t = q_i^t \quad \forall i \in I, \forall t \in T \quad (4)$$

$$f_{ij}^t \leq q_i^t x_{dj} \quad \forall i \in I_d, \forall d \in D_j, \forall j \in J^*, \forall t \in T \quad (5)$$

$$x_{dj} \leq y_j \quad \forall d \in D_j, \forall j \in J^* \quad (6)$$

$$n_\omega^t \leq y_j M_\omega^t \quad \forall \omega \in \Omega_j, \forall j \in J^*, t \in T \quad (7)$$

$$\sum_{i \in I_d, d \in D_j} f_{ij}^t = \sum_{\omega \in \Omega_j} u_{\omega j}^t \quad \forall j \in J^*, \forall t \in T \quad (8)$$

$$\sum_{j \in J_\omega} u_{\omega j}^t \leq Q_\omega n_\omega^t \quad \forall \omega \in \Omega, \forall t \in T \quad (9)$$

$$x_{dj} \in \{0, 1\} \quad \forall d \in D, \forall j \in J \quad (10)$$

$$y_j \in \{0, 1\} \quad \forall j \in J^* \quad (11)$$

$$n_\omega^t \in \mathbb{N} \quad \forall \omega \in \Omega, \forall t \in T \quad (12)$$

$$f_{ij}^t \in \mathbb{R} \quad \forall i \in I, \forall j \in J, \forall t \in T \quad (13)$$

$$u_{\omega j}^t \in \mathbb{R} \quad \forall \omega \in \Omega, \forall j \in J, \forall t \in T \quad (14)$$

The objective function (1) sums up the cost of FTL routes and the LTL distribution costs. Constraint (2) indicates the number  $p$  of RDCs to be selected. Constraints (3)

are single assignment constraints for each department. Constraints (4) ensure customer demand satisfaction. Note that the source  $j \in J$  of material flow to customer  $i \in I$  can be the RDC assigned to  $i$  and/or the production zone. Constraints (5) state that if RDC  $j \in J$  does not serve department  $d \in D_j$ , then there is no product flow from  $j$  to  $i \in I_d$ . Constraints (6) state that a department  $d$  can only be delivered by a selected RDC. Since  $j_0$  is always selected, this constraint applies only for  $j \in J^*$ . Constraints (7) state that if an RDC is non selected, then the FTL routes visiting this location are not operating. The value of  $M_\omega^t$  is set at the sum of all customers demands which can be delivered from all  $j \in J_\omega$  on date  $t$ , divided by the vehicle capacity:

$$M_\omega^t = \left\lceil \frac{\sum_{i \in I_d, d \in D_j, j \in J_\omega} q_i^t}{Q_\omega} \right\rceil.$$

Constraints (8) model flow conservation at an RDC  $j \in J^*$ . Constraints (9) model vehicle capacity. Constraints (10)–(14) define the nature of decision variables.

Note that constraints (5) can be aggregated into the following form:

$$\sum_{i \in I_d} f_{ij}^t \leq \sum_{i \in I_d} q_i^t x_{dj} \quad \forall d \in D_j, \forall j \in J^*, t \in T.$$

In practice, we observed a negligible difference in the solver's performance when using aggregated or disaggregated versions of these constraints.

#### 4.3. Additional real-life constraints

The model (1)–(14) ignores some business constraints which must be considered in order to capture all preferences of decision-makers. First, some candidate locations from the same region may directly compete with each other. Mutual exclusion between candidate locations can be formally modeled by *disjunctive constraints*. More formally, if at most  $\bar{p}$  candidate locations can be selected in a set  $\bar{J} \subset J^*$ , this can be formulated by the following constraints:

$$\sum_{j \in \bar{J}} y_j \leq \bar{p}. \quad (15)$$

Note that some disjunctive constraints can be *a priori* formulated by the decision makers. Other competing subsets of facilities may arise from the analysis of numerical results and be appended *a posteriori* to the mathematical model.

Secondly, a solution of the mathematical model is a selection of RDCs and an allocation of departments to selected RDCs. A potential problem is that the set of all departments allocated to an RDC does not always fit with the regions and the itineraries actually serviced by the LTL shipping companies. Thus, we formulate so-called *districting constraints* that eliminate unrealistic allocations. This idea was inspired by Rossi et al. (2012). Let us consider an RDC  $j \in J^*$  and the set  $D_j$  of all departments possibly allocated to  $j$ . If a

distant department  $d$  is possibly allocated to  $j$ , then an itinerary from  $j$  to  $d$  will form a chain of departments. The set of all chains forms a directed acyclic graph with root  $j$ , as illustrated in Figure 4.

