



**HAL**  
open science

## A supply chain network design problem under flow consolidation constraints

Mouna Kchaou Boujelben, Céline Gicquel, Michel Minoux

► **To cite this version:**

Mouna Kchaou Boujelben, Céline Gicquel, Michel Minoux. A supply chain network design problem under flow consolidation constraints. 9th International Conference on Modeling, Optimization & SIMulation, Jun 2012, Bordeaux, France. hal-00728636

**HAL Id: hal-00728636**

**<https://hal.science/hal-00728636>**

Submitted on 30 Aug 2012

**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.

## A SUPPLY CHAIN NETWORK DESIGN PROBLEM UNDER FLOW CONSOLIDATION CONSTRAINTS

**M. KCHAOU BOUJELBEN**

LGI / Ecole Centrale Paris  
Grande Voie des Vignes  
F-92 295 CHÂTENAYMALABRY.  
mouna.kchaou@ecp.fr

**C. GICQUEL**

LRI / Université Paris-Sud  
Bât 650 Université Paris-Sud 11  
91405 Orsay Cedex France  
celine.gicquel@iri.fr

**M. MINOUX**

Lip6 / Université Pierre & Marie  
Curie  
4, place Jussieu 75005 Paris  
michel.minoux@lip6.fr

**ABSTRACT:** We consider a supply chain network design problem featuring a 3-level structure, which uses intermediate distribution centers, the number and location of which are to be determined. A facility location model is proposed, including a special type of constraints called flow consolidation constraints in addition to maximum covering distance and limited capacity constraints. Our study is based on a concrete application in the automotive industry, namely, the outbound supply chain of a carmaker. Computational experiments are carried out on industrial data under different versions of the model. Our results show the benefits of integrating a clustering approach and several potentialities in terms of managerial scenarios analysis, such as the trade-off between distribution cost decrease and delivery time limit.

**KEYWORDS:** network design, flow consolidation, clustering, automotive

### 1 INTRODUCTION

In order to be competitive on the market, any company has to improve its supply chain management in order to satisfy its customers at the best service level and the lowest possible cost. It is therefore of key importance to optimize the logistics flow at different levels: strategic, tactical and operational. At the strategic level, network design problems have been widely studied by OR researchers, especially using facility location theory. Overviews of the related literature are provided among others by (Owen and Daskin, 1998), (Daskin and al., 2003) and (Melo and al, 2009). However, due to their various possible applications in real life, there are still many interesting research issues to explore within facility location problems (FLP). In this context, we are studying some shipment consolidation strategies in connection with facility location issues arising from a real life application in the automotive industry. The network under study (Figure 1) consists of three layers: assembly plants at the first one, distribution centers (DCs) at the second one and customers or car dealers at the third level.

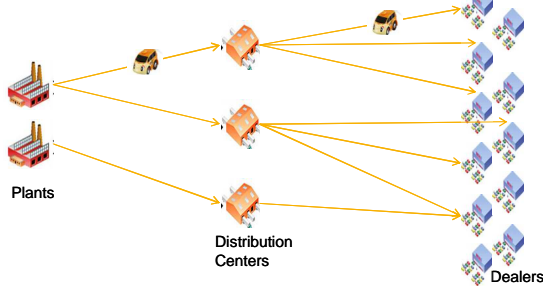


Figure 1: Outbound supply chain network of an automotive company

Vehicles are produced at assembly plants and are then transported to intermediate distribution centers (“primary transport”) where they are held for a short transit time (typically a few days) before being sent by truck to the car dealers (“secondary or last-mile transport”). As explained in (Eskigun A. and al., 2005), who discussed a similar application, intermediate DCs are created in order to “consolidate and distribute vehicles from different plants to dealers”. The company chose this organization to benefit from economies of scale for long-distances transport from manufacturing sites to DCs. Namely, cars are voluminous products, which can be transported only by specific, dedicated trucks of small capacity. Hence, while designing the logistics network, it is of primary importance to ensure that the use of full truckloads is possible at the tactical and operational levels. Our main focus in the present study is thus on designing the network so as to enable transport flow consolidation. Given the Supply Chain network described above and the average annual demand of each dealer, we aim at answering the following questions:

- (1) How many distribution centers (DCs) should be installed, and where should they be located
- (2) To which DC should each dealer be assigned and for which product?

In order to provide an answer, we propose a multi-product facility location model considering explicitly flow consolidation constraints together with distance and capacity constraints. The remainder of this paper is organized as follows. In section 2, we provide a review of the related literature. In Section 3, we discuss the modeling of costs and constraints before formulating the problem in section 4. Section 5 is devoted to the analysis of some computational experiments. Finally, Section 6 summarizes our conclusions and provides possible research perspectives.

## 2 LITERATURE REVIEW

(Melo and al, 2009) listed different kinds of FLP applications in the context of supply chain management. However, research works applied to the automotive industry appear to be rather scarce. For instance, (Eskigun and al., 2005) implemented a capacitated network design model using a non-linear waiting time whereas (Lin and al., 2006) modeled economies of scale in a context of vehicle distribution. Other works like (Grangeon and al., 2010) and (Nozick and Turnquist, 1998) proposed resolution methods and some results for integrated models. This consists in integrating tactical/operational decisions like inventory management or vehicle routing problems with the main strategic decision, which is the network design issue. The major purpose in this case is to study the trade-off between interdependent logistical components. This body of work includes the location routing problem (LRP) defined by (Nagy and Salhi, 2007) as “an approach to modeling and solving locational problem” where “we simultaneously need to solve a vehicle routing problem” in addition to the global location problem. The authors proposed an extensive literature review about LRP and classified the LRP heuristics according to a scheme of three classes based on how the location and the routing problems are related. Among these works, we find those using clustering-based methods such as (Srivastava and Benton, 1990), (Min, 1996) and (Barreto and al., 2007). The clustering approach that we are using in the present paper is similar to the one studied in the latter work. However, the majority of these works were in a mono-product context: see e.g. (Sajjadi, 2008), (Gunnarsson and al., 2006) and (Yi and Ozdamar, 2007) for multi-product examples. Furthermore, transport flow consolidation was seldom considered in the facility location literature. Authors usually presented concave cost functions to show economies of scale, see e.g. (Shen, 2005), (Lin and al., 2006), (Bucci, 2009), but few of them mentioned the trade-off between time and cost. Two papers (Guha and al., 2000) and (Karger and Minkoff, 2000) introduced the facility location problem with lower bound constraints in the context of internet access network design problem using caching strategies. They presented essentially the same approximation algorithm, which was improved by (Svitkina, 2010). In a supply chain management context, we found only one paper considering volume constraints in a transportation problem. (Lim and al., 2006) presented indeed a model where decisions involve allocation of cargos from customers to carriers taking into account an American regulation enforcing Minimum Quantity commitments (MQC) for carriers, if chosen.