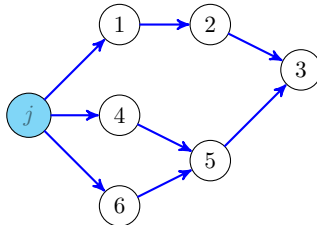


Figure 4: A graph illustrating districting constraints

The RDC  $j$  can deliver departments 1 to 6. A route from  $j$  to 5 crosses either 4 or 6. Districting constraints state that department 5 is allocated to  $j$  only if either 4 or 6 is also allocated to  $j$ . Similarly, department 3 is allocated to  $j$  only if one of its predecessors, 2 or 5, is allocated to  $j$ .

For each RDC, we define an acyclic graph with the main itineraries followed by vehicles. Then, we derive districting constraints of the following form:

$$x_{dj} \leq \sum_{d' \in \Gamma^-(d)} x_{d'j}, \quad (16)$$

where  $\Gamma^-(d)$  is the set of predecessors of  $d$  in the acyclic graph.

## 5. Experiments and case study

In this section, we describe the real-life case study which motivated this paper. Then, we describe the numerical experiments which resulted in the selection of a subset of candidate facilities.

### 5.1. The case study

The case study concerns the distribution of horticultural products in France. The suppliers form a cluster of 4 companies located in Western France; the maximal distance between two suppliers is only 60 km.

Companies that resort to horizontal collaboration are often competitors. They have to overcome many cultural, organizational and technical barriers, including the fear of sharing information, lack of trust, loss of flexibility, loss of their own control or extra cost of new IT tools. In our case study, all suppliers compete with each other, but all belong to the same world-class competitiveness cluster and have been leading common research and development projects together. Thus, it can be assumed that all prerequisites to establishing collaboration are already in place.

The customers are mainly specialized stores, supermarkets or retail stores, but also resellers, local authorities, schools, etc. They are spread over the whole territory, at an average distance of 380 km from the suppliers. The quantity of customer orders is generally very small. The average quantity per shipment is only 2.9 units. More precisely, specialized stores represent 65% of the customers and 77% of the sales. DIY stores represent 16% of the customers and 15% of the sales, large retailers 10% and 4%, and a large set of small customers (flower stores, tree nurseries, professionals, municipalities, schools) 9% and 4% respectively. One strong characteristic of this market is that no customer can strongly influence the distribution process. To give an idea, the first national brand in the area of horticultural products has more than 1000 shops, but only 8% of the market. Moreover, many delivery points are shops working under a franchise business model. Thus, they are all considered as distinct customers with separate procurement strategies.

Saving on the logistics costs is not the only incentive for horizontal collaboration. We could observe other collateral benefits, which are harder to precisely quantify. Collaborative distribution (i) increases the quality of service to customers, thanks to the joint deliveries instead of separate small-volume deliveries, (ii) enables delivering in remote regions that were considered “unreachable” beforehand from a commercial point of view, (iii) stabilizes the supplier-carrier relationships, by setting long-term agreements and giving repeated business to carriers, (iv) pushes the collaborating companies to improve their internal processes and to harmonize their information system and database.

The available data includes a list of 29 candidate RDCs and 67 potential FTL routes containing, at most, two RDCs. We have an exhaustive record of one complete year of the activity of the collaborating suppliers. This represents 42886 shipments to 3640 customers. The number of weekly deliveries is 3 during the high season, 2 during the mid-season and 1 during the low season. This results in 99 distinct shipping dates. The corresponding mathematical model has 2784 binary variables (29 variables  $y_j$ , 2755 variables  $x_{dj}$ ), 6633 integer variables  $n_{\omega}^t$ .

## 5.2. Experimental design

We used the mathematical model (1)-(14) with two approaches:

- **Scenarios  $p$  ( $1 \leq p \leq 29$ ):** each value of the parameter  $p$  in the model (1)-(14) defines a distinct optimization problem which is a generalization of the  $p$ -median problem. A first approach was to use the solver on each of these 29 problems ( $1 \leq p \leq 29$ ) independently, with a maximal computing time set at 1 hour.
- **Scenario 0:** we also defined a more global approach, consisting in relaxing constraints (2) in order to get a free number of selected locations. In order to avoid selecting dummy RDCs with no material flow associated, we added an arbitrarily small fixed cost for each selected RDC. We ran this model with a maximal computing time of 12 hours.