## 3 MODELING COSTS AND CONSTRAINTS

As our different decisions are based on a concrete case study, explaining the main assumptions underlying the model is essential before presenting the problem mathematical formulation.

### 3.1 Product representation

As explained in the introduction, we deal here with a multi-product network design problem: around 40 vehicle types, without considering colors and options, have to be distributed through the network. However, assembly plants are mostly specialized, i.e. they can manufacture various vehicle types but a given type is usually produced in a single plant. Moreover, transport decisions starting from plants are taken according to the final destinations of the vehicles and not to their types. It is thus meaningful to use an aggregate representation of the product types based on their sourcing plants: in what follows, we will consider the different types produced in a same plant as a single product.

### 3.2 Modeling primary transport and transit costs

#### 3.2.1 Modeling primary transport costs

As mentioned in the introduction, primary transport concerns the products moved from assembly plants to distribution centers (DCs). It is easy to compute the cost of a truck starting from a given plant and going directly to a DC, using fixed and kilometric components. However, the evaluation of the unit delivery cost is less straightforward as it strongly depends on the truckloads on the considered link and thus on the given volume and the targeted waiting time at the plant. Indeed, increasing the waiting time to improve the truck loading rates leads to transport costs decreasing but results in an additional inventory cost and a poor service time. Figure 2 shows the unit delivery cost on a plant-DC link as a function of the average daily volume to be distributed on this link. We assume a truck capacity of 10 vehicles and a maximum waiting time (delivery frequency) of 5 days.

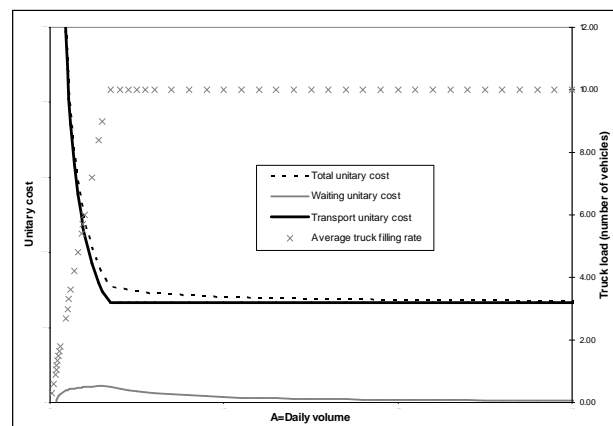


Figure 2: Primary transport unitary cost as a function of daily volume

Using the numerical data from our case study, we notice that the inventory cost (in grey) is insignificant compared to the transport cost (in bold black), that is why we do not consider it in the remainder of this paper. We also point out that the transport cost is steeply decreasing for volumes less than the volume limit and then becomes

constant. This volume limit consists in the minimum average daily volume needed to consolidate a full truckload within the maximum waiting time of 5 days. For a daily volume greater than this limit, deliveries are made using a full truckload. The unit delivery cost is thus equal to the cost of a truck divided by the truck capacity equal to 10 (crosses curve value equal to 10). Nevertheless, in reality, trucks never leave the plants unless they are completely full. Therefore, to achieve a trade-off between the two conflicting objectives, namely having full truckloads (to decrease costs) and meeting the delivery time, it is necessary to have much volume. The distribution network has thus to be designed so as to guarantee at least a volume equal to the volume limit for each primary transport link. This is done in our mathematical formulation by assigning a minimum volume constraint to each plant-DC link.

However, as this kind of constraint can be difficult to meet (it can lead to an infeasible problem if combined with other constraints), it was necessary to introduce artificial variables allowing, only in case of infeasibility, to fall under the volume limit. A penalty parameter, using these variables, is added to the objective function in order to measure the violation of the minimum volume constraints (see the model presented later in 3.4). Thus, it is possible to detect sources of infeasibility in order to adjust the volume limit, and therefore the maximum waiting time, per plant. In fact, in order to fill the trucks, it is tolerable to increase the waiting time at some plants, namely those manufacturing low volume.

### 3.2.2 Modeling transit costs

In the situation under study, most of the distribution centers and all the related activities (loading, unloading, etc.) are outsourced. Hence, there are no fixed opening costs to be incurred before using a distribution center. However, each time a vehicle goes through a DC, which is not belonging to the carmaker, a unitary cost called transit cost is paid to the third-party logistics. This cost is usually mentioned in a contract between the company and the third-party logistics where the two parties agree upon a minimum and a maximum transit volume during the planning period. We thus impose in our mathematical formulation minimum and maximum values for the annual transit volume through each DC if we decide to use it. We assume that the transit cost is constant for a value situated between these two bounds.

### 3.3 Modeling secondary transport costs

Secondary transport covers the transport of goods, by truck, from distribution centers to the final customers over distances usually shorter than primary transport distances. A given truck starting from a distribution center may have to visit several customers, as one customer has not necessarily enough demand to receive a full truckload within the allowed delivery time. Figure 3 illustrates the combination of primary and secondary transport.