All models were coded in Java using the Concert Technology framework of CPLEX 12.6. Tests were run on a PC Intel Core i7-3537 processor 2.0 Ghz with 8 Gb of memory under the System Windows 8.0. We tested different Cplex parameters and eventually played each scenario twice, with MIP emphasis set on feasibility and optimality respectively. Since in scenario 0, the solution space is larger than in any scenario  $p$ , scenario 0 yields a lower bound of the optimal value of scenario  $p$ . For each value of  $p$ , we have two upper bounds and four lower bounds (two obtained in scenario  $p$  and two obtained in scenario 0.) We denote  $u_p$  the lowest upper bound and  $l_p$  the highest lower bound obtained in scenario  $p$ . We also denote  $u_0$  and  $l_0$  the best upper and lower bounds obtained in scenario 0.

### 5.3. Computational results

Due to confidentiality issues, the real costs cannot be shown in detail. In particular, we are not authorized to disclose the savings estimates. However, we can evaluate both scenarios based on the lower bounds and feasible solutions provided by the solver. Figure 5 presents the evaluation of cost  $u_p$  for all values of  $p$ . Since there is no fixed cost associated with the selection of facilities, the optimal value of the  $p$ -median problems decreases as  $p$  increases. In practice, we observe that  $u_p$  decreases until  $p = 16$  and then it is quite stable. The relative difference between  $u_1$  and  $u_{16}$ , computed as  $(u_1 - u_{16})/u_{16} \times 100$ , is 38.3%. The relative difference between  $u_{12}$  and  $u_{16}$  is only 1.7%.

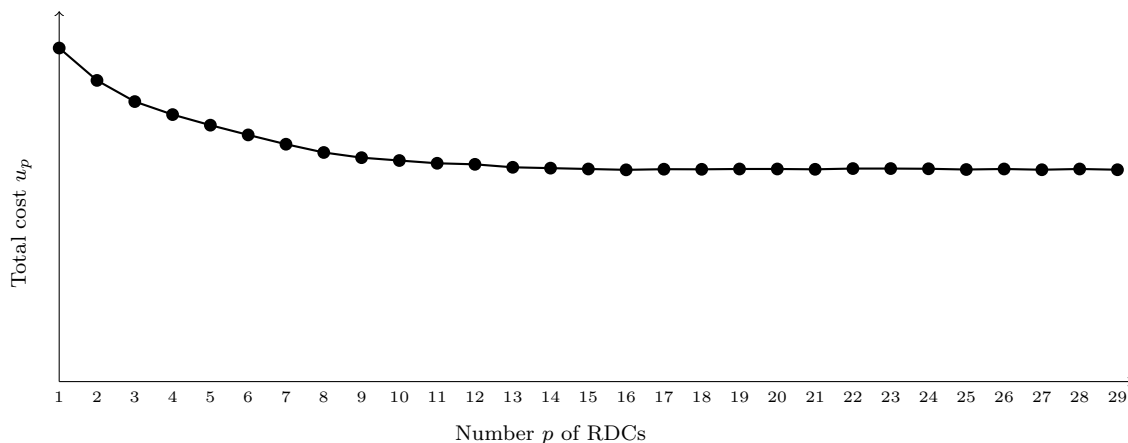


Figure 5: Evaluation of costs in scenarios  $1 \leq p \leq 29$

In scenario 0, the best solution  $u_0$  correspond to the value  $p = 18$ . This solution is better than  $u_{18}$  but slightly worse than  $u_{16}$  (the relative difference is 0.04%). If we assume a fixed cost of 50000 euros for selecting an RDC (a very conservative estimate), the model recommends selecting 8 RDCs. These observations can help decision-makers estimate the best possible value of  $p$ .

Figure 6 presents the relative gaps between  $u_p$  and the lower bound  $\hat{l}_p = \max(l_0, l_p)$ . This gap is calculated as  $(u_p - \hat{l}_p)/\hat{l}_p \times 100$ . For  $p \leq 3$ , the solver finds optimal solutions. The relative gap increases when  $p$  is between 4 and 12. For these instances, the lower bound  $\hat{l}_p$  corresponds to the value  $l_p$ . For  $p \geq 15$ , the gap remains in the interval  $[1.5\%, 2\%]$ . For these instances, the lower bound  $\hat{l}_p$  is given by  $l_0$ . The average gap over all 29 instances is 1.55%.

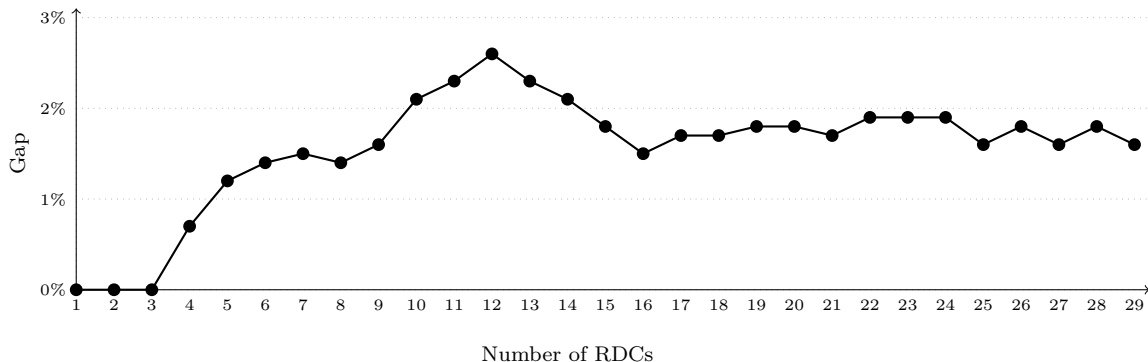


Figure 6: Evaluation of optimality gaps according to  $p$

To assess the interest of collaborating, we compared the pooled network with a situation in which all suppliers organize their transportation independently. We followed the following methodology: the data was split into four independent subsets corresponding to the orders of each supplier. We ran the facility location model for each supplier separately. For each value of  $p$ , let  $l'_p$  define the sum of the four corresponding lower bounds. We calculated the relative difference between upper bounds  $u_p$  of the pooled network cost and lower bounds  $l'_p$ . The gap  $(l'_p - u_p)/u_p \times 100$  represents a guaranteed cost decrease due to collaboration. The value of this gap is always positive, and increases regularly from 1 to 16. For confidentiality reasons, we do not give any precise estimate. We can simply indicate that for realistic number of RDCs (from 12 to 16), this gap does not contradict other case studies found in the literature (see Table 1).

#### 5.4. Analysis of the solutions

The decision-making process is based on the results of the MILP models and on an analysis of the material flow. One concern of decision-makers is how to ensure sufficient material flow in each selected RDC. Since there is no fixed cost associated with each RDC, our modeling can lead to solutions in which some RDCs are virtually selected but have no material flow. This situation is likely to happen mainly for large values of  $p$ .

For all scenario with  $p \leq 16$ , we observe that selected RDCs receive significant material flow. For  $p = 17$ , the solution found by CPLEX after 1h of calculation corresponds to the

same selection of candidate RDCs, plus an additional RDC with no flow. For  $p \geq 20$ , a dummy RDC is always found and the number of actual RDCs is between 18 and 20.

Although RDCs do not have strict capacity restrictions, it is assumed that significant quantities will be shipped to each operating facility. Figure 7 shows the material flow operated by a candidate facility (called  $A$ ) which is selected in almost all scenarios.