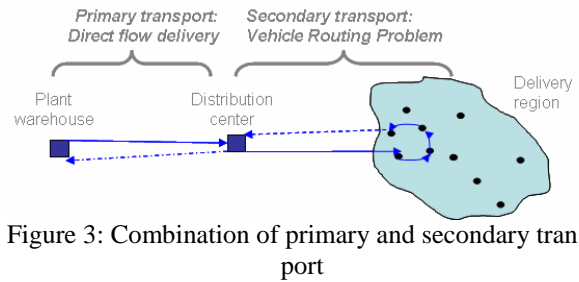


Figure 3: Combination of primary and secondary transport

As for primary transport, the main challenge, in this case, is to optimize the truckloads without increasing delivery times. We have though an additional difficulty related to the vehicle routing problem (VRP) described in Figure 3. This is an operational issue that we do not deal with in the present work but we have to consider it while designing the network as it could have a significant impact on location and assignment decisions. Namely, as was shown by (Shen and Qi, 2007), there could be a total cost reduction of more than 8% compared to a sequential approach (i.e. network design then routes calculation), if using an integrated routing distance approximation method in an uncapacitated location-inventory problem. (Javid and Azad, 2010) proved that calculating the routes without approximation leads to a cost improvement of 9-27% compared to Shen and Qi's model. In classical facility location problems, secondary transport costs are usually linear from source to destination and no effort is spent on how to get or to evaluate the unitary cost. In location-routing problems, the route decisions are explicitly modeled in the optimization problem, which results in large sized problems usually heuristically solved. The clustering-based method (Barreto and al, 2007) is a way to approximate the route lengths while keeping a simple and mid-sized model using linear costs. Accordingly, we have chosen this method for modeling our network design problem.

#### 3.3.1 The clustering-based method

The idea of the clustering-based method is to form groups of customers, which will be allocated to the same routes. In other words, we suppose that each time a truck leaves a distribution center to transport products intended for a group  $G$ , it stops at each customer of  $G$  to deliver a part of the truckload. Although this is not exactly what happens in reality, as routes are dynamic and not fixed for all the planning period, we consider that this static method provides us with a good estimate of secondary transport distances. Figure 4 illustrates the clustering-based method that we can summarize in four steps:

- (1) Construct clusters of customers with a minimum volume limit and other constraints.
- (2) Solve a Traveling Salesman Problem for each distribution route from DCs to clusters.
- (3) Compute the unitary cost per transported product for each route.
- (4) Solve a facility location problem using clusters as customers instead of original dealers

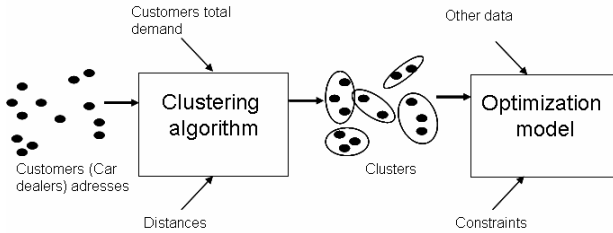


Figure 4: The clustering-based method

In step 1, we consider as input of the clustering algorithm the total demand of each customer, including all the product types. This is due to the possible combination of different car models in a same secondary transport truckload. In step 2, we use a simple enumeration to determine the shortest route starting from a DC, visiting all the customers of a cluster then coming back to the DC. In fact, a particular feature of vehicle distribution is that a route consists of a small number of customers (usually between one and four). This makes solving the TSP problem through enumeration possible. The specific trucks dedicated to cars transport can carry up to 8-10 products, according to the load factor of the vehicle type. Thus, it is theoretically possible to visit up to ten customers but from experience, the number is almost always less than or equal to four. In step 3, we use the same truck cost formula used for primary transport cost calculation in addition to a stop cost for each customer of the cluster.

### 3.3.2 The Clustering algorithm

The algorithm consists in clustering the close customers in a way to group those who have low demand with those who have important demand. The total demand of a cluster has to be sufficient to fill in average at least one truck within the allowed waiting time on the distribution center (transit time). The clustering algorithm has to meet, as far as possible, three constraints:

- A maximum distance between two customers (but if we have no other choice, we ignore this constraint)
- A maximum number of customers per cluster
- A minimum and a maximum total demand for each cluster

Input distances between customers are calculated using a Geographic Information System (GIS) to form a point-to-point distance matrix.

At the beginning of the algorithm, the cluster list is initialized to single-element clusters (i.e. each customer corresponds to a cluster). Let  $m$  be the current number of clusters remaining in the clustering algorithm, Figure 5 illustrates the first phase of the implemented algorithm.

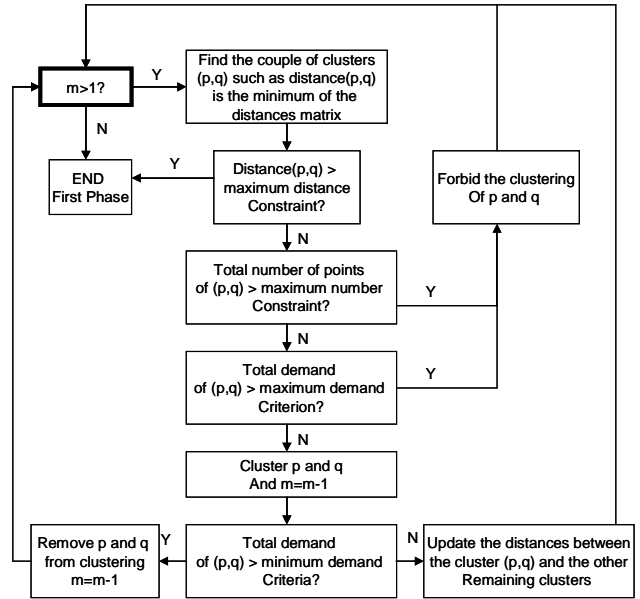


Figure 5: First phase of the clustering algorithm

We chose the following distance definition between two clusters  $p$  and  $q$ :

$$\text{Distance}(p, q) = \text{Min} \{ \text{Distance}(i, j), i \in C_p, j \in C_q \}$$

$C_p$  is the set of customers of the cluster  $p$ .