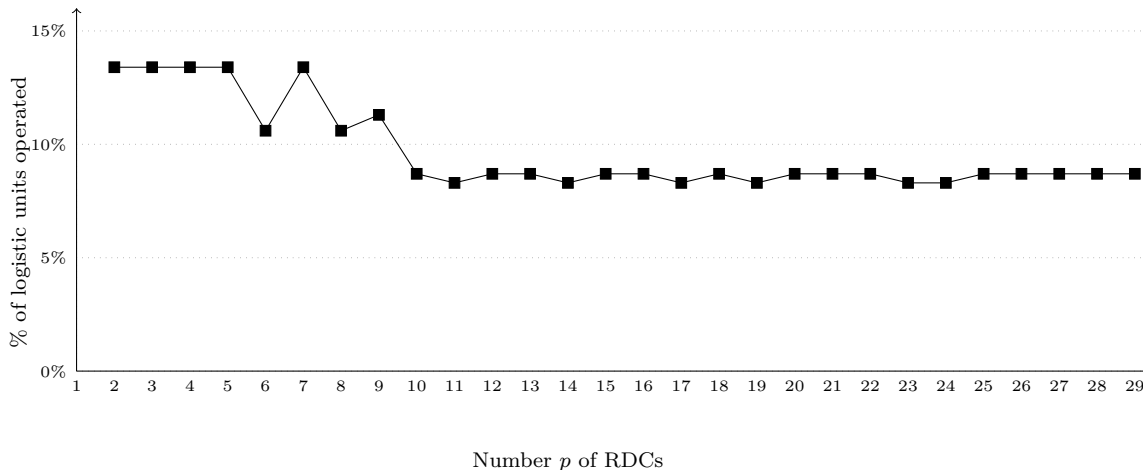


Figure 7: Evaluation of flow according to  $p$  for RDC  $A$

This flow is expressed as a percentage of the whole flow shipped by the collaborating suppliers.  $A$  is selected as soon as  $p = 2$  and is always selected for larger values of  $p$ . When  $p$  increases, the percentage is globally decreasing, but not monotonically, since the percentage also depends on the other facilities selected. The stable parts (between  $p = 2$  and 5, and for  $p \geq 10$ ) can be interpreted as follows. The newly selected RDC is not located in the same region as  $A$ , so that they do not really compete. Thus adding an RDC does not modify the set of customers delivered by  $A$ .

Let us now consider another example, with two neighboring candidate facility locations denoted by  $B$  and  $C$ . Figure 8 shows the percentage of material flow operated by  $B$  (triangles) and  $C$  (circles) in scenario  $p$ . It can be seen that for all  $p \geq 9$ , either  $B$  or  $C$  is selected, but never both simultaneously. Moreover, it seems that  $B$  and  $C$  deliver exactly the same departments, at similar cost and with the same percentage of material flow. These facilities can be considered direct competitors.

One more complex situation is illustrated in Figure 9. It shows the percentage of material flow operated by three candidate facility locations  $D$  (squares),  $E$  (triangles) and  $F$  (circles) in each instance of the  $p$ -median model. Facility  $D$  is located between  $E$  and  $F$ . Only facility  $D$  is selected for  $3 \leq p \leq 6$ . When  $7 \leq p \leq 14$ ,  $D$  is no longer selected but either  $E$  or  $F$  is selected. For  $p \geq 15$ ,  $E$  and  $F$  are both selected. This shows that  $D$  competes with  $E$  and  $F$  but  $E$  and  $F$  compete only partially with each other.

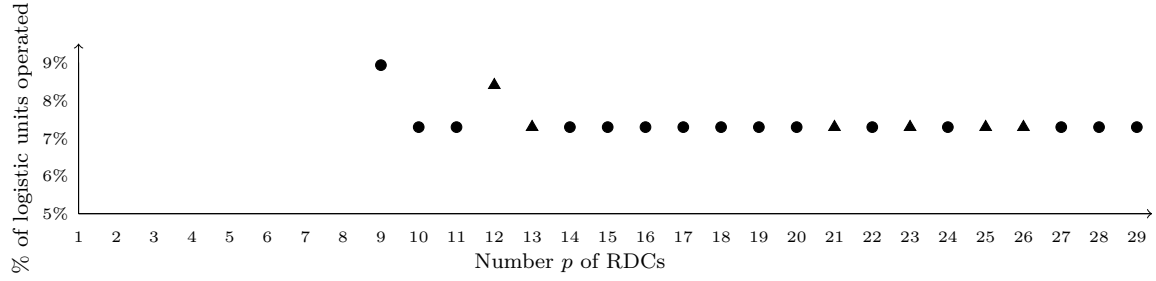


Figure 8: Percentage of material flow operated by RDCs  $B$  (triangles) and  $C$  (circles)

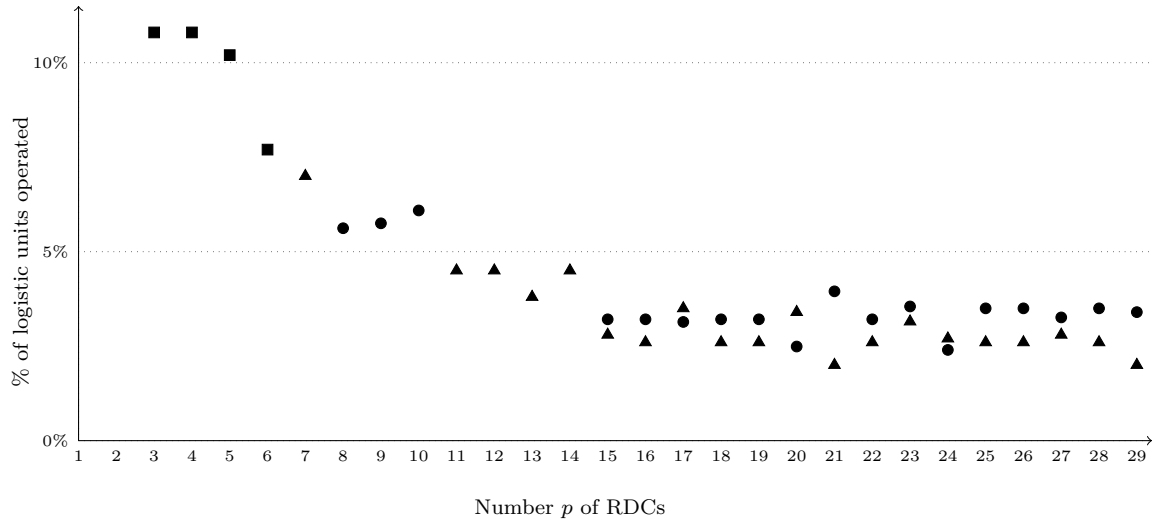


Figure 9: Percentage of material flow operated by RDCs  $D$  (square),  $E$  (triangle) and  $F$  (circle)

$p$	Impact of disjunctive constraints	Impact of districting constraints
5	+0.12 %	+0.12 %
6	0 %	+0.17 %
7	-0.03 %	+0.14 %
8	0%	+0.08 %
9	-0.08 %	+0.08 %
10	+0.2 %	-0.03 %
11	-0.36 %	-0.16 %
12	-0.35 %	-0.20 %

Table 4: Impact of disjunctive and districting constraints

The examples in Figures 8 and 9 show how it is possible to isolate subsets of competing candidate facilities, that can be modeled by disjunctive constraints of the form (15).

#### 5.5. Impact of disjunctive and districting constraints

In order to assess the impact of the disjunctive constraints (15) and the districtive constraints (16), we compared the numerical results with or without these constraints, with a fixed number of RDCs ranging from 5 to 12.

For each pair of neighboring candidate locations  $j$  and  $j'$ , we enriched the model with a disjunctive constraint of the form  $j + j' \leq 1$ . Then, we solved the model with a computing time limit of 1 hour by emphasizing feasibility in CPLEX. This results are presented in column 2 of Table 4.

Then, we enriched the original model with 190 chains and 331 districting constraints, and solved the model with a computing time limit of 1 hour by emphasizing feasibility in CPLEX. The results are presented in column 3 of Table 4.

In columns 2 and 3, the impact of additional constraints is measure as the ratio  $(z' - z)/z \times 100$ , where  $z'$  is the best solution found with the additional constraints considered and  $z$  is the best solution found with the original model (1)-(14).

It is worth noting that adding new constraints theoretically degrades the objective value of an optimal solution, but also helps the solver in its branching strategy. This can explain why some negative values appear in column 2. The important conclusion is that these constraints improve the modeling of decision-makers preferences without reducing the quality of the solutions.

The percentage in column 2 represents the impact of districting constraint on the value of the objective function. For the same reasons as in Table 4, some improvements could be observed although districting constraints are supposed to increase the value of the objective

function. Here again, we can conclude that districting constraints facilitate the allocation process without penalizing the economic aspects.

## 6. Conclusion

In this paper, we addressed the optimal location of intermediate facilities between a cluster of collaborating suppliers and a large set of customers. We proposed a mixed integer linear programming (MILP) formulation with uncapacitated facilities and no fixed cost associated with the selection of RDCs. We considered two versions of the MILP, in which the number  $p$  of selected facility locations is either free or set at a predefined value. This helps decision-makers identify a tight interval for  $p$ . The model was then enriched with disjunctive constraints that model the existence of conflicting candidates and so called districting constraints that favor compact regions around each RDC.

Several practical and theoretical extensions can be proposed to this work. As far as theoretical extensions are concerned, the model could include heterogeneous fleet on each route and be reinforced by valid inequalities (note that the disjunctive and districting constraints are not valid inequalities). Several theoretical extensions could also be motivated by practical considerations. First, setting or modifying a complex logistics network is generally a long process that cannot be implemented at once. All decisions can be regularly revised. In particular, RDC locations can be revised for example at the end of each transportation contract with the carriers, in case of commercial development in some new regions or if new partners join the coalition. This flexibility is made possible since the whole distribution system is subcontracted: the suppliers do not buy any facility, but only logistics and transportation services. This suggests dynamic facility location models that represent the deployment of a logistics network over several years. Another motivation for defining multi-periodic models is the possible introduction of flexibility in the network, with the existence of seasonal RDCs that are used only during specific seasons of each year. A slight adaptation of the model would be to propose FTL routes visiting one or two intermediate RDCs. Finally, the current model is considered deterministic although it uses uncertain data. Solving a stochastic model with uncertain future demand was not possible in practice due to the difficulty of modeling demand uncertainty. However, we think this is a challenging perspective both from a theoretical and practical point of view.