This is a proximity measure called the “single linkage” measure; see (Barreto and al, 2007) for a discussion of other possible measures.

After the first phase, a second phase has to be applied in order to consider the clusters discarded in the first phase due to different constraints. In the second phase, only the minimum volume and the maximum number of points constraints are imposed. We check that each cluster has a total demand greater than the minimum demand required. If it is not the case for some cluster  $q$  then we attach each of its customers to the nearest cluster having fewer customers than the maximum allowed number of customers.

### 3.3.3 Difficulties due to the multi-product, multi-sourcing aspects

The main objective of a clustering algorithm according to a minimum volume limit is to ensure full truckload routes for secondary transport. As mentioned before, we consider as an input of the clustering algorithm, the total demand of each customer, including all the product types. Thus, if we choose a single sourcing scenario, i.e. each cluster is assigned to exactly one DC, the criterion of minimum volume will be met as done in (Barreto and al, 2007). Nevertheless, if we consider a multiple sourcing scenario, then the total demand of each cluster could be assigned to more than one DC. Thus, the minimum volume limit could not be met and trucks would not be full. As we are concerned by the second case, the solution that we chose to implement in our model is to add a

minimum volume constraint to each of the DC-Cluster links. This will limit the number of DCs associated to each customer and will therefore guarantee a sufficient volume ensuring full truckloads within the maximum waiting time fixed by the user.

### 3.4 Constraint analysis

In addition to the minimum volume constraints for transport and transit flows, other constraints have to be added in order to properly fit our concrete application. For instance, a truck driver delivering vehicles to car dealers has to go back to his distribution center at the end of the working day. The traveled distance per delivery trip should not thus be greater than a given limit, allowing him to respect the legal daily driving time. This leads to a maximum covering distance constraint: assignment of customers to DC are forbidden if the corresponding distance is larger than a given value. Considering this additional constraint, together with the lower bounds for primary transport flows makes the location problem more difficult. Clearly, while the volume constraint pushes for limiting the number of DCs linked to each plant, the distance constraint tends to the opposite decision in order to be as close as possible to customers. If we also use volume constraints for secondary flows, the problem becomes even more complex as each of the two volume restrictions (primary and secondary) result in different decisions as far as the network structure is considered. Targeting a minimum quantity for secondary transport links encourages the decision of decreasing the number of DCs per cluster while the same constraint for primary transport leads to increasing this number. The former observation is intuitive but the latter is not, it was suggested by computational experiments and a possible explanation was deduced. In fact, increasing volume on each plant-DC link leads to decreasing the number of DCs associated to each plant. This means that all the DCs opened due to the distance constraint cannot deal with all the products. In this case, a cluster has to look at many centers in order to satisfy all its demand and this fact results in a higher number of centers per cluster.

In the field of logistics, there are also other rules that could be forced by third-party logistics or carriers. In our case, the vehicle distribution team has to cope with a "districting constraint", that is, assignment decisions are made on the basis of a given decomposition of the territory into districts and not for each single customer. This leads to a slightly different version of the mathematical model, which will be explained in paragraph 4.2.

### 3.5 Statement of the proposed model

One of the contributions of this paper consists in providing an approach to consider shipment consolidation within a network design problem. Two important ideas are studied in this context: introducing minimum volume constraints for transport and transit

flows, and using a clustering method for grouping customers. As far as transport is considered, to achieve a trade-off between two conflicting objectives, namely maximum delivery time limits and distribution cost decrease, it is necessary to have much volume; hence the idea of lower bounds for the primary and the secondary transport volumes. We also include this kind of constraint for the flows going through DCs instead of using fixed costs as usually done in the literature. The manufacturer does not pay fixed costs but the third-party logistics managing each DC asks for a minimum guaranteed volume to provide a transit area with a given tariff.

The clustering of customers into groups could also be viewed as a way for flow consolidation because it results in full truckloads in secondary transport. To the best of our knowledge, the clustering method in a multi-product, capacitated context with a multi-sourcing option (a customer is not necessarily associated to exactly one DC) and minimum volume constraints has not been addressed in the literature yet.

We provide also a concrete application dealing with a facility location problem in a supply chain context, a field for which application papers are scarce, as pointed out in (Melo and al., 2009).

## 4 PROBLEM FORMULATION

Using the modeling decisions detailed in paragraph 3, we formulate the mathematical model.

### 4.1 Mathematical model

Here we provide the problem formulation that will be the basis for other modifications.

#### 4.1.1 Data sets

$I$  set of plants ( $i=1..I$ )

$J$  set of potential DCs ( $j=1..J$ )

$Q$  set of clusters ( $q=1..Q$ )

#### 4.1.2 Parameters

$D_{qi}$ : Yearly demand of a cluster  $q$  for the product manufactured in plant  $i$

$minCap_j$ : Yearly minimum volume of vehicles going through DC  $j$  if it is chosen

$maxCap_j$ : Yearly maximum volume of vehicles going through DC  $j$  if it is chosen

$PTC_{ij}$ : Primary transport cost for a truck going from plant  $i$  to distribution center  $j$

$STC_{jq}$ : Secondary transport cost for a truck route starting at DC  $j$  and visiting all the customers of cluster  $q$

$TC_j$ : Unitary transit cost for the distribution center  $j$  (constant for a volume between  $minCap_j$  and  $maxCap_j$ )

$W_i$ : Average truck capacity for a truck starting its route from the plant  $i$  (primary transport)