Decision makers are often refrained from using mathematical models as a decision aiding tool for facility location decision. Among the practical impediments, we can cite the difficulty to collect the necessary data and to estimate all parameters. Thus, robustness of decisions is an issue in this kind of application. In our case, robustness was checked by an a posteriori analysis of solutions. We scanned the solutions obtained in each scenario and performed a flow analysis. The objective was to choose RDCs that were selected in most scenarios and that process a large quantity of goods. Although no minimal flow constraint exists in the mathematical model, this was an underlying requirement of decision makers.

Integrating robustness issues in the modeling and solution method would be a theoretical contribution with high practical impact.

## Acknowledgement

This research is part of the FUI 15 project Végésupply, which is partially funded by BPI France and labeled by the competitiveness clusters Novalog and Vegepolys.

## References

## References

- B. Adenso-Díaz, et al. (2014). ‘Analysis of the synergies of merging multi-company transportation needs’. *Transportmetrica A: Transport Science* **10**(6):533–547.
- S. Alumur & B. Y. Kara (2008). ‘Network hub location problems: The state of the art’. *European Journal of Operational Research* **190**(1):1–21.
- J. Andersen, et al. (2011). ‘Branch and Price for Service Network Design with Asset Management Constraints’. *Transportation Science* **45**(1):33–49.
- J.-F. Audy, et al. (2012). ‘An empirical study on coalition formation and cost/savings allocation’. *International Journal of Production Economics* **136**(1):13–27.
- J.-F. Audy, et al. (2011). ‘Cost allocation in the establishment of a collaborative transportation agreement - an application in the furniture industry’. *Journal of the Operational Research Society* **62**(10):960–970.
- C. Bode & S. Wagner (2015). ‘Structural Drivers of Upstream Supply Chain Complexity and the Frequency of Supply Chain Disruptions’. *Journal of Operations Management* **36**:215–228.
- P. Buijs & J. H. Wortmann (2014). ‘Joint operational decision-making in collaborative transportation networks: the role of IT’. *Supply Chain Management: An International Journal* **19**(2):200–210.
- J. F. Campbell (1994). ‘Integer programming formulations of discrete hub location problems’. *European Journal of Operational Research* **72**(2):387–405.
- J. F. Campbell (2005). ‘Strategic Network Design for Motor Carriers’. In A. Langevin & D. Riopel (eds.), *Logistics Systems: Design and Optimization*, pp. 245–278. Springer US.
- I. Contreras (2015). ‘Hub location problems’. In G. Laporte, S. Nickel, & F. Saldanha da Gama (eds.), *Location Science*, pp. 311–344. Springer.

- T. G. Crainic, et al. (2014). ‘Service Network Design with Resource Constraints’. *Transportation Science (available online)* **0**(0):null.
- K. L. Croxton, et al. (2003). ‘Models and Methods for Merge-in-Transit Operations’. *Transportation Science* **37**(1):1–22.
- F. Cruijssen, et al. (2007a). ‘Joint route planning under varying market conditions’. *International Journal of Physical Distribution & Logistics Management* **37**(4):287–304.
- F. Cruijssen, et al. (2007b). ‘Horizontal cooperation in logistics: Opportunities and impediments’. *Transportation Research Part E* **43**(2):129–142.
- B. Dai & H. Chen (2012). ‘Profit allocation mechanisms for carrier collaboration in pickup and delivery service’. *Computers and Industrial Engineering* **62**(2):633–643.
- M. S. Daskin & K. L. Maass (2015). ‘The p-median problem’. In G. Laporte, S. Nickel, & F. Saldanha da Gama (eds.), *Location Science*, pp. 21–46. Springer.
- A. L. Erera, et al. (2013a). ‘Creating schedules and computing operating costs for LTL load plans’. *Computers & Operations Research* **40**(3):691–702.
- A. L. Erera, et al. (2013b). ‘Improved Load Plan Design Through Integer Programming Based Local Search’. *Transportation Science* **47**(3):412–427.
- Ö. Ergun, et al. (2007). ‘Shipper collaboration’. *Computers & Operations Research* **34**(6):1551–1560.
- R. Francis & T. J. Lowe (2015). ‘Demand point aggregation for some basic location models’. In G. Laporte, S. Nickel, & F. Saldanha da Gama (eds.), *Location Science*, pp. 487–506. Springer.
- M. Frisk, et al. (2010). ‘Cost allocation in collaborative forest transportation’. *European Journal of Operational Research* **205**(2):448–458.
- S. Gelareh & S. Nickel (2011). ‘Hub location problems in transportation networks’. *Transportation Research Part E: Logistics and Transportation Review* **47**(6):1092–1111.
- B. Groothedde, et al. (2005). ‘Towards collaborative, intermodal hub networks: A case study in the fast moving consumer goods market’. *Transportation Research Part E* **41**(6):567–583.
- J. Hill & M. Galbreth (2008). ‘A heuristic for single-warehouse multiretailer supply chains with all-unit transportation cost discounts’. *European Journal of Operational Research* **187**(2):473–482.



- M. Hingley, et al. (2011). ‘Using fourth-party logistics management to improve horizontal collaboration among grocery retailers’. *Supply Chain Management: An International Journal* **16**(5):316–327.
- M. A. Krajewska & H. Kopfer (2009). ‘Transportation planning in freight forwarding companies: Tabu search algorithm for the integrated operational transportation planning problem’. *European Journal of Operational Research* **197**(2):741–751.
- S. D. Lapierre, et al. (2004). ‘Designing Distribution Networks: Formulations and Solution Heuristic’. *Transportation Science* **38**(2):174–187.
- G. Laporte, et al. (2015). *Location Science*. Springer International Publishing.
- H. M. le Blanc, et al. (2006). ‘Factory gate pricing: An analysis of the Dutch retail distribution’. *European Journal of Operational Research* **174**(3):1950–1967.
- R. Leitner, et al. (2011). ‘Structural concepts for horizontal cooperation to increase efficiency in logistics’. *CIRP Journal of Manufacturing Science and Technology* **4**(3):332–337.
- K. A. Lindsey, et al. (2013). ‘A pickup and delivery problem using crossdocks and truckload lane rates’. *EURO Journal on Transportation and Logistics* **2**(1-2):5–27.
- R. Mason, et al. (2007). ‘Combining vertical and horizontal collaboration for transport optimisation’. *Supply Chain Management: An International Journal* **12**(3):187–199.
- M. T. Melo, et al. (2009). ‘Facility location and supply chain management - A review’. *European Journal of Operational Research* **196**(2):401–412.
- L. Mui Ann Chan, et al. (2002). ‘Effective Zero-Inventory-Ordering Policies for the Single-Warehouse Multiretailer Problem with Piecewise Linear Cost Structures’. *Management Science* **48**(11):1446–1460.
- E. Özkaya, et al. (2010). ‘Estimating and benchmarking Less-than-Truckload market rates’. *Transportation Research Part E*: **46**(5):667–682.
- S. Pan, et al. (2013). ‘The reduction of greenhouse gas emissions from freight transport by pooling supply chains’. *International Journal of Production Economics* **143**(1):86–94.
- S. Pan, et al. (2014). ‘Environmental and economic issues arising from the pooling of SMEs’ supply chains: case study of the food industry in western France’. *Flexible Services and Manufacturing Journal* **26**(1-2):92–118.
- M. Perona & G. Miragliotta (2004). ‘Complexity management and supply chain performance assessment. A field study and a conceptual framework’. *International Journal of Production Economics* **90**(1):103–115.

- A. Rossi, et al. (2012). ‘Connectivity-and-hop-constrained design of electricity distribution networks’. *European Journal of Operational Research* **218**(1):48–57.
- T. M. Simatupang & R. Sridharan (2002). ‘The collaborative supply chain’. *The International Journal of Logistics Management* **13**(1):15–30.
- J. C. Tyan, et al. (2003). ‘An evaluation of freight consolidation policies in global third party logistics’. *Omega* **31**(1):55–62.
- C. Vanovermeire & K. Sörensen (2014). ‘Measuring and rewarding flexibility in collaborative distribution, including two-partner coalitions’. *European Journal of Operational Research* **239**(1):157–165.
- T. Wu, et al. (2015). ‘A multi-period location model with transportation economies-of-scale and perishable inventory’. *International Journal of Production Economics* **169**:343 – 349.
- G. Z#apfel & M. Wasner (2002). ‘Planning and optimization of hub-and-spoke transportation networks of cooperative third-party logistics providers’. *International Journal of Production Economics* **78**(2):207–220.