$W$ : Average truck capacity for a truck starting its route from any distribution center (secondary transport)

$M$ : Big value

NWD: Number of working days in a planning period

$T_{max}(i)$ : Maximum waiting time at the plant  $i$  before shipping is made to distribution centers

$T$ : Maximum waiting time at a distribution center before shipping is made to dealers

$Max\_cap\_distance$ : Maximum route distance between a distribution center and a cluster  $q$

$PI$ : Plant-DC low volumes penalization amount

#### 4.1.3 Decision variables

We introduce the following decision variables:

*Binary variables:*

$$y_j = \begin{cases} 1 & \text{if distribution center } j \text{ is selected} \\ 0 & \text{if not} \end{cases}$$

$$z_{ij} = \begin{cases} 1 & \text{if route } [i,j] \text{ from plant } i \text{ to DC } j \text{ is selected} \\ 0 & \text{if not} \end{cases}$$

$$a_{jq} = \begin{cases} 1 & \text{if cluster } q \text{ is allocated to DC } j \text{ for at} \\ & \text{least 1 product} \\ 0 & \text{if not} \end{cases}$$

*Continuous variables:*

$x_{jq}$  = Proportion of the cluster  $q$  demand for the product manufactured in plant  $i$  and delivered by DC  $j$ .

$x_{jq}$  is defined only if:

$$\text{Route}(j,q) \leq \text{max\_cap\_distance}$$

$V'_{ij}, V''_{ij}$ : Artificial variables used to write minimum volume constraints for primary transport links.

$$V'_{ij} \geq V_{\min}(i)$$

$$V''_{ij} \leq V_{\min}(i)$$

With  $V_{\min}(i)$  equal to the minimum yearly volume ensuring a full truckload within  $T_{\max}(i)$  from plant  $i$  to any DC.

$$\Rightarrow V_{\min}(i) = \frac{W_i}{T_{\max}(i)} \times \text{NWD}$$

#### 4.1.4 Constraints

*Demand satisfaction constraint:*

$$\forall i \in I \quad \forall q \in Q \quad \text{such as } D_{qi} > 0 \quad \sum_{j=1}^J x_{jq} = 1 \quad (1)$$

Constraints (1) state that the demand of the cluster  $q$  for the product manufactured in plant  $i$  is satisfied.

*Minimum volume constraints in primary transport links:*

$$\forall i \in I \quad \forall j \in J \quad \sum_{q=1}^Q x_{jq} D_{qi} = V'_{ij} - V''_{ij} \quad (2)$$

$$V_{\min}(i) \times z_{ij} \leq V'_{ij} \leq M \times z_{ij} \quad (3)$$

$$V''_{ij} \leq V_{\min}(i) \times z_{ij} \quad (4)$$

With  $M = \min(\text{maxCap}_j, \text{Total Demand}(i))$

Constraints (2) state that the total volume going from plant  $i$  to distribution center  $j$  is expressed as a difference between two continuous variables. Constraints (3) and (4) mean that:

◆ If  $z_{ij} = 0$ , then the link  $[ij]$  is not selected, thus

$$V'_{ij} = 0 \text{ and } V''_{ij} = 0$$

◆ If  $z_{ij} = 1$ , then the link  $[ij]$  is selected, thus

$$V'_{ij} \geq V_{\min}(i), V'_{ij} \leq M \text{ and } V''_{ij} \leq V_{\min}(i)$$

$V''_{ij}$  is the allowed decrease of  $[i,j]$  volume compared to the volume limit  $V_{\min}(i)$ . It will be minimized, null if possible, as it is strongly penalized (see objective function). In fact, variables  $V'_{ij}$  and  $V''_{ij}$  are created to detect the possible sources of infeasibility.

*Distribution center maximum capacity constraint:*

$$\forall j \in J \quad \sum_{i=1}^I \sum_{q=1}^Q x_{jq} \times D_{qi} \leq y_j \times \text{maxCap}_j \quad (5)$$

Constraints (5) stipulate that if the DC  $j$  is selected ( $y_j=1$ ), then the flows transiting by  $j$  must not exceed its maximum capacity but, if the DC  $j$  is not selected ( $y_j=0$ ), then there is no flow transiting by it (all the  $x_{jq}$  are null).

*Distribution center minimum capacity constraint:*

$$\forall j \in J \quad \sum_{i=1}^I \sum_{q=1}^Q x_{jq} \times D_{qi} \geq y_j \times \text{minCap}_j \quad (6)$$

Constraints (6) stipulate that if the DC  $j$  is selected ( $y_j=1$ ), then the flows transiting by  $j$  have to be greater than its minimum volume limit but if the DC  $j$  is not selected ( $y_j=0$ ), then the quantity is positive or null.

*Minimum volume constraints in secondary transport links:*

$$\forall j \in J \quad \forall q \in Q \quad \sum_{i=1}^I x_{jq} \times D_{qi} \geq \frac{W}{T} \times \text{NWD} \times a_{jq} \quad (7)$$

$$\forall j \in J \quad \forall q \in Q \quad \sum_{i=1}^I x_{jq_i} \leq I \times a_{jq} \quad (8)$$

Constraints (7) ensure that secondary transport volume between warehouse  $j$  and cluster  $q$  is greater than a minimum volume (guaranteeing a full truckload within the maximum waiting time  $T$ ) if the link between  $j$  and  $q$  exists ( $a_{jq}=1$ ). Constraints (8) stipulate that all of the  $x_{jq_i}$  are zero if the link between  $j$  and  $q$  does not exist ( $a_{jq}=0$ ).

*Logical constraints:*

$$\forall i \in I \quad \forall j \in J \quad z_{ij} \leq y_j \quad (9)$$

Constraints (9) stipulate that route  $[i, j]$  cannot be selected if the distribution center  $j$  is not selected.

*Integrity and non-negativity constraints:*

$$\forall j \in J \quad y_j \in \{0,1\}$$

$$\forall i \in I \quad \forall j \in J \quad z_{ij} \in \{0,1\}$$

$$\forall j \in J \quad \forall q \in Q \quad a_{jq} \in \{0,1\}$$

$$\forall j \in J \quad \forall q \in Q \quad \forall i \in I \quad x_{jq_i} \in [0,1]$$

#### 4.1.5 Objective function

Min  $f(x)$ =Primary transport costs  
 + Secondary transport costs  
 + Distribution centers transit costs  
 + Penalization of low plant-DC volumes

With:

$$\text{primary transport costs} = \sum_{i=1}^I \sum_{j=1}^J \sum_{q=1}^Q \frac{PTC_{ij}}{W_i} x_{jq_i} D_{q_i}$$

$$\text{secondary transport costs} = \sum_{j=1}^J \sum_{q=1}^Q \sum_{i=1}^I \frac{STC_{jq}}{W} x_{jq_i} D_{q_i}$$

$$\text{Transit costs} = \sum_{j=1}^J \sum_{q=1}^Q \sum_{i=1}^I x_{jq_i} D_{q_i} TC_j$$

and

$$\text{penalization costs} = \sum_{i=1}^I \sum_{j=1}^J PI \times V''_{ij}$$

$V''_{ij}$  are artificial variables.

#### 4.2 Other versions of the model

##### 4.2.1 Adding the districting constraints

In the present work, we are dealing with a concrete application where the transport activities are outsourced

and the carrier imposes to divide the geographical area into districts using the zip codes. Assignment decisions as well as prices are thus given per district and not per single customer. It is therefore necessary to test another version of the model, considering a “districting constraint” which consists in applying three modifications to the modeling approach already described above. On the one hand, we have to use assignment variables  $x$  for districts instead of clusters and hence to constraint customers of a same cluster to belong to the same district. On the other hand, we must write the maximum distance condition in such a way that the distance between a DC and all the clusters of a district is less than the maximum allowed distance.

##### 4.2.2 Using the sequential approach

In order to study the contributions of the clustering approach, we also implement a sequential method where we first solve a facility location problem without regard to any clustering. We suppose that customers simply correspond to districts (demand aggregation) and use a unique reference location as representing each district when calculating distances from DCs. Then, using the location and assignment decisions, we apply the clustering algorithm to evaluate the route lengths for the deliveries allocated to each opened DC. Thus, we can review the secondary transport costs considering the obtained routes. Figure 6 summarizes the sequential approach:

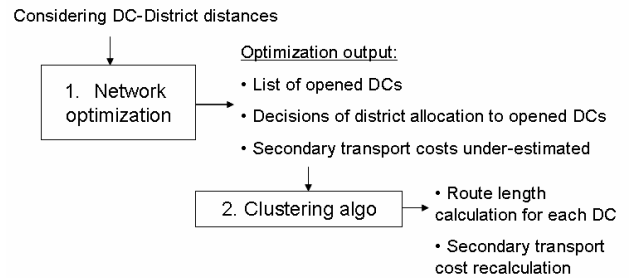


Figure 6: The sequential approach

## 5 COMPUTATIONAL TESTS

### 5.1 Case study description

Computational tests were carried out for one data set related to the concrete application we are studying, with 17 plants, 52 potential DCs and 563 customers. We considered a specific truck capacity (load factor) for each plant while calculating primary transport costs. On the contrary, we used the same truck capacity for all the secondary transport flows, as there is usually a mix of vehicle models in each truckload. This truck capacity was evaluated as the weighted average load factor using product volumes as weights. After applying the clustering algorithm according to the chosen grouping parameters (two truckloads as minimum volume, three truckloads as maximum volume and four as the maximum number of customers per cluster), we obtain 307 clusters (294 clusters if the districting constraint is applied). For



all the tests, the target waiting time at each plant was set to 5 days and the minimum volume on each DC-cluster link to one truckload.

## 5.2 Numerical results

The program was implemented in C++ and solved using ILOG Cplex academic version 12.1. All the tests were carried out on a Pentium Core 2 Duo (2.53 GHz) with 1.92Go of RAM, running under Windows XP. We employed a free batch (Batch géocodeur) to obtain the geospatial coordinates of the different points and a well-known Geographic Information System (GIS) to calculate route distances. We used the default settings of CPLEX solver for optimization. In the following subsections, we present the results of the computational experiments carried out to evaluate time, cost and managerial insights.

### 5.2.1 Time analysis

We intend here to study how the computational time depends on the value of the maximum covering distance constraint. Although we are able to reach the optimality in all the tests, we preferred to stop the algorithm at a gap equal to 1%. In fact, proving the optimality, in some cases, is very hard and the computational times exceed the limit of a manageable duration (3hours). Figure 7 illustrates the results obtained for the case study defined above.

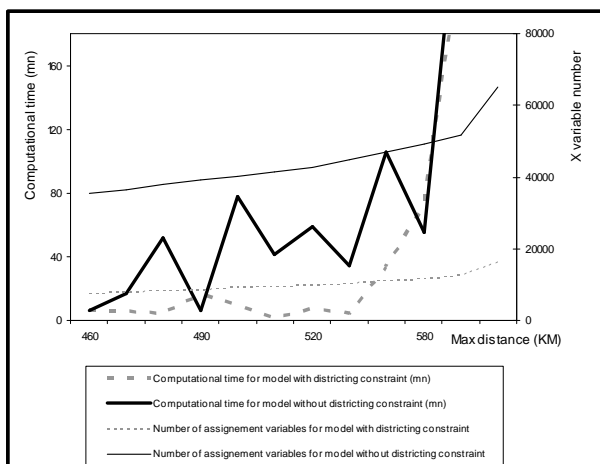


Figure 7: Time and variables number as a function of the maximum distance constraint value

It is seen that as the maximum distance value increases, the computational time increases. This observation seems to be obvious as increasing the distance results in a greater number of assignment variables  $x$  and thus in a problem of a larger size. However, it is worth pointing out that, in the case of our specific application, the problem becomes hard to manage starting from a distance constraint equal to 580km. It is noticeable also that, for a distance between 460 and 580km, the computational time is fluctuating and not regularly rising. It depends indeed on the potential links added due to the distance constraint relaxation.

### 5.2.2 Cost analysis

The decisions issuing from the three versions of the optimization model have different implications on the total distribution cost. For instance, considering that the assignment decisions have to be identical for all the clusters of a same district is very constraining. Thus, there is an additional cost to pay if the manager chooses (or is obliged to choose) this alternative. This cost increase is evaluated to 3.7% if testing the two model versions with a maximum distance constraint equal to 460KM and the same volume and waiting time constraints everywhere else.

As far as the clustering approach is concerned, we chose to compare the economic outputs of the integrated and the sequential method, considering the districting constraint. However, we face here a modeling difficulty concerning the maximum distance constraint formulation. In fact, in the integrated clustering approach, it is possible to include the distance constraint using the route distances between DCs and clusters, whereas in the sequential method, this is not possible, as clusters are not integrated in the optimization step. Thus, we have to use distances from DCs to the reference locations representing the districts, in the two cases, in order to compare the two models under the same hypothesis. Moreover, we add to the objective function two weighting factors, one for the primary transport cost and the other for the secondary transport cost. This aims at detecting the influence of the two costs proportions on the clustering benefits, the results are summarized in table 1.

Primary transport cost factor	Secondary transport cost factor	Secondary transport cost proportion*	Gain**
1	1	17.2%	1.68%
0.8	1.2	22.4%	2.08%
0.4	1.6	39.1%	4.60%
0.2	1.8	54.7%	6.02%
0.1	1.9	65.9%	6.96%
0.08	2.2	71.2%	7.24%
0.06	2.4	75.3%	7.31%
0.03	2.6	80.3%	8.02%

Table 1: The economic impact of using an integrated clustering approach

\* Calculated using the outputs of the integrated approach as secondary transport cost/objective function value

\*\* Calculated using the objective function value as (sequential model value - integrated model value)/sequential model value

As we can see, using an integrated clustering-approach leads to savings in the total distribution cost which become more significant when secondary transport costs account for the major part of the total cost. This could occur, for example, if distribution centers become closer to plants and farther from customers.

### 5.2.3 Managerial insights

As the network design model that we are studying is a decision-aid tool for supply chain managers, it is important to analyze some of the related managerial insights. In this context, we carried out a second test to compare the results obtained with the integrated and the sequential methods, this time formulating differently the maximum distance constraint (we use DC-clusters route

lengths for the first and DC-districts round-trip distances for the second). We consider then the results shown in table 2 (in the sequential case, maximum route lengths are calculated after the network optimization, using a clustering per DC). If a black entry is associated to a given DC, it means that this DC is not opened by the optimization approach corresponding to the considered line.

DC		1	2	3	4	5	6
Maximum route (KM)	Integrated approach	444.86	440.79		434.18	381.34	
	Sequential approach		625.299	406.854	626.452	361.55	570.908

DC		7	8	9	10	11	12
Maximum route (KM)	Integrated approach	456.82	409.88	451.37	354.26	444.29	
	Sequential approach	616.29	571.041	592.601	496.906	467.554	583.127

DC		13	14	15	16	17	18
Maximum route (KM)	Integrated approach	382.12	380.40	454.91	232.01	395.62	457.08
	Sequential approach	697.19	510.228	692.756	293.53	608.777	465.976

DC		19	20	21	22	23	24
Maximum route (KM)	Integrated approach	458.29	385.03	340.35	420.48	41.08	432.51
	Sequential approach	626.02	486.199	507.496	483.734	41.08	534.413

DC		25	26	27	28	29
Maximum route (KM)	Integrated approach	456.89	382.36	435.79		450.52
	Sequential approach	687.9	399.583		526.896	798.205

Table 2: Maximum route length for each opened DC in the cases of integrated and sequential approaches

We can obviously conclude that the clustering approach integration makes it possible to respect the maximum route distance constraint in all the assignment decisions (in table 2, first line, all the distances are less than 460KM). On the contrary, it is not the case in the sequential approach as the optimization model constraint is formulated globally for the whole district and not precisely for each cluster. Managers are also careful to the waiting times and their influence on the customer service rate. Further analyses of the results given by the integrated approach reveal that depending on the maximum route constraint value, waiting times at plants could deteriorate and even the problem could become infeasible if some parameters are not tuned. For instance, the test we made with the integrated approach, using a districting constraint and a maximum route length equal to 460KM, indicates that it is not possible to have a feasible problem with a maximum waiting time of five days at each plant as wishes the logistics manager. In fact, due to the maximum route constraint for secondary transport, each plant has to deliver its products to 25 opened DCs. This leads to the violation of the minimum volume constraints on three of the primary transport links. The constraint violation is detected by observing a high penalization parameter in the objective function value. In this case, it is mandatory to tune the problem parameters by increasing the maximum waiting time for the corresponding plants in order to avoid the penalization. Moreover, decreasing the maximum distance value from 460KM to 400KM makes the problem infeasible with the initial parameters. It is though possible to obtain a solution if we finely analyze the parameters to change. For example, in this case, we can detect a leak of capacity in one of the DCs and some customers that are situated at more than 400KM of the nearest DC.

It is also noticeable that the results given by the three implemented versions provide fractional assignment variables (about 9-25% of nonzero variables), that is, a product that comes from a given plant and intended for a given customer could go through different DCs. According to the logistics manager, this is not the case in practice but imposing an additional constraint of integer assignment variables leads to high computation times and even infeasibility in many cases.

## 6 CONCLUSIONS AND FURTHER RESEARCH

In this paper, we studied a facility location problem arising in the context of vehicle distribution. We proposed a modeling approach considering the flow consolidation in primary and secondary transport. For the first case, we introduced a lower bound, which is evaluated according to a trade-off between time delays and cost decrease for each plant-DC link. For the second case, we first grouped customers using a clustering algorithm in order to ensure full truckload deliveries. Then, we used the resulting clusters as final customers in our optimization program. Taking into account a fixed waiting time at DCs, we defined a lower bound for the volume going through each DC-cluster link. We assigned also a minimum transit volume constraint to each DC, instead of the traditional fixed costs. This led to the formulation of a large-size MILP with minimum volume, maximum distance and capacity constraints. Our computational experiments on a concrete case study showed encouraging results with the clustering approach. Compared to the sequential approach, the clustering-based one leads to a cost reduction, which increases as the secondary transport cost ratio increases. Furthermore, we analyzed some managerial insights such as the impact on cost of the districting constraint and the influence of the distance constraint on waiting times. As far as computational times are concerned, we concluded that the problem becomes hard to solve for higher values of the distance limit. It would thus be interesting to study a specific algorithm in the case of large-scale problems, which could happen if we relax the maximum distance constraint or if we consider many countries simultaneously. Besides, in real life situations, primary transport is multi-modal, using truck (road), train (rail), barge (river) or boat (sea). That is why, we suggest as a further consideration, to include transport mode choice in the model already defined.

## ACKNOWLEDGMENTS

The authors thank the company that supported this work.

## REFERENCES

- Barreto S., Ferreira C., Paixao J., Santos B. S., 2007. Using clustering analysis in a capacitated location-routing problem. *European Journal of Operational Research* 179, p. 968-977.

- Batch géocodeur,  
<http://www.batchgeocodeur.mapjnz.com/>
- Bucci Michael James, 2009. Solution Procedures for Logistic Network Design Models with Economies of Scale. A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.
- Daskin Mark S., Snyder Lawrence V., Berger Rosemary T., 2003. Facility location in supply chain design. Book "Logistics Systems: Design and Optimization" Chapter N°2.
- Eskigun E., Uzsoy R., Preckel P.V., Beaujon G., Krishnan S., Tew J.D., 2005. Outbound Supply Chain network design with mode selections, lead times, and capacitated vehicle distribution centers. *European Journal Of Operational Research*, 165, p.182-206.
- Grangeon Nathalie, Norre Sylvie, Gourguechon Olivier, Suon Médéric, 2010. Un problème de planification stratégique de type production-distribution avec économies d'échelle et technologies de production. *8e Conférence Internationale de MODélisation et SIMulation - MOSIM'10 Hammamet Tunisie.*
- Guha Sudipto, Meyerson Adam and Munagala Kamesh, 2000. Hierarchical Placement and Network Design Problems. *Proceedings of the 41th annual IEEE Symposium on Foundations of Computer Science*, p. 603-612.
- Gunnarsson H., Ronnqvist M. and Carlsson D, 2006. A combined terminal location and ship routing. *Journal of the Operational Research Society*, 57, p. 928-938.
- Javid Amir Ahmadi and Azad Nader, 2010. Incorporating location, routing and inventory decisions in supply chain network design. *Transportation Research Part E* 46, p. 582-597.
- Karger David R. and Minkoff Maria, 2000. Building Steiner Trees with Incomplete Global Knowledge. *Proceedings of the 41th annual IEEE Symposium on Foundations of Computer Science*, p. 613-623.
- Lim Andrew, Wang Fan and Xu Zhou, 2006. A Transportation Problem with Minimum Quantity Commitment. *Transportation Science, Volume 40 Issue 1*.
- Lin Jenn-Rong, Nozick Linda K., Turnquist Mark A., 2006. Strategic design of distribution systems with economies of scale in transportation. *Annals of Operations Research* 144, p. 161-180.
- Melo M.T., Nickel S. and Saldanha-Da-Gama F., 2009. Facility location and Supply Chain management, a review. *European Journal Of Operational Research* 196, p. 401-412.
- Min H., 1996. Consolidation terminal location-allocation and consolidated routing problems. *Journal of Business Logistics* 17 (2), 235-263.
- Nagy G. and Salhi S, 2007. Location-routing: Issues, models and methods. *European Journal of Operational Research* 117, pp.649-672.
- Nozick Linda K. and Turnquist Mark A., 1998. Integrating inventory impacts into a fixed-charge model for locating distribution centers. *Transportation Research Part E: Logistics and Transportation Review* 34 p. 173-186.
- Owen S.H. S.H., Daskin M.S, 1998. Strategic facility location: A review. *European Journal of Operational Research* 111 p.423-447.
- Sajjadi S.R., 2008. Integrated Supply Chain: Multi products location routing problem integrated with inventory under stochastic demand. *Dissertation submitted in partial fulfillments of the requirements for the degree of Doctor of Philosophy*, The department of industrial and manufacturing engineering and the faculty of the graduate school, Wichita state university.
- Shen Z.J., 2005. A multi-commodity supply chain design problem. *IIE Transactions*, vol.37, p.753-762.
- Shen Zuo-Jun Max and Qi Lian, 2007. Incorporating inventory and routing costs in strategic location models. *European journal of operational research* 179 p. 372-389.
- Srivastava R. and Benton W.C., 1990. The location-routing problem: Considerations in physical distribution system design. *Computers in Operations Research* 17, 427-435.
- Svitkina Zoya, 2010. Lower-bounded facility location. *ACM Transactions on Algorithms (TALG)* Volume 6 Issue 4.
- Yi Wei and Ozdamar Linet, 2007. A dynamic logistics coordination model for evacuation and support in disaster response activities. *European Journal of Operational Research*, 179, p.1177-1193.